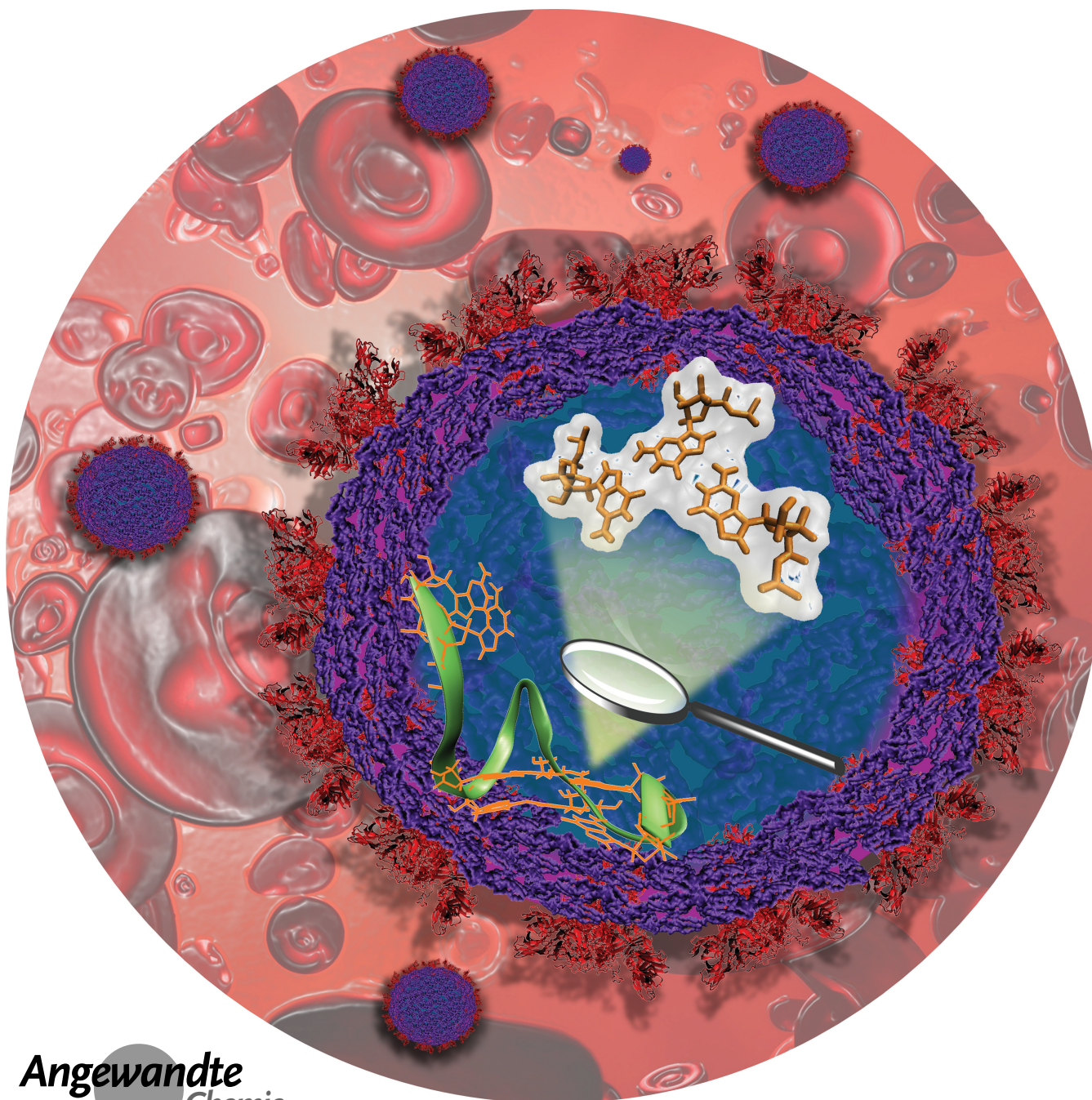


VIP **Biophysics** Very Important Paper

Experimental and Computational Evidence of a Stable RNA G-Triplex Structure at Physiological Temperature in the SARS-CoV-2 Genome

Marco Campanile, Roberto Improta, Luciana Esposito, Chiara Platella, Rosario Oliva, Pompea Del Vecchio, Roland Winter, and Luigi Petraccone*

Angewandte
International Edition
Chemie

Abstract: RG1 is a quadruplex-forming sequence in the SARS-CoV-2 genome proposed as possible therapeutic target for COVID-19. We demonstrate that the dominant conformation of RG1 under physiological conditions differs from the parallel quadruplex previously assumed. Through comprehensive investigations employing CD, UV, NMR, DSC, gel electrophoresis, MD simulations, in silico spectroscopy and the use of truncated RG1 sequences, we have identified this stable conformation as an RNA G-triplex composed of two G-triads. We believe this previously unreported RNA structure could serve as a novel therapeutic target. Our findings open new avenues for further studies on the presence and biological role of RNA G-triplexes in vivo.

