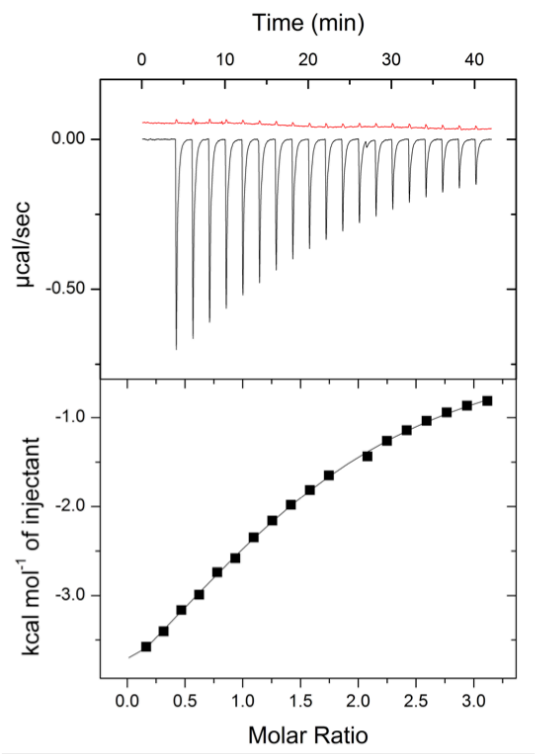
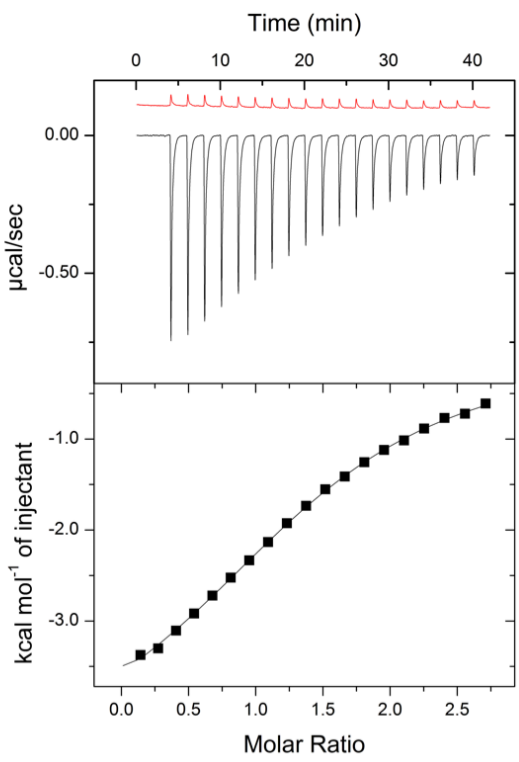
SUPPLEMENTARY FIGURES AND TABLES

