

## Effect of 2-thiouridine on RNA Conformation

Sarkar A.K.<sup>1</sup>, Sarzynska J.<sup>2</sup>, Lahiri A.<sup>1</sup>

<sup>1</sup>University of Calcutta, Kolkata, India

<sup>2</sup>Institute of Bioorganic Chemistry, Poznan, Poland

[almbg@caluniv.ac.in](mailto:almbg@caluniv.ac.in)

Diverse post-transcriptional modifications are present in different types of RNA, and apparently, the modifications influence RNA function by fine tuning the structure. The 2-thio modification of uridine and its other derivatives occur frequently at the 34 position of the anticodon domain of tRNA and presumably play a role in the discrimination of codons. In an NMR study of a pentanucleotide mimicking part of the anticodon loop of HIV reverse-transcription primer tRNA<sup>Lys</sup>, it was found that the sugar pucker of 2-thiouridine strongly preferred the C3'-endo conformation and altered the preference of the sugar pucker of the neighbouring uridines toward C3'-endo in comparison with the corresponding unmodified pentanucleotide. We have earlier showed that the existing force field parameters for 2-thiouridine in the AMBER molecular modeling suite did not reproduce the pucker distribution at the nucleoside level and proposed a revised set of parameters to address that. Here, we have carried out simulations of the pentanucleotide sequence with and without a 2-thiouridine modification with our revised parameters. The conformational ensemble of the sequence with 2-thiouridine showed the presence of a cluster with a structure closer to the corresponding sequence in the crystal structure of HIV reverse-transcription primer tRNA<sup>Lys</sup> in comparison with the unmodified pentanucleotide. Stacking analysis revealed that the sequence of three uridines is much better stacked when the first uridine has the 2-thio modification. Our results showed that the proposed parameters of 2-thiouridine reliably reproduced a more ordered A-RNA like structure in the presence of the modification.

*Key words: RNA, post-transcriptional modifications, molecular dynamics simulation, force field parameters, AMBER.*

### 1. Introduction

2-thiouridine (s<sup>2</sup>U) is a post-transcriptional modification of uridine and is found predominantly in the wobble position of tRNAs along with its C5-derivatives [1]. It has been shown that the thio modification substantially stabilizes the C3'-endo (north) sugar conformation in the case of s<sup>2</sup>U at the nucleoside and dinucleotide level [2–5]. Although direct evidence for this type of stabilization in tRNAs has not been obtained, the stabilizing properties of s<sup>2</sup>U have been confirmed for oligoribonucleotides by UV, CD and NMR spectroscopic studies on an RNA pentamer [6] whose sequence closely resembled the anticodon of HIV reverse-transcription primer tRNA<sup>Lys</sup> (PDB ID 1FIR) [7]. It has been observed that the RNA pentamer having s<sup>2</sup>U (5'Gps<sup>2</sup>UpUpUpC3') strongly preferred the C3'-endo sugar pucker as compared to its unmodified counterpart (5'GpUpUpUpC3') [6]. The effect of the s<sup>2</sup>U was found to propagate towards the 3' adjacent bases as their sugar pucker preferences were also changed to C3'-endo.

The present work aims to provide microscopic insight into the behaviour of the RNA pentamer with and without s<sup>2</sup>U along with participating in the ongoing efforts to devise and extensively benchmark

high-quality force field parameters for RNA and modified residues in RNA [8].

We have carried out 500 ns long molecular dynamics simulations of the 5'Gps<sup>2</sup>UpUpUpC3' and 5'GpUpUpUpC3' pentanucleotide systems in an explicit solvent model and analyzed their conformational characteristics utilizing the AMBER [9] FF12 force field and force fields incorporating parameters for s<sup>2</sup>U. Although force field parameters were available for a few RNA modifications, the first systematic development of parameters for the entire set of naturally occurring modifications was first attempted by Aduri et al. [8] based on the parameter building strategy in the AMBER [9] suite of molecular modeling program. Subsequently, Xu et al. [10] provided an equivalent set of parameters based on the CHARMM suite of molecular modeling tools. We have earlier demonstrated the inadequacies of the force field parameters developed by Aduri et al. [8] at the nucleoside level for a few modified uridines, which included s<sup>2</sup>U [11]. We have also shown that a revised version of the parameters containing the reoptimized glycosidic torsion parameter for s<sup>2</sup>U resulted in a much improved agreement with experimentally observed conformational distribution [12]. In this work, we have investigated the performance of these parameters in the context of the pentanucleotide (5'Gps<sup>2</sup>UpUpUpC3') and compare

