

# Modulation of the Conformational Space of SARS-CoV-2 RNA Quadruplex RG-1 by Cellular Components and the Amyloidogenic Peptides $\alpha$ -Synuclein and hIAPP

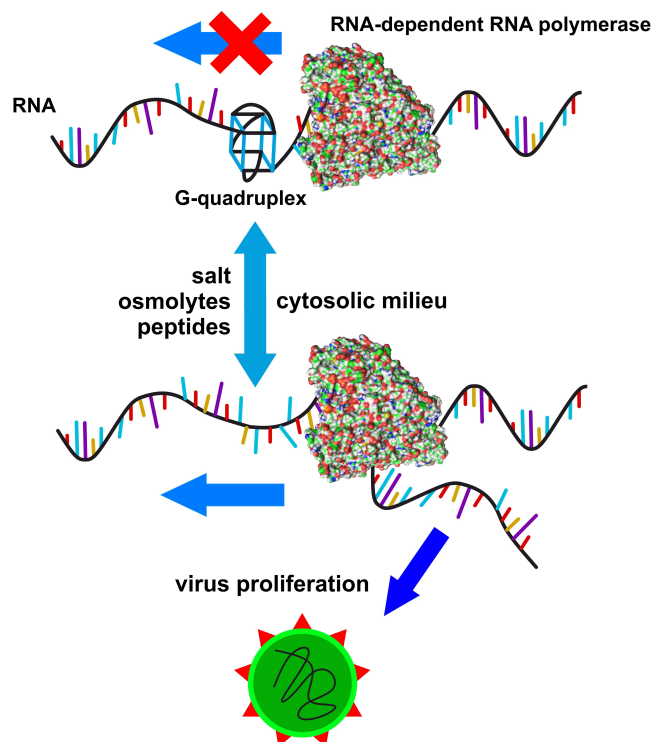
Sanjib K. Mukherjee<sup>+</sup>,<sup>[a]</sup> Jim-Marcel Knop<sup>+</sup>,<sup>[a]</sup> and Roland Winter<sup>\*[a]</sup>

**Abstract:** Given the emergence of the severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2), which particularly threatens older people with comorbidities such as diabetes mellitus and dementia, understanding the relationship between Covid-19 and other diseases is an important factor for treatment. Possible targets for medical intervention include G-quadruplexes (G4Qs) and their protein interaction partners. We investigated the stability and conformational space of the RG-1 RNA-G-quadruplex of the SARS-CoV-2 N-gene in the presence of salts, cosolutes, crowders and intrinsically disordered peptides, focusing on  $\alpha$ -Synuclein and the human islet amyloid polypeptide, which are involved in Parkinson's disease (PD) and type-II diabetes mellitus (T2DM), respec-

tively. We found that the conformational dynamics of the RG-1 G4Q is strongly affected by the various solution conditions. Further, the amyloidogenic peptides were found to strongly modulate the conformational equilibrium of the RG-1. Considerable changes are observed with respect to their interaction with human telomeric G4Qs, which adopt different topologies. These results may therefore shed more light on the relationship between PD as well as T2DM and the SARS-CoV-2 disease and their molecular underpinnings. Since dysregulation of G4Q formation by rationally designed targeting compounds affects the control of cellular processes, this study should contribute to the development of specific ligands for intervention.

## Introduction

The COVID-19 pandemic is caused by an enveloped RNA coronavirus called severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2). Next to the respiratory tract, the COVID-19 virus affects several other organs.<sup>[1–3]</sup> Currently, researchers are following three major approaches to develop therapeutics for fighting COVID-19: vaccines, neutralizing antibodies, and antiviral drugs, the main focus being on viral proteases and their inhibitors.<sup>[4–8]</sup> Recently, non-canonical nucleic acids structures, such as G-quadruplexes (G4Qs), have also been recognized as promising therapeutic targets.<sup>[7–11]</sup> The structure of G4Qs is characterized by the stacking of two or more planar arrangements of G-tetrads which are stabilized by lateral Hoogsteen-type hydrogen bonds and by the coordination of a monovalent cation, such as  $K^+$  (Figure 1).<sup>[11]</sup> Depending on the orientation of the G-tetrads, the structure of G4Qs can be parallel (with four G-



**Figure 1.** G-quadruplex formation can decrease transcription and translation of viral RNA and therefore reduce viral proliferation. The formation and stability of the G-quadruplex is highly dependent on its chemical environment, that is, the components of the cellular milieu, and on the presence of intrinsically disordered peptides.

[a] Dr. S. K. Mukherjee,<sup>+</sup> Dr. J.-M. Knop,<sup>+</sup> Prof. Dr. R. Winter  
Physical Chemistry I - Biophysical Chemistry  
Department of Chemistry and Chemical Biology  
TU Dortmund University  
Otto-Hahn Street 4a, 44227  
Dortmund (Germany)  
E-mail: roland.winter@tu-dortmund.de

[<sup>+</sup>] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/chem.202104182>

© 2021 The Authors. Chemistry - A European Journal 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.

tetrad in the same orientation), antiparallel (with two G-tetrad in opposite orientation with respect to the other two), or hybrid (with one G-tetrad in opposite orientation with respect to the other three) and further distinguished by their loop position (for example, antiparallel basket or chair conformation as well as different types of hybrid structures).<sup>[11]</sup> DNA/RNA polymerases coordinate their action with enzymes that unwind G4Qs, known as G4-helicases. G4Qs have therefore received considerable attention over the last twenty years due to their involvement in the regulation of cellular processes including replication, transcription and translation.<sup>[10–15]</sup> Dysregulation of G4Q formation and their binding proteins, which assist them in regulating the equilibrium between their structured and unstructured/unfolded forms, due to mutations or through the alteration of their stability by environmental factors (for example, by changes in intracellular solution conditions or by G4Q-stabilization induced by a ligand), have been found to contribute to many human pathologies, including neurodegenerative diseases, cancer, and microbial infections.<sup>[4–7]</sup>

In addition to eukaryotes and prokaryotes, G4Q motifs are also considered key elements in regulating the life cycle of viruses.<sup>[4–9]</sup> Their G4Qs repress the expression of viral proteins, some of which play immunomodulatory roles by limiting antigen presentation to cytotoxic T cells, thereby allowing the virus to survive in infected cells without being recognized by the host immune system. Hence, G4Q-specific compounds could be potential candidates for antiviral agents. Recently, it has been shown that G4Q-forming sequences in SARS-CoV-2 can form G4Q structures in living cells.<sup>[7,9]</sup> Hence, they would be novel targets for G4Q-specific compounds that could exert antiviral activity. The G4Q-forming sequence RG-1 is located in the coding sequence region of the SARS-CoV-2 nucleocapsid phosphoprotein (N-protein). In fact, the formation of RG-1 G4Q has been shown to decrease the amount of N-protein by inhibiting its translation. The RG-1 quadruplex was found to consist of two G-tetrad layers and to have a parallel G4Q structure, according to CD spectroscopy data.<sup>[7]</sup> Given the central role of the N-protein in controlling virus assembly and replication, RG-1 G4 may be a promising therapeutic target for SARS-CoV-2.<sup>[7]</sup> A prerequisite for the development of G4Q-related antiviral drugs against COVID-19 is a detailed knowledge of the structure and stability of the G4Q.

