







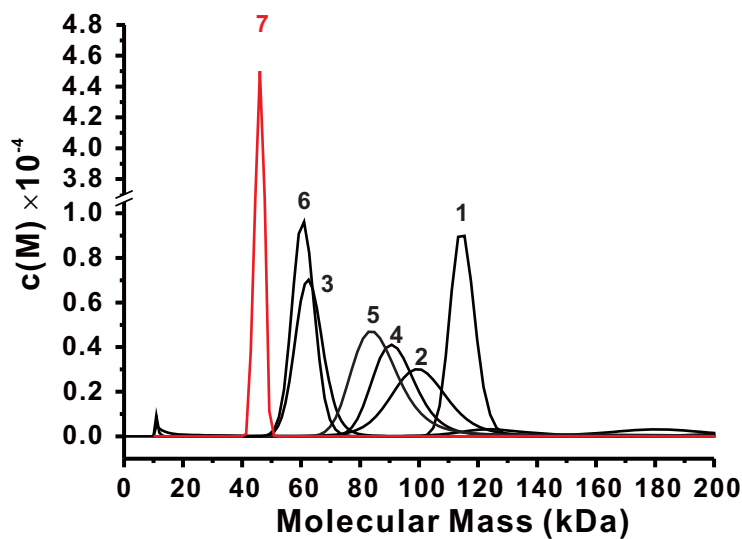


## Supplementary figure 1

**A**

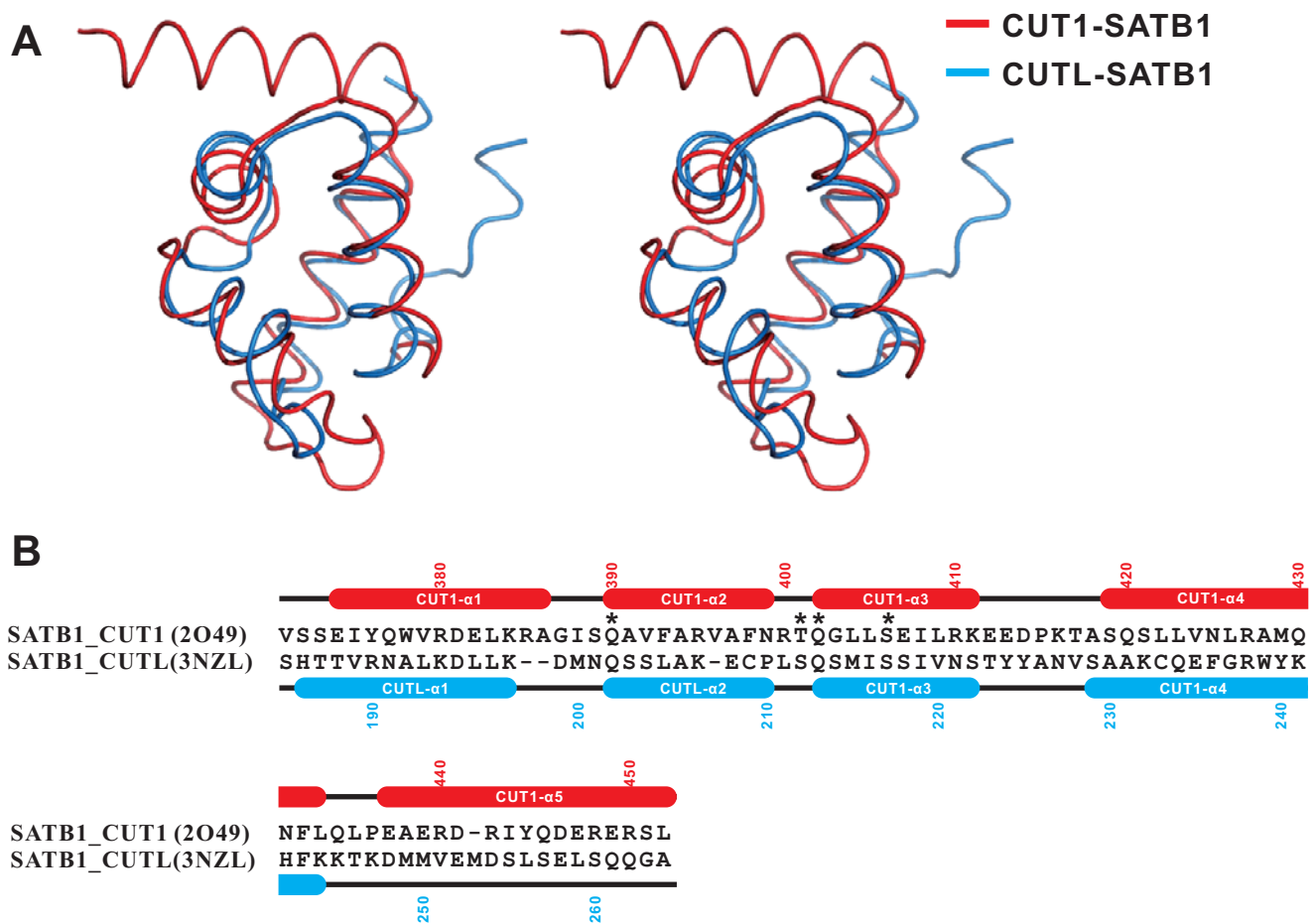
		Protein expression	Protein stability	Theoretical MW (kDa)	Fitted MW by SV (kDa)	crystal screening	PDB code
1-263		yes	+	29.5	113.3	N.A	N.A
60-263		yes	++	23.4	98.7	N.A	N.A
60-206		yes	++	16.9	61.3	yes	N.A
71-263		yes	++	22.1	90.7	N.A	N.A
71-248		yes	+++	20.5	82.9	yes	N.A
71-206		yes	++	15.6	61.0	yes	N.A
71-172		yes	++++	11.7	46.1	yes	3TUO
88-206		insoluble	N.A	13.7	N.A	N.A	N.A

