On the Conformational Preferences of 2-Selenouridine and its Derivatives

Pal R.[1], Mishra S.[1], Lahiri A.[1]

[1]*University of Calcutta, Kolkata, India*

albmbg@caluniv.ac.in, lahiri.ansuman@gmail.com

Modifications in RNA occur extensively and elucidating their functional role would be very much useful to obtain insight into RNA structure and function. The 5-position derivatives of 2-seleno-modified uridine have been observed in the wobble position of tRNAs in which they are expected to affect codon-anticodon interaction. An accurate theoretical prediction of the structural and dynamical consequences of such modifications can be facilitated by providing reliable force field parameters for molecular modeling. We show that the currently recommended AMBER parameters do not result in conformational properties that match experimental observations for 2-selenouridine and 5-aminomethyl-2-selenouridine. We also show that reparameterization of the glycosidic torsion by fitting quantum chemically obtained torsional energy profiles can improve the theoretical prediction of the conformational properties significantly.

Key words: RNA modification, wobble decoding, AMBER force field, parameter development.

1. Introduction

The first anticodon (wobble) position of tRNAs is known to be inordinately prone to be post-transcriptionally modified. Substitutions at the 5-position of uridine, 2-thiouridine (S2U) and 2-selenouridine (Se2U) are widely encountered in the wobble position of tRNAs in all the three domains of life. The functional consequences of these modifications are elusive and it is generally expected that elucidation of their structural effect may offer insight into their functional role. S2U and Se2U modifications and their 5-position derivatives are of great interest as they show conformational features that are significantly different from those of their canonical counterpart uridine and its 5-position derivatives. One of the reasons for this difference in conformational propensities lies in the fact that the van der Waals radii of S or Se atoms are much larger than that of O. This structural feature increases the steric interaction with the 2'-hydroxyl group of the sugar. As a result, the conformational preference of the sugar puckering predominantly shifts towards C3' endo character (71 % in 2-thiouridine, and 80 % in 2-selenouridine) compared to that of the unmodified uridine (53 %) [1-5]. It has also been reported that S or Se modification of uridine can significantly alter the interaction with complementary bases in an RNA chain. For example, the canonical uracil base shows Watson-Crick hydrogen bonding with adenine predominantly and also exhibits wobble base pairing with guanine. Interestingly, presence of S or Se at position 2 of the base ring enhances the thermodynamic stability of the base pairing with adenine and restricts the base pairing with guanine [6]. Molecular modeling and molecular dynamics simulations provide us with convenient and inexpensive tools to study the structural and dynamics consequences of these modifications. However, the reliability of the results depends very much on the accuracy of the force field parameters used. Here we report the results from our effort to obtain reliable force field parameters for Se2U and its 5-position derivatives.

2. Computational details

2.1. System preparation

For starting structures, the PDB file deposited by Aduri et al. [7] was adopted. The PDB files present in the database given by Aduri et al. [7] contained structures for 5-methylaminomethyl-2-selenouridine (a derivative of 2-selenouridine). This structure was edited to get the geometry of 2-selenouridine. The bond, angle, and dihedral values were taken from the crystal data deposited by Leszczynska et al. [8].

A total of four geometries were prepared following the protocol given by Yildirim et al. [9]. The conformations corresponding to these four schemes were termed as sc1, sc2, sc3 and sc4. Selective dihedral angles were kept fixed at particular values as described in the protocol by Yildirim et al. [9] in order to achieve the mentioned geometries. Doing this either facilitates or prohibits the base-sugar hydrogen bonding interactions keeping the sugar pucker either in C3'-endo or in C2'-endo conformation.

