Conformation of a Rare Nucleoside in the Anti-Codon Loop of tRNAs: Potential Energy Calculations for 2'-O-Methyl Cytidine

S. D. STELLMAN,* B. HINGERTY,** S. BROYDE,† and R. LANGRIDGE, Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, N.J. 08540

Synopsis

The classical potential energy of 2'-O-methyl cytidine was calculated using contributions from van der Waals', electrostatic, and torsional terms. All five conformational angles, namely $\chi$, $\Psi$, the methoxy angles $m_1$ and $m_2$ which are unique to O'-methylated nucleosides, and the sugar pucker $P$ were varied simultaneously, and the energy was minimized with respect to these parameters. An extensive search of conformation space was made, particularly with respect to the sugar pucker. At the predicted global minimum, $\chi$ and $\Psi$ are anti-gg, $P$ is C(2')-endo-C(3')-exo, and the methyl group is staggered with respect to the sugar. This calculated minimum agrees very well with the recently determined crystal structure of 2'-O-methyl cytidine. Thus, 2'-O-methylation still permits the conformational regions of the common nucleosides to be adopted. However, we infer that the predicted C(2')-endo sugar pucker results from the added methyl group, since cytidine is C(3')-endo in the crystal, and the common ribopyrimidine nucleosides generally favor C(3')-endo.

INTRODUCTION

Although RNA is ordinarily thought of as consisting of a linear sequence of the four common bases (adenine, guanine, cytidine, and uracil), it is now known to contain as much as 10–20% of so-called rare nucleosides, which are chemical modifications of these four. One important class of modified nucleosides is characterized by methylation at the 2'-hydroxyl position of the ribose group; all four RNA bases are known to undergo this modification in various RNAs.

Apart from the fact that presence of a methyl group at the 2' position of pyrimidine nucleosides blocks the action of pancreatic ribonucleases, little is known about the presumably subtle effects this type of nucleoside modification is likely to have on the conformations of the many RNA species in which it has been observed. In particular, one such py-
rimidine nucleoside, 2'-O-methyl cytidine (2'MeCyd) is a frequent component of the anticodon loop of several tRNAs. An examination of the influence of 2'-O-methylation on the conformation of tRNA might lead to a better picture of the codon-anticodon interaction, and hence to an improved understanding of the relation between structure and function in these molecules.

In this paper we describe results of a study of the conformational properties of 2'MeCyd as determined by semiempirical potential energy calculations. Recently Hingerty et al. solved the crystal structure of this molecule by X-ray diffraction methods. They found that 2'MeCyd crystallizes with two molecules in the asymmetric unit, each having a distinctly different conformation. Previous crystallographic analysis of the dinucleoside phosphates GpC and UpA also showed that these molecules exhibit flexibility even within one crystal. Semiempirical potential calculations have proved useful in explaining these observations and relating conformational features of these smaller molecules to allowable conformations of RNA.

In this paper we report the results of a similar investigation made for the molecule 2'MeCyd, in vacuo. This study is of particular interest in development of potential energy methods, because the five-dimensional conformation space of this molecule could be investigated with greater thoroughness than was possible for previous molecular studies, giving us confidence that we have found all the important minima for our model of 2'MeCyd. Our calculations predict a global minimum in which the cytidine base conformation is anti, the methyl group is staggered with respect to the sugar, the 5'-hydroxyl group is gauche to both C(1') and C(3'), and the sugar puckering is in the C(2')-endo-C(3')-exo twist form. This prediction is in excellent agreement with the X-ray crystallographic findings.

Method

The method follows that used in our previous work. The molecular structure of 2'MeCyd is shown in Figure 1, and the angle-naming convention in Table I. The molecular conformation is considered to depend on five variables: the dihedral angles $x$, $\Psi$, $m_1$, and $m_2$ (Table I), and the sugar pucker $P$. The angle $P$ is a pseudorotation coordinate whose values (from 0 to 360°) correspond to each of the ten envelope forms C(3')-endo, C(2')-exo, etc.) and intermediate twist forms, in the fashion described by Altona and Sundaralingam. For example, $P = \ldots$

<table>
<thead>
<tr>
<th>Table I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle-Naming Convention for 2'-O-Methyl Cytidine</td>
</tr>
<tr>
<td>$x$</td>
</tr>
<tr>
<td>$m_1$</td>
</tr>
<tr>
<td>$m_2$</td>
</tr>
<tr>
<td>$\Psi$</td>
</tr>
</tbody>
</table>
18° corresponds to pure C(3')-endo, while $\mathbf{P} = 0°$ corresponds to a C(3')-endo–C(2')-exo twist form.