**B**



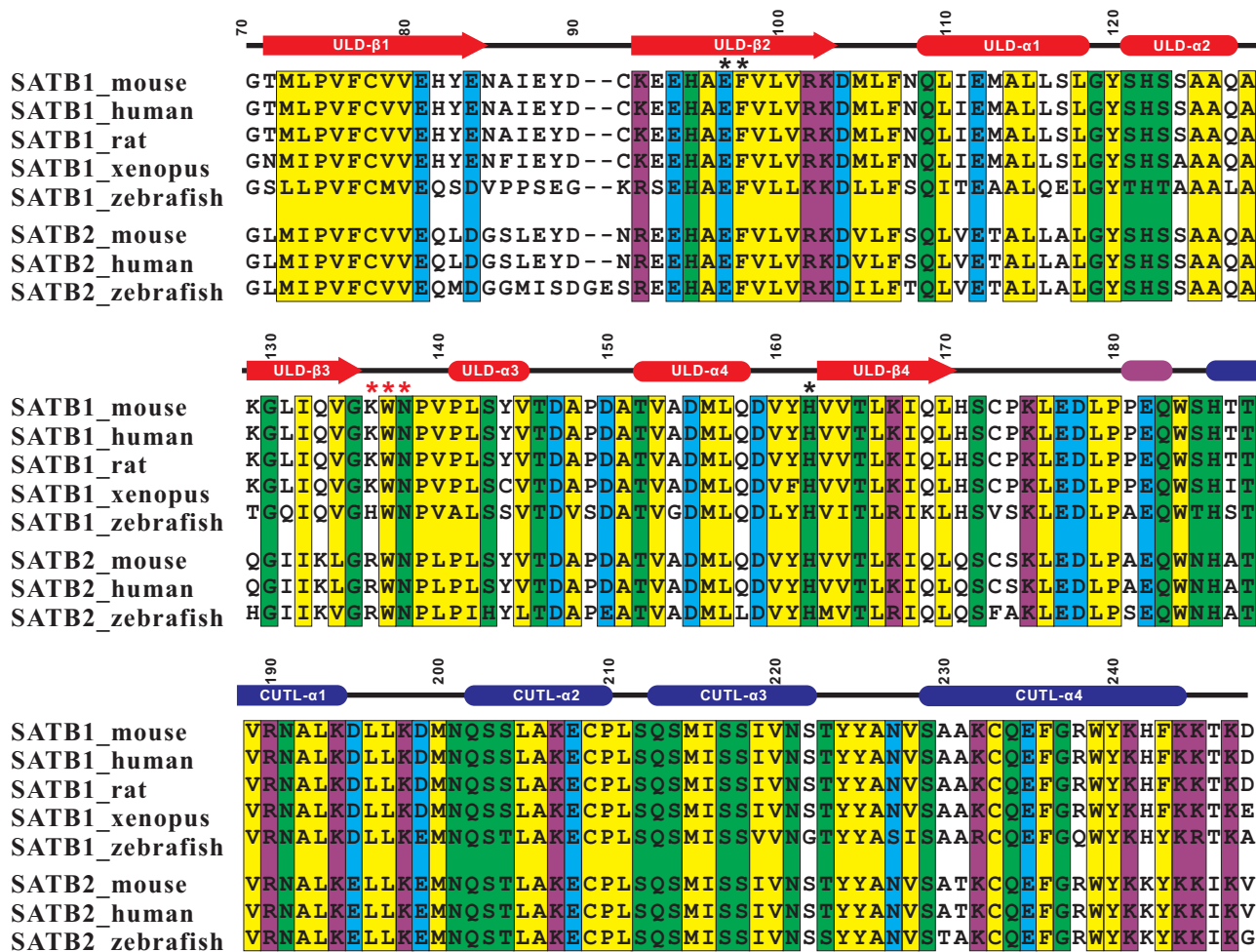
**Figure S1: Summary of the design of the N-terminal SATB1 constructs and tests for protein expression, molecular weight determination and crystallization. (A)** Greater numbers of “+”s correlate with more stable recombinant proteins. “N.A” indicates “Not Available”. The purified proteins from construct-3, -5, -6, and -7 were applied to crystal screening. Only the trial of construct-7 gave us hits for protein crystallization, and the structure of construct-7 was determined finally in this study by X-ray crystallography. **(B)** The molecular weights of the purified proteins were measured by analytical ultracentrifugation sedimentation velocity (SV), indicating that all purified proteins formed into tetramers. The labeled line numbers in (B) are consistent with the serial numbers in (A).

