

Review

# Cooperative Analysis of Structural Dynamics in RNA-Protein Complexes by Single-Molecule Förster Resonance Energy Transfer Spectroscopy

Nathalie Meiser <sup>†</sup>, Christin Fuks <sup>†</sup> and Martin Hengesbach <sup>\*†</sup>

Institute for Organic Chemistry and Chemical Biology, Goethe-University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany; meiser@nmr.uni-frankfurt.de (N.M.); fuks@nmr.uni-frankfurt.de (C.F.)

\* Correspondence: hengesbach@nmr.uni-frankfurt.de

† These authors contributed equally to this work.

Academic Editor: Quentin Vicens

Received: 25 February 2020; Accepted: 13 April 2020; Published: 28 April 2020



**Abstract:** RNA-protein complexes (RNPs) are essential components in a variety of cellular processes, and oftentimes exhibit complex structures and show mechanisms that are highly dynamic in conformation and structure. However, biochemical and structural biology approaches are mostly not able to fully elucidate the structurally and especially conformationally dynamic and heterogeneous nature of these RNPs, to which end single molecule Förster resonance energy transfer (smFRET) spectroscopy can be harnessed to fill this gap. Here we summarize the advantages of strategic smFRET studies to investigate RNP dynamics, complemented by structural and biochemical data. Focusing on recent smFRET studies of three essential biological systems, we demonstrate that investigation of RNPs on a single molecule level can answer important functional questions that remained elusive with structural or biochemical approaches alone: The complex structural rearrangements throughout the splicing cycle, unwinding dynamics of the G-quadruplex (G4) helicase RHAU, and aspects in telomere maintenance regulation and synthesis.

**Keywords:** RNA-protein complex (RNP); single-molecule Förster resonance energy transfer (smFRET) spectroscopy; RNP dynamics; spliceosome; G quadruplex helicase RHAU; telomerase

## 1. Introduction

Many cellular processes rely on the faithful interaction between proteins and RNA, which constitute most major molecular machineries. For some of these RNA-protein complexes (RNPs), structural data has been available for quite some time, the most prominent example being the ribosome [1]. However, the underlying molecular processes are highly complex and oftentimes exhibit an unexpected degree of conformational dynamics, rendering their analysis technically demanding. In addition, assessing molecular or conformational dynamic heterogeneity is as difficult as acquiring more exhaustive descriptions of such processes. For the latter, one current trend is to employ integrative structural biology approaches, combining analytical toolsets with structural and biochemical data to provide more detailed characterization of such processes. Such ensemble techniques are however limited when it comes to the description of heterogeneous conformations and dynamics; in these cases, a number of methods that allow for the dissection of such heterogeneous populations are required. Here, techniques that provide analytics on the level of individual molecules have proven to be of great value, as they allow the assessment of both heterogeneity and conformational dynamics at the same time. The most prominent method that has been used for quite a number of different biological systems is single molecule FRET spectroscopy.

FRET is based on the non-radiative energy transfer from a donor to an acceptor fluorophore that show a spectral overlap of emission and absorption spectra, respectively. This transfer is distance dependent to the sixth power and scales with the FRET pair specific Förster radius ( $R_0$ ), leading to high FRET efficiencies when the two fluorophores come into closer proximity. This allows applying FRET spectroscopy as a molecular ruler, covering distances ranging from approximately 2 to 10 nm. First FRET experiments were performed in an ensemble approach, therefore averaging over all molecules in potentially different conformations. However, applications in a single molecule setup allow observation of individual, non-synchronized molecules with different conformations and dynamics. Detection in single molecule resolution is achieved by applying total internal reflection fluorescence (TIRF) microscopy or confocal based setups. However, TIRF based approaches are the most widely used weapon of choice. This technique requires immobilization of the FRET-labeled constructs to the imaged surface. Independent of the imaging technique, strategic construct design is a crucial step to achieve significant results using smFRET. Labeling sites are chosen based on structural data (if available) corresponding to an expected high distance variance upon conformational changes preferably crossing the Förster radius. In studies on RNP dynamics, RNA or proteins can be targeted for fluorophore labeling [2–5]. The experimentally generated FRET trajectories then provide information on the conformational distribution in equilibrium, the time-resolved interconversion between different states, and can identify intermediates and folding pathways.

Within the last years, a wide array of approaches for various systems has been developed; ranging from ultrashort dynamics of unfolded proteins [6] over the analysis of individual RNAs to highly complex systems like helicases [7,8], the ribosome [9–11], argonaute proteins [12,13], or polymerases [5,14].

Here, we want to showcase selected recent examples on three RNPs that have been presented within the last years and immensely benefited from smFRET approaches: the spliceosome [15–18], a G-quadruplex resolving helicase [19–22], and telomere maintaining complexes [23–25]. All of these have in common that the insights from single molecule FRET experiments are complemented by structural and biochemical analysis, and we will emphasize how these approaches are mutually beneficial.

## 2. Spliceosome

Splicing is an essential process that allows eukaryotic cells to derive different protein coding mRNAs from the same genomic sequence. This is achieved by cutting and rejoining the pre-mRNA in a site-specific fashion. The RNP complex that catalyzes this highly regulated process is called spliceosome, which is in itself highly dynamic in its structure as well as its molecular composition; among the components are five unique spliceosomal RNAs (U1, U2, U4–6) and a varying number (~100) of proteins. The recent advances in structural biology have yielded the first models of the spliceosome core structure [26–28], aided by novel approaches that facilitated purification of spliceosomes at defined steps of the splicing cycle. In retrospect, the highly conformationally and compositionally dynamic nature of the splicing process was one of the major obstacles in many of the approaches that aimed at investigating the structure on the molecular level.

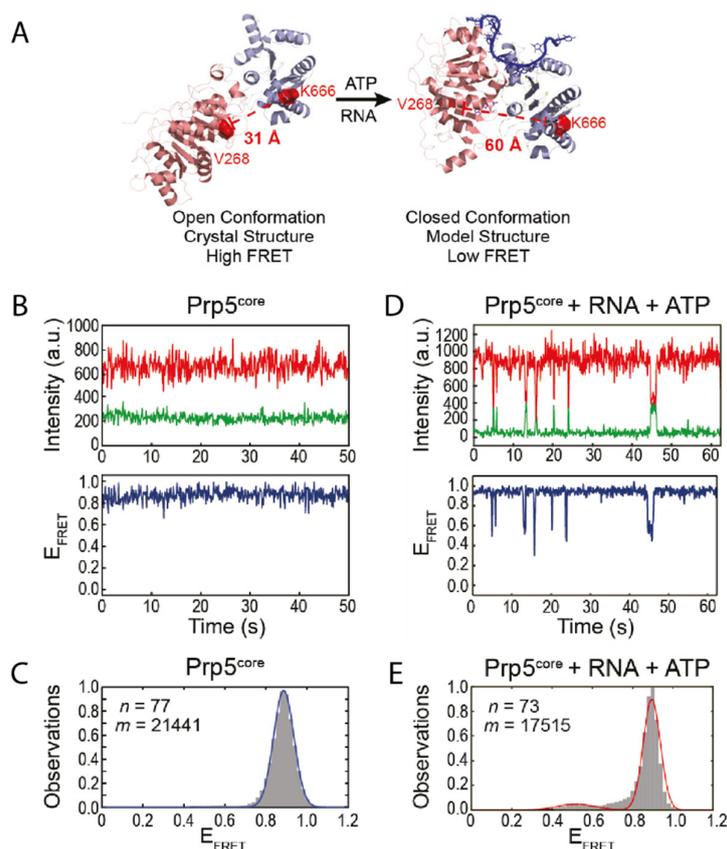
Several crucial steps can be defined that are required for efficient splicing: (1) Recognition of the 5' splice site (5'SS), which occurs by binding of the pre-mRNA to the U1 RNA (commitment complex 2), (2) Binding of U2 RNA to the branch site (A-complex), (3) Recruitment of the trimeric snRNP (tri-snRNP) into the complex (B-complex), (4) multiple conformational rearrangement and remodeling steps, during which U1 and U4 are removed from the complex (B<sup>act</sup>-complex), (5) formation of the catalytic core by U2 and U6, (6) transesterification and formation of a lariat structure (C-complex), (7) additional remodeling steps, and (8) second transesterification and release of the ligated exons.

While this list is neither exhaustive nor sufficiently detailed to fully describe the splicing process, it provides an impression of the complexity of the splicing cycle. It also illustrates the need for techniques that faithfully report on dynamics within each of these steps.

Recent work in several labs was focusing on different steps throughout the splicing cycle: During spliceosome assembly, the role of Prp5 on U2 and pre-mRNA was investigated [18,29]. For conformational rearrangements during the transesterification reaction, dynamics of U2 RNA and the effect of nucleotide modifications, magnesium ion binding, and the interaction with Cus2 were taken into focus [16,17]. To generate a more holistic picture of the splicing cycle, an unbiased classification method for structural dynamics was established [15].

### 3. The Role of Prp5 on Spliceosome Assembly

The DEAD-box ATPase Prp5 was found to be essential for initiating the splicing process [30]. It mediates recognition of the pre-mRNA branch site by the spliceosomal RNA U2 while undergoing a conformational switch [31]. This step provides proofreading functionality for the branch site recognition [32,33], while it is no longer needed in later steps of the splicing process [32]. The structure of Prp5 has been solved to 1.95 Å showing that it assumes a twisted open conformation (Figure 1A) [31]. To understand the mode of action of the protein it is inevitable to not only consider snapshots, but also to identify possible short-lived conformations and how they are influenced by certain factors. For Prp5 specifically this might be an additional, closed conformation as observed for other DEAD-box helicases [7,34].



