FAM181A and FAM181B, two new TEAD interactors.

Fedir Bokhovchuk[1#], Yannick Mesrouze[1#], Clara Delaunay[1], Typhaine Martin[1], Frédéric Villard[2], Marco Meyerhofer[1], Patrizia Fontana[1], Catherine Zimmermann[1], Dirk Erdmann[1], Pascal Furet[3], Clemens Scheufler[2], Tobias Schmelzle[1] and Patrick Chène[1]*.

[1]Disease Area Oncology, [2]Chemical Biology & Therapeutics and [3]Global Discovery Chemistry, Novartis Institutes for Biomedical Research, Basel, Switzerland.

Supplementary Table

	FAM181A$^{190-205}$:TEAD4$^{217-434}$ PDB 6SEN	FAM181B$^{220-235}$:TEAD4$^{217-434}$ PDB 6SEO
Data collection[a]		
Space group	C_2	$H3_2$
Cell (Å, deg)	a=66.5, b=132.1, c=62.0, β=115.9	a=b=135.9, c=88.7
Resolution (Å)	1.61 (1.73-1.61)[b]	2.52 (2.64-2.52)[b]
$R_{merge, all}$ (%)	3.8 (83.5)	11.2 (224.9)
Mean I/σ(I)	14.3 (1.4)	15.5 (1.1)
Completeness spherical (%) Completeness ellipsoidal (%)	77.0 (20.1) 93.7 (63.4)	72.5 (28.0) 95.0 (81.7)
Multiplicity	3.4 (3.5)	10.1 (10.2)
Refinement[c]		
Resolution (Å)	25 - 1.65	30 - 2.55
No. reflections	47141	7693
R_{work}/ R_{free} (%)	19.0 / 21.9	18.3 / 23.2
R.m.s deviations bond lengths (Å)	0.010	0.010
R.m.s deviations bond angles (deg)	1.09	1.2
Complex molecules in AU	2	1
Myristate	Covalent[d]	Covalent[d]

Supplementary Table 1. Crystallographic data collection and refinement statistics for the FAM181A/B:TEAD4 complexes. [a]. Values as reported in autoPROC [1]. [b]. Highest resolution shell is shown in parentheses. [c]. Values as defined in BUSTER (Global Phasing Ltd., UK). [d]. The mysristate was modeled covalently bound to Lys344^{TEAD4}.

Supplementary Figures

Supplementary Fig. S1. Sequences of the motifs used to interrogate the databases. See Material and Methods for explanations. The residues in red indicate the positions for which the identification of alternative amino acids is evaluated in the corresponding motif search.

Supplementary Fig. S2. Inhibition curves obtained in the TR-FRET assay. The ability of the different peptides or protein fragments to inhibit the YAP:TEAD interaction was measured in a TR-FRET assay. Twelve stepwise dilutions (dilution factor 3.33) of each peptide/protein fragment were used in the experiments. The highest concentrations present in the assay were 222 μM for FAM181A$^{190-205}$, FAM181B$^{220-235}$ or YAP$^{84-99}$; 250 μM for YAP$^{85-99}$; 55 μM for FAM181A$^{127-205}$; 125 μM for FAM181B$^{157-237}$. The IC_{50} values were estimated by nonlinear regression analysis with GraphPad Prism (GraphPad Software, San Diego, CA). The stars indicate the signal measured in the absence of the peptide or protein fragments.

Supplementary Fig. S3. LC-MS analyses of FAM181A$^{127-205}$ and FAM181B$^{157-237}$. The figure represents HPLC profiles obtained with the purified FAM181A$^{127-205}$ and FAM181B$^{157-237}$ fragments. The numbers indicate the mass measured by mass spectrometry. The theoretical molecular weight calculated from the primary sequence are given in brackets. The amount of sample used for the analysis is indicated.

Supplementary Fig. S4. Binding of FAM181B$^{157-237}$ and FAM181A$^{127-205}$ to wt^{TEAD4}, Val389Ala^{TEAD4} and Asp272Ala^{TEAD4}. The biotinylated N-Avitagged TEAD4 proteins were immobilized on sensor chips, and the binding of different concentrations of FAM181B$^{157-237}$ or FAM181A$^{127-205}$ was measured at 298°K by Surface Plasmon Resonance. The upper panels show representative sensorgrams and the lower panels the corresponding binding isotherms from which K_d values (at equilibrium) were derived (excepted for Asp272Ala^{TEAD4}). The sensorgrams were globally fitted with a 1:1 interaction model using the Biacore T200

evaluation software (Biacore, Sweden). The concentrations used are indicated. The signal measured at equilibrium (R_{max}^{eq}) and the calculated maximum feasible signal (R_{max}^{th}) are given. Only apparent K_d values were determined with FAM181A$^{127-205}$ (see text).

