G-Quadruplexes in Human Ribosomal RNA

Santi Mestre-Fos[1,2], Petar I. Penev[1,3], Suttipong Suttapitugsakul[2], Michael Hu[1,2], Chieri Ito[1,2], Anton S. Petrov[1,2], Roger M. Wartell[1,2,3], Ronghu Wu[2] and Loren Dean Williams[1,2,3]

1 - Center for the Origin of Life, *Georgia Institute of Technology, Atlanta, GA 30332-0400, USA*
2 - School of Chemistry and Biochemistry, *Georgia Institute of Technology, Atlanta, GA 30332-0400, USA*
3 - School of Biological Sciences, *Georgia Institute of Technology, Atlanta, GA 30332-0400, USA*

Correspondence to Loren Dean Williams: *Center for the Origin of Life, Georgia Institute of Technology, Atlanta, GA 30332-0400, USA. loren.williams@chemistry.gatech.edu*
https://doi.org/10.1016/j.jmb.2019.03.010
Edited by Philip C Bevilacqua

Abstract

rRNA is the single most abundant polymer in most cells. Mammalian rRNAs are nearly twice as large as those of prokaryotes. Differences in rRNA size are due to expansion segments, which contain extended tentacles in metazoans. Here we show that the terminus of an rRNA tentacle of *Homo sapiens* contains 10 tandem G-tracts that form highly stable G-quadruplexes *in vitro*. We characterized rRNA of the *H. sapiens* large ribosomal subunit by computation, circular dichroism, UV melting, fluorescent probes, nuclease accessibility, electrophoretic mobility shifts, and blotting. We investigated Expansion Segment 7 (ES7), oligomers derived from ES7, intact 28S rRNA, 80S ribosomes, and polysomes. We used mass spectrometry to identify proteins that bind to rRNA G-quadruplexes in cell lysates. These proteins include helicases (DDX3, CNBP, DDX21, DDX17) and heterogeneous nuclear ribonucleoproteins. Finally, by multiple sequence alignments, we observe that G-quadruplex-forming sequences are a general feature of LSU rRNA of Chordata but not, as far as we can tell, of other species. Chordata ribosomes present polymorphic tentacles with the potential to switch between inter- and intramolecular G-quadruplexes. To our knowledge, G-quadruplexes have not been reported previously in ribosomes.

Introduction

Cytosolic ribosomes are the most abundant assemblies in any cell, containing over 80% of cellular RNA [1]. Ribosomes are built on a "common core" [2] of rRNA with universal structure and function in all extant species. Common core rRNA is approximated by prokaryotic rRNA; around 90% of prokaryotic rRNA is contained in the common core.

The LSU of eukaryotic ribosomes contains additional rRNA in a secondary shell that surrounds the common core. The rRNA of the eukaryotic shell is composed of expansion segments (ESs) that attach to common core rRNA at a handful of specific sites [3–6]. ESs are the most variable rRNA structures over phylogeny. In *Saccharomyces cerevisiae*, ESs are important in ribosome biogenesis [7] and chaperone association [8,9]. Knorr and coworkers [10] recently described a three-dimensional structure in which *S. cerevisiae* ESs associate with and localize N-terminal acetylases.

ESs of chordates contain "tentacles" (Fig. 1) [11]. These tentacles reach a zenith in primates and birds, extending for around a hundred Ångstroms from the ribosomal surface. Here we observe G-quadruplexes in rRNA tentacles of human ribosomes. Tandem G-tracts are found in the tentacles of ESs 7 and 27 of *Homo sapiens* ribosomes ($ES7_{HS}$ and $ES27_{HS}$). $ES7_{HS}$ contains 10 tandem G-tracts in *tentacle a* and 4 in *tentacle b*. $ES27_{HS}$ contains six tandem G-tracts in *tentacle a*, two sets of three in *tentacle b*, and four within a base helix (Table 1).

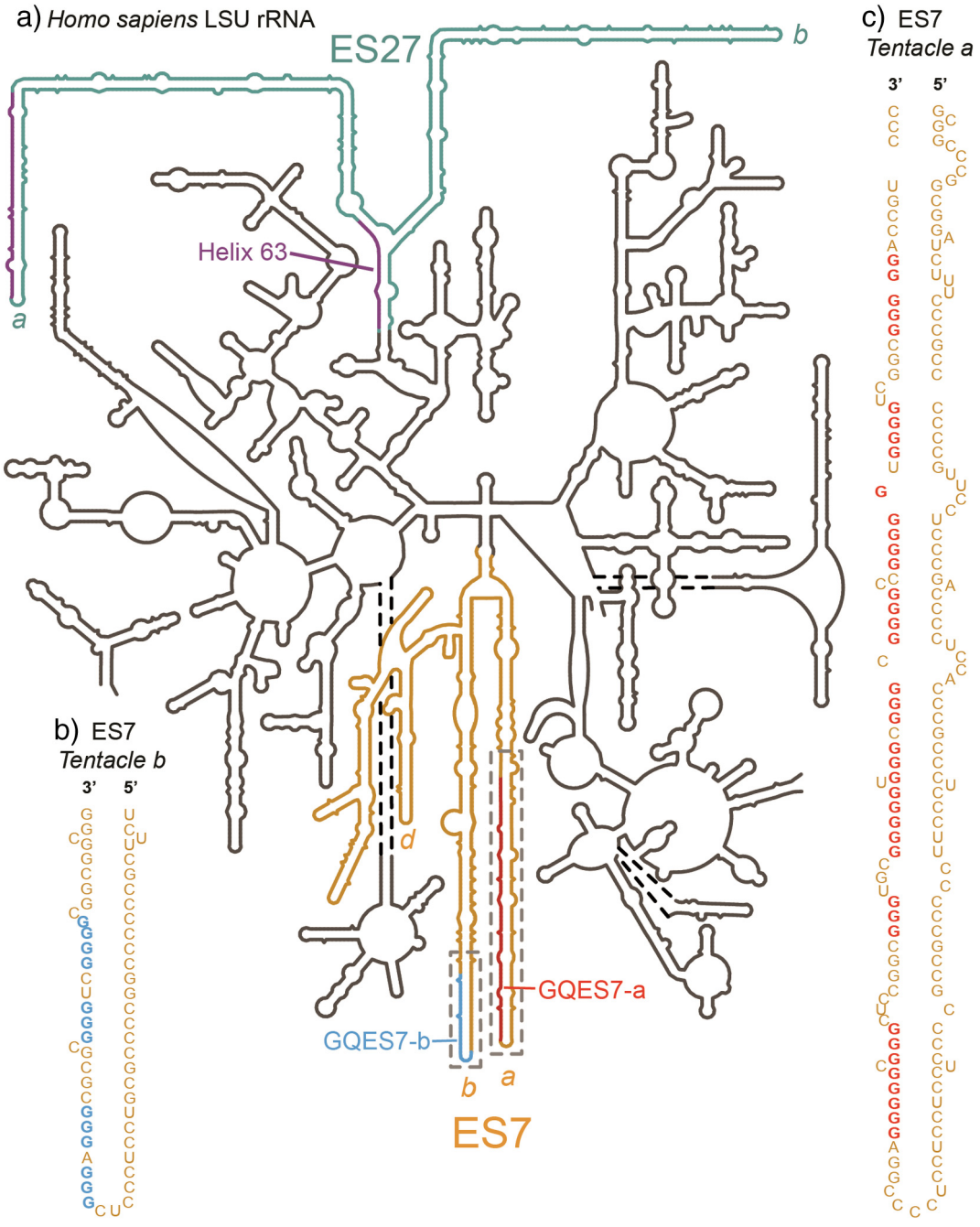

Fig. 1. Model of the secondary structure of the LSU rRNA of *H. sapiens*. G-quadruplex-forming regions (defined by $G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$) are highlighted. (a) Expansion segment $ES7_{HS}$ is orange. *Tentacles a*, *b*, and *d* of $ES7_{HS}$ are indicated. G-quadruplex-forming regions of $ES7_{HS}$ are GQES7-a (red, in *tentacle a*) and GQES7-b (cyan, in *tentacle b*) and are boxed by dashed lines. Expansion segment $ES27_{HS}$ is green with purple G-tracts. Helix 63, at the base of $ES27_{HS}$, contains a G-quadruplex motif (purple). *Tentacles a* and *b* of $ES27_{HS}$ are indicated. (b) An expanded view of GQES7-b indicates the nucleotide sequence. (c) An expanded view of GQES7-a indicates the nucleotide sequence.

G-quadruplexes are favored by tandem G-tracts separated by short non-specific sequences. To investigate the possibility that G-tracts in rRNA tentacles form G-quadruplexes, we used computation, circular dichroism, fluorescent probes, thermal melting, nuclease accessibility, electrophoretic mobility shift assays (EMSA), dot blotting, Western blotting, and pull-down assays combined with stable isotope labeling by amino acids in cell culture (SILAC) and mass spectrometry. To investigate

Table 1. G-quadruplex-forming regions within ES7 and ES27 rRNA

		Nucleotides
ES7		
Tentacle a		
Homo sapiens	**GGGGG**C**GGG**CUCCGGC**GGG**UGC**GGGGG**U**GGG**C**GGG**C**GGGG**CC**GGGGG**U**GGGG**UCGGC**GGGGG**	587–648
Pan troglodytes	**GGGGG**C**GGG**CUCCGGC**GGG**UGC**GGGGG**U**GGG**C**GGG**C**GGGG**CC**GGGGG**U**GGGG**UCGGC**GGGGG**	583–644
Mus musculus	**GGG**C**GGGG**CC**GGGGG**U**GGGG**UCGGC**GGGGG**	627–656
Gallus gallus	**GGGG**C**GGGG**C**GGG**CCCA**GGGGGGG**C**GGG**C**GGG**CC**GGGG**	557–594
Tentacle b		
Homo sapiens	**GGG**A**GGG**CGCGC**GGG**UC**GGGG**	829–849
Pan troglodytes	**GGG**A**GGG**CGCGC**GGG**UC**GGGG**	816–836
Tentacle d		
Mus musculus	**GGG**C**GGG**CGU**GGGGG**U**GGGGG**CC**GGG**	907–932
Gallus gallus	**GGGG**CGC**GGGGG**C**GGGGGGG**UC**GGG**	933–958
ES27		
Tentacle a		
Homo sapiens	**GGGGG**AGCGCCGCGU**GGGGG**CGGCGGC**GGGGGG**AGAA**GGG**UC**GGGG**CGGCA**GGGG**	3095–3149
Tentacle b		
Homo sapiens[a]	**GGGGG**C**GGGG**AGCGGUC**GGG**CGGCGGCGGUCGGC**GGG**CGGC**GGGG**C**GGGG**	3373–3422
Helix 63		
Homo sapiens	**GGG**CU**GGG**UCGGUC**GGG**CU**GGGG**	2896–2918

[a] This sequence falls outside the $G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$ motif.

phylogenetic distribution of ribosomal G-quadruplexes, we conducted multiple sequence alignments (MSAs) and database analysis.