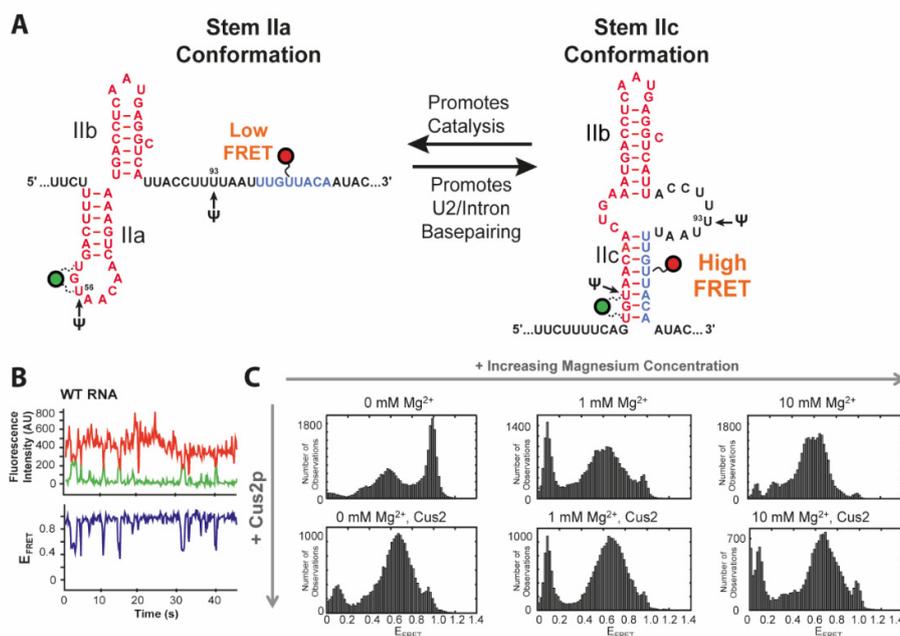
**Figure 1.** Conformational changes within Prp5. (A) Crystal structure of Prp5 in the twisted open conformation and modelled structure of closed conformation. The two RecA-like domains of Prp5 are shown in pink and purple and the RNA in the closed conformation in blue. The residues mutated for fluorophore introduction are indicated in red. (B) Exemplary Förster resonance energy transfer (FRET) trajectory of Prp5 alone and (C) the corresponding FRET histogram fitted with one peak (blue line). (D) FRET behavior of Prp5 in the presence of poly (A) RNA and ATP and (E) the corresponding FRET histogram fitted with two peaks (red line). This Figure was taken from Beier et al. [18] and is used under Creative Commons Non-Commercial License (CC-BY NC).

Based on the available structural information, Beier et al. designed a smFRET construct using cysteine mutations of Prp5 [18]. Within this construct, the twisted open conformation would lead to a small distance between the FRET pair, resulting in high FRET efficiency (Figure 1A). The assumed conformational change however would increase the distance, leading to a lower FRET efficiency in the closed conformation. In the absence of any cofactor or substrate, the protein assumes the open conformation as expected (Figure 1B,C); this is also in line with studies on other DEAD-box helicases [7,34]. Only when adding both ATP and substrate however, the construct underwent transient, short-lived transitions into the closed conformation (Figure 1D,E). The authors showed that their smFRET results can be coupled with other biochemical methods to gain an even more detailed understanding of the system. A biochemical ATPase assay confirmed that ATP hydrolysis was increased in presence of RNA. The conformational switch was also disfavored in presence of ADP, but the lifetime of the closed state was increased in presence of the nonhydrolyzable ADPNP. Mutations of Prp5 were introduced that either increased or decreased specificity of the branch site recognition. Conformational analysis of these mutants showed that there is a strong correlation between the structural dynamics of Prp5 mutants, i.e., the lifetime of their closed and open state, their ATP hydrolysis activity, and their impact on branch site recognition [18].

#### 4. U2 Toggling

The cold-sensitive U2 snRNA suppressor protein 2 (Cus2) assists Prp5 in prespliceosome activation and confers ATP dependency of the remodeling process [35]. It has been proposed that hydrolysis of ATP is required for the subsequent removal of Cus2 [32,35,36]. Throughout the splicing cycle, the U2 snRNA undergoes several conformational changes between the two transesterifications and rearrangements of the pre-mRNA substrate. This primarily affects the stem II region, which can adopt two conformations via folding of either hairpin IIa or IIc (Figure 2A) [37,38]. While IIc is responsible for both catalytic reactions, IIa is necessary for substrate binding and release [37,38]. The latter conformation is presumably bound and stabilized by Cus2 [39]. Splicing efficiency is also dependent on posttranscriptional modifications of U2, such as three conserved pseudouridines [40–42] and under distinct stress conditions additionally at positions 93 and 56 [43,44].

Rodgers et al. and van der Feltz et al. addressed this toggling of U2 between stems IIa and IIc using smFRET spectroscopy [16,17]. Discrimination between the two conformations was facilitated by the two states showing either a low FRET (stem IIa) or a high FRET (stem IIc) state (Figure 2C) as designed based on structural data. In the absence of  $Mg^{2+}$  analysis of their RNA construct showed coexistence of a high FRET state of 0.96 corresponding to IIc, a FRET efficiency of 0.54 for IIa and an additional, minor low FRET conformation of 0.35, all three dynamically interconverting into each other (Figure 2B,C). smFRET showed the titratable impact of  $Mg^{2+}$  or Cus2 as they both individually shifted the equilibrium in favor of the IIa conformation. Wild type transitions to the low FRET state vanished in the presence of Cus2. smFRET could not only reveal the influence of the environment on the ratio between subpopulations and their transition frequency. It can also be applied to evaluate the effects of nucleotide modifications as presented by van der Feltz et al. through introduction of pseudouridine at the stress inducible positions  $\Psi 56$  and  $\Psi 93$ . With pseudouridylation at both positions for instance, neither  $Mg^{2+}$  nor Cus2 impacted stem II conformation. Only a combination of  $Mg^{2+}$  with Cus2 was able to overcome this stabilization of the IIc conformation caused by the dual modification  $\Psi 56$  and  $\Psi 93$ .



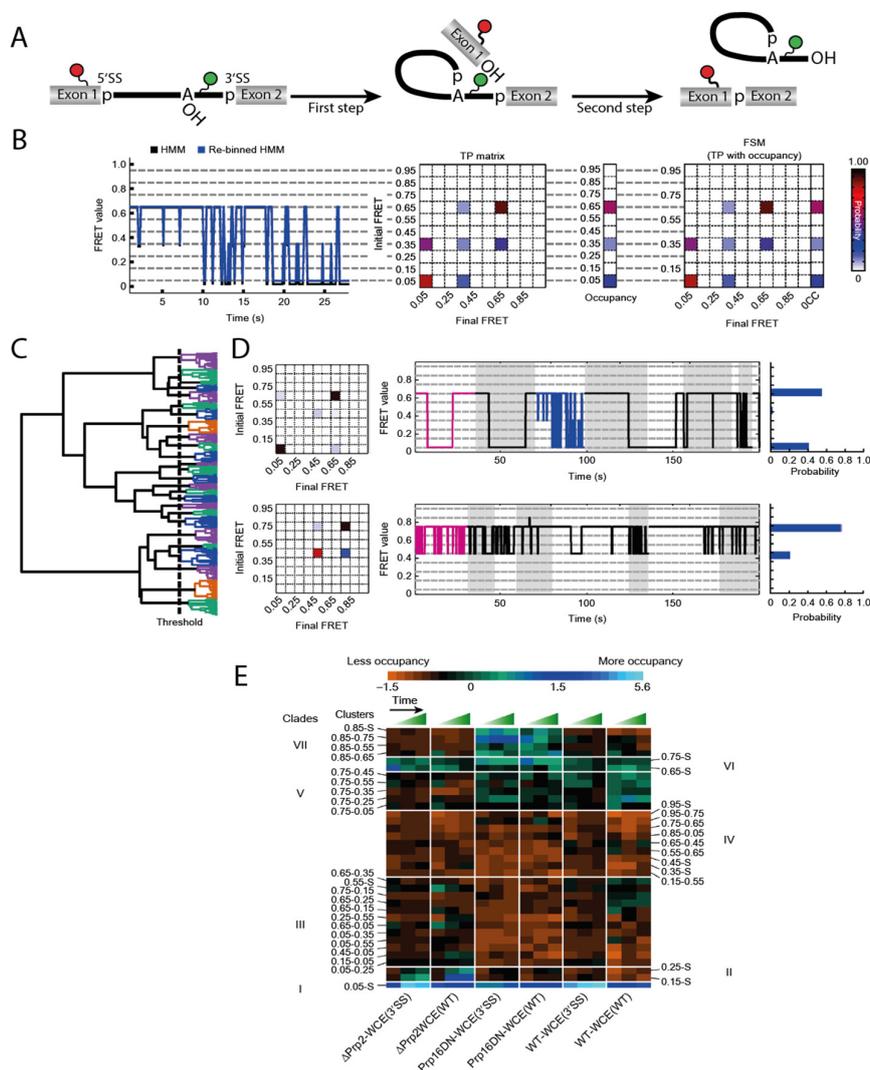
**Figure 2.** FRET construct and data for toggling of U2 snRNA. (A) The stem II of the U2 snRNA can switch between two conformations. Labeling positions in the two studies of Cy3 (green) and Cy5 (red) as well as Pseudouridylation sites are indicated. (B) Example FRET trajectory of U2 shows a spontaneous toggling between two FRET states. (C) FRET histograms of WT RNA in dependence of  $Mg^{2+}$  and Cus2 concentrations. This Figure was adapted and modified from Rodgers et al. [16] and used under Creative Commons Non-Commercial License (CC-BY NC).

