Structure of the super-elongation complex subunit AFF4 C-terminal homology domain reveals requirements for AFF homo- and heterodimerization

Ying Chen, Patrick Cramer

Supplementary information

Supplementary Tables: Table S1-S2

Supplementary Figures: Figure S1-S3

Supplementary Tables

Supplementary Table S1, Primer list and sequences

Nucleotide sequences used in fluorescence anisotropy

Name	**Sequence (5'-3')**
FBS-35mer G-quadruplex RNA	6-FAM-AAGAAGAGGACAAGGAGGAAGAGGACGGGAGGAGGCUUC
HIV1 LTR-III G-quadruplex DNA	6-FAM-GGGAGGCGTGGCCTGGGCGGGACTGGGG
Unstructured TAR RNA	6-FAM-UUAAGGAAUUAAGUCGUGCGUCUAAUAACCGGAGAGGGAACCCACU
Stem-loop-TAR RNA	6-FAM-ACCAGAUCUGAGCCUGGGAGCUCUCUGGCUAACUAGGGAACCCACU

Primers use to clone AFF4-CHD and AFF1-CHD to protein expression vector

Name	**Sequence (5'-3')**	**Description**
AFF4-899-LICv1-Fw	TACTTCCAATCCAATGCATCTAAGCCTCGGAGAACAAAGCTTG	His_6-tagged AFF4-CHD
AFF4-1163-LICv1-Rev	TTATCCACTTCCAATGTTATTAAGATATCAACTTGGCATCCTGGC	
AFF1-938-LICv1-Fw	TACTTCCAATCCAATGCAAATCCTTTTCCAGTGCCTTCTTTGC	His_6-MBP-tagged AFF1-CHD
AFF1-1210-LICv1- Rev	TTATCCACTTCCAATGTTATTAAGGTGTTTTGGTTAATTCTTGTAGCTGCTG	

Primers used to generate mutants of AFF4-CHD

Name	**Sequence (5'-3')**	**Description**
AFF4-H1090A-Fw	TTCAGTGACCATTCCACAGAAGATCGCACAGATGGCAGCCAGC	His_6-tagged AFF4-CHD-5MA mutant
AFF4-H1090A-Rev	GCTGGCTGCCATCTGTGCGATCTTCTGTGGAATGGTCACTGAA	

Name	Sequence (5'-3')	Description
AFF4-YV1096,1607AA-Fw	CAGATGGCAGCCAGCGCAGCACAGGTCACATCCAAC	
AFF4-YV1096,1607AA-Rev	GTTGGATGTGACCTGTGCTGCGCTGGCTGCCATCTG	
AFF4-FL1103,1134AA-Fw	CAGGTCACATCCAACGCAGCATATGCCACCGAAATTTGGGACCAAGC	
AFF4-FL1103,1134AA-Rev	GCTTGGTCCCAAATTTCGGTGGCATATGCTGCGTTGGATGTGACCTG	
AFF4-H1090D-Fw	AGTGACCATTCCACAGAAGATCGATCAGATGGCAGCCAGC	His_6-tagged AFF4-CHD-5MD mutant
AFF4-H1090D-Rev	GCTGGCTGCCATCTGATCGATCTTCTGTGGAATGGTCACT	
AFF4-YV1096,1097DD-Fw	CAGATGGCAGCCAGCGATGATCAGGTCACATCCAACGATGATTAT	
AFF4-YV1096,1097DD-Rev	ATAATCATCGTTGGATGTGACCTGATCATCGCTGGCTGCCATCTG	
AFF4-FL1103,1134DD-Fw	CAGGTCACATCCAACGATGATTATGCCACCGAAATTTGGGACCAAG	
AFF4-FL1103,1134DD-Rev	CTTGGTCCCAAATTTCGGTGGCATAATCATCGTTGGATGTGACCTG	

Supplementary Table S2, Curve fitting data for AFF4-CHD RNA and DNA binding as determined by fluorescence anisotropy. Fluorescence anisotropy data on Figure 3 and 4 were fitted with the quadratic single site binding equation in GraphPad Prism, see Experimental Procedure.
Curve fitting data for Figure 3B, the binding of FBS-35mer G-quadruplex RNA to AFF4-CHD in buffer containing different KCl concentration.