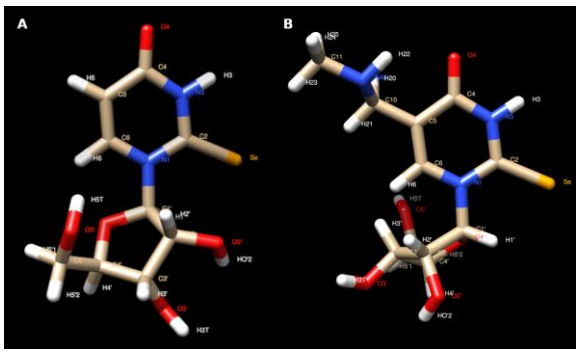

Fig. 1. Systems studied in this work. A. 2-selenouridine (Se2U) and B. 5-methylaminomethyl-2-selenouridine (mnm5Se2U).

2.2. Geometry optimization

After generating the initial geometries of the four schemes of Se2U, a quantum mechanical (QM) geometry optimization was done for each of the structures. The geometries were optimized using the GAUSSIAN09 software suite [10]. HF/6-31G (d) level of theory was used for the geometry optimization calculations. The specific dihedrals that are responsible for the construction of the four different schemes (as mentioned earlier) were kept frozen in their specific values [9] during the optimization process and the chi(χ) dihedral (O4'-C1'-N1-C6) was kept fixed at 0°.

2.3. Glycosidic torsion energy scanning and obtaining Quantum Mechanical (QM) energy profiles

QM optimized geometries corresponding to the four schemes were subjected to a gas phase potential energy surface (PES) scan around the glycosidic torsion angle. The glycosidic torsion χ (O4'-C1'-N1-C6) was given a rotation of total 360° by generating 72 conformations each having glycosidic torsion angle 5° more than its preceding ones. The dihedral angles that constitute the four conformational schemes were kept fixed at their specific values during the potential energy scanning process. Next, QM energy profiles (E_{QM}) were obtained with MP2/6-31G* level of theory for each conformations.

2.4. Calculation of partial charges

Before proceeding to the molecular mechanical calculations, the partial charges were revised using the Restrained Electrostatic Potential (RESP) fitting technique [11] using the RED-vIII.52.pl script of RED server [12]. A Multi-Conformational fitting approach was applied and all the four schemes as described earlier were considered for the calculations. Geometry optimization was done for each of the four conformations prior to the RESP calculations. For each scheme, the lowest energy structure of the QM energy profile was taken. During the calculations, the partial charges of the sugar atoms were restrained and only the partial charges of the base atoms were calculated.

2.5. Molecular Mechanical (MM) energy profiles

The AMBER molecular modeling suite AMBER 12[13] and AMBER 18 [14] were used to calculate the MM energies, E_{MM}, for the 72 QM optimized geometries. The nucleoside was described with the parameters provided in modrna08 [7]. Since the parameters given by Aduri et al (2007)[7] contained parameters for 5-methylamiomethyl 2 selenouridine (a derivative of 2-selenouridine), the parameters related to the bulky group at 5 position was discarded and only related to the Se modification was taken for calculations for Se2U. The parameter-topology (prmtop) files were created by the tleap module. The MM calculations were done with the newly obtained partial charges and keeping the chi (χ) torsion parameters zero. The MM energy minimizations were carried out by restraining the dihedral angles to the values of the optimized QM geometries. A long range cut-off of 8 Å was considered to include the non-bonded interactions during energy minimization in vacuum. The same method was followed during the MM energy calculation with the newly developed chi (χ) torsion parameters for comparison.

2.6. Energy profile fitting

The energy differences (E_{CHI}) between the QM energies (E_{QM}) and MM energies (E_{MM}) represent the potential energy due to glycosidic torsion angle and given by

$$E_{CHI} = E_{QM} - E_{MM} .$$

The 288 E_{CHI} values (72 for each of the 4 schemes) obtained from the above equation were fitted to the Fourier series below

$$E_{CHI} = \sum V_n \left(1 + \cos n\chi\right)$$

where χ is the glycosidic torsion angle, n varies from 1 to 4 and V_n is the potential energy barrier corresponding to the n^{th} term.