The combined results indicate that G-quadruplexes form in oligomers composed of sequences derived from $ES7_{HS}$. In addition, G-quadruplexes form in intact $ES7_{HS}$ and in purified human 28S rRNA. We present data supporting formation of G-quadruplexes in 80S ribosomes and in polysomes. SILAC experiments show that known G-quadruplex-binding proteins associate with the G-tracts of $ES7_{HS}$. MSAs indicate that G-quadruplex-forming sequences are found in ES7s of all chordates.

Results

ES7 and ES27 of the human LSU contain G-quadruplex-forming sequences

The propensity of an RNA to form G-quadruplexes can be estimated from sequence—by lengths of guanine tracts and the lengths and compositions of loops regions. The program QGRS Mapper [12] provides "G-scores", which quantitate this propensity. We have identified ES7 and ES27 as the primary regions in the human LSU with sequences that appear to be capable of forming G-quadruplexes. The computational results suggest that G-quadruplexes can form near the termini of the longest rRNA tentacles of these two ESs. In *tentacles a* and *b* of $ES7_{HS}$, two regions, here named GQES7-a and GQES7-b (Fig. 1), meet the G-quadruplex consensus($G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$). The G-scores of GQES7-a and GQES7-b are in the range of well-established RNA G-quadruplexes.

The sequence 5′ **GGGG**CC**GGGGG**U**GGG**-**G**UCGGC**GGGG** 3′ (nts 623–647, from within GQES7-a, Fig. 1, Table 1) gives a G-score of 60. The sequence 5′ **GGG**UGC**GGG**GGU**GGG**C**GGG** 3′ (nts 603–621, also within GQES7-a) gives a G-score of 40. The sequence 5′ **GGG**A**GGG**CGCGC**GGG**UC**GGGG** 3′ (nts 829–849 within GQES7-b, Fig. 1, Table 1) gives a G-score of 38. Differences in the number of tandem G-tracts (10 in GQES7-a and 4 in GQES7-b), the lengths of the G-tracts, and G-scores suggest more stable and more extensive G-quadruplex formation in GQES7-a than in GQES7-b. The G-tracts of GQES7-a are longer than those of GQES7-b, and there are more of them. A greater propensity of GQES7-a over GQES7-b for G-quadruplex formation is seen in all experiments below.

As a positive control for both computation and experiment, we used the G-quadruplex from the 5′-UTR of the mRNA of the ADAM10 metalloprotease [13]. This stable and well-characterized RNA G-quadruplex gives a G-score of 42. As negative controls, we used two mutant RNA oligomers (*mt*ES7-a and *mt*ES7-b) that are analogous to GQES7-a and GQES7-b in composition and length, with disrupted G-tracts (Table S.1). Neither gives a G-score. In several experiments, we used yeast-$tRNA^{phe}$ as an additional negative control.

We focused our experiments primarily on $ES7_{HS}$. However, the end of *tentacle a* of $ES27_{HS}$ contains the sequence 5′ **GGGG**AGAA**GGG**UC**GGGG**CGG-CA**GGG** 3′ (nts 3124–3148, *tentacle a*), which gives a G-score of 40 (Fig. 1, Table 1). $ES27_{HS}$ also

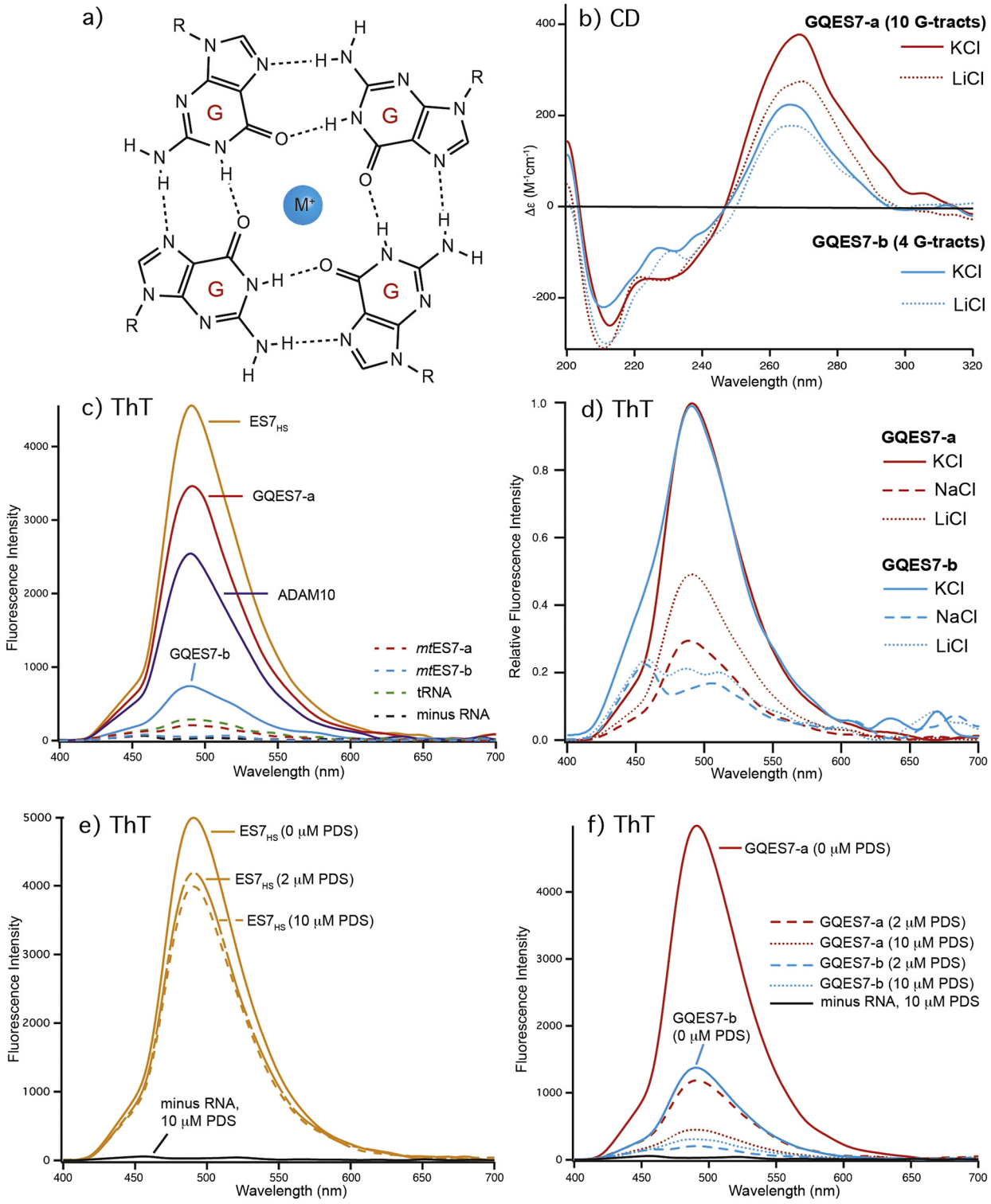

Fig. 2. Formation of G-quadruplexes by rRNA fragments GQES7-a, GQES7-b, and $ES7_{HS}$. (a) G-quadruplexes preferentially coordinate K^+. (b) CD spectra of GQES7-a and GQES7-b in the presence of either K^+ or Li^+. (c) Fluorescence emission at 487 nm of the G-quadruplex probe ThT in the presence of $ES7_{HS}$, GQES7-a, GQES7-b, or positive control ADAM10. Negative controls (dashed) are tRNA, *mt*ES7-a, *mt*ES7-b, and minus RNA. (d) ThT fluorescence emission of GQES7-a and GQES7-b in the presence of various monovalent cations. Intensities of GQES7-a and GQES7-b are normalized in the presence of K^+ to highlight cation-induced differences. (e) PDS competes with ThT in association with $ES7_{HS}$. (f) PDS competes with ThT in association with GQES7-a and GQES7-b.

contains a G-quadruplex-forming region within helix 63, near the junction of *tentacles a* and *b*. Based on the high G-scores and our experimental observation of G-quadruplexes within $ES7_{HS}$, we expect G-quadruplexes to form in $ES27_{HS}$.

Circular dichroism

CD spectra of GQES7-a and GQES7-b indicate G-quadruplex formation, with the expected dependence on type of counterion. CD is used widely to study RNA and DNA G-quadruplexes [14–18]. CD spectra of GQES7-a and GQES7-b (Fig. 2b) show the characteristic peak at 260 nm and trough at 240 nm. It is known that G-quadruplex formation is promoted by K^+ and is inhibited by Li^+ or Na^+ [19]. The intensities of the 260-nm peaks of both GQES7-a and GQES7-b are attenuated when the monovalent cation is switched from K^+ to Li^+. GQES7-a gives a more intense CD signal than GQES7-b under all conditions.

Thioflavin T fluorescence

Thioflavin T (ThT) is known to yield intense fluorescence at 487 nm upon association with G-quadruplexes [20,21]. ThT fluorescence results here suggest formation of G-quadruplexes in GQES7-a, GQES7-b, and intact $ES7_{HS}$ (Fig. 2c). Intact $ES7_{HS}$ and GQES7-a give more intense ThT fluorescence signals than the positive control (ADAM10). The GQES7-b signal is less than that of ADAM10 but is significantly greater than the negative controls. ThT fluorescence of GQES7-a and GQES7-b is attenuated when the monovalent counterion is switched from K^+ to Li^+. (Fig. 2d). Consistent with results of QGRS Mapper and CD spectroscopy, the ThT-induced fluorescence signal for GQES7-b is less than that of GQES7-a under all conditions.