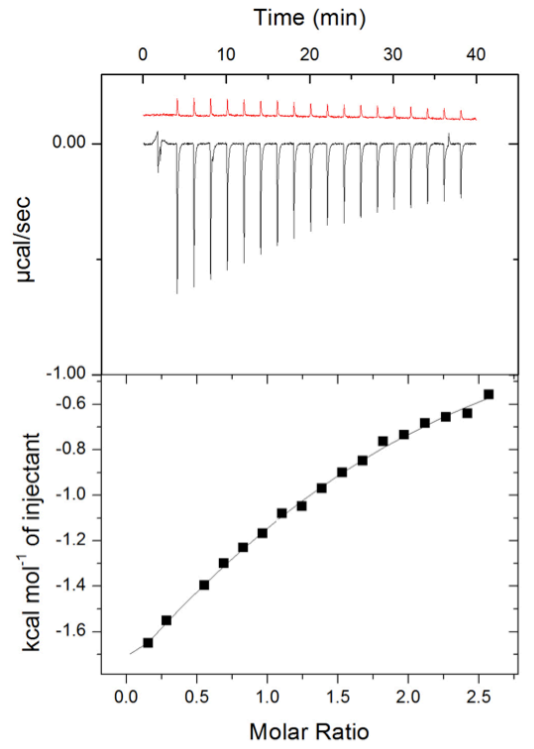
H3 4-mer (ARTK) vs BAZ2A PHD



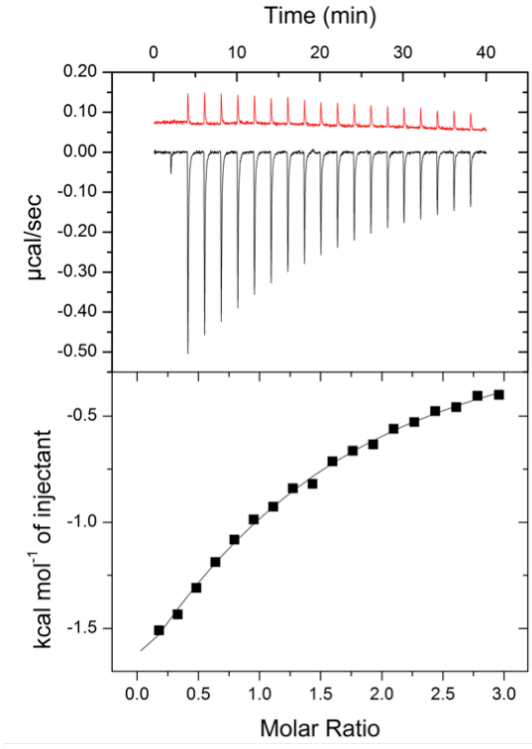
H3 4-mer (ARTA) vs BAZ2A PHD



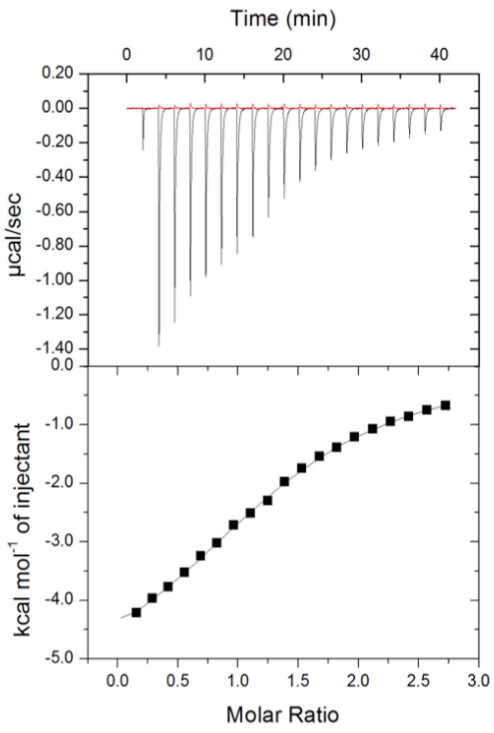
H3 5-mer (ARTKQ) vs BAZ2A PHD



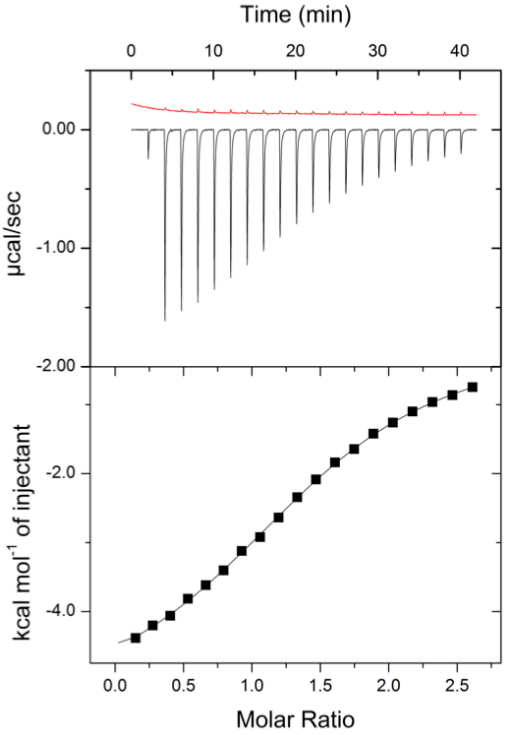
H3 5-mer (ARTKQ) vs BAZ2B PHD



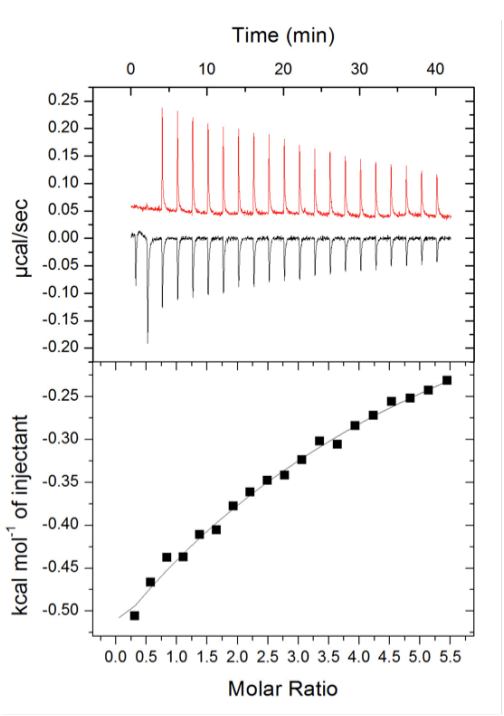
H3 10-mer (ARTKQTARKS) vs BAZ2A PHD



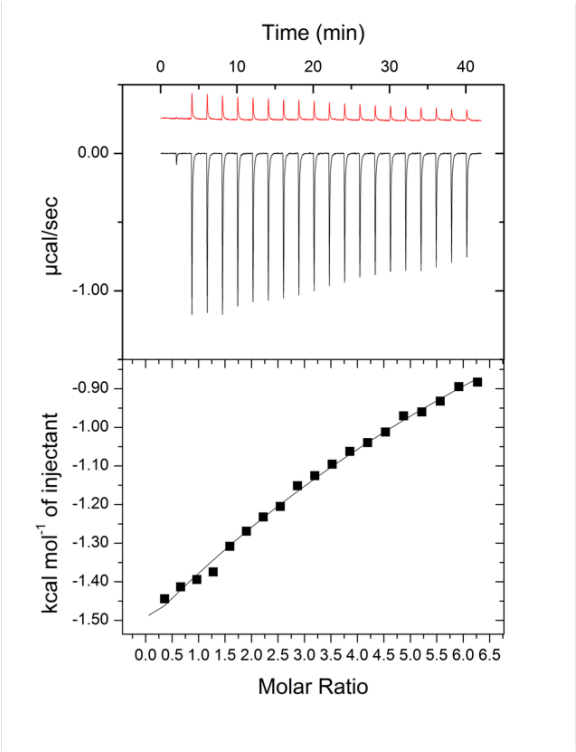
H3 10-mer (ARTKQTARKS) vs BAZ2B PHD