DNA and RNA G-rich sequences can form noncanonical G-quadruplex (G4) structures, formed by stacking of G-tetrad layers of four guanines locked by Hoogsteen-type hydrogen bonds and further stabilized by monovalent cations such as K^+ or Na^+ .^[1] The formation of such structures is involved in the regulation of gene expression as well as in many other cellular processes, including RNA metabolism, epigenetic states, and related diseases.^[2–4] Moreover, the biological role of G4s is not limited to humans but extends to yeasts, bacteria, and viruses.^[5] Therefore, G4s emerged as promising therapeutic targets for drug development.^[3,5–7] Bioinformatics analysis has been widely used to search for potential G-quadruplex forming sequences in the genomes of several organisms.^[8–11] The minimal requirement for a G-rich sequence to be a putative G-quadruplex forming sequence is to contain four G-tracts (regions of at least two adjacent guanines). However, a G-rich sequence containing only three G-tracts can potentially form a G-triplex, another non-canonical structure strictly related to the G-quadruplex.^[12,13] G-triplexes are formed by several layers of G-triads each containing Hoogsteen-like hydrogen bonds similar to the ones forming a G-tetrad.^[12–15] Although G-triplexes share similarities with G-quadruplexes, they offer unique structural and functional properties that deserve further investigation. Putative G-triplex sequences clearly include the number of the putative G-quadruplex sequences in a genome but reasonably exceed this number. Understanding their role in biological processes and their potential as therapeutic targets could open new avenues for drug development. Despite their potential biological role, studies on G-triplexes are extremely limited. The formation of a stable G-triplex structure has been demonstrated for few DNA sequences^[13–17] and is often invoked as intermediate in DNA G-quadruplex folding/unfolding processes.^[18–20] To our knowledge, stable G-triplexes have never been reported for an RNA sequence.

In this work, we investigated in detail the conformational behavior of the RG1 sequence (GGCUGG-CAAUGGCGG), an RNA G-rich sequence in the SARS-CoV-2 genome located in the coding sequence region of the viral nucleocapsid phosphoprotein (denoted N-protein).^[21] Previous experimental and computational studies suggested that RG1 forms a monomolecular two G-tetrads parallel quadruplex.^[21,22] Remarkably, our data reveals that RG1 forms a stable RNA G-triplex structure at physiological temperature whereas the corresponding G-quadruplex conformation is not significantly populated.

We firstly characterized the RG1 conformation in the presence of 10 mM phosphate buffer pH 7.4 and 100 mM of K^+ (K^+ buffer) at 0 °C, conditions that should stabilize a folded G-quadruplex conformation. The thermal UV/Vis difference spectrum (TDS) profile of RG1 is indeed characteristic of a G4 structure with maxima at 248 nm and 274 nm and a minimum around 297 nm (Figure 1A). Further, the value of the TDS factor $\Delta A_{240} \text{ nm} / \Delta A_{295} \text{ nm} > 4$ is consistent with the parallel topology.^[23] The circular dichroism (CD) spectrum is also consistent with a parallel G-quadruplex with a maximum at 268 nm and a minimum at 239 nm (Figure 1B, black spectrum).^[21] We further checked the molecularity of RG1 by means of native polyacrylamide gel electrophoresis (PAGE) experiments using as references

plexes, they offer unique structural and functional properties that deserve further investigation. Putative G-triplex sequences clearly include the number of the putative G-quadruplex sequences in a genome but reasonably exceed this number. Understanding their role in biological processes and their potential as therapeutic targets could open new avenues for drug development. Despite their potential biological role, studies on G-triplexes are extremely limited. The formation of a stable G-triplex structure has been demonstrated for few DNA sequences^[13–17] and is often invoked as intermediate in DNA G-quadruplex folding/unfolding processes.^[18–20] To our knowledge, stable G-triplexes have never been reported for an RNA sequence.