2.7. Molecular dynamics simulation and data analyses

The systems were subjected to a two-step minimization process. In the first step, the nucleosides were restrained with a force constant of 500 kcal/mol and only the water molecules were allowed to move. In this stage a total number of 1000 steps were done for minimization among which the first 500 steps followed the steepest descent and the rest followed the conjugate gradient minimization algorithm.

In the second step of the minimization process the restraining force on the nucleosides were removed and the entire systems were allowed to move. This part consisted of a total 2500 steps of minimization among which initial 1000 and the rest followed the steepest descent and the conjugate gradient minimization algorithm respectively. After energy minimization of the systems, the temperature of the systems were raised to 300 K in 20 ps keeping the nucleosides restrained

with a force constant of 10 kcal/mol followed by allowing the system to relax at that temperature for another 200 ps without any restraining force. Finally 16 replicas of the system were equilibrated in 16 different temperature windows (300.0 K, 305.8 K, 311.7 K, 317.8 K, 323.9 K, 330.2 K, 336.6 K, 343.1 K, 349.7 K, 356.5 K, 363.4 K, 370.5 K, 377.6 K, 384.9 K, 392.4 K, 400.0 K) for 1000 ps.

Replica exchange molecular dynamics (REMD) simulations were started with the geometries as starting structures obtained after the equilibrium process. The REMD production runs produced 12 ns of simulation time for each replica and yielded a total of 192 ns of simulation time in aggregate. Langevin dynamics with random velocity scaling with a collision frequency of 1 ps^{-1} was used to propagate the trajectories. The bonds involving the hydrogen atoms were constrained by the SHAKE algorithm [15] and the electrostatic interactions were treated with the particle mesh Ewald (PME) method [16] during all the equilibration steps and the REMD equilibration and production runs. During all the minimization, equilibration and production runs a long-range cut-off of 8 Å was used to include non-bonded interactions.

Table 1. Conformational properties obtained from the simulation data and comparison with available experimental results for Se2U

Properties studied	modrna08+ χ OL3 + parmbsc0	modrna08 + newly obtained χ parm+ parmbsc0	Experimental data[8]
P (%NORTH)	5.05	80.25	80.00
%ANTI of χ dihedral	14.27	99.25	ANTI
% g+, g-, TRANS, OTHER of γ dihedral	63.47, 33.18, 2.77, 0.58	3.15, 21.07, 73.37, 2.42	TRANS

Conventions suggested by Saenger [17] were followed for all atoms and dihedrals nomenclatures. To validate the simulation results with the NMR data for the equilibrium distribution of the pseudorotation angle (P) it was divided into C3'-endo or NORTH ($270^o \leq P < 90^o$) and C2'-endo or SOUTH ($90^o \leq P < 270^o$) sugar puckering regions as reported by Altona and Sundaralingam [18] and Foloppe and Nilsson [19].

The angular ranges 170°−300° and 30°−90° for the χ torsion angles with respect to O4′−C1′−N1−C2 correspond to the ANTI and SYN conformations, respectively. The values beyond these ranges were referred to as OTHERS. In the case of γ torsion angle, the conformational space with respect to the O5′−C5′−C4′−C3′ was divided into the g^+ (60° ± 30°), g^- (300° ± 30°), trans (180° ± 30°) and others. The ranges were considered as described by Foloppe and Nilsson [19]. The calculations were done using the *cpptraj* [20] utility of AMBERtools [21] and in-lab developed scripts.

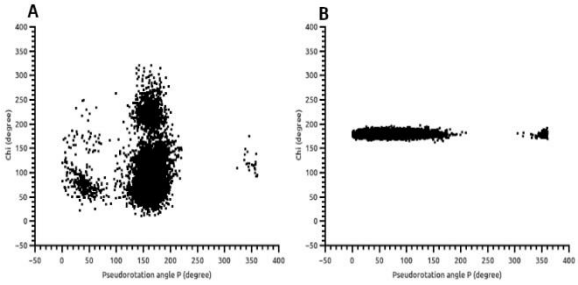

Fig. 2. Scatter plot of the Pseudorotation angle (P) vs Chi (χ) dihedral of Se2U from the data obtained from REMD simulations using A. modrna08 + χ OL3 + parmbsc0 and B. modrna08 + newly obtained chi (χ) parms + parmbsc0.