KCl concentration	**$K_{d, app}$ (μM)**	**R^2**	**B_{max}**
50 mM	0.422 ± 0.016	0.9935	124.4 ± 1.15
100 mM	0.645 ± 0.036	0.9861	91.54 ± 1.39
200 mM	0.691 ± 0.081	0.9436	65.09 ± 2.11

Curve fitting data for Figure 3C, the binding of different nucleotides to AFF4-CHD in buffer containing 100 mM KCl.

Nucleotides	**$K_{d, app}$ (μM)**	**R^2**	**B_{max}**
FBS-35mer G-quadruplex RNA	0.270 ± 0.026	0.9784	90.88 ± 2.00
HIV1 LTR-III G-quadruplex DNA	1.647 ± 0.176	0.9792	51.05 ± 1.87
Unstructured TAR RNA	0.228 ± 0.044	0.9058	69.66 ± 3.05
Stem-loop-TAR RNA	0.251 ± 0.026	0.9724	67.85 ± 1.60

Curve fitting for Figure 4C, the binding of FBS-35mer G-quadruplex RNA to phosphorylated and Apo AFF4-CHD.

Phosphorylation state	**$K_{d, app}$ (μM)**	**R^2**	**B_{max}**
Phosphorylated AFF4-CHD (P-TEFb WT)	0.693 ± 0.068	0.9545	81.98 ± 2.24
Apo AFF4-CHD (P-TEFb N-mut)	0.275 ± 0.019	0.9751	74.90 ± 1.20

Legends to Supplementary Figures

Figure S1. Alignments of AFF proteins from human, mouse, chicken, zebra fish and drosophila. The sequence ID of the proteins in non-redundant protein sequences databank are as the following: *Homo sapiens* AFF4 (NP_055238.1), *Mus musculus* AFF4 (NP_291043.1), *Gallus gallus* AFF4 (XP_004945087.1), *Danio rerio* AFF4 (XP_005173957.1), *Homo sapiens* AFF1 (NP_001160165.1), *Mus musculus* AFF1 (NP_001074267.1), *Gallus gallus* AFF1 (XP_025006145.1), *Homo sapiens* AFF3 (NP_002276.2), *Mus musculus* AFF3 (NP_001277743.1), *Gallus gallus* AFF3 (XP_05133278.1), *Homo sapiens* AFF2 (NP_001162594.1), *Mus pahari* AFF2 (021043331.1), *Danio rerio* AFF2 (XP_002664429), *Gallus gallus* AFF2 (XP_015134139.2), *Danio rerio* AFF3 (XP_009302946), *Drosophila melanogaster* Lillipitian (NP_523464.1). The alignment was made using full-length AFF proteins and only the C-terminal sequences are shown in the figure. Stars indicate phosphorylation sites identified by mass spectrometry. Black stars mark the residues that are not highly conserved among AFF proteins. Brown stars indicate the conserved residues that are identified as phosphorylation sites. Residues on dimerization interface are marked with black and orange squares. Orange squares indicate the residues that are mutated in 5MA and 5MD constructs.

Figure S2. (A)-(C) Superposition of AFF4-CHD with top hits from DALI search. (A) Superposition of AFF4-CHD onto 5G05-O, the anaphase-promoting complex subunit 5 (APC5) (30). AFF4-CHD is colored in dark green and APC5 is colored in gray. AFF4-CHD superposes well with the last 5 helices in the C-terminal of APC5, with an RMSD of 2.7 over 452 atoms (30). (B) Superposition of AFF4-CHD onto 4JHR-B (32), the G-protein-signaling modulator 2. AFF4-CHD is colored in dark green and G-protein modulator 2 is colored in gray. AFF4-CHD superposes well with the 5 helices in the N-terminal of G-protein modulator 2, with an RMSD of 2.216 over 424 atoms. (C) Superposition of AFF4-CHD onto 6HEP-B (in gray), the 14-3-3 protein β, in the presence of CFTR R-domain peptide pS753-pS768 as the ligand (6HEP-E, in black) (33). (D-E) SDS-PAGE of the elution fractions of analytical gel filtration. (D) Shows the fractions of analytical gel filtrations of AFF4-CHD-WT, AFF4-CHD-5MA and AFF4-CHD-5MD, corresponding to the elution profile in Figure 2B. The same fractions that cover the peaks of all three runs are shown on the SDS-PAGE and stained with instant blue (Expendeon). (E) shows the fractions of Figure 2D. The same fractions from all four runs were loaded to SDS-PAGE and stained with instant blue (Expendeon).

