Supplemental Experimental Procedures

CD spectroscopy

Proteins analyzed by CD spectroscopy (Mss116p, Mss116p/Δ598-664, Mss116p/Δ569-664, and Mss116p/Δ551-664) were purified in the same manner as Mss116p/Δ598-664, except that the final step was gel filtration through an S-200 column in the buffer used for CD measurements (10 mM Tris-HCl, pH 7.5, 500 mM KCl, and 10% glycerol). Proteins (200 µl in a 1-mm quartz cuvette; Mss116p, 2.5 µM; Mss116p/Δ598-664, 3 µM; Mss116p/Δ569-664, 2.5 µM; and Mss116p/Δ551-664, 2.5 µM) were analyzed with a Jasco J-815 CD spectropolarimeter with a Peltier attachment for temperature control. In melting experiments, the temperature was ramped from 20 to 90˚C, CD scans from 200 to 260 nm were recorded every 5˚C, and the CD at 220 nM was monitored continuously. The parameters used were as follows: data pitch, 0.1 nm; bandwidth, 1 nm; detector response time, 4 sec; and scanning speed, 50 nm/min. The temperature ramp was a step gradient with 4-min long CD scans that occurred every 5˚C and a 1˚C/min ramp in between. The melting data at 220 nm were fit by modeling a two-state transition of a monomer from a folded to unfolded state with the following equation:

$$f = \left\lfloor \frac{(u-l)*\exp((h/1.987*t))*((t/m)-1)}{1+\exp((h/(1.987*t))*((t/m)-1))} \right\rfloor + l$$

where u = mean residue ellipticity of the folded protein; l = mean residue ellipticity of the unfolded protein; h = starting enthalpy in cal/mol; t = temperature in Kelvin; and m = $T_{1/2}$ in Kelvin (Greenfield, 2006). We note that the unfolding of Mss116p was irreversible, likely due to aggregation of the unfolded protein, and $T_{1/2}$ is the temperature at the midpoint of the thermal transition.

Supplemental Figure Legends

Figure S1. Structure-based sequence alignment of Mss116p and other DEAD-box proteins. Structural information for Mss116p (PDB ID 3i5x; this work), Vasa (2db3), DDX19 (3fht) and eIF4AIII (2hyi) was used to align the sequences. The secondary structure elements (as defined by PyMOL) and residue numbers for Mss116p are shown above the alignment. Downward arrows indicate boundaries between domain 1, domain 2, and the CTE. Residues that contact ATP analogs and RNA are colored blue and red, respectively. The positions of the eleven conserved DEAD-box protein motifs (colored as in Figure 2B) are indicated below the alignment. The alignment was produced by combining the output of T-Coffee (3DCoffee option; O'Sullivan et al., 2004) and Dali (pairwise option; Holm et al., 2008).

Figure S2. Locating the Br atom in the Br-U_{10} complex. Anomalous difference density between the Br-U_{10}/AMP-PNP and the U_{10}/AMP-PNP complex data contoured at 5 σ (magenta) on the model of Mss116p viewed as in Figure 2. At this contour, a single peak exists corresponding to an atom bonded to the 5 position of a uridine base (inset), confirming that this is the 5-BrU at position 4 in the U_{10} oligonucleotide. At lower contours, no peaks were observed at the 5 positions for either U3 or U5, implying that the oligonucleotide does not bind in more than one register.

Figure S3. Electron density for the RNA in structures of Mss116p with different bound ATP analogs. Stereoviews of the RNA-binding site showing sigmaA-weighted $m|F_o| - D|F_c|$ electron density (black) contoured at 2 σ for the (A) AMP-PNP, (B) ADP-BeF_3^-, and (C) ADP-AlF_4^- complexes, and (D) at 1.8 σ for the Br-U_{10} complex. The RNA was removed from the model prior to map calculation. The view is similar to that in Figure 4. In the AMP-PNP, ADP-BeF_3^-, and ADP-AlF_4^- complexes, the electron density was continuous for U3-U8 and for the ribose and phosphate of U2, with poorer density for U9 and U10. The Br-U_{10} complex provided continuous

density for U3-U10 and the ribose and phosphate of U2. The ribose of U3 appeared to have two possible conformations in the AMP-PNP, ADP-BeF_3^-, and ADP-AlF_4^- complexes, but we modeled only one conformation for each complex. In all four complexes, the base of U2 was disordered, and there was density for the base of U1, but the conformation of the U1 ribose was unclear.