In this work, we investigated in detail the conformational behavior of the RG1 sequence (GGCUGG-CAAUGGCGG), an RNA G-rich sequence in the SARS-CoV-2 genome located in the coding sequence region of the viral nucleocapsid phosphoprotein (denoted N-protein).^[21] Previous experimental and computational studies suggested that RG1 forms a monomolecular two G-tetrads parallel quadruplex.^[21,22] Remarkably, our data reveals that RG1 forms a stable RNA G-triplex structure at physiological temperature whereas the corresponding G-quadruplex conformation is not significantly populated.

We firstly characterized the RG1 conformation in the presence of 10 mM phosphate buffer pH 7.4 and 100 mM of K^+ (K^+ buffer) at 0 °C, conditions that should stabilize a folded G-quadruplex conformation. The thermal UV/Vis difference spectrum (TDS) profile of RG1 is indeed characteristic of a G4 structure with maxima at 248 nm and 274 nm and a minimum around 297 nm (Figure 1A). Further, the value of the TDS factor $\Delta A_{240} \text{ nm} / \Delta A_{295} \text{ nm} > 4$ is consistent with the parallel topology.^[23] The circular dichroism (CD) spectrum is also consistent with a parallel G-quadruplex with a maximum at 268 nm and a minimum at 239 nm (Figure 1B, black spectrum).^[21] We further checked the molecularity of RG1 by means of native polyacrylamide gel electrophoresis (PAGE) experiments using as references

plexes, they offer unique structural and functional properties that deserve further investigation. Putative G-triplex sequences clearly include the number of the putative G-quadruplex sequences in a genome but reasonably exceed this number. Understanding their role in biological processes and their potential as therapeutic targets could open new avenues for drug development. Despite their potential biological role, studies on G-triplexes are extremely limited. The formation of a stable G-triplex structure has been demonstrated for few DNA sequences^[13–17] and is often invoked as intermediate in DNA G-quadruplex folding/unfolding processes.^[18–20] To our knowledge, stable G-triplexes have never been reported for an RNA sequence.

[*] M. Campanile, C. Platella, R. Oliva, P. Del Vecchio, L. Petraccone
Department of Chemical Sciences
University of Naples Federico II
Via Cintia 4, 80126 Naples, Italy
E-mail: luigi.petraccone@unina.it

R. Improta, L. Esposito
Institute of Biostructure and Bioimaging
National Research Council CNR
Via P. Castellino 111, 80131 Naples, Italy

R. Winter
Physical Chemistry I—Biophysical Chemistry, Department of
Chemistry and Chemical Biology
TU Dortmund University
Otto-Hahn Street 4a, 44227 Dortmund, Germany

© 2024 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

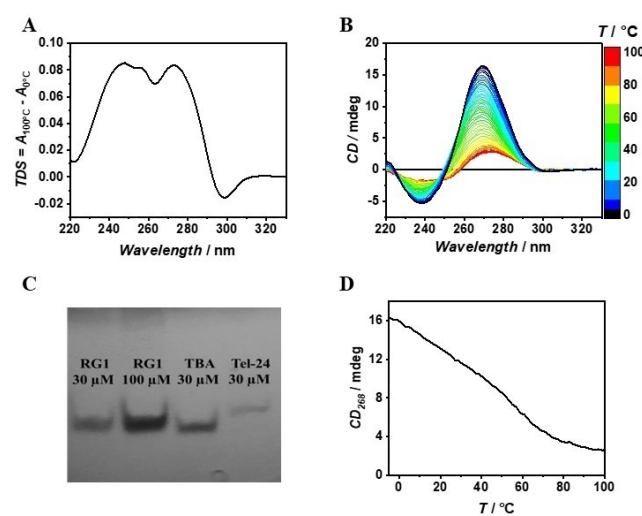


Figure 1. TDS (A), 3D melting (B), Native PAGE (C) and CD melting curve (D) of RG1. All the samples were prepared in K^+ buffer.

the two G-tetrads TBA and the three G-tetrads Tel-24 monomolecular quadruplexes.^[24] The gel electrophoresis results clearly show that RG1 migrates as a single band up to 100 μM strand concentration with a mobility similar to TBA (Figure 1C). Overall, our UV, CD and PAGE results indicate that RG1 forms a monomolecular two-G-tetrads parallel G-quadruplex in agreement with a previously proposed structural model.^[22]

