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 P2₁ 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 \( C(3')-endo \) 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**

\( \beta \)-4-Thiouridine (s\(^4\)U), 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\(^4\)U as substitute for U in biological systems [5–13].

NMR data have demonstrated that s\(^4\)U 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 the pyrimidine (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\(^4\)U as substitute for U in biological systems [5–13].

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**     | \( \text{C}_9\text{H}_{13}\text{N}_2\text{O}_5 \)S |
| **Space group** | \( \text{P}_2_1 \), 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**      | \( \rho_{\text{calc}} = 1.568 \text{ g/cm}^3 \) |
| **Z**            | 2 |
| **F (000)**      | 272 |
| **Number of measured data** | 895 |
| \( \theta_{\text{max}} \) | 120° |
| **R-factor**     | 5.6% |

---

* 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
Results and Discussion

The molecular conformation of s4U 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 IIa, 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(10 and N(l) in the anti form are “normal”, those in the syn form are widened, compare C(2')-C(l')-N(l) and C(l')-N(l)-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 s4U.

The torsion angles, (Table IV), however, differ greatly, especially  \( \gamma_{CN} \), defined by  \( \theta(1')-C(1')-N(1)-C(6) \), is \(-87.1^\circ \) in syn [13] but \(26.5^\circ \) in the anti form. Further, the ribose puckering modes are different, with pseudorotation parameters \( T_{\text{max}} = 41.0^\circ \), \( P = 34.2^\circ \) for syn corresponding to a C(3')-endo, C(4')-exo \( (T) \) twist pucker and for anti, \( T_{\text{max}} = 41.1^\circ \), \( P = 15.7^\circ \) refers to an ideal C(3')-endo \( (E) \) 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 \( T = \exp \left[-2 \pi^2(U_{11}h^2a^*2 + U_{22}k^2b^*2 + U_{33}l^2c^*2 + 2 U_{12}hk a^* b^* + 2 U_{13}h l a^* c^* + 2 U_{23}k l b^* c^*) \right] \). (a) Nonhydrogen atoms (all values are \( \times 10^4 \)). (b) Hydrogen atoms.