Figure S4. Flexible regions in Mss116p crystals. Both the thickness and color of the Mss116p C_α and U_{10} phosphate backbones are displayed according to atomic B factor from lowest (thin and blue) to highest (thick and red), using the B factor putty mode of PyMOL. B factors displayed are total B factors (residual + TLS) for the AMP-PNP complex as output by TLSANL (CCP4, 1994). The ADP-BeF_3^-, ADP-AlF_4^-, and Br-U_{10} complexes gave essentially the same results (data not shown).

Figure S5. CD spectroscopy of Mss116p and several C-terminal truncations. (A) Secondary structure estimation. Mean residue ellipticity (θ) as a function of wavelength from 200 to 260 nm at 30˚C. Each scan is the result of three accumulations. The secondary structure content of each protein, as predicted from the 200-240 nm region by the K2D2 server (Perez-Iratxeta and Andrade-Navarroare, 2008), is indicated. (B) Thermal unfolding. Mean residue ellipticity (θ) as a function of temperature at 220 nM. Each unfolding curve was fit by a two-state transition model (see Supplemental Experimental Procedures), and the $T_{1/2} \pm$ the standard error of the fit is indicated. The plots were created with SigmaPlot (Systat Software)

Figure S6. Comparison of the Mss116p structure with those of Vasa, eIF4AIII, and DDX19. Cartoons of (A) Vasa, (B) eIF4AIII, and (C) DDX19 superposed onto Mss116p by PyMOL. In all panels, helicase core domains 1 and 2 are pink for Mss116p and gray for the other proteins. The CTE of Mss116p is shown in red. In panel (B), two fragments of the Barentsz (Btz) protein

from the EJC structure are shown in blue, one of which occupies a similar position to α18 in the Mss116p CTE.

Supplemental References

CCP4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760-763.

Greenfield, N.J. (2006). Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat. Protocols *1*, 2527-2535.

Holm, L., Kääriäinen, S., Rosenström, P., and Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. Bioinformatics *24*, 2780-2781.

O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D.G., and Notredame, C. (2004). 3DCoffee: combining protein sequences and structures within multiple sequence alignments. J. Mol. Biol. *340*, 385-395.

Perez-Iratxeta, C., and Andrade-Navarroare, M.A. (2008). K2D2: Estimation of protein secondary structure from circular dichroism spectra. BMC Structural Biology *8*, 25.