The formation of G-quadruplexes by GQES7-a, GQES7-b, and intact $ES7_{HS}$ is supported by competition assays with pyridostatin (PDS) (Fig. 2e and f). PDS is a G-quadruplex stabilizer and a ThT competitor with a greater affinity than ThT for G-quadruplexes [22]. As expected if G-quadruplexes form in these rRNAs, PDS displaces ThT. In this series of experiments, *mt*ES7-a, *mt*ES7-b, and tRNA were used as negative controls, giving signals near background.

UV thermal melting

The melting of G-quadruplexes is distinguishable from melting of other RNA secondary structures. Melting of G-quadruplexes, but not other RNA structures, is accompanied by *hypochromicity* at 295 nm [23]. Melting of G-quadruplexes, but not other RNA structures, shows an acute dependence on type of monovalent counterion. The T_m's of G-quadruplex melting are expected to be greater in K^+ than in Li^+.

The UV melting profile of GQES7-b demonstrates the characteristic hypochromic shift at 295 mm and the expected salt dependence (Fig. 3a). The T_m of the melting transition is increased by around 9 °C when the counterion is switched from Li^+ ($T_m \cong 36$ °C) to K^+ (~45 °C). We were unable to observe a melting transition under any conditions for GQES7-a. Even under low concentrations of Li^+, the T_m of GQES7-a appears to be greater than experimentally accessible temperatures. The observed differences between melting behaviors of GQES7-a and GQES7-b are consistent with the greater G-quadruplex propensity of GQES7-a than of GQES7-b observed by QGRS Mapper, CD spectroscopy, and ThT fluorescence.

MBN cleavage

In the conventional secondary model of *H. sapiens* LSU rRNA (Fig. 1), G-quadruplex-forming sequences are represented as doubled-stranded and are paired with C-rich strands. If these tentacles form G-quadruplexes, the C-rich strands would be dissociated, presumably as single strands. As an additional test for G-quadruplexes in intact $ES7_{HS}$ rRNA, we examined cleavage by MBN (Fig. 3b). MBN preferentially cleaves single-stranded RNA or DNA and would cleave $ES7_{HS}$ rRNA more rapidly if G-quadruplexes form than if they do not. MBN cleavage has been used previously to test for G-quadruplexes in DNA [24]. The results here show that MBN cleaves $ES7_{HS}$ most rapidly under G-quadruplex-stabilizing conditions (Fig. 3b). Addition of K^+ increases the extent of cleavage (at constant time). Addition of PDS to K^+ further increases the extent of cleavage. The simplest interpretation of the MBN results is that $ES7_{HS}$ exists as a mixture of duplex and G-quadruplex forms and that the equilibrium is shifted by the type of counterion and by G-quadruplex stabilizers. In a negative control, extent of MBN hydrolysis of tRNA did not increase upon the addition of K^+ and/or PDS (Fig. S.2).

Antibody binding

BG4 is an antibody developed by Biffi and coworkers [25,26] that binds to a variety of G-quadruplex types but not to other nucleic acids such as RNA hairpins, or single-stranded or double-stranded DNA. Here, to test for G-quadruplex formation in GQES7-a, an EMSA was performed with BG4 (Fig. 3c). BG4 was also used for dot blotting experiments with GQES7-a, GQES7-b, and intact $ES7_{HS}$ (Fig. 4). We observe binding of BG4 to GQES7-a, GQES7-b, and intact $ES7_{HS}$ (Fig. 4a–b). Consistent with the results above, BG4 binds more tightly to GQES7-a than to GQES7-b.

The experiments presented above are consistent with *in vitro* formation of G-quadruplexes by $ES7_{HS}$ and by oligomers derived from $ES7_{HS}$. Below we

investigate whether intact *H. sapiens* 28S rRNA can form G-quadruplexes when protein-free or when assembled in ribosomes. The 28S rRNA was extracted from HEK293T cells and dot blotting was performed with BG4 (Fig. 4d). The results suggest that 28S rRNA forms G-quadruplexes.

To determine if 28S rRNA from *H. sapiens* forms G-quadruplexes when assembled in intact ribosomes, dot blotting was also performed with purified 80S human ribosomes and with polysomes (Fig. 4e and f). The results show that the BG4 antibody binds preferentially to intact human ribosomes and polysomes in a concentration-dependent manner. PDS enhances binding of the antibody, as expected for G-quadruplex formation. The observation of more extensive binding of the antibody to polysomes than to monomer ribosomes suggests formation of intermolecular G-quadruplexes in polysomes (i.e., that G-quadruplexes link tentacles of adjacent ribosomes).

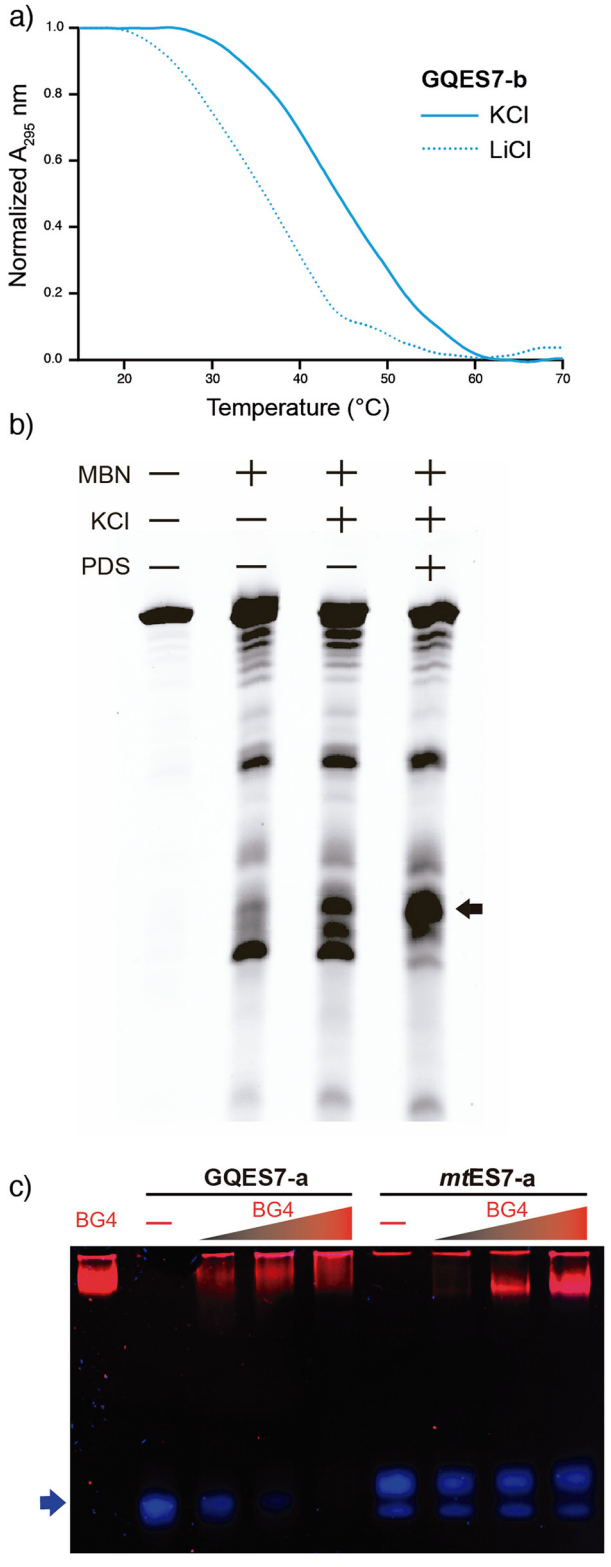

Fig. 3. (a) UV thermal melting profile of GQES7-b at 295 nm. Before melting RNA was annealed in the presence of either 100 mM KCl or 100 mM LiCl. (b) $ES7_{HS}$ cleavage by mung bean nuclease (MBN). ES7 was annealed with or without KCl and with or without PDS. The black arrow indicates cleaved rRNA. (c) EMSA of the BG4 antibody with GQES7-a and its non-G-quadruplex-forming mutant *mt*ES7-a, visualized on a native gel. GQES7-a and *mt*ES7-a RNAs were loaded at a constant strand concentration with increasing concentrations of BG4 antibody. The RNA (arrow) is blue and the protein is red.

G-quadruplex sequences in ribosomes throughout Chordata

Focusing specifically on translation, we have developed the SEREB Database [2], which contains fully curated and cross-validated sequences of rRNAs from all major phyla, yet samples the tree of life in a sparse, efficient, and unbiased manner. Here we extended the SEREB database, increasing the number of chordate species from 10 to 17, for a fine-grained analysis of ES7.

G-quadruplex-forming sequences in chordate ES7s

Our MSA confirms that the lengths of rRNA tentacles of eukaryotes are variable, reaching maxima in species such as *Gallus gallus* and *H. sapiens* (Fig. 5). Aligned sequences of relevant segments of ES7s of various eukaryotes demonstrate G-quadruplex-forming sequences in chordates, indicated by the motif $G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$ [number of G tracts (n) ≥ 3]. The motif is observed near the termini of ES7 tentacles in all warm-blooded chordates, although the exact locations and specific sequences are variable. The maximum number of tandem G-tracts in *tentacle a* of ES7 is 10 in human and chimpanzee and eight in rat and chicken. Fish, reptiles, and amphibians appear to lack the G-tract motif in ES7 tentacles.