It is well known that the structure and thus the function of RNAs such as RNA-G4Qs are very susceptible to changes in their environment, that is, they respond sensitively to the presence of salts, osmolytes, crowding agents, and to the binding of proteins, which are also abundant in the biological cell.<sup>[16–23]</sup> In general, the stability of quadruplex structures is determined by H-bonds between nucleobases,  $\pi$ -stacking interactions between neighboring base pairs, counterion condensation at the phosphate backbone by cations, hydration changes, and by changes in conformational entropy.<sup>[11]</sup> Recently, it has been observed that the stability of the G4Qs depends also on the number of G-quartets. For example, the polyethylene glycol polymeric crowding agent PEG200 was found to stabilize RNA-

G4Qs with three or four G-quartets but did not change the stability of G4Qs with two G-quartets.<sup>[11]</sup>

In this study, we explored the effects of central environmental factors, including salts, osmolytes, crowding agents and intrinsically disordered peptides (IDPs), on the conformational landscape of the RG-1 G4Q of SARS-CoV-2. As IDPs we selected  $\alpha$ -Synuclein ( $\alpha$ -Syn) and the human islet amyloid polypeptide (hIAPP).  $\alpha$ -Syn regulates neurotransmitter vesicle cycling, but is, under pathological conditions, closely associated with Parkinson's disease (PD). PD results from abnormal aggregation of  $\alpha$ -Syn, and these aggregates are found primarily in Lewy bodies, the hallmark of PD.<sup>[24–26]</sup> Interestingly, SARS-CoV-2 has also been found in neurons in different brain regions.<sup>[26–29]</sup> Recent reports of cases of Parkinson's disease in relatively young patients after a SARS-CoV-2 infection suggest that there may be a link between SARS-CoV-2-infections and the development of PD.<sup>[29–31]</sup> In fact, multiple indications of a relation between amyloidoses and viral infections have been reported in the past.<sup>[31]</sup> The human  $\alpha$ -Syn protein has 140 amino acid residues and consists of three distinct regions, which include an amphipathic N-terminal domain, a central hydrophobic region (the non-A $\beta$  component (NAC) region), and a highly negatively charged proline-rich C-terminal domain. The conformation of the polypeptide is strongly affected by cosolutes and lipid membranes, and  $\alpha$ -Syn has also a high propensity to interact with DNAs.<sup>[32–41]</sup> hIAPP (amylin) is a 37 amino acid hormone that is associated with the progression of type II diabetes mellitus (T2DM).<sup>[42–44]</sup> The peptide hormone misfolds to form amyloid deposits in and around the pancreatic islet  $\beta$ -cells that synthesize both insulin and hIAPP, leading to a decrease in  $\beta$ -cell mass in patients with the disease. Recent findings suggest that there is also a strong correlation between T2DM and SARS-CoV-2 diseases.<sup>[45–47]</sup> Therefore, studying the conformation of the SARS-CoV-2 RNA-G4Qs in the presence of  $\alpha$ -Syn and hIAPP is of immense importance, as interaction with these disordered peptides can significantly affect replication and transcription along with causing genetic instability in host cells.

## Results and Discussion

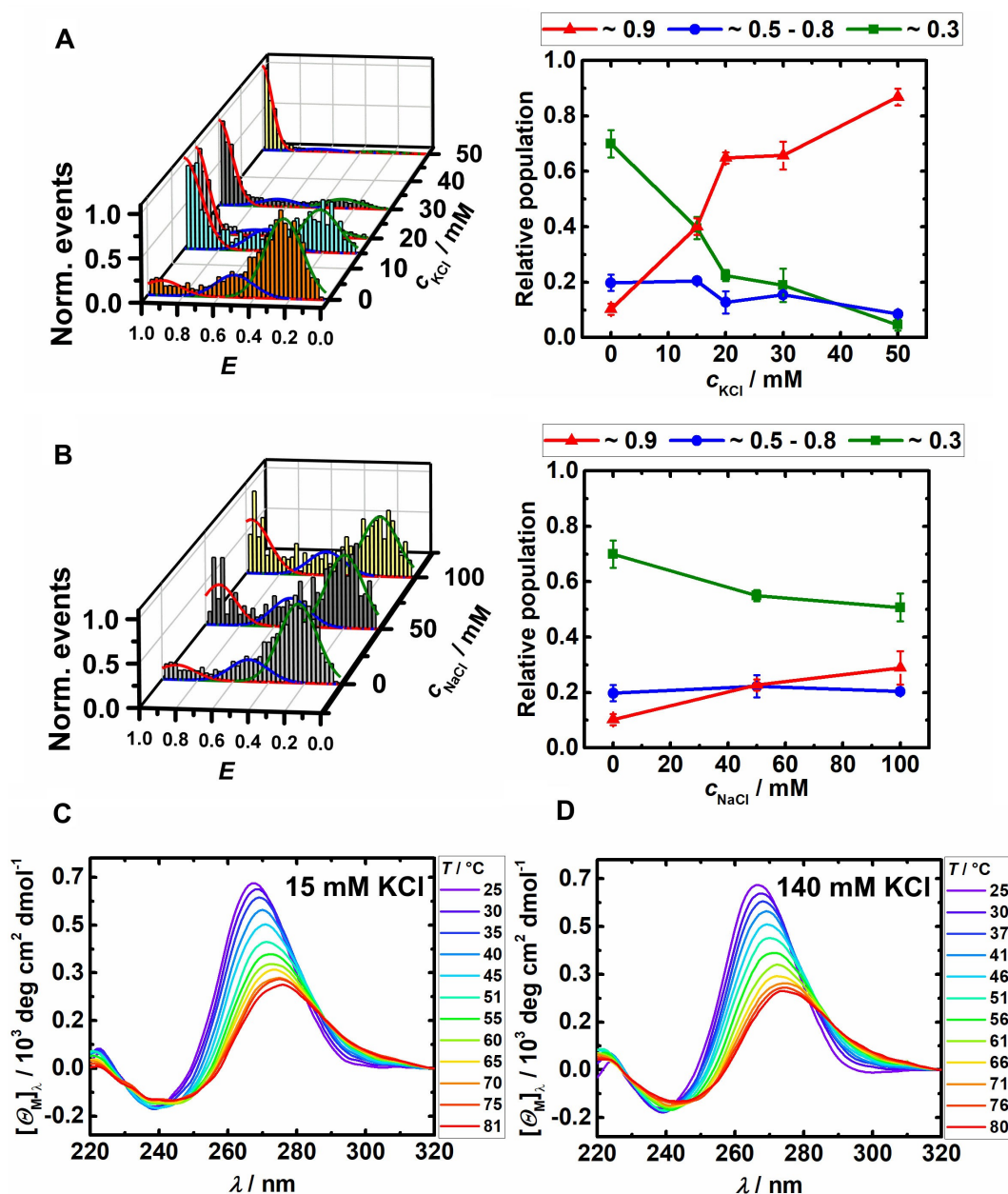
### Effect of salts, crowding and cosolvents on the conformation of the SARS-CoV-2 RG-1-RNA quadruplex

Since the main objective of this study was to investigate the effects of various environmental factors on the conformational landscape of the RG-1 RNA quadruplex, we performed conformation-sensitive single-molecule Förster resonance energy transfer (smFRET) experiments. This method avoids ensemble averaging and enabled us to elucidate the structure and conformational dynamics of the quadruplex and how the equilibrium between conformational substates is affected by the different solution conditions.<sup>[48,49]</sup> The peaks in the FRET efficiency histograms refer to conformations with different spatial separations,  $R$ , of the two attached dyes and thus different FRET efficiencies,  $E$ , because  $E = R_0^6 \cdot (R_0^6 + R^6)^{-1}$ .<sup>[21,48,49]</sup> The Förster radius,  $R_0$ , is the distance at which 50% of the

excited donor molecules will be deactivated ( $R_0=6.5$  nm for the fluorophores used, Atto 550 and Atto 647N). As it is well known that monovalent cations play an important role in the stability of the G4Qs, smFRET measurements were carried out in the presence of monovalent ( $K^+$ ,  $Na^+$ ) salts to quantify their effect on the conformational transitions of the RG-1 SARS-CoV-2 RNA. Of note, in diseased cells such as those of cancer and neurodegenerative diseases, the concentration of intracellular ions is typically changed by overexpression (or inactivation) of disease specific ion channel proteins. We observed that the stability of the RG-1 SARS-CoV-2 RNA-G4Q is highly susceptible to the concentration of monovalent  $K^+$ . In the absence of  $K^+$  cations

