



Data-informed reparameterization of modified RNA and the effect of explicit water models: application to pseudouridine and derivatives

Nivedita Dutta¹ · Indrajit Deb¹ · Joanna Sarzynska² · Ansuman Lahiri¹

Received: 29 June 2021 / Accepted: 4 March 2022 / Published online: 26 March 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

Pseudouridine is one of the most abundant post-transcriptional modifications in RNA. We have previously shown that the FF99-derived parameters for pseudouridine and some of its naturally occurring derivatives in the AMBER distribution either alone or in combination with the revised γ torsion parameters (parmbsc0) failed to reproduce their conformational characteristics observed experimentally (Deb et al. in *J Chem Inf Model* 54:1129–1142, 2014; Deb et al. in *J Comput Chem* 37:1576–1588, 2016; Dutta et al. in *J Chem Inf Model* 60:4995–5002, 2020). However, the application of the recommended bsc0 correction did lead to an improvement in the description not only of the distribution in the γ torsional space but also of the sugar pucker distributions. In an earlier study, we examined the transferability of the revised glycosidic torsion parameters (χ_{IDRP}) for Ψ to its derivatives. We noticed that although these parameters in combination with the AMBER FF99-derived parameters and the revised γ torsional parameters resulted in conformational properties of these residues that were in better agreement with experimental observations, the sugar pucker distributions were still not reproduced accurately. Here we report a new set of partial atomic charges for pseudouridine, 1-methylpseudouridine, 3-methylpseudouridine and 2'-O-methylpseudouridine and a new set of glycosidic torsional parameters (χ_{ND}) based on chosen glycosidic torsional profiles that most closely corresponded to the NMR data for conformational propensities and studied their effect on the conformational distributions using REMD simulations at the individual nucleoside level. We have also studied the effect of the choice of water model on the conformational characteristics of these modified nucleosides. Our observations suggest that the current revised set of parameters and partial atomic charges describe the sugar pucker distributions for these residues more accurately and that the choice of a suitable water model is important for the accurate description of their conformational properties. We have further validated the revised sets of parameters by studying the effect of substitution of uridine with pseudouridine within single stranded RNA oligonucleotides on their conformational and hydration characteristics.

Keywords RNA modification · Pseudouridine · Glycosidic torsion · Conformational characteristics · Hydration pattern · ssRNA structure

Introduction

Post-transcriptional modifications have been known to be crucial in the regulation of the structure, stability and function of RNA molecules. The MODOMICS database currently lists 172 such modifications [1]. Pseudouridine

(Ψ) was the first post-transcriptional modification discovered [2–4] and is one of the most abundant modifications. Pseudouridine, an isomer of uridine (U), was identified as 5-ribosyluracil and was called the fifth nucleoside [5–8]. This modified residue contains a C–C base-sugar bond, i.e., in the case of pseudouridine, the uracil base is attached to the sugar by a C1'–C5 bond unlike the C1'–N1 glycosidic linkage found in uridine (Fig. 1a). Hence, in contrast to uridine, pseudouridine contains an additional ring nitrogen atom (N1 imino atom) which acts as an additional hydrogen bond donor and is found to be protonated at physiological pH [3, 9].

Pseudouridine was reported to be the most commonly observed modification in the stable RNAs, i.e., tRNA, rRNA

✉ Ansuman Lahiri
albmbg@caluniv.ac.in

¹ Department of Biophysics, Molecular Biology and Bioinformatics, University of Calcutta, 92, Acharya Prafulla Chandra Road, Kolkata, West Bengal 700009, India

² Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

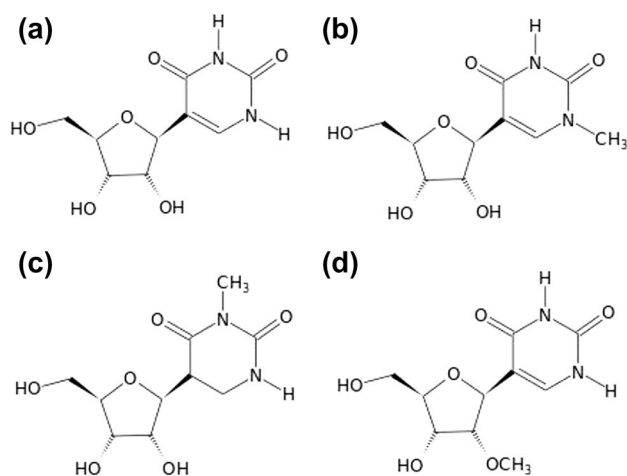


Fig. 1 Structures of **a** pseudouridine, Ψ (PSU); **b** 1-methylpseudouridine, $m^1\Psi$ (1MP); **c** 3-methylpseudouridine, $m^3\Psi$ (3MP); and **d** 2'-O-methylpseudouridine, Ψ_m (MRP)

and snRNA [3]. Further studies involving high-throughput sequencing methods and transcriptome mapping revealed the abundance of pseudouridine as an epigenetic modification, i.e. in mRNA as well as in long noncoding RNA (lncRNA) [10–14]. Several experimental and theoretical studies suggest the important contribution of pseudouridine to the structure, dynamics and thermal stability of RNA [15–21]. This modification has been found to reduce the motion of the neighbouring bases, stabilize the C3'-endo conformation and enhance the stability and the stacking propensity in a context-dependent manner [15, 20–23]. Newby and Greenbaum studied the interaction between Ψ and water in the pre-mRNA branch-site helix and reported that a water- Ψ HN1 hydrogen bond contributes to the stabilization of the unique observed architectural features of this helix [18].

In 2016, we reported that the reoptimized set of glycosidic torsion parameters (χ_{IDRP}) for pseudouridine developed by us, were sufficient to improve the description of the conformational distribution of the glycosidic torsion space but the description of the sugar pucker distribution for Ψ was still not accurate [24]. In another study in 2020, we checked the transferability of these parameters (χ_{IDRP}) to the derivatives of Ψ and observed that the χ_{IDRP} parameters combined with the AMBER FF99-derived parameters [25] and the revised set of γ torsional parameters predicted the conformational properties of these residues which were in general agreement with the experimental (NMR) data but failed to describe the sugar pucker distributions accurately [26].

In the present study we report a new set of partial atomic charges for pseudouridine (Ψ), 1-methylpseudouridine ($m^1\Psi$), 3-methylpseudouridine ($m^3\Psi$) and 2'-O-methylpseudouridine (Ψ_m) and a new set of glycosidic torsional

parameters (χ_{ND}) (Fig. 1). We have compared the results obtained with these parameters with those previously obtained with the FF99 parameters and the FF99 parameters in combination with the χ_{IDRP} parameters and bsc0 γ torsional parameters.

In the earlier studies, multiple schemes [27] and/or general schemes [28] were chosen for the quantum mechanical scan and the molecular mechanical energy profiles were fitted with those with the objective that the re-optimized parameters will be able to explore, preferentially, any of the four quadrants (NORTH/*syn*, NORTH/*anti*, SOUTH/*syn*, SOUTH/*anti*) of the conformational preferences. In the present work, we calculated the quantum mechanical glycosidic torsional energy profiles for five different initial conformations. Then a particular scheme was chosen which outperformed other schemes in reproducing QM profile that was in agreement with the experimentally observed conformational preference. Next, the MM profile was fitted to the chosen QM profile. Additionally, the partial charges were newly generated at the individual modification level before generating the MM profile to incorporate the effect of electrostatic interactions. As a proof of concept, we have chosen pseudouridine and three of its derivatives as a (small) closely related test set that includes molecules with different chemical moieties. For additional validation of our parameter sets, we examined their performance in predicting the conformational and hydration characteristics of the ssRNA trimers and tetramers containing pseudouridine.

Finally, the parameters developed in a targeted approach were further tested in different explicit water models to get insight into the correlation between the solvation model and conformational preferences. It has been reported in recent studies that the choice of water model has a significant impact on the predicted RNA structure and dynamics [29, 30]. Kùhrova et al., based on their study involving the simulation of canonical A-RNA duplexes using explicit water models, i.e. TIP3P [31], TIP4P/2005 [32], TIP5P [33] and SPC/E [34], reported that the TIP5P water model was not found to be optimal for simulating RNA systems [29]. Here, we have investigated the impact of the choice of explicit water models on the conformational characteristics and hydration pattern of Ψ , $m^1\Psi$, $m^3\Psi$, and Ψ_m .

Methods

Ab initio potential energy surface (PES) scan

Preparation of the initial geometries

For the initial geometries of the modified nucleosides Ψ (PSU), $m^1\Psi$ (1MP), $m^3\Psi$ (3MP), and Ψ_m (MRP), we have used the mean values for bonds, angles and dihedral

angles corresponding to the ribose sugar following Gelbin et al. (1996) [35] and considered planar geometries for the bases. The three-letter codes of the modified residues are according to Aduri et al. (2007) [25]. These structures were prepared using the molecular structure editor MOLDEN [36]. The geometries of the modified nucleosides were kept either in the C3'-endo/g⁺ conformation or in the C2'-endo/g⁺ conformation and for that the corresponding torsional angles were fixed at definite values. The value of the γ dihedral angle (O5'-C5'-C4'-C3') was fixed at 54° (which corresponds to the g⁺ conformation) as observed in the A-form RNA [37]. To compel the nucleoside geometries to stay in the C3'-endo conformation, the values of the δ (C5'-C4'-C3'-O3') and O4'-C1'-C2'-C3' dihedral angles were fixed at 81° and -24°, respectively. To constrain the geometries to the C2'-endo sugar pucker conformation the value of the δ (C5'-C4'-C3'-O3') and O4'-C1'-C2'-C3' dihedral angles were set to 140° and 32° respectively. Five initial geometries, i.e., SC1, SC2, SC3, SC4 and SC5 (Table S1) with constrained values of the H5T-O5'-C5'-C4' and C1'-C2'-O2'-HO2' torsional angles were prepared for each of the modified nucleosides, to either promote or restrict the base-sugar hydrogen bonding interactions by maintaining the nucleosides either in C3'-endo or in C2'-endo sugar pucker conformation. The schemes SC1–SC4 were chosen following the values of the torsional angles corresponding to the four schemes chosen in Yildirim et al. [27] and SC5 was chosen based on the *syn* scheme as mentioned in Deb et al. [24]. SC4 also corresponds to the *anti* scheme as mentioned in Deb et al. [24]. For the SC4 conformational scheme, the H5T-O5'-C5'-C4' and C1'-C2'-O2'-HO2' dihedrals were respectively constrained to 174° and 93° and due to that the O5'-H...O4 base-sugar hydrogen bonding interaction is restricted and O2'-H...O4 base-sugar hydrogen bonding interaction is facilitated and hence the geometries corresponding to PSU and its derivatives are compelled towards *anti* conformation which is not the predominant conformation for these nucleosides. For the SC5 scheme, the values of the H5T-O5'-C5'-C4' and C1'-C2'-O2'-HO2' dihedrals were respectively constrained to 60° and -153° to promote the O5'-H...O4 and restrict the O2'-H...O4 base-sugar hydrogen bonding interactions and hence to force a *syn* conformation which is predominant for PSU and its derivatives [38]. The SC1 and SC2 conformational schemes were kept in the C2'-endo conformation while SC3–SC5 were kept in the C3'-endo conformation. To prevent any hydrogen bonding interaction between H3T or O2' and base, so that these interactions cannot affect the glycosidic torsion energy profile, the C4'-C3'-O3'-H3T torsion was fixed at -148° for all the initial geometries. The initial structures corresponding to each of the five conformational schemes are shown in Fig. S1. The geometry

which corresponds to the SC5 conformational scheme for each of the modified nucleosides (along with the atom names) is shown in Fig. S2.

Quantum mechanical scan

All the quantum mechanical calculations were performed using the GAUSSIAN09 software suite [39]. For all the five initial geometries for each of the modified nucleosides, a gas phase PES scan was executed around the glycosidic torsion angle (O4'-C1'-C5-C6) with an increase in its value by 5° resulting in 72 conformations for each nucleoside geometry. Optimization of the structures, during the PES scan, was carried out using the HF/6-31G* level of theory. During the geometry optimization step, the dihedral angles mentioned in Table S1, were kept frozen with the objective of obtaining a smooth QM energy profile. The QM energies (E_{QM}) corresponding to each of the 72 geometry optimized conformations (for each scheme) were calculated using the MP2/6-31G* level of theory. Out of the five quantum mechanical energy (E_{QM}) profiles around χ , we have chosen one particular conformational scheme, i.e. SC5, because the lowest energy minimum for this scheme corresponded to the *syn* region of the glycosidic torsional space (Fig. S3) and experimental (NMR) studies for pseudouridine and its derivatives, under study, reported a preference for the *syn* conformation [54–56]. Additionally, the value of energy corresponding to the global minimum of that profile was found to be the least compared to other schemes (Fig. S3).

RESP fitting

The new set of partial atomic charges for each of the modified nucleosides was developed corresponding to the lowest energy conformation of the quantum mechanical energy profile of the chosen scheme, i.e. SC5, by RESP [40, 41] fitting (Restrained Electrostatic Potential fitting) method using the R.E.D. version III.52 perl program [42]. The partial atomic charges for the atoms of each of the nucleosides are listed in the supporting information (Table S2).