its conformational characteristics with those of the unmodified pentanucleotide (5'GpUpUpUpC3').

## 2. Materials and method

### 2.1. System preparation

The initial structures for the systems under study were produced using the online server <http://structure.usc.edu/make-na/server.html> in the canonical A-RNA form.

Hydrogens were added to the initial structures followed by neutralization and solvation. The TIP3P water model [13] was used to mimic the solvent environment. The individual systems and water molecules were put in a truncated octahedral solvent box. Neutralization was carried out by using Na<sup>+</sup> ions. We chose to use the FF12 [9] for the standard residues and  $\chi$ \_IDRP [12] and Aduri et al. [8] force field parameters individually in combination with the FF12 for the modified residue s<sup>2</sup>U. A total of 500 ns of MD simulations were performed for each of the pentanucleotides (5'Gps<sup>2</sup>UpUpUpC3' and 5'GpUpUpUpC3'). All simulations were performed using the PMEMD module of AMBER 12 and 16 [9].

### 2.2. Molecular dynamics simulations

We considered the distance of 10 Å as the closest distance between any atom of the solute and the edge of the periodic box. Prior to heating, two step energy minimizations were carried out. A 10 Å long-range nonbonded cutoff was used. The first step of minimization comprised of 500 steps of steepest descent followed by 500 steps of conjugate gradient minimization holding the nucleic acid systems restrained with a force of 500 kcal/mol·Å<sup>2</sup>. In the second step of minimization, 1000 steps of steepest descent followed by 1500 steps of conjugate gradient were carried out without any restraint. After minimization, the systems were heated from 0 K to 300 K in 20 ps with 2 fs time step using constant volume dynamics holding the solute with a restraint force of 10 kcal/mol·Å<sup>2</sup>. Langevin dynamics [14] with random velocity scaling with a collision frequency of 1 ps<sup>-1</sup> was used. SHAKE [15] was turned on to constrain bonds involving hydrogen atoms and the particle mesh Ewald (PME) [16] method was used for electrostatic interactions. Constant pressure dynamics for 200 ps with 2 fs time step was carried out at 300 K for initial equilibration. The reference pressure was 1 atm with a barostat relaxation time of 2 ps. During production run the same conditions were followed as in the equilibration step and continued upto 500 ns. Trajectories were saved at each 10 ps time step.

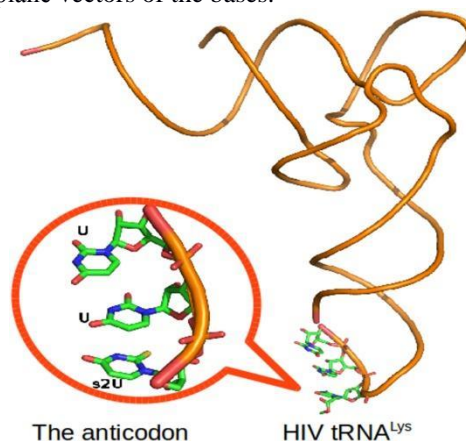
### 2.3. Structural analysis

All heavy atom RMSDs with respect to the structure of the corresponding sequence in the anticodon domain of the tRNA<sup>Lys</sup> [7] was calculated to compare the deviation of the two pentanucleotides

from their canonical conformation. The coordinates corresponding to the anticodon region was extracted from the crystal structure (PDB ID 1FIR) and the hyper-modified uridine (mcm<sup>5</sup>s<sup>2</sup>U) was changed to s<sup>2</sup>U for our case (Fig. 1). The end residues were excluded from subsequent analysis and only the three central uridines, being the residues of the anticodon region, were considered for all structural analysis like RMSD, radius of gyration, sugar pucker and base stacking.

