



Article scientifique

Article

2022

Appendix

Open Access

This file is a(n) Appendix of:

Mechanistic Insights into the Ligand-Induced Unfolding of an RNA G-
Quadruplex

Haldar, Susanta; Zhang, Yashu; Xia, Ying; Islam, Barira; Gervasio, Francesco Luigi; Liu, Sisi;
Mulholland, Adrian J; Waller, Zoë A E; Wei, Dengguo; Haider, Shozeb

This publication URL:

<https://archive-ouverte.unige.ch/unige:167126>

Publication DOI:

[10.1021/jacs.1c11248](https://doi.org/10.1021/jacs.1c11248)

Supplementary Information

Mechanistic insights into the Ligand-induced unfolding of an RNA G-quadruplex

Susanta Haldar^{1,2,†}, Yashu Zhang^{3,4,†}, Ying Xia,^{5,†} Barira Islam⁷, Sisi Liu,⁸ Francesco L Gervasio^{9,10,11}, Adrian J. Mulholland¹, Zoë A. E. Waller,⁵ Dengguo Wei^{3,4,12,13,14*}, Shozeb Haider^{5,6*}

¹ School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK

² D.E. Shaw India Private Ltd., Hyderabad, Telangana 500096, India

³ State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China

⁴ College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China

⁵ UCL School of Pharmacy, University College London, London, WC1N 1AX, UK

⁶ UCL Centre for Advanced Research Computing, University College London, London, WC1H 9RN, UK

⁷ Department of Pharmacy, School of Applied Sciences, University of Huddersfield, HD1 3DH, UK

⁸ College of Science, Huazhong Agricultural University, Wuhan, 430070, China

⁹ Department of Chemistry, University College London, London, WC1H 0AJ, UK

¹⁰ Pharmaceutical Sciences, University of Geneva, Geneva CH-1211, Switzerland

¹¹ Institute of Pharmaceutical Sciences of Western Switzerland (ISPSO), Geneva CH-1211, Switzerland

¹² National Reference Laboratory of Veterinary Drug Residues and MAO Key Laboratory for Detection of Veterinary Drug Residues, Huazhong Agricultural University, Wuhan, 430070, China

¹³ Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, 430070, China

¹⁴ Interdisciplinary Sciences Institute, Huazhong Agricultural University, Wuhan, 430070, China

* To whom correspondence should be addressed. Tel: (44)2077535883; Email: shozeb.haider@ucl.ac.uk and dgwei@mail.hzau.edu.cn

† joint first authors

	Contents	Page Number
	PC Analysis of the bound states	4
	PC Analysis on the native RNA	4
	PC Analysis on the top-face RNA-TMPyP ₄ bound state	6
	PC Analysis on the groove site RNA-TMPyP ₄ bound state	7
	Description of the free energy surface	9
Figure S1	1D Potential of mean force projected on distance, CV of the binding of TMPyP ₃ to RNA G ₄	10
Figure S2	FES showing the backbone fluctuation of native RNA G ₄ topology	11
Figure S3	FES showing the backbone fluctuations of top-face TMPyP ₄ -RNA complex	12
Figure S4	FES showing the backbone fluctuations of groove bound TMPyP ₄ -RNA complex	13
Figure S5	The FES (D vs T) showing the most populated clusters of RNA-TMPyP ₄ complex on the groove binding position	14
Figure S6	Construction of CVs	15
Figure S7	MD simulation statistics of TMPyP ₄ -groove bound states	16
Figure S8	Job Plot of RNA PQS ₁₈₋₁ and TMPyP ₄	17
Figure S9	CD melting and annealing of RNA PQS ₁₈₋₁	17

Figure S10	CD melting and annealing of RNA PQS18-1 and 2 eq TMPyP4	18
Figure S11	CD melting and annealing of RNA PQS18-1 and 5 eq TMPyP4	19
Figure S12	UV melting and annealing of RNA PQS18-1	20
Figure S13	UV melting and annealing of RNA PQS18-1 + TMPyP4 (2 eq)	20
Figure S14	UV melting and annealing of RNA PQS18-1 + TMPyP4 (5 eq)	21
Figure S15	UV melting and annealing of RNA PQS18-1 in the absence/presence of TMPyP4	21
Figure S16	UV titration spectra and plot of UV absorbance	22
Figure S17	UV absorbance at 440 nm	22
Figure S18	Fluorescence spectra of dual labelled RNA PQS18-1	23
Figure S19	Torsion definition	24
Table S1	List of atoms with parameters that are involved in building the distance and torsion CVs	24
Figure S20	N1-O6 and N2-N7 distances in quartets	25
Movie S1	The dynamics and rebinding mechanism of TMPyP4 to RNA G4	26
Movie S2	The RNA unfolding mechanism by TMPyP4	26

Supplementary Discussion

PC Analysis of the bound states

To get an insight into the dynamical behaviour of the RNA itself and further to understand the unfolding pathway found in the metadynamics simulation, we performed Principal Component Analysis (PCA) to each of the TMPyP₄ bound states from the unbiased MD simulation. The three most stable complexes between RNA and TMPyP₄ are shown in Figure 1, 2 and 3 in the main text. The most stable basin is the top-face bound state where TMPyP₄ interacts with A₁₄, A₁₄' and G₁₀, G₁₀', G₁₃ and G₁₃' bases. The next stable basins following the top-face one belong to the major groove bound states where TMPyP₄ interact individually with C₁₁ and C₁₁' bases. We performed ~1 μ s (800 ns) standard MD simulation to each of these binding poses to study their dynamics.

PC Analysis on the native RNA

PCA calculations were performed using Gromacs software package using gmx cluster facility; version-5.0.7. The starting structure of RNA G₄ for the MD simulation is taken from the X-ray crystal structure which shows an open-like backbone conformation. Only the RNA G₄ (without TMPyP₄) was used for PCA analysis. The PCA analysis on the native RNA after 800 ns of standard unbiased MD simulation is shown in Figure S2. The FES obtained using PCA₁ and PCA₂ revealed four stable basins which correspond to the changes in the conformations of the RNA backbone. A cluster analysis was performed using only the backbone heavy atoms and the most populated cluster from each basin was obtained. Basin A

depicts a X-ray crystal-like structure (PDB id: 6JH) where both the top Adenine bases, A₁₄ and A₁₄', stacked on the G₁₃ and G₁₃' bases of the first G-tetrad. Some fraying behaviour of A₁₄ and A₁₄' was also observed. In particular, A₁₄ base is flipped over G₁₃ base which disrupts the stable H-bond between both the Adenine bases allowing A₁₄ base to fray towards solvent. At the bottom of G-tetrad a crystal-like open conformation of the backbone was observed where capped bases such as C₅ and C₅' bases are engaged in H-bond formation while stacking on the G₆' and G₁ bases of the last/bottom G-tetrad, respectively (see Figure 1 from main text). In the next basin, basin B, where the backbone containing A₁₄ base is flipped and partially stacked with G₁₃ base from the groove site (Figure S2). It is worth mentioning that such conformational change is also seen in the metadynamics simulation when TMPyP₄ leave the top-face bound state for solvated state. Looking at the bottom site, one can again see an open like conformation where C₅ makes a π - π stacking interaction with C₃ base thereby increasing the H-bond distance from 2.8 Å to 3.7 Å between C₅ and C₅' base. In the open like conformation, both the C₅ and C₅' bases are intact with the formation of a H-bond interaction and with the increase or decrease of this H-bond distance, the RNA backbone goes far or come close to each other, respectively. In basin C, a large backbone conformational change is seen. Especially, two specific non-covalent interaction is formed: a H-bond between C₅' base and the oxygen atom of the backbone and a π - π stacking between C₅ and C₅' base. Since the specific H-bond between C₅ and C₅' base is now broken, and also due to the newly formed π - π stacking interaction, both the bases come close to each other bringing chain A and chain B close and forming the close like conformation. Such conformational change provides evidence of the native RNA being flexible over a short period of time. In the basin D, the conformation is more or

less similar to the basin C in which C₃ bases come closer to the C₅' base and formed a stable H-bond network. Moreover, the top-face conformation does not change throughout the simulation. Overall, the PCA analysis revealed the flexibility of RNA structure and hence it can adapt to a range of conformations.