To verify whether this structure is stable at physiological temperature, we then recorded the 3D melting profile by means of temperature-dependent CD measurements (Figure 1B). Remarkably, the CD melting profile at 268 nm shows a multiphasic shape, suggesting that RG1 unfolding is not a simple two-state process (Figure 1D). To determine the number of significant species needed to properly analyze the thermal unfolding, all the collected temperature dependent spectra were subjected to Singular Value Decomposition (SVD) procedure.^[25] SVD analysis clearly reveals that at least three main conformations are involved in the whole RG1 unfolding process (see section 8.1 in the Supporting Information for further discussion). We further verified the complexity of the RG1 unfolding by following the process by means of Differential Scanning Calorimetry (DSC). As expected, the obtained DSC profile clearly shows two transitions centered at about 25 $^{\circ}\text{C}$ and 60 $^{\circ}\text{C}$ confirming the biphasic behavior of the RG1 unfolding (Figure 2C). On the basis of DSC, CD and SVD results, we modeled the RG1 unfolding with a three-states mechanism (Figure 2A) involving an intermediate species (see section 8.2 in the Supporting Information for details). The proposed mechanism matched the experimental melting profiles with great accuracy (Figure 2B,C). Inspection of the thermodynamic parameters (Table 1) reveals that RG1 unfolds with a total enthalpy change of 210 kJ/mol, a value in agreement with the one expected for a two G-tetrads RNA quadruplex.^[26,27] Notably, about 65% of the enthalpy and entropy change is involved in the unfolding of the intermediate state suggesting that it is a well-structured species somehow resembling the RG1 G-quadruplex. Figure 2D shows the molar fraction of each species as function of temperature. Interestingly, the intermediate species represents by far the most populated one at physiological temperature (> 75%). This is significant as this RG1 conformation could be the most suitable target for a future antiviral strategy. For this reason, we explored the nature of the intermediate species in the RG1 unfolding in more detail.

First, we recorded the TDS spectrum, setting as a lower limit the temperature at which the intermediate state (**I**) of RG1 is at the highest molar fraction ($\sim 40^{\circ}\text{C}$), so that the resulting TDS should mainly reflect the structural features of this state. Interestingly, the TDS does not show the minimum at 297 nm (present at 0°C) revealing that **I** is not a quadruplex structure (Figure 3A). We then determined the CD spectrum of **I** by knowing the experimental CD spectra of the quadruplex and unfolded RG1 and the molar fraction of each species as a function of temperature (see section 8.3 in the Supporting Information for details). The spectrum of **I** (Figure 3B) is remarkably similar to that of the RG1 quadruplex revealing that the intermediate species has a

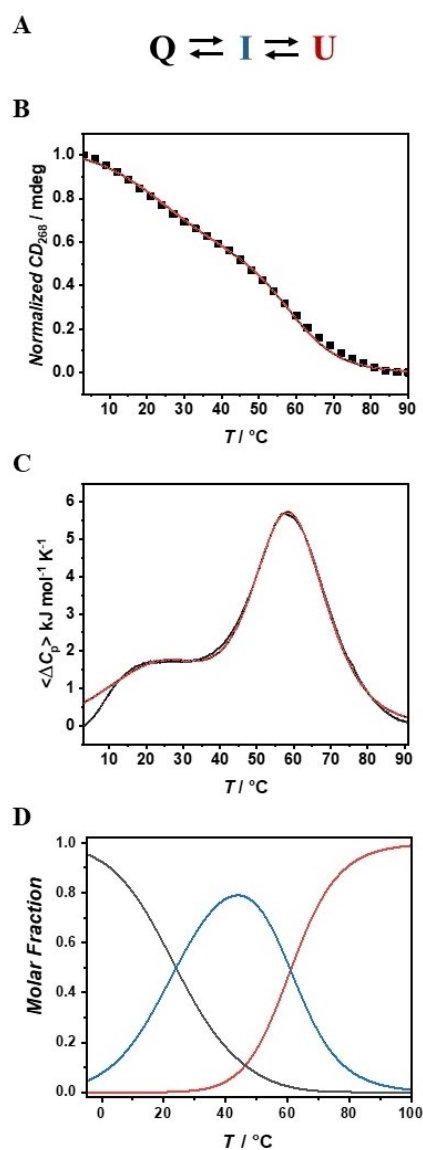


Figure 2. Three-states mechanism for the RG1 unfolding (A). CD (B) and DSC melting experiments (C); red lines represent the best fitting curves (with adjusted $R^2 > 0.99$). Temperature-dependent molar fractions (D).

