The OB-fold proteins of the *Trypanosoma brucei* editosome execute RNA-chaperone activity

Christin Voigt[1], Mateusz Dobrychłop[2], Elisabeth Kruse[1], Anna Czerwoniec[2], Joanna M. Kasprzak[2], Patrycja Bytner[2], Cristian Del Campo[1], W.-Matthias Leeder[1], Janusz M. Bujnicki[2,3,*] and H. Ulrich Göringer [ID][1,*]

[1]Molecular Genetics, Darmstadt University of Technology, Darmstadt, Germany, [2]Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland and [3]Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland

ABSTRACT

Sequence-deficient mitochondrial pre-mRNAs in African trypanosomes are substrates of a U-nucleotide-specific RNA editing reaction to generate translation-competent mRNAs. The reaction is catalyzed by a macromolecular protein complex termed the editosome. Editosomes execute RNA-chaperone activity to overcome the highly folded nature of pre-edited substrate mRNAs. The molecular basis for this activity is unknown. Here we test five of the OB-fold proteins of the *Trypanosoma brucei* editosome as candidates. We demonstrate that all proteins execute RNA-chaperone activity albeit to different degrees. We further show that the activities correlate to the surface areas of the proteins and we map the protein-induced RNA-structure changes using SHAPE-chemical probing. To provide a structural context for our findings we calculate a coarse-grained model of the editosome. The model has a shell-like structure: Structurally well-defined protein domains are separated from an outer shell of intrinsically disordered protein domains, which suggests a surface-driven mechanism for the chaperone activity.

INTRODUCTION

Editosomes are high molecular mass (0.8 MDa, 20S) protein complexes that catalyze the U-insertion/deletion RNA-editing reaction in African trypanosomes and other kinetoplastid organisms (1,2). The processing reaction is characterized by the site-specific insertion and to a lesser extent, deletion of exclusively U-nucleotides and converts cryptic primary transcripts into translatable mRNAs. Editosomes harbor one substrate-RNA binding site (3) and in-

teract with a large set of pre- and partially edited mitochondrial transcripts to execute the reaction. Recent structure probing experiments uncovered that the pre-edited mRNAs are unusually folded (4). The different transcripts have thermodynamic stabilities that resemble structural RNAs and due to a very high content of runs of G-nucleotides they contain multiple, up to five, G-quadruplex (GQ)-folds (5). Perhaps as a consequence, *Trypanosoma brucei* editosomes execute a complex-intrinsic RNA-chaperone activity (3). The activity acts by increasing the flexibility of predominantly U-residues to lower their base-pairing probability thereby generating a simplified RNA-folding landscape with a reduced energy barrier to facilitate the binding of gRNAs (4). Thus, the editosome-driven RNA-unfolding reaction is important for the editing cycle especially during the initiation and elongation phases of the process. However, the molecular nature of the chaperone activity is not understood.

Proteins with RNA-chaperone activity (6) represent a structurally diverse group of polypeptides that contribute to almost all RNA-driven biochemical processes in all domains of life (7). This includes many viral and bacterial proteins such as Ncp7 (8), StpA and Hfq (9,10) and a large number of ribosomal proteins of both, pro- and eukaryotic origin (11). Importantly, several of the RNA chaperones contain one or more OB-fold motif(s). OB refers to the general oligonucleotide/oligosaccharide-binding ability of the proteins, which is mediated by a five-stranded β-barrel fold (12). OB-folds have specifically been identified in bacterial and plant cold shock proteins (Csp) where they contribute to resolve misfolded RNA-species (13,14). Furthermore, OB-fold proteins are universally involved in RNA-remodeling steps prior to the initiation of protein biosynthesis with ribosomal protein S1 and *Escherichia coli* initiation factor 1 (IF1) containing multiple OB-domains (15,16).

The protein inventory of the *T. brucei* 20S editosome lists six OB-fold proteins termed TbMP81, TbMP63, TbMP42,

---

TbMP24, TbMP19 and TbMP18 following the nomenclature of Worthey *et al*. (17). TbMP stands for *T. brucei* mitochondrial protein followed by a number indicating the calculated molecular mass in kDa (for an alternative nomenclature see (18) and Supplementary Table S1). TbMP81, TbMP63 and TbMP42 further contain two C2H2-type Zn-fingers or C2H2-Zn-finger-like domains and all six proteins have been predicted to harbor long stretches of intrinsically disordered regions (IDR) (19). The domain structure of the different proteins, a plot of their disorder propensities (20) and all identified inter-OB-fold protein crosslinks within the 20S editosome (21) are shown in Figure 1. A 3D-consensus model of the OB-fold domain is presented in Supplementary Figure S1 and the characteristics of the ID-regions are summarized in Supplementary Table S2.

Although several of the editosomal OB-fold proteins have been investigated in gene ablation experiments and as recombinant proteins, no coherent function for any of the six polypeptides has emerged from these studies. TbMP24, TbMP18 and TbMP42 were identified as RNA-binding proteins (22–24). Recombinant TbMP42 was characterized to execute endo/exonuclease activity (24,25) and recombinant TbMP24 was shown to perform RNA-annealing activity (26). Gene knockdown of five of the proteins (TbMP18, TbMP24, TbMP42, TbMP63, TbMP81) identified an impact on the structural integrity of the editosome (23,27–30), which was supported by yeast two-hybrid experiments in combination with protein/protein interaction studies (31). The data suggest a scenario in which TbMP18, TbMP24, TbMP42, TbMP63 and TbMP81 interact with each other using the OB-folds as docking modules (32). In agreement with such a situation, recent interprotein chemical cross-linking data have identified cross-links between all six OB-fold proteins most of which within or in proximity to their OB-fold domains (21).

Importantly, clustered OB-folds in the yeast RRP44 protein have been demonstrated to catalyze the unwinding of dsRNA (33,34). As a consequence, Böhm *et al*. (3) suggested that the interacting OB-fold proteins of the *T. brucei* editosome might act in a similar fashion. Here, we show that five of the six *T. brucei* editosomal OB-fold proteins indeed catalyze a RNA-chaperone reaction. We map the protein-induced RNA-structure changes using SHAPE-chemical probing and demonstrate that the activities correlate to the surface areas of the different proteins. The data suggest a cumulative contribution of the individual proteins to the chaperone activity of the editosome and to provide a structural context for our finding we present a coarse-grained model of the 20S editosome. The model has shell-like characteristics: Structurally well-defined protein domains are separated from an outer layer of intrinsically disordered protein domains, which suggests a surface-driven mechanism for the chaperone activity.

MATERIALS AND METHODS

Expression of recombinant proteins

DNA-sequences for all editosomal OB-fold protein constructs were PCR-amplified from *T. brucei* genomic DNA (strain Lister 427) (35) using the DNA-oligonucleotide primers listed in Supplementary Table S3. Plasmid expression vectors were modifications of pCDF-1b and pET-33b (Novagen) with the original protease cleavage sites replaced by the cleavage site for tobacco etch virus (TEV) protease. Mitochondrial targeting sequences were predicted using MitoProt (36) and were omitted from the constructs. The proteins were expressed as full length (FL) and OB-fold-only polypeptides containing a cleavable N-terminal hexa-histidine (His$_6$)-tag: TbMP19-FL (aa 18–170), TbMP63-FL (aa 53–587), TbMP81-FL (aa 56–762), TbMP24-OB (aa 126–246), TbMP24-FL (aa 47–246), TbMP42-OB (aa 245–371), TbMP42-FL (aa 22–371), TbMP63-OB (aa 472–587) and TbMP81-OB (aa 626–762). The pRSF-duet plasmid for the expression of His$_6$-tagged TbMP18-FL (aa 19–164) was a gift from Wim Hol (University of Washington, Seattle). Recombinant plasmids were propagated in *E. coli* DH5α. For protein expression, the plasmids were transformed into Rosetta™(DE3)pLysS (Novagen). Cells were grown to late log phase (A_{600} app. 0.8) in Luria broth at 37°C. Protein expression was induced by adding 1 mM isopropyl 1-thio-β-D-galactopyranosid (IPTG) at 18°C overnight or at 37°C for 4 h. Cells were harvested by centrifugation and processed either directly or stored at –20°C until further use.

Protein purification and size exclusion chromatography

Escherichia coli cells were lysed in the presence of 0.25 mg/ml lysozyme by three freeze/thaw cycles and sonication in binding buffer (BB): 20 mM Tris–HCl, pH 7.6, 0.3 M NaCl, 20 mM imidazole containing 1 mM phenyl-methanesulfonyl fluoride (PMSF) and 1 μg/ml leupeptin. In the case of the Zn-finger-containing proteins (TbMP42, TbMP63, TbMP81) 0.1 mM ZnSO$_4$ was added. Proteins were purified by immobilized metal affinity chromatography (IMAC) using a Ni-Sepharose HP-resin (GE Healthcare). Protein binding was performed in batch. All subsequent washing and elution steps (with 250 mM imidazole in BB) were performed 'in-column'. His$_6$-tags were removed by treating the eluates with TEV-protease and 1mM DTT overnight (4°C) followed by a second binding step to the affinity-resin. Final protein solutions were supplemented with DTT (2 mM), shock frozen in liquid nitrogen, either directly or after adding 50% (v/v) glycerol and stored at –80°C. The oligomerization state of the different protein constructs was analyzed by size exclusion chromatography using a Superdex 200 HR 10/30 column in 20 mM Tris–HCl pH 7.6, 300 mM NaCl, 1 mM DTT. Elution profiles were monitored by absorbance measurements at 280 nm (A_{280}). Molecular masses were determined by calibration to globular proteins. Protein fractions were precipitated with 10% (w/v) trichloroacetic acid (TCA) and analyzed in 10–15% (w/v) SDS-containing polyacrylamide gels. All recombinant protein preparations were analyzed by mass spectrometry to exclude the co-purification of known *E. coli* chaperone proteins. For that the protein isolates were carbamidomethylated followed by trypsination and nano-LC separation of the generated peptides. Peptides were analyzed by MS/MS spectrometry (Bruker Maxis Impact Q-TOF) and Mascot database searched.
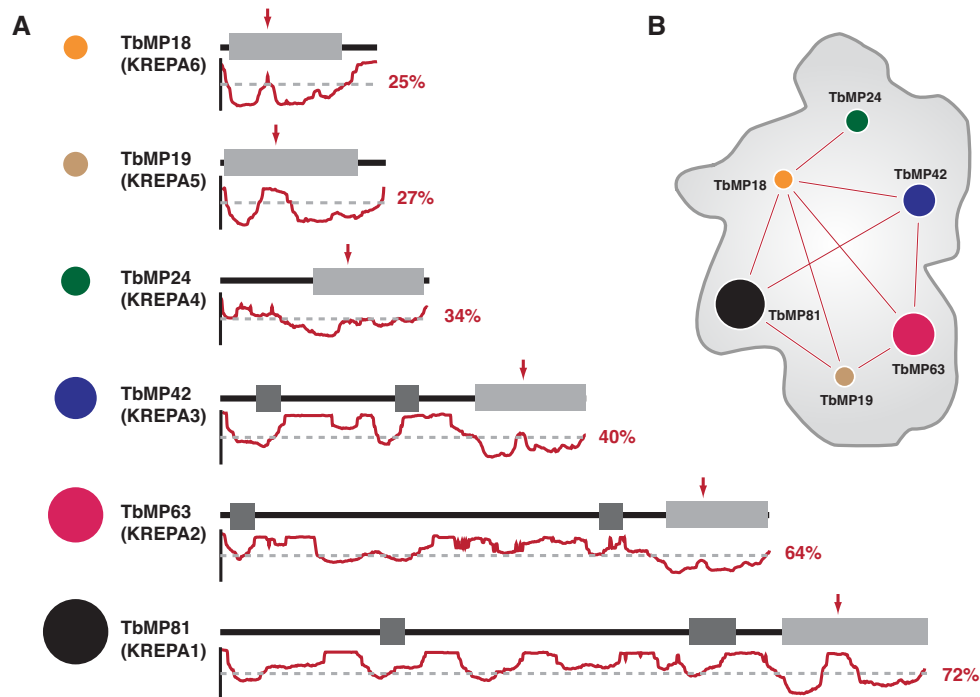
Figure 1. Domain structure of the OB-fold proteins of the *T. brucei* editosome. (**A**) OB-fold domains: boxes in light grey. C2H2-zinc fingers and C2H2-zinc finger-like domains: boxes in dark gray. Disorder propensity plots (between 0 and 1) for each protein are shown in red. The dashed grey line indicates a value of 0.5. Average disorder propensities are listed as percent values in red. Left: the different proteins are depicted as colored dots with diameters reflecting their molecular masses. Arrows in red: intrinsically disordered regions (IDR) within the OB-fold domains. (**B**) Outline of the cryo-EM structure of the *T. brucei* 20S editosome (49) with the 6 OB-fold proteins arbitrarily positioned in the form of colored dots. Red lines indicate established inter-protein crosslinks between the different polypeptides (21).