Supplementary Fig. S5. Molecular dynamics simulation of the FAM181B$^{154-236}$:TEAD4$^{217-434}$ complex. The simulation of molecular dynamics was run using the Desmond module (Maestro package, Schrodinger Inc., Cambridge, MA). Ten snapshots of the 10 ns molecular dynamics simulation are shown. The initial, intermediate and final conformations of FAM181B$^{154-236}$ are in blue, orange and green, respectively. TEAD4$^{217-434}$ is represented in gray. The different secondary structure elements of FAM181B$^{154-236}$ are indicated. The picture was drawn with PyMOL (Schrödinger Inc., Cambridge, MA).

Supplementary Fig. S6. Structural study of FAM181A/B. **A.** Circular dichroism (CD) analysis of FAM181A$^{127-205}$ and FAM181B$^{157-237}$. The proteins were dialyzed in 20 mM phosphate buffer pH 7.4, 100 mM KF, 0.25 mM TCEP and diluted in this buffer to 0.2 mg.mL^{-1}. Far-UV CD spectra were recorded as previously described [2]. TEAD4$^{217-434}$ was used as an example of folded protein. The figure represents the average CD spectrum obtained from two independent experiments. **B.** PrDos analysis of full length FAM181A/B. The primary sequences of FAM181A (UniProt Q8N9Y4) and FAM181B (UniProt A6NEQ2) were analyzed with the protein disorder prediction server PrDOS (http://prdos.hgc.jp) [3] using the default settings. The upper panels are graphic representations of the disorder probability calculated for each residue. A probability score higher than 0.5 suggests that the residue is located in a disordered area. The lower panels correspond to the primary sequence of the two proteins with the residues located in regions predicted to be disordered indicated in red. The green arrows indicate FAM181A$^{127-205}$ and FAM181B$^{157-237}$.

References

1. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G (2011) Data processing and analysis with the autoPROC toolbox. Acta Crystallogr Sect D 67:293-302.