For a given fixed conformation the energy is calculated as the sum:

$$E = E_{\text{nb}} + E_{\text{el}} + E_{\text{tors}} + E_{\text{rib}}$$  \hspace{1cm} (1)

where $E_{\text{nb}}, E_{\text{el}},$ and $E_{\text{tors}}$ are, respectively, the nonbonded, electrostatic, and torsional contributions, and $E_{\text{rib}}$ is the energy of the unmethylated ribose, and includes the strain energy. Parameters used for the first three terms were those used in previous work, and a value of 4.0 was used for the dielectric constant. The ribose component was originally calculated by T. Sato (unpublished work) as a function of $P$. A graph of $E_{\text{rib}}$ versus $P$ is given in Figure 2.

Interaction parameters for O(2') and its methyl group were chosen from chemically similar atoms elsewhere in the molecule. Thus O(2')
was assigned the same charge as O(1'), -0.162. The carbon and hydrogen atoms of the methyl group were assigned charges of 0.114 and 0.053, respectively, like O(2') and H(2'). The barrier heights for rotation about \( m_1 \) and \( m_2 \) were both taken to be 2.0 kcal/mole.

The total energy, \( E \), was minimized starting from 180 different points in the conformation space of the molecule. These starting positions may be summarized as follows:

\[ P \text{ 0–324° in 36° intervals} \]
\[ \chi \text{ 230° (syn) and 15° (anti, C(3')-endo region) or 55° (anti, C(2')-endo region)} \]
\[ m_1 \text{ 60° (gt), 180° (tg), 300° (gg)} \]
\[ m_2 \text{ 60° (staggered)} \]
\[ \psi \text{ 60° (gg), 180° (gt), 300° (tg)} \]

For the minimizations, the modified Powell algorithm BOTM was used, in which the required accuracy at the minimum was set to 1° in each parameter, and the maximum variation in any angle at any step was 100°.\(^{11a}\)

**RESULTS**

From 180 starting conformations, 64 different local minima were reached, with highest at about 38 kcal/mole above the global minimum. The first 13 of these have relative energies under 2.0 kcal/mole, and are represented in Table II. As a first approximation, we take as the degeneracy of a minimum, \( D \), the number of starting conformations leading to that minimum, and we find the average degeneracy is approximately 2.8. Since the starting conformations themselves are nearly evenly distributed throughout the conformation space of the molecule, \( D \) should give a rough picture of the width of each local minimum. Thus, the

**TABLE II**

Thirteen Lowest Minimum Energy Conformations Calculated for 2'-O-Methyl Cytidine

<table>
<thead>
<tr>
<th>Minimum no.</th>
<th>Conformational angles, deg</th>
<th>Relative energy</th>
<th>( D_i )</th>
<th>( P_i ), %</th>
<th>Description(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60 60 55 165 0.00 4 13.0 A ( ^1E ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>31 64 57 22 0.04 4 18.3 A ( ^1E ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>59 60 299 180 0.06 4 11.8 A ( ^1T ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>58 61 178 180 0.47 4 5.9 A ( ^1T ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>25 179 61 57 180 0.52 6 8.2 A ( ^1T ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>31 68 295 11 0.52 5 6.8 A ( ^1E ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>238 61 297 162 0.54 11 14.5 S ( ^1E ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>237 61 297 162 0.68 7 7.3 S ( ^1E ) gg gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>29 65 179 11 0.78 5 4.5 A ( ^1E ) gt gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>33 179 299 183 1.06 6 3.3 A ( ^1T ) gt gg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>239 61 179 162 1.31 6 2.2 S ( ^1E ) gt gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>28 178 178 189 1.33 6 2.1 S ( ^1T ) gt gg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>237 61 179 162 1.69 4 0.9 S ( ^1E ) gt gt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( ^a \) In order: \( \chi, P, \Psi, m_1, S = \text{syn}, A = \text{anti}, \ ^1E = \text{C(2')-endo}, \ ^1T = \text{C(2')-endo–C(3')-exo}, \ ^2E = \text{C(3')-endo}. \) For \( \Psi, gg = 60°, gt = 180°, tg = 300°. \) For \( m_1, gg = 300°, gt = 60°, tg = 180°. \) Energies in kcal/mole.
first 13 minima have an average degeneracy of 5.69, and account for 41% of all minima found.

A more meaningful estimate of the size of each region is a computed statistical weight, $w_i$, defined as

$$w_i = Z^{-1} D_i \exp \{-E_i/RT\}$$  \hspace{1cm} (2)$$

where

$$Z = \sum_{i=1}^{s_4} D_i \exp \{-E_i/RT\}$$  \hspace{1cm} (3)$$

In Eqs. (2) and (3), $E_i$ is the energy and $D_i$ the degeneracy of the $i$th minimum, and $Z$ is the partition function. The probability in percent, $P_i$, for the $i$th minimum is $100w_i$. Probabilities, $P_i$, are also included in Table II, using a temperature of 300°K. In this calculation, $Z = 1.191 \times 10^{-2}$. The probabilities for these first 13 minima sum to 98.8%, and thus represent the great majority of significant conformations.