Table 2. Conformational properties obtained from the simulation data and comparison with available experimental results for mnm5Se2U

Properties studied	modrna08 + χ OL3+ parmbsc0	Experimental data [8]
P (%NORTH)	4.47	72.00
%ANTI of χ dihedral	13.67	–
% g+, g-, TRANS, OTHER of γ dihedral	56.43, 38.25, 4.97, 0.35	–

3. Results and discussion

The results of the conformational properties for Se2U from the REMD simulations using the force field parameter combinations of modrna08 [7]+ χ OL3 [22]+ parmbsc0 [23] and modrna08 [7]+newly obtained chi (χ) parm + parmbsc0 [23] along with available experimental data [8] are given in table 1. It is clear that the combination of modrna08+ χ OL3+parmbsc0 failed to reflect the experimental observations.

The newly obtained revised χ parameters for Se2U were also validated against the experimental data to check whether any improvement is achieved or not. It is clear from the observations listed in table 1 that the newly revised χ parameters in combination with modrna08 and parmbsc0 performed reasonably well to reflect the experimental observations.

Figure 2 shows the scatter plots from the REMD simulations of Se2U with the two force fields. The coupled distribution of the pseudorotation angle and the glycosidic torsion angle does not conform to the experimentally observed conformational preference for the NORTH, ANTI region in the case of the currently recommended AMBER force field combination (figure

2A). The distribution became much more constrained and corresponded much more closely to experimental observations when our reoptimized parameters were used.

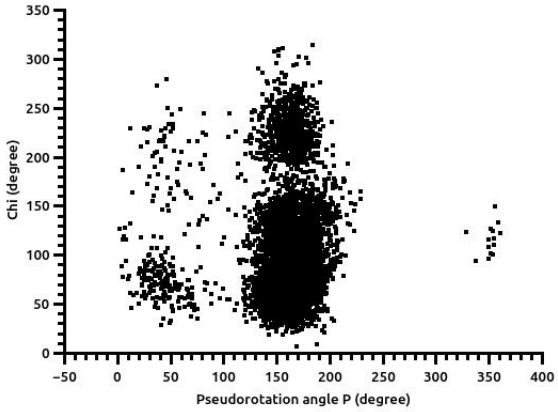

Fig. 3. Scatter plot of the Pseudorotation angle (P) vs Chi (χ) dihedral of mnm5Se2U from the data obtained from REMD simulations using modrna08 + χ OL3 + parmbsc0.

In table 2 and figure 3 we have reported the results from the REMD simulation of mnm5Se2U using the currently recommended force field combination in AMBER for modified RNA residues. As is quite clear, neither the average value of P, nor the correlated distribution of the pseudorotation angle and the glycosidic torsion corresponded to the expected NORTH, ANTI conformation.

In the foregoing we have presented preliminary results from our ongoing investigation on the conformational preferences of Se2U and one of its 5-position derivatives, namely, mnm5Se2U. We showed that, use of the currently recommended force field parameters available from the AMBER distribution leads to a gross underestimation of the propensity of the sugar pucker for the NORTH conformation when compared with experimental data. Also, the propensity for the ANTI conformation was found to be quite low although, at least for Se2U, experimental results show an ANTI conformation. We therefore reparametrized the modified residues following our previously tested protocol [24] and observed a significant improvement in the reproduction of experimental observations for Se2U. The validation of the new parameters of mnm5Se2U is ongoing along with the development of parameters for a few other naturally occurring 5-position derivatives of Se2U.