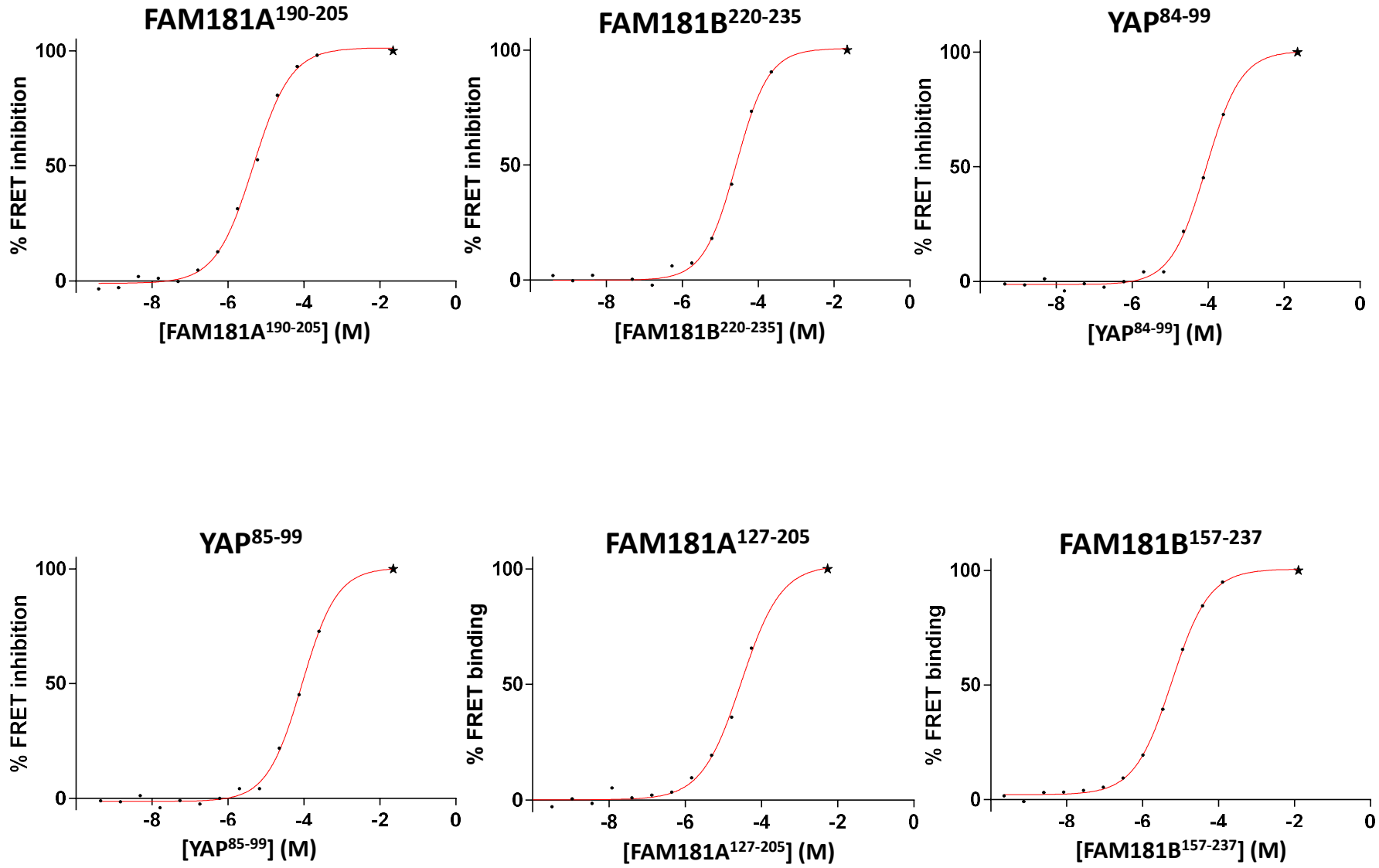
2. Mesrouze Y, Bokhovchuk F, Meyerhofer M, Fontana P, Zimmermann C, Martin T, Delaunay C, Erdmann D, Schmelzle T, Chène P (2017) Dissection of the interaction between the intrinsically disordered YAP protein and the transcription factor TEAD. eLife 6:e25068.

3. Ishida T, Kinoshita K (2007) PrDOS: prediction of disordered protein regions from amino acid sequence. Nucl Acids Res 35:W460-W464.