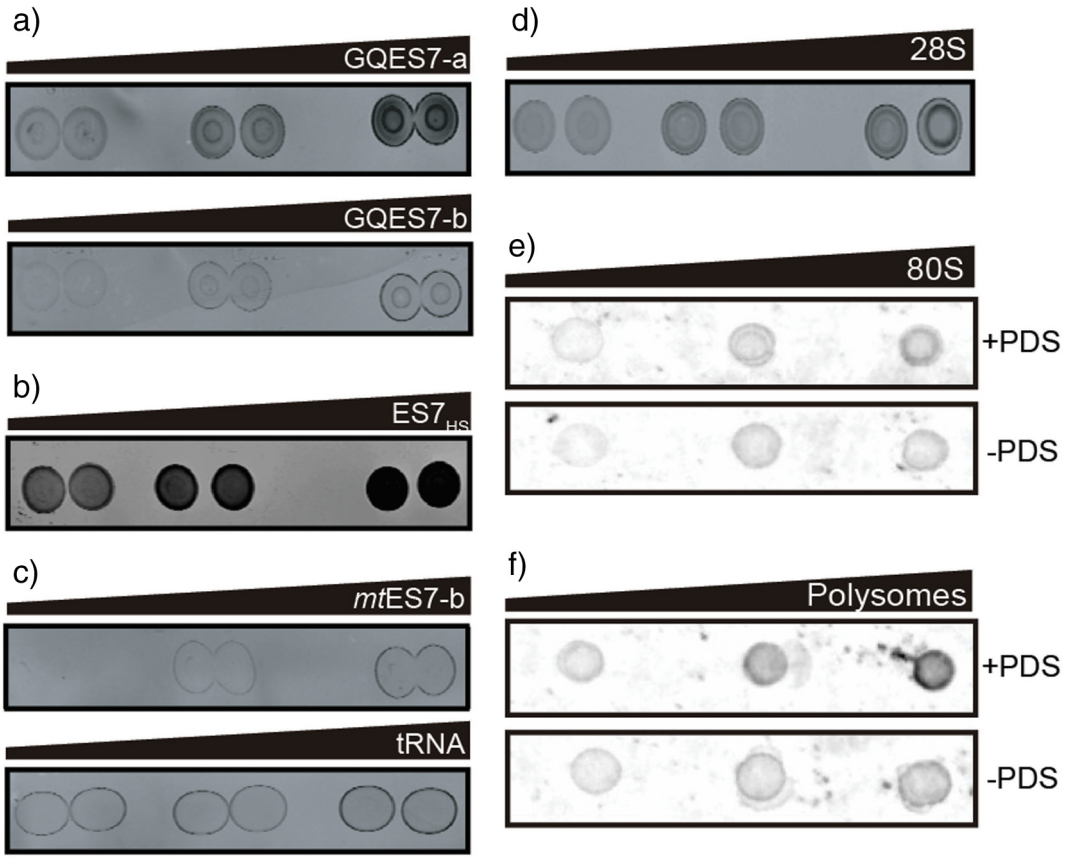

Fig. 4. Dot blots performed with the BG4 antibody on (a) GQES7-a and GQES7-b, (b) intact $ES7_{HS}$, (c) the negative controls *mt*ES7-b and tRNA, (d) the 28S rRNA extracted from HEK293T cells, and (e) human 80S ribosomes and (f) polysomes purified from HEK293 cells. All samples were incubated in the presence of 50 mM KCl, and ribosomes and polysomes were further analyzed with or without 10 μM PDS, which stabilizes G-quadruplexes. Samples were loaded onto the membrane in increasing amounts from left to right.

G-quadruplex-forming sequences in chordate ESs other than ES7

The extended SEREB Sequence Database suggests that G-quadruplex-forming sequences are universal to chordates (Table S.2). Several chordate species present G-quadruplex-forming sequences in tentacles other than ES7 (Table S.2). These G-quadruplex-forming sequences are not shown in Fig. 5. In addition, repeated G-tracts outside of the motif can form G-quadruplexes [27,28]. Therefore, additional G-quadruplexes cannot be excluded in these ribosomes. It is possible that G-tracts with $n < 4$ form intermolecular G-quadruplexes with other tentacles or with other ribosomes as in polysomes.

Absence of G-quadruplex sequences in non-chordate rRNAs

To determine the phylogenic distribution of G-quadruplexes in LSU rRNA, we inspected highly curated sequences of 20 non-chordate eukaryotes from the SEREB database [2]. Thus far, we can find no evidence of G-quadruplex-forming sequences in ribosomes of non-chordate eukaryotes.

RNA remodeling proteins bind to rRNA G-quadruplex sequences

The localization of G-quadruplex-forming sequences to ribosomal tentacles suggests the possibility of interaction with non-ribosomal proteins. To identify the proteins that bind to rRNA G-quadruplexes, we performed pull-down experiments and used stable isotope labeling by amino acids in cell culture (SILAC) for protein quantification. We focused on GQES7-a, the longest and most stable G-quadruplex-forming region in human rRNA (Fig. 6). GQES7-a rRNA was linked on the 3′ end to biotin (GQES7-a-Biotin), and associated proteins in human cell lysates were pulled down and analyzed by mass spectrometry. The biotinylation of GQES7-a does not disrupt the G-quadruplexes (Fig. S.3).

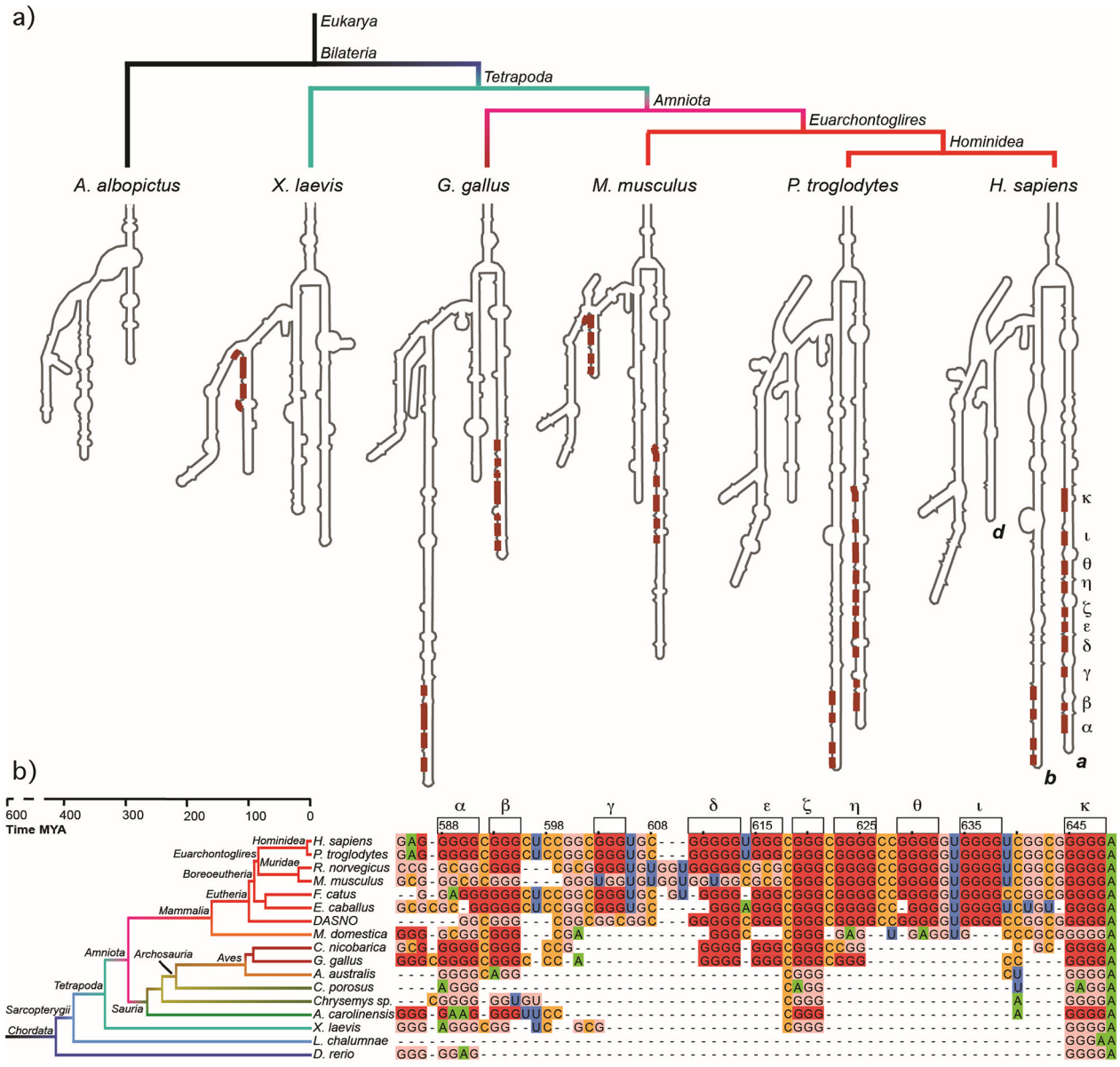

Fig. 5. G-tracts are observed in ES7 tentacles. (a) Conventional secondary structural models of ES7 from various eukaryotes. G-tracts within the $G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$ motif are highlighted in red. (b) Sequence alignment of ES7 *tentacle a* showing G-quadruplex-forming sequences is common in chordates. Individual G-tracts in both panels are labeled with Greek symbols. Nucleotides are colored by type. G's within G-tracts are dark red. Other G's are pink. All nucleotides are numbered in accordance with *H. sapiens* 28S rRNA. Sizes of eukaryotic ES7 secondary structures are not to scale. Complete species nomenclature is provided in Table S.3.

Known G-quadruplex-binding proteins were pulled down by this assay including CNBP, YBOX1, hnRP F, hnRP H, DDX21, DDX17, and DDX3 [29–35]. A significant number of helicases were identified (DDX3, CNBP, DDX21, DDX17), all of which have been reported to unfold G-quadruplexes. In addition, a significant number of heterogeneous nuclear ribonucleoproteins (hnRNPs) were shown in this experiment to bind to GQES7-a, including hnRNP G-T/RMXL2, hnRNP M, hnRNP G/RBMX, hnRNP H2, hnRNP H, hnRNP F, hnRNP H3, and FUS. hnRNPs are a family of RNA-binding proteins with functions including pre-mRNA processing and transport of mRNAs to ribosomes [36]. Several of these proteins have been previously identified as ribosome-binding proteins [37].