Figure S3. (A)-(B) SDS-PAGE of the AFF4-CHD before and after phosphorylation by P-TEFb and P-TEFb N mutant. This figure shows the phosphorylation of the protein used in Figure 4C and 4D. (A) shows the protein involved in the phosphorylation of replication 1 (R1). (B) shows the protein involved in the phosphorylation of replication 2 and 3 on the same PAGE (R2 and R3). (C) Calibration of analytical gel filtration column Superdex 200 3.2/300 increase (GE Healthcare). The protein standards and their molecular weight are as indicated in the figure. The Log Mw (y) of each protein was plotted against the elution peak volume of each protein (x) in GraphPad Prism. Linear regression curve was fitted to the data points. The fitting equation and the R^2 of the fitting are indicated on the panel. The molecular weight of an unknown protein can be calculated using the given equation and the elution volume (x). (D)-(E) Static light scattering results of AFF4-CHD-WT, AFF4-CHD-5MD, MBP-AFF1-CHD and MBP-AFF1-CHD/AFF4-CHD-WT complex. In each panel, the elution profile of a SEC run (A280, curve in black) and the molecular weight measured from light scattering (Mw, curve in red) are plotted against elution volume. The values are indicated as left and right y-axis respectively. For panel G, to form the heterodimeric MBP-AFF1-CHD/AFF4-CHD-WT complex, MBP-AFF1-CHD and AFF4-CHD-WT were mixed in 1:3 molar ratio and incubated overnight before injecting to the column. The molecular weight of the heteromeric peak and the excess AFF4 are both analyzed and indicated on the figure.