Molecular mechanical (MM) energy minimization

For the calculation of the molecular mechanical (MM) energies (E_{MM}) corresponding to the 72 quantum mechanically (QM) optimized geometries, we have used the AMBER16 software package [43] (Fig. 2). During the MM energy minimizations, the dihedral angles (as mentioned in Table S1) were restrained to the values corresponding to the QM optimized geometries by applying a force constant of 1500 kcal/mol Å². The starting structures for the MM energy minimization step were the structures equivalent to the QM optimized geometries obtained from the PES scan. The

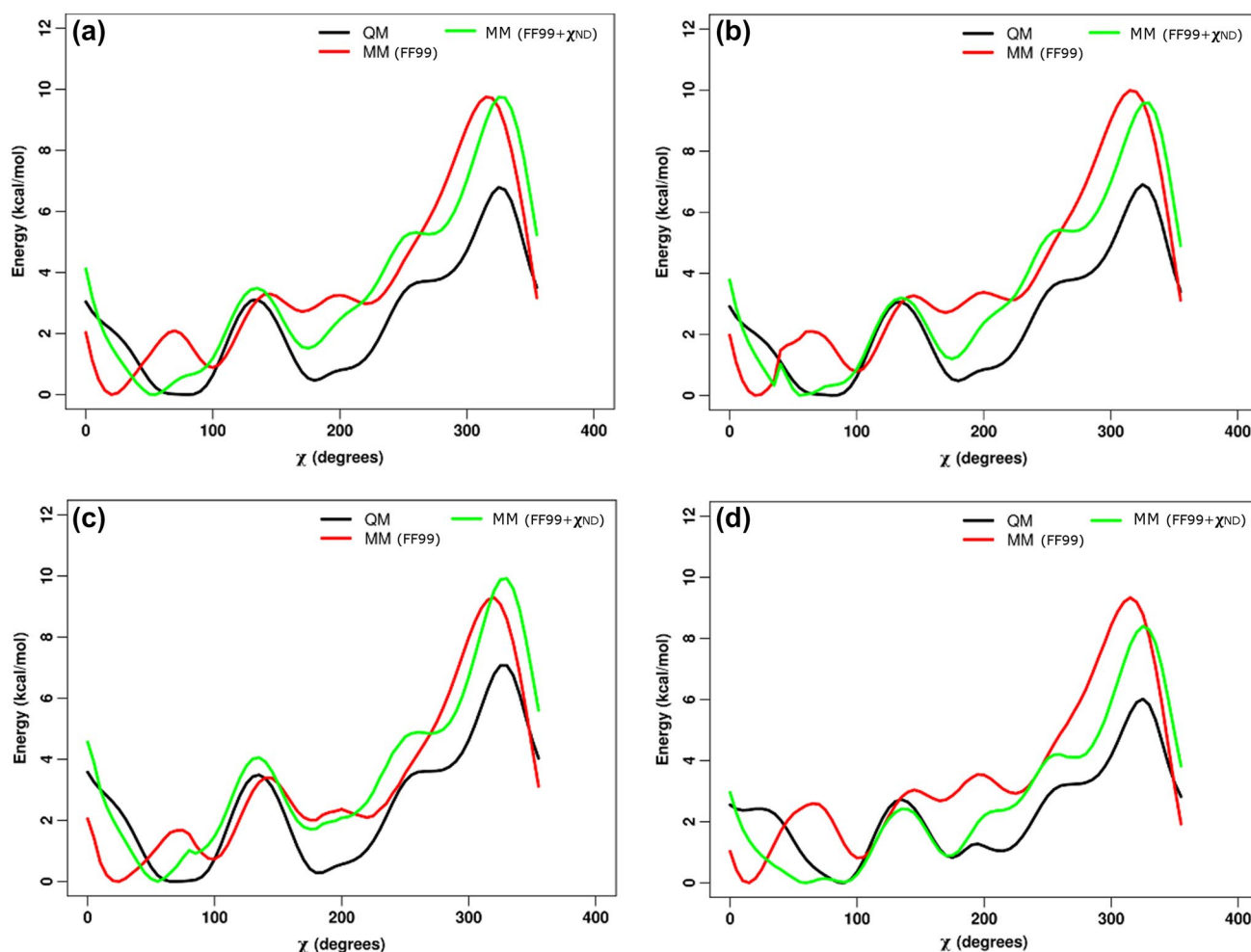


Fig. 2 Energy profiles of the χ torsional angles (O4'-C1'-C5-C4) for **a** Ψ (PSU), **b** $m^1\Psi$ (1MP), **c** $m^3\Psi$ (3MP) and **d** Ψ_m (MRP) residues corresponding to QM calculations (black), MM calculations with the FF99 parameter sets keeping the glycosidic torsion parameters zero (red) and MM calculations with the FF99 parameter sets combined with the newly derived χ torsional parameters and the newly devel-

oped partial atomic charges (FF99- χ_{ND}) (green) by fitting the difference between the QM and MM energies. The minimum energies were set to zero for convenience. The ranges 30°–90° and 170°–300° for the χ torsional angles along the X-axis, correspond to the *syn* and *anti* base orientations respectively

5'-phosphate group was replaced with a hydrogen (5'-OH) and a hydrogen atom (3'-OH) was added to the 3' end of the original topology provided by Aduri et al. [25] to create the topologies for all the modified nucleosides used in this study with the parameters corresponding to the 5'-OH and 3'-OH groups taken from the FF99 force field parameter set [44]. During the MM energy minimization, all the glycosidic torsion parameters corresponding to the Aduri et al. [25] parameter set were set to zero for all the modified nucleosides. Minimizations were carried out using the steepest descent method followed by the conjugate gradient method in order to obtain a smooth glycosidic torsional energy profile for each residue. To incorporate the non-bonded interactions during the energy minimization in vacuum, a long range cut-off of 12 Å was used.

Fitting χ torsion potentials

The potential energy due to the glycosidic torsion angle is represented by the difference (E_{CHI}) between the QM energy (E_{QM}) and MM energy (E_{MM}) and is given by the following equation:

$$E_{CHI} = E_{QM} - E_{MM} \quad (1)$$

The 72 values for E_{CHI} obtained from Eq. (1) were fitted to the Fourier series as shown in Eq. (2):

$$E_{CHI} = \sum_{n=1}^4 [V_n \{1 + \cos(n\chi - \phi_n)\}] \quad (2)$$

where χ represents the glycosidic torsion angle, i.e. the dihedral around (O4'–C1'–C5–C6) and V_n represents the potential energy barrier around the glycosidic torsion angles (χ) and ϕ_n is the phase angle.

Replica exchange molecular dynamics simulations of the modified nucleosides

System preparation

The starting structures in this study were taken from the original PDB format files for each of the four modified ribonucleoside residues corresponding to their quantum mechanically optimized geometries provided by Aduri et al. [25], and available in the AMBER 2018 package. These initial structures were in a NORTH/anti/g⁺ conformation. The FF99_χ_{IDRP}_bsc0 [24] parameter set for Ψ was obtained from Deb et al. [24], and FF99_χ_{IDRP}_bsc0 [24] parameter sets for m¹Ψ, m³Ψ, and Ψm residues were obtained from Dutta et al. [26]. The FF99_χ_{ND}_bsc0 parameter sets for Ψ, m¹Ψ, m³Ψ, and Ψm residues were prepared by combining our newly derived χ torsional parameters (χ_{ND}) and the revised γ parameters developed by Pérez et al. [45] (parmbsc0) with the required bond, angle and torsional parameters for each modification from the AMBER provided parameters derived from Aduri et al. parameters [25]. The revised γ torsional parameters were incorporated by replacing the atom type that described the terms corresponding to the γ torsion in the default topology files with the torsional terms provided in the revised parmbsc0 force field. The newly developed partial atomic charges for the atoms (as mentioned in the supporting information) of each of the four modified ribonucleosides were introduced replacing the partial atomic charges of these atoms in the preparatory file (prepin) provided by Aduri et al. [25]. We used these revised parameter sets for energy minimization and MD simulation

steps. The revised force field parameter sets for Ψ, m¹Ψ, m³Ψ, Ψm (FF99_χ_{ND}_bsc0) are given in the supporting information. The modified ribonucleosides Ψ, m¹Ψ, m³Ψ, and Ψm were separately simulated using the FF99_χ_{IDRP}_bsc0 and FF99_χ_{ND}_bsc0 parameters respectively. Detailed description of the force field parameters used in this study are provided in Table 1. The newly derived glycosidic (χ) torsion parameters are listed in Table 2.

REMD simulations

All replica exchange molecular dynamics (REMD) simulations [46] were performed using the multi-sander approach in AMBER 16 [43] in explicit water. To study the effect of the water model on the conformations of these nucleosides,

Table 2 Revised χ torsion parameters for Ψ, m¹Ψ, m³Ψ, and Ψm

Modified nucleosides	Torsional angle	n	V _n	φ
Ψ (PSU)	O4'–C1'–C5–C6	1	1.00081	180.738622
		2	0.757522	0.758978
		3	1.41029	180.459801
		4	0.340171	1.19649
m ¹ Ψ (1MP)		1	0.761663	180.683487
		2	0.660492	0.713806
		3	1.4897	180.419626
		4	0.347886	1.10914
m ³ Ψ (3MP)		1	0.751305	180.505381
		2	0.871664	0.801897
		3	1.33006	180.466825
		4	0.402064	1.14129
Ψm (MRP)		1	1.06577	180.37802
		2	0.872603	0.711823
		3	1.52327	180.474343
		4	0.322335	1.07613

Table 1 Relevant details of the force fields used for the modified nucleosides in this study

Force fields	Applied revised parameters for torsions	Definition of the applied revised torsional terms
FF99	None	AMBER provided parameters for Ψ, m ¹ Ψ, m ³ Ψ, Ψm nucleosides developed by Aduri et al. [25]
FF99_bsc0	γ	AMBER provided parameters for Ψ, m ¹ Ψ, m ³ Ψ, Ψm nucleosides developed by Aduri et al. [25] in combination with revised γ torsion parameters developed by Pérez et al. [45] (parmbsc0)
FF99_χ _{IDRP} _bsc0	χ and γ	For Ψ, FF99_χ _{IDRP} _bsc0 parameters obtained from Deb et al. [24] and for its three derivatives (m ¹ Ψ, m ³ Ψ, and Ψm), FF99_χ _{IDRP} _bsc0 parameters [24, 25, 45] modified by the introduction of required bond, angle and torsional parameters for each modification from the AMBER provided parameters derived from Aduri et al. parameters [25] (obtained from Dutta et al. [26])
FF99_χ _{ND} _bsc0	χ and γ	Revised glycosidic torsion parameters (χ _{ND}) for Ψ, m ¹ Ψ, m ³ Ψ, and Ψm nucleosides and revised γ torsion parameters developed by Pérez et al. [45] (parmbsc0) in combination with the required bond, angle and torsional parameters for each modification from the AMBER provided parameters derived from Aduri et al. parameters [25] along with the newly developed set of partial atomic charges for each of these modified nucleosides

REMD simulations were carried out using the combination of the FF99- χ_{IDRP} -bsc0 and FF99- χ_{ND} -bsc0 force fields with each of the TIP3P [31], TIP4P-Ew [47] and SPC/E [34] water models and the hydration patterns for pseudouridine and its three derivatives corresponding to the different force field-water model combinations were analyzed. The modified nucleoside residues Ψ , $m^1\Psi$, $m^3\Psi$, Ψm were solvated with TIP3P or TIP4P-Ew or SPC/E water molecules in truncated octahedral boxes with a closest distance of 9 Å between any solute atom and the edge of the box. The computational details of the energy minimization, equilibration and REMD production runs are provided in the supporting information ('METHODS' section).

Molecular dynamics simulations of single stranded RNA (ssRNA) oligomers

System preparation

The initial model structures of the single stranded RNA trinucleotides AUA and A Ψ A and tetra nucleotides AAUA and AA Ψ A were prepared in the canonical A-RNA form using AMBERTools18. Topology generation, addition of hydrogens, neutralization and solvation of each system with TIP3P [31] water molecules in a cubic solvent box, were carried out using AmberTools18. The closest distance between any solute atom and the edge of the periodic box was set to 10 Å. The systems were neutralized using Na⁺ ions.

MD simulations

All molecular dynamics (MD) simulations were performed using AMBER16 software suite [43]. MD simulations were performed using the FF99- χ_{OL3} -bsc0 [25, 28, 45] force field parameter combinations for the canonical nucleosides (adenosine (A) and uridine (U)) and the FF99- χ_{IDRP} -bsc0 or FF99- χ_{ND} -bsc0 parameters for pseudouridine (Ψ). The computational details of the molecular dynamic simulations of the single stranded RNA oligomers are provided in the supporting information ('METHODS' section).

Analysis of the conformational ensembles of the modified nucleosides

For the analysis of the simulated ensembles we calculated the distribution of sugar pucker conformations, distribution of the *syn* or *anti* conformations of the glycosidic torsion angle (χ) and the distribution of the γ torsional angle over different conformational states.

The convention followed for the atom names and the dihedral angle nomenclatures was as given in Saenger [38]. The magnitude of the pseudorotation angle was calculated following Altona and Sundaralingam [48].