## 5. Observation of the Full Splicing Cycle

The first three studies discussed in this review characterized the dynamic behavior of isolated steps of the splicing process. Throughout the splicing cycle, the spliceosome shows a drastic level of compositional as well as conformational complexity. A thorough description of at least aspects of the splicing cycle therefore requires tools for investigating the system in toto. The heterogeneous occupancy of states, the ability to study unsynchronized processes and their dynamics in real time render a single molecule approach the method of choice for this type of spliceosome analysis.

Blanco et al. described an algorithm to tackle this problem [15]. Their method Single Molecule Cluster Analysis (SiMCAn) is based on ‘hierarchical clustering of Hidden Markov modeling-fitted smFRET trajectories’ [15]. Each individual FRET trace is first fitted via Hidden Markov modeling (HMM) [45] and subsequently re-binned to one of ten evenly spaced FRET states from 0.05 to 0.95 (Figure 3B). Using the HMM traces, a transition probability (TP) matrix is generated for each molecule which is then additionally described with the occupancy to a FRET similarity matrix (FSM) (Figure 3B). After separating out the static molecules, the dynamic FSMs allow clustering of several thousand molecules according to their FRET behavior, their state occupancy, and transition kinetics. This clustering results in a hierarchical tree sorting the molecules according to similarity in their dynamic behavior (Figure 3C). Each individual cluster contains an average TP matrix, a collection of traces, and an occupancy probability distribution (Figure 3D). While this technique can already identify dynamics that cannot be found by manual sorting, it is necessary to connect the numerous clusters to their function in order to understand a mechanistically complex system such as the spliceosome. In order to obtain data that is sufficient to provide such an understanding, it is necessary to trap the process under investigation at predefined stages to accumulate certain intermediates. Those roadblocks could come in the form of deletions or mutations of components, or the use of specific inhibitors. By generating such intermediates, additional data on the existing clusters can be obtained. In a second round of clustering, the static and dynamic clusters combined are summarized in clades if they show

similar enrichment or depletion under the same condition (Figure 3E). This allows for the assignment of the FRET dynamics to the individual steps of the entire system by means of their clades.



**Figure 3.** Workflow of the Single Molecule Cluster Analysis (SiMCAn) approach on the full splicing cycle [15]. (A) pre-mRNA FRET construct used in this study. During the first splicing step a lariat structure is formed and after splicing the exons are ligated and the intron-lariat released. (B) Hidden Markov modeling (HMM) fitted exemplary trace (black) and re-binned trace (blue) with the resulting transition probability (TP) matrix and the FRET similarity matrix (FSM). (C) Clustering of the molecules resulting in a hierarchical tree with every leaf representing a single molecule and sorting into 25 clusters indicated by the same color. (D) Representation of 2 clusters with their TP matrix, exemplary traces with the best trace in magenta and the occupancy in each state. (E) Second round of clustering of the clusters into 7 clades which show similar behavior under changing conditions. © 2015 Nature America, Inc. Used with permission.

The spliceosome substrate, the pre-mRNA, is present at all stages of the splicing cycle and undergoes many conformational as well as chemical changes through the process. A construct that had previously been used for smFRET analysis of the first splicing step dynamics [29] was a pre-mRNA that was labeled at Exon 1 near the 5'SS and near the branchpoint, expecting high FRET when those sites come in close proximity (Figure 3A). The same construct was now been used with a nuclear extract conferring splicing activity in smFRET analysis. While these data are too complex for a manual evaluation because molecules occupy different steps of the splicing cycle, the authors emphasize the

advantages of SiMCAn and the amount of information this method provides. When analyzing the highly dynamic and complex data obtained from these experiments, the SiMCAn algorithm found ten static and 25 dynamic clusters, which are further clustered to seven distinct clades [15].

Applying different roadblocks throughout the splicing cycle and observing the (time-dependent) accumulation, the authors could characterize the clade deriving from the commitment complex 2, the A-complex, the B<sup>act</sup> complex and the post first step C-complex. SiMCAn uncovered that the splicing cycle is not strictly directed forwards. A reversible switching back to the commitment complex 2 and A-complex was identified through dynamic clades. In addition to resolving all the biochemically described steps mentioned above, the SiMCAn technique provided evidence for the existence of a novel proofreading and discarding mechanism for incorrect 3'SS sequences after the first reaction. smFRET in combination with SiMCAn is a powerful tool to analyze multistep reactions as shown by the spliceosome. This technique was not only able to characterize distance changes throughout the full cycle, but also delivered kinetic data as every cluster also contains information about transition rates. Furthermore, new reaction steps were identified in the case of the spliceosome.

In summary, the studies presented here demonstrated how conformational analysis of parts of the spliceosome afford a functional connection between individual components of this highly complex RNP. The level of complexity requires both of the approaches presented here: bottom-up, where individual (sub-) domains of RNAs and proteins are put into focus, and top-down, where a less well-defined splicing activity conferred by cellular extracts are analyzed to yield insights from both directions.

## 6. G-Quadruplex Helicase

The key structural element of the complex protecting the telomeres are intricate DNA structures consisting of three guanosine tetrads, termed G-quadruplexes (G4s) [46]. These structures have been described for a large number of nucleic acid species, including not only chromosomal DNA termini, but presumably also in various other functionally relevant genomic regions and RNAs [47–50]. They are however evolutionarily depleted in bacterial transcriptomes [51].

The structure of these G-quadruplex elements depends on various factors, such as sequence composition, loop size, the cation bound into the quadruplex core, and the number of molecules involved in formation of quadruplex structures, i.e., inter- or intramolecular structures [47,49,52]. In general, all quadruplexes exhibit very high thermal stability, with RNA G4s usually exceeding the stability of DNA G4s. At the same time, RNA G4s are usually in a parallel orientation and show lower structural diversity than DNA G4 structures [53,54].

Aside from their function within telomeres, these structures have been described to be functionally relevant in processes like transcriptional pausing, replication, translation, and RNA processing. They exert their effect by utilizing the intrinsic thermal stability, leading to steric interference, competition with other conformations, or as recruiting platform for other proteins [54].

The broad spectrum of structures G4s are able to adopt, and their broad variety of binding partners as well as their intrinsic dynamics render these structures extremely difficult to study in their functional context. It is therefore not surprising that studies focusing on G4 structures in complex with e.g., protein binding partners are scarce.

There have been a number of proteins described that are able to unfold G4 structures in DNA as well as in RNA [53]. The most prominent member of RNA G4 helicases is the G4 resolvase RHAU (also termed DHX36 or G4 resolvase 1) which was first identified in 1997 [55]. It is able to resolve both inter- and intramolecular G4 in DNA and RNA [56–58] and its physiological importance is exemplified by its role in telomere reverse transcription. RHAU unfolds the intramolecular G4 in the 5'-region of human telomerase RNA template (hTR) and thereby promotes formation of the P1 helix, which prevents misincorporation during reverse transcription by serving as physical barrier [59].

Recent research on the RNA G4 helicase RHAU is focusing on the binding mode and unwinding mechanism on intramolecular DNA and RNA G4s. However, structural data obtained from crystal structures and small angle X-ray scattering (SAXS) even in combination was not sufficient to characterize

the unwinding process, as it provided only snapshots. The authors therefore opted to complement their structures of different complex compositions (i.e., in absence and presence of cofactors) by employing smFRET to investigate the conformationally dynamic nature of the unwinding process.

Previous structural and biochemical data revealed selective binding of RHAU to parallel DNA as well as RNA G4 [52,60]. smFRET data on structurally diverse G4s supported this finding, identifying RHAU binding only to the fraction of G4s corresponding to a parallel G4 fold [19]. For this study, an approach using immobilized substrate G4s was applied. To assess the binding and unwinding of the G4, donor and acceptor fluorophore were attached at opposite ends of the nucleotide strand (Figure 4A). A completely folded G4 brings the two fluorophores in close proximity and therefore exhibits high FRET efficiency values. Binding and unfolding of the G4 structure leads to a decrease in FRET efficiency, resulting in lower FRET efficiency values for an unfolded state. Several later studies used analogous labeling schemes to assess the unwinding mechanism of RHAU on G4s as they capture conformational changes in G4s along the whole process [20–22].

Moreover, a single stranded overhang of at least 9 nucleotides is required for efficient RHAU binding to G4s [19,57]. Later studies on bovine RHAU however indicate the possibility of recognizing antiparallel G4s as well [20,60], and binding of G4s regardless of their structure was shown for RHAU from *Drosophila*. In general, there are a number of differences described between bovine and *Drosophila* RHAU, showing why a detailed analysis at the level of individual molecules is of advantage.

The co-crystal structure of the DEAH-box helicase RHAU and a G4 DNA individually describes the importance of several protein domains (RecA1 and 2, DSM and OB like, see Figure 4B) for G4 binding as well as unwinding [20,22]. The G4 core is located in a positively charged cage, stacked under a DSM domain helix (Figure 4C). This structural element is essential for substrate binding in bovine, but not in *drosophila* RHAU [20,22,60]. smFRET assays analyzing this binding mode in detail used immobilized DNA and RNA G4s together with several mutants targeting the DSM helix of bovine RHAU. Dissociation of RHAU from DNA G4s upon buffer flow demonstrate the requirement for an intact DSM helix for tight DNA binding. However, in the same experimental setup, no dissociation from RNA G4s was detected, hinting towards a different role of DSM in RNA G4 binding [20,21].