Circular dichroism spectroscopy

CD-spectra were recorded between 320 and 190 nm at 27°C in 17 mM $K_xH_yPO_4$ pH 7.5, 10 mM $MgCl_2$ and 0.5 mM DTT at protein concentrations of 0.1–0.2 mg/ml in a 0.1 cm cuvette. Data were collected every 0.2 nm at a scan speed of 50 nm/min. Five spectra were binned and background subtracted. Signals from 190 to 240 nm were used to calculate secondary structure contents using CDSSTR (37) with reference data sets 4 and 7 (38) on the DichroWeb server (39,40). The calculated values of the two reference sets were averaged. The fractions of regular and distorted types of α-helices and β-sheets and the fraction of turns and unordered residues were summed up.

RNA-chaperone assay and protein surface area calculation

The RNA-chaperone activity of the different protein constructs was analyzed as in (4) using pre-edited CYb-RNA (1 nM) and 'gDNA' CYb-5 (100 nM) as a pre-mRNA/gDNA-pair. Experiments were conducted using minimally six different protein concentrations (10–500 nM) in at least two independent experiments to yield ≥12 data points. The amount of cleaved RNA was determined by peak-integration and was normalized for each experiment. Editing-active 20S editosome preparations served as positive controls (4) and were prepared as in (49). Protein preparations derived from non-induced *E. coli* cells that harbor a OB-fold protein expression plasmid served as negative, i.e. background (BG)-controls. None of the BG-

samples showed any measurable RNA-chaperone activity (see Supplementary Figure S2D). Data points were fitted to a sigmoidal function: $y = L_{low} + L_{up}/[1 + e^{(c_{1/2}-x/\text{rate})}]$ (L_{low} = lower limit, L_{up} = upper limit). The concentration of half maximal RNA-cleavage ($c_{1/2}$) represents a proxy for the RNA-chaperone activity. (Note: since the amount of active protein in the different preparations is not known, all $c_{1/2}$-values must be considered apparent RNA-chaperone activities.) Normalized surface areas were calculated based on the assumption that the proteins have globular shapes. Volume (V) calculations used an average partial specific volume of 0.73 cm^3/g from the molecular masses determined by size exclusion chromatography. Surface areas (A) were calculated as: $A = (36\pi V^2)^{1/3}$. All values were normalized to the surface of TbMP81-OB as the smallest protein.

Isothermal differential absorption spectroscopy

Pre-edited A6-mRNA (~100 pmol) was denatured in 15 µl 10 mM Tris–HCl pH 8, 0.1 mM EDTA for 2 min at 95°C and snap cooled. The RNA was transferred into 0.45 ml 20 mM HEPES pH 7.5, 30 mM KCl, 10 mM $MgCl_2$, 0.1 mM $ZnSO_4$ and equilibrated for 30 min at 27°C. Hereafter, molar equivalents of TbMP81-FL or TbMP42-FL were added to the blank and the sample. After 5 min of equilibration, UV-spectra were recorded between 210 and 350 nm at a scanning rate of 400 nm/min. Data acquisition was every 0.7 nm. The resulting spectra were Rayleigh-scattering corrected (41,42) and normalized by setting the peak at 260 nm to 1 and the baseline between 350 and 320 nm to zero.

Isothermal difference (ID)-spectra (43) were calculated as the difference (Δ) between the spectra recorded in presence and absence of protein. BSA as an unrelated protein served as a negative control. All ID-profiles were Savitzky-Golay smoothed (44) and the dominant negative peak around 278 nm was normalized to –1.

RNA-synthesis and RNA-structure probing

Pre-edited transcripts of apocytochrome b (CYb), ribosomal protein S12 (RPS12) and subunit 6 of the mitochondrial ATPase (A6) were generated by run off *in vitro* T7-transcription following standard procedures. (^{32}P)-labeled RNA preparations were generated by adding α-[^{32}P]-UTP to the transcription mix. Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) was conducted as in (4) using 1-methyl-7-nitroisatoic anhydride (1M7) as the modification reagent. After the RNA-refolding step 300 pmol of TbMP18-FL or TbMP81 was added to the reaction mix and incubated at 27°C for 30 min. Raw electrophoretic traces were analyzed using SHAPEfinder (45) utilizing the boxplot approach to determine the number of statistical outliers. Normalized SHAPE-reactivities were generated by averaging a minimum of three independent experiments and were used as pseudo-Gibbs free energies to calculate RNA 2D-structures for a temperature of 27°C using RNAstructure v5.6 (46,47). ShapeKnots (48) was used to search for pseudoknots using the default parameters $p1 = 0.35$ kcal/mol and $p2 = 0.65$ kcal/mol.

Structure modeling of the 20S editosome

Models of the *T. brucei* 20S editosome structure were generated using PyRy3D (http://genesilico.pi/pyry3d/) in combination with the cryo-EM density map EMDB:1595 (49). TbMP18 and TbMP19 were included in the set of proteins. To identify the most likely position of the different proteins, the conformational space was sampled using a Monte Carlo Simulated Annealing method (50) aiming at maximizing the volume of the map while at the same time minimizing the presence of protein chains outside of the map. Furthermore, restraint violations and steric clashes between the individual components were minimized. Protein sequences predicted to be disordered were modeled in a coarse-grained representation and treated as flexible shapes able to change their conformation during the modeling procedure. PyRy3D was used with its default parameters: Simulated annealing algorithm, starting temperature $T_0 = 10$ in dimensionless units, temperature decrease during the simulation according to $T_n = T_0 \times 0.999^n$ (n = number of the simulation step, 300 000 steps, grid size 1.8 Å). The 0.0419 density threshold was used to define the map volume (49). The approximate position of TbMP52 was derived from the electron density map of the *Leishmania tarentolae* editing complex (51). TbMP52 was pre-oriented to the base of the apex and was allowed to move freely. The RNaseIII-domain containing proteins TbMP61 and TbMP44 as well as TbMP67 and TbMP46 were modelled as dimers (21). Three hundred individual simulations were performed. The top 100 models were clustered according to their RMSD-values to obtain groups of similar solutions. The biggest cluster contained 54 models and was used to calculate a medoid structure. The quality of the fit of the medoid to the cryo-EM map was measured using the cross-correlation coefficient implemented in the 'Fit In Map'-procedure of the UCSF Chimera viewer (52). Superposition of electron density maps was performed using ADP_EM (53). Images of molecular structures were generated using PyMOL 1.5.0.4 (http://www.pymol.org) and Chimera (52). Manipulations of coordinates were performed in Swiss-PDBViewer (54) and PyMOL 1.5.0.4.

Order/disorder analysis of the editosome model

The editosome model was analyzed for the positioning of disordered and structured amino acid residues inside and at the surface of the cryo-EM map of the editosome. For that, the volume of the map (at a contour level of 0.0419) was partitioned into an inner region of high electron density and an outer region of low electron density. The position of a residue in the map was defined by its Cα-atom. The number of structured and disordered amino acids inside the map with contour levels of 0.0419 and 0.249 (which corresponds to 50% of the volume of the 0.0419 contour level), and outside the map were calculated. 'Inner' residues were defined as residues positioned inside the 0.249 value. 'Outer' residues are residues inside the 0.0419 but outside the 0.249 boundary. In addition, a correction was made to account for residues that are inside the shell but point to the inner cavity. By using the 'vop scale'-tool implemented in UCSF Chimera, density values at each point of the electron density map were multiplied by –1, to generate an 'imprint' of the editosome representing the regions of negative electron density (at a contour level of 0.0419). This 'imprint' was segmented using the immersed watershed algorithm (55). All segments not connected to the largest cavity were deleted to obtain the surface of the cavity's volume. Minimal distances of the Cα-atoms to the surface of the obtained cavity were calculated. Structured and disordered residues in each region: 'inner-editosome', 'outer-editosome' and 'outside the map' were divided in two categories with respect to their distance to the cavity surface. The threshold was set to 1.2 nm. The structured and disordered residues that are on the editosome shell and in the editosome core were calculated as: shell = outer residues (above threshold) + outside map (above threshold); core = inner residues (below + above threshold) + outer residues (below threshold) + outside map (below threshold).