To support results of the pull-down experiments, Western blotting was performed with four of the

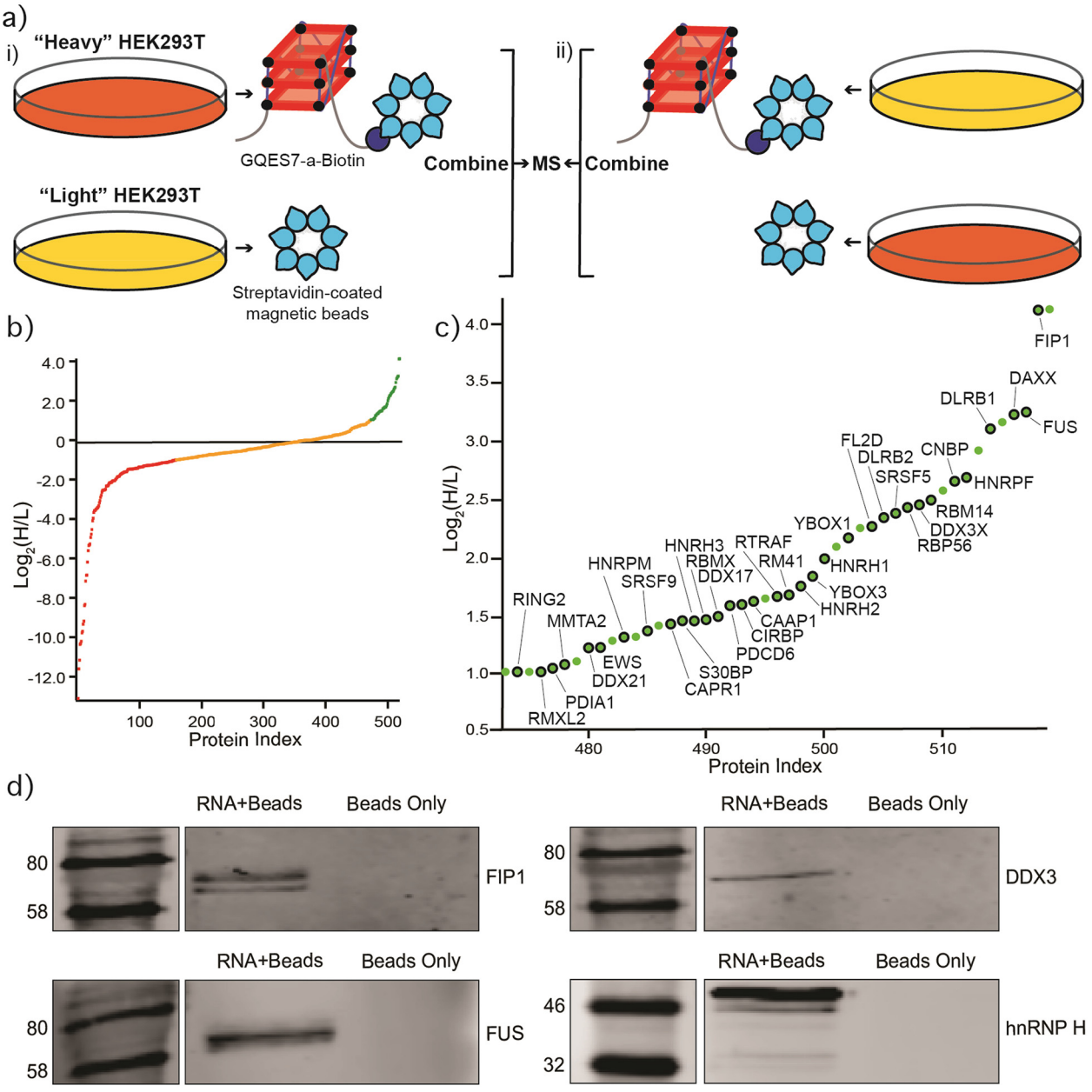

Fig. 6. Identification of GQES7-a-binding proteins. (a) Scheme of the SILAC experiment. The "RNA + beads" samples were combined in HEK293T grown in heavy media (panel i). The "Beads Only" control sample was combined in HEK293T grown in light media. To verify the proteins identified by this method, the experiment was performed using reverse labeling (panel ii). (b) Scatter plot representing fold enrichment of the proteins binding to GQES7-a in "Heavy" HEK293T. Color representation indicates specific proteins that bound more tightly to GQES7-a than to the beads (green) or to the beads than to GQES7-a (red), or bound to the beads and GQES7-a to a similar extent (orange). (c) A close-up view of the green region of the scatter plot represented in panel b. Dots with a black contour are used to indicate proteins that appeared in the green region of the two replicate experiments described in panel a. (d) Western blotting analyses of the eluted proteins from the RNA pull-down of HEK293T. All four blotted proteins (FIP1, FUS, DDX3, and hnRNP H) eluted from the GQES7-a sample (RNA + Beads) but not from the control (Beads Only), confirming the SILAC results.

proteins obtained in the pull-down experiments (Fig. 6d). We assayed a DEAD-box RNA helicase (DDX3), a heterogeneous nuclear ribonucleoprotein (hnRNP H), the RNA-binding protein FUS, and a pre-mRNA polyadenylation stimulator (FIP1). hnRNP H and DDX3 have been previously identified as G-quadruplex-binding proteins. All four proteins bind to GQES7-a in the Western blot, suggesting that we have tapped an uncharacterized pool of G-quadruplex-binding proteins.

Discussion

The results presented here suggest that G-quadruplexes are far more profuse than previously conceived. rRNA is the most abundant macromolecule in most cells. Human LSU rRNA and rRNA of other chordates contain sequences with strong propensity to form G-quadruplexes (Fig. 1). We have identified 10 tandem G-tracts in *tentacle a* of human ES7 rRNA and four in *tentacle b*. These tandem G-tracts form stable G-quadruplexes under a variety of conditions *in vitro*. Computation, ThT fluorescence, CD spectroscopy, UV melting, EMSAs, nuclease digestion, and blotting with a G-quadruplex antibody provide a consistent picture of the propensities of various regions of 28S rRNA to form G-quadruplexes.

G-quadruplex-forming sequences have been shown previously to cluster within regulatory mRNA regions such as 5′ and 3′ untranslated regions [38] and within the first intron [39]. The extent to which such G-quadruplexes form *in vivo* remains uncertain. Guo and Bartel [40] suggest that mRNA G-quadruplexes are globally unfolded by unwinding factors in eukaryotic cells. By contrast, Yang *et al*. [41] support a model in which G-quadruplexes continuously form and unfold *in vivo*. However, rRNAs were explicitly excluded from both of these investigations; the extent of G-quadruplex formation *in vivo* remains an open question.

In our view, the following merit a ribosome centric reinvestigation of G-quadruplex structure and function *in vivo:*

i. inherent flexibility and polymorphism of rRNA tentacles,
ii. their ability to extend hundreds of Ångströms from the ribosomal surface,
iii. the large number of tandem G-tracts on some rRNA tentacles,
iv. high stability of tentacle G-quadruplexes *in vitro*, and
v. extreme concentrations of rRNA on the rough ER and in polysomes.

Our results suggest that nature's most complex organisms have evolved long rRNA tentacles with unexpected structural polymorphism, including the ability to form G-quadruplexes (Fig. 7).

G-quadruplex-forming rRNA sequences appear to be a universal feature of tentacles of chordate ribosomes. We have inferred the locations of tandem rRNA G-tracts in various species by MSAs. The specific sequences and exact locations of the G-quadruplexes on tentacles are variable across phylogeny. We searched the SEREB database and thus far could find no evidence of G-quadruplex-forming sequences outside of the Chordata phylum. The SEREB database is specifically designed for analysis of rRNA and includes species from all major phyla and samples the tree of life in a sparse, efficient, and accurate manner [2]. It contains complete and highly curated rRNA sequences.

The preferential localization of rRNA G-quadruplexes near the termini of rRNA tentacles suggests that these regions are loci for association of specific cytosolic proteins or for assembly of ribosomes. Here we identified multiple human RNA helicases and other RNA remodeling proteins that bind to rRNA G-quadruplexes (Fig. 6). These proteins could be participants in G-quadruplex regulation on ribosomes.

Our observation here that polysomes appear to form more extensive G-quadruplexes than monomer ribosomes suggests a role for inter-ribosomal G-quadruplexes in closely associated ribosomes (Fig. 7b). Our work points to the possibility that, inside cells, ribosomes present polymorphic tentacles that can switch between unimolecular and multimolecular G-quadruplexes and duplex forms. In this model, surfaces of ribosomes contain multivalent docking sites for G-quadruplex-specific proteins and for nucleic acid assemblies, including in polyribosomes. It has been shown that G-quadruplexes can form phase separated RNA gels [42]. It is conceivable that densely packed ribosomes on the rough ER and in polysomes are surrounded by phase-separated G-quadruplexes gels composed of rRNA tentacles.

G-quadruplex-forming sequences have been described in genes encoding rRNA, where they are proposed to influence transcription [43] and bind to the nucleolar protein nucleophosmin [44]. These studies have focused on the external and internal transcribed regions (ETS and ITS) and are not part of the assembled ribosome. Kim and co-workers [45] reported tetramerization of an oligomer containing a single G-tract, derived from *Escherichia coli* 5S rRNA. The results here highlight potential ribosomal functionality associated with the large rRNA tentacles of mammals and birds. The conservation of rRNA G-quadruplex-forming sequences throughout Chordata, albeit at various locations in rRNA tentacles, suggests significant functions.

Methods

$ES7_{HS}$ spans nucleotides 436 to 1311 of the *H. sapiens* LSU rRNA and contains sequences that we call GQES7-a (nts 583–652) and GQES7-b (nts 825–853). RNAs corresponding to $ES7_{HS}$, GQES7-a, and GQES7-b were synthesized *in vitro* by transcription (HiScribe™ T7 High Yield RNA Synthesis Kit; New England Biolabs). For cation-dependent experiments, precautions were used to remove contaminating cations that stabilize G-quadruplexes. RNAs were ethanol precipitated from 800 mM LiCl and 10 mM Tris–HCl (pH 7.5). The RNA pellets were resuspended in 1 mM LiCl and 10 mM Tris–HCl (pH 7.5) and were dialyzed