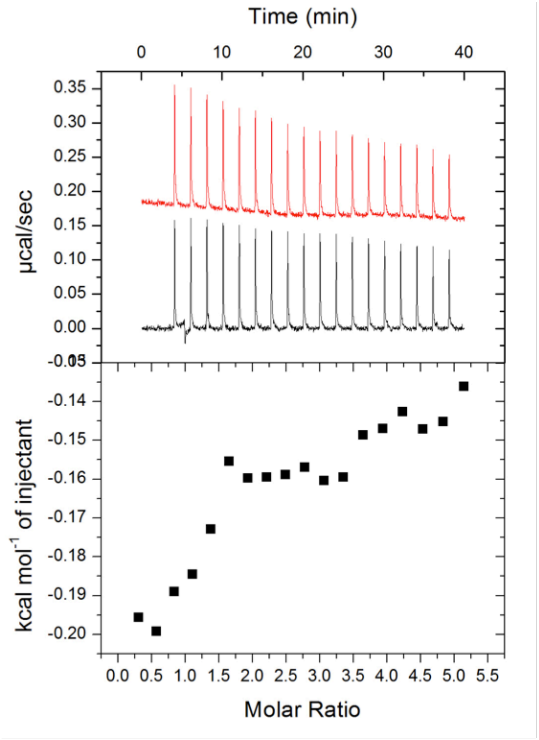
H3 10-mer (AATKQTARKS) vs BAZ2A PHD



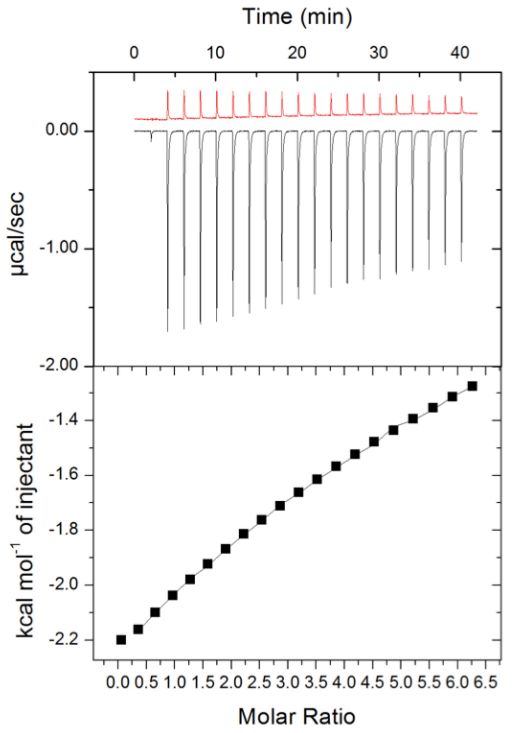
H3 10-mer (AATKQTARKS) vs BAZ2B PHD



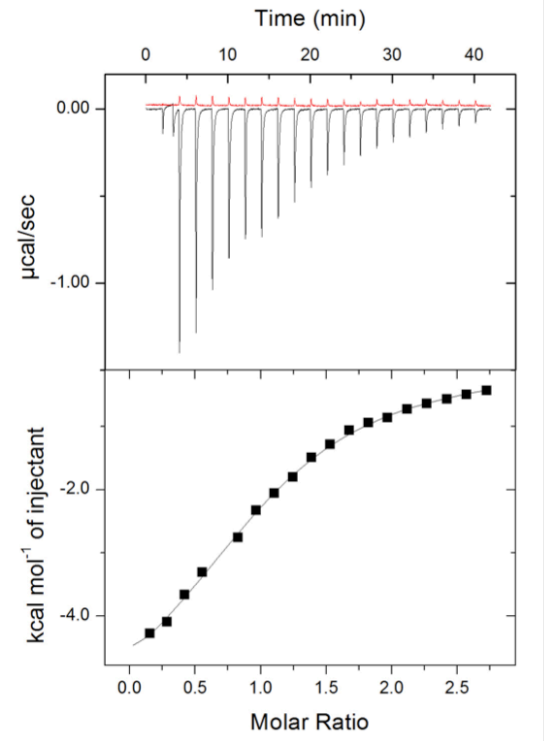
H3 10-mer (ARAKQTARKS) vs BAZ2A PHD



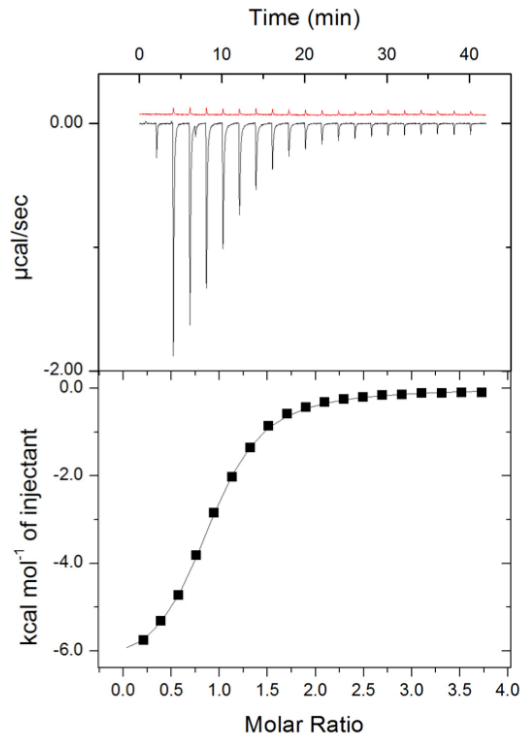
H3 10-mer (ARAKQTARKS) vs BAZ2B PHD



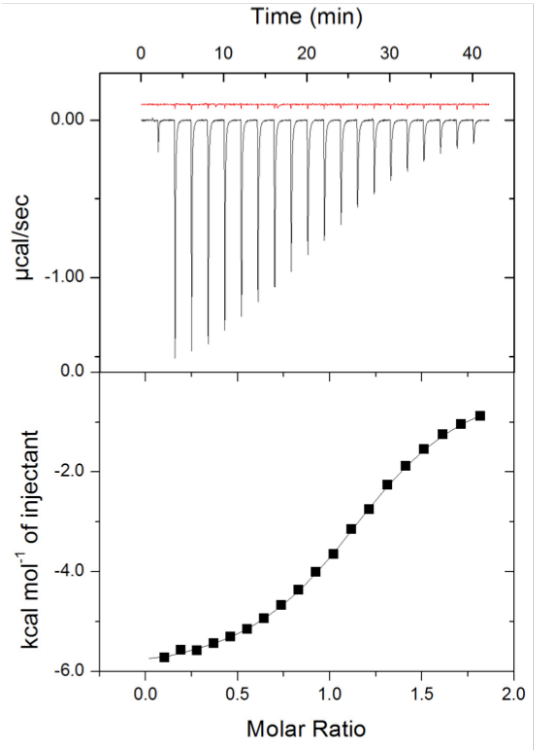
H3 10-mer (ARTAQTARKS) vs BAZ2A PHD



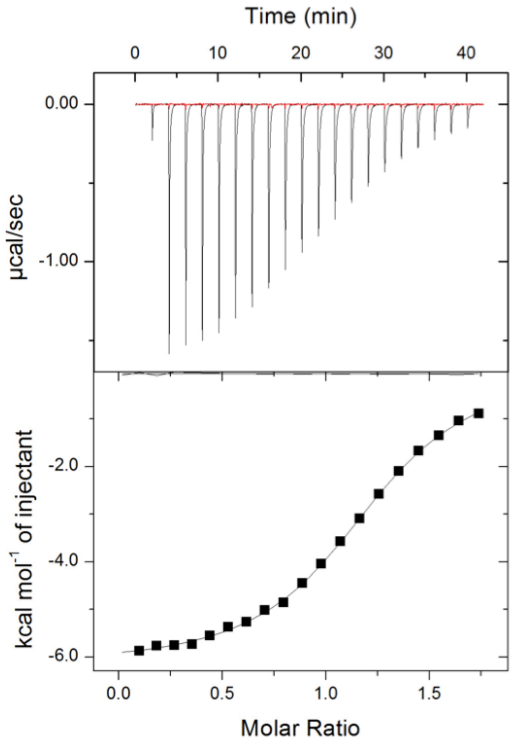
H3 10-mer (ARTAQTARKS) vs BAZ2B PHD