In Table III we give a breakdown of the minimum-energy conformational angles among the 13 most prominent minima. It can be seen at once that these angles take on very narrow ranges of values, giving rise to a small, discrete set of possible molecular conformations.

The glycosidic angle $\chi$ is located in either a narrow syn range ($S$), or a much broader anti range ($A$), with the latter preferred by a probability margin of 3:1. The sugar puckering is distributed almost equally among three basic conformations: C(2')-endo ($^2E$), C(3')-endo ($^3E$), and twist C(2')-endo-C(3')-exo ($^2T_{33}$). The first and third of these, which might otherwise be lumped into the same category, are inferred to be different for the purposes of this study because of the narrowness of their respective ranges and the gap between them, as seen in Table III and Figure 3.

The two pendant groups, 5'-hydroxyl and 2'-O-methyl, also take on plausible values. For $\Psi$, the probabilities are distributed about equally
between gg and tg, with a much smaller gt component. However, the orientation of this angle is linked to the conformation of the base. The gg conformation of Ψ is associated exclusively with bases anti (P_i = 39.5%), while tg occurs with equal probability for anti (P_i = 21.9%) and syn (P_i = 21.8%) bases. The methoxy angle m_1 favors gt (60°) over tg (180°) by nearly 4:1, and the gg (300°) conformation is improbable. Furthermore, the sugar pucker influences this angle: for C(3')-endo pucker, m_1 is only gt in the first thirteen minima (P_i = 29.6%), while the C(2')-endo ranges favor gt (P_i = 47.4%) but permit tg (P_i = 21.8%). The methyl conformation m_2 is found to be exclusively staggered (56–62°).

It is useful to compare the observations noted above for the most important minima with the results for the higher energy local minima, even though numbers 14–64 collectively contribute to only 1.2% of the total conformational probabilities at 300°K. In Figure 3 we plot histograms for observations on all local minima with relative energies of 7.0 kcal/mole or less. This set accounts for 150 of the total 180 observations. (Observations with energies above 7 kcal/mole were ignored for convenience—they give rise to occasional anomalous angles and their overall contribution is effectively zero.) Once again, the glycosidic
angle is confined to syn and anti, although the latter region is somewhat broader than before (0–60°).

When higher energy conformations are considered, the sugar puckering can now take on several new puckering values, including C(4')-exo (4E) and C(1')-exo–C(2')-endo (1T²). Otherwise, the three puckering which dominated the 13 lowest minima are still the most important ones here also. The hydroxyl angle Ψ is still narrowly confined strictly to its three major ranges, in nearly equal proportions. Finally, the methoxy angle m₁ is still found in the gt and tg conformations, although for conformations above 3.5 kcal/mole in energy it is now capable of assuming the gg angle. As illustrated in Figure 4, to accomodate this angle the sugar conformation is sometimes distorted from the pure envelope or twist forms, particularly when the puckering is close to C(3')-endo.

In general, then, we observe from Figure 3 that the basic preferred conformational angles are preserved, although the ranges are broadened when minima above the first 13 are considered. In other words, the introduction of conformational stress into the cytidine molecule in the
form of 2'-O-methylation may lead to some distortion, but it is not predicted to force the molecule into unusual conformations.

DISCUSSION

Comparison with Previous Potential Studies

Our findings are consistent with predictions made in previous potential energy studies for the common nucleosides and nucleotides. For example, Lakshminarayanan and Sasisekheran\textsuperscript{12} calculated the potential energy of common nucleosides as a function of $\chi$ for various fixed sugar puckering, and found that for pyrimidines $anti$ was favored over $syn$ by about 3 kcal/mole for C(3')-endo sugars, while for C(2')-endo sugars the energies of $anti$ and $syn$ were nearly equal, except for cytidine, in which $anti$ was favored by about 0.5 kcal/mole. (Anti was also favored for both C(2')-exo and C(3')-exo puckering.) We make the sim-
ilar observations for 2′MeCyd that the lowest syn conformer is 0.5 kcal/mole above the lowest anti, with sugar pucker C(2′)-endo, while for C(3′)-endo pucker syn was not observed. Calculations by these authors on the sugar-phosphate unit concluded\(^\text{13}\) that low-energy ranges for \(\Psi\) were 60°, 180°, and 300° (gg, gt, tg). It is noteworthy that their results were essentially undisturbed by the presence or absence of the 2′ hydroxyl group.