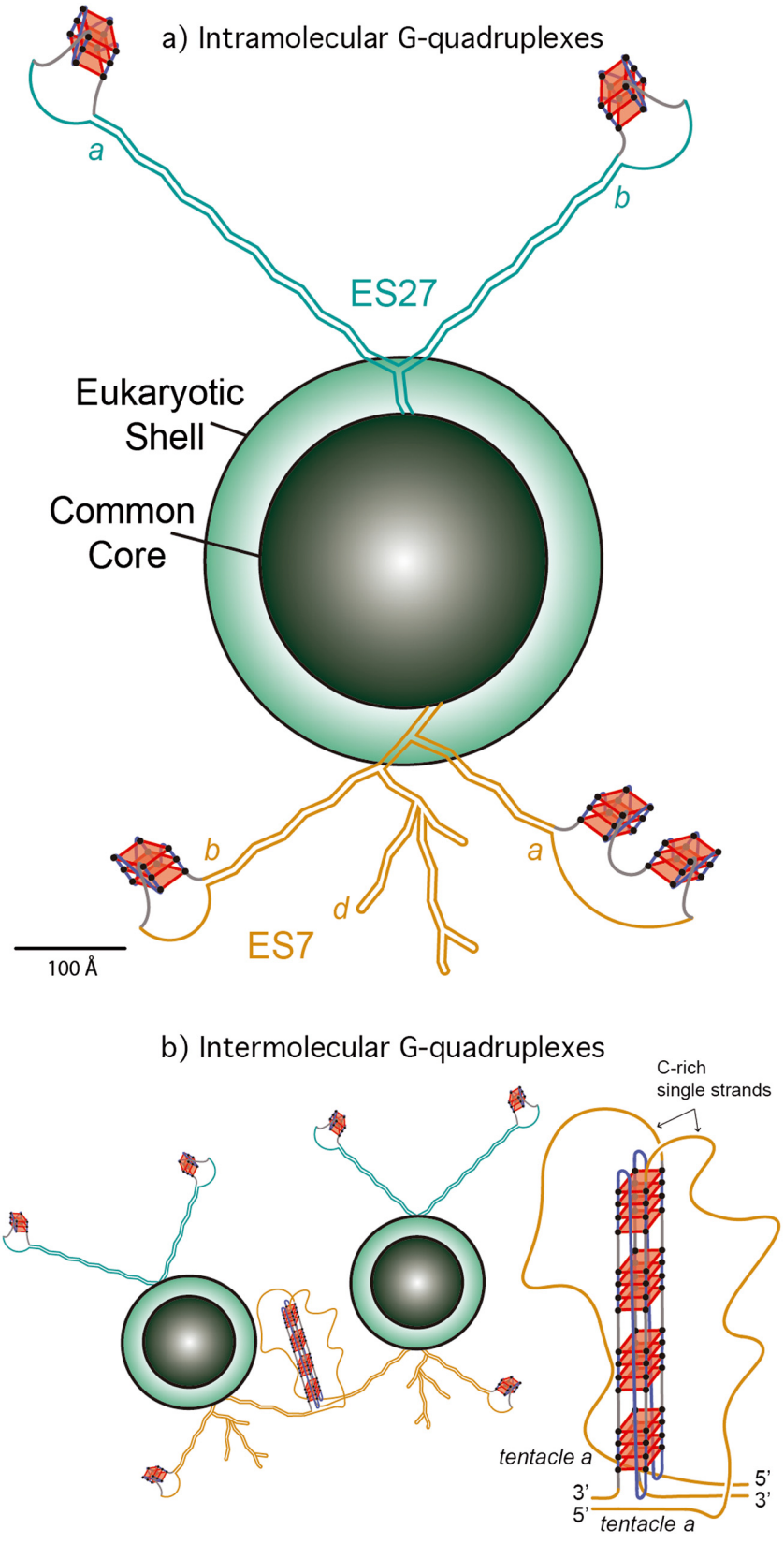

Fig. 7. (a) Schematic representation of the common core, the eukaryotic shell, and the tentacles of the LSU of the *H. sapiens* ribosome. G-quadruplexes are indicated on ES7 and ES27. The lengths of $ES7_{HS}$ (orange) and $ES27_{HS}$ (green) tentacles are roughly scaled to the size of the common core. The G-quadruplexes represented in *tentacle b* of $ES27_{HS}$ do not fall within the $G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$ motif and are speculative. The G-quadruplex region found in helix 63 of $ES27_{HS}$ is not indicated here. (b) Schematic representation of interactions between ribosomes *via* intermolecular G-quadruplexes. G-tracts on ES7 *tentacle a* from different ribosomes contribute to the formation of G-quadruplexes (an expanded view is presented on the right).

extensively against the same buffer using Slide-a-Lyzer dialysis cassettes (MWCO 3500; Pierce) at 4 °C. *mt*ES7-a and *mt*ES7-b were ordered as RNA oligomers. Baker's yeast tRNAs were purchased from Roche. RNA purity was monitored by 8 M Urea 5% acrylamide gel in TBE buffer. Complete sequences of $ES7_{HS}$, GQES7-a, GQES7-b, *mt*ES7-a, and *mt*ES7-b are contained in Table S.1.

HEK293T 28S rRNA extraction and purification

HEK293T cells were grown to 60% confluency after which total RNA was extracted with TRI Reagent® (Sigma-Aldrich). 28S rRNA was extracted from an agarose gel by running the rRNA into wells in the center of the gel, where the rRNA was extracted with a pipette. The rRNA was precipitated in 5 M ammonium acetate-acetic acid (pH 7.5) with excess ethanol. 28S rRNA purity was monitored on 1% agarose gels (Fig. S.1).

ThT fluorescence

RNAs were prepared at a final concentration of 1 μM (strand) and annealed in 150 mM KCl, NaCl or LiCl, 10 mM Tris–HCl (pH 7.5), and 2 μM ThT by cooling from 90 °C to 25 °C at 1 °C/min. RNAs were incubated at 4 °C for 10 min and were loaded onto a Corning® 384 Well Flat Clear Bottom Microplate. Fluorescence from 300 to 700 nm and exciting at 440 nm were acquired on a BioTek Synergy™ H4 Hybrid plate reader. When appropriate, PDS was added to the desired concentration after the RNA was annealed.

Circular dichroism

RNA at 1 μM (strand) in 150 mM KCl or LiCl and 10 mM Tris–HCl (pH 7.5) was annealed as described above. CD spectra were acquired at 20 °C on a Jasco J-810 spectropolarimeter using 1-mm cuvettes. Data from 200 to 320 nm were acquired at a rate of 100 nm/min with 1-s response, a bandwidth of 5 nm, and averaged over three measurements. The buffer spectrum was subtracted. Smoothing was performed with Igor Pro. The observed ellipticity (θ, mdeg) was normalized [46] using the expression $\Delta\varepsilon = \theta/(32{,}980 \times c \times l)$, where c is the molar strand concentration of the RNA and l is the path length of the cuvette in centimeters.

UV thermal melting

Absorbance measurements were collected at 295 nm using a Varian Cary-1E UV spectrophotometer. RNA samples (800 μL, final OD_{260} of 0.50 units) in 10 mM Tris–HCl (pH 7.5) and 100 mM KCl or LiCl were annealed as described above and added to 1 cm path-length quartz cuvettes. Samples were then heated from 15 °C to 90 °C and cooled at the same rate at 0.5 °C/min. Data was recorded every 0.5 °C.

EMSAs

The anti-G-quadruplex BG4 antibody was purchased from Absolute Antibody (catalog no. Ab00174-1.1). GQES7-a (3 μM) rRNA or the negative control *mt*ES7-a RNA was annealed in 20 mM Hepes–Tris (pH 7.5) and 50 mM KCl. GQES7-a rRNA or *mt*ES7-a RNA was combined with various concentrations of BG4 at a final RNA concentration of 1 μM RNA (strand). RNA–protein mixtures were incubated at room temperature for 20 min in 50 mM KCl. RNA–protein interactions were analyzed by 5% native-PAGE. Gels were visualized following a dual fluorescent dye protocol [47] with an Azure imager c400 (Azure Biosystems).

rRNA–BG4 antibody dot blotting

RNAs were annealed in the presence of 50 mM KCl and were diluted 1×, 2×, and 4×. GQES7-a, GQES7-b, *mt*ES7-b, and tRNA: 3.2, 1.6, and 0.8 μM. $ES7_{HS}$: 1.4, 0.7, and 0.35 μM. 28S rRNA: 55, 27.5, and 13.7 nM. RNAs were loaded onto nitrocellulose membranes and dried at room temperature for 30 min. The membranes were blocked for 1 h at room temperature using Odyssey Blocking Buffer (Li-COR). BG4 antibody was added (1:2000 dilution) and incubated with gentle rocking for 60 min at room temperature. The membrane was washed for 10 min twice with 1 × TBST and incubated for 60 min with an appropriate fluorescent secondary antibody anti-mouse (1:10,000 dilution) (Biotium, no. 20065-1). The membrane was washed for 10 min twice with 1 × TBST and was imaged on a Li-Cor Odyssey Blot Imager. Intact 80S ribosomes and polysomes were purified from HEK293, which were incubated 5 min in 10 μg/mL cycloheximide at 37 °C. Lysis buffer [10 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris–HCl (pH 7.5), 1% Triton X-100, 1% sodium deoxycholate, 0.2 U/mL DNase I, RNase inhibitor, 1 mM dithiothreitol, 10 μg/mL cycloheximide] was used to scrap the cells. Nuclei and cell debris were removed by centrifugation, and the supernatant was transferred to a 15%–50% sucrose gradient containing 100 mM NaCl, 10 mM $MgCl_2$, and 30 mM Tris–HCl (pH 7.5) and centrifuged by ultracentrifugation. Purified 80S ribosomes and polysomes were then incubated at room temperature for 15 min in the presence of 50 mM KCl with or without 10 μM PDS. Ribosomes or polysomes were added iteratively in 30-min intervals to the same site on a nitrocellulose membrane (0.9 μg, 2.7 μg, 4.5 μg). The membrane was then treated as described above. BG4 was added to a final dilution of 1:1000, and the secondary antibody was added to a final dilution of 1:5000.

MBN probing

$ES7_{HS}$ and tRNA were prepared at 100 ng/μL and annealed in the presence/absence of 100 mM KCl and 15 mM Tris–HCl (pH 7.5) by cooling from 90 °C to 25 °C at 1 °C/min. PDS was added to the annealed RNA to a final concentration of 2 μM. One unit of MBN was added per microgram of RNA, and samples were incubated at 30 °C for 30 min. SDS was added to a final concentration of 0.01% to denature the nuclease, and RNA was purified by ethanol precipitation. The extent of RNA cleavage was determined on an 8 M urea 5% acrylamide gel (19:1 acrylamide/bisacrylamide) stained with ethidium bromide.

ES7 secondary structures

*S*econdary structures of human and *Drosophila melanogaster* ES7 were obtained from RiboVision [48]**.** Nucleotides of G-quadruplex regions in *Pan troglodytes*, *Mus musculus*, and *G. gallus* (Table 1) were numbered as in Bernier [2], subtracting the nucleotides from the 5.8S rRNA.