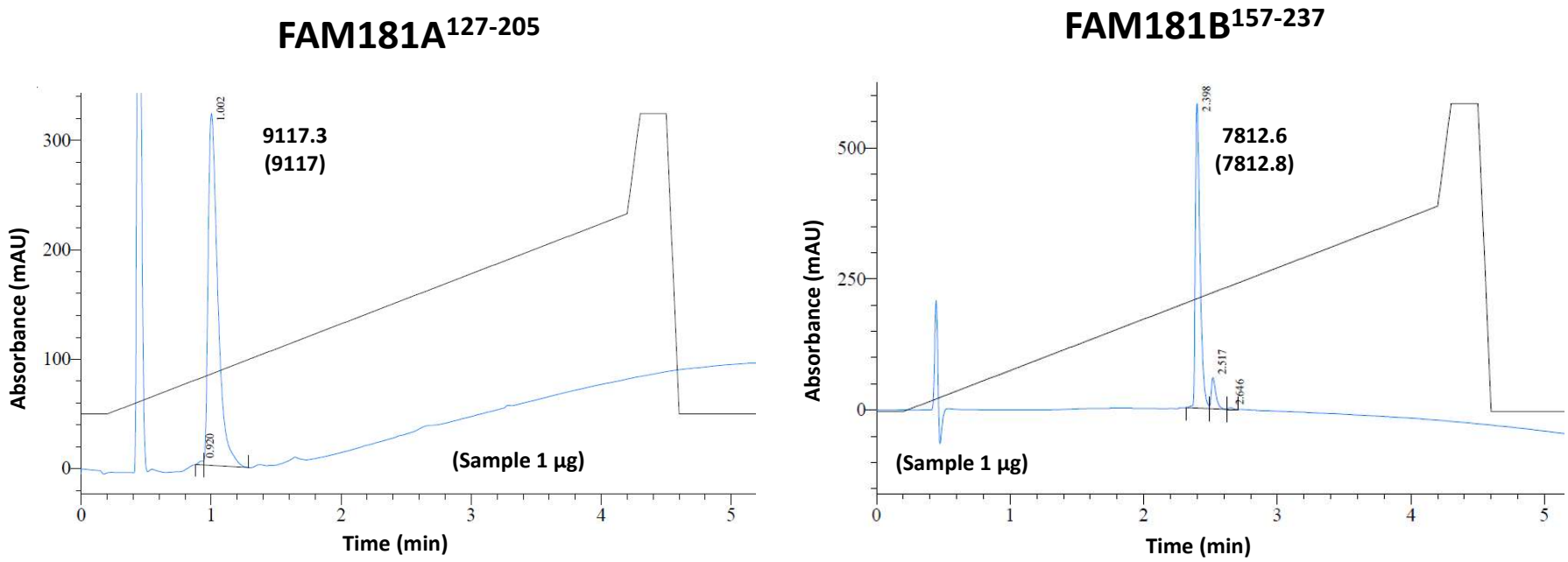
Supplementary Fig. S1

	Corresponding positions in YAP$^{85-99}$														
	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
Motif_01	P	x	R	x	R	x	[LM]	P	x	{S}	F	F	x	x	P
Motif_02	P	x	R	x	R	x	[LM]	P	x	S	{F}	F	x	x	P
Motif_03	P	x	R	x	R	x	[LM]	P	x	S	F	{F}	x	x	P
Motif_04	P	x	R	x	R	x	{LM}	P	x	S	F	F	x	x	P
Motif_05	{P}	x	R	x	R	x	[LMF]	P	x	S	F	[FW]	x	x	P
Motif_06	P	x	{R}	x	R	x	[LMF]	P	x	S	F	[FW]	x	x	P
Motif_07	P	x	[RKHS]	x	{R}	x	[LMF]	P	x	S	F	[FW]	x	x	P
Motif_08	P	x	[RKHS]	x	R	x	[LMF]	P	x	{S}	F	[FW]	x	x	P
Motif_09	P	x	[RKHS]	x	R	x	[LMF]	P	x	S	{F}	[FW]	x	x	P
Motif_10	P	x	[RKHS]	x	R	x	[LMF]	{P}	x	S	F	[FW]	x	x	P
Motif_11	{P}	x	[RKHS]	x	R	x	[LMF]	P	x	S	F	[FW]	x	x	P
Motif_12	P	x	[RKHS]	x	R	x	[LMF]	P	x	S	F	[FW]	x	x	{P}
Motif_13	P	x	[RKHS]	x	R	x	[LMF]	P	x	S	F	{FW}	x	x	P
Motif_14	P	x	[RKHS]	x	R	x	{LMF}	P	x	S	F	[FW]	x	x	P
Final	P	x	[RKHS]	x	R	x	[LMF]	P	x	S	F	[FW]	x	x	P