Calculations by Yathindra and Sundaralingam\(^\text{14}\) on 5′-nucleotides predict \(\chi\) and \(\Psi\) to be anti-gg as the global minimum for 5′-AMP, 5′-CMP, and 3′,5′-CDP. Furthermore, no significant contributions from syn conformations were found, and their calculated energy maps for 5′-CMP show no contours for the syn region below 5 kcal/mole. We find, on the other hand, a significant fraction (25%) of syn conformations for 2′MeCyd arising mostly from minima 7 and 8, and having a combined degeneracy of 18, indicating a very broad local minimum. Moreover, the syn conformation of 2′MeCyd permits \(\Psi\) to be gt, or less likely, tg, but not gg. This accounts for the higher probability for \(\Psi = gt\) in
2'MeCyd compared to 5'-CMP and 3',5'-CDP, where the probabilities for \( gg \) are 95% and 84%, respectively. The elimination of the \( \text{syn} \) range for 5'-CMP is attributed to the presence of the 5'-phosphate group. We therefore expect that this region would also be reduced or eliminated in the corresponding methylated 5'-nucleotide.

The present calculations on 2'MeCyd differ from these earlier calculations on nucleosides and nucleotides, and from a recent conformational study of a methylated ribopurine by Prusiner et al.,\(^{15}\) in that (a) all conformational angles were permitted to be simultaneously variable, and (b) an assessment was made of the probabilities of the various sugar puckering. The C(2')-endo regions, \((\beta_E)\) and \((\beta_3T)\), represent over 2/3 of the calculated statistical weights. A comparison of this prediction with experimental observations is made below.

**Comparison with Experimental Crystal Structures**

Our calculations are consistent with the structure of 2'MeCyd determined experimentally by X-ray crystallographic analysis. In Table IV we compare the values of parameters observed by Hingerty et al.,\(^5\) for the two independent molecules of 2'MeCyd with the predictions given in Table II. The global minimum for these calculations coincides with the predominant observation for molecule 1, while predicted minimum no. 5, lying only 0.5 kcal above no. 1, is similar to and represented as molecule 2 in the crystal. These conformers are depicted in Figure 5. It is of further interest to note that for calculated conformations having sugar pucker in the C(2')-endo ranges, \((\beta_E)\) and \((\beta_3T)\), with \( \chi = \text{anti} \), the probability ratio of \( gt \) to \( tg \) conformations for \( m_1 \) is nearly 2:1, which is very close to the observed statistical distribution of this disordered angle within molecule 1. Closer agreement than this could not be expected, in view of (a) the large number of calculated local minima lying within 2 kcal of the lowest one, and (b) the fact that as a scantily hydrated crystal, 2'MeCyd might easily be subject to crystal packing forces sufficient to distort the molecule significantly, or at least rearrange the order of the minima listed in Table IV.

**TABLE IV**

Comparison of Predicted and Observed Conformations of 2'-O-Methyl Cytidine

<table>
<thead>
<tr>
<th>Conformation</th>
<th>Ref.</th>
<th>Dihedral angle, deg Pucker, ( \chi ) ( \psi )</th>
<th>( m_1 )</th>
<th>( m_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecule 1</td>
<td>5</td>
<td>175 47 56 95 (65%)(^a) (60)(^b)</td>
<td>156 (35%)</td>
<td></td>
</tr>
<tr>
<td>Molecule 2</td>
<td>5</td>
<td>170 44 52 168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum 1</td>
<td>this work</td>
<td>165 60 55 60 56</td>
<td>60 56</td>
<td></td>
</tr>
<tr>
<td>Minimum 5</td>
<td>this work</td>
<td>180 25 57 179 61</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( ^a \) Statistical disorder.
\( ^b \) Disordered.
\( ^c \) Uncertain.
The predicted and observed C(2')-endo sugar pucker for 2'MeCyd probably reflects the influence of the 2'-O-methyl group. The ribose of cytidine, in the crystal, is in the C(3')-endo–C(2')-exo twist conformation; furthermore, common ribopyrimidine nucleosides generally favor C(3')-endo puckering, although C(2')-endo does occur.

CONCLUSION

Classical potential energy calculations show that 2'-O-methylation of cytidine does not alter the commonly preferred conformational regions for χ and Ψ, namely anti and gg, respectively. However, a high probability for C(2')-endo sugar puckering is predicted as a consequence of this modification in the nucleoside structure. The methoxy angles m₁ and m₂ which are unique to these unusual nucleosides are both permitted the 60° range, and m₁ can adopt the 180° conformation as well.

The authors thank Prof. R. M. Wartell, School of Physics, Georgia Institute of Technology, for stimulating discussions. The molecular modeling was performed at the Princeton University Computer Graphics Laboratory, supported by NIH Grant RR-00578. Other support was received from NIH GM-16539 and NSF GB-28021.

References


Received January 10, 1975
Accepted April 7, 1975