The pseudorotation angular space was divided into C3'-endo/NORTH ($270^\circ \leq P < 90^\circ$) and C2'-endo/SOUTH ($90^\circ \leq P < 270^\circ$) regions of sugar puckering [49], which allowed us to directly compare simulated conformational distributions and the equilibrium distributions of the pseudorotation angle (P) as reported in the NMR data.

In our analysis, the χ torsional angle is defined by the atoms O4'-C1'-C5-C4 (for all the modified nucleosides) and was considered to be in the *anti* conformation if its magnitude was within the angular range of 170° – 300° and in the *syn* conformation if it was within the angular range of 30° – 90° [35, 50, 51]. The values that were beyond these ranges were referred to as others [35, 50, 51].

For the calculation of the γ torsional angle, the conformational space with respect to the torsional angle consisting of the atoms O5'-C5'-C4'-C3' was divided into the conformations g^+ (for $60^\circ \pm 30^\circ$), g^- (for $300^\circ \pm 30^\circ$), *trans* ($180^\circ \pm 30^\circ$) and others (outside the ranges mentioned for the other conformations) [38].

We analysed the hydrogen bonding characteristics and radial distribution function (RDF) for each of the four modified residues and the distribution of the θ torsion angle (H2'-C2'-O2'-HO2') for the Ψ , $m^1\Psi$, $m^3\Psi$ residues. For the calculation of the pseudorotation angle P , the χ , γ , and θ torsion angles, hydrogen bonds and RDFs, cpptraj tool from AmberTools18 [52] was used. RDFs of water oxygen atoms around the HN1 atom was calculated for each of the Ψ , $m^3\Psi$ and Ψm residues and RDFs of water oxygen atoms around the HN3 atom was calculated for each of the Ψ , $m^1\Psi$ and Ψm residues. Hydrogen bond formations were taken into account if the distance between the donor and the acceptor atoms was ≤ 3 Å and the donor-hydrogen-acceptor angle was $\geq 135^\circ$. The water occupancy maps around the average MD structure of Ψ corresponding to the FF99- χ_{ND} -bsc0 and TIP3P force field and water model combination were calculated using the grid routine in cpptraj tool and visualization was done using UCSF-Chimera [53]. The average MD structures were obtained from 600 frames corresponding to each of the four conformations, i.e. NORTH/*syn*, SOUTH/*syn*, NORTH/*anti* and SOUTH/*anti* conformations from a set of 16 ns REMD simulations.

Analysis of the simulated trajectories of the ssRNA oligomers

The cpptraj module of AmberTools18 [52] was used to analyze the MD trajectories of the ssRNA trimers and tetramers. The distributions of the pseudorotation angle P , the χ and γ torsion angles, hydrogen bonds, waterbridges, RDFs were calculated from the 3 μ s concatenated trajectory while conformational entropy (based on the quasi-harmonic approach) of each of the oligomers was calculated using individual

trajectories for each of the ssRNA oligomers using suitable utilities available with the cpptraj module.

The glycosidic (χ) torsion angle defined by the atoms O4'–C1'–C5–C4 for pseudouridine and the χ torsional angle defined by the atoms O4'–C1'–N9–C4 and O4'–C1'–N1–C2 for the canonical nucleosides adenosine (A) and uridine (U), respectively, were considered for the calculation of the *anti* and *syn* conformations. The ranges considered for the definition of the different conformations of the pseudorotation angle P, the χ and γ torsion angles were same as mentioned earlier for the single modified nucleosides. Formation of hydrogen bonds was considered if the donor–acceptor distance was ≤ 3 Å and the donor–hydrogen–acceptor angle was $\geq 135^\circ$.

Results and discussion

Conformational and hydration characteristics of modified nucleosides

In an earlier study [26] we validated the revised parameter sets for pseudouridine (Ψ) (FF99_ χ_{IDRP} _bsc0) [24] and checked the transferability of these parameters to the four pseudouridine derivatives, i.e. m¹ Ψ , m³ Ψ , Ψ m and m¹acp³ Ψ and our observations indicated that the revised parameters for Ψ were transferable to the Ψ derivatives. In the present study we developed new sets of partial atomic charges individually for Ψ and its three derivatives m¹ Ψ , m³ Ψ and Ψ m and reoptimized the parameters for the glycosidic torsion angle for each of these residues and compared the conformational ensembles. The REMD simulations were carried out using the combination of the force fields, i.e. FF99_ χ_{IDRP} _bsc0 and FF99_ χ_{ND} _bsc0 with the TIP3P, TIP4P-Ew and SPC/E water models.

Pseudorotation angle (P)

With the AMBER FF99 parameter sets, the distribution of the pseudorotation angle was observed to have a smaller population of the NORTH sugar pucker conformation compared to the experimentally observed population for each of the modified residues except for Ψ [26] (Table S3, Fig. 3). Inclusion of the revised γ torsion parameters (parmbsc0) with the AMBER FF99 parameter sets resulted in an improvement in the propensity of the NORTH sugar pucker conformation for all the Ψ -derivatives. But with the FF99_bsc0 parameters, the propensity of the NORTH sugar pucker conformation for Ψ was significantly lower than the experimentally observed value [26].

For Ψ , the FF99_ χ_{ND} _bsc0 force field in combination with the TIP3P and the SPC/E models generated a population of the NORTH sugar pucker conformation which were

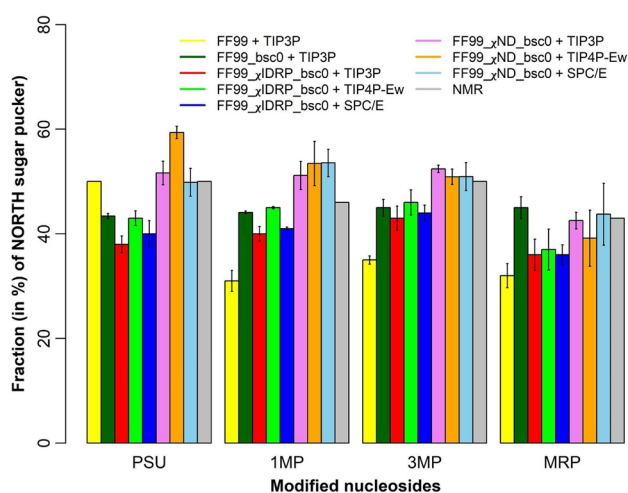


Fig. 3 Fraction (in %) of NORTH sugar pucker in equilibrium ensembles of Ψ (PSU), m¹ Ψ (1MP), m³ Ψ (3MP), and Ψ m (MRP) at 300 K. The values reported here are the averages (along with the standard deviations) calculated from three independent sets of 16 ns REMD simulations

in general much closer to the experimentally observed population than those generated by the FF99_ χ_{IDRP} _bsc0 force field in combination with each of the water models in this study. But FF99_ χ_{ND} _bsc0 in combination with the TIP4P-Ew water model generated a much greater population of the NORTH conformers of Ψ than the other force field and water model combinations and also the experimentally observed population. However, FF99_ χ_{IDRP} _bsc0 + TIP4P-Ew reproduced the experimental value of the NORTH population for m¹ Ψ better than all the other force field–water model combinations. In the case of m³ Ψ , it was observed that, the FF99_ χ_{ND} _bsc0 force field in combination with each of the three water models generated a population of the NORTH sugar pucker conformation which agreed better with the NMR results than what was observed with the FF99_ χ_{IDRP} _bsc0 force field in combination with each of the water models in this study. For Ψ m, the FF99_ χ_{ND} _bsc0 + TIP3P and FF99_ χ_{ND} _bsc0 + SPC/E combinations generated population of the NORTH conformers in better agreement with the NMR results than those generated by the other force field–water model combinations.

Glycosidic torsion angle (χ)

For each of the modified nucleosides under this study, experimental (NMR) studies reported a preference for the *syn* conformation [54–56]. The FF99 and FF99_bsc0 parameters, for each of the modified residues, predicted an excess population of *anti* conformers ($> 90\%$) [26, 57]. Earlier, we reported that FF99_ χ_{IDRP} _bsc0 + TIP3P shifted the equilibrium towards the *syn* conformation. The

FF99- $\chi_{\text{IDRP}}\text{-bsc0}$ parameter sets in combination with each of the TIP4P-Ew and SPC/E water models also generated a much greater population of *syn* conformation in good agreement with the NMR data than that obtained with the FF99 parameter sets (Table S4, Fig. 4). With the revised parameter sets FF99- $\chi_{\text{ND}}\text{-bsc0}$ in combination with each of the three water models, the modified residues adopted a much greater population of the *syn* conformation than what was predicted by the default AMBER parameters. But for each modified nucleoside, the population of *syn* conformers predicted by the FF99- $\chi_{\text{ND}}\text{-bsc0}$ parameters were lower than what was predicted by the FF99- $\chi_{\text{IDRP}}\text{-bsc0}$ parameters for each of the water models under this study.

Gamma torsion angle (γ)

In our earlier studies, we reported that, with the FF99 parameter sets, the g^+ population was much lower than the experimentally observed population for pseudouridine and its derivatives [26, 57]. In the present study, it was observed that all the force field and water model combinations predicted the g^+ population greater than what was predicted with the FF99 parameter sets, but also than the experimentally observed population (Table S5, Fig. 5). As was reported earlier [26], in the present study also, we observed that the inclusion of the revised γ torsion parameters developed by Pérez et al. [45] (parmb_{sc0}) shifted the

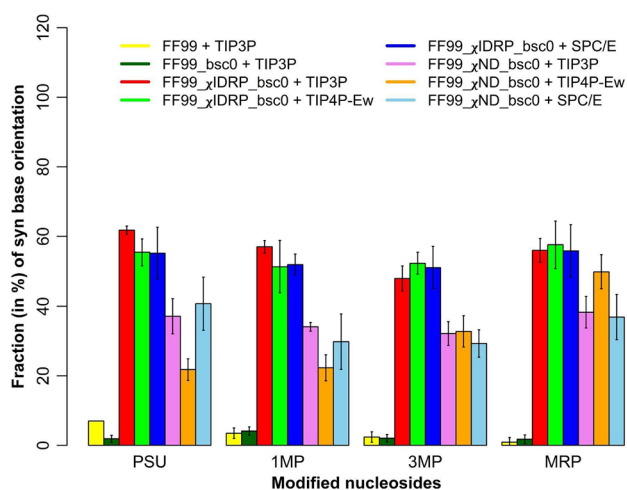


Fig. 4 The fraction (in %) of *syn* base orientation in the equilibrium ensembles of Ψ (PSU), $m^1\Psi$ (1MP), $m^3\Psi$ (3MP), and Ψ_m (MRP) at 300 K. The values reported here are the averages (along with the standard deviations) calculated from three independent sets of 16 ns REMD simulations. The modified nucleosides Ψ , $m^1\Psi$ (1MP), $m^3\Psi$ (3MP), and Ψ_m (MRP) have been reported to prefer the *syn* conformation [54–56]

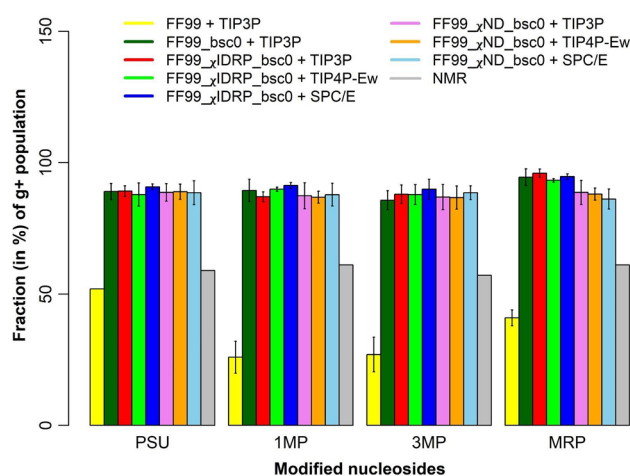


Fig. 5 The fraction (in %) of g^+ population in the equilibrium ensembles of Ψ (PSU), $m^1\Psi$ (1MP), $m^3\Psi$ (3MP), and Ψ_m (MRP) at 300 K. The values reported here are the averages (along with the standard deviations) calculated from three independent sets of 16 ns REMD simulations

equilibrium almost exclusively towards the g^+ conformation ($\sim 90\%$).

Correlation of the pseudorotation equilibrium with the glycosidic torsion angle (χ)

The two-dimensional scatter correlation plots of pseudorotation angle (P) vs glycosidic torsion angle (χ) revealed that for all the ribonucleosides in this study, with FF99- $\chi_{\text{IDRP}}\text{-bsc0}$ + TIP3P there was a significantly large population of the SOUTH/*syn* conformations (Figs. S4–S6). With FF99- $\chi_{\text{IDRP}}\text{-bsc0}$ + TIP4P-Ew, there were almost equal populations of SOUTH/*syn* and NORTH/*syn* conformations for all the four modified nucleosides, but the population of the SOUTH/*syn* conformers was a little higher in each case. The FF99- $\chi_{\text{IDRP}}\text{-bsc0}$ + SPC/E force field-water model combination also predicted a higher population of SOUTH/*syn* conformers than the others. In general, with the FF99- $\chi_{\text{ND}}\text{-bsc0}$ force field in combination with the TIP3P and the SPC/E water models, almost equal populations of the SOUTH/*syn* and NORTH/*anti* conformers were observed for each of the modified residues. The combination FF99- $\chi_{\text{ND}}\text{-bsc0}$ + TIP4P-Ew predicted a large population of NORTH/*anti* conformers for Ψ and $m^1\Psi$ nucleosides. But for $m^3\Psi$, this force field-water model combination predicted almost equal populations of the SOUTH/*syn* and NORTH/*anti* conformers. With FF99- $\chi_{\text{ND}}\text{-bsc0}$ + TIP4P-Ew, Ψ_m preferentially adopted the SOUTH/*syn* conformation.