The two state model was chosen to divide the pseudorotation space of sugar pucker into North/C3'-endo (270° ≤ P < 90°) and South/C2'-endo (90° ≤ P < 270°) partitions to facilitate comparison with NMR studies. Pseudorotation angles (P) have been determined following Altona and Sundaralingam [17].

Stacking propensity, in percentage, between the successive bases were calculated according to the stacking criteria employed by Vokacova et al. [18]. We have taken the distance parameter as the distance (R) between the geometric centers of the ring atoms of the bases. The two angular geometric parameters were defined as (i) the angle (Φ1) between the in-plane vectors determined by the atoms N9 and C6 for purines and N1 and C4 for pyrimidines and (ii) the angle (Φ2) between the base normals calculated from the plane vectors of the bases.



**Fig. 1.** Clover leaf representation of the HIV reverse-transcription primer tRNA<sup>Lys</sup>. The anticodon region is highlighted in red circle. The mcm<sup>5</sup>s<sup>2</sup>U residue has been replaced by s<sup>2</sup>U.

## 3. Results and discussion

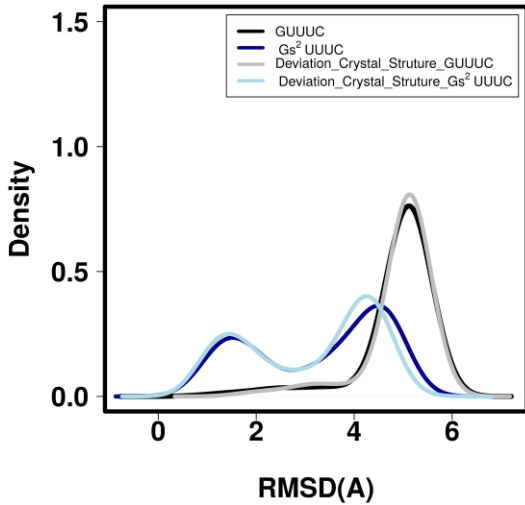
### 3.1. Conformational preferences of s<sup>2</sup>U

The nucleoside s<sup>2</sup>U shows a strong preference for the C3'-endo sugar pucker along with anti conformation for the glycosidic torsion. Unfortunately, the currently available parameter set for s<sup>2</sup>U in the AMBER (Aduri et al.) [8] was earlier shown by us to be insufficient to reproduce experimentally observed conformational distributions in solution [11].

Subsequently, Deb et al. [12] have developed a new set of parameters (called  $\chi$ \_IDRP parameter) for

$s^2U$  and some other modifications of uridine by reoptimizing the glycosidic torsion parameter and additionally showed that fine tuning the van der Waals parameters resulted in even better agreement with experiments.

Since these calculations were carried out for single nucleosides in solution, the relevant question remained how these parameters would behave in the context of larger systems. In this work, we have obtained results from extensive simulations of a pentanucleotide with and without the  $s^2U$  modification that enabled us to carry out adequate sampling while providing a system complex enough to demonstrate the neighbouring-sequence effects of  $s^2U$ .



**Fig. 2.** Distribution of RMSD of 5'GUUUC3' and 5'Gs<sup>2</sup>UUUC3 with respect to the initial structure and crystal structure of the anticodon region of the HIV reverse-transcription primer tRNA<sup>Lys</sup>.

### 3.2. Comparison of sugar pucker between NMR and simulations

In the present study we extended the validation of  $\chi$ \_IDRP force field [12] to the oligoribonucleotide level. Two RNA pentanucleotides, 5'GUUUC3' and 5'Gs<sup>2</sup>UUUC3' were modeled using AMBER FF12 [9] and FF12 +  $\chi$ \_IDRP force field parameters respectively and their dynamics were compared.