## Supplementary figure 2



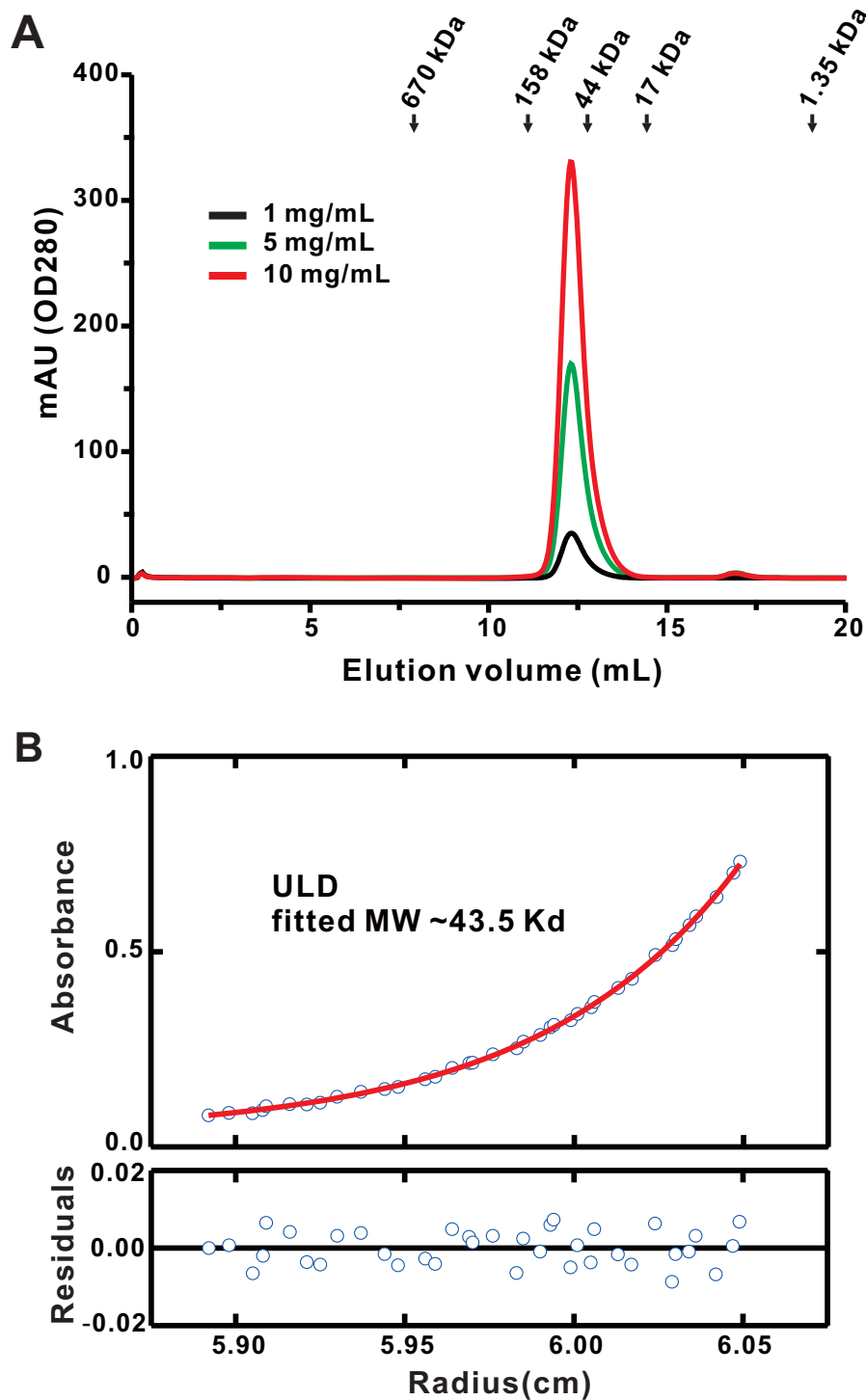
**Figure S2: SATB1 CUT1 and CUTL domain have similar 3D structure and conserved amino acids for DNA-binding.** (A) Stereo view showing the superimposed structures of SATB1 CUT1 (PDB code: 2O49) and CUTL domain (PDB code: 3NZL). (B) Structure based sequence alignment of SATB1 CUT1 and CUTL domain. The secondary structures of CUT1 and CUTL are derived from PDB code (2O49) and (3NZL) respectively. The conserved amino acids involved in DNA binding are marked with black asterisks. The amino acid residue number is also indicated in the figure.