PC Analysis on the Top-face RNA-TMPyP₄ bound state

The PCA analysis of the RNA-TMPyP₄ top-face bound state is performed after ~1 μ s (800 ns) MD simulation. The starting structure of the RNA G₄ TMPyP₄ complex for the MD simulation is again taken from the X-ray crystal structure (PDB id : 6JJH) which shows an open-like backbone conformation. The two-dimensional FES using PCA₁ and PCA₂ reveals three most stable basins (Figure S3). Basin A represents the starting conformation where the RNA backbone retains its open like conformation which is similar to the conformation found in basin A for native RNA PCA analysis (Figure S2). At the bottom, one can see that an open-like conformation exists following the two typical non-covalent interaction: a H-bond between C₅ and C₅' base and a π - π stacking interaction between the C₅' and C₃' base. TMPyP₄ is bound on top of the RNA being sandwiched in between the fraying bases A₁₄, A₁₄' and G₁₀, G₁₀', G₁₃ and G₁₃' bases from the first G-tetrad employing π - π stacking interactions. In basin B, TMPyP₄ maintained the same conformation which suggests that the ligand stabilizes the top-face loop fluctuation incorporating A₁₄ and A₁₄' bases relative to native RNA (Figure S1 for the corresponding loop fluctuation in the absence of TMPyP₄). At the bottom, the backbone still maintains an open-like conformation as basin A. The only difference is found to be a slight increase in the H-bond distance between C₅ and C₅' base which helps to maintain open backbone conformation between two single RNA strands (chain A and chain B). In

basin C, the top-face is again seen to be stable without any significant change in the structure relative to basin A and B. However, looking at the bottom, a notable difference can be seen in terms of the absolute position of the bases. In particular, the U₄ base flips from its fraying position in basin B to the π - π stacking conformation with the G₁ base allowing the C₅ base to shift its position from being stacking with G₁ base to G₁' base. These rather large conformational shifts by both U₄ and C₅ bases disrupt the π - π stacking interaction between C₅' and C₃' base, thereby allowing the backbone to shift back and finally maintain an open-like conformation again. Overall, the PCA analysis on the ~ 1 μ s MD simulation revealed that RNA was stabilized by the complex formation with TMPyP₄ which in particular do not allow large conformational changes. Moreover, RNA was maintained in an open-like conformation throughout the simulation.

PC Analysis on the Groove site RNA-TMPyP₄ bound state

We have performed the PCA analysis on the RNA-TMPyP₄ groove site binding state to compare with the native and top-face complexed state. Specifically, C₁₁ (RNA)...TMPyP₄ groove bound state has been thoroughly analysed. The FES obtained after ~ 1 μ s (800 ns) unbiased MD simulation employing PCA₁ and PCA₂ reveals three most stable basins (Figure S4). A cluster analysis was carried out as earlier using the backbone heavy atoms to obtain the representative structures of the basins. Finally, the populated clusters are thoroughly represented along the FES from each basin. Basin A represent a rather close-like conformation where both the C₅ and C₅' bases are neither engaged in the H-bond formation nor they are in the π - π stacking interaction between them; rather they are engaged in the H-bond formation with backbone oxygen atoms from opposite RNA chains while the ligand is bound to the C₁₁ base. Additionally, C₅' base forms an extra H-bond

with C₃ base. These large H-bond networks bring both the RNA strand close to each other resulting in a close-like conformation. Such a typical H-bond network is already seen in the native RNA PCA analysis (see basin D, Figure S2). In basin B, one can see a slight change in the conformation of the C₃ base and thus the backbone of overall RNA. Specifically, the C₃ base flips over the G₁ base which disrupts the H-bond between C₅' and C₃ base but the RNA still remains in its close-like conformation. In basin C, a rather large shift in some specific bases were observed compared to the overall conformational change in the backbone. The H-bond network in the top-face was broken due to a large conformational shift of the A₁₄' base, however at the same time, A₁₄ base maintained its original position by stacking on the G₁₃ base of the first G-tetrad. These findings suggests that the top-face loop is stabilized only by the formation of RNA-TMPyP₄ top-face complex. Further, at the bottom, both the C₅ and C₅' bases are engaged in the π - π stacking interaction maintaining the overall close-like conformation. Such a conformation again resembles with one of the native RNA conformations found earlier (see basin C and D, Figure S2). Overall, it is worth mentioning that the starting structure for the RNA-TMPyP₄ groove-bound complex, taken from the WT-MetaD simulation, is a close-like conformational state which stabilizes the RNA by the complex formation and maintains its conformation throughout the simulation. Moreover, it is also evident that TMPyP₄ groove-bound states always maintain a close-like conformation which is opposite to the top-face bound states.

Description of the Free Energy Surface

The D vs T FES does not separate the groove binding sites with top-face ones (Figure S5). Three most populated clusters are shown which represent the basin M1 in the groove position of the FES ranges from 1.4 nm to 2.1 nm as shown in Figure 1 (see main text). Cluster 1 represents the RNA-TMPyP₄ groove binding mode where TMPyP₄ makes π - π stacking with C₁₁ base from chain A. In particular, the indole rings from the porphyrin ring in TMPyP₄ is taking part in the π - π stacking with C₁₁ base from RNA. Cluster 2 appear for an RNA-TMPyP₄ complex with top-face binding conformation. Here the ligand is stacking on the top-face of the RNA interacting with fraying adenine base such as A₁₄' and partially with A₁₄. Cluster 3 constitute for another RNA-TMPyP₄ groove binding conformation where again a π - π stacking is seen in between C₁₁' base from chain B and TMPyP₄.

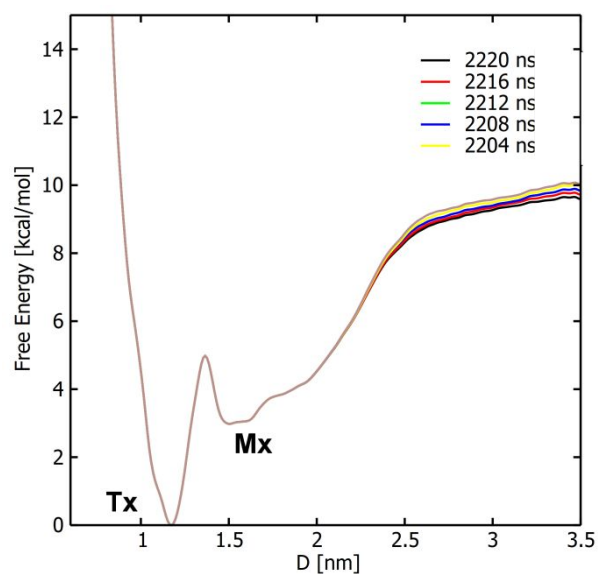


Figure S1: The one-dimensional potential of mean force (PMF) projected on the distance (D), CV of the binding of TMPyP4 to RNA G4. The free energy profile is monitored in the last 16 ns going from 2204 to 2220 ns showing the convergence of the free energy with the difference in the free energy being <0.4 kcal/mol. The last replica (walker-4) is ignored due to sampling of the unfolded region.