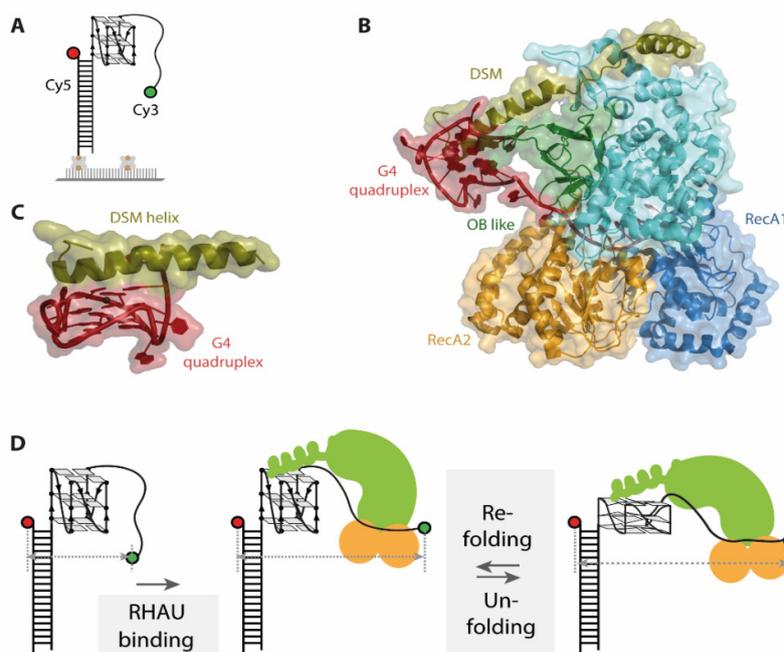
The conformational dynamics of the G4 unwinding mechanism by RHAU were then investigated via smFRET. In fact, different unwinding pathways were reported for DNA and RNA G4s [20–22]. For DNA, binding of the helicase to the substrate G4 is followed by repetitive unfolding and refolding of the quadruplex as evidenced by switching between two FRET states at 0.35 and 0.6, respectively [19] (Figure 4D). The observed FRET efficiency oscillations were elegantly shown to arise from a single RHAU molecule rather than protein dissociation and rebinding, as immobilizing the protein instead of the substrate did not alter the observed FRET behavior. An additional dwell time analysis of FRET states uncovered that the unfolding and refolding time of this process is closely related to the thermal stability of the G4, with faster refolding rates for more stable G4s. The repetitive unfolding mechanism would have remained elusive by structural approaches, which only captured the intermediate conformations in this cycle.

Comparison of the unwinding dynamics to structural data impressively augmented the understanding of the unwinding process. The unfolding and refolding represents a repetitive pulling mechanism on the 3'-most G in a direction outwards of the quadruplex. This pulling reorganizes the canonical G4 and results in formation of an ATGG quartet at the 5' end of the quadruplex, which is stacked on top of two canonical guanosine tetrads. Within the 3' single stranded region of the G4 this repetitive pulling results in an alternating stacking between 4 and 5 nucleotides in the protein tunnel [20,22].

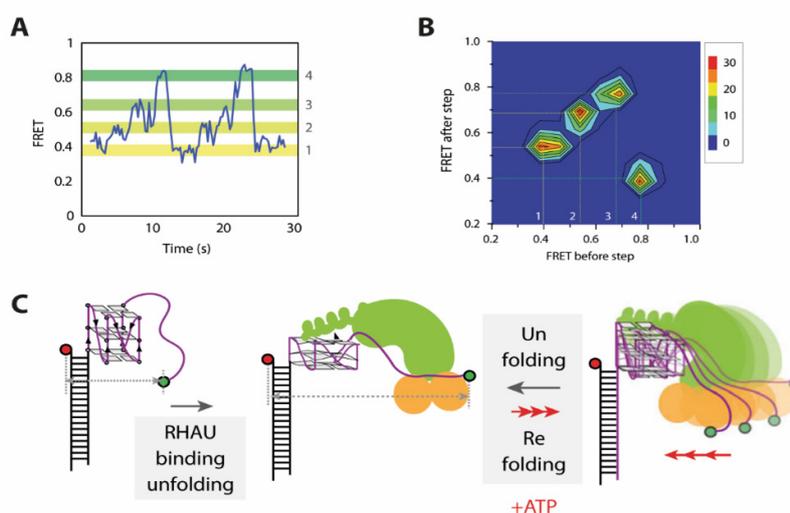
The abovementioned mechanism is however ATP independent, and only leads to partial unfolding of the G4 structure [20,22]. Nevertheless, Tippana et al. could show by smFRET assays that surprisingly, RHAU-mediated pulling of one single nucleotide from the G4 structure is sufficient to destabilize the quadruplex to such an extent, that a complementary cis-strand can efficiently anneal. This is indicated by a stable low FRET value, matching the measured FRET value of the completely unfolded

G4. Moreover, addition of ATP did not change the annealing or unfolding behavior [19]. The role of ATP in this unwinding process was reported to be different for bovine and drosophila RHAU [20,22].

Based on smFRET studies, the unwinding mechanism of RNA G4s by RHAU was very recently investigated by Tippana et al., describing a significantly different conformationally dynamic nature than that of DNA G4 unfolding by RHAU [21]. Binding of RHAU to RNA G4s results in a two-step decrease from a high FRET state at 0.8 to an intermediate state at 0.6, and then to a low FRET value of 0.4, which up to this point is similar to the ATP independent partial unfolding of DNA. However, after binding ( $E_{\text{FRET}} 0.6$ ) and partial disruption ( $E_{\text{FRET}} 0.4$ ) of the G4, no FRET fluctuations between the two states were detected, which indicates that there is no refolding step in the absence of ATP. Upon addition of ATP, refolding of the G4 by RHAU was observed in three distinct steps. The last step of refolding was followed by a rapid decrease back to a FRET state at 0.4, which was assigned to be a second single-step unfolding (Figure 5A). Transition density plots (TDPs) of the FRET traces were able to confirm that this is a closed cycle, showing no additional interconversions (Figure 5B). In contrast to the faster DNA G4 repetitive unwinding by RHAU, the significantly slower unfolding and refolding cycle of RNA G4 was observed repeatedly in the presence of ATP, prior to RHAU dissociation from the substrate RNA (Figure 5A,C). Similar to studies on DNA [19], an approach using immobilized RHAU confirmed that this motion arises from a single RHAU molecule and is not due to repeated docking and dissociation. This is a perfect example of how smFRET spectroscopy is able to complement existing structural data with relevant kinetic information and interconversion pathways between intermediates.



**Figure 4.** (A) Schematic representation of one G-quadruplex (G4) single molecule Förster resonance energy transfer (smFRET) construct used in the described studies. (B) Co-crystal structure (pdb 5VHE) of RHAU in complex with a c-Myc G-quadruplex, color-coded for different important protein domains. (C) The DSM helix of RHAU stacks as planar, unipolar surface on top of the G-quadruplex. (D) Unfolding mechanism of DNA G4s by RHAU derived from smFRET studies. Binding of RHAU to the G4 results in FRET decrease, followed by ATP independent, repetitive partial unfolding and refolding of DNA G4. Figures A and D were adapted and modified from Tippana et al. and are used under Creative Commons Attribution 4.0 International License (CC-BY 4.0).



**Figure 5.** Unfolding mechanism of RNA G4s by RHAU. **(A)** FRET traces of ATP dependent repetitive unfolding and stepwise refolding. **(B)** Transition density plots (TDPs) generated from FRET traces. **(C)** Unwinding mechanism of RNA G4s by RHAU developed by smFRET studies. Initial binding of RHAU induces partial unfolding, followed by ATP dependent, repetitive stepwise refolding and successive unfolding. This Figure was taken from Tippiana et al. and used Creative Commons Attribution 4.0 International License (CC-BY 4.0).

The ATP dependence of the RNA G4 refolding step by RHAU arises from the presence of a 3' single stranded RNA overhang [21]. Innovative smFRET constructs built from RNA-DNA chimeras (DNA G4–RNA overhang, RNA G4–DNA overhang) showed significant differences in RHAU-mediated unfolding dynamics. Constructs with a DNA overhang and an RNA G4 exhibited the FRET fluctuations known from DNA G4 unfolding by RHAU, whereas constructs with an RNA overhang displayed movements similar to RNA unfolding dynamics. The behavior upon ATP addition is also analogous, showing an ATP independence for unwinding of the RNA G4–DNA overhang chimera, and a strict ATP dependence for unwinding of the DNA G4–RNA overhang chimera.

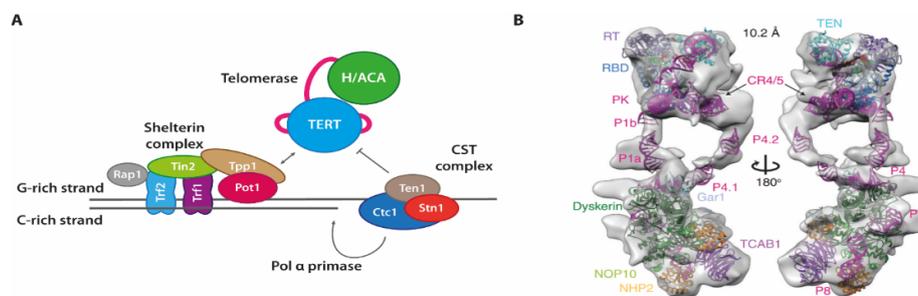
The de facto unwinding of the RNA G4 by RHAU was proven by cis-strand annealing to the RNA. Under ATP hydrolysis, efficient annealing was observed, indicated by a stable low FRET state that was persisting under SDS-buffer flow. Without addition of ATP or with ATPase deficient RHAU mutants, SDS-buffer flow induced a refolding of the G4 due to dissociation of RHAU from the substrate, shifting the FRET values back to high FRET.