Correlation of the pseudorotation equilibrium with the gamma torsion angle (γ)

From the two-dimensional correlation maps (two-dimensional scatter plots), it was observed that the FF99- χ_{IDRP} -bsc0 force field in combination with each of the three water models in the present study, predicted a greater population of the SOUTH/ g^+ conformers followed by that of the NORTH/ g^+ conformers for each of the modified nucleosides (Figs. S7–S9). In general, with the FF99- χ_{ND} -bsc0 parameter sets in combination with each of the three water models, we observed that there were almost equal populations of the NORTH/ g^+ and SOUTH/ g^+ conformers for all the residues. With FF99- χ_{ND} -bsc0 + TIP4P-Ew, Ψ preferentially adopted the NORTH/ g^+ conformation while Ψ_m preferentially adopted the SOUTH/ g^+ conformation. The populations of the g^- and trans conformers were extremely low due to the inclusion of the γ torsion parameters developed by Pérez

et al. [45] (parmbsc0) as was observed in our earlier study [26].

Hydrogen bonding

The hydrogen bonds except O5'–H5T...O4 (Fig. 6) and O2'–HO2'...O4 hydrogen bonds were observed to be negligible (Tables 3 and 4). With each of the force field-water model combinations, for all the modified residues (not applicable to Ψ_m), it was observed that the number of conformers with O2'–HO2'...O4 hydrogen bonding interaction were very small and much lesser than that of the O5'–H5T...O4 hydrogen bonding interaction.

For Ψ , FF99- χ_{IDRP} -bsc0 + TIP3P, FF99- χ_{ND} -bsc0 + TIP3P and FF99- χ_{ND} -bsc0 + SPC/E force field-water model combinations predicted greater populations of conformers with O5'–H5T...O4 hydrogen bonding interaction than what were predicted by the other force field-water

Fig. 6 Snapshots of **a** the O5'–H5T...O4 hydrogen bond in Ψ (PSU) residue in the *syn* conformation of the glycosidic torsion and NORTH conformation of sugar pucker; **b** the O5'–H5T...O4 hydrogen bond in Ψ (PSU) residue in the *syn* conformation of the glycosidic torsion and SOUTH conformation of sugar pucker; **c** the O2'–HO2'...O4 hydrogen bond in Ψ (PSU) residue in the *anti* conformation of the glycosidic torsion and NORTH conformation of sugar pucker; **d** the O2'–HO2'...O4 hydrogen bond in Ψ (PSU) residue in the *anti* conformation of the glycosidic torsion and SOUTH conformation of sugar pucker. All the snapshots are taken from a set of 16 ns REMD simulations corresponding to the FF99- χ_{ND} -bsc0 force field and TIP3P water model combination

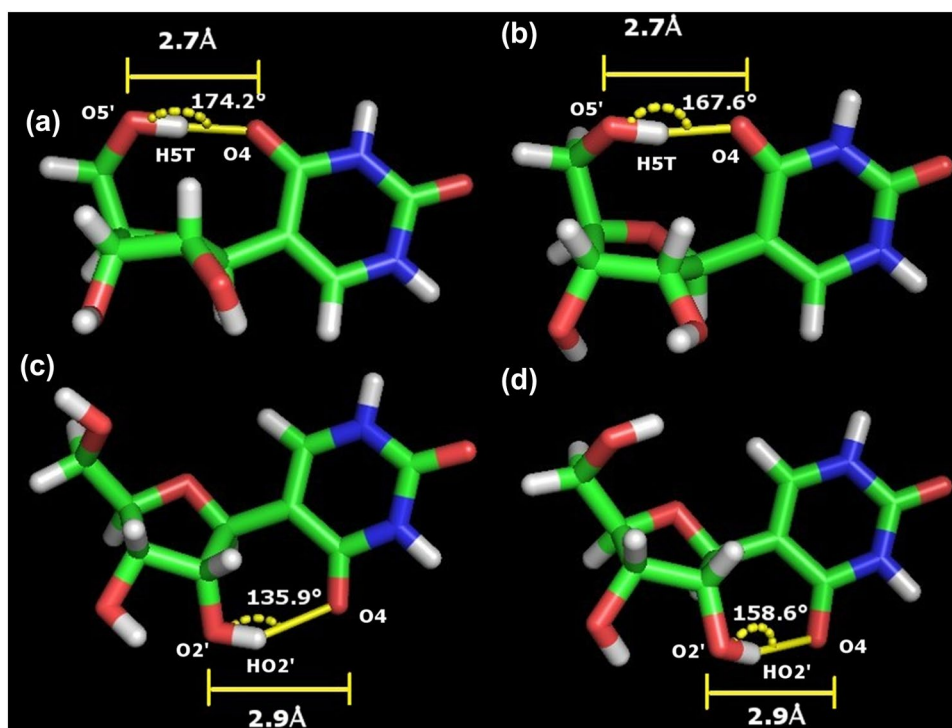


Table 3 Percent (%) occurrence of O5'–H5T...O4 hydrogen bond (at 300 K)

Force fields	Water models	Hydrogen bonding atoms	Ψ (PSU)	($m^1\Psi$) 1MP	($m^3\Psi$) 3MP	(Ψ_m) MRP
FF99- χ_{IDRP} -bsc0	TIP3P	O5'–H5T...O4	20 ± 0.82	18 ± 1.3	16 ± 3.0	19 ± 2.2
FF99- χ_{IDRP} -bsc0	TIP4P-Ew		15 ± 2.3	14 ± 2.2	15 ± 2.6	16 ± 1.4
FF99- χ_{IDRP} -bsc0	SPC/E		17 ± 2.3	15 ± 0.41	15 ± 2.0	17 ± 2.3
FF99- χ_{ND} -bsc0	TIP3P		20 ± 2.1	22 ± 3.0	17 ± 2.3	23 ± 3.1
FF99- χ_{ND} -bsc0	TIP4P-Ew		10 ± 0.9	12 ± 2.2	18 ± 3.0	28 ± 2.8
FF99- χ_{ND} -bsc0	SPC/E		21 ± 4.3	15 ± 5.3	17 ± 1.4	22 ± 4.0

Table 4 Percent (%) occurrence of O2'–HO2'···O4 Hydrogen bond (at 300 K)

Force fields	Water models	Hydrogen bonding atoms	Ψ (PSU)	(m ¹ Ψ) 1MP	(m ³ Ψ) 3MP
FF99_ χ_{IDRP} _bsc0	TIP3P	O2'–HO2'···O4	2.4 ± 0.34	2.7 ± 0.58	3.1 ± 0.09
FF99_ χ_{IDRP} _bsc0	TIP4P-Ew		2.5 ± 0.21	2.4 ± 0.25	3.0 ± 0.18
FF99_ χ_{IDRP} _bsc0	SPC/E		2.2 ± 1.1	2.6 ± 0.27	2.9 ± 0.72
FF99_ χ_{ND} _bsc0	TIP3P		3.4 ± 0.27	4.0 ± 0.12	3.5 ± 0.25
FF99_ χ_{ND} _bsc0	TIP4P-Ew		4.8 ± 0.53	6.0 ± 0.83	3.2 ± 0.64
FF99_ χ_{ND} _bsc0	SPC/E		3.1 ± 0.26	4.2 ± 0.58	3.1 ± 0.15

model combinations. With FF99_ χ_{ND} _bsc0 + TIP4P-Ew, the number of O5'–H5T···O4 hydrogen bonding interactions for Ψ were significantly lower than what were predicted by the other force field-water model combinations. For m¹ Ψ , FF99_ χ_{ND} _bsc0 + TIP3P predicted a greater number of O5'–H5T···O4 hydrogen bonding interactions than those predicted by the other force field-water model combinations. All the force field-water model combinations generated almost equal populations of conformers with O5'–H5T···O4 hydrogen bonding interaction for the m³ Ψ residue. For Ψ m, FF99_ χ_{ND} _bsc0 force field parameters predicted greater number of O5'–H5T···O4 hydrogen bonding interactions than what was predicted by FF99_ χ_{IDRP} _bsc0 in combination with each of the water models in this study. For Ψ and m¹ Ψ , the FF99_ χ_{ND} _bsc0 force field predicted slightly greater population of conformers with O2'–HO2'···O4 hydrogen bonding interaction than the FF99_ χ_{IDRP} _bsc0 parameter sets with each of the water models. But for m³ Ψ , all the force field-water model combinations generated similar populations of conformers with O2'–HO2'···O4 hydrogen bonding interaction.

Radial distribution function

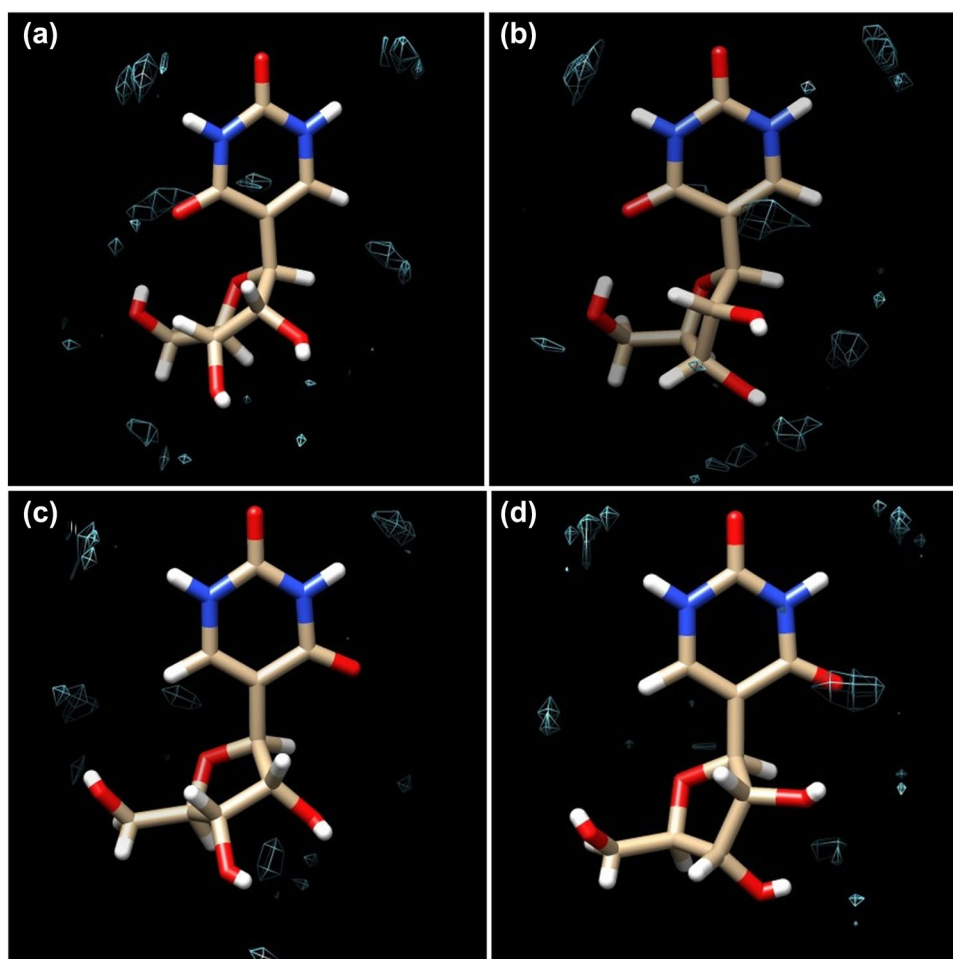
From the RDF plot of water oxygen atoms with respect to the HN1 atom of Ψ , it was observed that the FF99_ χ_{IDRP} _bsc0 force field in combination with each of the water models predicted the formation of a well-defined first hydration shell between 1.5 and 2.5 Å having a maximum at ~ 2 Å (Figs. S10–S12). This observation was consistent with that of the recent report by Deb et al. [21]. The FF99_ χ_{ND} _bsc0 force field in combination with each of the water models also predicted the formation of a well-defined first hydration shell between 1.5 and 2.5 Å having a maximum at ~ 2 Å around the Ψ -HN1 atom. For the HN1 atoms of m³ Ψ and Ψ m also, all the force field and water model combinations predicted the formation of a well-defined first hydration shell between 1.5 and 2.5 Å having a maximum at ~ 2 Å. For the Ψ and m³ Ψ residues, the concentration of the water molecules around the HN1 atom was observed to be slightly higher with the FF99_ χ_{ND} _bsc0 force field than what was observed with the FF99_ χ_{IDRP} _bsc0 force field for each of the water

models while for Ψ m the concentration was observed to be slightly lower with the FF99_ χ_{ND} _bsc0 force field than what was observed with the FF99_ χ_{IDRP} _bsc0 force field for each of the water models. From the RDF plots of water oxygen atoms with respect to the HN3 atoms of Ψ , m¹ Ψ and Ψ m nucleosides, with each of the force field and water model combinations, the formation of a well-defined first hydration shell was observed between 1.5 Å to 2.5 Å having a maximum at ~ 2 Å. For Ψ and m³ Ψ , the concentration of the water molecules around the HN3 atom was observed to be similar with each of the force field and water model combinations. Interestingly, for Ψ m, the FF99_ χ_{ND} _bsc0 force field parameters predicted a higher concentration of water molecules around the HN3 atom than what was predicted by FF99_ χ_{IDRP} _bsc0 in combination with each of the water models. The hydration pattern around pseudouridine (Ψ) corresponding to the FF99_ χ_{ND} _bsc0 and TIP3P force field and water model combination is shown in Fig. 7.