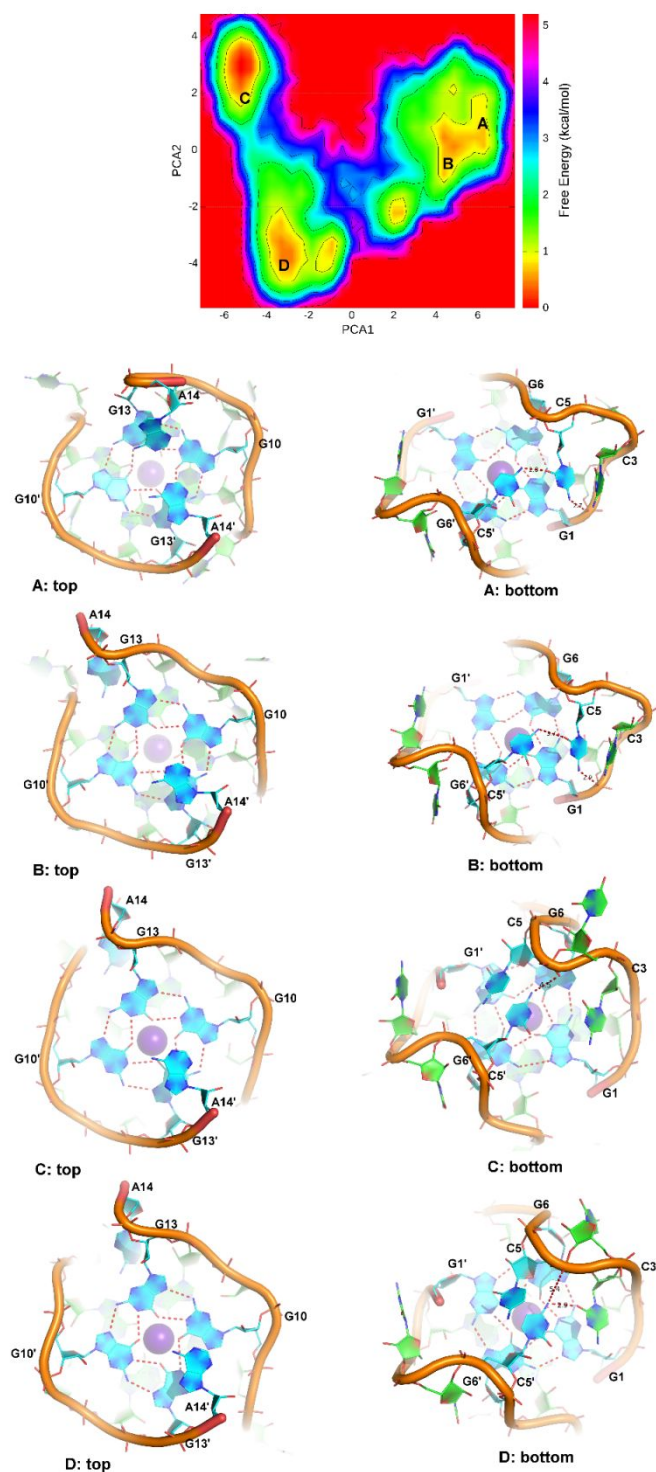


Figure S2: FES showing the backbone fluctuations of native RNA G₄ topology mapped with PCA₁ and PCA₂. A variation from open to close-like conformation has been sampled in the long 800 ns of MD simulation showing the flexibility of RNA G₄ over a short period of time.

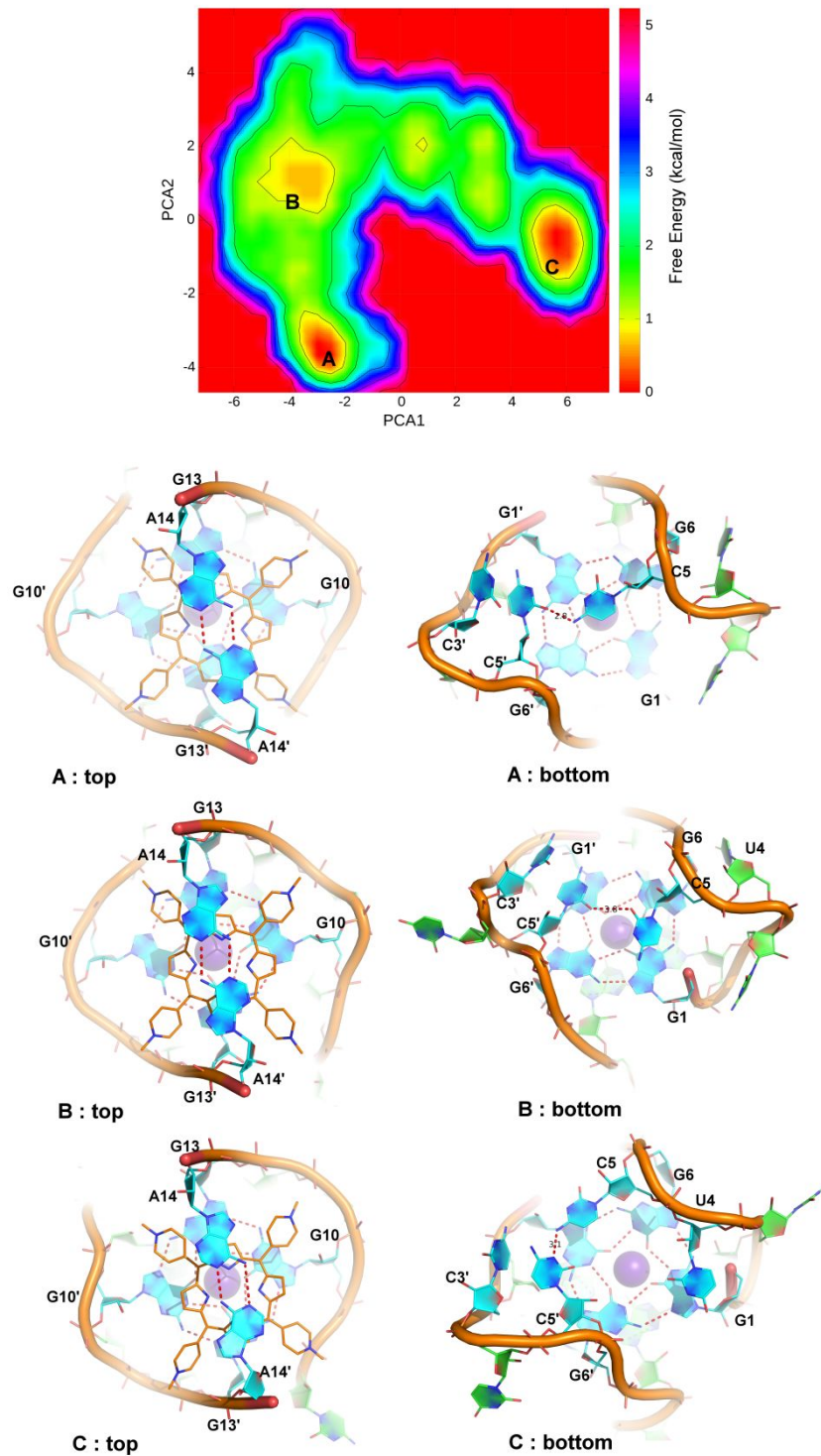


Figure S3: FES showing the backbone fluctuations of top-face TMPyP₄-RNA complex mapped with PCA₁ and PCA₂.

TMPyP₄ mostly seen to stabilize the RNA G₄ binding on top-face through stacking interaction with the G-quartet.