When looking at the studies on G4 unwinding presented here, smFRET based findings were able to complement the structural data to a remarkably clearer understanding of the conformationally dynamic nature of the unwinding process, experimentally linking individual structures to a distinct mechanism: First, the conformational dynamics of the G4 structures throughout the unwinding process show that the interactions between the protein and the G4 are more complex than anticipated and differences kinetics between RNA G4s and DNA G4s. The differences in ATP-dependence was also surprising, and the analysis using RNA-DNA chimera constructs elegantly isolated the individual contribution of structural elements on the unwinding dynamics. All of these findings nicely complement the information on the same complexes available from structural analysis.

## 7. Telomere Maintenance

Each of the chromosomes in eukaryotic cells consists of linear double-stranded DNA (ds DNA), which also contains termini of the linear DNA at each end. In order to prevent undesired recombination or other erroneous processing of these DNA ends (as they may mistakenly be recognized as double-strand breaks), complexes termed shelterin are placed onto and shield each of the chromosome ends [61]. These proteins also help regulate the selective elongation of both terminal DNA strands, and

are thus a key player in maintaining a specific length of the telomeres [62] (Figure 6). To selectively evade the shortening of telomeric sequences during cell division, cells use a specialized system consisting of two complexes each synthesizing one strand of the genomic DNA. For the G-rich strand that is synthesized de novo, an RNP containing a reverse transcriptase with an integral RNA template form the telomerase enzyme. For the C-rich strand, the G-rich strand is used as a template. Here, the CST-complex consisting of three proteins Ctc1, Stn1 and Ten1, recruits the Pol  $\alpha$ -Primase, and also inhibits telomerase.



**Figure 6.** (A) Protein and RNP complexes at the chromosome ends. Whereas the shelterin complex binds to protect the chromosome ends, telomerase as well as the CST complex act on the G-rich or C-rich strand, respectively, to facilitate their synthesis. (B) Cryo-EM-based model structure of the human telomerase holoenzyme, taken from [63]. Copyright of Figure 6B by Wang et al. is with CSHP—Copyright for use has been requested from the publisher.

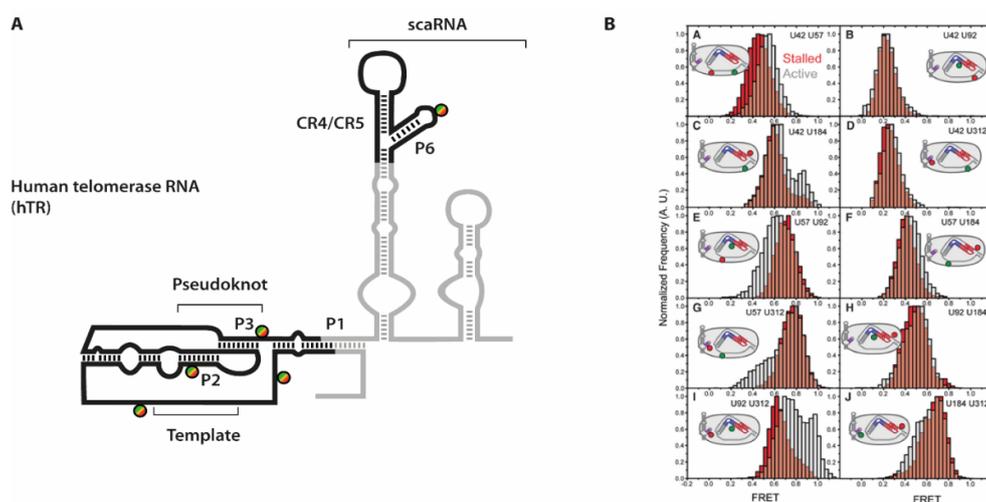
Telomerase holoenzyme RNP complexes are assembled and processed in the nucleus and nucleolus before being selectively localized at chromosome ends [64]. The use of an integral RNA template to add repetitive DNA sequences requires that a short duplex of RNA template and the newly synthesized DNA are shifted for each nucleotide added to allow for consecutive addition of bases (“nucleotide addition processivity”). After reaching the end of the RNA template, the duplex has to be realigned to allow for the synthesis of a new repeat (“repeat addition processivity”).

The unique architecture and functionality of telomerase enzymes has made a detailed functional characterization as well as structural biology approaches difficult for a long time. Initial, partial structures of protein domains from diverse organisms including the reverse transcriptase TERT [65,66], accessory proteins like the Tetrahymena p65 [67], or parts of RNA elements within the RNA component [68,69] have provided some insight. Fully assembled complexes however were extremely difficult to obtain [70]. This was changed by more recent cryogenic electron microscopy (cryoEM) structures first on the Tetrahymena holoenzyme [71], and later on high-resolution structures of Tetrahymena telomerase [72] as well as the human telomerase holoenzyme [73].

Together with a number of single molecule-based approaches, the structures also shed light on aspects of telomerase subunit folding [74–77], as well as on the general conformational dynamics that are required for telomerase activity [78,79].

Especially for the human telomerase holoenzyme however, the overall arrangement of the RNA component relative to TERT was difficult to establish. One particular obstacle was the discontinuous set of structural and functional elements within hTR (see Figure 7).

In order to elucidate the orientation of these domains towards each other as well as upon binding to TERT, Parks et al. combined an extensive set of smFRET experiments with a modeling approach [24]. They placed FRET pairs on five positions into different parts of hTR, reconstituted these RNAs with TERT into functional constructs and derived distance constraints from these experiments. The distances were strategically placed into each of the hTR domains: close to the P1 helix, two positions on either side adjacent to the pseudoknot, 3' of the template sequence, and within the CR4/CR5 element P6 helix.



**Figure 7.** (A) Secondary structure of human telomerase RNA hTR with labeling sites. The parts in grey are representing the full RNA, while the parts in black have been used in the study [24]. (B) smFRET histograms derived from the ten labeling site combinations. Data derived from the stalled complex is depicted in red, while complexes selected for activity are in grey. Figure 7B is taken from Parks et al. and used under Creative Commons Non-Commercial License (CC-BY NC).

To efficiently immobilize fully reconstituted telomerase holoenzyme species, a specific primer substrate showing high affinity towards the assembled complex was used [80]. In absence of dNTP substrates, these complexes are referred to as “stalled” complexes. In order to then obtain data throughout the telomere synthesis cycle, the “active” complexes were imaged after addition of substrate dNTPs [81]. Conformational dynamics that occur within hTR would then result in a different distance to be reported in FRET experiments. Indeed, for several of the ten labeling combinations, differences between “stalled” and “active” complexes were observed. The conformational shifts expectedly were not quantitative (see e.g., Figure 7), as several studies showed that a large fraction of human telomerase complexes that can be isolated from cells are actually catalytically inactive [64,73]. However, the additional FRET populations allowed deriving of an additional set of distances within the telomerase holoenzyme. Interestingly, the differences were mostly observed in the orientation of either the template sequence and the pseudoknot structure, or between the pseudoknot structure and the CR4/CR5 hairpin, respectively.

The additional information derived in this study was however obtained by including the distance restraints derived from FRET experiments into a structural model of hTR architecture that was in line with previous experimental datasets [62,79]. Considering functional restraints, a clustered Rosetta modeling approach was used to obtain a likely fold for hTR in complex with TERT. The authors showed that only when considering the FRET restraints, these models were convergent (Figure 8). These models were in line with the models available for the Tetrahymena telomerase at that time. Rewardingly, these models were also in excellent agreement with the structural data of human telomerase generated by cryoEM later on [73]. Even considering dynamic movements within the holoenzyme, the study produced viable models for conformations that can be assumed throughout the active cycle.

While telomerase provides one factor to extend the successively shortened G-rich strand, the C-rich complementary strand also has to be selectively synthesized (Figure 6). In order to exert control over both quadruplex formation and telomerase, human cells employ the CST complex, consisting of proteins Ctc1, Stn1, and Ten1. This complex limits the activity of telomerase [82], unwinds the G-rich DNA quadruplex structures [83] to allow for the “fill-in” reaction of the C-rich strand by Pol $\alpha$ , and by combination of these processes limits the length of the overhang formed by the G-rich strand [84].

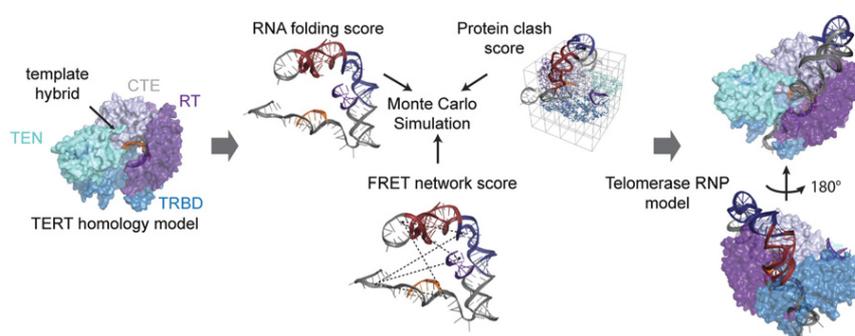
How the CST complex is able to make the G-rich strand accessible for further DNA synthesis is less well understood. Here, an approach combining biochemical and smFRET analysis of DNA

duplexes with different overhang lengths and positions in presence of the CST complex provided deeper insight [25]: In electrophoretic mobility shift assay experiments, high affinity of CST to telomeric overhang structures shows that the site where the dsDNA transitions into single stranded (ss) DNA at the G-rich strand is specifically recognized and destabilized by CST. As a biochemical assessment of these heterogeneous systems was difficult, the authors used an smFRET approach where – similar to other G4 enzymes (Figure 5)–a folded DNA was immobilized and G4 stability monitored by FRET. Analysis of these biochemically identified structures unveiled that CST is also able to unwind the G4 structure in the context of the ds-to-ss-DNA junction (Figure 9A). This was also the case when there was a non-telomeric overhang present, showing that the unfolding was structure- rather than sequence-specific.