(pure buffer solution), two peaks were observed in the FRET histograms, located at  $E \approx 0.3$ , representing the unfolded conformation) and at  $E \approx 0.5-0.8$ , representing a metastable or partially folded conformation, respectively (Figure 2A). At and beyond 15 mM KCl, a further conformation appeared, located at  $E \approx 0.9$  in the histogram, which represents the folded conformation of the RNA G4Q. Conformational switching from open to closed states was observed between 0 and 30 mM KCl. At high  $K^+$  concentrations (140 mM KCl), only the fully folded conformation is retained (Figure 2A).

Complementary CD measurements at 15 mM KCl showed a positive band around 267 nm and a negative peak close to



**Figure 2.** FRET efficiency and relative population of conformational states of the elongated and labeled RG-1 (~ 100 pM) RNA in 20 mM TrisHCl-buffer at pH 7.5 and 25 °C at different A) KCl and B) NaCl concentrations. Temperature dependent CD-spectra of the unlabeled sequence (20  $\mu\text{M}$ ) in 20 mM  $\text{Na}_2\text{H}_2\text{P}_2\text{O}_7$ -buffer, pH 7.5, with C) 15 mM and D) 140 mM KCl.

240 nm, with no significant changes at higher salt concentrations (140 mM KCl), which can be attributed to a parallel topology of the G4Q (Figure 2C and D) and is similar to the CD spectrum of the core structure of the G4Q.<sup>[7]</sup> Unfolding of the RG-1 RNA-G4Q at high temperature as observed in the CD measurements was also observed by corresponding smFRET measurements of the elongated RG-1 G4Q (Figure S2). Interestingly, the Na<sup>+</sup> cation was found to be unable to stabilize the folded conformation of the RG-1 RNA-G4Q (Figure 2B), even at 100 mM concentration. Conversely, the human telomeric (htel) DNA G-quadruplex adopts a folded antiparallel topology in Na<sup>+</sup> salt solution.<sup>[49]</sup>

Various factors contribute to the stability of G-quadruplexes and determine their folding topology, including  $\pi$ -stacking interactions, H-bonding, cation binding, and hydration changes.<sup>[11,50,51]</sup> Since the stacking interactions partly account for the net energetic gain, increasing the number of quartets is energetically favorable and provides more stability.<sup>[51]</sup> Along with that, stacking of other non-tetrad bases can also contribute to the overall stabilization of the G4Q structure.<sup>[51]</sup> Recent studies revealed that RG-1 G4Q exhibits a parallel quadruplex topology which is composed of two superimposed tetrads and a flexible peripheral loop composed of adenines and cytosines, serving as a linker to the guanines involved in tetrad formation.<sup>[7,52]</sup> Cation coordination is another essential factor in the stabilization of G4Qs.<sup>[53]</sup> Cations bind to the negatively charged phosphates groups of the G4Qs both through non-specific interactions and through specific (site-bound) coordination to the O-6 lone-pair electrons of the guanines of the tetrads after loss of their hydration sphere. Binding of the cations to the negatively charged phosphate backbone reduces the electrostatic repulsions and thus provides stability to the folded conformation.<sup>[54,55]</sup> As the ionic radius of K<sup>+</sup> is larger (1.33 Å) compared to Na<sup>+</sup> (0.95 Å), we may assume that Na<sup>+</sup> should be small enough to be coordinated within the plane of a G-quartet. Nevertheless, the free energy of binding to the RG-1 RNA quadruplex seems to be much lower for K<sup>+</sup>, promoting the formation of a stable parallel-folded G-quadruplex beyond 15 mM KCl. For comparison, the htel telomeric DNA quadruplex adopts a hybrid-stranded conformation in the presence of K<sup>+</sup> and an antiparallel basket structure in the presence of Na<sup>+</sup>.<sup>[47,55]</sup> The htel DNA G4Q consists of three superimposed tetrads. Therefore, we can assume that the binding of Na<sup>+</sup> by an additional stacking interaction across the third tetrad provides sufficient stability, which is not possible in the case of the RG-1 G4Q.

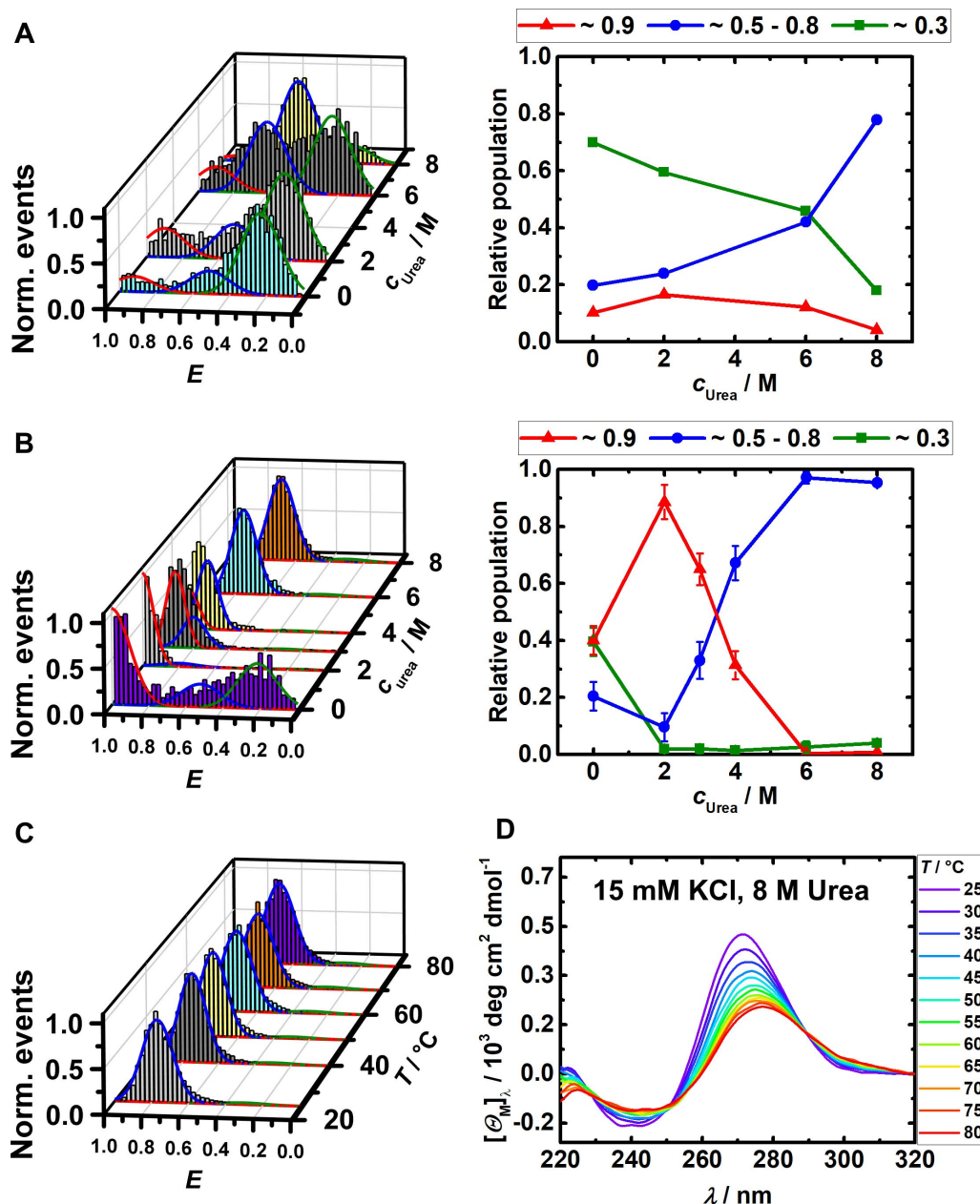
The intracellular environment is crowded with biomolecules that occupy a significant portion (up to 30%–40%) of the cellular volume, which is expected to affect the free energy and conformational landscape of biomacromolecules.<sup>[21–23,56–58]</sup> We used 30 wt% Ficoll, a polysaccharide of about 5 nm size, to mimic cellular macromolecular crowding conditions. Figure S3 shows that crowding stabilizes the folded conformation of the RG-1 G4Q. This effect can be understood invoking an entropy-driven excluded volume effect. Due to the reduction of the solvent accessible surface area in the folded state, the