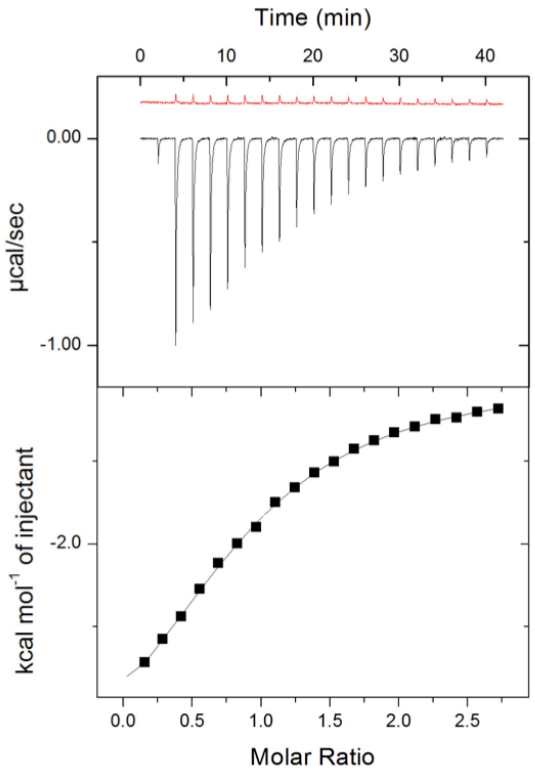
H3 10-mer (ARTKATARKS) vs BAZ2A PHD



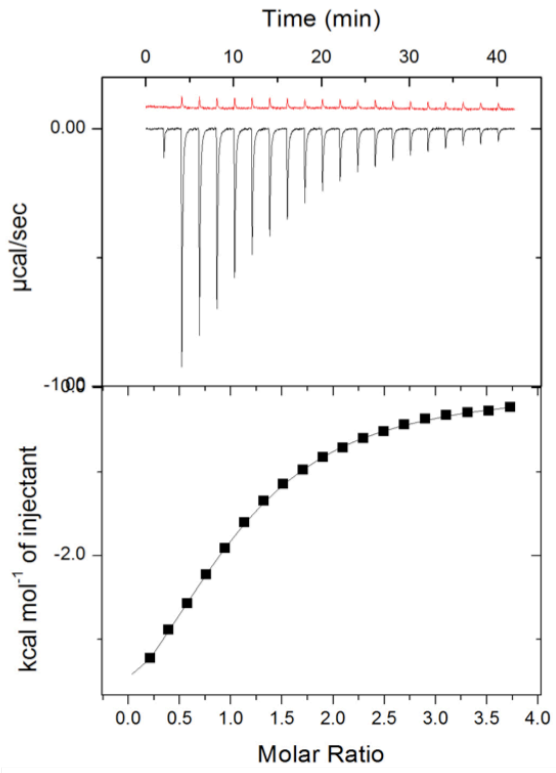
H3 10-mer (ARTKATARKS) vs BAZ2B PHD



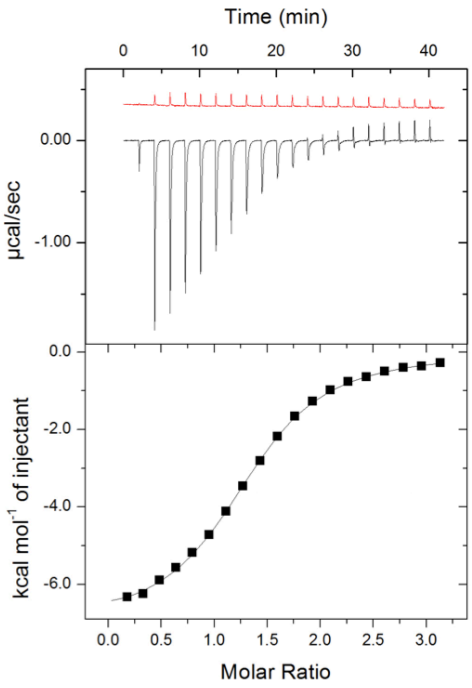
H3 10-mer (ARTKQAARKS) vs BAZ2A PHD



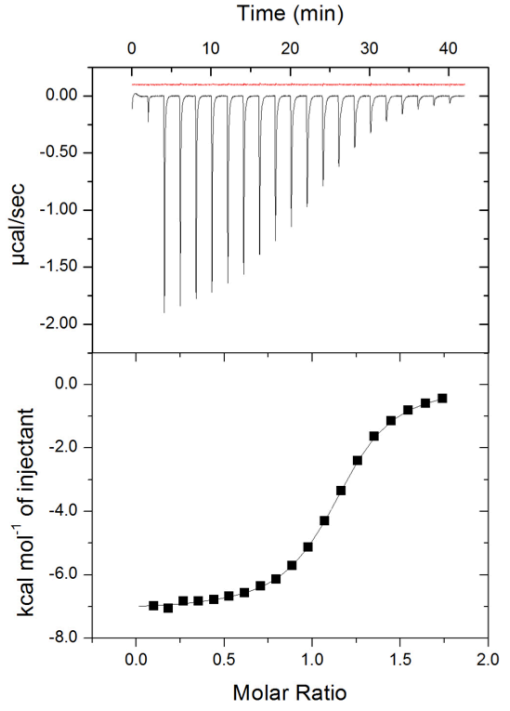
H3 10-mer (ARTKQAARKS) vs BAZ2B PHD



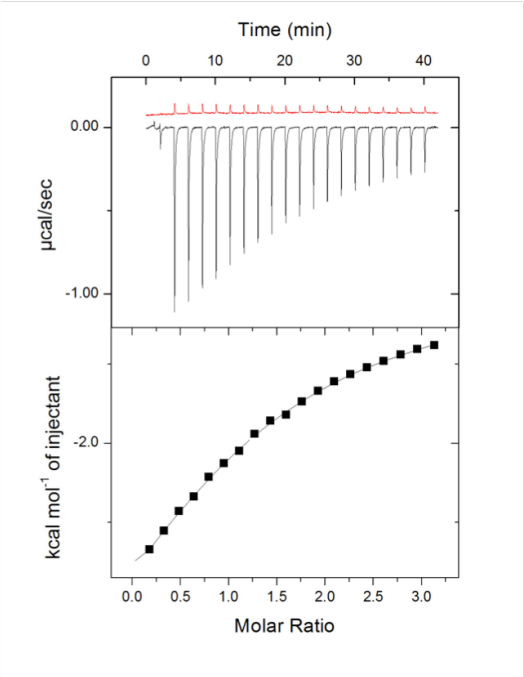
H3 10-mer (ARTAATARKS) vs BAZ2A PHD



H3 10-mer (ARTAATARKS) vs BAZ2B PHD



H3 10-mer (ARTGGTARKS) vs BAZ2A PHD