As suggested by the NMR observations [6], presence of the  $s^2U$  induces strong preference for north sugar pucker at subsequent uridines in 5'Gs<sup>2</sup>UUUC3' as compared to 5'GUUUC3'. We found a good agreement in sugar pucker distribution between MD simulation and NMR when 5'Gs<sup>2</sup>UUUC3' was simulated with the FF12 +  $\chi$ \_IDRP force field [12] (Table 1). On the other hand, the Aduri et al. parameters [8] were found to be consistent in its inefficiency in reproducing the observed sugar pucker distribution (Table 1). In the following, we report results from simulations with FF12 and FF12 +  $\chi$ \_IDRP force field parameters only.

### 3.3. Effect of $s^2U$ on the ensemble of 5'GUUUC3'

To have an idea how the replacement of the uridine with  $s^2U$  affected the conformational ensemble of the 5'GUUUC3', we have compared the root mean square deviation (RMSD) of all heavy atoms of 5'GUUUC3' with that of 5'Gs<sup>2</sup>UUUC3' (Fig. 2).

While the 5'GUUUC3' showed quite a large deviation from the initial energy minimized structure (RMSD  $\sim$  5 Å), the deviation in the case of 5'Gs<sup>2</sup>UUUC3' was found to be relatively less (Fig. 2) and bimodal with maxima around  $\sim$  1.8 Å and 4.2 Å. This indicated the possibility that  $s^2U$  induced some kind of conformational rigidity that helped to maintain the structure near the A-RNA form at least for a subpopulation of the molecule.

**Table 1:** Comparison of the sugar pucker among the NMR, FF12 +  $\chi$ \_IDRP [11] and FF12 + Aduri et al. [8]

Sequence	Sugar pucker (% North)		
	NMR [6]	$\chi$ _IDRP	Aduri et al.
GU/(s <sup>2</sup> U)UUC	51 <sup>1</sup> , 62 <sup>2</sup>	82 <sup>2</sup>	95 <sup>2</sup>
GU/(s <sup>2</sup> U)UUC	42 <sup>1</sup> , 75 <sup>2</sup>	62 <sup>2</sup>	32 <sup>2</sup>
GU/(s <sup>2</sup> U)UUC	41 <sup>1</sup> , 77 <sup>2</sup>	62 <sup>2</sup>	50 <sup>2</sup>
GU/(s <sup>2</sup> U)UUC	41 <sup>1</sup> , 75 <sup>2</sup>	61	78 <sup>2</sup>
GU/(s <sup>2</sup> U)UUC	59 <sup>1</sup> , 59 <sup>2</sup>	34 <sup>2</sup>	36 <sup>2</sup>

<sup>1</sup>data for 5'GUUUC3'.

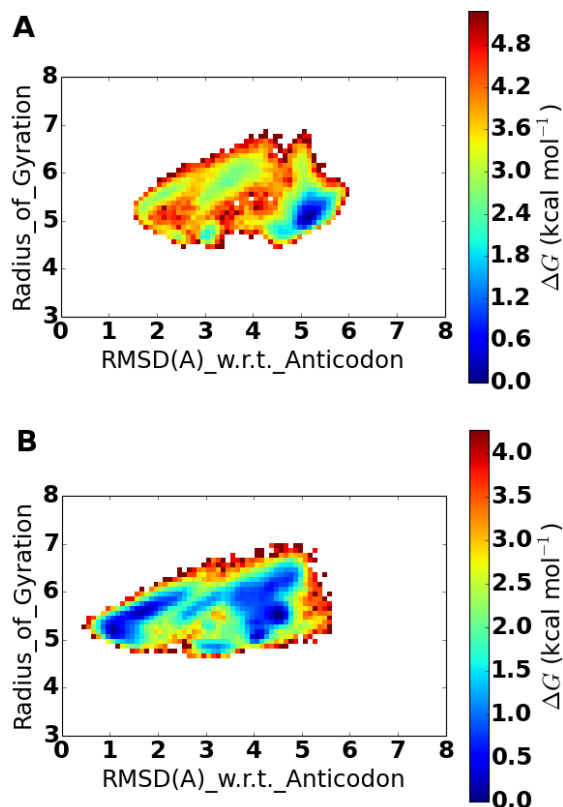
<sup>2</sup>data for 5'Gs<sup>2</sup>UUUC3'.