equilibrium shifts toward the folded quadruplex structure upon addition of the crowding agent.

As a representative of a chaotropic cellular cosolute that generally interacts nonspecifically and unfavorably with proteins and nucleic acids (NAs), urea is a good choice. Urea is a natural organic cosolute with implications in maintaining osmotic pressure in cells and is present in rather large concentrations in various living organisms.<sup>[59]</sup> We observed that urea strongly modulates the equilibrium between different RG-1 G4Q conformations. Unexpectedly, in buffer with 15 mM KCl, urea concentrations beyond 2 M shifts the equilibrium towards the folded state (Figure 3B). At concentrations of 3–4 M urea, co-existence of various conformations was observed in the smFRET histograms. We can tentatively assign the peak at  $E \approx 0.9$  in the  $E$ -histograms to a parallel folded conformation, the peak at  $E \approx 0.5$ – $0.8$  to a partially folded conformation, and that at  $E \approx 0.3$  to the unfolded state structure. At high urea concentrations (6–8 M), a single conformation is mainly seen in the FRET histogram only, which can be attributed to a partially folded state. The partially folded conformer is steadily stabilized upon increasing the urea concentrations, reaching population values around 80%. No drastic changes are seen in the CD spectrum of the partially folded state compared to the folded structure. For example, the CD spectrum of RG-1 G4Q in 15 mM K<sup>+</sup> and 8 M urea at room temperature showed a positive CD band around 270 nm, which is little red-shifted compared to the pure buffer solution (15 mM K<sup>+</sup>, Figure 2C), along with a reduced amplitude and a negative peak close to 240 nm, implying that RG-1 adopts a partially unfolded topology in accordance with the smFRET measurements that showed a single  $E$ -distribution peak located at  $E \approx 0.8$  (Figure 3C).

An increase in temperature lead to a further conformational change of RG 1 G4Q as manifested in a redshift and decrease in the amplitude of the positive CD peak, indicating unfolding. The combined CD and smFRET data at ambient temperature suggest that high urea concentrations cause the RG-1 quadruplex topology to transform into a structure in which only the overhang conformation is altered, while the core of the RG-1 G4Q structure remains essentially unchanged. Heat-induced denaturation, as visible in the CD-spectra at high urea concentration (Figure 3D), results in a small shift of the FRET efficiency only, from  $E \approx 0.8$  to  $E \approx 0.7$  (Figure 3C), indicating that unfolding of the core structure at temperatures above  $\sim 50^\circ\text{C}$  keeps the dyes still in close proximity.

It is generally expected that urea leads to favorable interactions with the rings and functional groups of nucleic acid bases when becoming exposed to the solvent in the unfolded state.<sup>[59,60]</sup> However, the mechanism that urea exerts on the stability of G4Qs is still not well understood. Moore and coworkers showed that the structure of a four-stranded RNA-quadruplex remains stable even at 8 M urea concentration.<sup>[61]</sup> A pronounced accumulation of urea molecules at short distances around a DNA G-quadruplex was seen in a study of Smiatek et al.<sup>[62]</sup> Such accumulation of urea around the G4Q surface implies that water molecules are mainly replaced by urea. We can therefore assume that in the RG-1 G4Q the number of hydrogen bonds between the G4Q and urea predominates over

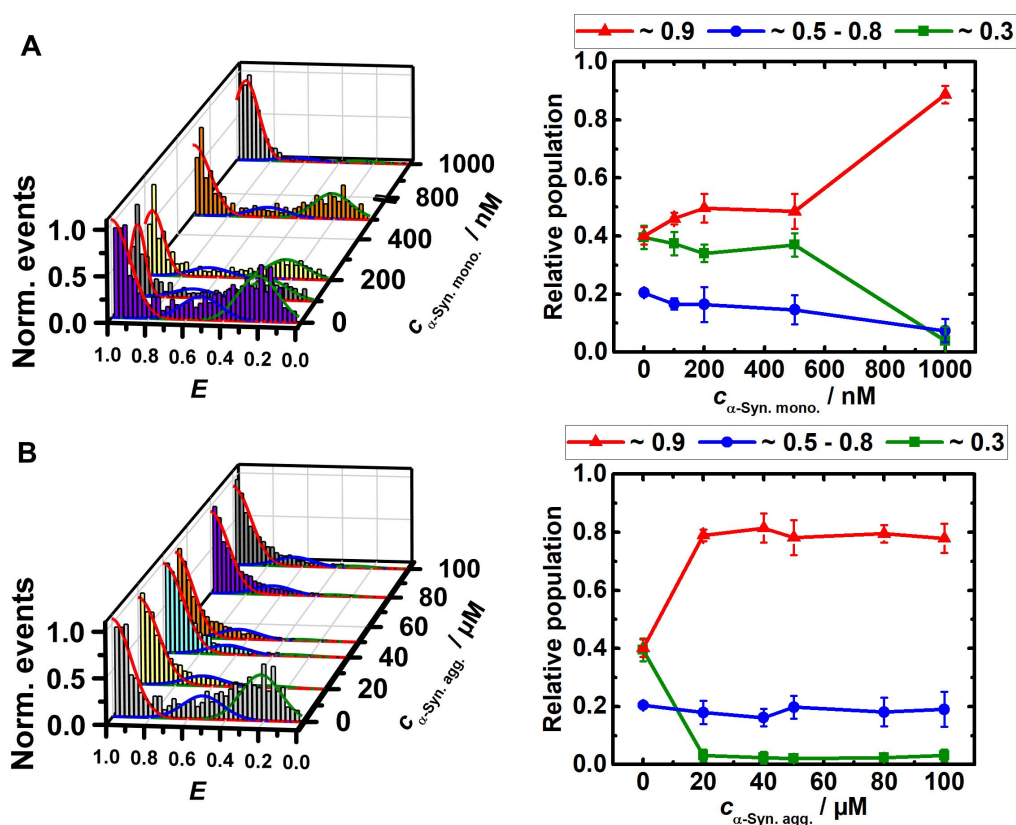


**Figure 3.** A and B) FRET efficiency and relative conformational population of the elongated and labeled RG-1 RNA (~100 pM) in 20 mM TrisHCl buffer at pH 7.5 and 25 °C with increasing urea concentration. Samples A) contained no salt, samples B) contained 15 mM KCl. C) shows the temperature dependent FRET efficiency histograms with 15 mM KCl and 8 M urea. D) depicts the corresponding temperature dependent CD spectra (20 μM RNA).