Orientation of the 2'-hydroxyl group of the Ψ , m¹ Ψ and m³ Ψ nucleosides

The orientation of the 2'-hydroxyl groups of RNA has been reported to have a significant contribution to the stability of the A-form RNA helices [58] and also in RNA–protein interactions [59]. The A-RNA duplex has been suggested to be stabilized by a network consisting of water-mediated hydrogen bonds mediated by the 2' hydroxyl groups and also the extensive individual hydration of the 2' hydroxyl groups [60, 61]. Kùhrova et al. (2014) reported that the choice of water model has a significant effect on the orientation of the 2'-OH atoms of nucleotides and hence also on the entire RNA structure [29]. The θ torsion angle (H2'–C2'–O2'–HO2') populates three regions, the O3' domain (value of θ between 50° and 140°), the O4' domain (value of θ between 175° and 230°) and the base domain (value of θ between 270° and 345°) for the C3'-endo sugar pucker conformation [29, 62]. It has been reported that the 2'-OH group when oriented towards the base domain can act as a hydrogen bond donor to a water molecule and when it is oriented towards the O3' domain it can accept a hydrogen bond from the same water molecule

Fig. 7 Hydration pattern of Ψ (PSU) in **a** NORTH/*syn* **b** SOUTH/*syn*, **c** NORTH/*anti* and **d** SOUTH/*anti* conformations corresponding to the FF99- χ_{ND} -bsc0 force field and TIP3P water model. Water occupancy contoured at equivalent levels visualized using UCSF-chimera [53]



[58, 60]. NMR studies at low temperatures suggested that the 2'-OH group can be oriented either towards the O3' domain or towards the base domain and the predominant orientation of the 2'-OH group is reported to be towards the O3' domain [63]. In the present study, we checked the effect of the combinations of the FF99- χ_{IDRP} -bsc0 and FF99- χ_{ND} -bsc0 force fields with the three different water models TIP3P, TIP4P-Ew and SPC/E, on the orientation of the 2'-OH atom corresponding to the Ψ , $m^1\Psi$, and $m^3\Psi$ residues (Fig. 8). The distribution of the θ torsion angle (H2'-C2'-O2'-HO2') angle was similar for each of the three water models for each modified residue. But the distribution differed between the two force fields. The O3'-domain was predominantly sampled (followed by the base-domain) by all the force field-water model combinations in agreement with the experimental and theoretical studies [29, 63]. The population of the conformers with the 2'-OH atom oriented towards the O4'-domain were significantly lower than the population of the conformers with the 2'-OH atom oriented towards the other two domains. While a prominent peak was observed at the O4'-domain with the FF99- χ_{IDRP} -bsc0 force field in combination with

each of the three water models, the FF99- χ_{ND} -bsc0 force field did not predict the same.

Effect of the substitution of U with Ψ in ssRNA oligomers

For further validation of the force field parameters, we have carried out molecular dynamics simulation of ssRNA trinucleotides AUA and A Ψ A and tetranucleotides AAUA and AA Ψ A and analyzed the effect of substitution of U with Ψ in ssRNA. The initial structures of the ssRNA trimers and tetramers in this study are shown in Fig. S16.

Conformational characteristics of ssRNA oligomers

Pseudouridine (Ψ) has been reported to preferentially adopt the *anti* base orientation within a polynucleotide chain context [16, 17]. Although this modification has the same basic topology as uridine in RNA, it has been observed to confer rigidity upon both dsRNA and ssRNA regions [4, 16, 17, 23, 64, 65]. It has been established that, generally, the population (in %) of the 3'-endo/NORTH conformers of an

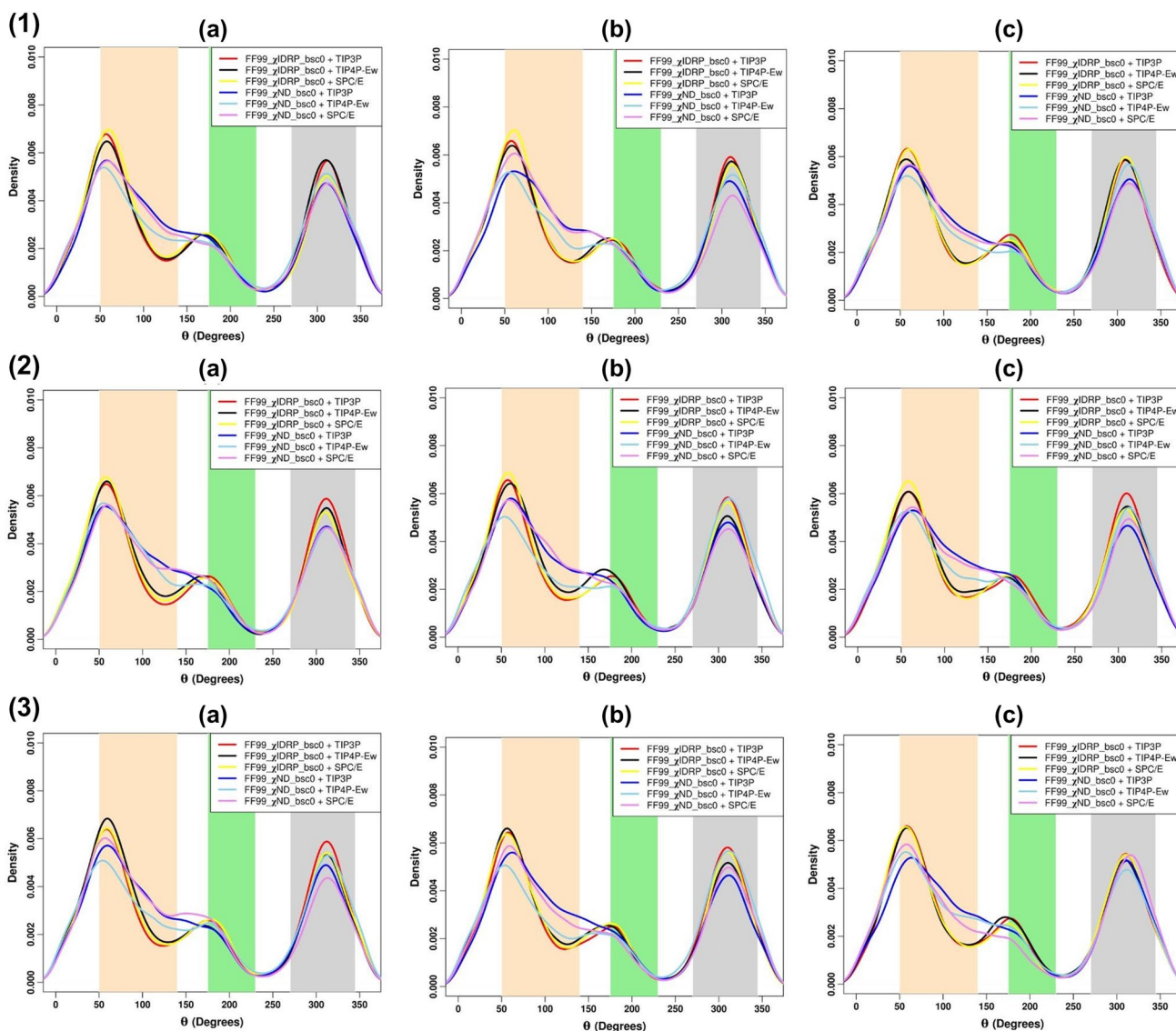


Fig. 8 Distribution of the θ torsion angle ($H2'-C2'-O2'-HO2'$) for **a** Ψ (PSU), **b** $m^1\Psi$ (1MP), and **c** $m^3\Psi$ (3MP), corresponding to FF99_ χ_{IDRP_bsc0} and TIP3P, FF99_ χ_{IDRP_bsc0} and TIP4P-Ew, FF99_ χ_{IDRP_bsc0} and SPC/E, FF99_ χ_{ND_bsc0} and TIP3P, FF99_ χ_{ND_bsc0} and TIP4P-Ew, and FF99_ χ_{ND_bsc0} and SPC/E force field and water model combinations for the (1) first set, (2) second set, (3) third set of

16 ns REMD simulations (at 300 K). The θ torsion angle ($H2'-C2'-O2'-HO2'$) populates the following three regions: O3' domain (value of θ between 50° – 140°) (indicated in bisque rectangle); O4' domain (value of θ between 175° – 230°) (indicated in light green rectangle) and the base domain (value of θ between 270° – 345°) (indicated in grey rectangle), for the C3'-endo sugar pucker conformation [29, 62]

individual nucleotide, indicates the extent of base stacking corresponding to that nucleotide position within RNA oligonucleotides [66, 67]. NMR study by Davis (1995) revealed that the substitution of U with Ψ promoted the C3'-endo/NORTH sugar pucker conformation for Ψ and also for the nucleotides preceding the modified residue in the ssRNA oligoribonucleotides and hence results in enhanced base stacking and stabilization throughout the ssRNA molecule [22]. Ψ adopts a low *anti* conformation, constrains the base to an *anti* and axial conformation and hence facilitates the C3'-endo/NORTH and axial conformation of the nucleotide

corresponding to the A-form RNA helix and rotation about the glycosidic bond from this C3'-endo/NORTH and *anti* conformation has been suggested to be highly disfavored [16, 22, 38, 66, 67].

In the present study we have theoretically investigated the effect of substituting U with Ψ in the AUA and AAUA oligomers and compared the results with the experimental observation [22, 68]. It was observed that, in general, both the revised parameter sets FF99_ χ_{IDRP_bsc0} and FF99_ χ_{ND_bsc0} for Ψ in combination with the FF99_ χ_{OL3_bsc0} for A/U reproduced the experimentally observed

conformational properties of the ssRNA oligomers. In general, the FF99_ χ_{OL3} _bsc0 parameter sets predicted somewhat greater population of the the NORTH conformers of the canonical nucleotides in the AUA and AAUA oligomers particularly in the 5' direction than that observed in the experiment [22, 68]. But, with each of the revised parameter sets, the A Ψ A and the AA Ψ A oligomers, adopted more stabilized A-form RNA structures compared to the canonical AUA and AAUA oligomers. For the oligomers under this study, we observed that the stabilization of the sugar pucker conformation was more prominent towards the 5' direction than the 3' direction which is consistent with the experimental data [22]. For the A Ψ A trimer, the FF99_ χ_{ND} _bsc0 parameters predicted a greater population of the NORTH conformers as well as that of the *anti* conformers for the 5'A and the Ψ nucleotides than what were predicted by the FF99_ χ_{IDRP} _bsc0 parameter sets (Tables S6–S7, Fig. 9). For the AA Ψ A tetramer, the sugar pucker conformations predicted by each of the revised parameter sets, were similar and of the parameter sets predicted a greater population of the NORTH conformers for Ψ and the preceding A nucleotide than the populations of the NORTH conformers

of U and the preceding A nucleotide in the AAUA tetramer (Table S6, Fig. 10). But with the FF99_ χ_{ND} _bsc0 parameter sets, the population of the syn conformers corresponding to the 5'A nucleotide in the AA Ψ A tetramer was greater than what was observed with the FF99_ χ_{IDRP} _bsc0 parameter sets (Table S7, Fig. 10). The deviation of the terminal A residues (especially A(4)) from the A-RNA conformation is expected since the terminal residues in the AA Ψ A tetramer showed a tendency to deviate from the ideal A-form RNA helical conformation at 298 K in the NMR experiment reported by Davis (1995) [22]. In general, with each of the revised parameter sets, Ψ in A Ψ A trimer and AA Ψ A tetramer was observed to preferentially adopt a low *anti* conformation in agreement with the experimental data. For all the oligomers under this study, each of the constituent nucleotides preferentially adopted a g^+ conformation of the gamma torsion (Table S8). From the quasi-harmonic (QH) conformational entropy calculations we observed that the conformational entropies (S) of the A Ψ A and the AA Ψ A oligomers were slightly lower than those of the AUA and the AAUA oligomers suggesting the stabilizing effect of Ψ over U (Table S9).