### 3.4. Importance of $s^2U$ in the anticodon region

To investigate the importance of  $s^2U$  in maintaining the shape of the anticodon region, the RMSD of all heavy atoms was further quantified with respect to the anticodon region of the HIV reverse-transcription primer tRNA<sup>Lys</sup> (PDB ID 1FIR) [7]. As evident from the RMSD distribution plot (Fig. 2), the RNA pentamer containing  $s^2U$  deviated less from the crystal structure as compared to the RNA pentamer without  $s^2U$  (Fig. 2). This clearly indicated that the presence of the  $s^2U$  was quite important to maintain the shape of the anticodon region.

To further probe the shape of the anticodon region in the presence and absence of  $s^2U$ , free energy landscapes (FELs) were constructed based on two reaction coordinates, namely, the radius of gyration and RMSD from the anticodon region (Fig. 3). The FELs were significantly different for 5'GUUUC3' and 5'Gs<sup>2</sup>UUUC3'. Unlike in the case of 5'GUUUC3', the FEL of 5'Gs<sup>2</sup>UUUC3' have two major basins separated by a 2.5 Kcal/mol energy barrier. The basin for the 5'GUUUC3' was observed around RMSD  $\sim$  5.5 Å and radius of gyration  $\sim$  5.5 Å, while, for the 5'Gs<sup>2</sup>UUUC3' the two basins were located at radius of gyration  $\sim$  5.5 Å and RMSDs  $\sim$  1.5 Å and  $\sim$  4.5 Å respectively. The radius of gyration of the anticodon of HIV reverse-transcription primer tRNA<sup>Lys</sup> was found to be  $\sim$  5.5 Å indicating that

central 5'UUU3' or 5'<sup>s</sup>UUU3' of both the RNA pentanucleotides were as extended as the anticodon. However, as observed from the FELs, it seems that the A-RNA like shape was relatively better maintained in 5'Gs<sup>s</sup>UUUC3' for a major subpopulation. Since the A-RNA-like structure is characterized by significant stacking interaction between successive bases, a consideration of the base overlap should provide further insight into the conformational difference between the oligonucleotides.

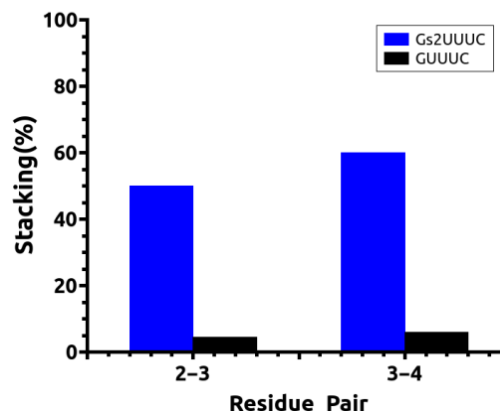


**Fig. 3.** Free energy landscape projected on RMSD from the anticodon region and radius of gyration (in Å) for A. GUUUC and B. Gs<sup>s</sup>UUUC.

### 3.5. Base stacking in maintaining shape of the anticodon region

Our simulations with the  $\chi$ -IDRP force field revealed a relatively strong preference for stacking for the s<sup>2</sup>U/U and U/U pairs in the 5'Gs<sup>s</sup>UUUC3' system as compared to the two U/U pairs in 5'GUUUC3' (Fig. 4). This showed the ability of s<sup>2</sup>U in stabilizing the stacking interaction.

Comparison between the stacking data and the FELs also indicated that although the extended conformation has been maintained in 5'GUUUC3', the central bases were very poorly stacked. On the other hand, a considerable amount of stacking was observed between successive bases in the central 5'<sup>s</sup>UUU3' region of 5'Gs<sup>s</sup>UUUC3' RNA.