RESULTS

Recombinant expression of the OB-fold proteins of the *T. brucei* editosome

To test whether the editosome-inherent RNA-chaperone activity is mediated by one or several of the editosomal OB-fold proteins we expressed the six proteins as recombinant polypeptides in *E. coli*. TbMP81, TbMP63, TbMP42 and TbMP24 were expressed as full-length (FL) and as OB-fold-only (OB) constructs (Figure 2A). TbMP19 and TbMP18, due to their small size, were expressed as FL-constructs only. While the different constructs expressed well upon induction, the majority of polypeptides remained insoluble
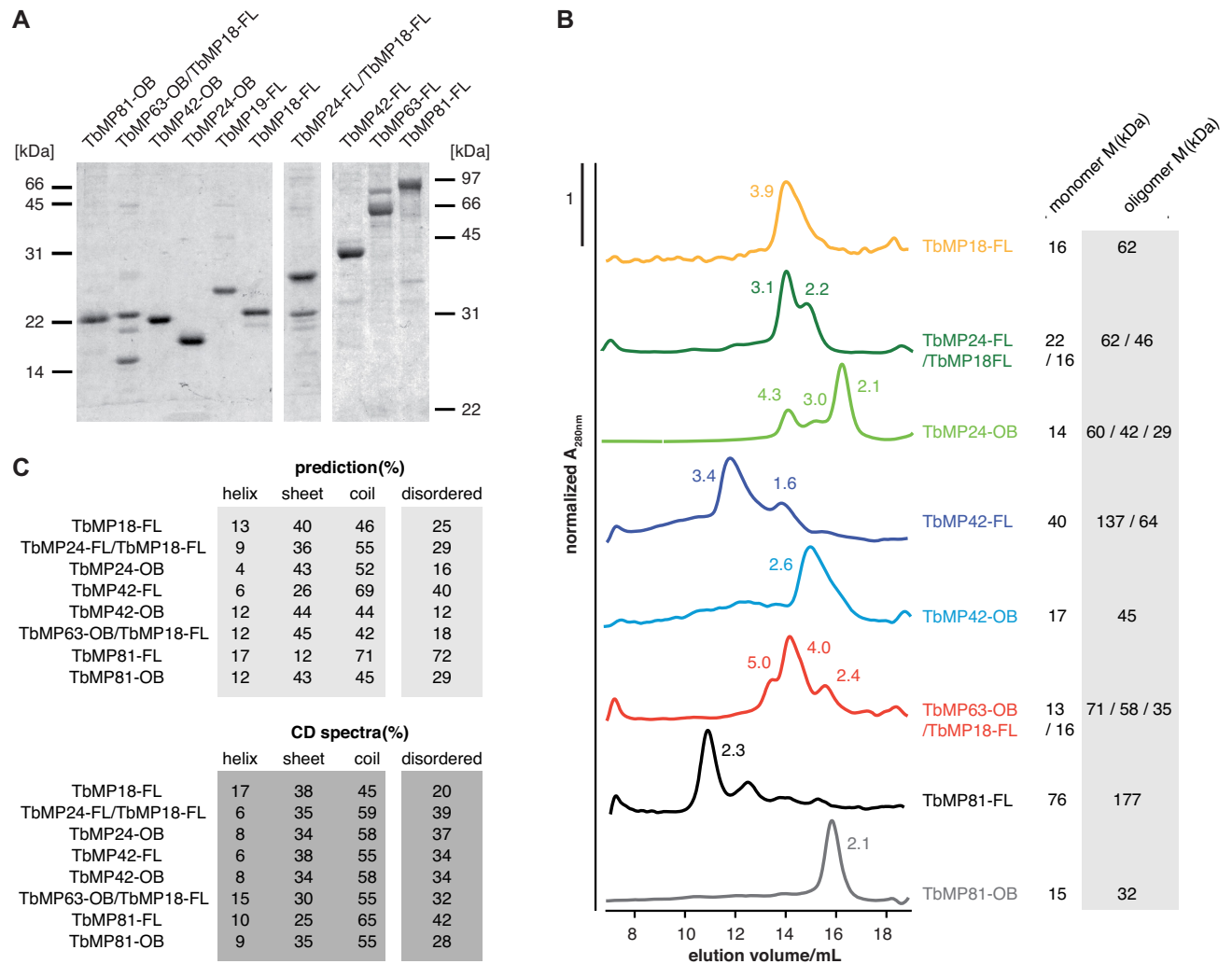
Figure 2. Characterization of recombinant editosomal OB-fold proteins. (**A**) Gel electrophoretic analysis of purified protein preparations in 15% (w/v) (left/center) and 10% (w/v) (right) SDS-containing polyacrylamide gels. Apparent molecular masses are in kDa. (**B**) Normalized size exclusion chromatography profiles of purified OB-fold proteins. Peak numbers indicate oligomerization state(s), which were determined as outlined in the Materials and Methods section. The molar ratios of the co-expressed protein constructs were determined as 2:1 and 1.5:1 for the TbMP24-FL/TbMP18-FL construct and as 3:2, 2:2 and 1:1 for the TbMP63-OB/TbMP18-FL complex. (**C**) Circular dichroism (CD)-derived secondary structure contents of the recombinant OB-fold proteins. Predicted secondary structure contents were derived from the homology models. The structure of non-modeled protein domains was predicted using PSIPRED (68). The molar ratios of the co-expressed constructs was determined as 1.1:1 for the TbMP24-FL/TbMP18-FL complex and as 1:1.5 for the TbMP63-OB/TbMP18-FL complex. Disorder propensities were calculated using MetaDisorderMD2 (20).

at an expression temperature of 37°C. Lowering the temperature to 18°C improved the solubility, however, all attempts to express soluble TbMP19-FL and TbMP63-FL failed. For TbMP63-OB and TbMP24-FL we devised a co-expression regime together with TbMP18-FL as a previously identified interaction partner (21). In the end, 8 of the 10 recombinant protein constructs were available in yields between 5 and 40 mg soluble protein/l *E. coli* culture.

Notably, all expressed protein constructs formed multiple homo-oligomeric higher order assemblies with one dominant assembly state (Figure 2B). TbMP81-FL and TbMP81-OB predominantly formed homodimers, TbMP42-FL, TbMP42-OB and TbMP24-OB homotrimers and TbMP18-FL homotetramers. Importantly, the identified oligomers of the FL and OB-only constructs were invariably identical, indicating that the interaction

surfaces involve the OB-fold domains of the different proteins. In the case of the two co-expressed constructs (TbMP24-FL/TbMP18-FL, TbMP63-OB/TbMP18-FL) we exclusively identified hetero-oligomeric assemblies (Figure 2B), suggesting an enhanced stability and/or solubility of the hetero-oligomeric complexes over the homo-oligomeric complexes.

The purified protein preparations were analyzed for their secondary (2D)-structure content using circular dichroism (CD) spectroscopy (Figure 2C). As expected and in line with structure prediction algorithms, all proteins display a high fraction of unstructured sequence stretches. The α-helical content of the different proteins is as low as 6–17%, the β-sheet content varies between 25 and 38% and the amount of coil structures varies between 45 and 65%. On average 33% of the protein sequences are disordered.

The OB-fold proteins of the *T. brucei* editosome execute RNA-remodeling activity

To quantitatively assess the potential RNA-remodeling activity of the different OB-fold proteins we relied on the recently described RNaseH-based RNA-chaperone assay (4). The assay is based on the rationale that a remodeled, i.e. structurally 'open' target RNA should bind a guide RNA-mimicking, short complementary DNA-oligonucleotide more readily when compared to a structurally constrained RNA. The formed pre-mRNA/DNA-oligonucleotide hybrid molecules can be identified by RNaseH cleavage followed by an electrophoretic separation of the resulting RNA-fragments (Figure 3A, Supplementary Figure S2). As a representative *T. brucei* mitochondrial transcript we selected the pre-edited mRNA of apocytochrome b (CYb-mRNA). The RNA-chaperone activity of 20S editosomes (3,4) served as a positive control (Figure 3B, Supplementary Figure S2) and protein isolates from non-induced *E. coli* cells were analyzed as a background control (Supplementary Figure S2). Importantly, all experiments were performed with varying amounts of protein to use the protein concentration required to achieve half-maximal RNA-cleavage ($c_{1/2}$) as a metric for the RNA-chaperone activity. All purified OB-fold homo-oligomers, co-expressed hetero-oligomers as well as the *in vitro* formed TbMP24-OB/TbMP18-FL hetero-oligomer were tested. Figure 3C shows the dose-dependent curves for all protein isolates. A summary of the data is shown in Figure 3D. Surprisingly, all protein constructs show RNA-remodeling activity. The measured $c_{1/2}$-values range from 57 ± 13 nM (TbMP81-FL) to 164 ± 16 nM (TbMP81-OB) with a mean of 108 ± 12 nM. This represents a 10–30-fold higher protein concentration in comparison to 20S editosomes ($c_{1/2}$ 5 ± 0.4 nM). Thus, although all protein constructs are capable of remodeling the pre-mRNA, they are less active than 20S editosomes. This suggests that the RNA-chaperone activity of the editosome likely represents a cumulative trait, to which each of the different OB-fold proteins contributes a defined increment. Furthermore, we noticed that all FL-proteins are more active than the OB-fold-only constructs. This indicates that protein sequences outside the OB-fold domains, i.e. the ID-regions of the proteins are for the most part responsible for the chaperone activity.

Importantly, for intrinsically disordered RNA-binding proteins it has been shown that the size of their RNA-binding surfaces correlate with the overall sizes of the proteins (56). As a consequence, we analyzed whether the RNA-chaperone activities of the different OB-fold proteins can be correlated to their surface areas. For that we calculated relative surface areas for all proteins and analyzed them in relation to the measured $c_{1/2}$-values (Figure 3D). The two parameters are inversely correlated: the larger the surface area of the protein, the lower its $c_{1/2}$-value. Low $c_{1/2}$-values indicate a high RNA-chaperone activity and thus plotting the normalized RNA-chaperone activities as a function of the normalized suface areas (Figure 3E) results in a saturation-type curve with a correlation coefficient (r^2) of 0.87. Thus, a major determinant of the RNA-chaperone activity of the different OB-fold proteins is their surface area. Proteins with larger surfaces such as TbMP81-FL ex-

ecute a higher chaperone activity and *vice versa*. The high chaperone activity of the editosome ($c_{1/2} = 5 \pm 0.4$ nM) is the result of the large surface of the 0.8 MDa complex.

The RNA-chaperone activity is incremental and cumulative

To analyze the possible cumulative characteristics of the RNA-chaperone activity of the recombinant OB-fold proteins we performed isothermal differential absorption spectroscopy (IDS) measurements (43). For that we exploited the UV-signature absorbances of the four G-quadruplex (GQ)-folds containing pre-edited A6-mRNA (4) and compared the dose-dependent RNA-chaperone reaction of TbMP81-FL with that of stoichiometric mixtures of TbMP81-FL and TbMP42-FL. The principle of the IDS-measurement is sketched out in Figure 4A. Figure 4B (left panel) shows the normalized difference spectra of pre-edited A6-mRNA in the absence of protein and in the presence of BSA (at a 5-fold molar excess) as an unrelated control. Figure 4B (central panel) shows the normalized difference spectra of pre-edited A6-mRNA in the presence of up to six molar equivalents of TbMP81-FL. The protein causes a concentration-dependent signal increase of the negative signal at 278 nm due to the chaperone-induced solvent exposure of the pre-mRNA (43). Mixtures of TbMP81-FL and TbMP42-FL behave qualitatively and quantitatively similar (Figure 4B, right panel). The addtion of either of the two proteins induces an stepwise increase of the signal and as expected, the individual increments are smaller for TbMP42-FL than for TbMP81-FL due to its smaller surface area. The signal increase in both cases is cumulative thereby demonstrating the additive characteristics of the RNA-chaperone activity.

