Kevin D. Raney et al. DEAD-box RNA helicases Dbp2, Ded1 and Mss116 bind to G-quadruplex nucleic acids and destabilize G-quadruplex RNA.
DEAD-box RNA helicases Dbp2, Ded1 and Mss116 bind to G-quadruplex nucleic acids and destabilize G-quadruplex RNA†

Jun Gao,†a Alicia K. Byrd,†a Boris L. Zybaiov,†a John C. Marecki,†a Michael J. Guderyon,†a Andrea D. Edwards,†a Shubeena Chib,†a Kirk L. West,a Zachary J. Waldrip,†a Samuel G. Mackintosh,†a Zhaofeng Gao,†b Andrea A. Putnam,b Eckhard Jankowsky,b and Kevin D. Raney*†a

We identified 29 G-quadruplex binding proteins by affinity purification and quantitative LC-MS/MS. We demonstrated that the DEAD-box RNA helicases Dbp2, Ded1 and Mss116 preferentially bind to G-quadruplex nucleic acids in vitro and destabilize RNA quadruplexes, suggesting new potential roles for these helicases in disruption of quadruplex structures in RNA.

G-quadruplexes (G4) are non-canonical, four-stranded nucleic acid structures, consisting of stacked planar guanine tetrads, in which guanines are held together by Hoogsteen hydrogen bonds.1 G4 structures are of important biological significance in processes such as regulation of gene transcription and translation.2,3 They are also associated with genetic instability and human diseases.4,5 Several proteins have been reported to bind to G4. These proteins include but are not limited to the RNA helicases DHX36/RHAU,6 DHX9,7 and DDX21,8 and the DNA helicases XBP,9 XPD,9 BLM,10 and Pif111,12 and the nucleolar protein nucleolin,13 transcription factor SP1,14 yeast tRNA binding protein Arc1/G4p1,15 yeast transcription coactivator Sub1 and its human homolog PC4.16 Herein, we report that DEAD-box RNA helicases Dbp2, Ded1, and Mss116 preferentially bind to both G4DNA and G4RNA and destabilize G4RNA.

Using affinity purification from Saccharomyces cerevisiae whole cell lysates with a G4DNA bait followed by quantitative LC-MS/MS analysis (Fig. 1), we systemically identified 29 proteins as G4 binding proteins (Table 1). The complete list of proteins identified by mass spectrometry is in Table S2 (ESI†). Arc1/G4p1 is a known G4DNA binding protein,15 and we recently characterized the G4DNA interaction of Sub1, a global transcriptional coactivator, and confirmed that it preferentially binds to G4DNAs.16 The sequences of G-quadruplex DNA interacting proteins identified in our proteomics screen were analysed with the MEME package,17 an online motif finder tool. A discriminatory search was performed using non-G4 binding sequences.18 The RGG motif was the top scoring motif found in seven of the G4DNA binding proteins [Fig. S1, ESI†], which is consistent with previous reports that RGG motifs in proteins such as nucleolin and FMRP can interact with G-quadruplexes.19–21 The next two most highly enriched motifs are conserved motifs from the DEAD-box family of RNA helicases Fig. S1, ESI†, which is consistent with previous reports that RGG motifs in proteins such as nucleolin and FMRP can interact with G-quadruplexes.19–21 The next two most highly enriched motifs are conserved motifs from the DEAD-box family of RNA helicases [Fig. S1, ESI†]. In Table 1, the most highly enriched protein in the G4DNA sample was the DEAD-box RNA helicase Dbp2. In addition to Dbp2, three other DEAD-box RNA helicases (Ded1, Dbp1, and Mss116) were enriched in the G4DNA pull-down samples.

DEAD-box RNA helicases, characterized by the conserved “Asp-Glu-Ala-Asp” (DEAD) motif, form a large protein family which is found in nearly all organisms.23,24 They play important roles in numerous essential cellular processes23,24 such as gene transcription, mRNA processing and translation, stress granule and P-body formation,25 and ribonucleoprotein (RNPs) remodelling.26,27 This family of helicases is also involved in disease (such as cancer) development and progression, viral infection and replication,10,11 as well as host innate immune responses.32,33

We therefore purified recombinant Dbp2, Ded1 and Mss116, and measured their binding affinity to different conformations of DNA (Fig. 2a and Fig. S2a, ESI†). Dbp2 binds to tailed cMYC G4DNA very tightly, with a dissociation...
cMYC G4DNAs is also tighter than that for dsRNA (tailless cMYC G4DNA ($K_d$ binds to both the tailed cMYC G4DNA ($K_d$ binds to tailless ScTEL G4DNA (Fig. 2d). However, Mss116 shows affinity to duplex DNA nor to hTEL G4DNA. These results partially to G4DNAs, including parallel and hybrid G4DNAs. 