3. Acknowledgments

Authors thank the high performance computing center, Department of Microbiology, University of Calcutta for providing computational resources. RP would like to thank CSIR for providing fellowship (09/028(1041)/2018-EMR-1). SM thanks UGC for providing fellowship (NTA Ref. No.:191620022802). Funding from SERB (EMR/2016/007753) is gratefully acknowledged.

4. Reference

1. Sierzputowska-Gracz H., Sochacka E., Malkiewicz A. et al. *Journal of the American Chemical Society*. 1987. V. 109. P. 7171–7177.
2. Testa S.M., Disney M.D., Turner D.H., Kierzek R. *Biochemistry*. 1999. V. 38. P. 16655–16662.
3. Bartos P., Ebenryter-Olbinska K., Sochacka E., Nawrot B. *Bioorganic & Medicinal Chemistry*. 2015. V. 23. P. 5587–5594.
4. bo Zhang R., Eriksson L.A. *Physical Chemistry Chemical Physics*. 2010. V. 12. P. 3690–3697.
5. Kogami M., Davis D.R., Koketsu M. *Heterocycles*. 2016. V. 92. P. 64–74.
6. Sun H., Sheng J., Hassan A.E.A., Jiang S., Gan J., Huang Z. *Nucleic Acids Research*. 2012. V. 40. P. 5171–5179.
7. Aduri R., Psciuk B.T., Saro P., Taniga H., Schlegel H.B., SantaLucia J. *Journal of Chemical Theory and Computation*. 2007. V. 3. P. 1464–1475.
8. Leszczynska G., Cypryk M., Gostynski B. et al. *International journal of molecular sciences*. 2020. V. 21. P. 2882.
9. Yildirim I., Stern H.A., S.D. Kennedy, Tubbs J.D., Turner D.H. *Journal of Chemical Theory and Computation*. 2010. V. 6. P. 1520–1531.
10. Frisch M.J.E.A., Trucks G.W., Schlegel H.B. et al. *Gaussian 09, revision D. 01*. 2009.
11. Bayly C.I., Cieplak P., Cornell W., Kollman P.A. *The Journal of Physical Chemistry*. 1993. V. 97. P. 10269–10280.
12. Vanquelef E., Simon S., Marquant G. et al. *Nucleic Acids Research*. 2011. V. 39. Iss. suppl_2. P. W511–W517.
13. Case D.A., Darden T., Simmerling C., Cheatham T. *AMBER 12*. San Francisco: University of California, 2012.
14. Case D.A., Walker R.C., Cheatham T.E. III et al. *AMBER 18*. San Francisco: University of California, 2018.
15. Ryckaert J.-P., Ciccotti G., Berendsen H.J.C. *Journal of Computational Physics*. 1977. V. 23. P. 327–341.
16. Darden T.D.Y., Lee P. *The Journal of Chemical Physics*. 1993. V. 98. P. 10089–10092.
17. Saenger W. In: *Principles of nucleic acid structure*. New York: Springer, 1984. P. 9–28.
18. Altona C.T., Sundaralingam M. *Journal of the American Chemical Society*. 1972. V. 94. P. 8205–8212.
19. Foloppe N., Nilsson L. *The Journal of Physical Chemistry B*. 2005. V. 109. P. 9119–9131.
20. Roe D.R., Thomas E.C. III. *Journal of Chemical Theory and Computation*. 2013. V. 9. P. 3084–3095.

21. Case D.A., Walker R.C., Cheatham T.E. III et al. *AMBER 2021*. San Francisco: University of California, 2021.
22. Zgarbová M., Otyepka M., Šponer J. et al. *Journal of Chemical Theory and Computation.* 2011. V. 7. P. 2886–2902.
23. Pérez A., Marchán I., Svozil D. et al. *Biophysical Journal.* 2007. V. 92. P. 3817–3829.
24. Deb I., Pal R., Sarzynska J., Lahiri A. *Journal of Computational Chemistry.* 2016. V. 37. P. 1576–1588.