<table>
<thead>
<tr>
<th>Atom</th>
<th>( x/a )</th>
<th>( y/b )</th>
<th>( z/c )</th>
<th>U11</th>
<th>U22</th>
<th>U33</th>
<th>U23</th>
<th>U13</th>
<th>U12</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(1)</td>
<td>-353 (4)</td>
<td>5324 (13)</td>
<td>7130 (4)</td>
<td>116</td>
<td>19</td>
<td>249</td>
<td>24</td>
<td>299</td>
<td>24</td>
</tr>
<tr>
<td>C(2)</td>
<td>649 (5)</td>
<td>5512 (15)</td>
<td>6344 (5)</td>
<td>173</td>
<td>24</td>
<td>353</td>
<td>30</td>
<td>221</td>
<td>24</td>
</tr>
<tr>
<td>N(3)</td>
<td>1686 (4)</td>
<td>3726 (13)</td>
<td>6582 (4)</td>
<td>162</td>
<td>21</td>
<td>304</td>
<td>25</td>
<td>223</td>
<td>21</td>
</tr>
<tr>
<td>C(4)</td>
<td>1867 (5)</td>
<td>1965 (16)</td>
<td>7508 (5)</td>
<td>137</td>
<td>24</td>
<td>286</td>
<td>28</td>
<td>283</td>
<td>26</td>
</tr>
<tr>
<td>C(5)</td>
<td>816 (6)</td>
<td>1911 (17)</td>
<td>8312 (5)</td>
<td>236</td>
<td>26</td>
<td>451</td>
<td>31</td>
<td>293</td>
<td>26</td>
</tr>
<tr>
<td>C(6)</td>
<td>-242 (6)</td>
<td>3581 (17)</td>
<td>8093 (5)</td>
<td>254</td>
<td>26</td>
<td>399</td>
<td>30</td>
<td>250</td>
<td>26</td>
</tr>
<tr>
<td>O(2)</td>
<td>397 (4)</td>
<td>7125 (11)</td>
<td>5517 (4)</td>
<td>258</td>
<td>20</td>
<td>350</td>
<td>25</td>
<td>328</td>
<td>21</td>
</tr>
<tr>
<td>S(4)</td>
<td>3213 (1)</td>
<td>23 (6)</td>
<td>7658 (1)</td>
<td>238</td>
<td>8</td>
<td>412</td>
<td>10</td>
<td>373</td>
<td>9</td>
</tr>
<tr>
<td>C(1')</td>
<td>-1505 (6)</td>
<td>7182 (15)</td>
<td>6874 (5)</td>
<td>177</td>
<td>24</td>
<td>348</td>
<td>30</td>
<td>289</td>
<td>26</td>
</tr>
<tr>
<td>C(2')</td>
<td>2580 (6)</td>
<td>5882 (15)</td>
<td>6616 (5)</td>
<td>167</td>
<td>24</td>
<td>392</td>
<td>30</td>
<td>192</td>
<td>24</td>
</tr>
<tr>
<td>C(3')</td>
<td>3489 (6)</td>
<td>4583 (15)</td>
<td>6966 (5)</td>
<td>175</td>
<td>23</td>
<td>262</td>
<td>29</td>
<td>305</td>
<td>26</td>
</tr>
<tr>
<td>C(4')</td>
<td>3440 (6)</td>
<td>6809 (16)</td>
<td>7919 (5)</td>
<td>185</td>
<td>24</td>
<td>327</td>
<td>28</td>
<td>269</td>
<td>26</td>
</tr>
<tr>
<td>O(1')</td>
<td>-2061 (4)</td>
<td>7772 (11)</td>
<td>7979 (4)</td>
<td>207</td>
<td>19</td>
<td>341</td>
<td>23</td>
<td>300</td>
<td>20</td>
</tr>
<tr>
<td>O(2')</td>
<td>3227 (4)</td>
<td>8116 (11)</td>
<td>5343 (4)</td>
<td>212</td>
<td>19</td>
<td>381</td>
<td>24</td>
<td>279</td>
<td>20</td>
</tr>
<tr>
<td>O(3')</td>
<td>-4792 (4)</td>
<td>4242 (12)</td>
<td>6403 (4)</td>
<td>208</td>
<td>20</td>
<td>456</td>
<td>27</td>
<td>453</td>
<td>22</td>
</tr>
<tr>
<td>O(5')</td>
<td>3758 (7)</td>
<td>5904 (17)</td>
<td>9155 (6)</td>
<td>308</td>
<td>27</td>
<td>532</td>
<td>33</td>
<td>367</td>
<td>28</td>
</tr>
<tr>
<td>O(3')</td>
<td>-3028 (5)</td>
<td>3507 (13)</td>
<td>9511 (4)</td>
<td>500</td>
<td>24</td>
<td>490</td>
<td>26</td>
<td>287</td>
<td>21</td>
</tr>
</tbody>
</table>
Table D b.

<table>
<thead>
<tr>
<th>Atom</th>
<th>x/a</th>
<th>y/b</th>
<th>z/c</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>H (3)</td>
<td>2291</td>
<td>3910</td>
<td>6094</td>
<td>500</td>
</tr>
<tr>
<td>H (5)</td>
<td>864</td>
<td>659</td>
<td>9042</td>
<td>500</td>
</tr>
<tr>
<td>H (6)</td>
<td>-892</td>
<td>3446</td>
<td>8552</td>
<td>500</td>
</tr>
<tr>
<td>H (1)</td>
<td>-1198</td>
<td>9067</td>
<td>6471</td>
<td>500</td>
</tr>
<tr>
<td>H (2)</td>
<td>-2286</td>
<td>4607</td>
<td>5386</td>
<td>500</td>
</tr>
<tr>
<td>H (3)</td>
<td>-2912</td>
<td>2749</td>
<td>7360</td>
<td>500</td>
</tr>
<tr>
<td>H (4')</td>
<td>-4096</td>
<td>8367</td>
<td>7570</td>
<td>500</td>
</tr>
<tr>
<td>H (5')</td>
<td>-4643</td>
<td>5164</td>
<td>9133</td>
<td>500</td>
</tr>
<tr>
<td>H (5)</td>
<td>-3614</td>
<td>6958</td>
<td>10072</td>
<td>500</td>
</tr>
</tbody>
</table>

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 $\chi_{CN}$ and pseudorotation parameters $P$ and $T_{\text{max}}$:

U(A) $\chi_{CN} = 18.3^\circ$, $T_{\text{max}} = 40.4^\circ$, $P = 3.7^\circ$;

U(B) $\chi_{CN} = 24.3^\circ$, $T_{\text{max}} = 42.4^\circ$, $P = 13.8^\circ$.