Table 1: Thermodynamic parameters for the RG1 unfolding obtained from the global fit analysis of the CD and DSC data

Transition	T/ $^{\circ}\text{C}$	$\Delta H^{\circ}/\text{kJ mol}^{-1}$	$\Delta S^{\circ}/\text{kJ K}^{-1} \text{mol}^{-1}$
Q \rightarrow I	26 \pm 2	71 \pm 8	0.24 \pm 0.03
I \rightarrow U	59.2 \pm 0.4	139 \pm 12	0.42 \pm 0.04
Q \rightarrow U	\	210 \pm 14	0.66 \pm 0.05

structure similar to the one of the parallel G-quadruplex. However, its CD intensity is about 40% lower than the one of the RG1 quadruplex, suggesting that the intermediate state is not composed of fully formed parallel G-tetrads. All these observations argue in favor of a G-triplex as a reasonable structural candidate for **I**, as such structure could be virtually obtained by removing just two consecutive

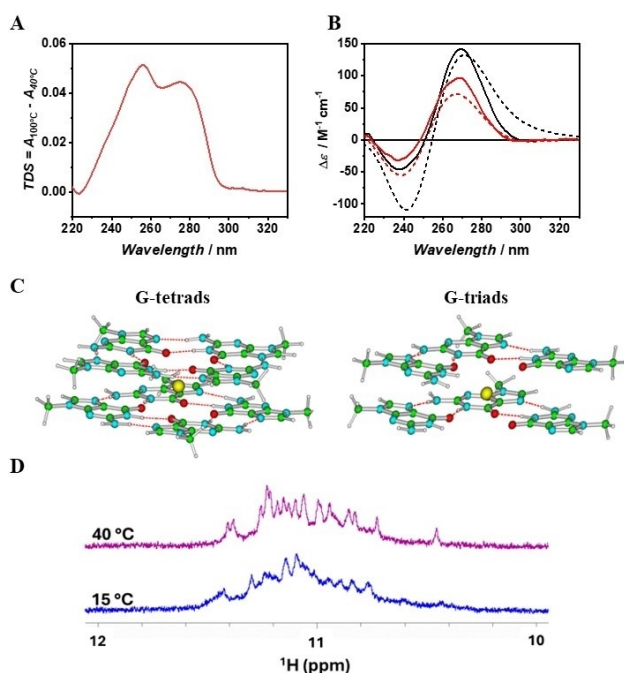


Figure 3. TDS of the intermediate species (A). Experimental (continuous lines) and theoretical (short dash lines) CD spectra of the RG1 quadruplex (in black) and its unfolding intermediate (in red) (B). Schematic drawing of the $G4X2-K^+$ (on the left) and $G3X2-K^+$ (on the right) models (C). 1H NMR spectra of RG1 in the region of Hoogsteen imino peaks (D).

guanine bases from the G-quadruplex. The resulting parallel G-triplex, as observed in our experiment, should have a similar shape but a less intense CD spectrum.^[17] To support our hypothesis, we simulated by means of an hybrid QM/MM approach (see section 10.1 in the Supporting Information for computational details) the expected CD spectral changes for a G-quadruplex–G-triplex transition assuming that the CD spectrum of each species is mainly due to the G-core. Briefly, we simulated the ECD spectrum for the two model systems depicted in Figure 3C, containing two guanine tetrads ($2 \times G4-K^+$) or two triads ($2 \times G3-K^+$) coordinating one central K^+ ion. As Figure 3B shows, the simulated spectra well reproduce the differences between the spectra of RG1 quadruplex and **I**. Remarkably, the computed spectra of $2 \times G4-K^+$ and $2 \times G3-K^+$ have a similar shape but the intensity of the former is $\sim 45\%$ higher, in agreement with the observed intensity change experimentally observed. We also verified that the introduction of the deoxyribose backbone and the other bases do not qualitatively change the electronic circular dichroism (ECD) spectra (see section 10.2 in the Supporting Information for further discussion). To prove the existence of G-triads in the species **I**, we looked for the presence of Hoogsteen hydrogen bonds by recording 1H NMR spectra over the 15–40 °C temperature range (Figure 3D and Figure S3). The observed signals in the 10.4–11.5 ppm range are consistent with the ones expected for imino protons characteristic of the presence of G-tetrads and/or G-triads.^[12,14] Interestingly, the imino signals at 40 °C are more resolved than at 15 °C, thus