H3 10-mer (ARTGGTARKS) vs BAZ2B PHD

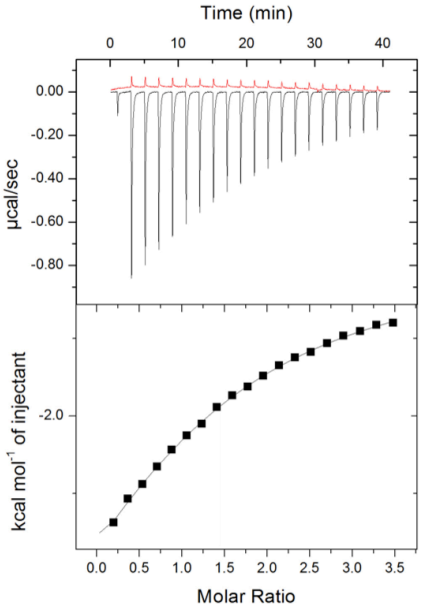
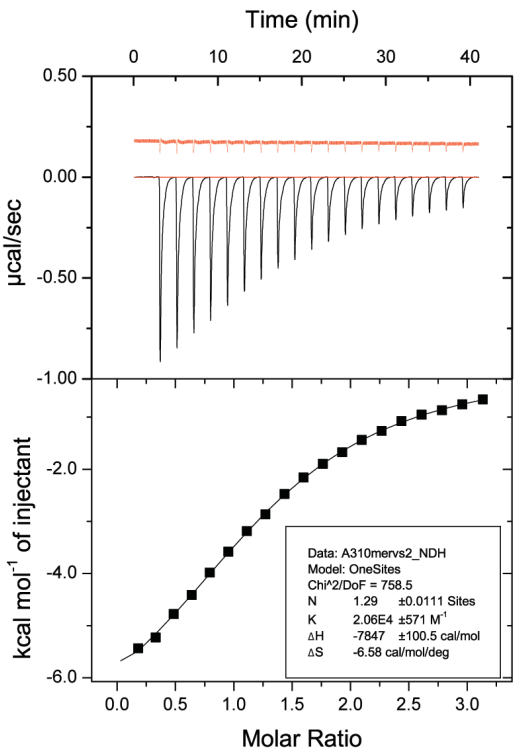
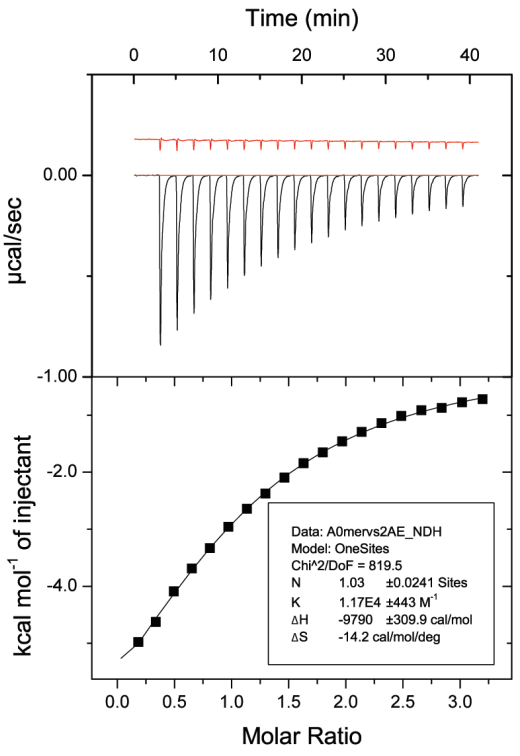


Figure S1. ITC binding curves of BAZ2A/B PHD fingers with different H3-derived peptides are shown in black and the relevant reference titrations (peptide into buffer) are shown in red in the upper panel. The integrated ΔH (kcal/mol) values are plotted versus the peptide/protein molar ratio and shown in the lower panel.