## Supplementary figure 3



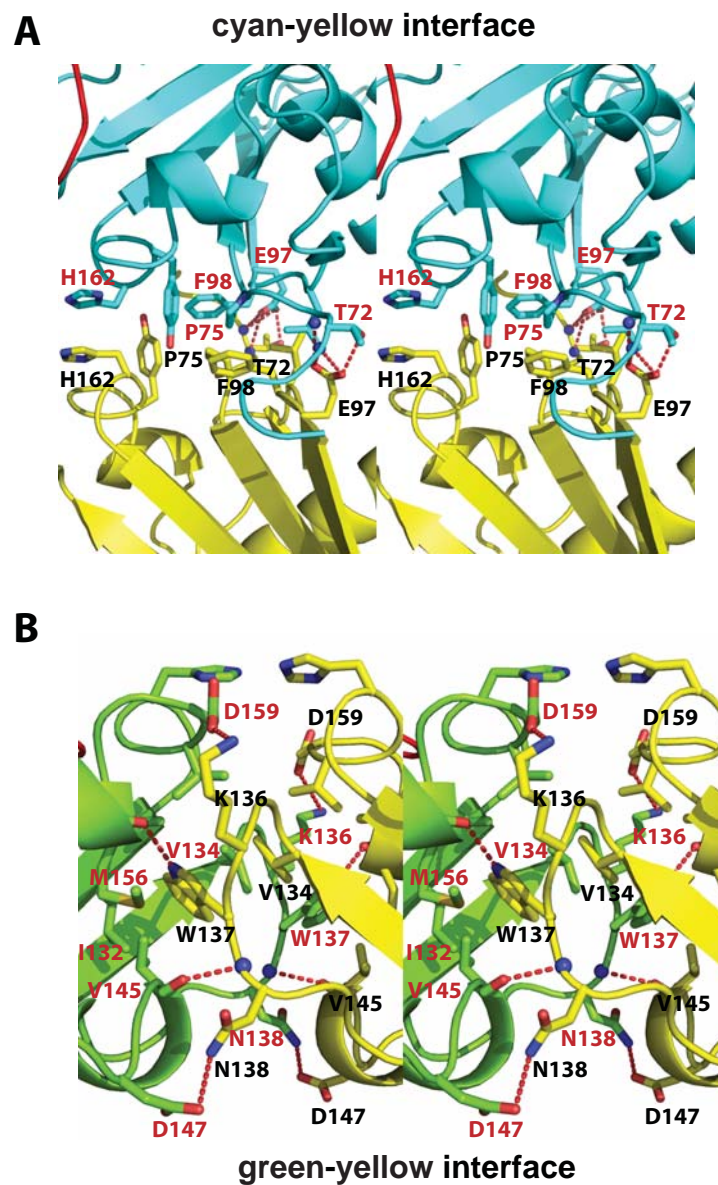
**Figure S3: Structure-based sequence alignment of SATB ULD-CUTLs from different species.** In this alignment, the secondary structures of the SATB1 ULD-CUTL are shown at the top of each sequence. The secondary structures of CUTL are derived from PDB code (3NZZ). The conserved hydrophobic residues of the ULD-CUTLs are in yellow; the conserved hydrophilic residues are in green; the conserved negatively charged residues are in cyan; and the conserved positively charged residues are in purple. The <sup>97</sup>E<sup>98</sup>F<sup>162</sup>H and <sup>136</sup>K<sup>137</sup>W<sup>138</sup>N motifs are marked with black and red asterisks, respectively. The amino acid residue number of mouse SATB1 ULD-CUTL is also indicated in the figure.