suggesting a reduction of the conformational heterogeneity of RG1 on increasing the temperature, in agreement with the calculated molar fractions (Figure 2D). Notably, the persistence of signals in the 10.4–11.5 ppm range and the appearance of a new signal at 10.4 ppm on increasing the temperature are fully consistent with a G-quadruplex–G-triplex transition with the latter becoming the prevalent conformation at 40 °C.^[14,19] To gain additional insights into the guanines involved in the G-triplex formation, we characterized the truncated sequences of RG1, 3-RG1 and 5-RG1, obtained by removing two consecutive guanines from the 3' or 5' end respectively. Both truncated sequences form, under our experimental conditions, a monomolecular structure with an electrophoretic mobility comparable to that of RG1 (see Figure S4).

To note that these truncations cannot form an intramolecular G-quadruplex but could potentially form an intramolecular G-triplex. Remarkably, the spectrum of 3-RG1 is almost superimposable to the spectrum of the intermediate species of RG1 unfolding, whereas the 5-RG1 spectrum has a lower intensity and an additional negative peak at about 292 nm (Figure 4A). Moreover, the 3-RG1 shows a cooperative two-state unfolding with increasing temperature whereas 5-RG1 does not, suggesting that 3-RG1 forms a well folded compact structure (Figure 4B). Further, the 3-RG1 unfolding enthalpy is remarkably close to the enthalpy changes for the RG1 intermediate species unfolding (Figure S5). Overall, these results suggest that the main conformation of the 3-RG1 sequence is very similar to the one of the intermediate state in the RG1 unfolding. Note that 3-RG1 is a very short oligonucleotide (13 nucleotides long) and it is difficult to predict intramolecular structures other than a G-triplex capable of accounting for

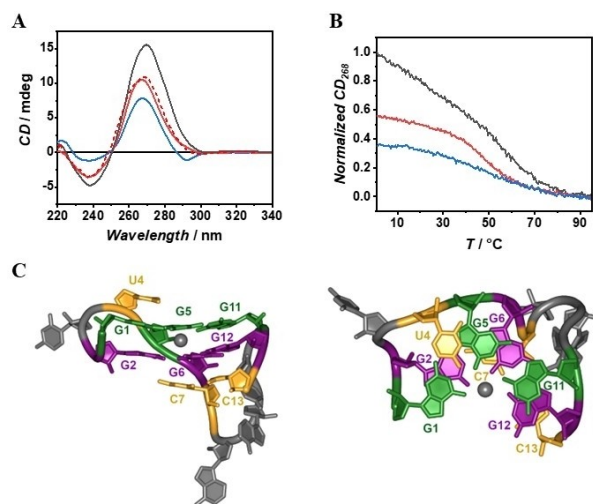


Figure 4. CD spectra of the RG1 quadruplex (black curve), 3-RG1 (red curve), experimental RG1 intermediate (red dashed curve) and 5-RG1 (blue curve) (A). Normalized CD melting curves of RG1 (black curve), 3-RG1 (red curve) and 5-RG1 (blue curve) (B). Orthogonal views of 3-RG1 G-triplex structure obtained by MD analysis. The K^+ ion is depicted as a grey sphere. For clarity, only sugar and base rings are shown as sticks (C).

its CD spectrum and the observed unfolding enthalpy changes. The conformational behavior of both 3-RG1 and 5-RG1 was further explored by means of molecular dynamics (MD). The MD simulations confirm that 5-RG1 is significantly less structured than 3-RG1, which adopts a compact structure where some Hoogsteen base pairings are maintained between adjacent strands, including a significant population (>40%) of G-triplex (Figure 4C, see also section 10.3 in the Supporting Information for further details). Further, the predicted ECD spectra for the most populated cluster extracted from the MD simulations are indeed consistent with the experimental CD spectrum of 3-RG1 (Figure S10 and S11). Overall, both our experimental and computational results reveal that the intermediate species in RG1 unfolding is a G-triplex formed by the detachment of the two 3'-end guanines from the corresponding G-tetrads.