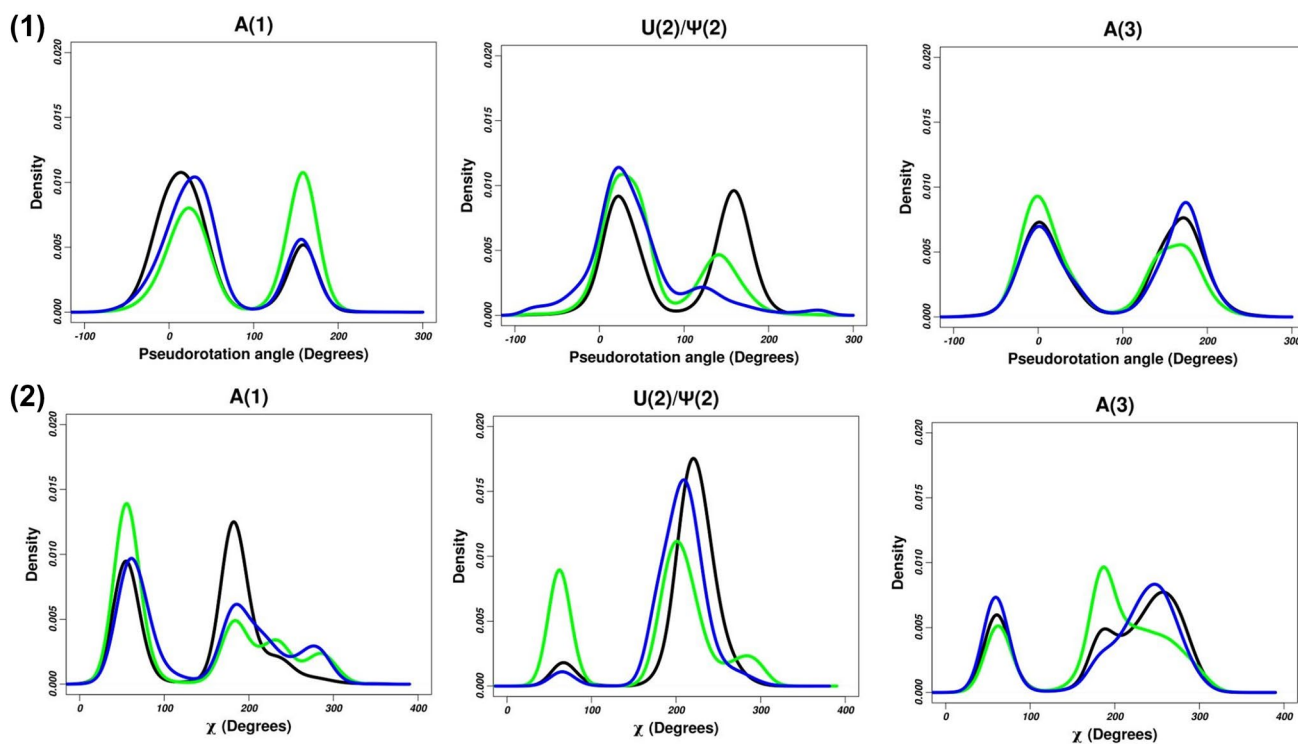


Fig. 9 Population distributions of (1) pseudorotation angle (density vs pseudorotation angle) and (2) glycosidic torsion angle (density vs glycosidic torsion angle (χ)) of the individual nucleotides in AUA corresponding to the simulation with the FF99_ χ_{OL3} _bsc0 force field parameters (black) and those in A Ψ A trimers corresponding to the simulations with the FF99_ χ_{OL3} _bsc0 force field parameters for A and FF99_ χ_{IDRP} _bsc0 force field parameters for Ψ (green) and

with the FF99_ χ_{OL3} _bsc0 force field parameters for A and FF99_ χ_{ND} _bsc0 force field parameters for Ψ (blue). The numbers given within braces beside the residue names indicate the residue numbers from the 5' end to the 3' end of the ssRNA trimers. The scale for pseudorotation angle (P) along the X-axis was converted to -90° to 270° to better delineate the NORTH ($270^\circ \leq P < 90^\circ$) and SOUTH ($90^\circ \leq P < 270^\circ$) regions

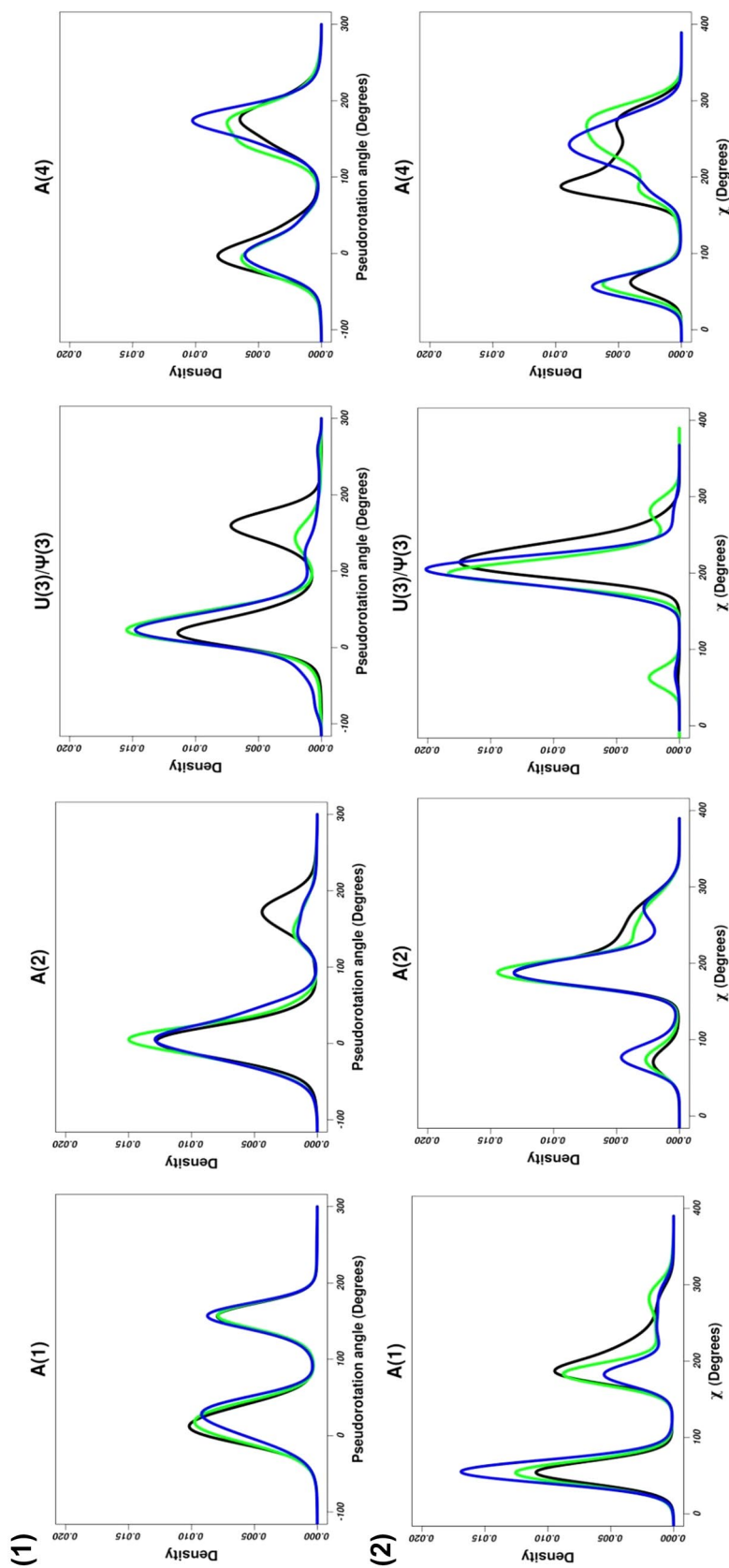


Fig. 10 Population distributions of (1) pseudorotation angle (density vs pseudorotation angle) and (2) glycosidic torsion angle (density vs glycosidic torsion angle (χ)) of the individual nucleotides in AAUA tetramer corresponding to the simulation with the FF99- χ_{0L3} -bsc0 force field parameters (black) and those in AA Ψ A tetramer corresponding to the simulations with the FF99- χ_{0L3} -bsc0 force field parameters for A and FF99- χ_{0L3} -bsc0 force field parameters for Ψ (green) and with the FF99- χ_{0L3} -bsc0 force field parameters for A and FF99- χ_{ND} -bsc0 force field parameters for Ψ (blue). The numbers given within braces beside the residue names indicate the residue numbers from the 5' end to the 3' end of the ssRNA tetramers. The scale for pseudorotation angle (P) along the X-axis was converted to -90° to 270° to better delineate the NORTH ($270^\circ \leq P < 90^\circ$) and SOUTH ($90^\circ \leq P < 270^\circ$) regions

The analysis of the intramolecular hydrogen bonding interaction revealed that in general, the propensity of the formation of intramolecular hydrogen bonds was greater in the oligomers containing the Ψ modification than their canonical counterparts (Tables S10–S11).

Hydration characteristics of ssRNA oligomers

The presence of an additional (N1H imino proton) hydrogen bond donor coordinating structural water molecules, has been reported to contribute to the structural and thermodynamic stabilization observed upon substitution of U with Ψ in RNA [9, 15, 16, 64, 69, 70]. Auffinger and Westhof [30], reported the presence of long-lived water bridges between the OP2 atoms of Ψ 32 and its preceding residue (C31) and the HN1 atom of Ψ 32, from the results of MD simulations of the tRNA^{Asp} anticodon hairpin performed using the AMBER FF94 force field in combination with SPC/E water and NH₄⁺ ions. Interestingly, a string of two water molecules was observed to be involved in the water-bridging interaction between the N1H imino proton and the backbone phosphates of Ψ and the preceding residue within the context of a A-form RNA duplex from MD simulations using the FF99- χ_{IDRP_bsc0} parameter sets for Ψ and the TIP3P water model [21]. In the present study we observed certain changes in the hydration pattern of the oligomers upon replacement of U with Ψ .

From the analysis of the hydrogen bonding interactions of the N3H imino group of uridine (U) and that of the N1H and N3H imino groups of pseudouridine (Ψ) with the solvent molecules it was observed that, the number of occurrences of the N3–HN3···O hydrogen bonds for U were similar in the AUA and AAUA oligomers (Table S12). For the A Ψ A trimer, with each of the revised parameter sets, the number of occurrences of the Ψ N3– Ψ HN3···O(water) hydrogen bonds were observed to be close to that of the UN3–UHN3···O(water) hydrogen bonds observed in AUA. The FF99- χ_{ND_bsc0} parameters predicted somewhat greater number of occurrences of Ψ N1– Ψ HN1···O(water) hydrogen bonds than that predicted by the FF99- χ_{IDRP_bsc0} parameters. For the AA Ψ A tetramer, the number of occurrences of the Ψ N3– Ψ HN3···O(water) hydrogen bonds were lesser than that of the UN3–UHN3···O(water) hydrogen bonds observed in AAUA. With the FF99- χ_{ND_bsc0} parameters, the number of occurrences of the Ψ N1– Ψ HN1···O(water) hydrogen bonds was observed to be much greater than that predicted by FF99- χ_{IDRP_bsc0} for the AA Ψ A tetramer.

In general, the number of occurrences of water bridges involving the backbone atoms were low in all the oligomers under this study. The propensities of water bridge formation between the backbone atoms of Ψ with the backbone atoms of the neighbouring adenosine residues in the A Ψ A trimer were not much different than what was observed

for U in the AUA trimer (Table S13). But with the FF99- χ_{ND_bsc0} parameters, the number of occurrences of water bridge interactions between the backbone atoms of Ψ (2) and A(3) were slightly higher than what was predicted by the FF99- χ_{IDRP_bsc0} parameters. The presence of additional water bridging interactions involving the N1H imino group and the backbone atoms of Ψ were observed in the A Ψ A trimer, which are not possible for the AUA trimer due to the absence of the N1H imino group in U. The propensities of water bridge formation between the backbone atoms of Ψ with the backbone atoms of the immediate neighbour adenosine residues in the AA Ψ A tetramer were a little greater with the FF99- χ_{IDRP_bsc0} parameters than those predicted by the FF99- χ_{ND_bsc0} parameters (Table S14). The presence of additional water bridging interactions involving the N1H imino group Ψ (3) and the backbone atoms of Ψ (3) and its preceding adenosine (A(2)) residue were observed in the AA Ψ A tetramer and such interactions are not possible for the canonical AAUA tetramer. With each of the revised parameter sets, it was observed that the Ψ (3)OP2-W- Ψ (3)HN1 water bridges were more frequent than the Ψ (3)OP1-W- Ψ (3)HN1 water bridges and the A(2)OP2-W- Ψ (3)HN1 water bridges were more frequent than the A(2)OP1-W- Ψ (3)HN1 water bridges. Each of the revised parameter sets predicted the formation of the water bridge between the HN1 atom of Ψ and the OP2 (phosphate oxygen) atoms of Ψ and its preceding A nucleotide (i.e. the A(2)-OP2-W- Ψ (3)-OP2 & Ψ (3)-HN1 water bridge interaction mentioned in Table S14) within the AA Ψ A tetramer. The snapshots of a coordinating water molecule hydrogen-bonded to the OP2 atom of A(2) and the OP2 and HN1 atoms of Ψ (3) in the AA Ψ A tetramer corresponding to the simulations with the revised parameter sets FF99- χ_{IDRP_bsc0} and FF99- χ_{ND_bsc0} for Ψ are given in Fig. S17.