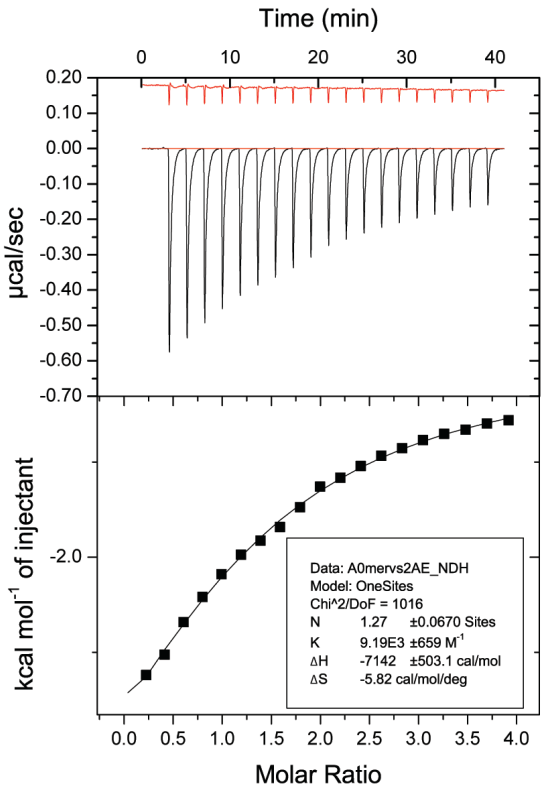
H3 10-mer vs BAZ2A PHD wt



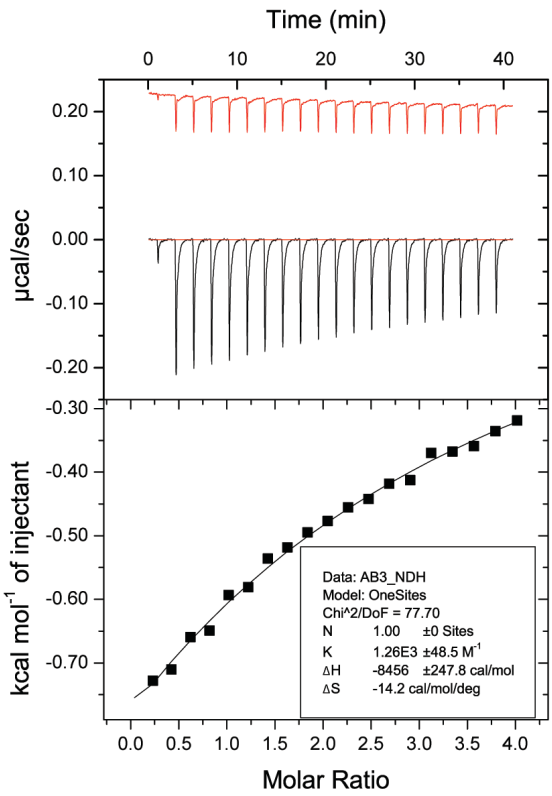
H3 10-mer vs BAZ2A PHD E1689Q



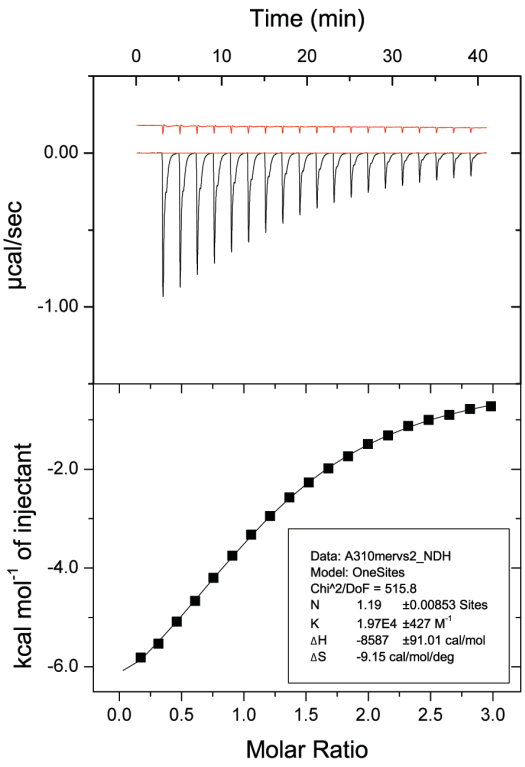
H3 10-mer vs BAZ2A PHD E1689K



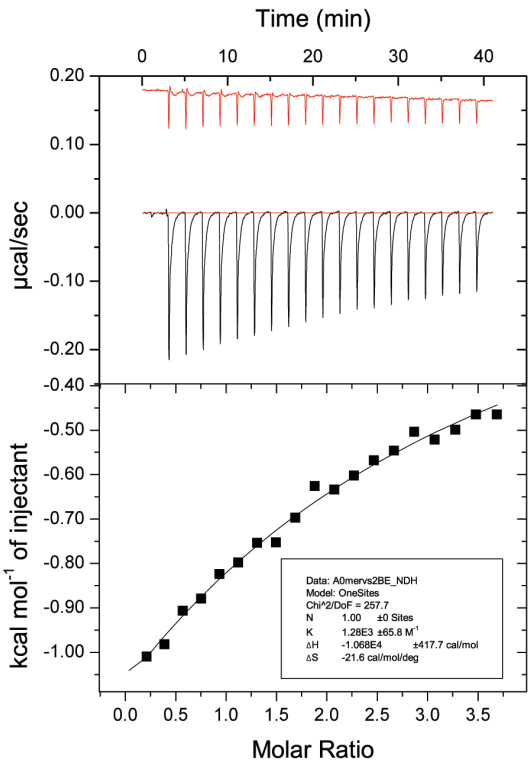
H3 10-mer vs BAZ2A PHD D1688N/E1689Q



H3 10-mer vs BAZ2B PHD wt



H3 10-mer vs BAZ2B PHD E1689Q



H3 10-mer vs BAZ2B PHD E1689K

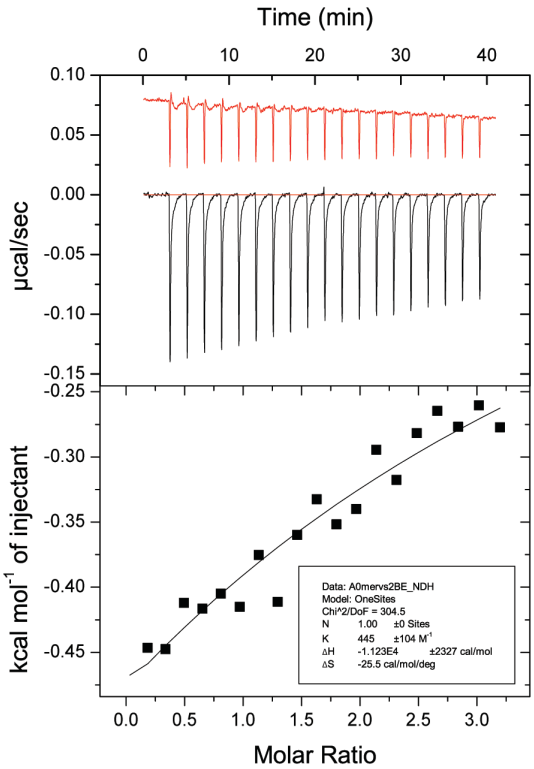


Figure S2. Representative ITC binding curves of wild-types and mutants BAZ2A/B PHD fingers with H3 10 mer peptide (ARTKQTARKS) are shown in black and the relevant reference titrations (peptide into buffer) are shown in red. In the lower panel, the integrated ΔH (kcal/mol) values plotted versus the peptide/protein molar ratio.