Structural characterization of the RNA-remodeling activity

As a follow up of the results described above we asked the question how the RNA-remodeling activities of the different OB-fold proteins manifest on a structural level and how they compare to the structural changes catalyzed by 20S editosomes (4). For that, we mapped the chaperone-induced structure changes with nucleotide resolution using SHAPE (Selective 2′-Hydroxyl Acylation analyzed by Primer Extension) chemical probing (57,58). The experiments were performed with TbMP18-FL and TbMP81-FL as the smallest and the largest of the *T. brucei* OB-fold proteins and with pre-edited RPS12 and A6 as substrate mRNAs. Controls were performed with BSA as an unrelated protein and with 20S editosomes as a positive control. Figure 5A shows the SHAPE-profiles of the RPS12-RNA in the absence of protein and in the presence of BSA. The two modification traces are identical with Pearson correlation coefficients (r) ranging between 0.8 and 0.98 in different experiments. This is further evidenced in the difference (Δ)SHAPE-profile in Figure 5A: BSA as an unrelated protein does not alter the 2D-structure of the pre-edited RPS12-transcript even at a 10-fold molar excess over RNA. Figure 5B shows the SHAPE-reactivity profiles of RPS12-mRNA in the presence of TbMP18-FL and in the presence of 20S editosomes. As anticipated, TbMP18-FL induces an increase in the mean SHAPE-reactivity (18%), similar
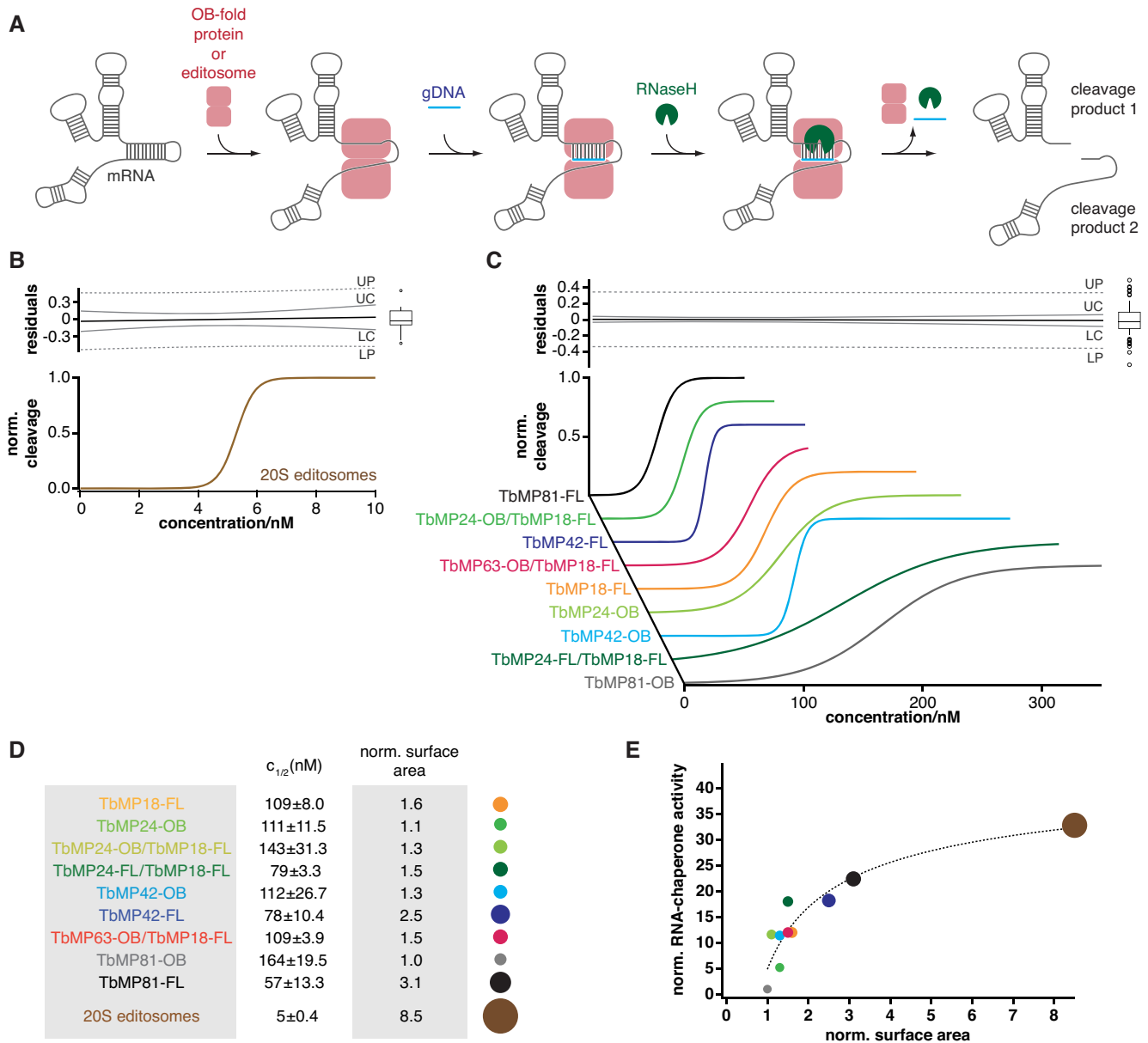
Figure 3. RNA-chaperone activities of the recombinant editosomal OB-fold proteins. (**A**) Sketch of the RNaseH-based RNA-chaperone assay as in ([4](#)). The drawing illustrates the OB-fold protein-induced change in RNA-structure, which allows the binding of a complementary DNA-oligonucleotide. The generated mRNA/DNA-hybrid is then cleaved by RNaseH to create two RNA-cleavage fragments. Whether the DNA-oligonucleotide annealing step is also mediated by the OB-fold proteins is currently not known. (**B**) Dose-dependent RNA-chaperone activity of 20S editosomes (0–10 nM) using pre-edited CYb-mRNA as a substrate mRNA. Individual data points (see Supplementary Figure S2) are fitted to a sigmoidal function to derive half-maximal cleavage concentrations ($c_{1/2}$) as a metric for the RNA-chaperone activity. Top panel: error analysis of the data. UC/LC: upper/lower confidence intervals, UP/LP: upper/lower prediction bands, black line: linear fit of all residuals. Box-plots of the data are shown on the right. (**C**) Dose-dependent RNA-chaperone activity of all recombinant OB-fold proteins and complexes using pre-edited CYb-mRNA as a substrate mRNA. Data fitting (see Supplementary Figure S2) and error analysis are as above. (**D**) Summary of all $c_{1/2}$-values (±SD) and normalized surface areas of the tested editosomal OB-fold constructs in comparison to 20S editosomes. (**E**) Plot of the normalized RNA-chaperone activities of all OB-fold proteins/complexes in relation to their normalized surface areas.

to the chaperone reaction of the editosome (26%). Twenty two percent of the nucleotide positions in the RPS12-transcript are responsive to the chaperone reaction and as evidenced in the ΔSHAPE-profile, the majority of affected nucleotides (71%) increase in flexibility. Again, this represents a value comparable to the remodeling activity of the 20S editosome (72%). Identical results were collected for

TbMP81-FL and the pre-edited A6-mRNA as another OB-fold protein/mRNA pair (Figure [5](#)C).

Using the SHAPE-reactivities as pseudo free energies we calculated minimal free energy (MFE)-structures for the two pre-mRNAs in both conformational states (plus/minus protein) (Figure [6](#)). As free RNAs, the two transcript are characterized by Gibbs free energies (ΔG) of –152 kcal/mol
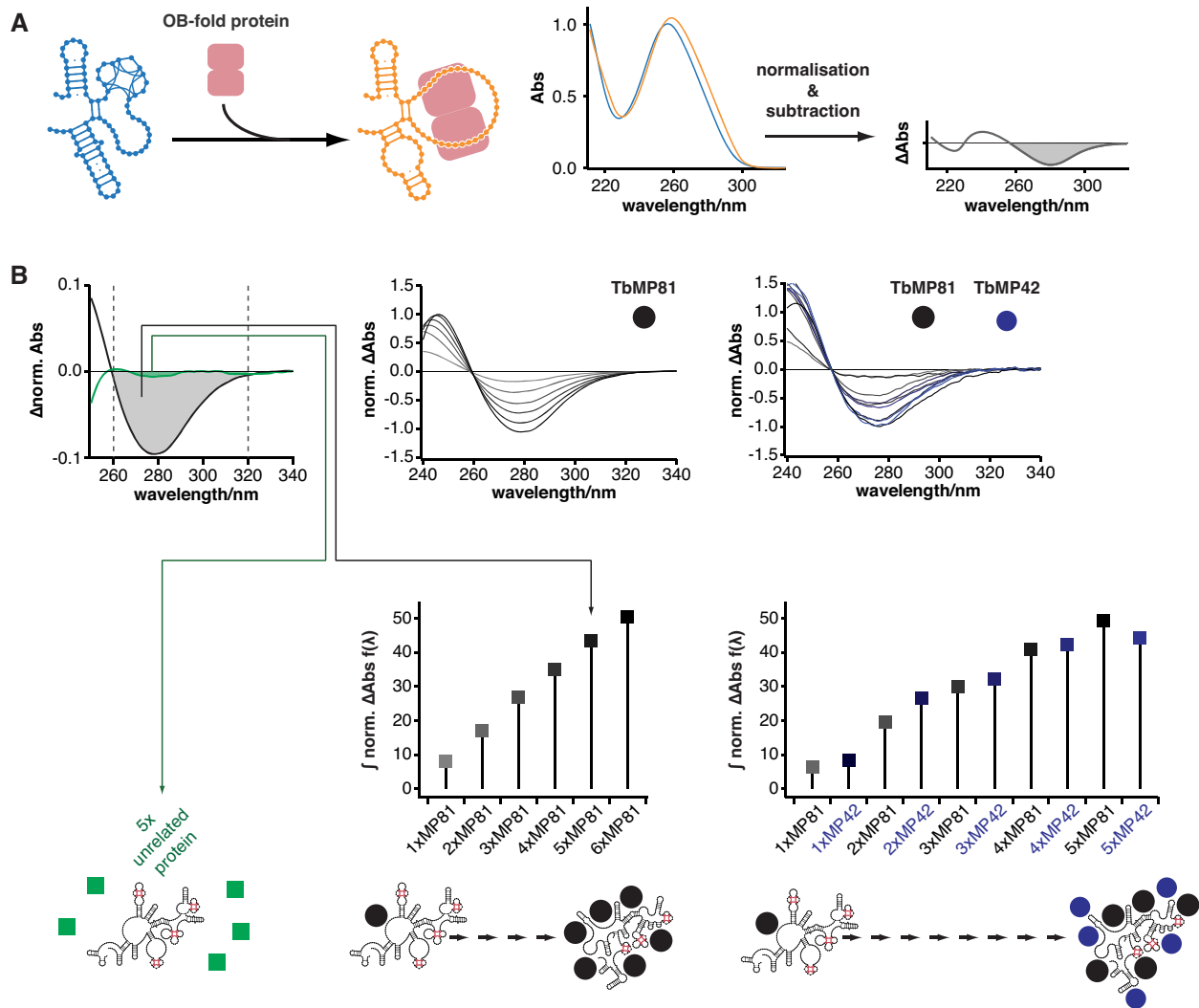
Figure 4. Cumulative characteristics of the RNA-chaperone activities of the *T. brucei* OB-fold proteins TbMP81-FL and TbMP42-FL. (**A**) Sketch and workflow of the isothermal differential absorption spectroscopy (IDS)-measurements (43). (**B**) Left panel: normalized ID-spectrum of pre-edited A6-mRNA in the absence of protein (black line). The dominant neagtive peak at app. 278 nm is shaded in grey. Green: ID-spectrum of pre-edited A6-mRNA in the presence of a 5-fold molar excess of BSA as a non-specific control. Please note the absence of any signal between 260nm and 320nm. Central panel: normalized ID-spectra of pre-edited A6-mRNA in the presence of TbMP81-FL up to a 6-fold molar excess over RNA. Right panel: normalized ID-spectra of pre-edited A6-mRNA in the presence of TbMP81-FL and TbMP42-FL. Bottom panel: quantification of the dose-dependent signal increase and sketch of the RNA/protein interaction. The data demonstrate the cumulative characteristics of the chaperone activity and provide an additional, independent measurement for the RNA-chaperone activity of the proteins.