The plots (Fig. S18) of the radial distribution function (RDF) of the water oxygen atoms with respect to the HN1 atom of Ψ in both the A Ψ A and AA Ψ A oligomers, indicated the formation of a well-defined first hydration shell between 1.5 and 2.5 Å having a maximum at ~2 Å and this observation was similar to what was observed for the Ψ nucleoside as well as for Ψ within duplex RNA [21]. The first hydration peak corresponding to the H5 atom of U was observed to be located at a further distance and with a lower concentration of water molecules. The RDF plots of water oxygen atoms about the geometric center of the OP2 and H5 atoms of U and that about the geometric center of the OP2 and HN1 atoms of Ψ indicated the presence of water molecules at a shorter distance between the base and the backbone of Ψ compared to that observed between the base and the backbone of U and the concentration of water molecules predicted by the FF99- χ_{IDRP_bsc0} parameters were greater than that predicted by the FF99- χ_{ND_bsc0} parameters (Fig. S19). The formation of a well-defined first hydration shell

with a peak at 0.55 Å, from the RDF plot of water oxygen atoms about the geometric center of the OP2 atom of A(2) and HN1 atom of $\Psi(3)$ in the AA Ψ A tetramer was observed for each of the revised sets of parameters (Fig. S20a). But with FF99- χ_{IDRP_bsc0} , the first hydration shell was located between 0 and 1.75 Å and had a much greater number of water molecules than that with FF99- χ_{ND_bsc0} , located between 0.25 and 1.45 Å. The RDF plot of water oxygen atoms about the geometric center of the OP2 atom of A(2) and the OP2 and HN1 atoms of $\Psi(3)$ in the AA Ψ A tetramer showed formation of a well-defined first hydration shell between 0 and 1.05 Å with a peak at 0.45 Å, with the FF99- χ_{IDRP_bsc0} parameters (Fig. S20b). But with the FF99- χ_{ND_bsc0} parameters, the first hydration shell was observed between 0.15 and 1.15 Å with a peak at 0.55 Å, having a lower number of water molecules than what was observed with the FF99- χ_{IDRP_bsc0} parameters. In general, this observation also suggested the presence of the coordinating structural water between the HN1 and OP2 atoms of $\Psi(3)$ and the OP2 atom of its preceding adenosine (A(2)) residue.

Conclusions

In the present study we derived a revised set of partial atomic charges and glycosidic torsional parameters (χ_{ND}) for the nucleosides Ψ , $m^1\Psi$, $m^3\Psi$, and Ψ_m following a data-informed approach. At the individual nucleoside level, the partial atomic charges and glycosidic torsional parameters (χ_{ND}) were calculated by applying RESP fitting method to the lowest energy conformation of the quantum mechanical energy profile of a chosen conformational scheme and fitting the molecular mechanics energy profile to that scheme-specific quantum mechanical energy profile, respectively. The choice of a particular conformational scheme was dictated by the NMR results that reported a preference for the *syn* conformation for pseudouridine and its derivatives under study [54–56] and thereafter looking for the scheme that had the lowest energy value for the *syn* conformation.

The consequences of the application of the revised set of glycosidic torsional parameters (χ_{ND}) in combination with the revised γ torsion parameters (parmb $sc0$) developed by Pérez et al. [45] and the AMBER FF99-derived parameters [25] for these modified nucleosides were analysed using replica exchange molecular dynamics simulations. The newly derived parameters were validated by comparing the simulated conformational preferences with the available experimental (NMR) data as well as with the observations in Dutta et al. [26]. REMD simulations were carried out using the FF99- χ_{IDRP_bsc0} [24] and FF99- χ_{ND_bsc0} force fields in combination with each of the TIP3P, TIP4P-Ew and SPC/E

water models. Three independent REMD simulations (each of 16 ns) were carried out in 16 temperature windows ranging from 300 to 400 K, resulting in 768 ns of simulation time in total for each system.

It was observed that there were significant differences in the description of the conformational properties of each of the modified nucleosides by different combinations of force fields and water models. The revised force field parameter sets (FF99- χ_{ND_bsc0}) with the TIP3P water model was able to closely reproduce the experimentally observed sugar pucker preferences for each of the modified nucleosides in this study. The accuracy of the prediction of the population of the C3'-endo/NORTH conformers might be important for accurate reproduction of the C3'-endo/NORTH pucker conformation associated with the A-form RNA structures.

In general, the newly developed force field parameters (FF99- χ_{ND_bsc0}) in combination with each of the water models under this study shifted the distribution of the base orientation for each of the modified nucleosides towards the *syn* conformation in contrast to the excess of *anti* conformations predicted by the AMBER FF99 and AMBER FF99- $bsc0$ parameters [25, 45]. But the population of the *syn* conformers predicted by the FF99- χ_{ND_bsc0} force field was observed to be less than that predicted by the FF99- χ_{IDRP_bsc0} force field parameters. The choice of water model was not found to influence the description of the base orientation to a significant extent for the FF99- χ_{IDRP_bsc0} force field parameters. However, the FF99- χ_{ND_bsc0} force field in combination with the TIP4P-Ew water model resulted in a somewhat smaller population of the *syn* conformers in the case of Ψ and $m^1\Psi$ nucleosides and a significantly greater population of the *syn* conformers for Ψ_m than what were observed with the other two water models.

In earlier studies from our group [24, 26], we reported that, at the single nucleoside level, the inclusion of the revised γ torsion parameters (parmb $sc0$) developed by Pérez et al. [45] along with the FF99- χ_{IDRP} parameter sets did not reproduce the experimentally observed population of the g^+ conformers, but predicted a much larger g^+ population for pseudouridine and its derivatives. We also noted that the large population of g^+ conformers observed with the FF99- χ_{IDRP_bsc0} parameters might be necessary to maintain the g^+ conformation of a nucleotide as is observed in the standard A-form of RNA [37]. In the present study, we observed that the newly derived FF99- χ_{ND_bsc0} parameter sets also predicted a large population of the g^+ conformation for each of the modified residues. The populations of g^+ conformers for all the nucleosides under this study, predicted by each of the force field and water model combinations were similar and were much larger than that predicted with the FF99 parameters.

The observations from the calculations of the number of O5'-H5T...O4 and O2'-HO2'...O4 hydrogen bonding

interactions for each of the modified nucleosides in this study, suggested that O5'–H5T···O4 hydrogen bonding interaction contributes to the stabilization of the *syn* base orientation [38] while the O2'–HO2'···O4 hydrogen bonding interaction may facilitate the *anti* base orientation. The orientation of the 2'-OH atom was observed to be similar for each of the modified residues under this study. All the force field water model combinations predicted the predominant orientation of the 2'-OH atom towards O3' which was consistent with previous NMR results [63]. Interestingly, with the FF99_χ_{IDRP}_bsc0 force field, a prominent peak in the O4'-domain was observed in the distribution of the θ torsion angle (H2'–C2'–O2'–HO2').

Overall, the results from the simulations of the ssRNA oligomers in this study suggest that both the revised sets of parameters (FF99_χ_{IDRP}_bsc0 and FF99_χ_{ND}_bsc0) for Ψ in combination with the FF99_χ_{OL3}_bsc0 parameters were able to reproduce the conformational and hydration properties of the oligomers. We observed that the substitution of uridine with pseudouridine in the ssRNA oligomers resulted in the formation of a more stable A-form RNA helix conformation. With the application of the reoptimized partial atomic charges and the revised glycosidic torsion parameters (χ_{ND}) developed in this study, Ψ, within each of the AΨA and AAΨA oligomers, was observed to predominantly adopt the experimentally observed *anti*/NORTH conformation and the preferences were stronger than what was observed with the χ_{IDRP} parameters. In the present study we also highlighted the effect of the presence of the pseudouridine modification in the hydration characteristics of the AΨA and AAΨA oligomers compared to their canonical counterparts. From our simulations, we observed and the occurrence of the water bridging interactions between the N1H imino proton and the phosphate oxygen atoms of Ψ and the preceding A residue in the AAΨA tetramer. The presence of a structural water molecule hydrogen bonded to the HN1 atom of Ψ(n) coordinating the phosphate oxygen atoms of Ψ(n) and A(n–1), is consistent with the experimental (crystallographic) evidence [70] and the model proposed in NMR studies [22, 64, 65, 69] and might play a significant role in the stabilization of the A-form RNA structure.

Several studies have suggested the importance of the 1-methyl derivative of Ψ (m¹Ψ) in mRNA therapeutics and this modification has a crucial contribution to the efficiency of the COVID-19 mRNA vaccines [71–75]. In our ongoing study, we are further validating our revised sets of parameters for the methyl derivatives of pseudouridine (i.e. m¹Ψ, m³Ψ and Ψm) based on simulations of the ssRNA oligomers containing these derivatives. It would be interesting to investigate the effect of the presence of the methyl derivatives of pseudouridine within single stranded as well as double stranded RNA context on the conformational and hydration characteristics of such RNA structures, both experimentally

and theoretically. The revised sets of parameters might be useful in the theoretical studies involving RNA structures containing such modifications and also in designing therapeutic antisense oligonucleotides.

Supporting information

METHODS: Replica exchange molecular dynamics simulations of the modified nucleosides; Molecular dynamics simulations of the ssRNA oligomers. AMBER preparatory files including the newly developed partial atomic charges for pseudouridine (Ψ) [PSU_ND.prepin], 1-methylpseudouridine (m¹Ψ) [1MP_ND.prepin], 3-methylpseudouridine (m³Ψ) [3MP_ND.prepin], 2'-O-methylpseudouridine (Ψm) [MRP_ND.prepin]. Revised parameter sets, i.e. AMBER frcmod files for pseudouridine (Ψ) [PSU_Chi_ND_bsc0.frcmod], 1-methylpseudouridine (m¹Ψ) [1MP_Chi_ND_bsc0.frcmod]; 3-methylpseudouridine (m³Ψ) [3MP_Chi_ND_bsc0.frcmod] and 2'-O-methylpseudouridine (Ψm) [MRP_Chi_ND_bsc0.frcmod]. Table S1. Frozen and restrained dihedrals during QM optimization in PES scan and MM energy minimizations. Table S2. Partial atomic charges for Ψ, m¹Ψ, m³Ψ, and Ψm ribonucleosides. Table S3. Propensity (in %) for NORTH sugar puckering of Ψ, m¹Ψ, m³Ψ, and Ψm ribonucleosides. Table S4. Fraction (in %) of base orientation states for Ψ, m¹Ψ, m³Ψ, and Ψm ribonucleosides. Table S5. Fraction (in %) of γ conformational states for Ψ, m¹Ψ, m³Ψ, and Ψm ribonucleosides. Table S6. Propensity (in %) for NORTH (C3'-endo) sugar pucker conformation of the individual nucleotides in the ssRNA oligomers Table S7. Fraction (in %) of the *anti* and *syn* base orientation states of the individual nucleotides in the ssRNA oligomers Table S8. Fraction (in %) of g⁺ conformation of the γ torsion angle of the individual nucleotides in the ssRNA oligomers Table S9. Conformational (quasi-harmonic entropy) entropy of the ssRNA oligomers Table S10–S11. Percent (%) of occurrence of intramolecular hydrogen bonding interaction of the ssRNA oligomers Table S12. Percent (%) of occurrence of hydrogen bonding interaction of the N3H imino group of uridine (U) and that of the N1H and N3H imino groups of pseudouridine (Ψ) with the solvent molecules for the ssRNA oligomers Tables S13–S14. Percent (%) of occurrence of water bridges formed by the backbone atoms in the ssRNA oligomers Fig. S1. Structures of Ψ, m¹Ψ, m³Ψ, and Ψm ribonucleosides corresponding to the five conformational schemes. Fig. S2. Conformational scheme (SC5) used in this work for the Ψ, m¹Ψ, m³Ψ, and Ψm ribonucleoside residues (along with the atom names). Figure S3. Energy profiles around χ torsional angles from QM calculations corresponding to the five schemes for

the Ψ , $m^1\Psi$, $m^3\Psi$, and Ψ_m ribonucleoside residues. Figs. S4–S6. Population distribution of the Ψ , $m^1\Psi$, $m^3\Psi$, and Ψ_m residues corresponding to the different force field and water model combinations with the pseudorotation angle (P) along the x-axis and the glycosidic torsion angle (χ) along the y-axis for the three independent sets of 16 ns REMD simulations respectively. Figs. S7–S9. Population distribution of the Ψ , $m^1\Psi$, $m^3\Psi$, and Ψ_m residues corresponding to the different force field and water model combinations with the pseudorotation angle (P) along the x-axis and the gamma torsion angle (γ) along the y-axis for the three independent sets of 16 ns REMD simulations respectively. Figs. S10–S12. RDFs of water oxygen atoms around the HN1 atom of the Ψ , $m^3\Psi$, and Ψ_m residues corresponding to the different force field and water model combinations for the three independent sets of 16 ns REMD simulations respectively. Figs. S13–S15. RDFs of water oxygen atoms around the HN3 atom of the Ψ , $m^1\Psi$, and Ψ_m residues corresponding to the different force field and water model combinations for the three independent sets of 16 ns REMD simulations respectively. Fig. S16. Initial structures of the ssRNA trimers and tetramers in this study Fig. S17. Snapshots of a coordinating water molecule hydrogen-bonded to the OP2 atom of A(2) and the OP2 and HN1 atoms of Ψ (3) in the AA Ψ A tetramer Fig. S18. RDF of water oxygen atoms about H5 atom of U (for AUA and AAUA) and about HN1 atom of Ψ (for A Ψ A and AA Ψ A). Fig. S19. RDF of water oxygen atoms about the geometric center of the OP2 and H5 atoms of U (for AUA and AAUA) and about the geometric center of the OP2 and HN1 atoms of Ψ (for A Ψ A and AA Ψ A) Fig. S20. RDF of water oxygen atoms about the geometric center of the OP2 atom of A(2) and H5 atom of U(3) (for AAUA) and the geometric center of the OP2 atom of A(2) and HN1 atom of Ψ (3) (for AA Ψ A) and RDF of water oxygen atoms about the geometric center of the OP2 atom of A(2) and OP2 and H5 atoms of U(3) (for AAUA) and the geometric center of the he OP2 atom of A(2) and OP2 and HN1 atoms of Ψ (3) (for AA Ψ A).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10822-022-00447-4>.