F) Supplemental Figure 1

Figure S1

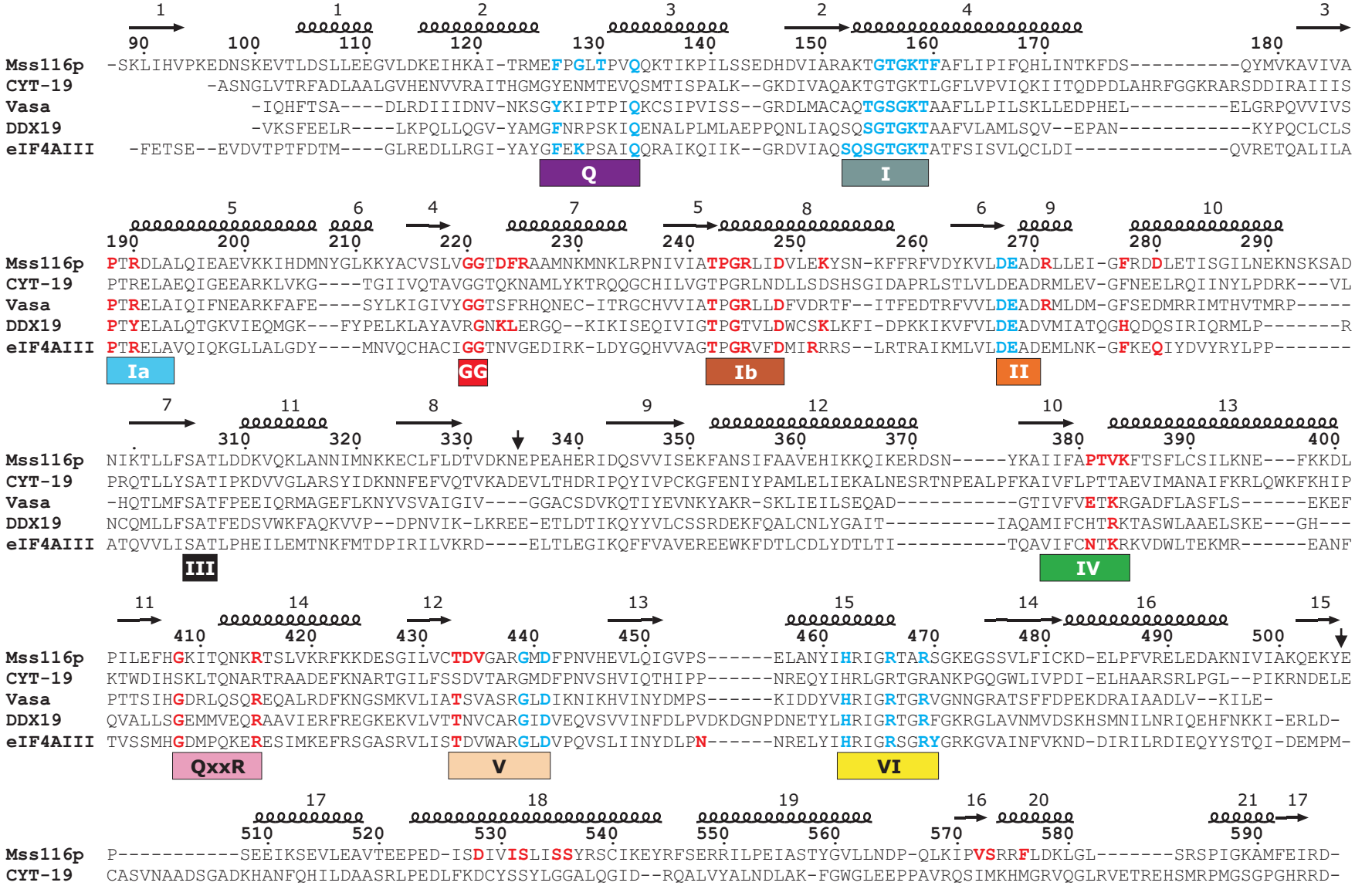


Figure S2

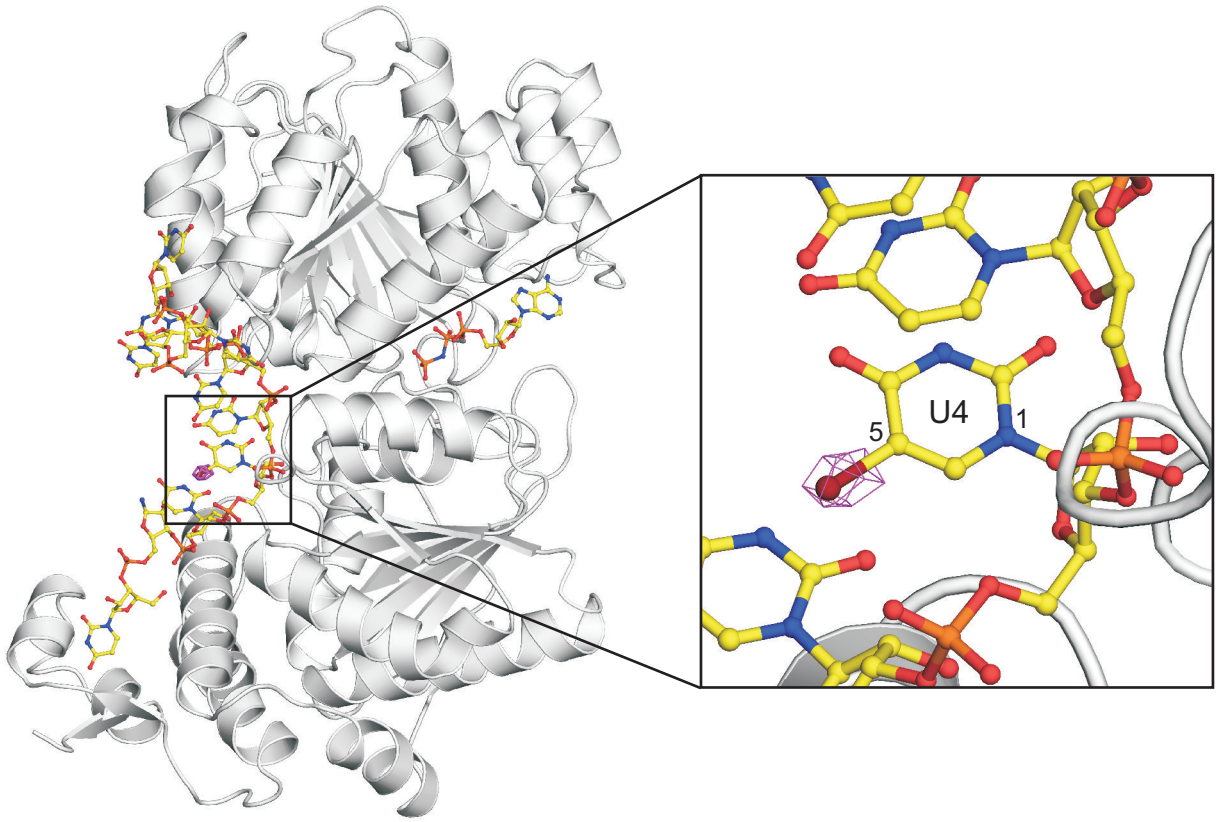


Figure S3

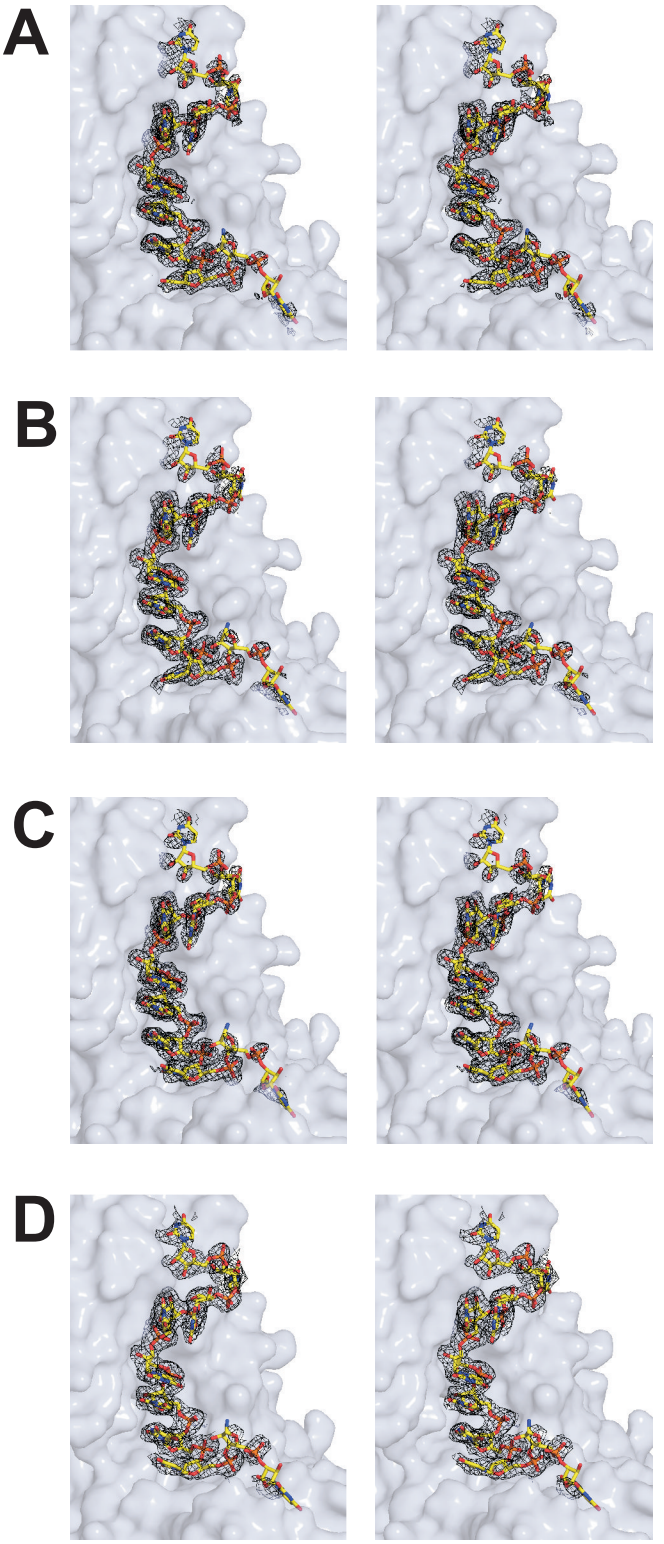