(RPS12-RNA) and –137 kcal/mol (A6-RNA). In the presence of protein they adopt 2D-folds of only –122 kcal/mol (RPS12-RNA) and –111 kcal/mol (A6-RNA). This demonstrates a structural destabilization with $\Delta\Delta G$'s in the range of 25–30 kcal/mol, which compares favorable to the destabilization induced by 20S editosomes (25 kcal/mol for the RPS12-RNA and 24 kcal/mol for the A6-transcript). A comparison of all probed structures (Figure 6) identified that ∼70% of the structural details are shared between the RNA-folds in the presence and absence of proteins. Differences map predominantly to the 5'- and 3'-termini in both RNAs, to a pseudoknot in the RPS12-transcript and to one of the GQ-elements in the A6-mRNA. Together, the SHAPE-probing data corroborate the results of the RNaseH-based RNA-chaperone assay and demonstrate

that both, TbMP18-FL and TbMP81-FL refold and destabilize pre-edited transcripts with qualitative and quantitative characteristics similar to 20S editosomes.

Modeling the RNA-refolding domain(s) of the 20S editosome

Lastly, we asked the question whether the collected data can be used to derive a structural model of the RNA-refolding domain(s) of the *T. brucei* 20S editosome. For that we performed computational modeling using the program PyRy3D (http://genesilico.pi/pyry3d/) specifically designed to calculate low-resolution models of high molecular mass complexes that contain intrinsic disorder. The program allows the usage of experimentally as well as computationally-derived atomic coordinates together with flexible shapes to include disordered substructures. It has been used to model
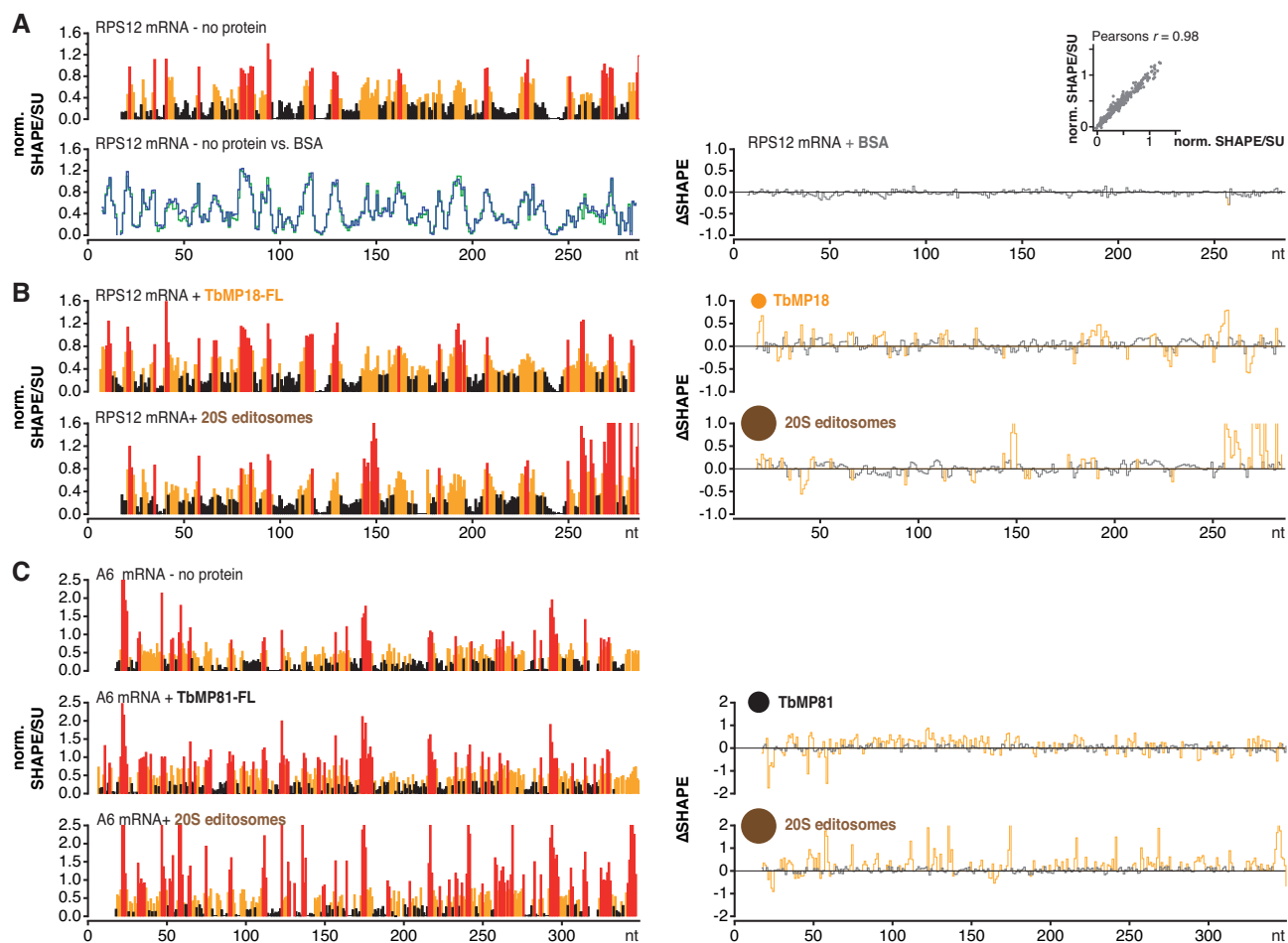
Figure 5. Mapping the structural consequences of the RNA-chaperone activity by SHAPE-chemical modification. (**A**) Upper panel: Normalized SHAPE-reactivity profile of pre-edited RPS12 mRNA in the absence of protein. Black: unreactive (SU < 0.35), yellow: moderately reactive (0.35 ≤ SU < 0.8), red: highly reactive (SU ≤ 0.8) nucleotide positions. SU = SHAPE-unit. Lower panel: Overlay plot of the SHAPE-reactivity profiles of RPS12-RNA in the absence of protein (green) and in the presence of BSA as an unrelated protein (blue). Right panel: Difference (△)SHAPE-reactivity profile and Pearson plot of the data showing a correlation coefficient (*r*) of 0.98. (**B**) Normalized SHAPE-reactivity profiles of RPS12-mRNA in the presence of TbMP18-FL (upper panel) and in the presence of 20S editosomes (lower panel). Colours are as above. Panel to the right: difference (△)SHAPE-reactivity plots. Yellow: signals ≥|0.2|SU. Gray: signals <|0.2|SU. (**C**) Normalized SHAPE-reactivity profiles of A6 pre-mRNA in the absence of protein (upper panel), in the presence of TbMP81-FL (middle panel) and in the presence of 20S editosomes (lower panel). Colours are as above. Panel to the right: △SHAPE-reactivity profiles. Yellow: signals ≥|0.2|SU. Gray: signals <|0.2|SU.

the structures of several proteins and complexes involved in RNA metabolism including the CCR4-NOT complex (59). PyRy3D-simulations rely on spatial restraints to implement experimentally-derived interaction data and a scoring function fits the individual components into the contour map of the complex. For that we used the cryo-EM structure of the *T. brucei* 20S editosome (49). As spatial restraints we implemented all published binary interactions of editosome components including the chemical cross-linking data of the *T. brucei* OB-fold proteins (21,31,32) (see Supplementary Table S4). Structure coordinates for the individual proteins were taken from Czerwoniec *et al.* (19 and references therein) and all disordered regions were simulated as coarse-grained flexible shapes. Three hundred independent simulations (300 000 steps each) were performed and the resulting 300 models were clustered as outlined in the Materials and Methods section. Figure 7A,B show the medoid structure of the largest cluster (54 models). The different proteins fit

tightly into the available volume of the 20S editosome EM-density map with a cross-correlation coefficient of 0.64. Interestingly, the OB-fold proteins cluster in a band-like arrangement in the center of the cryo-EM structure (Figure 7C) thereby bridging the two quasi global sub-domains of the editosome (49). This area coincides nicely with the identified high-affinity substrate RNA binding site of the complex visualized by TEM (3). A dissection of the structurally well-defined parts of the model from the positions of the intrinsically disordered domains of the different proteins is shown in Figure 8A and B. This identifies that structured protein domains are preferentially located inside the 0.8MDa complex, while all IDP-domains seem to be preferentially located in peripheral regions of the particle (Figure 8C). This suggests a general bi-partite domain composition of the editosome: an outer shell of intrinsically disordered proteins, which act as RNA-contact and RNA-remodeling
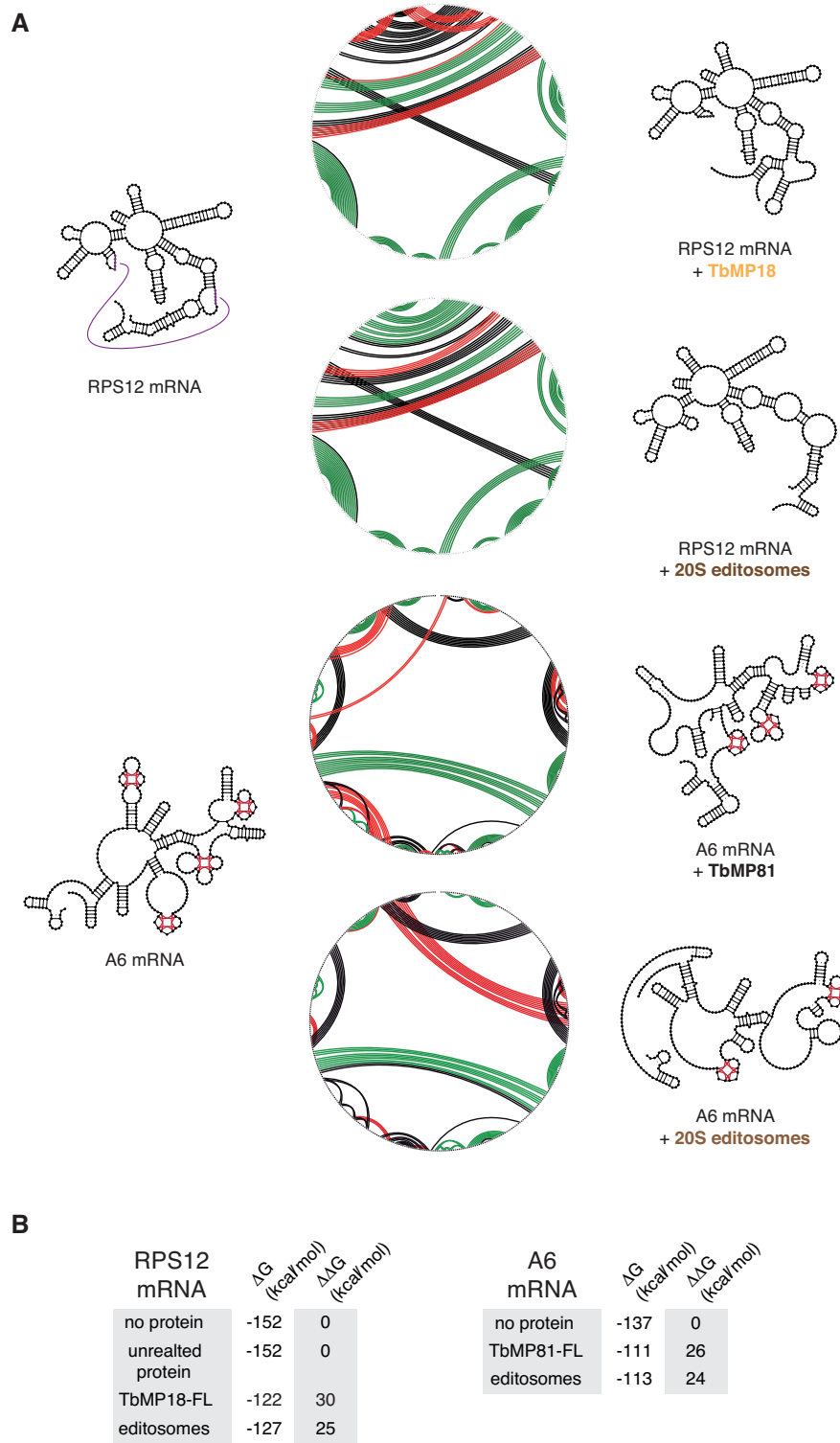
A