## Supplementary figure 4



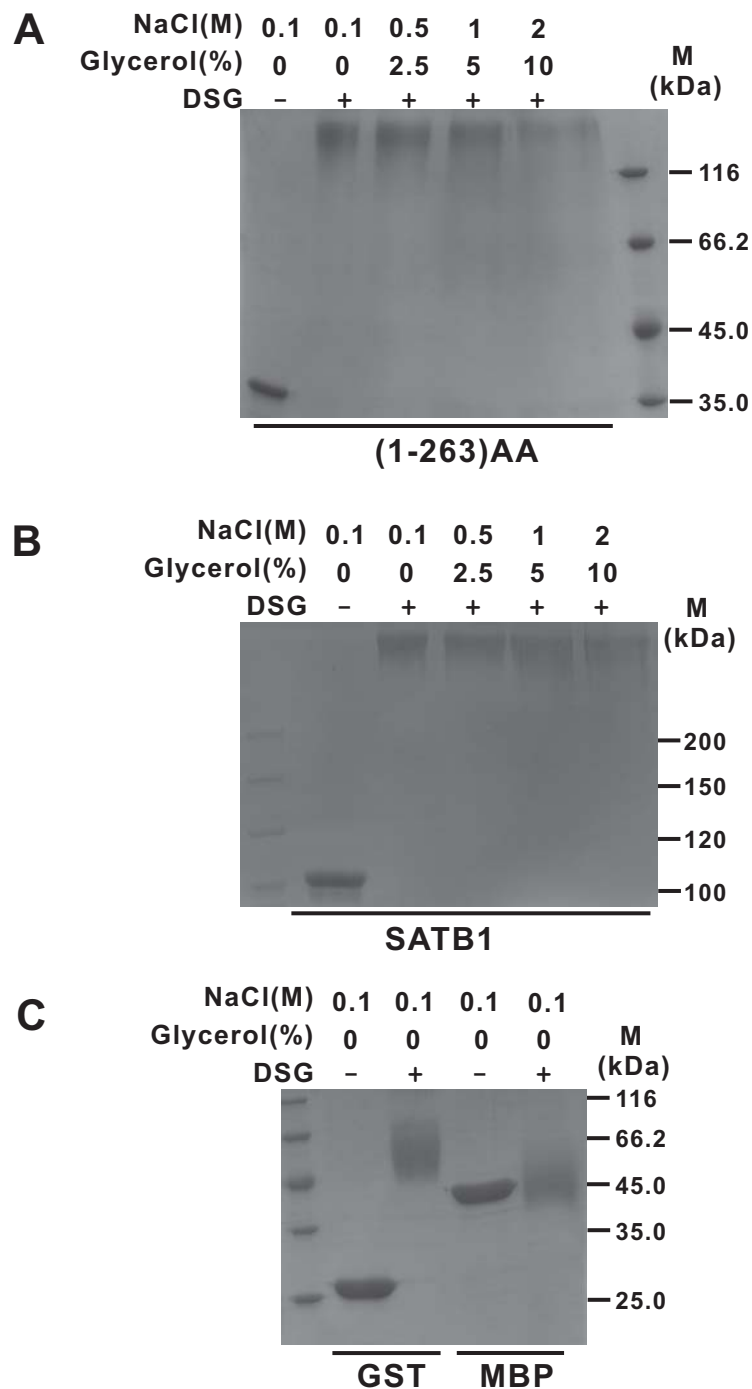
**Figure S4: The SATB1 ULD forms a stable homotetramer in solution. (A)** Analytical gel filtration profiles of the SATB1 ULD at different loading concentrations. The elution volumes of the molecular mass standards are marked at the top of the panel. **(B)** Representative sedimentation equilibrium profile for the SATB1 ULD derived from a global analysis of data at a range of loading concentrations and rotor speeds with a global fitted MW  $\sim 43.5$  kDa, indicating that the SATB1 ULD assembled into a tetramer in solution.