RNA-water interactions at high concentrations of urea, but this does not lead to a significant destabilization of the G4Q structure, but only stabilizes a partially folded state, most likely through interactions with the loop structure. The unfolding of the core structure of the G4Q is observed at high temperatures. However, the unfolded state is more structured in the presence of urea than the heat-induced state in pure buffer.

#### Effect of disordered amyloidogenic peptides on the conformation of RG-1 RNA G4Q

Intrinsically disordered proteins that bind to nucleic acids with high affinity are able to modulate the conformational dynamics of nucleic acids in a cellular environment and may hence have significant biological effects on gene expression and regulation patterns. We observed that the conformation of the SARS-CoV 2 RG-1 RNA G4Q is very susceptible to the presence of the amyloidogenic disordered peptides  $\alpha$ -Syn and hIAPP. As clearly seen in Figure 4, monomeric  $\alpha$ -Syn facilitates folding of the RG-1 G4Q in the presence of 15 mM KCl already at sub-μM



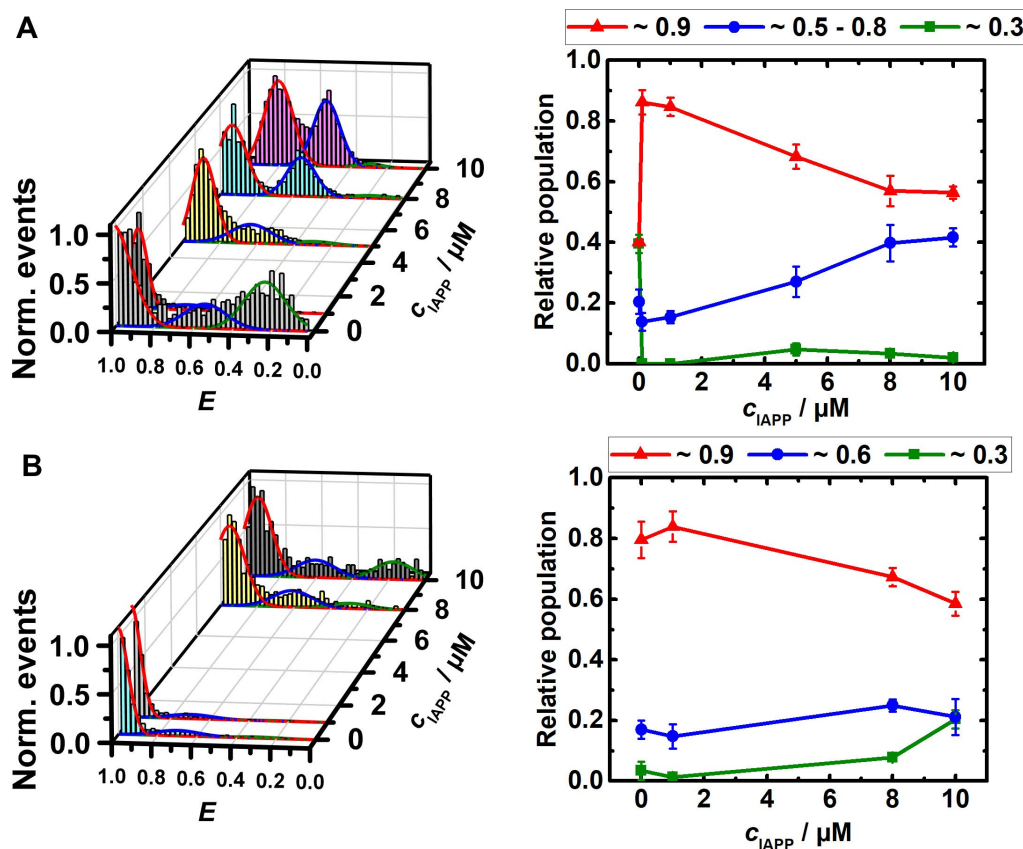
**Figure 4.** FRET efficiency and relative conformational population of the elongated and labeled RG-1 RNA (~100 pM) in 15 mM KCl, 20 mM TrisHCl buffer at pH 7.5 and 25 °C with increasing concentration of A)  $\alpha$ -Syn monomers and B)  $\alpha$ -Syn aggregates.

concentrations, probably by acting as a macromolecular counterion that effectively reduces charge repulsion of the phosphate backbone of the nucleic acid. In the absence of  $\alpha$ -Syn, the RG-1 G4Q adopts folded and unfolded conformations to a similar extent (~40%), along with smaller amounts of partially unfolded conformations (~20%, Figure 4A). With increasing concentration of monomeric  $\alpha$ -Syn (100 nM to 1  $\mu$ M), the amount of folded RG-1 G4Q conformers increases to ~90%. Remarkably, the conformational switching occurs already in the nanomolar concentration range. Increasing the  $\alpha$ -Syn concentration from 1  $\mu$ M up to 100  $\mu$ M has no further effect (Figure S4). In the presence of aggregated  $\alpha$ -Syn, a similar stabilization of the folded RG-1 G4Q structure was observed. A completely different scenario was found in the case of human telomeric DNA G4Q.<sup>[49]</sup> The conformation of the htel G4Q was largely unaffected by monomeric  $\alpha$ -Syn even at very high concentrations (200  $\mu$ M). The antiparallel conformation of the htel G4Q gradually shifted towards a more unfolded conformation in the presence of oligomeric species of  $\alpha$ -Syn, but not in the presence of monomeric  $\alpha$ -Syn.

Owing to the correlation found between type II diabetes and the SARS-CoV-2 disease, we may expect an effect of hIAPP on the replication and translation of the virus RNA.<sup>[45–47]</sup> Hence we studied also the interaction between hIAPP and the RG-1 G4Q from the SARS-CoV-2 N-gene. The smFRET data demonstrate that the conformation of RG-1 G4Q is also strongly

affected by hIAPP. Figure 5(A) shows that the equilibrium between unfolded and folded conformers of RG-1 G4Q is strongly shifted to the folded state already at low hIAPP concentrations (0.1–1  $\mu$ M). Moreover, a partially folded conformation (located at  $E \approx 0.5$ –0.6) is populated at higher hIAPP concentrations (5–10  $\mu$ M). For comparison, the conformation of htel DNA-G4Q was found to be less affected by the hIAPP (Figure 5B). Upon addition of 8  $\mu$ M hIAPP, a small change from the antiparallel conformation ( $E \approx 0.9$ ) to the hybrid topology ( $E \approx 0.6$ ) was observed. In the presence of 10  $\mu$ M hIAPP, also a small amount of unfolded DNA ( $E \approx 0.3$ ) was detected.