Binding of CST *in situ* caused an immediate shift into a low-FRET, G4 unfolded conformation (Figure 9B), and a number of additional experiments where the protein trimer was added during the experiment showed that this process was fully reversible (Figure 9C). This finding immediately puts forward a model where the specific recognition of the overhang junction and the immediate destabilization of the G4 structure facilitate the elongation of the C-rich strand, using the unstructured G-rich strand as a template.

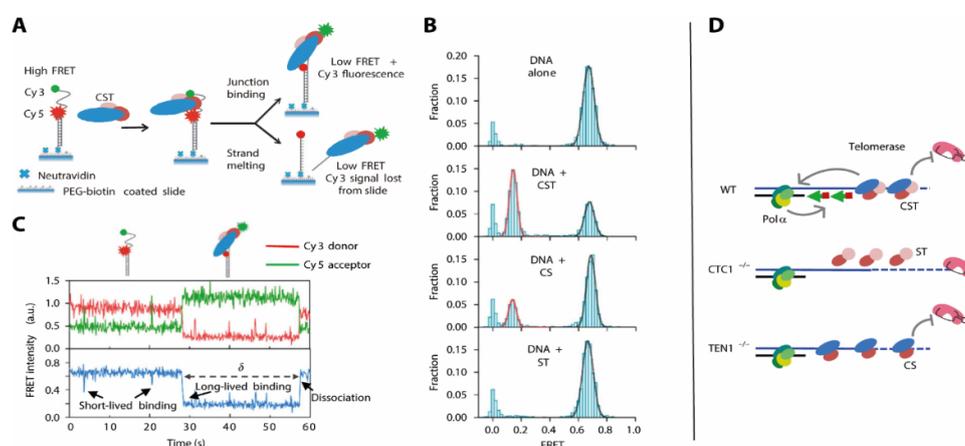
These findings however were derived from the CST trimer in absence of in-depth functional knowledge about any of the proteins. These individual contributions were then addressed in a follow-up study [23], showing that for Ctc1-depleted cells, there is both an increase in length for the G-rich strand and a length decrease in the C-rich strand.

To test how these different observations were exerted by the different subunits of the CST complex, the authors expressed and purified “CS” and “ST” complexes containing Ctc1-Stn1 and Stn1-Ten1, respectively. They assessed the G4 unfolding activity of these complexes in smFRET experiments and compared them to the results using the full CST complex (Figure 9). They found that the ST complex was unable to resolve the G4 structure within the G-rich strand, indicated by an unchanged high-FRET population. The CS complex had activity that was reduced, which was indicated by a difference in short-lived binding events between the DNA G4 and the CS complex. The resulting unfolded state of the DNA however was comparable to that generated by the CST complex. This is in perfect agreement with the results using cell line analysis: While CST both recruits Pol $\alpha$  and inhibits telomerase activity, these functions can be separated for the subcomplexes (Figure 9D). Here, the ST complex (in absence of Ctc1) allows for G-rich strand elongation by losing its telomerase inhibitor function. The CS complex in absence of Ten1 however diminishes telomerase activity, but lacks the feature to recruit Pol $\alpha$  in order to fill in the C-rich strand.



**Figure 8.** Workflow to derive structural models of higher quality by incorporating the FRET network score into Rosetta modeling. With existing structural data from homology models, the authors combined restraints obtained from RNA folding, steric calculations and smFRET data into a novel model of the hTR-TERT RNP. This Figure is taken from Parks et al. and used under Creative Commons Non-Commercial License (CC-BY NC).

This again provides an elegant example how carefully crafted smFRET experiments are able to fill in gaps in the functional understanding of such processes by comparative analysis of subcomplexes, and the combination of such experiments with biochemical assessment of substrate recognition.



**Figure 9.** (A) smFRET construct design to investigate G4 resolvase activity of the CST complexes. (B) smFRET data depicting the efficient unfolding of the G4 construct by CST, less efficient unfolding in presence of CS, and no unfolding with ST. (C) Time-resolved unfolding and identification of intermediates in presence of CST. (D) Model for the interplay between the CST complex and Pol $\alpha$ . This Figure is adapted from Feng et al. [23] and used under Creative Commons Non-Commercial License (CC-BY).

## 8. Conclusions

Taken together, the studies on various RNP complexes presented here all clearly show that strategic smFRET experiments can significantly improve the understanding of complex systems, on their conformationally dynamic nature but just as much on their structural composition.

The research on the spliceosome summarized here shows how smFRET was able to shed light onto the complex rearrangements necessary throughout the splicing cycle, where structure determination such as crystallization and cryo-EM only delivered snapshots. With different labeling schemes on either protein, snRNA, or substrate RNA it was possible to detect the RNA- and ATP-dependent closing of Prp5, the spontaneous or assisted conformational toggling of U2 snRNA in dependence of ions, proteins and modification as well as pre-mRNA dynamics throughout the whole splicing cycle. While some of these principles have been hypothesized previously, single-molecule data could provide crucial evidence for their existence. Furthermore, these smFRET studies were able to additionally find and visualize novel intermediates and pathways.

The presented studies on G4 unwinding by RHAU show the benefit of single molecule experiments when combining them with structural data on RHAU. The identification of distinct states throughout the unwinding process, the dynamics that come along with both the interaction with the enzyme and the ATP hydrolysis step and, finally, the impact of mutations on such dynamics all would have been extremely difficult to rationalize from structural or biochemical data alone.

The unique telomerase mechanism proved to be especially challenging to decipher. The smFRET data gave significantly deeper insights into many of the aspects in telomerase folding, assembly and activity, and demonstrated how questions that were again difficult to tackle by structural biology means were more readily amenable by single molecule techniques. Telomerase holoenzyme structures published in the last years confirmed that several previously hypothesized insights into the mechanism of telomerase-catalyzed DNA synthesis were correct. This mostly concerned the binding locations of proteins and their interactions with the RNA, but also the general architecture and arrangement of the RNA within the telomerase holoenzyme. They also further structurally explained how the

telomerase cycles during DNA synthesis, combining nucleotide addition processivity and repeat addition processivity to efficiently add significant amounts of DNA onto the chromosome ends.

By looking at strategically reconstituted subcomplexes, it was again the view into individual molecules that precisely identified functional roles for each protein within the CST complex.

In summary, the work showcased here provides a number of prominent examples in how functional questions on RNP complexes can be tackled by combining and comparing structural biology techniques with FRET spectroscopy on the level of individual molecules.

**Author Contributions:** N.M., C.F. and M.H. discussed the enclosed literature and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by DFG in SFB902 “Molecular Mechanisms of RNA-mediated regulation”, and by Goethe University Frankfurt.

**Acknowledgments:** The authors would like to thank Harald Schwalbe for constant support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yusupov, M.M.; Yusupova, G.Z.; Baucom, A.; Lieberman, K.; Earnest, T.N.; Cate, J.H.D.; Noller, H.F. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **2001**, *292*, 883–896. [[CrossRef](#)]
2. Jaliha, A.P.; Lund, P.E.; Walter, N.G. Coming together: Rnas and proteins assemble under the single-molecule fluorescence microscope. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032441. [[CrossRef](#)]
3. Hanspach, G.; Trucks, S.; Hengesbach, M. Strategic labelling approaches for RNA single-molecule spectroscopy. *RNA Biol.* **2019**, *16*, 1119–1132. [[CrossRef](#)]
4. Milles, S.; Tyagi, S.; Banterle, N.; Koehler, C.; Vandelinder, V.; Plass, T.; Neal, A.P.; Lemke, E.A. Click strategies for single-molecule protein fluorescence. *J. Am. Chem. Soc.* **2012**, *134*, 5187–5195. [[CrossRef](#)] [[PubMed](#)]
5. Jakob, L.; Gust, A.; Grohmann, D. Evaluation and optimisation of unnatural amino acid incorporation and bioorthogonal bioconjugation for site-specific fluorescent labelling of proteins expressed in mammalian cells. *Biochem. Biophys. Rep.* **2019**, *17*, 1–9. [[CrossRef](#)]
6. Schuler, B.; Soranno, A.; Hofmann, H.; Nettels, D. Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu. Rev. Biophys.* **2016**, *45*, 207–231. [[CrossRef](#)]
7. Andreou, A.Z.; Klostermeier, D. EIF4B and eIF4G jointly stimulate eIF4A ATPase and unwinding activities by modulation of the eIF4A conformational cycle. *J. Mol. Biol.* **2014**, *426*, 51–61. [[CrossRef](#)]
8. Andreou, A.Z.; Harms, U.; Klostermeier, D. Single-stranded regions modulate conformational dynamics and ATPase activity of eIF4A to optimize 5'-UTR unwinding. *Nucleic Acids Res.* **2019**, *47*, 5260–5275. [[CrossRef](#)]
9. Munro, J.B.; Altman, R.B.; O'Connor, N.; Blanchard, S.C. Identification of Two Distinct Hybrid State Intermediates on the Ribosome. *Mol. Cell* **2007**, *25*, 505–517. [[CrossRef](#)]
10. Prokhorova, I.; Altman, R.B.; Djumagulov, M.; Shrestha, J.P.; Urzhumtsev, A.; Ferguson, A.; Chang, C.W.T.; Yusupov, M.; Blanchard, S.C.; Yusupova, G.; et al. Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E10899–E10908. [[CrossRef](#)]
11. Wang, L.; Wasserman, M.R.; Feldman, M.B.; Altman, R.B.; Blanchard, S.C. Mechanistic insights into antibiotic action on the ribosome through single-molecule fluorescence imaging. *Ann. N. Y. Acad. Sci.* **2011**, *1241*, E1. [[CrossRef](#)] [[PubMed](#)]
12. Cui, T.J.; Joo, C. Facilitated diffusion of Argonaute-mediated target search. *RNA Biol.* **2019**, *16*, 1093–1107. [[CrossRef](#)] [[PubMed](#)]
13. Cui, T.J.; Klein, M.; Hegge, J.W.; Chandradoss, S.D.; van der Oost, J.; Depken, M.; Joo, C. Argonaute bypasses cellular obstacles without hindrance during target search. *Nat. Commun.* **2019**, *10*, 1–11. [[CrossRef](#)]
14. Kramm, K.; Endesfelder, U.; Grohmann, D. A Single-Molecule View of Archaeal Transcription. *J. Mol. Biol.* **2019**, *431*, 4116–4131. [[CrossRef](#)] [[PubMed](#)]
15. Blanco, M.R.; Martin, J.S.; Kahlscheuer, M.L.; Krishnan, R.; Abelson, J.; Laederach, A.; Walter, N.G. Single Molecule Cluster Analysis dissects splicing pathway conformational dynamics. *Nat. Methods* **2015**, *12*, 1077–1084. [[CrossRef](#)] [[PubMed](#)]
16. Rodgers, M.L.; Tretbar, U.S.; Dehaven, A.; Alwan, A.A.; Luo, G.; Mast, H.M.; Hoskins, A.A. Conformational dynamics of stem II of the U2 snRNA. *RNA* **2016**, *22*, 225–236. [[CrossRef](#)]