Noteworthy, the RG1 G-triplex is much more stable than the DNA two-triads G-triplex previously reported^[14] whose melting temperature is below the physiological temperature, suggesting that RNA G-triplexes may be intrinsically more stable and biological relevant than the DNA G-triplex. Further, the observed capability of RG1 to adopt two alternative non-canonical nucleic acids structures can be the basis of a more complex regulatory switch as previously proposed for DNA.^[28] In summary, our data unambiguously demonstrate that under physiological conditions the dominant conformation of RG1 is different from the parallel quadruplex previously assumed but it forms an RNA G-triplex containing two-triads. Such previously unreported RNA structure could represent a new antiviral target. Since G-rich regions are highly abundant in human mRNAs, telomeric RNA and in viral RNA genomes, our study paves the way for further investigations into the presence of RNA G-triplex in vivo and their biological role.

Supporting Information

The authors have cited additional references within the Supporting Information.^[29–56]

Acknowledgements

This work was supported by PNRR: CN00000041 National Center for Gene Therapy and Drugs based on RNA Technology and the European Union-NextGeneration EU. We acknowledge the CINECA award under the ISCRA initiative for the availability of high-performance computing re-sources and support (Project HP10 C6CJ49). R.W. acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2033—Projektnummer 390677874. Open Access publishing facilitated by Università degli Studi di Napoli Federico II, as part of the Wiley - CRUI-CARE agreement.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Biophysics · G-quadruplex · G-triplex · Molecular Modelling · SARS-CoV-2

- [1] J. Jana, S. Mohr, Y. M. Vianney, K. Weisz, *RSC Chem. Biol.* **2021**, *2*, 338–353.
- [2] P. Kharel, P. Ivanov, *Trends Cell Biol.* **2024**, S0962892424000059.
- [3] B.-X. Zheng, J. Yu, W. Long, K. H. Chan, A. S.-L. Leung, W.-L. Wong, *Chem. Commun.* **2023**, *59*, 1415–1433.
- [4] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.
- [5] S. Ciaco, R. Aronne, M. Fiabane, M. Mori, *ACS Omega* **2024**, *9*, 24163–24180.
- [6] N. Kosiol, S. Juranek, P. Brossart, A. Heine, K. Paeschke, *Mol. Cancer* **2021**, *20*, 40.
- [7] L. Petraccone, G. Barone, C. Giancola, *Curr. Med. Chem. Anti-Cancer Agents* **2005**, *5*, 463–475.
- [8] S. Galli, G. Flint, L. Růžičková, M. Di Antonio, *RSC Chem. Biol.* **2024**, *5*, 426–438.
- [9] F. Wu, K. Niu, Y. Cui, C. Li, M. Lyu, Y. Ren, Y. Chen, H. Deng, L. Huang, S. Zheng, L. Liu, J. Wang, Q. Song, H. Xiang, Q. Feng, *Commun. Biol.* **2021**, *4*, 98.
- [10] A. Bedrat, L. Lacroix, J.-L. Mergny, *Nucleic Acids Res.* **2016**, *44*, 1746–1759.
- [11] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat. Biotechnol.* **2015**, *33*, 877–881.
- [12] L. Cerofolini, J. Amato, A. Giachetti, V. Limongelli, E. Novellino, M. Parrinello, M. Fragai, A. Randazzo, C. Luchinat, *Nucleic Acids Res.* **2014**, *42*, 13393–13404.
- [13] H.-X. Jiang, Y. Cui, T. Zhao, H.-W. Fu, D. Koirala, J. A. Punnoose, D.-M. Kong, H. Mao, *Sci. Rep.* **2015**, *5*, 9255.
- [14] V. Limongelli, S. De Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, M. Parrinello, *Angew. Chem. Int. Ed.* **2013**, *52*, 2269–2273.
- [15] D. Koirala, T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, H. Sugiyama, *Chem. Commun.* **2012**, *48*, 2006.
- [16] H. Zhou, Z.-F. Wu, Q.-J. Han, H.-M. Zhong, J.-B. Peng, X. Li, X.-L. Fan, *Anal. Chem.* **2018**, *90*, 3220–3226.
- [17] L.-L. Zhao, T. Cao, Q.-Y. Zhou, X.-H. Zhang, Y.-L. Zhou, L. Yang, X.-X. Zhang, *Anal. Chem.* **2019**, *91*, 10731–10737.
- [18] M. Bončina, J. Lah, I. Prislán, G. Vesnaver, *J. Am. Chem. Soc.* **2012**, *134*, 9657–9663.
- [19] R. D. Gray, R. Buscaglia, J. B. Chaires, *J. Am. Chem. Soc.* **2012**, *134*, 16834–16844.
- [20] A. Rajendran, M. Endo, K. Hidaka, H. Sugiyama, *Angew. Chem. Int. Ed.* **2014**, *53*, 4107–4112.
- [21] C. Zhao, G. Qin, J. Niu, Z. Wang, C. Wang, J. Ren, X. Qu, *Angew. Chem. Int. Ed.* **2021**, *60*, 432–438.
- [22] T. Miclot, C. Hognon, E. Bignon, A. Terenzi, M. Marazzi, G. Barone, A. Monari, *J. Phys. Chem. Lett.* **2021**, *12*, 10277–10283.
- [23] A. I. Karsiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed.* **2011**, *50*, 10645–10648.