Protein-RNA interactions are a vital mechanism inside the cell to perform RNA synthesis, translation and degradation.<sup>[63]</sup> In general, protein interactions with nucleic acids are based on a series of molecular contacts, including hydrogen bonding (direct or mediated by water molecules), ionic/dipolar interactions,  $\pi$ -interactions, van der Waals and hydrophobic interactions. Although some amino acid residues interact specifically with NAs,<sup>[64]</sup> the diversity of structural motifs in proteins and NAs makes it difficult to comprehend the precise mechanism of protein-NA interactions at the molecular level.<sup>[65–69]</sup> Reynaldo and coworkers showed that RNA binding occurs either through an  $\alpha$ -helix or loop of the protein or some amino acids involved in a  $\beta$ -sheet structure which potentially interact with unpaired RNA bases.<sup>[67]</sup> NMR measurements showed that a 12-mer RNA aptamer forms a quadruplex structure capable of binding prion



**Figure 5.** FRET efficiency and relative conformational population of the elongated and labeled RG-1 RNA ( $\sim 100$  pM) in 15 mM KCl, 20 mM TrisHCl buffer at pH 7.5 and 25 °C with increasing concentration of A) hIAPP and, for comparison, of the htel G4Q used in a previous study<sup>[49]</sup> at the same solution conditions with increasing concentration of B) hIAPP.

protein (PrP) peptides from the PrP's N-terminal region and blocks the structural conversion of PrP through both specific and nonspecific interactions.<sup>[67]</sup>

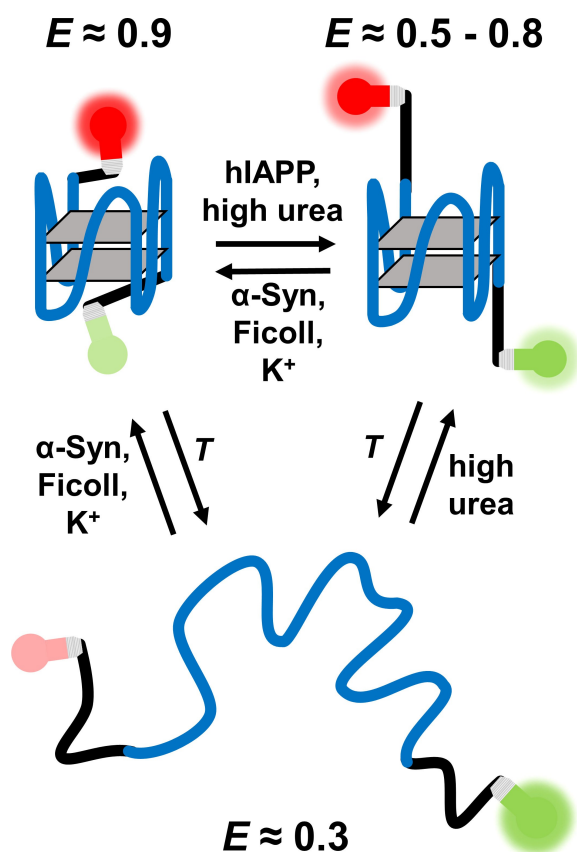
$\alpha$ -Syn is a natively disordered protein that is positively charged (+4) at the N-terminus and contains mainly lysine residues that can interact with the negatively charged phosphate groups of the nucleic acid, along with other amino acid residues that nonspecifically interact with the nucleobases of the RNA-G4Q.  $\alpha$ -Syn has 17 glutamate residues along with 3 asparagines and 6 glutamines. Earlier reports suggested that lysine and arginine strongly favor interaction with guanine bases and largely account for the abundance of hydrogen bonding interactions with the bases.<sup>[65]</sup> On the contrary, glutamine and asparagines are able to form a large number of hydrogen bonds with adenine bases, and glutamate is responsible for a maximum number of hydrogen bonds with cytosine. In addition, glutamate strongly interacts with cytosine through van der Waals interactions. Since RG-1 G4Q has a flexible peripheral loop composed of adenines and cytosines, such interactions would help stabilize the folded state of the RG-1 G4Q in the presence of  $\alpha$ -Syn. In the case of hIAPP, only 6 asparagines and 1 glutamine residue are present and able to form hydrogen bonds with adenine, and no additional stability is gained by other interactions. This might explain stabilization

of a partially folded conformation at high concentrations of hIAPP, only.

## Conclusion

In recent years, G-quadruplexes have received considerable attention as therapeutic targets for chemical intervention in biological functions. Since modulation of the stability and conformation of G4Qs has been found to play a significant role in many human pathological and neurodegenerative diseases, the underlying mechanisms need to be explored to be able to target G4Qs. We have shown that the stability and conformational landscape of the RNA G4Q-forming sequence RG-1, which is located in the coding sequence region of the SARS-CoV-2 nucleocapsid phosphoprotein (N), is broadly affected by salt, osmolyte, crowding and intrinsically disordered peptides.

Figure 6 gives an overview of the results obtained. Further, we have seen that significant differences exist between the stability and topology of the RG-1 RNA-G4Q and the human telomeric G4Q, providing novel insights into the mechanisms dictating the conformational preference of the different quadruplex structures.  $\alpha$ -Synuclein, the crowder Ficoll and  $\text{K}^+$  cations stimulate formation of the fully folded structure, while hIAPP as



**Figure 6.** Schematic representation of the different effects salt cations, urea, crowder and the IDPs  $\alpha$ -Syn and hIAPP impose on the structure of the RG-1 RNA G4Q. The blue line represents the G4Q forming NA, while the black lines represent the additional nucleotides. The gray planes are mimicking the G-tetrads and the bulbs represent the fluorophores Atto 550 (green) and Atto 647N (red) used in the smFRET study.

well as high concentrations of urea induce a partially folded conformation of the RG-1 RNA-G4Q. Only high temperature leads to full unfolding of RG-1. The interaction with salts, cosolutes, crowders and peptides are factors that affect the conformational dynamics and stability of the G4Qs and helped us explore the overall conformational landscape of the different quadruplexes accessible. Since dysregulation of G4Q formation by rationally designed targeting compounds is detrimental in regulating cellular processes, this knowledge should be helpful in designing specific ligands for intervention. Next to the number of tetrads that determines the stability of the G4Qs, the loop sequences are unique and not only provide a structural distinction responsible for the diverse topologies observed, but may also have the ability to specifically interact with designer ligands.

To date, the link between SARS-CoV-2 and the metabolic and pathological diseases observed is not well understood. The formation of RG-1 G4Q has been shown to reduce levels of N-protein by inhibiting its translation, hence might be a promising therapeutic target to fight SARS-CoV-2. We have seen that both amyloidogenic peptides,  $\alpha$ -Syn and hIAPP, affect the conformational equilibrium between folded and unfolded states, to a

different extent, however, in a sequence-specific way. Considerable changes are observed with respect to their interaction with htel G4Qs. hIAPP stabilized a partially folded state of the RG-1 RNA-G4Q, whereas  $\alpha$ -Syn strongly stabilized the folded state of the quadruplex. Since G-quadruplexes and their interaction with regulatory proteins play an important role as on/off-switches that modulate polymerase activity, the interaction with these amyloidogenic peptides may alter expression profiles of disease-modifying genes. Therefore, these results may shed more light on the relationship between PD and T2DM and the SARS-CoV-2 disease and its molecular basis.

## Experimental Section

**Materials:** Ficoll 70 was purchased from Carl Roth GmbH + Co. KG (Karlsruhe, Germany) and was used as received without further purification. Human IAPP (islet amyloid polypeptide or amylin) was purchased from GenScript (Leiden, Netherland). Before use, the IAPP peptide was dissolved in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) and incubated for at least 30 min to ensure an aggregate- and seed-free sample. The buffer solution used in the measurements contains 20 mM Tris-HCl (pH 7.4) and was filtered by a 0.45  $\mu$ m sterile Whatman Puradisc 30 syringe filter. The RNA and DNA sequences with and without fluorophore labels were purchased from biomers (Ulm, Germany). For the exact sequences, please refer to the Supporting Information. The expression and purification of  $\alpha$ -synuclein was carried out as described before<sup>[70,71]</sup> and is also described in the Supporting Information.