RPS12 mRNA

RPS12 mRNA
+ TbMP18

RPS12 mRNA
+ **20S editosomes**

A6 mRNA

A6 mRNA
+ **TbMP81**

A6 mRNA
+ **20S editosomes**

B

RPS12 mRNA	ΔG (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
no protein	-152	0
unrealted protein	-152	0
TbMP18-FL	-122	30
editosomes	-127	25

A6 mRNA	ΔG (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
no protein	-137	0
TbMP81-FL	-111	26
editosomes	-113	24

Figure 6. (**A**) SHAPE-derived minimum free energy (MFE) 2D-structures and circle-plots of pre-edited RPS12- and A6-mRNAs in the presence/absence of TbMP18-FL, TbMP81-FL and 20S editosomes. GQ-folds are shown as 'leaf-like' structures in red and a pseudoknot in the pre-edited RPS12-mRNA is annotated as a purple line. Base-pairs (bp) in the circle-plots are represented as colored lines: Black: bp unique to the minus protein/editosome structure. Red: bp unique to the plus protein/editosome structure. Green: bp present in both, plus/minus protein/editosome. (**B**) Gibbs free energies (ΔG) and $\Delta\Delta G$-values were calculated for a temperature of 27°C and have a standard deviation (SD) of ±4 kcal/mol.
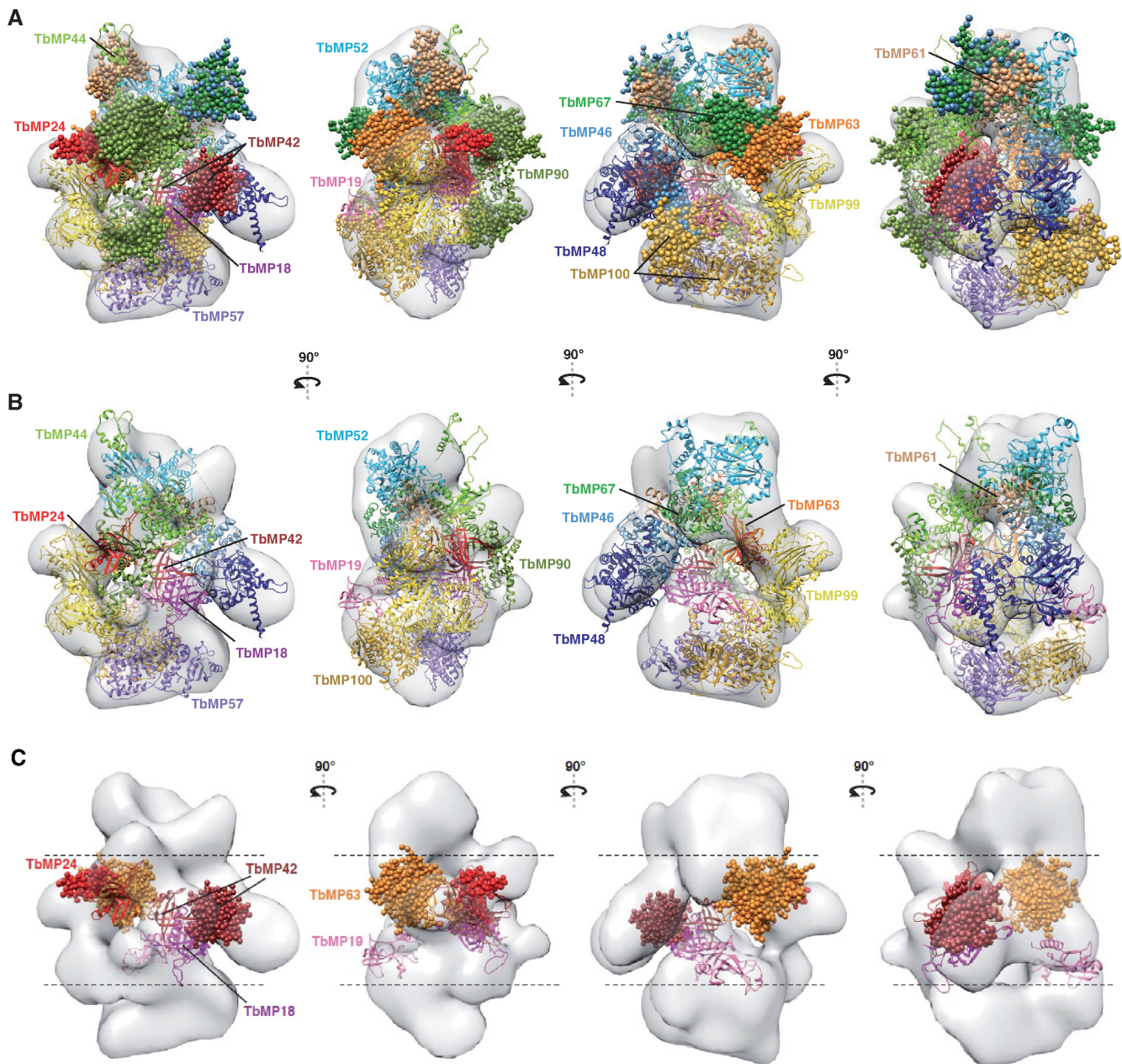
Figure 7. Structure model of the *T. brucei* 20S editosome. (**A**) Rotational images (90°) of the PyRy3D-derived medoid structure of the 20S editosome. The cryo-EM-derived shape of the editosome (49) is shown as a gray surface in the background at a contour level of 0.0419. Individual proteins are colour coded and labelled. Intrinsically disordered protein domains are shown as clustered spheres with each sphere representing a Cα-position (diameter 0.19nm). All distance restraints used in the calculation are listed in Supplementary Table S4. PyRy3D can be accessed at (http://genesilico.pi/pyry3d/). The model has been deposited at figshare (DOI: 10.6084/m9.figshare.6860705.v3). (**B**) 20S editosome model as in (A) omitting all intrinsically disordered protein domains for clarity. (**C**) Localization of the OB-fold proteins TbMP18, TbMP19, TbMP24, TbMP42 and TbMP63. The different proteins are colour coded and labelled. Intrinsically disordered regions are shown as clustered spheres as above. Dashed lines mark the central, band-like zone in which the different proteins are clustered.

elements and an inner core of structurally well-defined proteins, which mediate the catalytic reactions of the complex.

DISCUSSION

Editosomes execute a complex-inherent RNA-chaperone activity to remodel the highly folded structures of mitochondrial substrate pre-mRNAs (3,4). The activity has been characterized to simplify the folding landscape of the different RNAs to facilitate the annealing of gRNAs as templates in the reaction. While several of the protein components of the editosome have been identified to catalyze defined steps of the editing cycle (1) no editosomal protein has as of yet been recognized to mediate RNA-chaperone functionality.

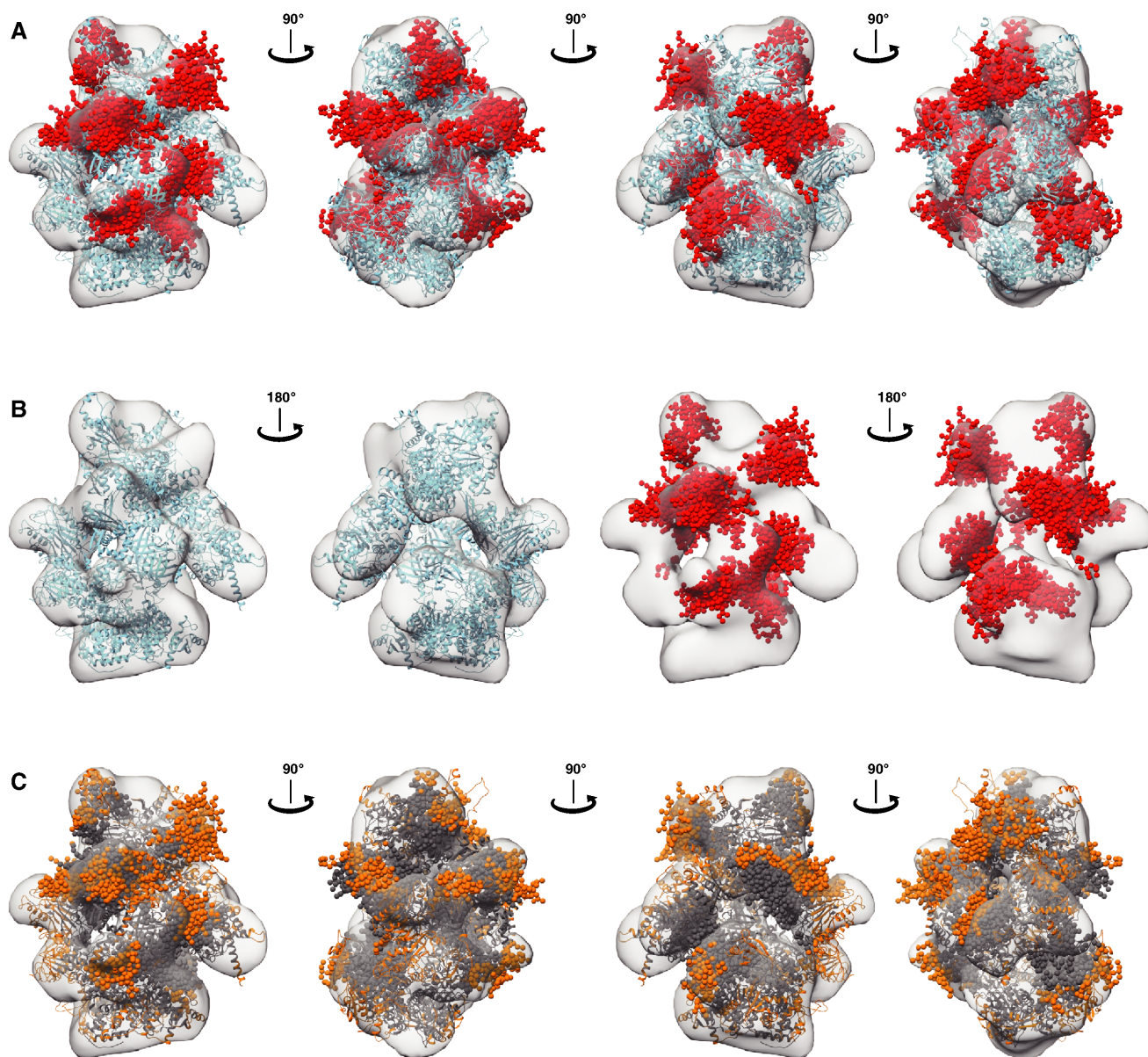Here, we provide evidence that the OB-fold proteins of the editosome execute RNA-chaperone activity. For that