## Supplemental figure 5



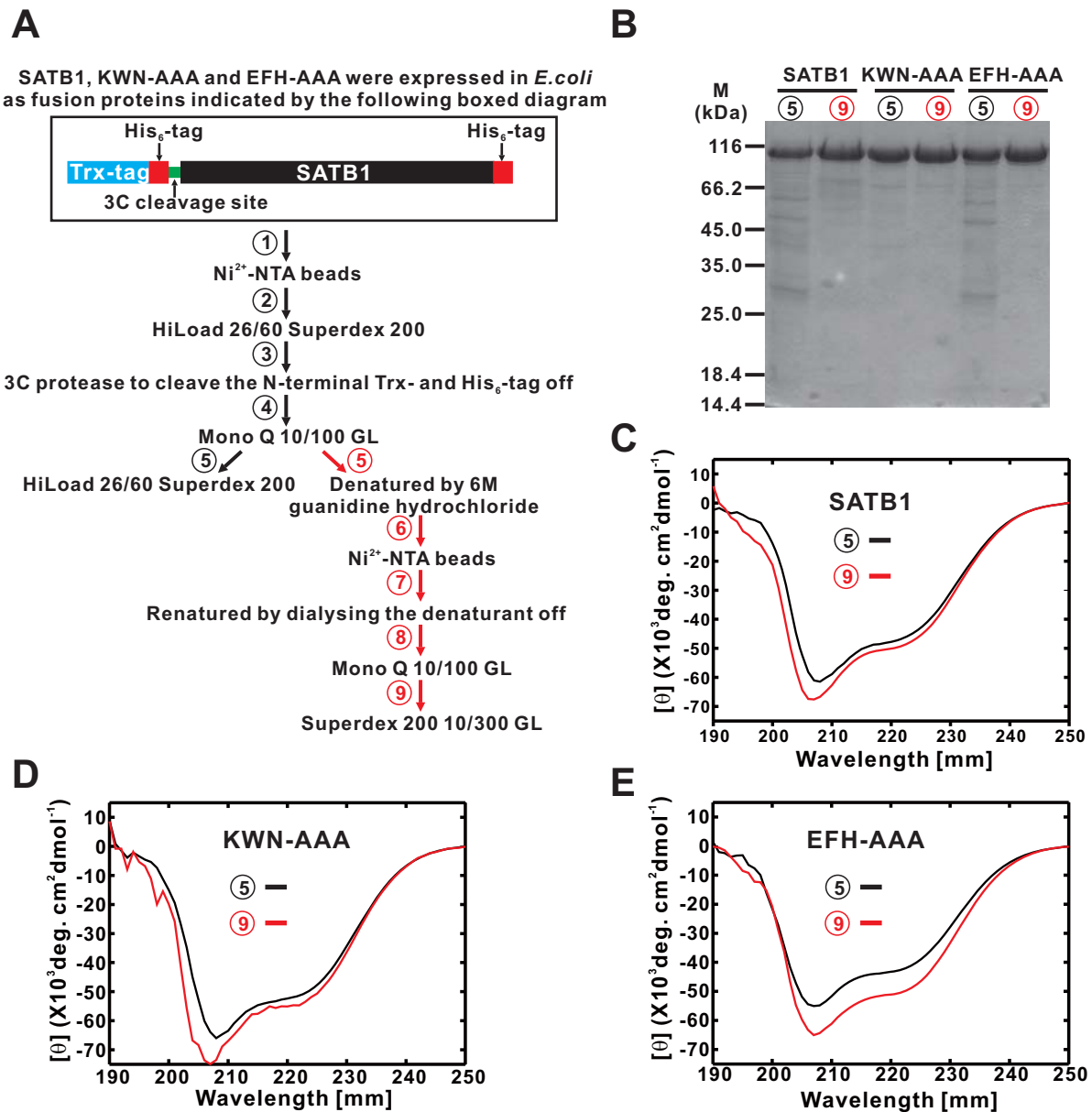
**Figure S5: Stereo view of the detailed interface within two ULD dimers.** The cyan–yellow interface (B) and the green–yellow interface (C) represent  $^{97}\text{E}^{98}\text{F}^{162}\text{H}$  and  $^{136}\text{K}^{137}\text{W}^{138}\text{N}$  motifs mediated ULD dimerization, respectively. The color of the monomer ULD is consistent with figure 2A. Oxygen and nitrogen atoms are colored red and blue, respectively. Hydrogen bonds are shown as red dotted lines.