Figure S1

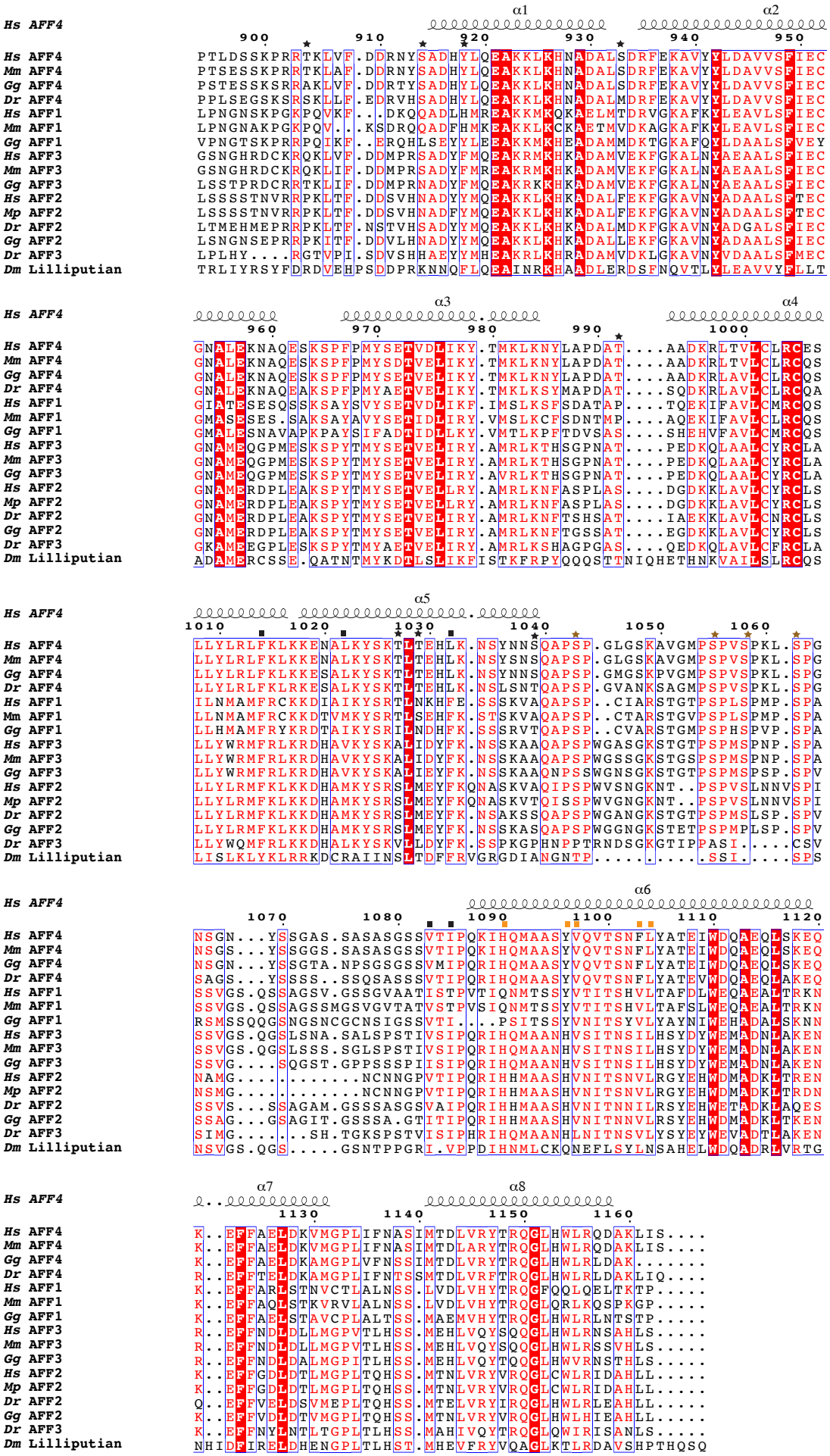
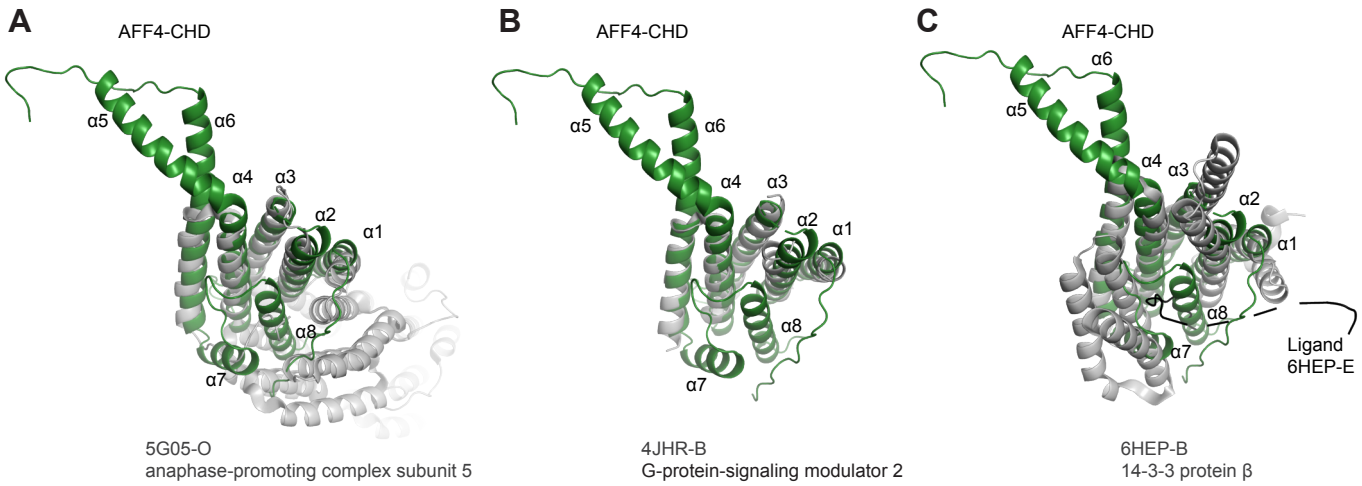
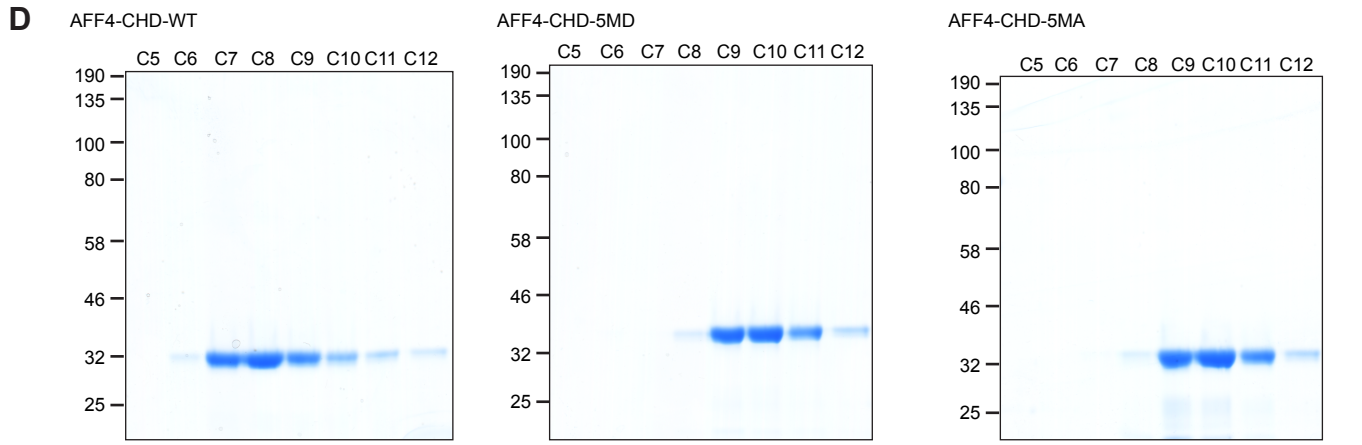