**Single-molecule FRET (smFRET) measurements and data analysis:** SmFRET measurements were carried out using a confocal fluorescence microscope (MicroTime 200, PicoQuant) under freely diffusing conditions. The pulsed interleaved excitation (PIE) FRET technique was used to separate dually labeled from singly labeled species.<sup>[72]</sup> Briefly, in the PIE FRET technique, both the donor and acceptor are alternatively excited by the laser pulse. First, a laser pulse of suitable wavelength excites the donor and then another laser pulse excites the acceptor independently from FRET after a certain time delay (50 ns, 20 MHz repetition rate), allowing us to calculate the photon stoichiometry,  $S$ , which is the ratio of photons emitted after donor excitation and the sum of total photons emitted after donor and direct acceptor excitation. For a donor-only species,  $S=1$ , and for an acceptor-species only,  $S=0$ . A green laser pulse at 560 nm (LDH series, PicoQuant) and a red laser pulse at 635 nm (LDH series, PicoQuant) were used to excite the donor Atto 550 and acceptor Atto 647N, respectively. A quad band dichroic mirror (ZT 405/488/561/640, Chroma) was used to reflect both the green and red laser light to the entrance port of the fluorescence microscope. Donor and acceptor fluorescence signals were separated to two different detection channels, first by using a dichroic mirror (FF 650 Di01, Semrock), followed by band pass filters FF 01-593/40 (Semrock) and FF 01-676/29 (Semrock). Two SPCM-AQR series single photon avalanche diodes (SPAD) were used as detection channels for the donor and acceptor fluorescence. Time-correlated single photon counting (TCSPC) data were first analyzed with Multiparameter Fluorescence Detection (MFD)<sup>[73]</sup> script from Symphotime 5.3.2.2 software. The resulting histograms were normalized for the highest number of events and fitted with a three Gaussian function using Matlab R2016a with a least squares fitting method. For determining the relative populations of conformers, the percentage of the integrals of the single Gaussian functions were calculated and plotted.



**CD measurements:** The topology of G-quadruplexes can be determined by monitoring the positive and/or negative circular dichroism (CD) signals at specific wavelengths. The CD measurements were performed using a JASCO J-715 spectropolarimeter in 20 mM  $\text{Na}_2\text{H}_2\text{PO}_4$  buffer at pH 7.5 in the presence of 15 mM or 140 mM KCl solution. The concentration of the RNA construct was 20  $\mu\text{M}$ . The molar ellipticity was calculated using the strand concentration. The pathlength of the quartz-glass cell was 1 mm.

## Acknowledgements

This project received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801459 – FP-RESOMUS and was funded by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy – EXC 2033 – 390677874 – RESOLV. Open Access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Keywords:**  $\alpha$ -synuclein · hIAPP · RNA G-quadruplex · SARS-CoV-2 · single-molecule FRET