Figure 8. Order/disorder distribution in the 20S editosome. (**A**) Rotational images (90°) of the PyRy3D-derived 20S editosome model specifically emphasizing the intrinsically disordered regions as clustered spheres in red and structurally ordered regions in light blue. The cryo-EM-based shape of the complex (49) is shown as a gray surface (contour level 0.0419). (**B**) Rotational images (180°) of the editosome model showing only the ordered (light blue) or disordered (red) regions of the complex. Ordered protein domains have a surface/core-ratio of 30/70. IDP-domains are characterized by a value of 40/60. (**C**) Rotational images (90°) of the editosome model emphasizing surface-localized protein domains in orange and structurally 'buried' residues in grey.

we generated recombinant versions of the different proteins and identified that they are highly unstructured. Furthermore, our CD-measurements support computational predictions suggesting that the OB-fold proteins of the *T. brucei* editosome are in large parts intrinsically disordered (19). The disordered regions not only locate outside of the OB-fold domains, they can also be found in the loop regions of the different OB-folds and here specifically in the loop sequences between β-sheets β2 and β3 (L_{23}) (Supplementary Figure S1). We also confirmed previous observations (21,31,32) that the different proteins have a high propensity to homo- and hetero-oligomerize using their OB-folds as in-

teraction domains. Interestingly, all identified pair-wise interactions involve TbMP18, the smallest of the editosomal OB-fold proteins, possibly implicating a key role in the assembly of the editosomal complex.

Importantly, all homo- and hetero-oligomeric complexes showed RNA-chaperone activity. The activities correlate with the surface size of the different proteins indicating that the process is primarily surface-driven with perhaps cumulative characteristics: every OB-fold protein provides a defined value of RNA-chaperone activity, which together generates the overall activity of the editosome. This is supported by the fact that editing-active editosome

preparations show the highest activity. In addition, all FL-constructs display higher chaperone activities in comparison to the OB-fold-only polypeptides suggesting that the chaperone activity is primarily located within the disordered regions of the proteins. This represents a different situation to the yeast RRP44 protein, where the RNA-unwinding reaction is mediated by a cluster of three OB-folds (33,34). However, as anticipated, the chaperone reaction induces a destabilization of the substrate RNAs, thereby generating RNA-minimal free energy structures of reduced thermodynamic stability. As such, the individual OB-fold proteins execute a chaperone reaction that is biochemically and structurally equivalent to the activity of the editosomal complex.

As a consequence of the described results, we propose a scenario in which the OB-folds and the IDP-sequences of the different proteins execute separate but interdependent functions: The OB-folds primarily act as docking modules to assemble within the editosome as previously discussed (32). This positions the IDP-sequences of the different proteins on the surface of the catalytic complex where they function to bind and refold mitochondrial pre-mRNAs. Thus, our data add to the functional assignment of the protein inventory of the editosome and within this context it is important to note that about half of the 12 core proteins of the editosome are allocated to remodel RNA. This indicates the importance of the process within the editing reaction cycle. Of course, we can not exclude that IDP-sequences of other, non-OB-fold editosomal proteins such as TbMP90, TbMP67, TbMP61 or TbMP100 might contribute as well (19).

Intrinsic disorder is a common feature of many RNA-binding proteins (60,61). It provides the functional advantage of allowing multiple contacts during the initial binding reaction while at the same time enabling sufficient conformational flexibility to target different RNA-ligands (61). In addition, intrinsic disorder has a kinetic advantage since an increased capture radius results in a higher binding probability. This mechanism is known as 'fly-casting mechanism' (62). It accounts for the fact that structurally flexible protein sequences can enlarge their effective 'reach', which is of special importance for membrane-bound macromolecular complexes especially in crowded solvent conditions. RNA-editing takes place in the highly crowded mitochondria of African trypanosomes and preliminary evidence for an attachment of the editosome to mitochondrial ribosomes and/or the mitochondrial inner-membrane have been discussed (63,64). We propose that the energy for the RNA-refolding reaction carried out by the editosome follows an 'entropy transfer'-mechanism (60). The binding of a substrate RNA induces a structural re-orientation in the OB-fold proteins, which results in a loss of conformational entropy. This entropy is used to remodel the bound RNA in agreement with the observation that the reaction does not require ATP (3). Our data further suggest that the pre-edited substrate mRNAs likely interact with the editosome by two successive RNA-binding modes. As in the case of the CBP2/bI5 group I intron interaction (65) an initial, rather non-specific interaction induces a set of conformational fluctuations in the bound RNA, which is followed by a slow, specific binding mode that stabilizes the editing-

competent RNA conformation. Support for such a scenario comes from the observation that flexibility restrictions in model editing substrate RNAs inhibit RNA editing *in vitro* (66).

As a consequence of our hypothesis we postulate that the IDP-domains of the different OB-fold proteins are located on the surface of the 20S editosome. To provide a structural context for the assumption we performed a 'coarse-grained' modeling of the structure of the *T. brucei* editosome using published cryo-EM data (49) and all OB-fold interaction data as spatial restraints (21,31,32). Interestingly, while the model agrees with the anticipated surface location of the different IDP-domains on the editosome, it also suggests a clustering of all structurally well-defined proteins in the center of the complex. Thus, the complex separates an inner core of catalytically active editosome components from an outer shell of ID-proteins that possibly act as RNA-contact and RNA-remodeling elements. Interestingly, a similar separation of structurally defined protein regions from intrinsically disordered domains was also identified for the human spliceosome (67). Thus, the described organization might reflect a more general structural scenario of macromolecular RNP-machineries.

Lastly, we would like to emphasize that it will be necessary to experimentally validate the RNA-chaperone activity of the different OB-fold proteins *in vivo*. In addition, it will be mandatory to study the crosstalk and perhaps fine-tuning of the activity by all complexes known to interact with the editosome as part of the reaction cycle. This includes the MRB1/GRBC-complex (69–72), the polyadenylation machinery as well as the ribosome (73). Using the well studied bacterial RNA-chaperone Hfq as a benchmark (74) it should be possible to decipher the intricate interplay of structured and disordered protein domains of the editosomal OB-fold proteins to execute an adaptable and dynamic chaperone response.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Göringer,H.U. (2012) 'Gestalt', composition and function of the *Trypanosoma brucei* editosome. *Annu. Rev. Microbiol.*, **66**, 65–82.
2. Aphasizhev,R. and Aphasizheva,I. (2014) Mitochondrial RNA editing in trypanosomes: small RNAs in control. *Biochimie*, **100**, 125–131.
3. Böhm,C., Katari,V.S., Brecht,M. and Göringer,H.U. (2012) *Trypanosoma brucei* 20S editosomes have one RNA substrate-binding site and execute RNA unwinding activity. *J. Biol. Chem.*, **287**, 26268–26277.
4. Leeder,W.M., Voigt,C., Brecht,M. and Göringer,H.U. (2016) The RNA chaperone activity of the *Trypanosoma brucei* editosome raises the dynamic of bound pre-mRNAs. *Sci. Rep.*, **6**, 19309.
5. Leeder,W.M., Hummel,N.F. and Göringer,H.U. (2016) Multiple G-quartet structures in pre-edited mRNAs suggest evolutionary driving force for RNA editing in trypanosomes. *Sci. Rep.*, **6**, 29810.
6. Rajkowitsch,L., Chen,D., Stampfl,S., Semrad,K., Waldsich,C., Mayer,O., Jantsch,M.F., Konrat,R., Bläsi,U. and Schroeder,R. (2007) RNA chaperones, RNA annealers and RNA helicases. *RNA Biol.*, **4**:118–130.
7. Semrad,K. (2011) Proteins with RNA chaperone activity: A world of diverse proteins with a common task-impediment of RNA misfolding. *Biochem. Res. Int.*, **2011**, 532908.
8. Bernacchi,S., Stoylov,S., Piémont,E., Ficheux,D., Roques,B.P., Darlix,J.L. and Mély,Y. (2002) HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J. Mol.Biol.*, **317**, 385–399.
9. Zhang,A., Derbyshire,V., Salvo,J.L. and Belfort,M. (1995) *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly *in vitro*. *RNA*, **1**, 783–793.
10. Moll,I., Leitsch,D., Steinhauser,T. and Bläsi,U. (2003) RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep.*, **4**, 284–289.
11. Semrad,K., Green,R. and Schroeder,R. (2004) RNA chaperone activity of large ribosomal subunit proteins from *Escherichia coli*. *RNA*, **10**, 1855–1860.
12. Murzin,A.G. (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861–867.
13. Bae,W., Xia,B., Inouye,M. and Severinov,K. (2000) *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 7784–7789.
14. Nakaminami,K., Karlson,D.T. and Ima,R. (2006) Functional conservation of cold shock domains in bacteria and higher plants. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 10122–10127.
15. Duval,M., Korepanov,A., Fuchsbauer,O., Fechter,P., Haller,A., Fabbretti,A., Choulier,L., Micura,R., Klaholz,B.P., Romby,P. *et al.* (2013) *Escherichia coli* ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol.*, **11**, e1001731.
16. Croitoru,V., Semrad,K., Prenninger,S., Rajkowitsch,L., Vejen,M., Laursen,B.S., Sperling-Petersen,H.U. and Isaksson,L.A. (2006) RNA chaperone activity of translation initiation factor IF1. *Biochimie*, **88**, 1875–1882.
17. Worthey,E.A., Schnaufer,A., Mian,I.S., Stuart,K. and Salavati,R. (2003) Comparative analysis of editosome proteins in trypanosomatids. *Nucleic Acids Res.*, **31**, 6392–6408.
18. Stuart,K.D., Schnaufer,A., Ernst,N.L. and Panigrahi,A.K. (2005) Complex management: RNA editing in trypanosomes. *Trends Biochem. Sci.*, **30**, 97–105.
19. Czerwoniec,A., Kasprzak,J.M., Bytner,P., Dobrychłop,M. and Bujnicki,J.M. (2015) Structure and intrinsic disorder of the proteins of the *Trypanosoma brucei* editosome. *FEBS Lett.*, **589**, 2603–2610.
20. Kozlowski,L.P. and Bujnicki,J.M. (2012) MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinformatics*, **13**, 111.
21. McDermott,S.M., Luo,J., Carnes,J., Ranish,J.A. and Stuart,K. (2016) The architecture of *Trypanosoma brucei* editosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E6476–E6485.
22. Tarun,S.Z. Jr, Schnaufer,A., Ernst,N.L., Proff,R., Deng,J., Hol,W. and Stuart,K. (2008) KREPA6 is an RNA-binding protein essential for editosome integrity and survival of *Trypanosoma brucei*. *RNA*, **14**, 347–358.
23. Salavati,R., Ernst,N.L., O'Rear,J., Gilliam,T., Tarun,S. Jr and Stuart,K. (2006) KREPA4, an RNA binding protein essential for editosome integrity and survival of *Trypanosoma brucei*. *RNA*, **12**, 819–831.
24. Brecht,M., Niemann,M., Schlüter,E., Müller,U.F., Stuart,K. and Göringer,H.U. (2005) TbMP42, a protein component of the RNA editing complex in African trypanosomes, has endo-exoribonuclease activity. *Mol. Cell*, **17**, 621–630.
25. Niemann,M., Brecht,M., Schlüter,E., Weitzel,K., Zacharias,M. and Göringer,H.U. (2008) TbMP42 is a structure-sensitive ribonuclease that likely follows a metal ion catalysis mechanism. *Nucleic Acids Res.*, **36**, 4465–4473.
26. Kala,S. and Salavati,R. (2010) OB-fold domain of KREPA4 mediates high-affinity interaction with guide RNA and possesses annealing activity. *RNA*, **16**, 1951–1967.
27. Drozdz,M., Palazzo,S.S., Salavati,R., O'Rear,J., Clayton,C. and Stuart,K. (2002) TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J.*, **22**, 1791–1799.
28. Huang,C.E., O'Hearn,S.F. and Sollner-Webb,B. (2002) Assembly and function of the RNA editing complex in *Trypanosoma brucei* requires band III protein. *Mol. Cell. Biol.* **22**, 3194–3203.
29. Guo,X., Ernst,N.L. and Stuart,K.D. (2008) The KREPA3 zinc finger motifs and OB-fold domain are essential for RNA editing and survival of *Trypanosoma brucei*. *Mol. Cell. Biol.*, **28**, 6939–6953.
30. Law,J.A., O'Hearn,S. and Sollner-Webb,B. (2007) In *Trypanosoma brucei* RNA editing, TbMP18 (Band VII) is critical for editosome integrity and for both insertional and deletional cleavages. *Mol. Cell. Biol.*, **27**, 777–787.
31. Schnaufer,A., Wu,M., Park,Y.J., Nakai,T., Deng,J., Proff,R., Hol,W.G. and Stuart,K.D.A. (2010) Protein-protein interaction map of trypanosome 20S editosomes. *J. Biol. Chem.*, **285**, 5282–5295.
32. Park,Y.J., Pardon,E., Wu,M., Steyaert,J. and Hol,W.G. (2012) Crystal structure of a heterodimer of editosome interaction proteins in complex with two copies of a cross-reacting nanobody. *Nucleic Acids Res.*, **40**, 1828–1840.
33. Lorentzen,E., Basquin,J., Tomecki,R., Dziembowski,A. and Conti,E. (2008) Structure of the active subunit of the yeast exosome core, Rrp44. Diverse modes of substrate recruitment in the RNase II nuclease family. *Mol. Cell*, **29**, 717–728.
34. Bonneau,F., Basquin,J., Ebert,J., Lorentzen,E. and Conti,E. (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell*, **139**, 547–559.
35. Cross,G.A. (1975) Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology*, **71**, 393–417.
36. Claros,M.G. (1995) MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput. Appl. Biosci.*, **11**, 441–447.
37. Johnson,W.C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins*, **35**, 307–312.
38. Sreerama,N. and Woody,R.W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.*, **287**, 252–260.
39. Whitmore,L. and Wallace,B.A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.*, **32**, W668–W673.
40. Whitmore,L. and Wallace,B.A. (2007) Protein secondary structure analysis from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*, **89**, 392–400.