Hs AFF4
Hs AFF4
Mm AFF4
Gg AFF4
Dr AFF4
Hs AFF1
Mm AFF1
Gg AFF1
Hs AFF3
Mm AFF3
Gg AFF3
Hs AFF2
Mp AFF2
Dr AFF2
Gg AFF2
Dr AFF3
Dm Lilliputian
α1
α2
α3
α4
α5
α6
α7
α8

Figure S2


A
AFF4-CHD
α5
α6
α4
α3
α2
α1
α8
α7
5G05-O
anaphase-promoting complex subunit 5
B
AFF4-CHD
α5
α6
α4
α3
α2
α1
α8
α7
4JHR-B
G-protein-signaling modulator 2
C
AFF4-CHD
α6
α5
α4
α3
α2
α1
α8
α7
Ligand
6HEP-E
6HEP-B
14-3-3 protein β


D
AFF4-CHD-WT
AFF4-CHD-5MD
AFF4-CHD-5MA
C5 C6 C7 C8 C9 C10 C11 C12
190
135
100
80
58
46
32
25

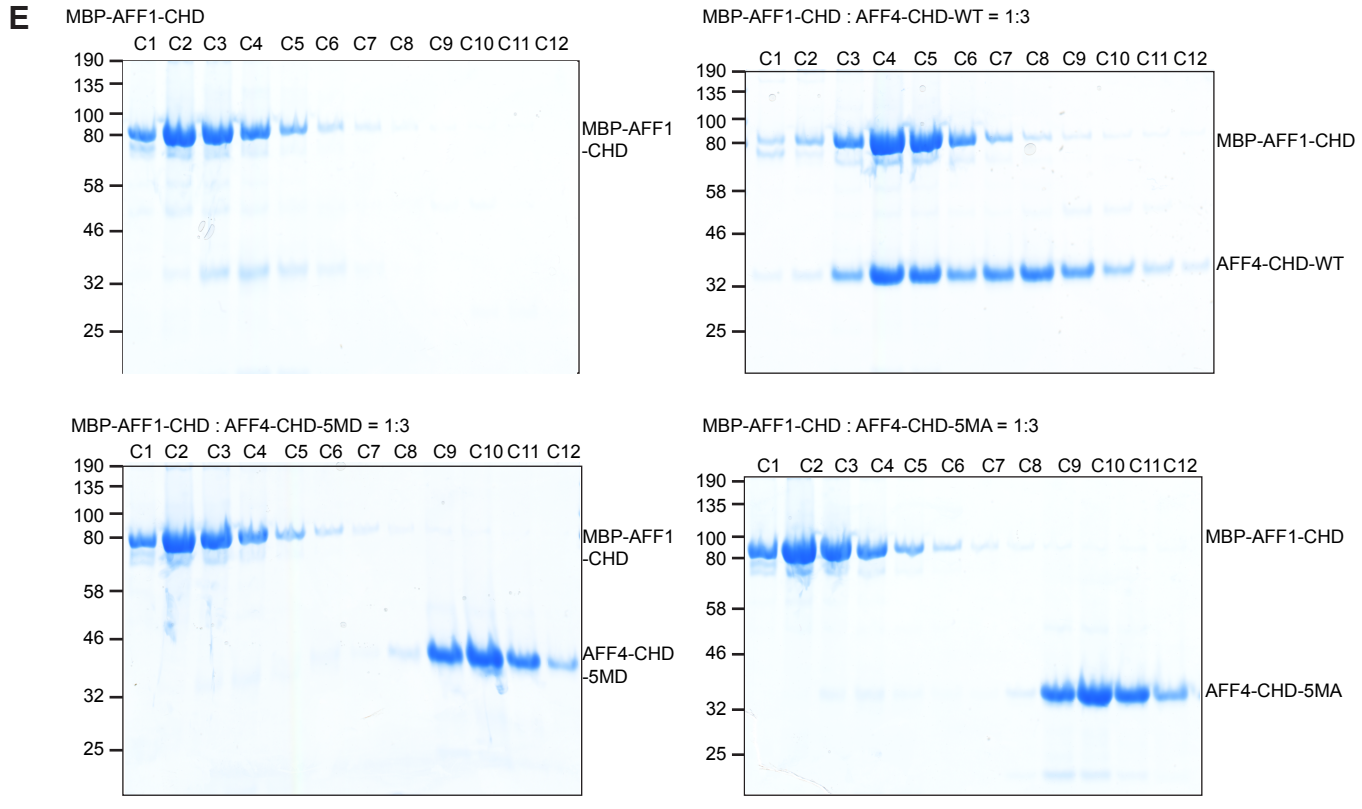
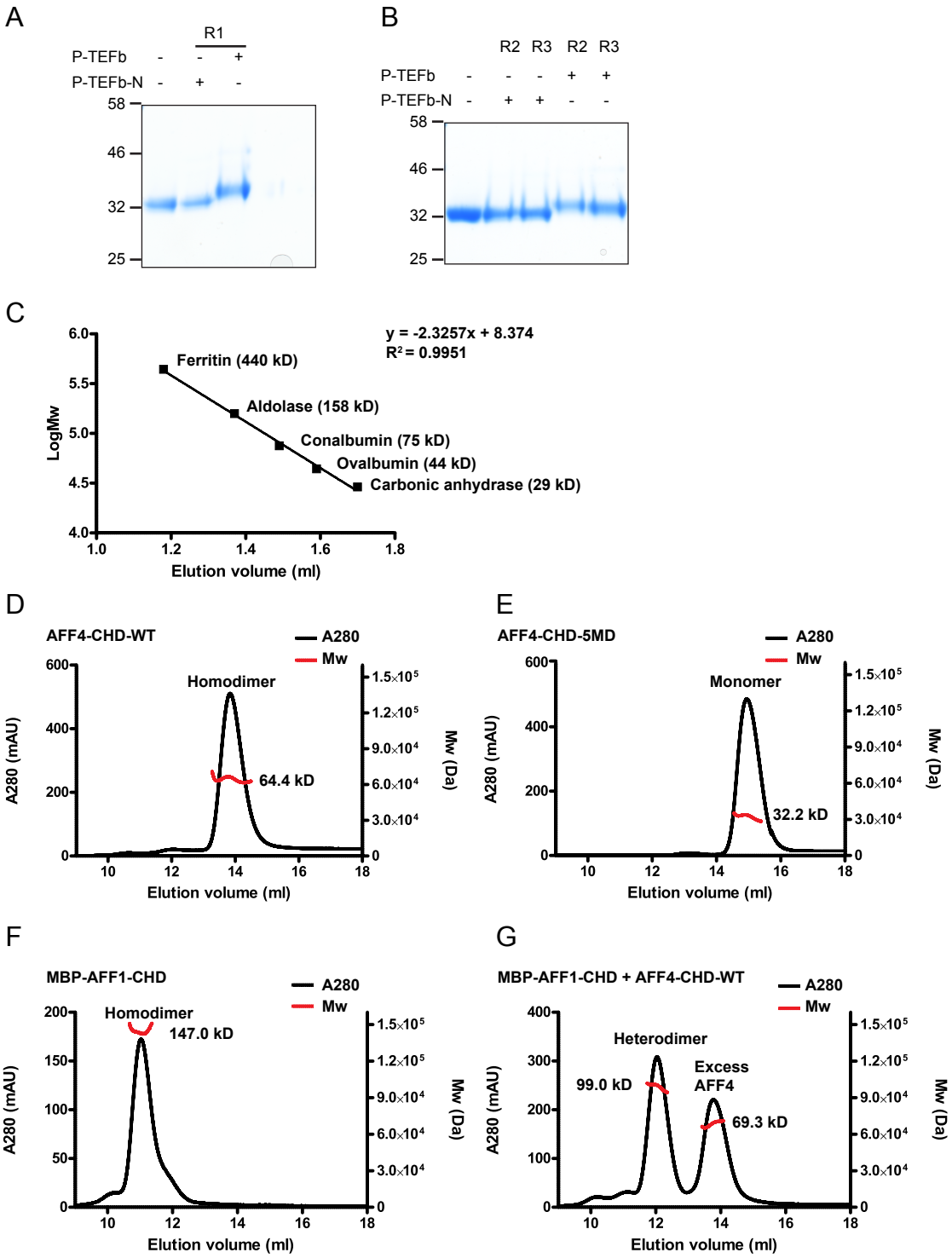