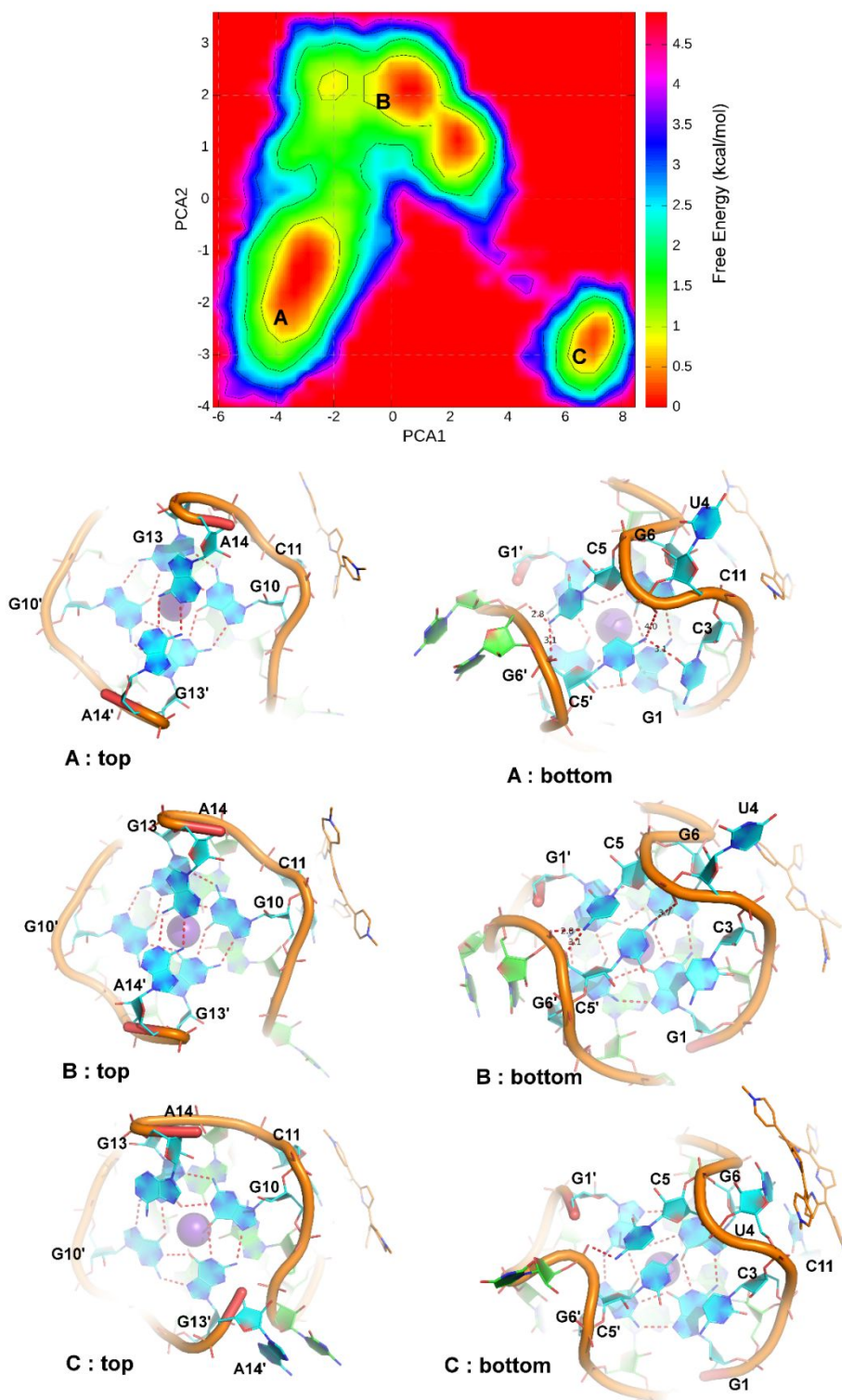


Figure S4: FES showing the backbone fluctuations of groove bound TMPyP₄-RNA complex mapped with PCA₁ and PCA₂.

TMPyP₄ seen to stabilize the RNA G₄ through groove binding as RNA G₄ topology remains to be in the close-like conformation throughout the 800ns MD simulation.

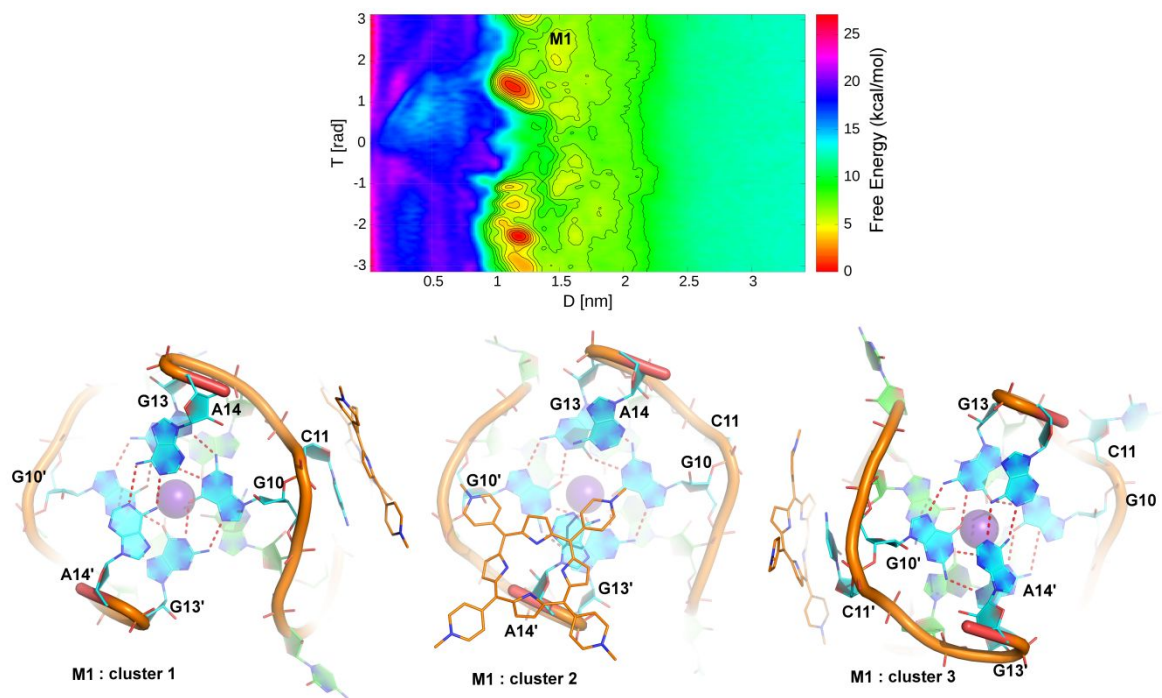
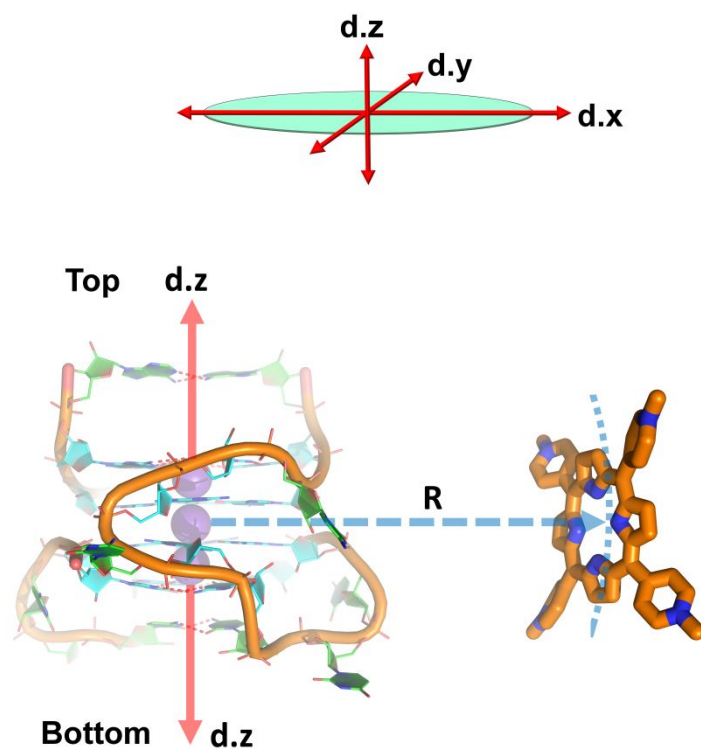


Figure S5: The FES (D vs T) showing the three most populated clusters of the RNA-TMPyP₄ complex on the groove binding position.



$$R = \sqrt{(d.x)^2 + (d.y)^2}$$

Figure S6: Construction of the CVs such as POA (Position on the axis) and DFA (Distance from the axis). POA defines as $d.z$ in nm where the z -axis is aligned with the RNA helix. DFA defines as R in nm where it is calculated as the distance of TMPyP₄ from the $d.z$ axis (on the xy plane) thereby describing both the top-face and groove bound states.