41. Jacobsen,C. and Steensgaard,J. (1979) Measurements of precipitin reactions by difference turbidimetry: a new method. *Immunology*, **36**, 293–298.

42. Deverill,I. and Reeves,W.G. (1980) Light scattering and absorption – developments in immunology. *J. Immunol. Methods*, **38**, 191–204.

43. Gray,R.D. and Chaires,J.B. (2012) Isothermal folding of G-quadruplexes. *Methods*, **57**, 47–55.

44. Savitzky,A. and Golay,M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.*, **36**, 1627–1639

45. Vasa,S.M., Guex,N., Wilkinson,K.A., Weeks,K.M. and Giddings,M.C. (2008) ShapeFinder: a software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA*, **14**, 1979–1990.

46. Deigan,K.E., Li,T.W., Mathews,D.H. and Weeks,K.M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 97–102.

47. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 129.

48. Hajdin,C.E., Bellaousov,S., Huggins,W., Leonard,C.W., Mathews,D.H. and Weeks,K.M. (2013) Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc. Nat. Acad. Sci. U.S.A.*, **110**, 5498–5503.

49. Golas,M.M., Böhm,C., Sander,B., Effenberger,K., Brecht,M., Stark,H. and Göringer,H.U. (2009) Snapshots of the RNA editing machine in trypanosomes captured at different assembly stages *in vivo*. *EMBO J.*, **28**, 766–778.

50. Metropolis,N. and Ulam,S. (1949) The Monte Carlo method. *J. Am. Stat. Assoc.*, **44**, 335–341.

51. Li,F., Ge,P., Hui,W.H., Atanasov,I., Rogers,K., Guo,Q., Osato,D., Falick,A.M., Zhou,Z.H. and Simpson,L. (2009) Structure of the core editing complex (L-complex) involved in uridine insertion/deletion RNA editing in trypanosomatid mitochondria. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 12306–12310.

52. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.

53. Garzon,J.I., Kovacs,J., Abagyan,R. and Chacon,P. (2007) ADP_EM: fast exhaustive multi- resolution docking for high-throughput coverage. *Bioinformatics*, **23**, 427–433.

54. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.

55. Pintilie,G.D., Zhang,J., Goddard,T.D., Chiu,W. and Gossard,D.C. (2010) Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J. Struct. Biol.*, **170**, 427–438.

56. Wu,Z., Hu,G., Yang,J., Peng,Z., Uversky,V.N. and Kurgan,L. (2015) In various protein complexes, disordered protomers have large per-residue surface areas and area of protein-, DNA-, RNA-binding interfaces. *FEBS Lett.*, **589**, 2561–2569.

57. Merino,E.J., Wilkinson,K.A., Coughlan,J.L. and Weeks,K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.*, **127**, 4223–4231.

58. Low,J.T. and Weeks,K.M. (2010) SHAPE-directed RNA secondary structure prediction. *Methods*, **52**, 150–158.

59. Ukleja,M., Cuellar,J., Siwaszek,A., Kasprzak,J.M., Czarnocki-Cieciura,M., Bujnicki,J.M., Dziembowski,A. and Valpuesta,J.M. (2016) The architecture of the *Schizosaccharomyces pombe*CCR4-NOT complex. *Nat. Commun.*, **7**, 10433.

60. Tompa,P. and Csermely,P. (2004) The role of structural disorder in the function of RNA and protein chaperones. *FASEB J.*, **18**, 1169–1175.

61. Varadi,M., Zsolyomi,F., Guharoy,M. and Tompa,P. (2015) Functional advantages of conserved intrinsic disorder in RNA-binding proteins. *PLoS One*, **10**, e0139731.

62. Shoemaker,B.A., Portman,J.J. and Wolynes,P.G. (2000) Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 8868–8873.

63. Aphasizheva,I., Maslov,D., Wang,X., Huang,L. and Aphasizhev,R. (2011) Pentatricopeptide repeat proteins stimulate mRNA adenylation/uridylation to activate mitochondrial translation in trypanosomes. *Mol. Cell*, **42**, 106–117.

64. Katari,V.S., van Esdonk,L. and Göringer,H.U. (2013) Molecular crowding inbitis U-insertion/-deletion RNA-editing *in vitro*: consequences for the *in vivo* reaction. *PLoS One*, **12**, e83796.

65. Bokinsky,G., Nivón,L.G, Liu,S., Chai,G., Hong,M., Weeks,K.M. and Zhuang,X. (2006) Two distinct binding modes of a protein cofactor with its target RNA. *J. Mol. Biol.*, **361**, 771–784.

66. Leeder,W.M., Reuss,A.J., Brecht,M., Kratz,K., Wachtveitl,J. and Göringer,H.U. (2015) Charge reduction and thermodynamic stabilization of substrate RNAs inhibit RNA editing. *PLoS One*, **10**, e0118940.

67. Korneta,I. and Bujnicki,J.M. (2012) Intrinsic disorder in the human spliceosomal proteome. *PloS Comput. Biol.*, **8**, e1002641.

68. McGuffin,L.J., Bryson,K. and Jones,D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404–405.

69. Panigrahi,A.K., Zíková,A., Dalley,R.A., Acestor,N., Ogata,Y., Anupama,A., Myler,P.J. and Stuart,K.D. (2008) Mitochondrial complexes in *Trypanosoma brucei*: a novel complex and a unique oxidoreductase complex. *Mol. Cell. Proteomics*, **7**, 534–545.

70. Hashimi,H., Zimmer,S.L., Ammerman,M.L., Read,L.K. and Lukeš,J. (2013) Dual core processing: MRB1 is an emerging kinetoplast RNA editing complex. *Trends Parasitol.*, **29**, 91–99.

71. Aphasizheva,I., Zhang,L., Wang,X., Kaake,R.M., Huang,L., Monti,S. and Aphasizhev,R. (2014) RNA binding and core complexes constitute the U-insertion/deletion editosome. *Mol. Cell. Biol.*, **34**, 4329–4342.

72. Weng,J., Aphasizheva,I., Etheridge,R.D., Huang,L., Wang,X., Falick,A.M. and Aphasizhev,R. (2008) Guide RNA-binding complex from mitochondria of trypanosomatids. *Mol. Cell*, **32**, 198–209.

73. Aphasizheva,I. and Aphasizhev,R. (2016) U-Insertion/deletion mRNA-editing holoenzyme: definition in sight. *Trends Parasitol.*, **32**, 144–156.

74. Santiago-Frangos,A. and Woodson,S.A. (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip. Rev. RNA*, **9**, e1475.

75. Piovesan,D., Tabaro,F., Mičetić,I., Necci,M., Quaglia,F., Oldfield,C.J., Aspromonte,M.C., Davey,N.E., Davidović,R., Dosztányi,Z. *et al.* (2017) DisProt 7.0: a major update of the database of disordered proteins. *Nucleic Acids Res.*, **45**, D1123–D1124.

76. The UniProt Consortium. (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res.*, **45**, D158–D169.