## Supplementary Figure S6



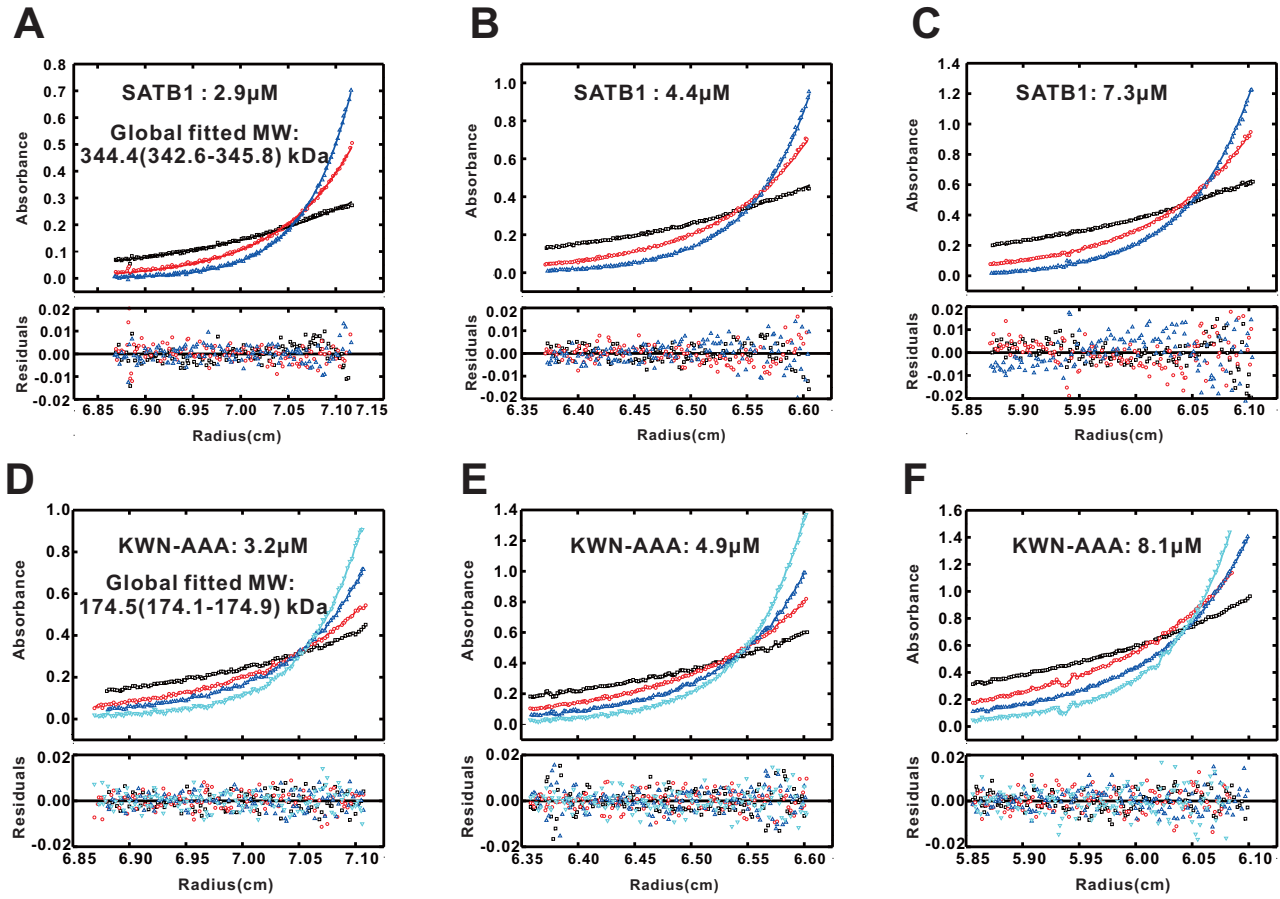
**Figure S6. Hydrophobic solvents and ionic salts didn't disrupt ULD-mediated SATB1 tetramerization.** DSG cross-linking assay showing the tetramerization of (1-263)AA fragment (A) and SATB1 full length (B) in buffer containing various concentration of NaCl and glycerol. GST and MBP protein (C) were served as positive and negative control, respectively.