Phylogeny and MSAs

The SEREB MSA [2] was used as a seed to align additional eukaryotic ES7 sequences to increase the density of eukaryotic species in the MSA. The 28S rRNA sequences in the SEREB MSA were used to search [49] the NCBI databases [50] for LSU rRNA sequences. The SEREB database has sequences from 10 chordate species; 7 additional chordate species from 7 new orders were added to the ES7 *tentacle a* MSA (Fig. 5 Table S.3). Sequences without intact ES7 *tentacle a* were excluded. Sequences with partial 28S rRNA were marked as partial. Sequences inferred from genomic scaffolds were marked as predicted (Table S.3). The extended database was queried for G-quadruplex-forming sequences.

Sequences were incorporated into the SEREB-seeded MSA using MAFFT [51] and adjusted manually using BioEdit [52]. Manual adjustments incorporated information from available secondary structures. In some cases, the positions of G-tracts in sequences with large gaps relative to *H. sapiens* are not fully determined, as they can be aligned equally well with flanking G-tracts in the MSA. Alignment visualization was done with Jalview [53]. The phylogenetic tree and the timeline of clade development were inferred from TimeTree [54].

Analysis of the entire LSU was performed on SEREB sequences, which are highly curated and always complete. This procedure ensured that negative results indicate absence of G-quadruplex-forming sequences from intact rRNAs rather than absence from rRNA fragments that lack the appropriate regions. G-quadruplex-forming sequences are not detected in any of the 20 non-chordate members of the SEREB database.

SILAC

HEK293T cells were cultured in SILAC media —"heavy" or "light" Dulbecco's modified Eagle media (Thermo Scientific) supplemented with 10% dialyzed fetal bovine serum (Corning) and 1% penicillin–streptomycin solution (Sigma) in a humidified incubator at 37 °C with 5% carbon dioxide. The heavy media contained 0.798 mM L-lysine ($^{13}C_6$ and $^{15}N_2$; Cambridge Isotope Laboratories) and 0.398 mM L-arginine ($^{13}C_6$; Cambridge Isotope Laboratories). The light media had the same concentrations of normal lysine and arginine (Sigma). Media were supplemented with 0.2 mg/mL proline (Sigma) to prevent arginine-to-proline conversion. Heavy and light cells were grown for at least six generations. Once the confluency reached 80%, cells were harvested by scraping, washed twice with ice-cold PBS (Sigma), lysed in a buffer containing 10 mM Hepes (pH 7.4), 200 mM KCl, 1% Triton X-100, 10 mM $MgCl_2$ (all from Sigma), and 1 pill/10 mL cOmplete ULTRA tablet protease inhibitor (Roche), and incubated on an end-over-end shaker at 4 °C for 1 h. Lysates were centrifuged at 25,830*g* at 4 °C for 10 min, and the supernatants were collected and kept on ice.

Ten micrograms of GQES7-a-Biotin RNA were annealed as described above in the presence of 10 mM Tris–HCl (pH 7.5) and 100 mM KCl. Twenty microliters of magnetic streptavidin-coated beads (GE Healthcare) was washed with the lysis buffer [10 mM Hepes (pH 7.4), 200 mM KCl, 1% Triton X-100, 10 mM $MgCl_2$, protease inhibitors]. Annealed RNA was then added to the washed beads and incubated at 4 °C for 30 min with gentle shaking. For control experiments, no RNA was added. SILAC cell lysates were incubated with 0.5 mg *E. coli* tRNA per 1 mg protein at 4 °C for 30 min with gentle shaking. "RNA + Beads" and control "Beads Only" samples were transferred into the SILAC cell lysates: "RNA + Beads" were added to the Heavy cell lysate, and "Beads Only" was added to "Light" HEK293T cell lysate. As a replicate, "RNA + Beads" was added to "Light" cell lysate, and "Beads Only" was added to "Heavy" cell lysate. RNasin (200 U/mL) was added, and the lysates were incubated at 4 °C for 2 h with gentle shaking. Samples were centrifuged, the supernatant was discarded, and the pelleted beads were washed with lysis buffer with increasing KCl concentrations (0.4 M, 0.8 M, 1.6 M). After the three washes, 100 μL of the elution buffer [100 mM Tris–HCl (pH 7.4), 1% SDS, 100 mM DTT] was added to one of the two samples and then combined with the beads of the corresponding sample. "RNA + Beads" in "Heavy" lysates were combined with "Beads Only"

in “Light” lysates, and “RNA + Beads” in “Light” lysates were combined with “Beads Only” in “Heavy” lysates. The combined samples were boiled and then briefly centrifuged. Beads were discarded, and samples were analyzed with an online LC–MS system.

Mass spectrometry

Eluted proteins were diluted 10 times with 50 mM Hepes (pH 7.4) and were alkylated with 28 mM iodoacetamide (Sigma) for 30 min at room temperature in the dark. Proteins were precipitated by methanol–chloroform, and the pellets were resuspended in digestion buffer containing 50 mM Hepes (pH 8.8), 1.6 M urea, and 5% acetonitrile (ACN) (all from Sigma). After digestion with sequencing-grade modified trypsin (Promega) at 37 °C for 16 h, reactions were quenched with 1% trifluoroacetic acid (Fisher Scientific) and purified with StageTip. Peptides were dissolved in 10 μL 5% ACN and 4% FA solution, and 1 μL was loaded to a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific) with a microcapillary column packed in-house with C18 beads (Magic C18AQ, 3 μm, 200 Å, 75 μmx16 cm). A 110-min gradient of 3%–22% ACN containing 0.125% FA was used. The peptides were detected with an LTQ Orbitrap Elite Hybrid Mass Spectrometer (Thermo Fisher Scientific) controlled by Xcalibur software (version 3.0.63). MS/MS analysis was performed with a data-dependent Top20 method. For each cycle, a full MS scan in the Orbitrap with the automatic gain control target of 10^6 and the resolution of 60,000 at 400 m/z was followed by up to 20 MS/MS scans in the ion trap for the most intense ions. Selected ions were excluded from further sequencing for 90 s. Ions with singly or unassigned charge were not sequenced. Maximum ion accumulation times were 1000 ms for each full MS scan and 50 ms for each MS/MS scan. The spectra were searched against a human protein database downloaded from UniProt using the SEQUEST algorithm (version 28) [55]. The following parameters were used: 20 ppm precursor mass tolerance; 1.0 Da fragment ion mass tolerance; trypsin digestion; maximum of three missed cleavages; differential modifications for methionine oxidation (+15.9949 Da), heavy lysine (+8.0142 Da), and heavy arginine (+6.0201 Da); and fixed modification for cysteine carbamidomethylation (+57.0215 Da). The false discovery rates (FDRs) were evaluated and controlled by the target-decoy method. Linear discriminant analysis was used to filter the peptides to <1% FDR based on parameters such as XCorr, ΔC_n, and precursor mass error. An additional filter was used to control the protein FDR to <1%. For SILAC quantification, the S/N ratios of both heavy and light peptides must be greater than 3. Otherwise, one of the two versions of the peptides (heavy or light) must have the S/N ratio greater than 10. Other peptides that did not pass these criteria were removed. The final protein ratio was calculated from the median value of the peptides from each parent protein. The raw files are publicly accessible at http://www.peptideatlas.org/PASS/PASS01260, Username: PASS01260, Password: TL3854zn.

Western blotting

Samples were electrophoresed on 12% SDS-PAGE gels and transferred to a nitrocellulose membrane overnight. Membranes were blocked for 1 h at room temperature with gentle shaking and then incubated for another hour with primary antibodies: 1:200 dilution of FIP1 (mouse monoclonal, sc-398,392), DDX3 (mouse monoclonal, sc-81,247), FUS (mouse monoclonal, sc-47,711), or hnRNP H (mouse monoclonal, sc-32,310). Antibodies were obtained from Santa Cruz Biotechnology. Membranes were washed three times with 1× TBST, and secondary antibody CF680 goat anti-mouse IgG (H + L) (Biotium, 20,065) was added (1:5000 dilution). Membranes were washed three times with 1× TBST and imaged on a Li-Cor Odyssey Blot Imager.

Acknowledgments

The authors thank Drs. Jonathan B. Chaires and Aaron Engelhart for helpful discussions, Dr. Lizzette M. Gómez Ramos for designing the negative G-quadruplex controls, and Ahmad M. Mohamed for the technical support on CD and UV spectroscopy. Purified 80S ribosomes and polysomes were a gift from Immagina BioTechnology. This work was supported by NASA (NNX16AJ28G, NNX16AJ29G, 80NSSC18K1139 to L.D.W.) and the National Institutes of Health (R01GM118803 to R.W.).

Conflict of Interest: The authors declare that they have no conflict of interest with the contents of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.03.010.

Received 16 November 2018;
Received in revised form 7 March 2019;
Accepted 8 March 2019
Available online 15 March 2019

Keywords:
rRNA;
expansion segments;
Chordates;
helicases;
polysomes

Abbreviations used:
ES, expansion segment; EMSA, electrophoretic mobility shift assay; SILAC, stable isotope labeling by amino acids in cell culture; MSA, multiple sequence alignment; PDS, pyridostatin; MBN, mung bean nuclease; hnRNP, heterogeneous nuclear ribonucleoprotein; ThT, thioflavin T; ACN, acetonitrile; FDR, false discovery rate.