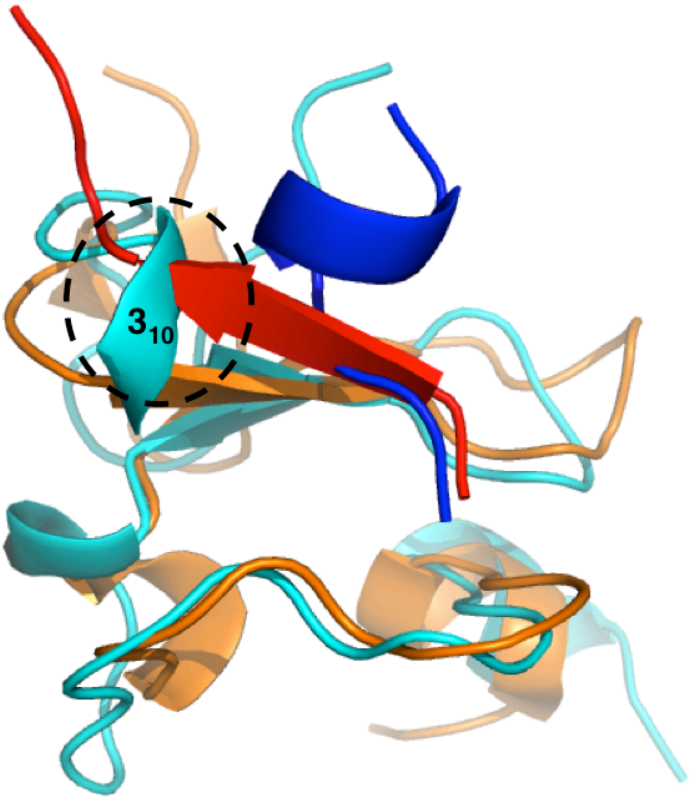


Figure S3. Structural superposition of the PHD fingers of BAZ2A (shown in cyan, PDB 5T8R) and ING2 (shown in orange, PDB: 2G6Q) in complex with an H3 N-terminal tail peptide (shown in blue for BAZ2A PHD and in red for ING2 PHD). A dotted black circle highlights the region where the H3 N-terminal peptide would clash with the 3_{10} helix of BAZ2A PHD if the peptide assumed an extended conformation as the one observed in complex with ING2 PHD.