# Supplementary Figure S7



**Figure S7: Outline for the purification of wild-type and mutant SATB1 proteins. (A)** Schematic diagram of the protein purification protocol. Black arrows show the general purification steps to produce native proteins, and the red arrows show the further purification steps by denaturation and renaturation to produce refolded proteins. **(B)** SDS-PAGE gel showing the final protein purities from the two approaches described in (A). **(C-E)** The circular dichroism spectra of various proteins in their native state or after refolding. The native (black curve) and refolded proteins (red curve) had very similar CD spectrum for wild-type SATB1 (C), KWN-AAA (D), and EFH-AAA (E).

## Supplementary Figure S8

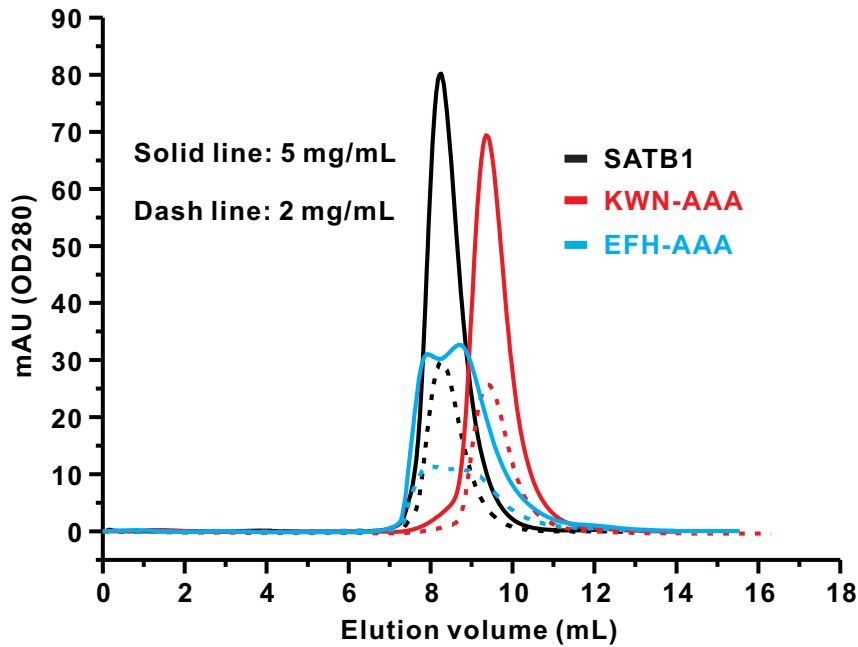


### Figure S8: SATB1 is a tetramer, and the KWN-AAA mutant is a dimer in solution.

Sedimentation equilibrium analysis of SATB1 (A -C) and the KWN-AAA mutant (D-F) at three different concentrations and three rotor speeds for wild-type SATB1 and four rotor speeds for the KWN-AAA mutant. The molecular weights of SATB1 and the KWN-AAA mutant were derived from the global analysis of the data at a range of loading concentrations and rotor speeds with a global fit of MW 344.4 kDa ( $1\sigma$  confidence interval 342.6-345.8 kDa) and 174.5 kDa ( $1\sigma$  confidence interval 174.1-174.9 kDa), respectively, indicating that SATB1 forms a tetramer, and the KWN-AAA mutant exists as a dimer in solution. The theoretical MW of SATB1 is 85.9 kDa. In each figure, different colored lines represent different rotor speeds at the same protein concentration. In figure A-C, black, red and blue represent 4,300, 6,400 and 8,000 r.p.m., respectively. In figure D-E, black, red, blue and cyan represent 6,000, 8,000, 10,000 and 12,000 r.p.m., respectively. The inserts present the individual protein concentrations in each experiment, and the global fit MW for SATB1 and the KW N-AAA mutant are indicated in (A) and (D), respectively.



## Supplementary Figure S9



**Figure S9: Analytical gel filtration profiles of wild-type and mutant SATB1 at different concentrations.** Wild-type SATB1 and the KWN-AAA mutant eluted as a single sharp peak by gel filtration whereas the EFH-AAA mutant eluted as two seriate broad peaks, indicating that wild-type SATB1 and the KWN-AAA mutant exist in a unique form of oligomers whereas the EFH-AAA mutant exists in multiple forms of oligomers in solution.