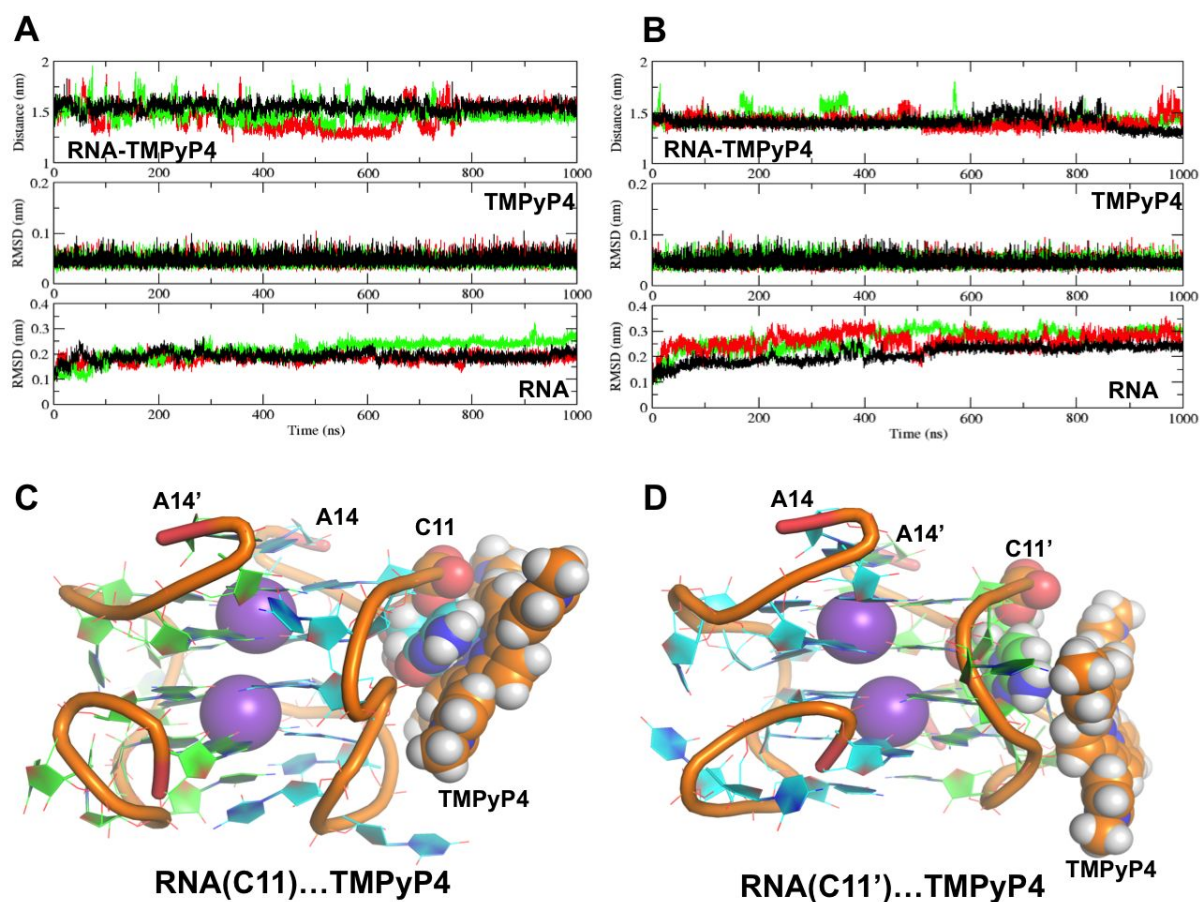


Figure S7: The MD simulation statistics of TMPyP₄-groove bound states. (a) TMPyP₄-RNA(C₁₁) G₄: The distance between TMPyP₄ and RNA G₄ with an average distance of 1.65 nm (top), TMPyP₄ RMSD relative to the starting position with an average RMSD being 0.05 nm (middle), and backbone RMSD of RNA G₄ relative to the starting position with an average RMSD being 0.19 nm (bottom). (b) TMPyP₄-RNA (C_{11'}) G₄: The distance between TMPyP₄ and RNA G₄ with an average distance of 1.72 nm (top), TMPyP₄ RMSD relative to the starting position with an average RMSD being 0.05 nm (middle), and backbone RMSD of RNA G₄ relative to the starting position with an average RMSD being 0.22 nm (bottom). (c) and (d) a representative structure of TMPyP₄...RNA G₄ complex bound to C₁₁ and C_{11'} groove bound site.

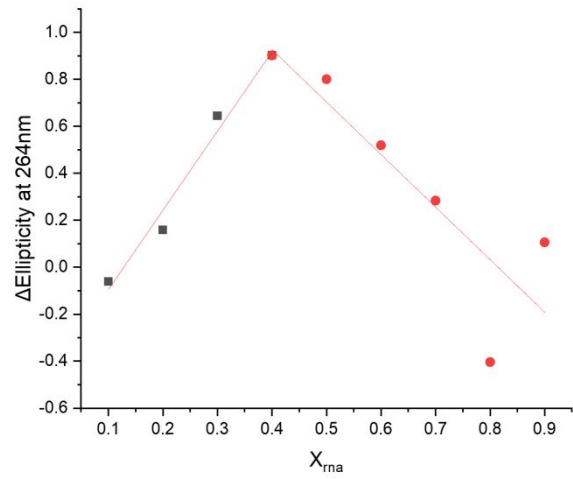


Figure S8. Example Job plot of RNA PQS18-1 and TMPyP4. Experiments performed using CD in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl.

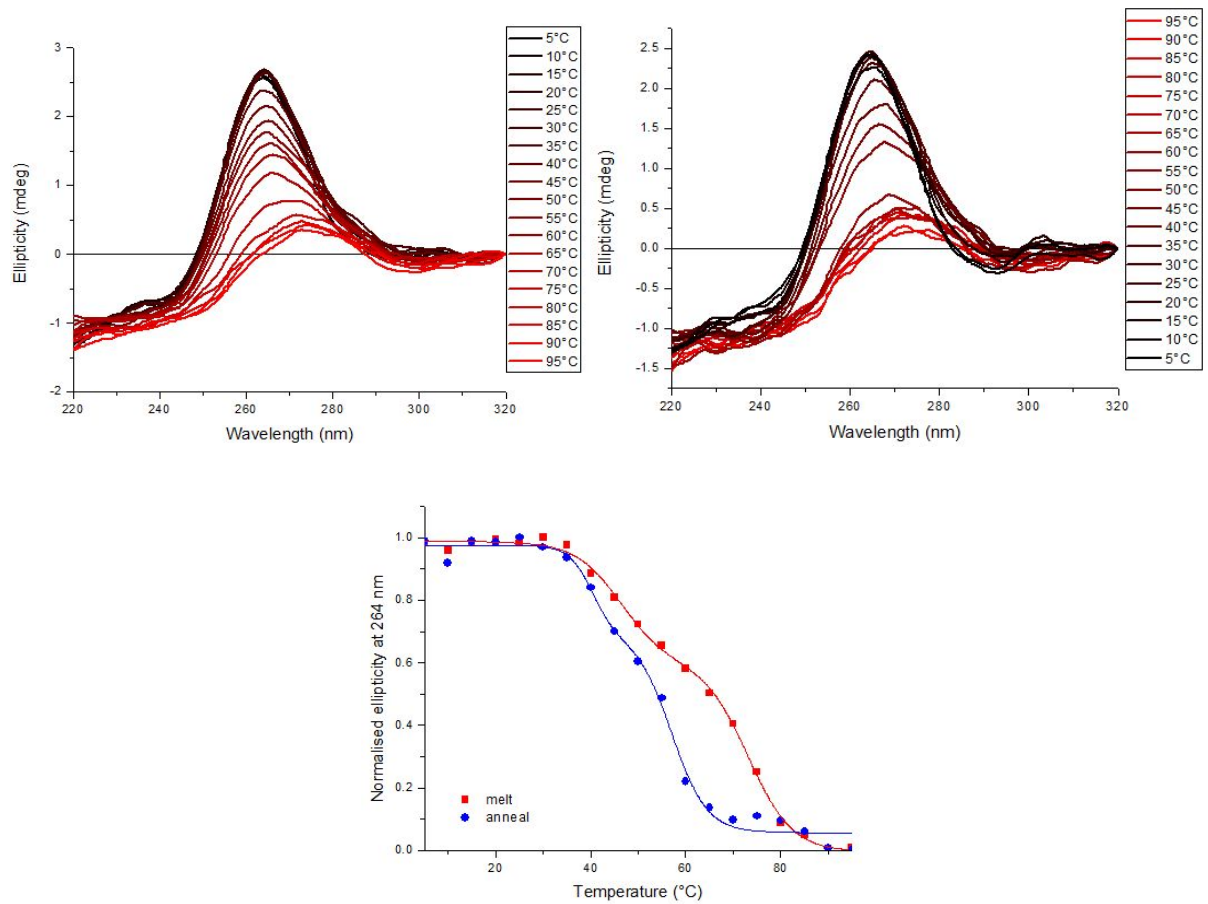


Figure S9. Example CD melting (Left) and annealing (right) of RNA PQS18-1 and melting/annealing curves (bottom). Experiments performed at 10 μ M RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl.