Thus, Mss116 preferentially binds to parallel G4DNA.

Tailless cMYC G4DNA (Fig. 2), suggesting the tail or the junction between the tail and G4DNA may also contribute to the binding. However, each of these helicases also bind the tailless cMYC G4DNA with high affinity indicating that the G4 structure itself recognized by the helicases. Although these three DEAD-box helicases bind to G4DNAs preferentially, none of them is capable of unfolding a G4DNA substrate containing either a 5′- or 3′-ssDNA extension (Fig. S3, ESI†).

Next, we conducted a bioinformatic search of yeast transcriptome, and found that the most frequently occurring G4RNA sequences in the yeast transcriptome are two-tetrad G4DNAs (Table S3, ESI†). The YNL098C G4RNA is the most common G4RNA sequence present in the yeast transcriptome, and it forms a parallel G4 structure in vitro (Fig. S2b and c, ESI†).

We measured the binding affinities of these three DEAD-box RNA helicases to G4RNA vs. ssRNA (Fig. 3). Dpb2 binds to YNL098C G4RNA very tightly ($K_d$ = 0.84 ± 0.12 nM), more than 200 fold tighter than that of ssRNA ($K_d$ = 210 ± 20 nM) (Fig. 3b). Ded1 also binds to YNL098C G4RNA very tightly ($K_d$ = 1.18 ± 0.5 nM), 14 fold tighter than that of ssRNA ($K_d$ = 16.6 ± 3.3 nM) (Fig. 3c). For Mss116, binding to YNL098C G4RNA ($K_d$ = 43 ± 10 nM, Fig. 3d) is approximately 4 fold tighter than binding to the ssRNA ($K_d$ = 180 ± 10 nM, Fig. S4, ESI†). Thus, Dpb2, Ded1, and Mss116 bind preferentially to G4RNA.

We then examined whether binding of these DEAD-box RNA helicases to G4RNA destabilizes the G4RNA structures constant ($K_d$ of 0.55 ± 0.1 nM (Fig. 2b). Binding to tailless G4DNA is about three-fold tighter than binding to the tailless cMYC G4DNA ($K_d$ = 1.9 ± 0.3 nM), 10 fold tighter than binding to the ssDNA ($K_d$ = 5.4 ± 0.9 nM), and 62 fold tighter than its binding affinity for the duplex DNA ($K_d$ = 34 ± 11 nM). Dpb2 also tightly binds to tailless ScTEL G4DNA ($K_d$ = 0.72 ± 0.28 nM), and tailless hTEL G4DNA ($K_d$ = 3.0 ± 0.4 nM). Thus, Dpb2 binds preferentially to G4DNAs, including parallel and hybrid G4DNAs.

Ded1 also binds to tailless cMYC G4DNA ($K_d$ = 1.3 ± 0.2 nM), 2.5 fold tighter than its binding affinity to the tailless cMYC G4DNA ($K_d$ = 3.4 ± 0.7 nM), 4.5 fold tighter than binding to the ScTEL G4DNA ($K_d$ = 5.8 ± 2.5 nM), and 21 fold tighter than binding to the ssDNA ($K_d$ = 27 ± 3 nM (Fig. 2c). Ded1 does not show affinity to duplex DNA nor to hTEL G4DNA. These results demonstrate that Ded1 preferentially binds to G4DNAs, including parallel and some hybrid G4DNAs.

Mss116 protein shows little to no binding to the ssDNA, duplex DNA, tailless ScTEL or tailless hTEL (Fig. 2d). However, Mss116 binds to both the tailless cMYC G4DNA ($K_d$ = 10 ± 1 nM) and the tailless cMYC G4DNA ($K_d$ = 19 ± 5 nM). The affinity of Mss116 for cMYC G4DNAs is also tighter than that for dsRNA ($K_d$ ≈ 120 nM). Thus, Mss116 preferentially binds to parallel G4DNA.