17. Van der Feltz, C.; DeHaven, A.C.; Hoskins, A.A. Stress-induced Pseudouridylation Alters the Structural Equilibrium of Yeast U2 snRNA Stem II. *J. Mol. Biol.* **2018**, *430*, 524–536. [[CrossRef](#)]
18. Beier, D.H.; Carrocci, T.J.; van der Feltz, C.; Tretbar, U.S.; Paulson, J.C.; Grabowski, N.; Hoskins, A.A. Dynamics of the DEAD-box ATPase Prp5 RecA-like domains provide a conformational switch during spliceosome assembly. *Nucleic Acids Res.* **2019**, *47*, 10842–10851. [[CrossRef](#)]
19. Tippana, R.; Hwang, H.; Opresko, P.L.; Bohr, V.A.; Myong, S. Single-molecule imaging reveals a common mechanism shared by G-quadruplex-resolving helicases. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 8448–8453. [[CrossRef](#)]
20. Chen, M.C.; Tippana, R.; Demeshkina, N.A.; Murat, P.; Balasubramanian, S.; Myong, S.; Ferré-D'amaré, A.R. Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* **2018**, *558*, 465–483. [[CrossRef](#)]
21. Tippana, R.; Chen, M.C.; Demeshkina, N.A.; Ferré-D'amaré, A.R.; Myong, S. RNA G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. *Nat. Commun.* **2019**, *10*, 1855. [[CrossRef](#)]
22. Chen, W.F.; Rety, S.; Guo, H.L.; Dai, Y.X.; Wu, W.Q.; Liu, N.N.; Auguin, D.; Liu, Q.W.; Hou, X.M.; Dou, S.X.; et al. Molecular Mechanistic Insights into Drosophila DHX36-Mediated G-Quadruplex Unfolding: A Structure-Based Model. *Structure* **2018**, *26*, 403–415.e4. [[CrossRef](#)] [[PubMed](#)]
23. Feng, X.; Hsu, S.J.; Bhattacharjee, A.; Wang, Y.; Diao, J.; Price, C.M. CTC1-STN1 terminates telomerase while STN1-TEN1 enables C-strand synthesis during telomere replication in colon cancer cells. *Nat. Commun.* **2018**, *9*, 1–12. [[CrossRef](#)] [[PubMed](#)]
24. Parks, J.W.; Kappel, K.; Das, R.; Stone, M.D. Single-molecule FRET-Rosetta reveals RNA structural rearrangements during human telomerase catalysis. *Rna* **2017**, *23*, 175–188. [[CrossRef](#)] [[PubMed](#)]
25. Bhattacharjee, A.; Wang, Y.; Diao, J.; Price, C.M. Dynamic DNA binding, junction recognition and G4 melting activity underlie the telomeric and genome-wide roles of human CST. *Nucleic Acids Res.* **2017**, *45*, 12311–12324. [[CrossRef](#)]
26. Wan, R.; Yan, C.; Bai, R.; Wang, L.; Huang, M.; Wong, C.C.L.; Shi, Y. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: Insights into spliceosome assembly and catalysis. *Science* **2016**, *351*, 466–475. [[CrossRef](#)]
27. Agafonov, D.E.; Kastner, B.; Dybkov, O.; Hofele, R.V.; Liu, W.T.; Urlaub, H.; Lührmann, R.; Stark, H. Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* **2016**, *351*, 1416–1420. [[CrossRef](#)]
28. Nguyen, T.H.D.; Galej, W.P.; Bai, X.C.; Oubridge, C.; Newman, A.J.; Scheres, S.H.W.; Nagai, K. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* **2016**, *530*, 298–302. [[CrossRef](#)]
29. Krishnan, R.; Blanco, M.R.; Kahlscheuer, M.L.; Abelson, J.; Guthrie, C.; Walter, N.G. Biased Brownian ratcheting leads to pre-mRNA remodeling and capture prior to first-step splicing. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1450–1457. [[CrossRef](#)]
30. Dalbadie-McFarland, G.; Abelson, J. PRP5: A helicase-like protein required for mRNA splicing in yeast. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4236–4240. [[CrossRef](#)]
31. Zhang, Z.M.; Yang, F.; Zhang, J.; Tang, Q.; Li, J.; Gu, J.; Zhou, J.; Xu, Y.Z. Crystal Structure Of Prp5P Reveals Interdomain Interactions That Impact Spliceosome Assembly. *Cell Rep.* **2013**, *5*, 1269–1278. [[CrossRef](#)] [[PubMed](#)]
32. Liang, W.W.; Cheng, S.C. A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes Dev.* **2015**, *29*, 81–93. [[CrossRef](#)] [[PubMed](#)]
33. Xu, Y.Z.; Query, C.C. Competition between the ATPase Prp5 and Branch Region-U2 snRNA Pairing Modulates the Fidelity of Spliceosome Assembly. *Mol. Cell* **2007**, *28*, 838–849. [[CrossRef](#)] [[PubMed](#)]
34. Karow, A.R.; Klostermeier, D. A conformational change in the helicase core is necessary but not sufficient for RNA unwinding by the DEAD box helicase YxiN. *Nucleic Acids Res.* **2009**, *37*, 4464–4471. [[CrossRef](#)] [[PubMed](#)]
35. Perriman, R.; Barta, I.; Voeltz, G.K.; Abelson, J.; Ares, M. ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13857–13862. [[CrossRef](#)] [[PubMed](#)]
36. Talkish, J.; Igel, H.; Hunter, O.; Horner, S.W.; Jeffery, N.N.; Leach, J.R.; Jenkins, J.L.; Kielkopf, C.L.; Ares, M. Cus2 enforces the first ATP-dependent step of splicing by binding to yeast SF3b1 through a UHM-UlM interaction. *Rna* **2019**, *25*, 1020–1037. [[CrossRef](#)]
37. Hilliker, A.K.; Mefford, M.A.; Staley, J.P. U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. *Genes Dev.* **2007**, *21*, 821–834. [[CrossRef](#)] [[PubMed](#)]