Comparison with s4 U shows that anti-s4 U is very similar to U(B).

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

<table>
<thead>
<tr>
<th>Atoms</th>
<th>syn [13]</th>
<th>anti</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-B-C</td>
<td>Distance</td>
<td>Angle</td>
</tr>
<tr>
<td>N(1)-C(2)-N(3)</td>
<td>1.382</td>
<td>115.3</td>
</tr>
<tr>
<td>C(2)-N(3)-C(4)</td>
<td>1.373</td>
<td>114.5</td>
</tr>
<tr>
<td>N(3)-C(4)-C(5)</td>
<td>1.433</td>
<td>119.8</td>
</tr>
<tr>
<td>C(5)-C(6)-N(1)</td>
<td>1.346</td>
<td>122.5</td>
</tr>
<tr>
<td>C(6)-N(1)-C(2)</td>
<td>1.370</td>
<td>120.8</td>
</tr>
<tr>
<td>O(2)-C(2)-N(1)</td>
<td>1.221</td>
<td>123.4</td>
</tr>
<tr>
<td>S(4)-C(4)-N(3)</td>
<td>1.662</td>
<td>121.4</td>
</tr>
</tbody>
</table>

The crystal packing patterns of s4 U 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-s4 U) 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-s4 U but at an angle 44.3° in anti-s4 U, 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-s4 U (Fig. 2) where S(4) accepts

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

<table>
<thead>
<tr>
<th>Atoms</th>
<th>syn [13]</th>
<th>anti</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>Angle</td>
<td>Distance</td>
</tr>
<tr>
<td>N(1)-C(2)-N(3)</td>
<td>1.382</td>
<td>115.3</td>
</tr>
<tr>
<td>C(2)-N(3)-C(4)</td>
<td>1.373</td>
<td>114.5</td>
</tr>
<tr>
<td>N(3)-C(4)-C(5)</td>
<td>1.433</td>
<td>119.8</td>
</tr>
<tr>
<td>C(5)-C(6)-N(1)</td>
<td>1.346</td>
<td>122.5</td>
</tr>
<tr>
<td>C(6)-N(1)-C(2)</td>
<td>1.370</td>
<td>120.8</td>
</tr>
<tr>
<td>O(2)-C(2)-N(1)</td>
<td>1.221</td>
<td>123.4</td>
</tr>
<tr>
<td>S(4)-C(4)-N(3)</td>
<td>1.662</td>
<td>121.4</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Atoms</th>
<th>syn</th>
<th>anti</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(1')</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>C(2')</td>
<td>0.286</td>
<td>-0.050</td>
</tr>
<tr>
<td>O(2')</td>
<td>-0.969</td>
<td>-1.435</td>
</tr>
<tr>
<td>C(3')</td>
<td>0.819</td>
<td>0.591</td>
</tr>
<tr>
<td>O(3')</td>
<td>0.676</td>
<td>0.155</td>
</tr>
<tr>
<td>C(4')</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>C(5')</td>
<td>0.491</td>
<td>0.763</td>
</tr>
<tr>
<td>O(1')</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
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 further 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).

<table>
<thead>
<tr>
<th>Bond</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(3)-O(2')</td>
<td>2.855</td>
</tr>
<tr>
<td>S(4)-O(3')</td>
<td>3.263</td>
</tr>
<tr>
<td>S(4)-O(5')</td>
<td>3.289</td>
</tr>
<tr>
<td>O(2')-O(3')</td>
<td>2.835</td>
</tr>
</tbody>
</table>

**Conclusions**

The two different crystal structures of *syn* and *anti* sU, 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*-sU 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 sU 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.

sU constitutes a good experimental basis for further theoretical investigations.
References