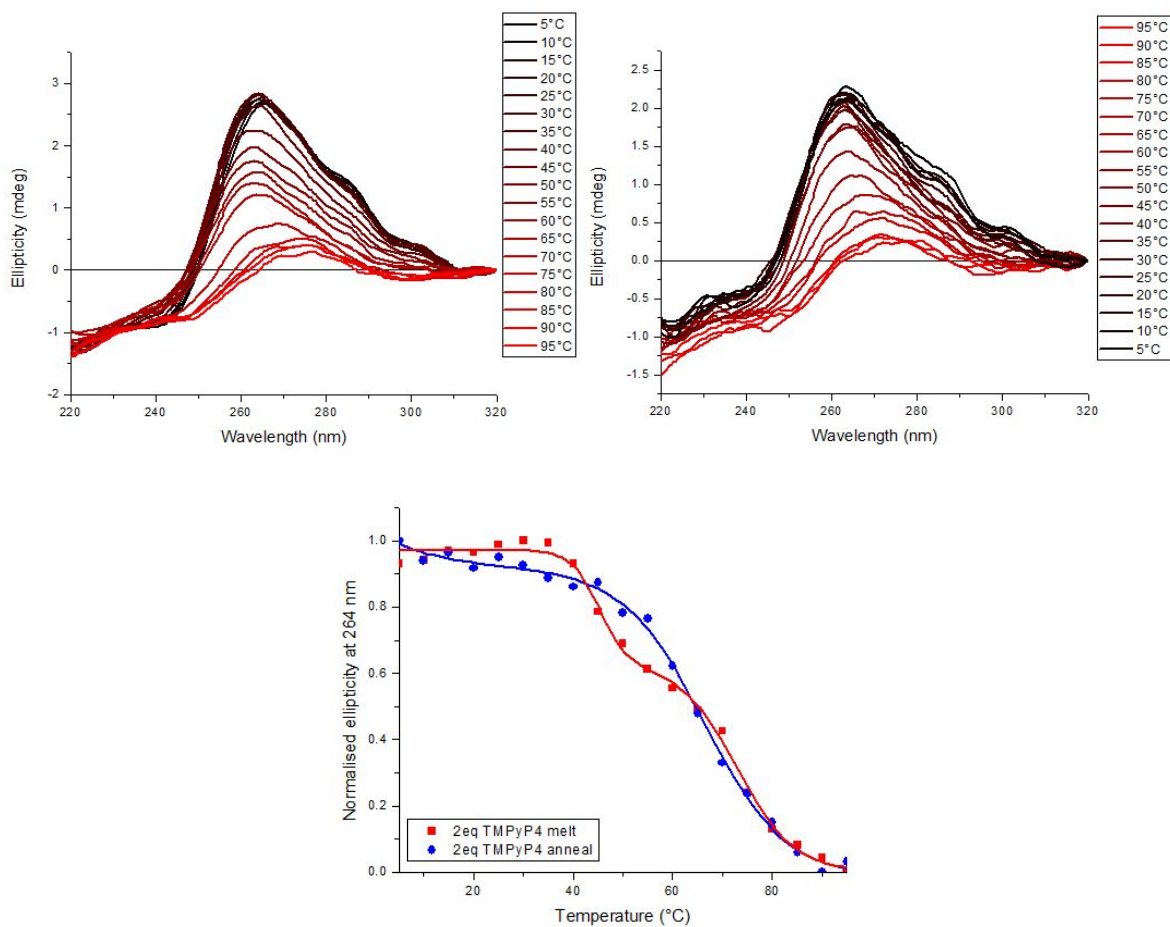


Figure S10. Example CD melting (Left) and annealing (right) of RNA PQS₁₈₋₁ and 2 eq TMPyP₄ and melting/annealing curves (bottom). Experiments performed at 10 μ M RNA with 20 μ M TMPyP₄ in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl.

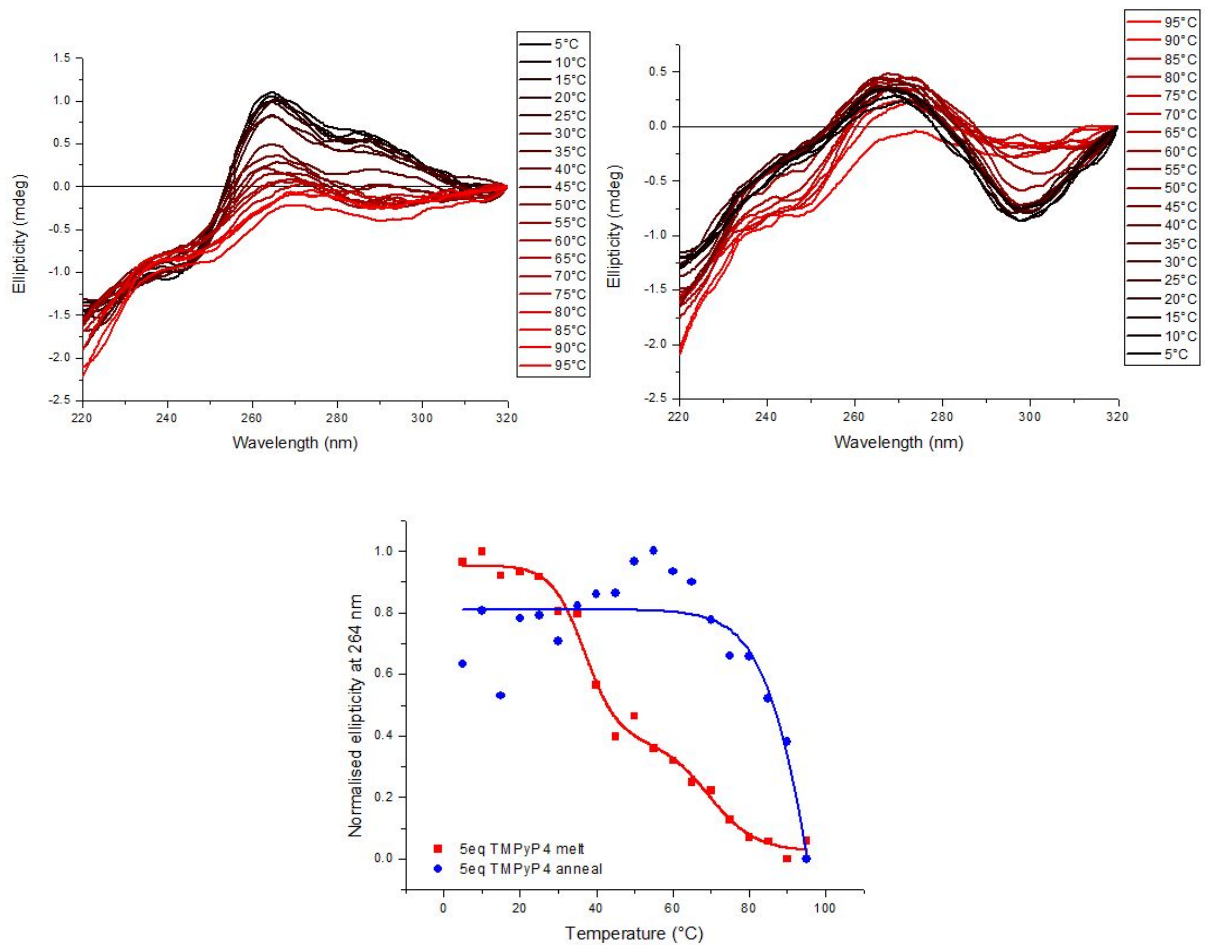


Figure S11. Example CD melting (Left) and annealing (right) of RNA PQS₁₈₋₁ and 5 eq TMPyP₄ and melting/annealing curves (bottom). Experiments performed at 10 μ M RNA with 50 μ M TMPyP₄ in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl.

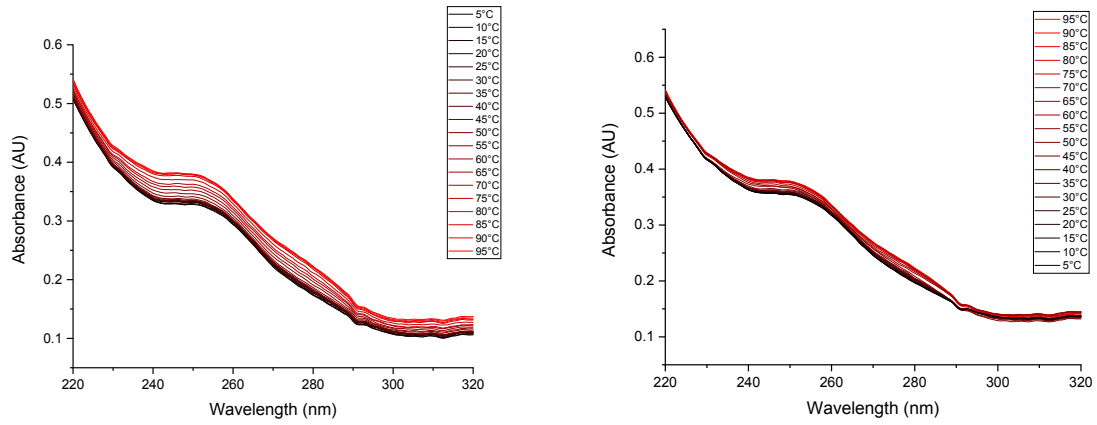


Figure S12. Example UV melting and annealing of RNA PQS18-1. Experiments performed at 10 μ M RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl.