38. Perriman, R.J.; Ares, M. Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. *Genes Dev.* **2007**, *21*, 811–820. [[CrossRef](#)]
39. Yan, D.; Perriman, R.; Igel, H.; Howe, K.J.; Neville, M.; Ares, M. CUS2, a Yeast Homolog of Human Tat-SF1, Rescues Function of Misfolded U2 through an Unusual RNA Recognition Motif. *Mol. Cell. Biol.* **1998**, *18*, 5000–5009. [[CrossRef](#)]
40. Newby, M.L.; Greenbaum, N.L. A conserved pseudouridine modification in eukaryotic U2 snRNA induces a change in branch-site architecture. *Rna* **2001**, *7*, 833–845. [[CrossRef](#)]
41. Wu, G.; Adachi, H.; Ge, J.; Stephenson, D.; Query, C.C.; Yu, Y. Pseudouridines in U2 snRNA stimulate the ATPase activity of Prp5 during spliceosome assembly. *EMBO J.* **2016**, *35*, 654–667. [[CrossRef](#)] [[PubMed](#)]
42. Yang, C.; McPheeters, D.S.; Yu, Y.T. Ψ35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **2005**, *280*, 6655–6662. [[CrossRef](#)] [[PubMed](#)]
43. Wu, G.; Radwan, M.K.; Xiao, M.; Adachi, H.; Fan, J.; Yu, Y.T. The TOR signaling pathway regulates starvation-induced pseudouridylation of yeast U2 snRNA. *Rna* **2016**, *22*, 1146–1152. [[CrossRef](#)] [[PubMed](#)]
44. Wu, G.; Xiao, M.; Yang, C.; Yu, Y.T. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J.* **2011**, *30*, 79–89. [[CrossRef](#)] [[PubMed](#)]
45. Bronson, J.E.; Fei, J.; Hofman, J.M.; Gonzalez, R.L.; Wiggins, C.H. Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. *Biophys. J.* **2009**, *97*, 3196–3205. [[CrossRef](#)]
46. Gilbert, D.E.; Feigon, J. Multistranded DNA structures. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305–314. [[CrossRef](#)]
47. Bochman, M.L.; Paeschke, K.; Zakian, V.A. DNA secondary structures: Stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **2012**, *13*, 770–780. [[CrossRef](#)]
48. Lam, E.Y.N.; Beraldi, D.; Tannahill, D.; Balasubramanian, S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* **2013**, *4*, 1–8. [[CrossRef](#)]
49. Rhodes, D.; Lipps, H.J. Survey and summary G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **2015**, *43*, 8627–8637. [[CrossRef](#)]
50. Luke, B.; Lingner, J. TERRA: Telomeric repeat-containing RNA. *EMBO J.* **2009**, *28*, 2503–2510. [[CrossRef](#)]
51. Guo, J.U.; Bartel, D.P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **2016**, *353*, aaf5371. [[CrossRef](#)] [[PubMed](#)]
52. Tippiana, R.; Xiao, W.; Myong, S. G-quadruplex conformation and dynamics are determined by loop length and sequence. *Nucleic Acids Res.* **2014**, *42*, 8106–8114. [[CrossRef](#)]
53. Mendoza, O.; Bourdoncle, A.; Boulé, J.B.; Brosh, R.M.; Mergny, J.L. G-quadruplexes and helicases. *Nucleic Acids Res.* **2016**, *44*, 1989–2006. [[CrossRef](#)] [[PubMed](#)]
54. Millevoi, S.; Moine, H.; Vagner, S. G-quadruplexes in RNA biology. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 495–507. [[CrossRef](#)]
55. Harrington, C.; Lan, Y.; Akman, S.A. The Identification and Characterization of a G4-DNA Resolvase Activity. *J. Biol. Chem.* **1997**, *272*, 24631–24636. [[CrossRef](#)] [[PubMed](#)]
56. Creacy, S.D.; Routh, E.D.; Iwamoto, F.; Nagamine, Y.; Akman, S.A.; Vaughn, J.P. G4 Resolvase 1 Binds Both DNA and RNA Tetramolecular Quadruplex with High Affinity and Is the Major Source of Tetramolecular Quadruplex G4-DNA and G4-RNA Resolving Activity in HeLa Cell Lysates. *J. Biol. Chem.* **2008**, *283*, 34626–34634. [[CrossRef](#)]
57. Giri, B.; Smaldino, P.J.; Thys, R.G.; Creacy, S.D.; Routh, E.D.; Hantgan, R.R.; Lattmann, S.; Nagamine, Y.; Akman, S.A.; Vaughn, J.P. G4 Resolvase 1 tightly binds and unwinds unimolecular G4-DNA. *Nucleic Acids Res.* **2011**, *39*, 7161–7178. [[CrossRef](#)]
58. Lattmann, S.; Stadler, M.B.; Vaughn, J.P.; Akman, S.A.; Nagamine, Y. The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme. *Nucleic Acids Res.* **2011**, *39*, 9390–9404. [[CrossRef](#)]
59. Chen, J.-L. Template boundary definition in mammalian telomerase. *Genes Dev.* **2003**, *17*, 2747–2752. [[CrossRef](#)]
60. Heddi, B.; Cheong, V.V.; Martadinata, H.; Phan, A.T. Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide–quadruplex complex. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 9608–9613. [[CrossRef](#)]

61. De Lange, T. Shelterin-Mediated Telomere Protection. *Annu. Rev. Genet.* **2018**, *52*, 223–247. [[CrossRef](#)] [[PubMed](#)]
62. Chan, H.; Wang, Y.; Feigon, J. Progress in Human and Tetrahymena Telomerase Structure Determination. *Annu. Rev. Biophys.* **2017**, *46*, 199–225. [[CrossRef](#)]
63. Wang, Y.; Sušac, L.; Feigon, J. Structural biology of telomerase. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032383. [[CrossRef](#)] [[PubMed](#)]
64. Egan, E.D.; Collins, K. Biogenesis of telomerase ribonucleoproteins. *RNA* **2012**, *18*, 1747–1759. [[CrossRef](#)] [[PubMed](#)]
65. Mitchell, M.; Gillis, A.; Futahashi, M.; Fujiwara, H.; Skordalakes, E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat. Struct. Mol. Biol.* **2010**, *17*, 513–518. [[CrossRef](#)]
66. Jansson, L.I.; Akiyama, B.M.; Ooms, A.; Lu, C.; Rubin, S.M.; Stone, M.D. Structural basis of template-boundary definition in Tetrahymena telomerase. *Nat. Struct. Mol. Biol.* **2015**, *22*, 883–888. [[CrossRef](#)]
67. Singh, M.; Wang, Z.; Koo, B.K.; Patel, A.; Cascio, D.; Collins, K.; Feigon, J. Structural Basis for Telomerase RNA Recognition and RNP Assembly by the Holoenzyme La Family Protein p65. *Mol. Cell* **2012**, *47*, 16–26. [[CrossRef](#)]
68. Zhang, Q.; Kim, N.K.; Peterson, R.D.; Wang, Z.; Feigon, J. Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18761–18768. [[CrossRef](#)]
69. Kim, N.-K.; Theimer, C.A.; Mitchell, J.R.; Collins, K.; Feigon, J. Effect of pseudouridylation on the structure and activity of the catalytically essential P6.1 hairpin in human telomerase RNA. *Nucleic Acids Res.* **2010**, *38*, 6746–6756. [[CrossRef](#)]
70. Sauerwald, A.; Sandin, S.; Cristofari, G.; Scheres, S.H.W.; Lingner, J.; Rhodes, D. Structure of active dimeric human telomerase. *Nat. Struct. Mol. Biol.* **2013**, *20*, 454–460. [[CrossRef](#)]
71. Jiang, J.; Chan, H.; Cash, D.D.; Miracco, E.J.; Loo, R.R.O.; Upton, H.E.; Cascio, D.; Johnson, R.O.B.; Collins, K.; Loo, J.A.; et al. Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science* **2015**, *350*, aab4070. [[CrossRef](#)] [[PubMed](#)]
72. Jiang, J.; Wang, Y.; Sušac, L.; Chan, H.; Basu, R.; Zhou, Z.H.; Feigon, J. Structure of Telomerase with Telomeric DNA. *Cell* **2018**, *173*, 1179–1190.e13. [[CrossRef](#)] [[PubMed](#)]
73. Nguyen, T.H.D.; Tam, J.; Wu, R.A.; Greber, B.J.; Toso, D.; Nogales, E.; Collins, K. Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* **2018**, *557*, 190–195. [[CrossRef](#)]
74. Mihalusova, M.; Wu, J.Y.; Zhuang, X. Functional importance of telomerase pseudoknot revealed by single-molecule analysis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20339–20344. [[CrossRef](#)]
75. Stone, M.D.; Mihalusova, M.; O'Connor, C.M.; Prathapam, R.; Collins, K.; Zhuang, X. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **2007**, *446*, 458–461. [[CrossRef](#)]
76. Chen, G.; Wen, J.D.; Tinoco, I. Single-molecule mechanical unfolding and folding of a pseudoknot in human telomerase RNA. *RNA* **2007**, *13*, 2175–2188. [[CrossRef](#)]
77. Hengesbach, M.; Kim, N.K.; Feigon, J.; Stone, M.D. Single-molecule FRET reveals the folding dynamics of the human telomerase rna pseudoknot domain. *Angew. Chem. Int. Ed.* **2012**, *51*, 5876–5879. [[CrossRef](#)] [[PubMed](#)]
78. Akiyama, B.M.; Loper, J.; Najarro, K.; Stone, M.D. The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. *RNA* **2012**, *18*, 653–660. [[CrossRef](#)]
79. Berman, A.J.; Akiyama, B.M.; Stone, M.D.; Cech, T.R. The RNA accordion model for template positioning by telomerase RNA during telomeric DNA synthesis. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1371–1375. [[CrossRef](#)]
80. Wallweber, G.; Gryaznov, S.; Pongracz, K.; Pruzan, R. Interaction of human telomerase with its primer substrate. *Biochemistry* **2003**, *42*, 589–600. [[CrossRef](#)]
81. Parks, J.W.; Stone, M.D. Coordinated DNA dynamics during the human telomerase catalytic cycle. *Nat. Commun.* **2014**, *5*. [[CrossRef](#)] [[PubMed](#)]
82. Chen, L.Y.; Redon, S.; Lingner, J. The human CST complex is a terminator of telomerase activity. *Nature* **2012**, *488*, 540–544. [[CrossRef](#)] [[PubMed](#)]

83. Zhang, M.; Wang, B.; Li, T.; Liu, R.; Xiao, Y.; Geng, X.; Li, G.; Liu, Q.; Price, C.M.; Liu, Y.; et al. Mammalian CST averts replication failure by preventing G-quadruplex accumulation. *Nucleic Acids Res.* **2019**, *47*, 5243–5259. [[CrossRef](#)] [[PubMed](#)]
84. Giraud-Panis, M.-J.; Teresa Teixeira, M.; Gé li, V.; Gilson, E. Molecular Cell Review CST Meets Shelterin to Keep Telomeres in Check. *Mol. Cell* **2010**, *39*, 665–676. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).