**Fig. 4.** The extent of base stacking for the three central nucleotides of 5'GUUUC3' and 5'Gs<sup>s</sup>UUUC3'.

## 4. Conclusion

NMR studies on the two RNA pentanucleotides, 5'GUUUC3' and 5'Gs<sup>s</sup>UUUC3' observed a strong stabilization of the north sugar conformation in s<sup>2</sup>U which propagated along the 3' uridines [6]. This led to the expectation that the presence of the s<sup>2</sup>U may stabilize the A-RNA-like helical and stacked structure necessary in the anticodon region of the anticodon loop.

In our 500 ns long simulations of 5'GUUUC3' and 5'Gs<sup>s</sup>UUUC3' in explicit solvent, we observed that the newly developed  $\chi$ -IDRP force field [12] provided much better agreement with the NMR data as compared to the force field parameters proposed by Aduri et al. [8] and included in the current AMBER distribution for modified RNA residues.

## 5. Acknowledgements

This research work was partially supported by the departmental DST-FIST and UGC-DSA programs. The UGC RFSMS and CSIR SRF programs are also acknowledged for providing fellowship to one of the authors (AKS).

## 6. References

1. Yokoyama S., Nishimura S. In: *tRNA. Structure, Biosynthesis, and Function*. Eds. Soll D., Bhandary R. Washington: ASM Press, 1995. P. 207–224.
2. Yokoyama S., Watanabe T., Murao K., Ishikura H., Yamaizumi Z., Nishimura S., Miyazawa T. *Proc. Natl. Acad. Sci. USA*. 1985. V. 82. P. 4905–4909.
3. Agris P.F., Sierzputowska-Gracz H., Smith W., Malkiewicz A., Sochacka E., Nawrot B. *J. Am. Chem. Soc.* 1992. V. 114. P. 2652–2656.
4. Sierzputowska-Gracz H., Sochacka E., Malkiewicz A., Kuo K., Gehrke C.W., Agris P.F. *J. Am. Chem. Soc.* 1987. V. 109. P. 7171–7177.

5. Smith W.S., Sierzputowska-Gracz H., Sochacka E., Malkiewicz A., Agris P.F. *J. Am. Chem. Soc.* 1992. V. 114. P. 7989–7997.
6. Kumar R.K., Davis D.R. *Nucleic Acids Res.* 1997. V. 25. P. 1272–1280.
7. Bénas P., Bec G., Keith G., et al. *RNA*. 2000. V. 6. P. 1347–1355.
8. Aduri R., Psciuk B.T., Saro P., Taniga H., Schlegel H.B., Santa Lucia J.Jr. *J. Chem. Theory Comput.* 2007. V. 3. P. 1464–1475.
9. Salomon-Ferrer R., Case D.A., Walker R.C. *WIREs Comput. Mol. Sci.* 2013. V. 3. P. 198–210.
10. Xu Y., Vanommeslaeghe K., Aleksandrov A., MacKerell A.D., Nilsson L. *J. Comput. Chem.* 2016. V. 37. P. 896–912.
11. Deb I., Sarzynska J., Nilsson L., Lahiri A. *J. Chem. Inf. Model.* 2014. V. 54. P. 1129–1142.
12. Deb I., Pal R., Sarzynska J., Lahiri A. *J. Comput. Chem.* 2016. V. 37. P. 1576–1588.
13. Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R., Klein M.L. *J. Chem. Phys.* 1983. V. 79. P. 926–935.
14. Davidchack R.L., Handel R., Tretyakov M.V. *The Journal of Chemical Physics.* 2009. V. 130. P. 234101.
15. Ryckaert J.P., Ciccotti G., Berendsen H.J.C. *J. Comput. Phys.* 1977. V. 23. P. 327–341.
16. Darden T., York D., Pedersen L. *J. Chem. Phys.* 1993. V. 98. P. 10089–10092.
17. Altona C., Sundaralingam M. *J. Am. Chem. Soc.* 1972. V. 94. P. 8205–8212.
18. Vokacova Z., Budesinsky M., Rosenberg I., Schneider B., Sponer J., Sychrovsky V. *J. Phys. Chem. B.* 2009. V. 113. P. 1182–1191.