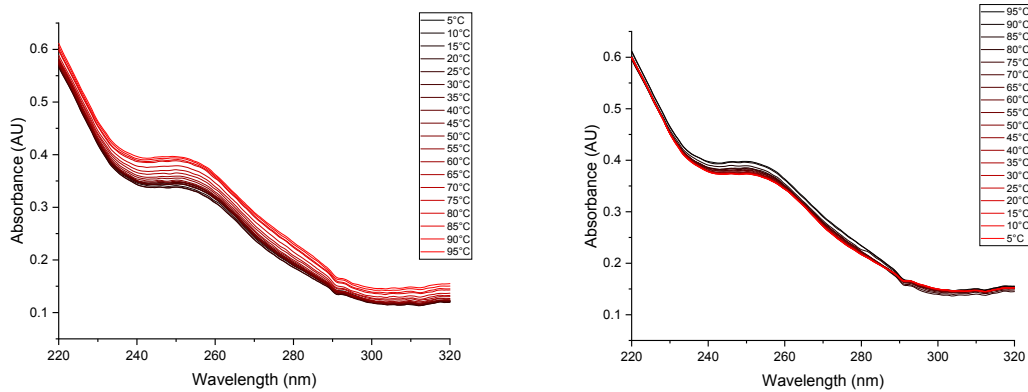


Figure S13. Example UV melting and annealing of RNA PQS18-1 in the presence of TMPyP4. Experiments performed at 10 μ M RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl and 2 eq of TMPyP4 (20 μ M).

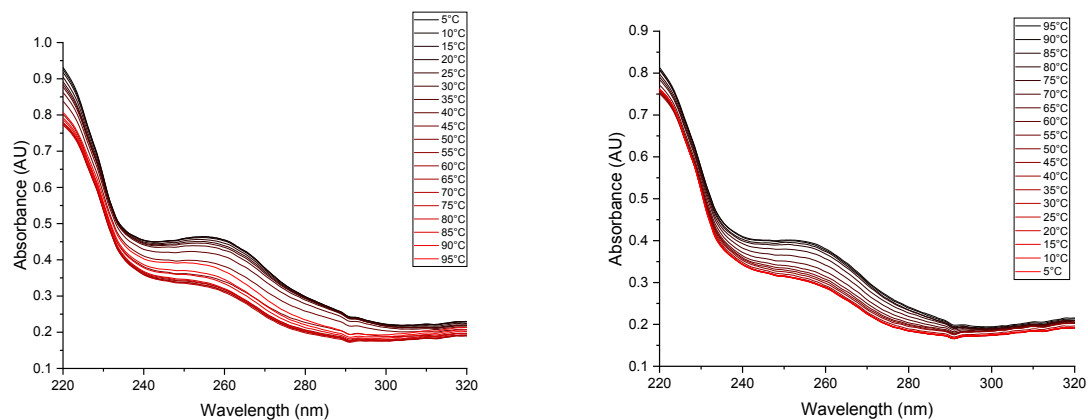


Figure S14. Example UV melting and annealing of RNA PQS18-1 in the presence of TMPyP4. Experiments performed at 10 μ M RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl and 5 eq of TMPyP4 (50 μ M).

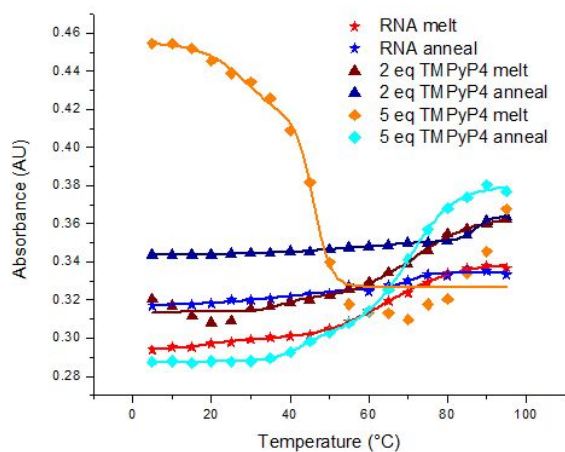


Figure S15. Example UV melting and annealing of RNA PQS18-1 in the absence and presence of TMPyP4. Experiments performed at 10 μ M RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl and 2 or 5 eq of TMPyP4 (20 and 50 μ M respectively) absorbance measured at 260 nm.

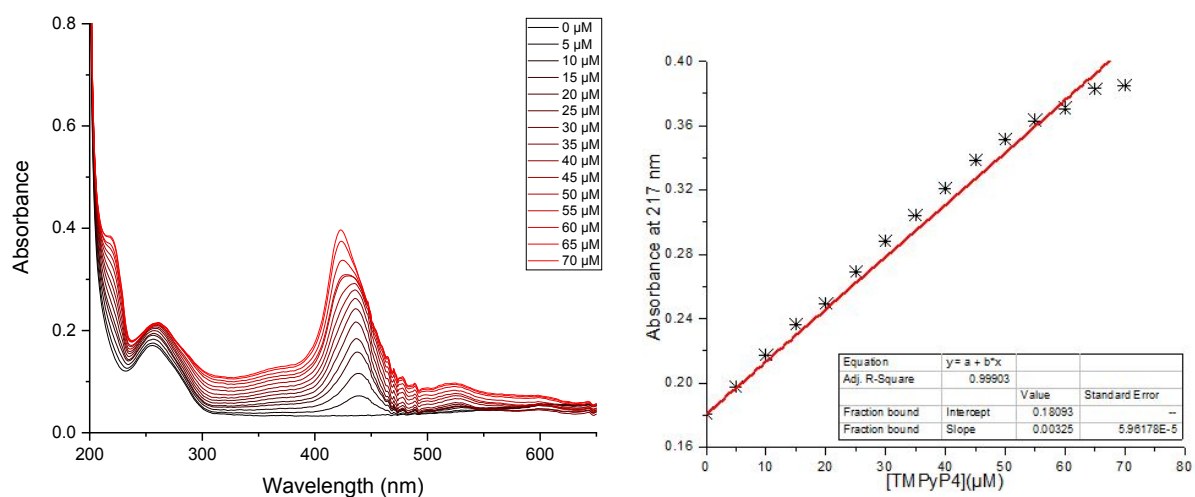


Figure S16. Example UV titration spectra and plot of UV absorbance at 217 nm of RNA PQS₁₈₋₁ in the presence of TMPyP₄.

Experiments performed at 10 μM RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl and 0 to 70 μM of TMPyP₄. Data fitted to a linear function.

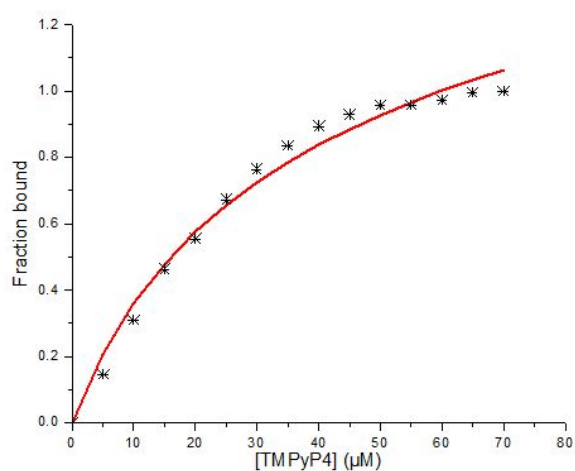


Figure S17. Example plot of UV absorbance at 440 nm of RNA PQS₁₈₋₁ in the presence of TMPyP₄. Experiments performed

at 10 μM RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl and 0 to 70 μM of TMPyP₄.

Data fitted to a two inequivalent sites binding model.