All three DEAD-box helicases (Dpb2, Ded1, and Mss116) bind to tailless cMYC G4DNA with higher affinity than that of tailless G4DNA (Fig. 2), suggesting the tail or the junction between the tail and G4DNA may also contribute to the binding. However, each of these helicases also bind the tailless cMYC G4DNA with high affinity indicating that the G4 structure itself recognized by the helicases. Although these three DEAD-box helicases bind to G4DNAs preferentially, none of them is capable of unfolding a G4DNA substrate containing either a 5′- or 3′-ssDNA extension (Fig. S3, ESI†).

Next, we conducted a bioinformatic search of yeast transcriptome, and found that the most frequently occurring G4RNA sequences in the yeast transcriptome are two-tetrad ribonucleotide sequences (Table S3, ESI†). We selected the top four G4RNA sequences from yeast transcriptome, and confirmed that they form stable G4RNA structures by circular dichroism spectra analysis (Fig. S2b, ESI†). The YNL098C G4RNA is the most common G4RNA sequence present in the yeast transcriptome, and it forms a parallel G4 structure in vitro (Fig. S2b and c, ESI†).

We measured the binding affinities of these three DEAD-box RNA helicases to G4RNA vs. ssRNA (Fig. 3). Dpb2 binds to YNL098C G4RNA very tightly ($K_d$ = 0.84 ± 0.12 nM), more than 200 fold tighter than that of ssRNA ($K_d$ = 210 ± 20 nM) (Fig. 3b). Ded1 also binds to YNL098C G4RNA very tightly ($K_d$ = 1.18 ± 0.5 nM), 14 fold tighter than that of ssRNA ($K_d$ = 16.6 ± 3.3 nM) (Fig. 3c). For Mss116, binding to YNL098C G4RNA ($K_d$ = 43 ± 10 nM, Fig. 3d) is approximately 4 fold tighter than binding to the ssRNA ($K_d$ = 180 ± 10 nM, Fig. S4, ESI†). Thus, Dpb2, Ded1, and Mss116 bind preferentially to G4RNA.
The diversity of proteins enriched on G4DNA implies that G4 structures play critical roles in multiple biological functions. We demonstrated three DEAD-box RNA helicases (Dbp2, Ded1, and Miss16) bind preferentially to G4DNA and G4RNA in vitro. Binding of Dbp2, Ded1 or Miss16 to G4DNA does not affect G4DNA structure. However, their binding destabilizes G4RNA in an ATP-independent manner. Our observation that binding of these proteins to G4RNA is sufficient to unfold the structure, and provides one possible explanation for the recent observation that G4RNAs are globally unfolded in eukaryotic cells. A recent X-ray crystal structure and single-molecule analysis captured a G4DNA which was partially unfolded in a transient ATP-independent manner due to binding of the RNA helicase, DHX36. Our findings suggest possible new roles for DEAD-box RNA helicases through interaction with G4DNAs and disruption of G-quadruplex structures in RNAs.

This work was supported by National Institutes of Health Grants (R01 GM098822, R01 GM117439, and R35 GM122601 to K. D. R.) and the UAMS Research Council. The University of Arkansas for Medical Sciences (UAMS) Proteomics Core and DNA Sequencing Core are supported by the Arkansas IDeA Network for Biomedical Research Excellence (National Institutes of Health Grant P20 GM103429), the University of Arkansas Center for General Medical Sciences (UAMS) Proteomics Core and DNA Sequencing Core are supported by the Arkansas IDeA Network for Biomedical Research Excellence (National Institutes of Health Grant P20 GM103429), the University of Arkansas Center for...
Protein Structure and Function (National Institutes of Health Grant P30 GM103450), the UAMS Center for Microbial Pathogenesis and Host Inflammatory Responses (National Institutes of Health Grant P20 GM103625), and the UAMS Translational Research Institute (National Institutes of Health Grant UL1TR000039). We thank Dr Elizabeth Tran for the pET28a-DBP2 plasmid and helpful suggestions, and Dr Galina Glazko and Dr Stephanie Byrum for helpful discussions and technical consultation.

Conflicts of interest

There are no conflicts to declare.

Notes and references