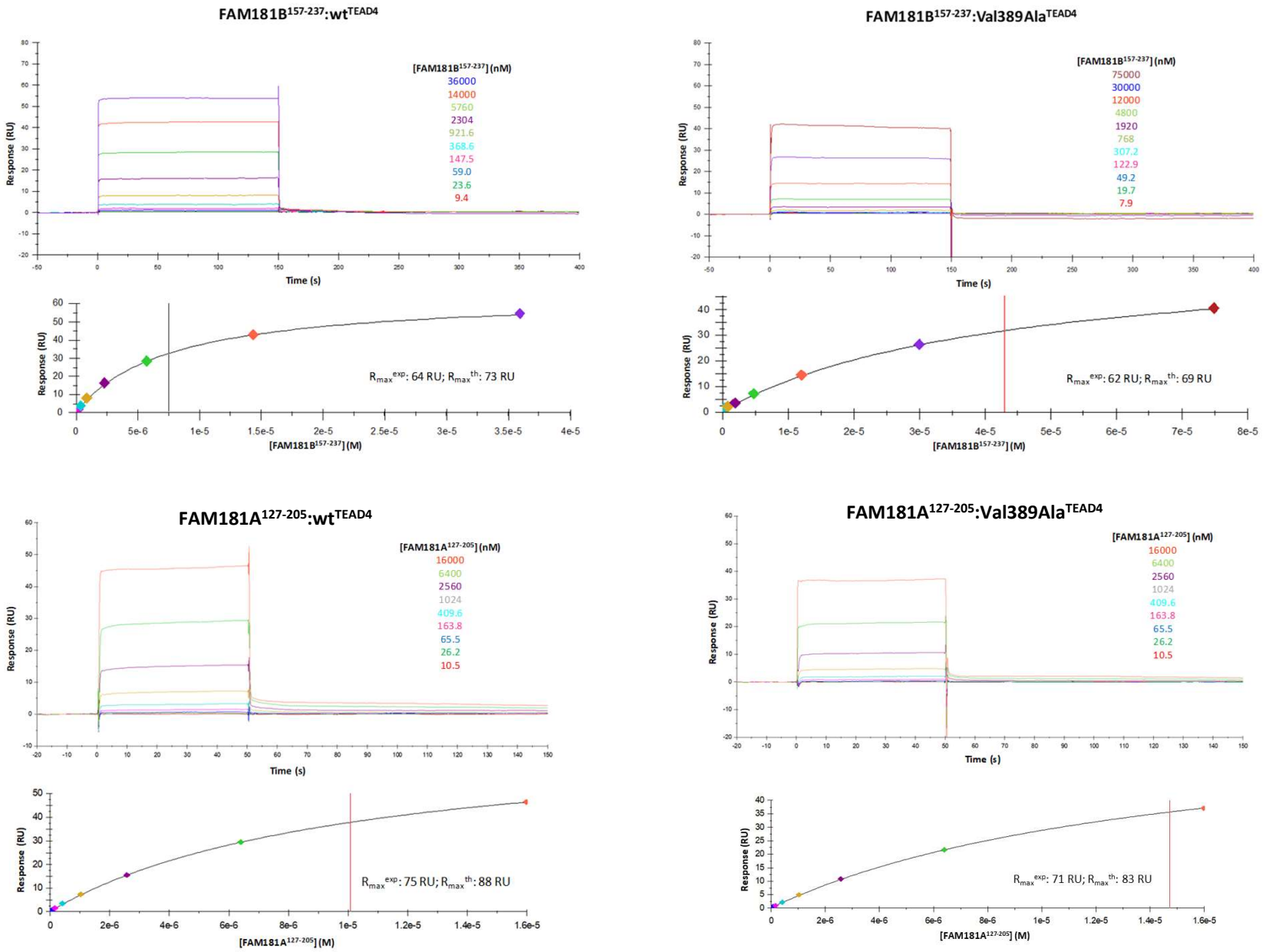
Supplementary Fig. S2



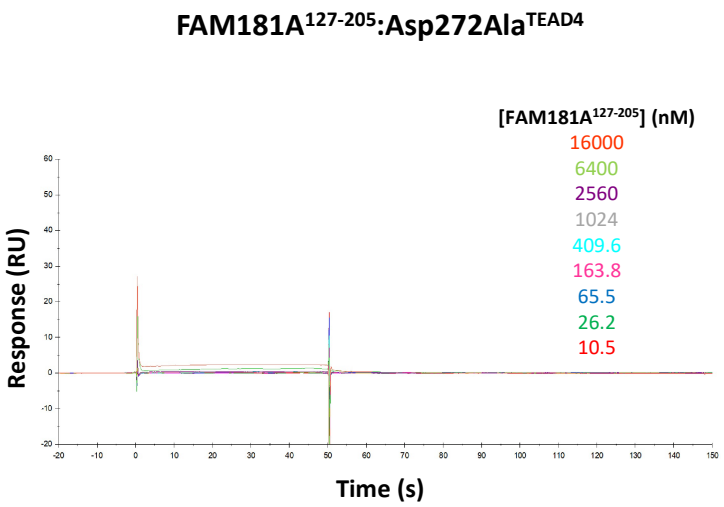
Supplementary Fig. S3

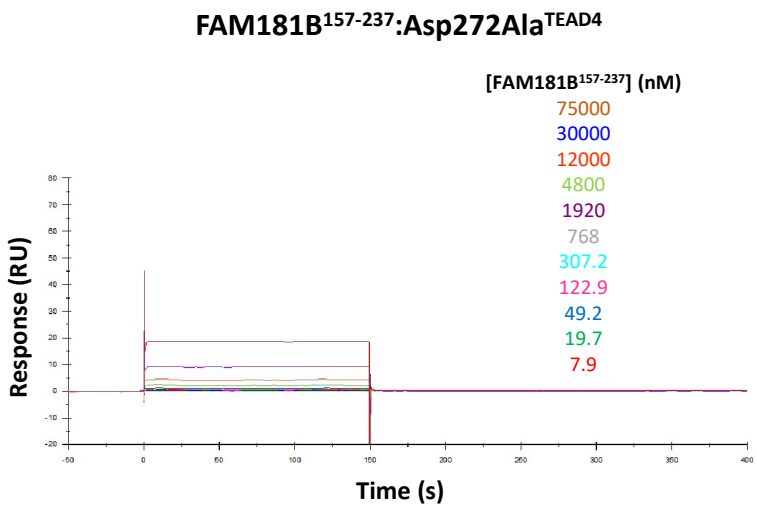


Supplementary Fig. S4

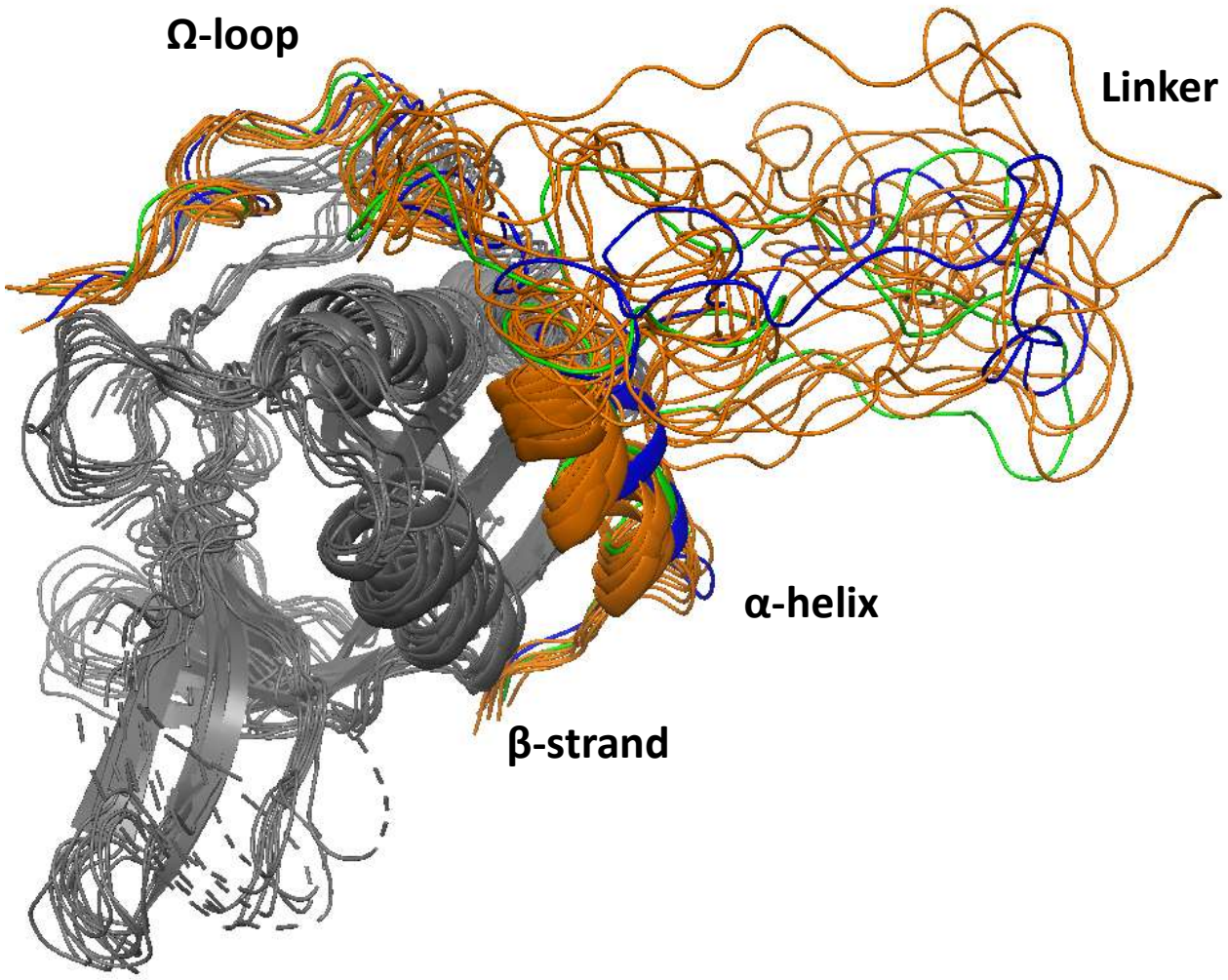


Supplementary Fig. S4

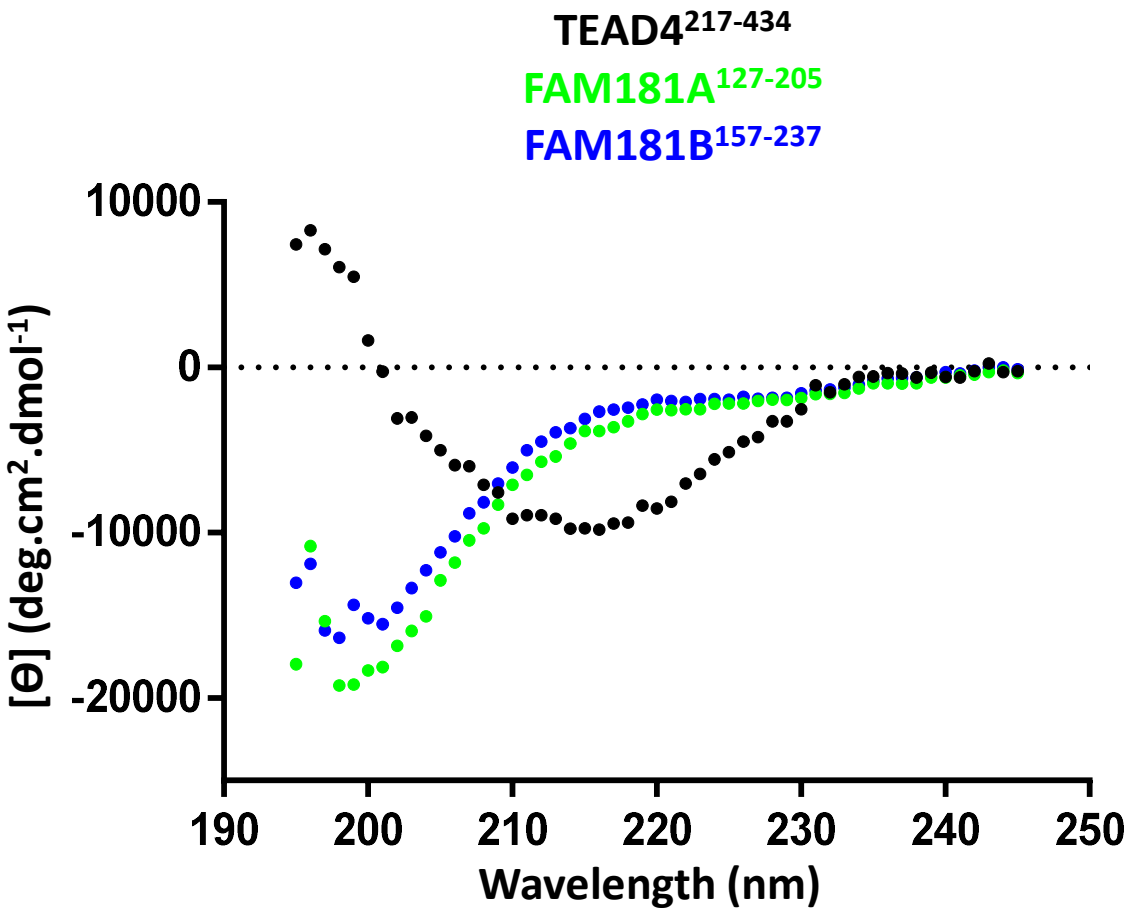




Supplementary Fig. S5

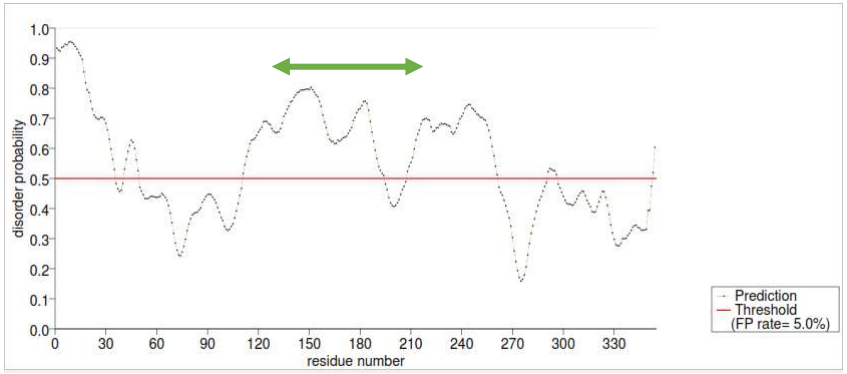


Supplementary Fig. S6A