References

[1] R. Milo, R. Phillips, Cell Biology by the Numbers, Garland Science, New York, New York, 2016.
[2] C.R. Bernier, A.S. Petrov, N.A. Kovacs, P.I. Penev, L.D. Williams, Translation: the universal structural core of life, Mol. Biol. Evol. 34 (2018) 2065–2076.
[3] V.C. Ware, B.W. Tague, C.G. Clark, R.L. Gourse, R.C. Brand, S.A. Gerbi, Sequence analysis of 28S ribosomal DNA from the amphibian *Xenopus laevis*, Nucleic Acids Res. 11 (1983) 7795–7817.
[4] C.G. Clark, B.W. Tague, V.C. Ware, S.A. Gerbi, *Xenopus laevis* 28S ribosomal RNA: a secondary structure model and its evolutionary and functional implications, Nucleic Acids Res. 12 (1984) 6197–6220.
[5] N. Hassouna, B. Michot, J.P. Bachellerie, The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes, Nucleic Acids Res. 12 (1984) 3563–3583.
[6] S.A. Gerbi, Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA, in: R.A. Zimmermann, A.E. Dahlberg (Eds.), Ribosomal RNA—Structure, Evolution, Processing, and Function in Protein Synthesis, CRC Press, Boca Raton, FL 1996, pp. 71–87.
[7] M. Ramesh, J.L. Woolford Jr., Eukaryote-specific rRNA expansion segments function in ribosome biogenesis, RNA. 22 (2016) 1153–1162.
[8] C. Leidig, G. Bange, J. Kopp, S. Amlacher, A. Aravind, S. Wickles, et al., Structural characterization of a eukaryotic chaperone—the ribosome-associated complex, Nat. Struct. Mol. Biol. 20 (2013) 23–28.
[9] A. Gumiero, C. Conz, G.V. Gesé, Y. Zhang, F.A. Weyer, K. Lapouge, et al., Interaction of the cotranslational hsp70 ssb with ribosomal proteins and rRNA depends on its lid domain, Nat. Commun. 7 (2016), 13563. .
[10] A.G. Knorr, C. Schmidt, P. Tesina, O. Berninghausen, T. Becker, B. Beatrix, et al., Ribosome-nata architecture reveals that rRNA expansion segments coordinate n-terminal acetylation, Nat. Struct. Mol. Biol. 26 (2018) 35–39.
[11] S. Melnikov, A. Ben-Shem, N. Garreau de Loubresse, L. Jenner, G. Yusupova, M. Yusupov, One core, two shells: bacterial and eukaryotic ribosomes, Nat. Struct. Mol. Biol. 19 (2012) 560–567.
[12] O. Kikin, L. D'Antonio, P.S. Bagga, QGRS mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences, Nucleic Acids Res. 34 (2006) W676–W682.
[13] S. Lammich, F. Kamp, J. Wagner, B. Nuscher, S. Zilow, A.-K. Ludwig, et al., Translational repression of the disintegrin and metalloprotease ADAM10 by a stable G-quadruplex secondary structure in its 5′-untranslated region, J. Biol. Chem. 286 (2011) 45063–45072.
[14] A. von Hacht, O. Seifert, M. Menger, T. Schutze, A. Arora, Z. Konthur, et al., Identification and characterization of RNA guanine–quadruplex binding proteins, Nucleic Acids Res. 42 (2014) 6630–6644.
[15] P. Fratta, S. Mizielinska, A.J. Nicoll, M. Zloh, E.M. Fisher, G. Parkinson, et al., *C9orf72* hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes, Sci. Rep. 2 (2012) 1016.
[16] K. Halder, M. Wieland, J.S. Hartig, Predictable suppression of gene expression by 5′-UTR-based RNA quadruplexes, Nucleic Acids Res. 37 (2009) 6811–6817.
[17] A. Arora, M. Dutkiewicz, V. Scaria, M. Hariharan, S. Maiti, J. Kurreck, Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif, RNA. 14 (2008) 1290–1296.
[18] S. Kumari, A. Bugaut, J.L. Huppert, S. Balasubramanian, An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation, Nat. Chem. Biol. 3 (2007) 218–221.
[19] K.H. Choi, B.S. Choi, Formation of a hairpin structure by telomere 3′ overhang, Biochim. Biophys. Acta 1217 (1994) 341–344.
[20] S. Xu, Q. Li, J. Xiang, Q. Yang, H. Sun, A. Guan, et al., Thioflavin T as an efficient fluorescence sensor for selective recognition of RNA G-quadruplexes, Sci. Rep. 6 (2016).
[21] A. Renaud de la Faverie, A. Guédin, A. Bedrat, L.A. Yatsunyk, J.L. Mergny, Thioflavin T as a fluorescence light-up probe for G4 formation, Nucleic Acids Res. 42 (2014) e65-e.
[22] S. Zhang, H. Sun, H. Chen, Q. Li, A. Guan, L. Wang, et al., Direct visualization of nucleolar G-quadruplexes in live cells by using a fluorescent light-up probe, Biochim. Biophys. Acta 1862 (2018) 1101–1106.
[23] J.-L. Mergny, A.-T. Phan, L. Lacroix, Following G-quartet formation by UV-spectroscopy, FEBS Lett. 435 (1998) 74–78.
[24] K.-W. Zheng, Z. Chen, Y.-h. Hao, Z. Tan, Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA, Nucleic Acids Res. 38 (2009) 327–338.
[25] G. Biffi, M. Di Antonio, D. Tannahill, S. Balasubramanian, Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells, Nat. Chem. 6 (2014) 75–80.
[26] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, Quantitative visualization of DNA G-quadruplex structures in human cells, Nat. Chem. 5 (2013) 182–186.
[27] X. Cang, J. Šponer, T.E. Cheatham III, Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes, Nucleic Acids Res. 39 (2011) 4499–4512.
[28] X.-M. Li, K.-W. Zheng, J.-Y. Zhang, H.-H. Liu, B.-F. Yuan, Y.-h. Hao, et al., Guanine-vacancy-bearing G-quadruplexes responsive to guanine derivatives, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 14581–14586.
[29] D. Benhalevy, S.K. Gupta, C.H. Danan, S. Ghosal, H.-W. Sun, H.G. Kazemier, et al., The human CCHC-type zinc

finger nucleic acid-binding protein binds G-rich elements in target mRNA coding sequences and promotes translation, Cell Rep. 18 (2017) 2979–2990.
[30] S. Khateb, P. Weisman-Shomer, I. Hershco-Shani, A.L. Ludwig, M. Fry, The tetraplex (CGG)n destabilizing proteins hnRNP A2 and CBF-A enhance the in vivo translation of fragile X premutation mRNA, Nucleic Acids Res. 35 (2007) 5775–5788.
[31] M.J. Matunis, J. Xing, G. Dreyfuss, The hnRNP f protein: unique primary structure, nucleic acid-binding properties, and subcellular localization, Nucleic Acids Res. 22 (1994) 1059–1067.
[32] M. Caputi, A.M. Zahler, Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) h/h′/f/2h9 family, J. Biol. Chem. 276 (2001) 43850–43859.
[33] E.K. McRae, E.P. Booy, G.P. Padilla-Meier, S.A. McKenna, On characterizing the interactions between proteins and guanine quadruplex structures of nucleic acids, J. Nucleic Acids 2017 (2017).
[34] E. Dardenne, M.P. Espinoza, L. Fattet, S. Germann, M.-P. Lambert, H. Neil, et al., RNA helicases DDX5 and DDX17 dynamically orchestrate transcription, miRNA, and splicing programs in cell differentiation, Cell Rep. 7 (2014) 1900–1913.
[35] B. Herdy, C. Mayer, D. Varshney, G. Marsico, P. Murat, C. Taylor, et al., Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts, Nucleic Acids Res. 46 (21) (2018) 11592–11604, https://doi.org/10.1093/nar/gky861.
[36] A. Chaudhury, P. Chander, P.H. Howe, Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: focus on hnRNP e1's multifunctional regulatory roles, RNA. 16 (2010) 1449–1462.
[37] D. Simsek, G.C. Tiu, R.A. Flynn, G.W. Byeon, K. Leppek, A. F. Xu, et al., The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity, Cell. 169 (2017) 1051–1065 (e18).
[38] J.L. Huppert, A. Bugaut, S. Kumari, S. Balasubramanian, G-quadruplexes: the beginning and end of UTRs, Nucleic Acids Res. 36 (2008) 6260–6268.
[39] J. Eddy, N. Maizels, Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes, Nucleic Acids Res. 36 (2008) 1321–1333.
[40] J.U. Guo, D.P. Bartel, RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria, Science. 353 (2016) aaf5371.
[41] S.Y. Yang, P. Lejault, S. Chevrier, R. Boidot, A.G. Robertson, J.M. Wong, et al., Transcriptome-wide identification of transient RNA G-quadruplexes in human cells, Nat. Commun. 9 (2018) 4730.
[42] A. Jain, R.D. Vale, RNA phase transitions in repeat expansion disorders, Nature. 546 (2017) 243–247.
[43] D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, et al., Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis, Cancer Res. 69 (2009) 7653–7661.
[44] S. Chiarella, A. De Cola, G.L. Scaglione, E. Carletti, V. Graziano, D. Barcaroli, et al., Nucleophosmin mutations alter its nucleolar localization by impairing G-quadruplex binding at ribosomal DNA, Nucleic Acids Res. 41 (2013) 3228–3239.
[45] J. Kim, C. Cheong, P.B. Moore, Tetramerization of an RNA oligonucleotide containing a GGGG sequence, Nature. 351 (1991) 331.
[46] R. del Villar-Guerra, R.D. Gray, J.B. Chaires, Characterization of quadruplex DNA structure by circular dichroism, Curr. Protoc. Nucleic Acid Chem. (2017) 17.8. 1–8. 6.
[47] D. Shcherbakov, W. Piendl, A novel view of gel-shifts: analysis of RNA–protein complexes using a two-color fluorescence dye procedure, Electrophoresis. 28 (2007) 749–755.
[48] C. Bernier, A.S. Petrov, C. Waterbury, J. Jett, F. Li, L.E. Freil, et al., Ribovision: visualization and analysis of ribosomes, Faraday Discuss. 169 (2014) 195–207.
[49] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[50] Database resources of the national center for biotechnology information, Nucleic Acids Res. 45 (2017) D12–d7.
[51] K. Katoh, D.M. Standley, A simple method to control over-alignment in the MAFFT multiple sequence alignment program, Bioinformatics. 32 (2016) 1933–1942.
[52] T.A. Hall, Bioedit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/nt, Nucleic Acids Symp. Ser, , Information Retrieval Ltd., London, 1999 95–98 [c1979-c2000].
[53] A.M. Waterhouse, J.B. Procter, D.M. Martin, M. Clamp, G.J. Barton, Jalview version 2—a multiple sequence alignment editor and analysis workbench, Bioinformatics. 25 (2009) 1189–1191.
[54] S.B. Hedges, J. Dudley, S. Kumar, Timetree: a public knowledge-base of divergence times among organisms, Bioinformatics. 22 (2006) 2971–2972.
[55] J.K. Eng, A.L. McCormack, J.R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989.