- [24] R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, J. Feigon, *Proc. Nat. Acad. Sci.* **1993**, *90*, 3745–3749.
- [25] R. D. Gray, J. B. Chaires, *Curr. Protoc. Nucleic Acid Chem.* **2011**, *45*, DOI 10.1002/0471142700.nc1704s45.
- [26] M. A. Mullen, S. M. Assmann, P. C. Bevilacqua, *J. Am. Chem. Soc.* **2012**, *134*, 812–815.
- [27] S. Matsumoto, H. Tateishi-Karimata, S. Takahashi, T. Ohyama, N. Sugimoto, *Biochemistry* **2020**, *59*, 2640–2649.
- [28] I. M. A. del Mundo, M. Zewail-Foote, S. M. Kerwin, K. M. Vasquez, *Nucleic Acids Res.* **2017**, *45*, 4929–4943.
- [29] R. W. Hendler, R. I. Shrager, *J. Biochem. Biophys. Methods* **1994**, *28*, 1–33.
- [30] Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.* **2008**, *41*, 157–167.
- [31] L. Martínez-Fernández, F. Santoro, R. Improta, *Acc. Chem. Res.* **2022**, *55*, 2077–2087.
- [32] R. Improta, F. Santoro, L. Blancafort, *Chem. Rev.* **2016**, *116*, 3540–3593.
- [33] L. Martínez-Fernández, L. Esposito, R. Improta, *Photochem. Photobiol. Sci.* **2020**, *19*, 436–444.
- [34] H. Asha, J. A. Green, L. Esposito, F. Santoro, R. Improta, *Chirality* **2023**, *35*, 298–310.
- [35] H. Asha, J. A. Green, L. Martinez-Fernandez, L. Esposito, R. Improta, *Molecules* **2021**, *26*, 4789.
- [36] J. A. Green, H. Asha, F. Santoro, R. Improta, *J. Chem. Theory Comput.* **2021**, *17*, 405–415.
- [37] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian 16 Revision C.01, **2016**, Gaussian Inc., Wallingford, CT, USA.
- [38] M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. I. Cheatham, P. Jurečka, *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.
- [39] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999–3094.
- [40] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* **2003**, *24*, 669–681.
- [41] G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, *132*, 114110.
- [42] V. Karunakaran, K. Kleineremanns, R. Improta, S. A. Kovalenko, *J. Am. Chem. Soc.* **2009**, *131*, 5839–5850.
- [43] F. J. Avila Ferrer, J. Cerezo, E. Stendardo, R. Improta, F. Santoro, *J. Chem. Theory Comput.* **2013**, *9*, 2072–2082.
- [44] D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, G. A. Cisneros, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O’Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao, P. A. Kollman, Amber, **2022**, University of California, San Francisco, CA, USA.
- [45] H. Asha, J. A. Green, L. Esposito, L. Martinez-Fernandez, F. Santoro, R. Improta, *J. Phys. Chem. B* **2022**, *126*, 10608–10621.
- [46] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21*, 1049–1074.
- [47] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Loughton, M. Orozco, *Biophys. J.* **2007**, *92*, 3817–3829.
- [48] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [49] I. S. Joung, T. E. I. Cheatham, *J. Phys. Chem. B* **2008**, *112*, 9020–9041.
- [50] J. Åqvist, P. Wennerström, M. Nervall, S. Bjelic, B. O. Brandsdal, *Chem. Phys. Lett.* **2004**, *384*, 288–294.
- [51] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- [52] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.
- [53] D. R. Roe, T. E. I. Cheatham, *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
- [54] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33–38.
- [55] Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8.2.0, **2015**, New York, NY, USA.
- [56] R. Improta, *Chem. Eur. J.* **2014**, *20*, 8106–8115.

Manuscript received: August 13, 2024

Accepted manuscript online: October 4, 2024

Version of record online: November 6, 2024