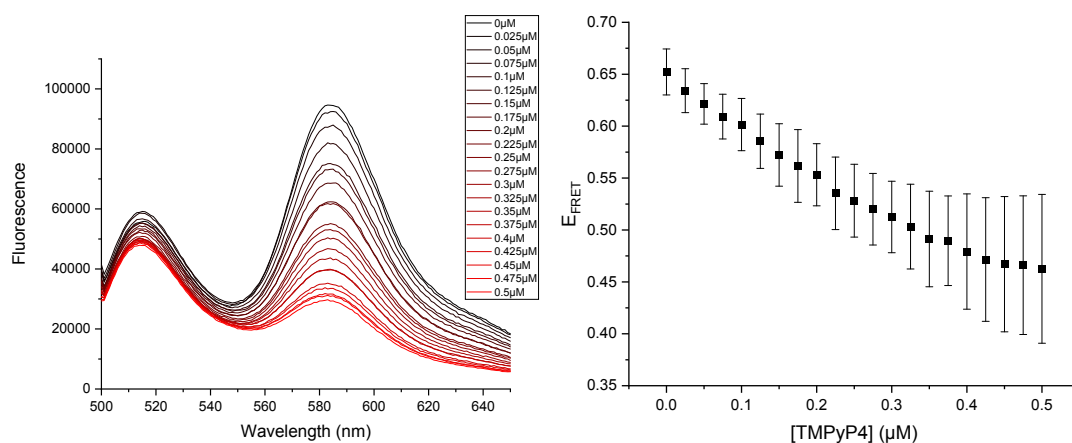


Figure S18. Example fluorescence spectra of dual-labelled RNA PQS18-1 in the presence of TMPyP4 (left) Relative FRET efficiency with increasing concentration of TMPyP4 (right). E_{FRET} Experiments performed at 0.1 μ M RNA in 10 mM lithium cacodylate buffer, pH 7.0 supplemented with 100 mM of KCl and 0 to 0.5 μ M of TMPyP4.

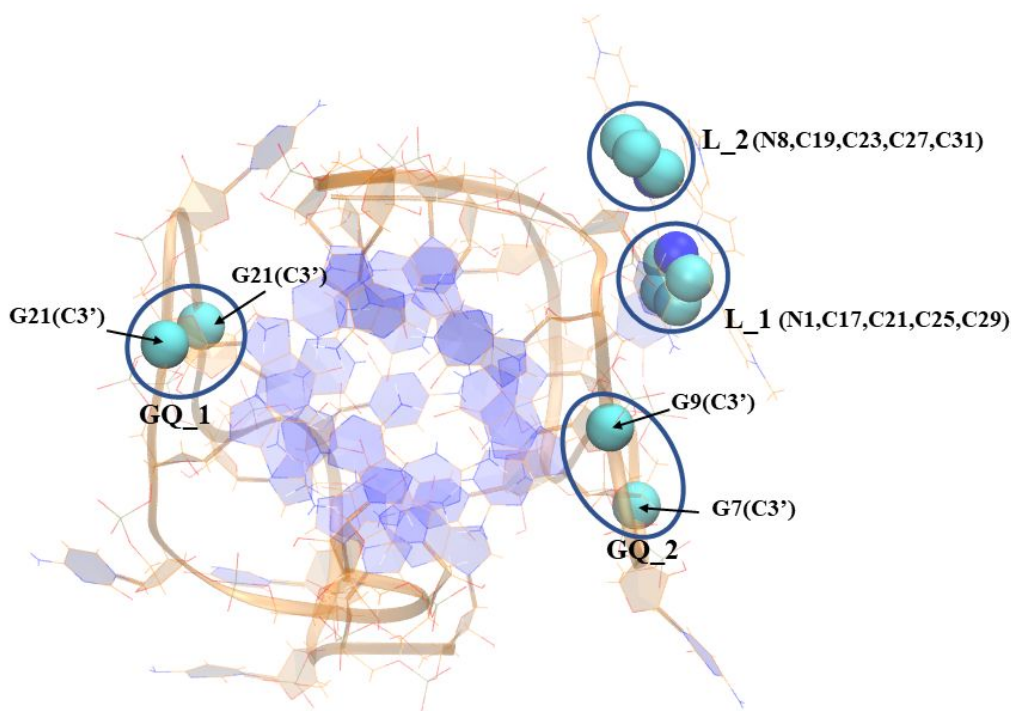


Figure S19: The Torsion definition of GQ₁, GQ₂, L₁, L₂ and their corresponding atom definition for the COM representation.

Table S1: List of atoms with parameters that are involved in building the Distance (D) and Torsion (T) CVs in the WT-MetaD simulation. The plumed.dat and the input.gro file associated with CVs is also attached in the Supplementary Information. COM stands for Center of Mass. GQ₁, GQ₂ and L₁, L₂ represents two points each from the RNA G4 and TMPyP4, respectively.

Collective Variables (CV)	Parameters	CVs
Distance (D) in the WT-MetaD	RNA G4: COM{[G7(C3')], [G9(C3')], [G21(C3')], [G23(C3')]} PMPyP4: COM (all heavy atoms)	D: RNA G4.....TMPyP4
Torsion (T) in the WT-MetaD	GQ₁: COM{[G21(C3')], [G23(C3')]} GQ₂: COM{[G7(C3')], [G9(C3')]} L₁: COM{N1, C17, C21, C25, C29} L₂: COM{N8, C19, C23, C27, C31}	T: GQ ₁ , GQ ₂ , L ₁ , L ₂

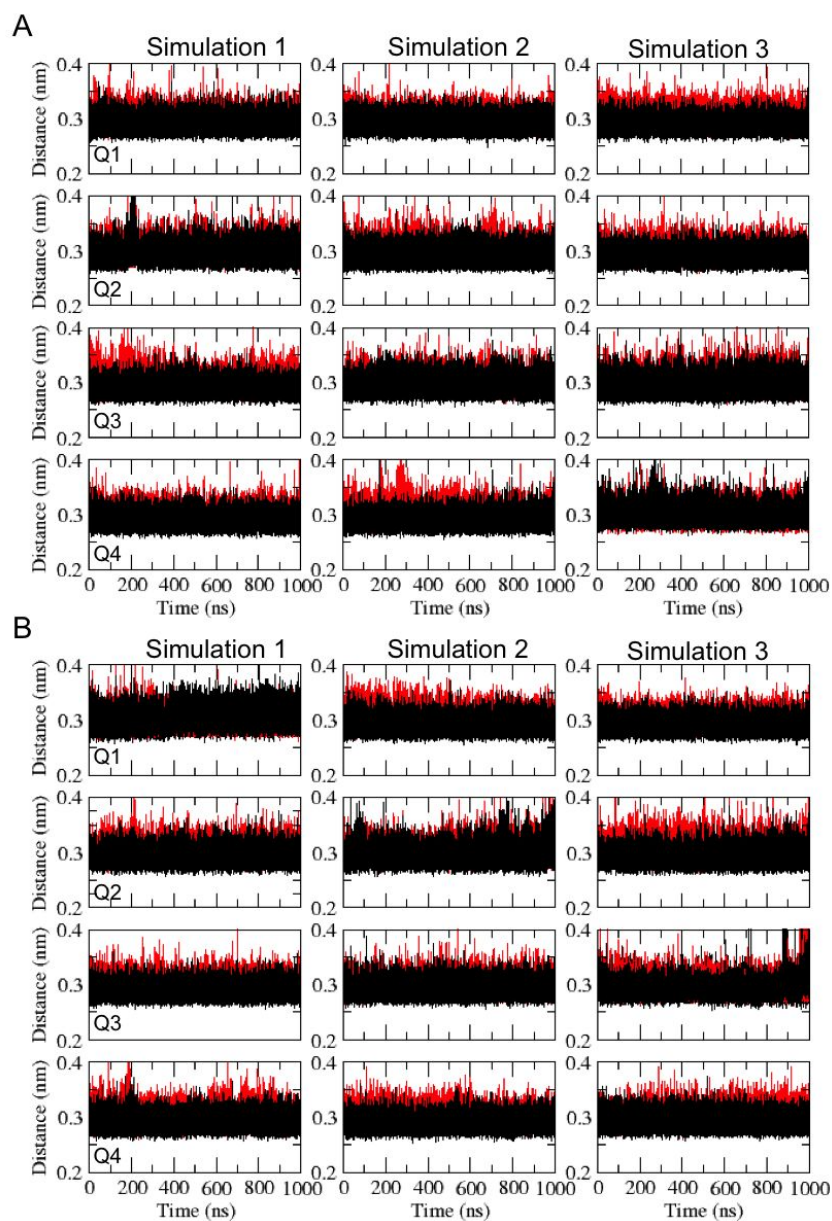


Figure S20: Hydrogen bond distances between N1-O6 (black) and N2-N7 (red) atoms in the G4 quartets. The distances have been calculated from the unbiased simulations when TMPyP₄ is stacked with (A) C₁₁ and (B) C₁₁' bases. Each row represents a quartet, while each column is a replicate simulation.

Movie S1: *The dynamics and the rebinding mechanism of TMPyP₄ to RNA G₄. TMPyP₄ is being ctached by C₁₁ base from the bulk and at the end reaching to the top-face bound state.*

Movie S2: *The RNA G₄ unfolding mechanism by TMPyP₄ through major π - π stacking with Guanine bases from G-tetrads.*