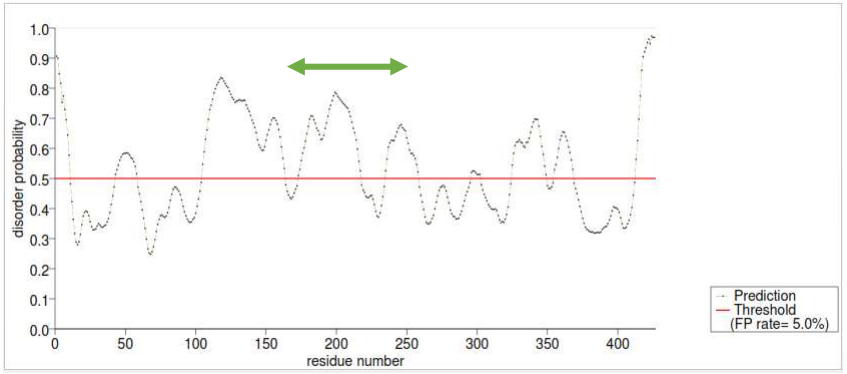
Supplementary Fig. S6B

FAM181A




```
1   MPLEERRSSG  ERNDAAPTNH  RRPGEKRAST  AKQVSSVPFL  GAAGHQQSLP  50
51  SSWKASCSGP  LVMASDSDVK  MLLNFVNLAS  SDIKAALDKS  APCRRSVDHR  100
101 KYLQKQLKRF  SQKYSRLPRG  LPGRAAEPYL  KRGSEDRPRR  LLLDLGPDSS  150
151 PGGGGGCKEK  VLRNPYREEC  LAKEQLPQRQ  HPEAAQPGQV  PMRKRQLPAS  200
201 FWEEPRPTHS  YHVGLEGGLG  PREGPPYEGK  KNCKGLEPLG  PETTLVSMSP  250