E
MBP-AFF1-CHD
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12
190
135
100
80
58
46
32
25
MBP-AFF1
-CHD
MBP-AFF1-CHD : AFF4-CHD-WT = 1:3
MBP-AFF1-CHD
AFF4-CHD-WT
MBP-AFF1-CHD : AFF4-CHD-5MD = 1:3
AFF4-CHD
-5MD
MBP-AFF1-CHD : AFF4-CHD-5MA = 1:3
AFF4-CHD-5MA

Figure S3


A
R1
P-TEFb - - +
P-TEFb-N - + -
58
46
32
25
B
R2 R3 R2 R3
P-TEFb - - - + +
P-TEFb-N - + + - -
58
46
32
25
C
y = -2.3257x + 8.374
R² = 0.9951
Ferritin (440 kD)
Aldolase (158 kD)
Conalbumin (75 kD)
Ovalbumin (44 kD)
Carbonic anhydrase (29 kD)
LogMw
Elution volume (ml)
D
AFF4-CHD-WT
A280
Mw
Homodimer
64.4 kD
A280 (mAU)
Mw (Da)
Elution volume (ml)
E
AFF4-CHD-5MD
A280
Mw
Monomer
32.2 kD
A280 (mAU)
Mw (Da)
Elution volume (ml)
F
MBP-AFF1-CHD
A280
Mw
Homodimer
147.0 kD
A280 (mAU)
Mw (Da)
Elution volume (ml)
G
MBP-AFF1-CHD + AFF4-CHD-WT
A280
Mw
Heterodimer
99.0 kD
Excess AFF4
69.3 kD
A280 (mAU)
Mw (Da)
Elution volume (ml)