Acknowledgements The authors acknowledge the support provided by the SERB funding (EMR/2016/007753). N.D. acknowledges support from the DST INSPIRE Junior Research Fellowship (DST/INSPIRE Fellowship/2018/IF180895). The simulations and calculations were performed at the Poznan Supercomputing and Networking Center. The authors would like to thank Aditya Kumar Sarkar, Rupak Pal and Shinjini Ghosh for technical help.

Author contributions ND, and AL jointly conceived and designed the study. ND generated the data and analyzed the results. All authors reviewed the manuscript and approved the final version.

Funding Department of Science and Technology – Science and Engineering Research Board funding (EMR/2016/007753). Department of Science and Technology – Innovation in Science Pursuit for Inspired Research Junior Research Fellowship (DST/INSPIRE Fellowship/2018/IF180895).

Data availability All data and software are available. Input files required to reproduce the raw data and the output data are available in the supporting information. Additional data which support the findings in this study are available upon request from the corresponding author. The software and web servers used in this study are as follows AMBER16: <http://ambermd.org/>. AmberTools18: <http://ambermd.org/>. Gaussian09: <https://gaussian.com/>. PyRED: <https://upjv.q4md-forcefieldtools.org/REDServer-Development/>. Pymol: <https://pymol.org/2/>. ACD/ChemSketch (Freeware version): <https://www.acdlabs.com/>. Avogadro 1.1.1: <https://avogadro.cc/>. R 3.5.0: <https://cran.r-project.org/>.

Code availability Not applicable.

Declarations

Conflict of interest There are no conflicts to declare.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Boccaletto P, Machnicka MA, Purta E et al (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* 46:D303–D307
2. Cohn WE, Volkin E (1951) Nucleoside-5'-phosphates from ribonucleic acid. *Nature* 167:483–484
3. Davis FF, Allen FW (1957) Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem* 227:907–915
4. Charette M, Gray MW (2000) Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* 49:341–351
5. Scannell JP, Crestfield AM, Allen FW (1959) Methylation studies on various uracil derivatives and on an isomer of uridine isolated from ribonucleic acids. *Biochem Biophys Acta* 32:406–412
6. Yu C-T, Allen FW (1959) Studies of an isomer of uridine isolated from ribonucleic acids. *Biochem Biophys Acta* 32:393–406
7. Cohn WE (1959) 5-ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids. *Biochim Biophys Acta* 32:569–571
8. Cohn WE (1960) Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. *J Biol Chem* 235:1488–1498
9. Hall KB, McLaughlin LW (1992) Properties of pseudouridine N1 imino protons located in the major groove of an A-form RNA duplex. *Nucleic Acids Res* 20:1883–1889
10. Carlile TM, Rojas-Duran MF, Zinshteyn B et al (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515:143–146
11. Schwartz S, Bernstein DA, Mumbach MR et al (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159:148–162
12. Li X, Zhu P, Ma S et al (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat Chem Biol* 11:592–597

13. Zaringhalam M, Papavasiliou FN (2016) Pseudouridylation meets next-generation sequencing. *Methods* 107:63–72
14. Lei Z, Yi C (2017) A radiolabeling-free, qPCR-based method for locus-specific pseudouridine detection. *Angew Chem Int Ed Eng* 56:14878–14882
15. Hall KB, McLaughlin LW (1991) Properties of a U1/mRNA 5' splice site duplex containing pseudouridine as measured by thermodynamic and NMR methods. *Biochemistry* 30:1795–1801
16. Davis DR, Veltri CA, Nielsen L (1998) An RNA model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in tRNA^{Lys}, tRNA^{His} and tRNA^{Tyr}. *J Biomol Struct Dyn* 15:1121–1133
17. Yarian CS, Basti MM, Cain RJ et al (1999) Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39. *Nucleic Acids Res* 27:3543–3549
18. Newby MI, Greenbaum NL (2002) Investigation of Overhauser effects between pseudouridine and water protons in RNA helices. *Proc Natl Acad Sci USA* 99:12697–12702
19. Hudson GA, Bloomingdale RJ, Znosko BM (2013) Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. *RNA* 19:1474–1482
20. Kierzek E, Malgowska M, Lisowiec J et al (2014) The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res* 42:3492–3501
21. Deb I, Popenda L, Sarzyńska J et al (2019) Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair. *Sci Rep* 9:16278
22. Davis DR (1995) Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res* 23:5020–5026
23. Auffinger P, Westhof E (1998) Effects of pseudouridylation on tRNA hydration and dynamics: a theoretical approach. In: Grosjean H, Benne R (eds) *Modification and editing of RNA*. Am. Soc. Microbiol, Washington, DC, pp 103–112
24. Deb I, Pal R, Sarzynska J, Lahiri A (2016) Reparameterizations of the χ Torsion and Lennard-Jones σ Parameters improve the conformational characteristics of modified uridines. *J Comput Chem* 37:1576–1588
25. Aduri R, Psciuk BT, Saro P et al (2007) AMBER force field parameters for the naturally occurring modified nucleosides in RNA. *J Chem Theory Comput* 3:1464–1475
26. Dutta N, Sarzynska J, Lahiri A (2020) Molecular dynamics simulation of the conformational preferences of pseudouridine derivatives: improving the distribution in the glycosidic torsion space. *J Chem Inf Model* 60:4995–5002
27. Yildirim I, Stern HA, Kennedy SD et al (2010) Reparameterization of RNA χ torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine. *J Chem Theory Comput* 6:1520–1531
28. Zgarbová M, Otyepka M, Sponer J et al (2011) Refinement of the Cornell et al nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* 7:2886–2902
29. Kührová P, Otyepka M, Šponer J, Banáš P (2014) Are waters around RNA more than just a solvent? An insight from molecular dynamics simulations. *J Chem Theory Comput* 10:401–411
30. Auffinger P, Westhof E (1997) RNA hydration: three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA Asp anticodon hairpin 1 Edited by J. Karn. *J Mol Biol* 269:326–341
31. Jorgensen WL, Chandrasekhar J, Madura JD et al (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
32. Abascal JLF, Vega C (2005) A general purpose model for the condensed phases of water: TIP4P/2005. *J Chem Phys* 123:234505
33. Mahoney MW, Jorgensen WL (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, non-polarizable potential functions. *J Chem Phys* 112:8910–8922
34. Berendsen HJC, Grigera JR, Straatsma TP (1987) The missing term in effective pair potentials. *J Phys Chem* 91:6269–6271
35. Gelbin A, Schneider B, Clowney L et al (1996) Geometric parameters in nucleic acids: sugar and phosphate constituents. *J Am Chem Soc* 118:519–529
36. Schaftenaar G, Noordik JH (2000) *J Comput Aided Mol Des* 14:123–134
37. Richardson JS, Schneider B, Murray LW et al (2008) RNA backbone: consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). *RNA* 14:465–481
38. Saenger W (1984) *Principles of nucleic acid structure*. Springer advanced texts in chemistry. Springer, New York
39. Frisch MJ, Trucks GW, Schlegel HB et al (2009) *Gaussian 09*. Gaussian Inc, Wallingford, CT
40. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 97:10269–10280
41. Cieplak P, Cornell WD, Bayly C, Kollman PA (1995) Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins. *J Comput Chem* 16:1357–1377
42. Vanquleuf E, Simon S, Marquant G et al (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res* 39:W511–W517
43. Case DA, Betz RM, Cerutti DS et al (2005) *Amber 2016*. University of California, San Francisco
44. Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074
45. Pérez A, Marchán I, Svozil D et al (2007) Refinement of the AMBER Force field for nucleic acids: improving the description of α/γ conformers. *Biophys J* 92:3817–3829
46. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151
47. Horn HW, Swope WC, Pitner JW et al (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* 120:9665–9678
48. Altona C, Sundaralingam M (1972) Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J Am Chem Soc* 94:8205–8212
49. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) (1983) Abbreviations and symbols for the description of conformations of polynucleotide chains. *Recommendations* 1982. *Eur J Biochem* 131(1):9–15
50. Foloppe N, Hartmann B, Nilsson L, MacKerell AD Jr (2002) Intrinsic conformational energetics associated with the glycosyl torsion in DNA: a quantum mechanical study. *Biophys J* 82(3):1554–1569
51. Foloppe N, Nilsson L (2005) Toward a full characterization of nucleic acid components in aqueous solution: simulations of nucleosides. *J Phys Chem B* 109(18):9119–9131
52. Case DA, Ben-Shalom IY, Brozell SR et al (2018) *AMBER 2018*. University of California, San Francisco
53. Pettersen EF, Goddard TD, Huang CC et al (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612

54. Chang Y-C, Herath J, Wang TH-H, Chow CS (2008) Synthesis and solution conformation studies of 3-substituted uridine and pseudouridine derivatives. *Bioorg Med Chem* 16:2676–2686
55. Neumann JM, Bernassau JM, Guéron M, Tran-Dinh S (1980) Comparative conformations of uridine and pseudouridine and their derivatives. *Eur J Biochem* 108:457–463
56. Desaulniers J-P, Chui HM-P, Chow CS (2005) Solution conformations of two naturally occurring RNA nucleosides: 3-methyluridine and 3-methylpseudouridine. *Bioorg Med Chem* 13:6777–6781
57. Deb I, Sarzynska J, Nilsson L, Lahiri A (2014) Conformational preferences of modified uridines: comparison of AMBER derived force fields. *J Chem Inf Model* 54:1129–1142
58. Fohrer J, Hennig M, Carlomagno T (2006) Influence of the 2'-hydroxyl group conformation on the stability of A-form helices in RNA. *J Mol Biol* 356:280–287
59. Hou Y-M (2001) An important 2'-OH group for an RNA-protein interaction. *Nucleic Acids Res* 29:976–985
60. Egli M, Portmann S, Usman N (1996) RNA hydration: a detailed look. *Biochemistry* 35:8489–8494
61. Olsen GL, Bardaro MF Jr, Echodu DC et al (2009) Hydration dependent dynamics in RNA. *J Biomol NMR* 45:133–142
62. Auffinger P, Westhof E (1997) Rules governing the orientation of the 2'-hydroxyl group in RNA. *J Mol Biol* 274:54–63
63. Nozinovic S, Fürtig B, Jonker HRA et al (2010) High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res* 38:683–694
64. Davis DR (1998) Biophysical and conformational properties of modified nucleosides in RNA (nuclear magnetic resonance studies). In: Grosjean H, Benne R (eds) *Modification and editing of RNA*. ASM Press, Washington, pp 85–102
65. Durant PC, Davis DR (1999) Stabilization of the anticodon stem-loop of tRNA^{Lys,3} by an A+-C base-pair and by pseudouridine. *J Mol Biol* 285:115–131
66. Lee CH, Tinoco I Jr (1977) Studies of the conformation of modified dinucleoside phosphates containing 1, N6-ethenoadenosine and 2'-O-methylcytidine by 360-MHz ¹H nuclear magnetic resonance spectroscopy. Investigation of the solution conformations of dinucleoside phosphates. *Biochemistry* 16:5403–5414
67. Lee CH, Tinoco I Jr (1980) Conformation studies of 13 trinucleoside diphosphates by 360 MHz PMR spectroscopy. A bulged base conformation. I. Base protons and H1' protons. *Biophys Chem* 11:283–294
68. Dalluge JJ, Hashizume T, Sopchik AE et al (1996) Conformational flexibility in RNA: the role of dihydrouridine. *Nucleic Acids Res* 24:1073–1079
69. Davis DR, Poulter CD (1991) ¹H–¹⁵N NMR studies of *Escherichia coli* tRNA(Phe) from hisT mutants: a structural role for pseudouridine. *Biochemistry* 30:4223–4231
70. Arnez JG, Steitz TA (1994) Crystal structure of unmodified tRNA(Gln) complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* 33:7560–7567
71. Andries O, Mc Cafferty S, De Smedt SC et al (2015) N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *J Control Release* 217:337–344
72. Nelson J, Sorensen EW, Mintri S et al (2020) Impact of mRNA chemistry and manufacturing process on innate immune activation. *Sci Adv* 6:eaaz6893
73. Nance KD, Meier JL (2021) Modifications in an emergency: the role of N1-methylpseudouridine in COVID-19 vaccines. *ACS Cent Sci* 7:748–756
74. Park JW, Lagniton PNP, Liu Y, Xu R-H (2021) mRNA vaccines for COVID-19: what, why and how. *Int J Biol Sci* 17:1446–1460
75. Morais P, Adachi H, Yu Y-T (2021) The critical contribution of pseudouridine to mRNA COVID-19 vaccines. *Front Cell Dev Biol* 9:789427

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.