251 RALAEKEPLK  MPGVSLVGRV  NAWSCCPFQY  HGQPIYPGPL  GALPQSPVPS  300
301 LGLWRKSPAF  PGELAHLCKD  VDGLGQKVCR  PVVLKPIPTK  PAVPPPIFNV  350
351 FGYL                                                      400
```

FAM181B




```
1   MAVQAALLST  HPFVPFGFGG  SPDGLGGAFG  ALDKGCCFED  DETGAPAGAL  50
51  LSGAEGGDVR  EATRDLLSFI  DSASSNIKLA  LDKPGKSKRK  VNHRKYLQKQ  100
101 IKRCSGLMGA  APPGPPSPSA  ADTPAKRPLA  APSAPTVAAP  AHGKAAPRRE  150
151 ASQAAAAASL  QSRSLAALFD  SLRHVPGGAE  PAGGEVAAPA  AGLGGAGTGG  200
201 AGGDVAGPAG  ATAIPGARKV  PLRARNLPPS  FFTEPSRAGG  GGCGPSGPDV  250
251 SLGDLEKGAE  AVEFFELLGP  DYGAGTEAAV  LLAAEPLDVF  PAGASVLRGP  300
301 PELEPGLFEP  PPAVVGNLLY  PEPWSVPGCS  PTKKSPLTAP  RGGLTLNEPL  350
351 SPLYPAAADS  PGGEDGRGHL  ASFAPFFPDC  ALPPPPPPHQ  VSYDYSAGYS  400
401 RTAYSSLWRS  DGVWEGAPGE  EGAHRD                              450
```