Figure S4

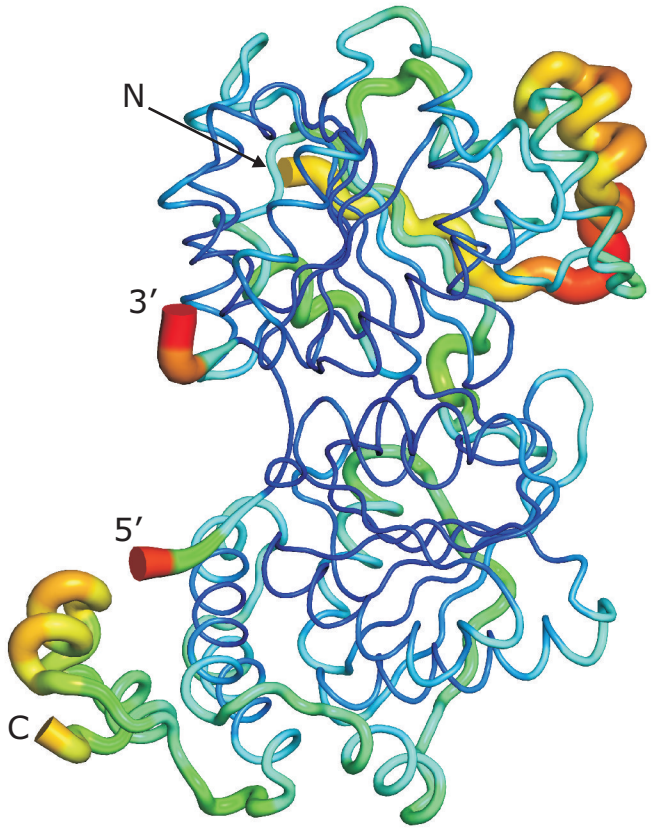


Figure S5

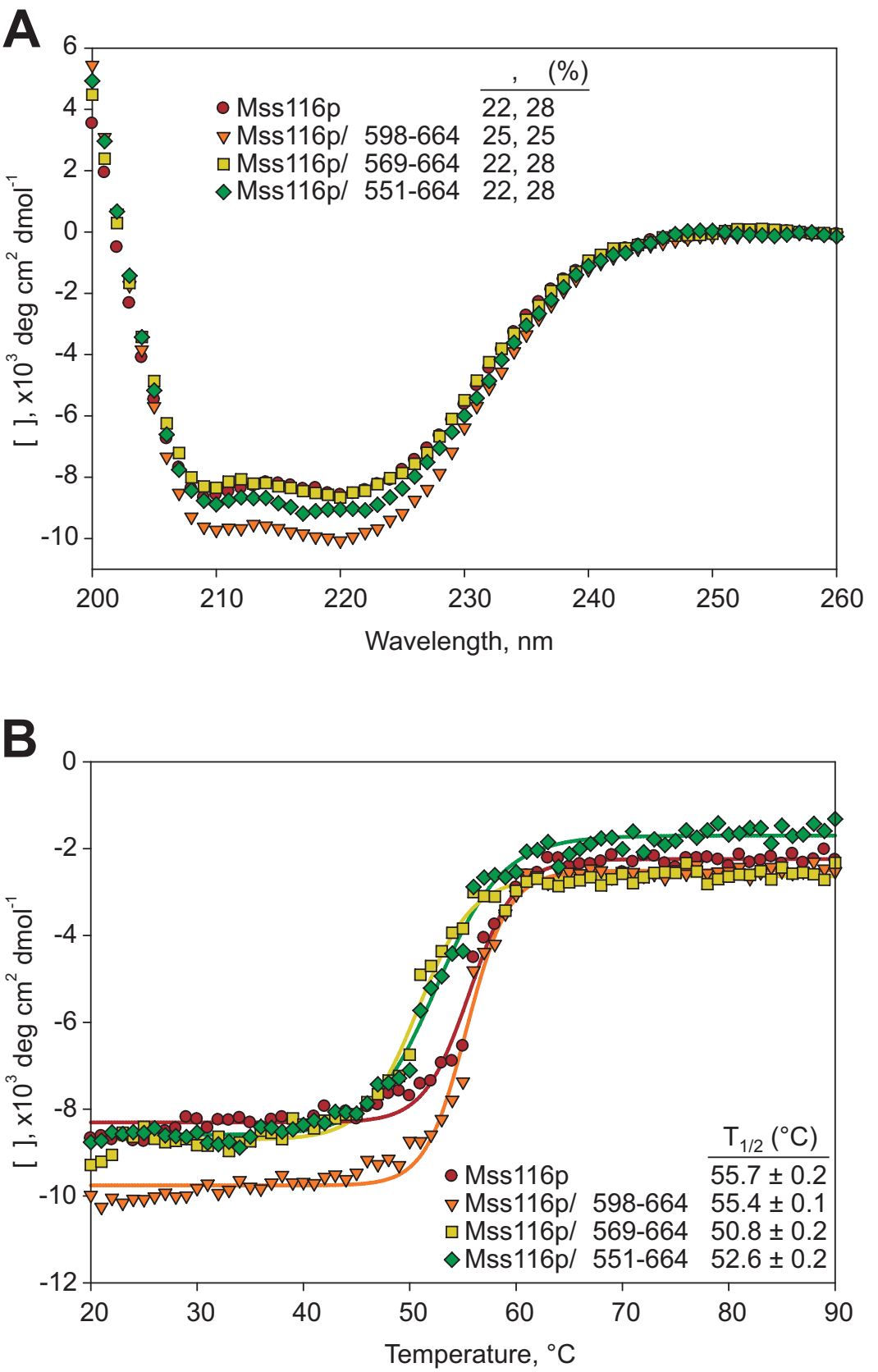
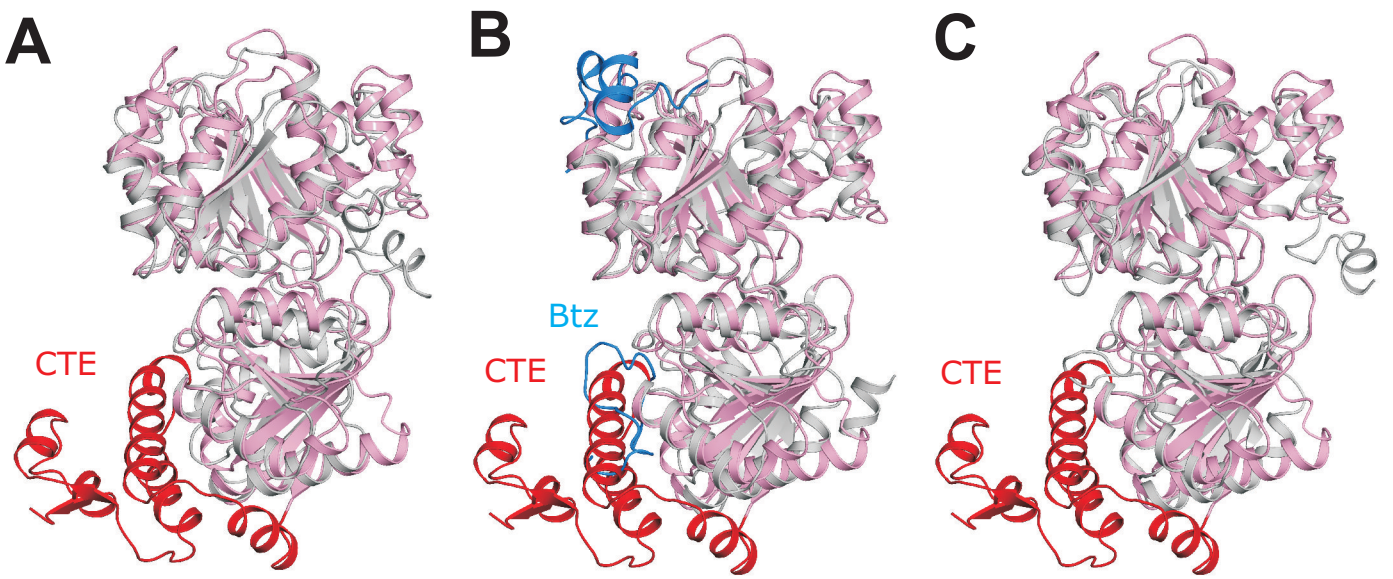

A
[ ], x10^3 deg cm^2 dmol^-1
Wavelength, nm
, (%)
Mss116p 22, 28
Mss116p/ 598-664 25, 25
Mss116p/ 569-664 22, 28
Mss116p/ 551-664 22, 28
B
[ ], x10^3 deg cm^2 dmol^-1
Temperature, °C
T1/2 (°C)
Mss116p 55.7 ± 0.2
Mss116p/ 598-664 55.4 ± 0.1
Mss116p/ 569-664 50.8 ± 0.2
Mss116p/ 551-664 52.6 ± 0.2

Figure S6



Mss116p & Vasa

Mss116p & eIF4AIII

Mss116p & DDX19