- [1] Y. Zhang, X. Geng, Y. Tan, Q. Li, C. Xu, J. Xu, L. Hao, Z. Zeng, X. Luo, F. Liu, H. Wang, *Biomed. Pharmacother.* **2020**, *127*, 110195.
- [2] U. M. Ashraf, A. A. Abokor, J. M. Edwards, E. W. Waigi, R. S. Royfman, S. A.-M. Hasan, K. B. Smedlund, A. M. G. Hardy, R. Chakravarti, L. G. Koch, *Physiol. Genomics* **2021**, *53*, 51–60.
- [3] Z. Song, L. Bao, P. Yu, F. Qi, S. Gong, J. Wang, B. Zhao, M. Liu, Y. Han, W. Deng, J. Liu, Q. Wei, J. Xue, W. Zhao, C. Qin, *Front. Microbiol.* **2021**, *11*, 1–15.
- [4] S. Sreeramulu, et al. *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 19191–19200.
- [5] H. Su, F. Zhou, Z. Huang, X. Ma, K. Natarajan, M. Zhang, Y. Huang, H. Su, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 9789–9802.
- [6] M. Sun, et al. *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 10266–10272.
- [7] C. Zhao, G. Qin, J. Niu, Z. Wang, C. Wang, J. Ren, X. Qu, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 432–438.
- [8] G. Bezzi, E. J. Piga, A. Binolfi, P. Armas, *Int. J. Mol. Sci.* **2021**, *22*, 2614.
- [9] E. Belmonte-Reche, I. Serrano-Chacón, C. Gonzalez, J. Gallo, M. Bañobre-López, *PLoS One* **2021**, *16*, e0250654.
- [10] S. Takahashi, A. Kotar, H. Tateishi-Karimata, S. Bhowmik, Z.-F. Wang, T.-C. Chang, S. Sato, S. Takenaka, J. N. Sugimoto, *J. Am. Chem. Soc.* **2021**, *143*, 16458–16469.
- [11] S. Takahashi, N. Sugimoto, *Acc. Chem. Res.* **2021**, *54*, 2110–2120.
- [12] D. Varshney, J. Spiegel, K. Zyner, D. Tannahill, S. Balasubramanian, *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 459–474.
- [13] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.
- [14] J. Mitteaux, et al. *J. Am. Chem. Soc.* **2021**, *143*, 12567–12577.
- [15] L. Dumas, P. Herviou, E. Dassi, A. Cammas, S. Millevoi, *Trends Biochem. Sci.* **2021**, *46*, 270–283.
- [16] S. Matsumoto, H. Tateishi-Karimata, S. Takahashi, T. Ohyama, N. Sugimoto, *Biochemistry* **2020**, *59*, 2640–2649.
- [17] L. Petraccone, B. Pagano, C. Giancola, *Methods* **2012**, *57*, 76–83.
- [18] D. Kilburn, J. H. Roh, L. Guo, R. M. Briber, S. A. Woodson, *J. Am. Chem. Soc.* **2010**, *132*, 8690–8696.
- [19] A. C. Hautke, S. Ebbinghaus, *ChemSystemsChem* **2020**, *2*, e2000052.
- [20] J. C. Miner, A. E. Angel, *J. Phys. Chem. B* **2017**, *221*, 3734–3746.
- [21] S. Patra, V. Schuabb, I. Kiesel, J. M. Knop, R. Oliva, R. Winter, *Nucleic Acids Res.* **2019**, *47*, 981–996.
- [22] M. Gao, D. Gnutt, A. Orban, B. Appel, F. Righetti, R. Winter, F. Narberhaus, S. Müller, S. Ebbinghaus, *Angew. Chem. Int. Ed.* **2016**, *55*, 3224–3228; *Angew. Chem.* **2016**, *128*, 3279–3283.
- [23] M. Gao, C. Held, S. Patra, L. Arns, G. Sadowski, R. Winter, *ChemPhysChem* **2017**, *18*, 2951–2972.
- [24] R. Jakes, M. G. Spillantini, M. Goedert, *FEBS Lett.* **1994**, *345*, 27–32.
- [25] D. W. Dickson, *Curr. Opin. Neurol.* **2001**, *14*, 423–432.
- [26] B. C. V. Campbell, Q. X. Li, J. G. Culvenor, P. Jäkälä, R. Cappai, K. Beyreuther, C. L. Masters, C. A. McLean, *Neurobiol. Dis.* **2000**, *7*, 192–200.
- [27] E. Song, et al. *J. Exp. Med.* **2021**, *218*, e20202135.
- [28] A. Ramani, et al. *EMBO J.* **2020**, *39*, e106230.
- [29] M. Merello, K. P. Bhatia, J. A. Obeso, *Lancet Neurol.* **2021**, *20*, 94–95.
- [30] K. Ezzat, et al. *Nat. Commun.* **2019**, *10*, 1–16.
- [31] S. A. Semerdzhiev, M. A. A. Fakhree, I. Segers-Nolten, C. Blum, M. M. A. E. Claessens, *BioRxiv* reprint <https://doi.org/10.1101/2021.04.12.439549>.
- [32] C. Galvagnion, D. Topgaard, K. Makasewicz, A. K. Buell, S. Linse, E. Sparr, C. M. Dobson, *J. Phys. Chem. Lett.* **2019**, *10*, 7872–7877.
- [33] J. K. Hannestad, S. Rocha, B. Agnarsson, V. P. Zhdanov, P. Wittung-Stafshede, F. Höök, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 14178–14186.
- [34] C. R. Bodner, C. M. Dobson, A. Bax, *J. Mol. Biol.* **2009**, *390*, 775–790.
- [35] I. Jahan, S. M. Nayeem, *Biophys. J.* **2019**, *117*, 1922–1934.
- [36] A. J. Schaser, et al. *Sci. Rep.* **2019**, *9*, 10919.
- [37] D. Cherny, W. Hoyer, V. Subramaniam, T. M. Jovin, *J. Mol. Biol.* **2004**, *344*, 929–938.
- [38] M. L. Hegde, P. Vasudevaraju, K. J. Rao, *Front. Biosci.-Landmark* **2010**, *15*, 418–436.
- [39] A. Louka, E. Zacco, P. A. Temussi, G. G. Tartaglia, A. Pastore, *Nucleic Acids Res.* **2020**, *48*, 11880–11889.
- [40] B. Macedo, T. A. Millen, C. A. Braga, M. P. Gomes, P. S. Ferreira, J. Kraineva, R. Winter, J. L. Silva, Y. Cordeiro, *Biochemistry* **2012**, *51*, 5402–5413.
- [41] K. Jiang, S. Rocha, A. Westling, S. Kesari-mangalam, K. D. Dorfman, P. Wittung-Stafshede, F. Westerlund, *Chem. Eur. J.* **2018**, *24*, 15685–15690.
- [42] M. C. Owen, D. Gnutt, M. Gao, S. K. T. S. Wärmländer, J. Jarvet, A. Gräslund, R. Winter, S. Ebbinghaus, B. Strodel, *Chem. Soc. Rev.* **2019**, *48*, 3946–3996.
- [43] J. Seeliger, F. Evers, C. Jeworrek, S. Kapoor, K. Weise, E. Andreetto, M. Tolan, A. Kapurniotu, R. Winter, *Angew. Chem. Int. Ed.* **2012**, *51*, 679–683; *Angew. Chem.* **2012**, *124*, 703–707.
- [44] F. Evers, C. Jeworrek, S. Tiemeyer, K. Weise, D. Sellin, M. Paulus, B. Struth, M. Tolan, R. Winter, *J. Am. Chem. Soc.* **2009**, *131*, 9516–9521.
- [45] S. Corrao, K. Pinelli, M. Vacca, M. Raspanti, C. Argano, *Front. Endocrinol.* **2021**, *12*, 609470.
- [46] E. Varghese, S. M. Samuel, A. Liskova, P. Kubatka, D. Büsselberg, *PLoS Pathog.* **2021**, *17*, e1009634.
- [47] H. Y. Fan, Y. L. Shek, A. Amiri, D. N. Dubins, H. Heerklotz, R. B. MacGregor, T. V. Chalikian, *J. Am. Chem. Soc.* **2011**, *133*, 4518–4526.
- [48] J. M. Knop, S. Patra, B. Harish, C. A. Royer, R. Winter, *Chem. Eur. J.* **2018**, *24*, 14346–14351.
- [49] J. M. Knop, S. K. Mukherjee, R. Oliva, S. Möbitz, R. Winter, *J. Am. Chem. Soc.* **2020**, *142*, 18299–18303.
- [50] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 4301–4309.
- [51] S. Amrane, R. W. L. Ang, Z. M. Tan, C. Li, J. K. C. Lim, J. M. W. Lim, K. W. Lim, A. T. Phan, *Nucleic Acids Res.* **2009**, *37*, 931–938.
- [52] T. Miclot, C. Hognon, E. Bignon, A. Terenzi, M. Marazzi, G. Barone, A. Monari, *J. Phys. Chem. Lett.* **2021**, *12*, 10277–10283.
- [53] J. R. Williamson, M. K. Raghuraman, T. R. Cech, *Cell* **1989**, *59*, 871–880.
- [54] J. C. Bowman, T. K. Lenz, N. V. Hud, L. D. Williams, *Curr. Opin. Struct. Biol.* **2012**, *22*, 262–272.
- [55] R. D. Gray, J. B. Chaires, *Biophys. Chem.* **2011**, *159*, 205–209.
- [56] S. Nakano, D. Miyoshi, N. Sugimoto, *Chem. Rev.* **2014**, *114*, 2733–2758.
- [57] S. Ghosh, S. Takahashi, T. Ohyama, T. Endoh, H. Tateishi-Karimata, N. Sugimoto, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 14194–14201.
- [58] H.-X. Zhou, G. N. Rivas, A. P. Minton, *Annu. Rev. Biophys.* **2008**, *37*, 375–397.
- [59] L. Aslanyan, J. Ko, B. G. Kim, I. Vardanyan, Y. B. Dalyan, T. V. Chalikian, *J. Phys. Chem. B* **2017**, *121*, 6511–6519.
- [60] I. V. Smirnov, R. H. Shafer, *Biopolymers* **2006**, *85*, 91–101.

- [61] J. Kim, C. Cheong, P. B. Moore, *Nature* **1991**, *351*, 331–332.
- [62] E. A. Oprzeska-Zingrebe, J. Smiatek, *Phys. Chem. Chem. Phys.* **2021**, *23*, 1254–1264.
- [63] S. Lim, J. H. Bae, H.-S. Kwon, M. A. Nauck, *Nat. Rev. Endocrinol.* **2021**, *17*, 11–30.
- [64] K. Nagai, *Curr. Opin. Struct. Biol.* **1996**, *6*, 53–61.
- [65] N. M. Luscombe, R. A. Laskowski, J. M. Thornton, *Nucleic Acids Res.* **2001**, *29*, 2860–2874.
- [66] R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, R. S. Mann, *Annu. Rev. Biochem.* **2010**, *79*, 233–69.
- [67] D. E. Draper, L. P. Reynaldo, *Nucleic Acids Res.* **1999**, *27*, 381–388.
- [68] T. Mashima, A. Matsugami, F. Nishikawa, S. Nishikawa, M. Katahira, *Nucleic Acids Res.* **2009**, *37*, 6249–6258.
- [69] Y. Cordeiro, B. Macedo, J. L. Silva, M. P. B. Gomes, *Biophys. Rev. Lett.* **2014**, *6*, 97–110.
- [70] M. J. Volles, P. T. Lansbury, *J. Mol. Biol.* **2007**, *366*, 1510–1522.
- [71] R. Shaltiel-Karyo, M. Frenkel-Pinter, N. Egoz-Matia, A. Frydman-Marom, D. E. Shalev, D. Segal, E. Gazit, *PLoS One* **2010**, *5*, e13863.
- [72] S. Ruettinger, V. Buschmann, B. Kraemer, S. Orthaus, F. Koberling, FRET analysis of Pulsed Interleaved Excitation (PIE) using the MicroTime 200. *Appl. note, PicoQuant GmbH*, **2013**, 1–8.
- [73] P. J. Rothwell, S. Berger, O. Kensch, S. Felekyan, M. Antonik, B. M. Wöhrle, T. Restle, R. S. Goody, C. A. M. Seidel, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1655–1660.

---

Manuscript received: November 21, 2021

Accepted manuscript online: December 9, 2021

Version of record online: January 5, 2022