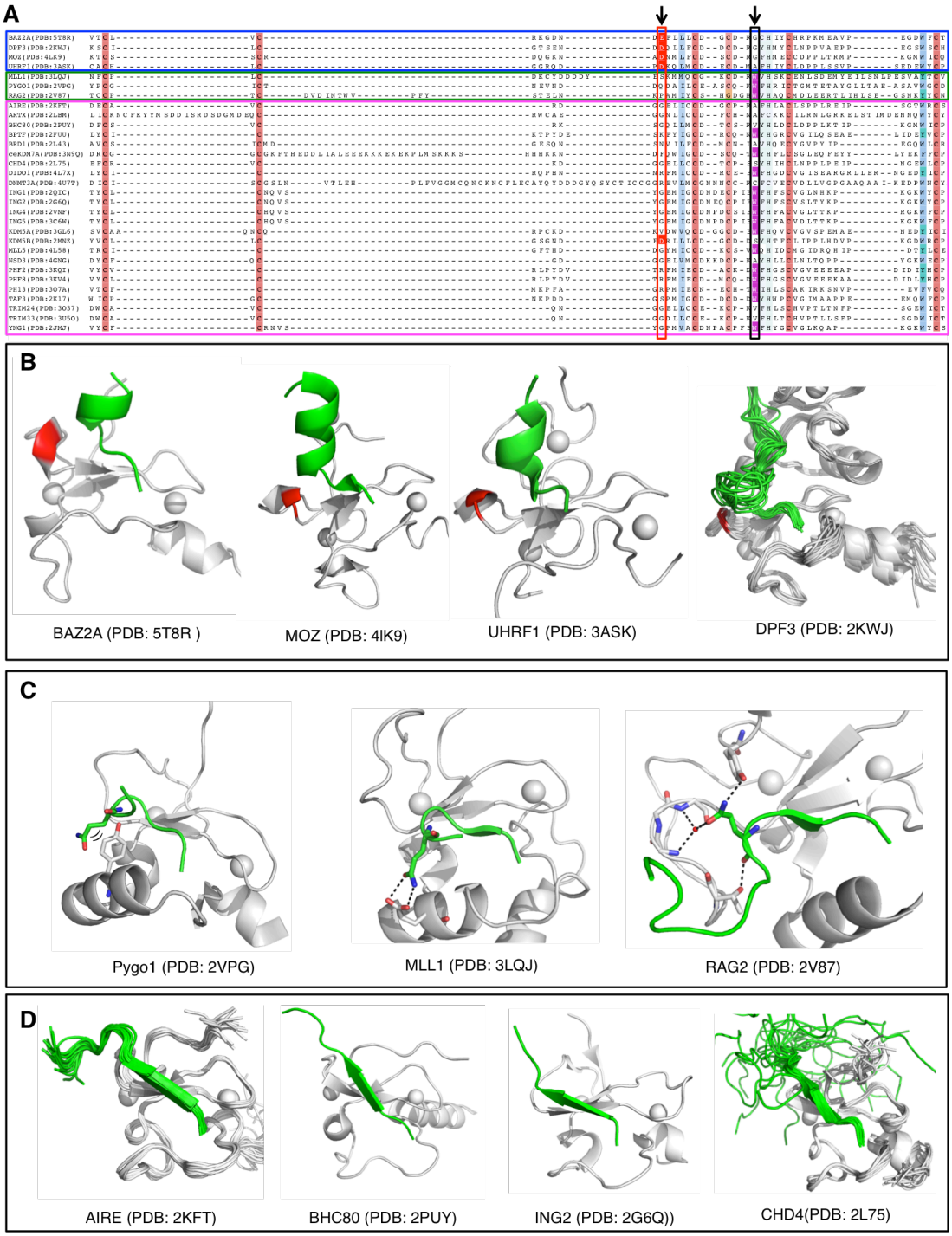
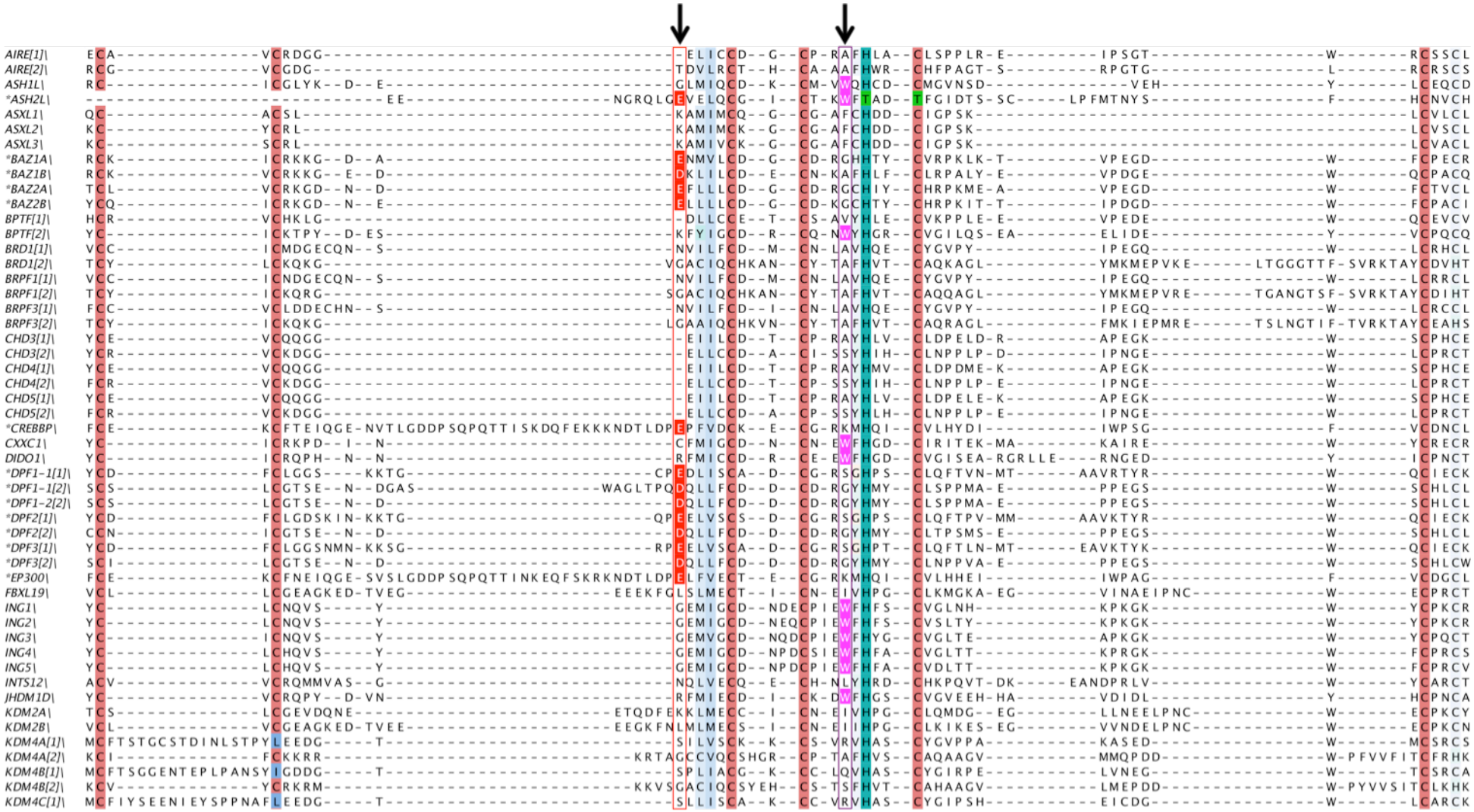
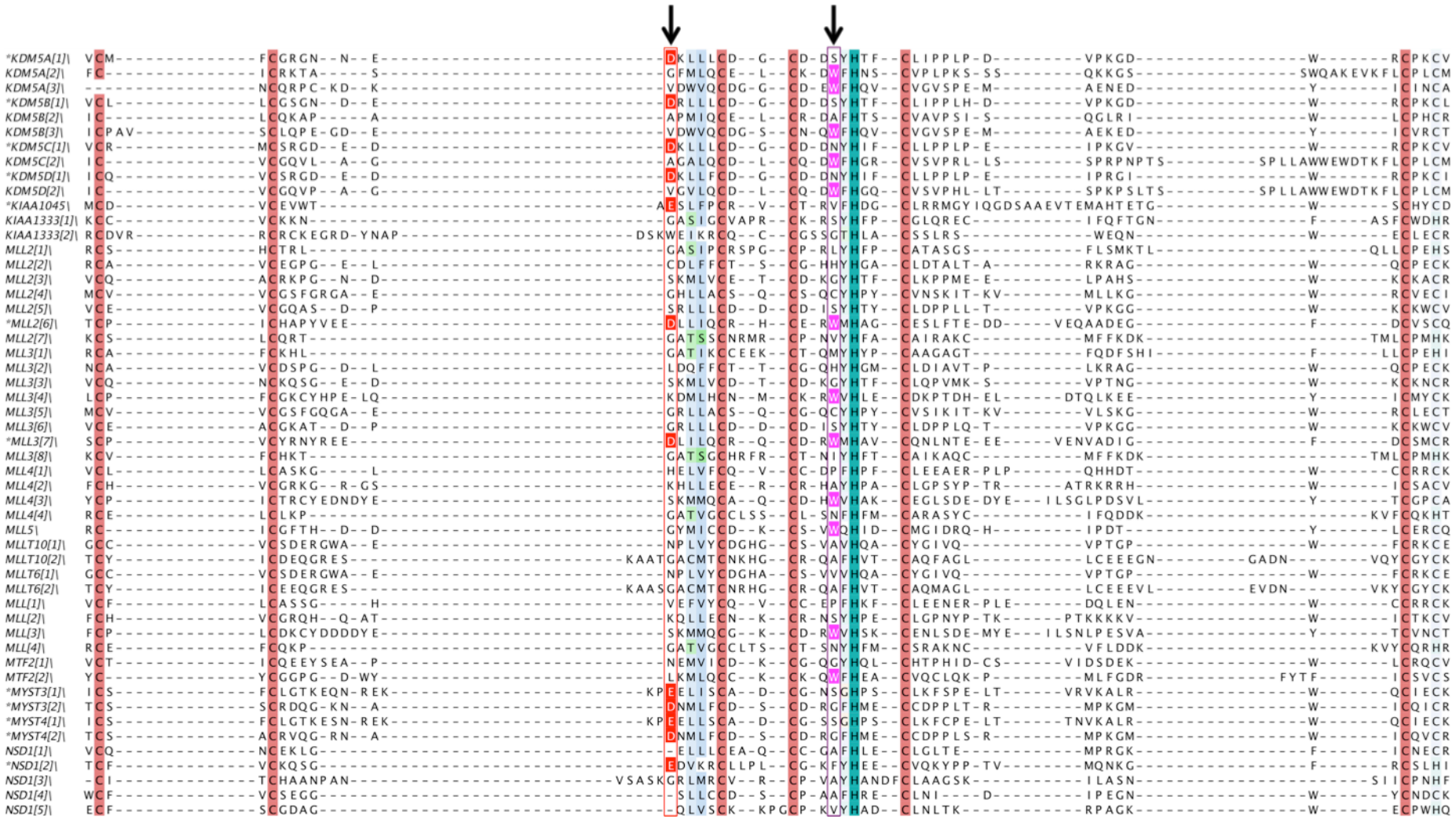
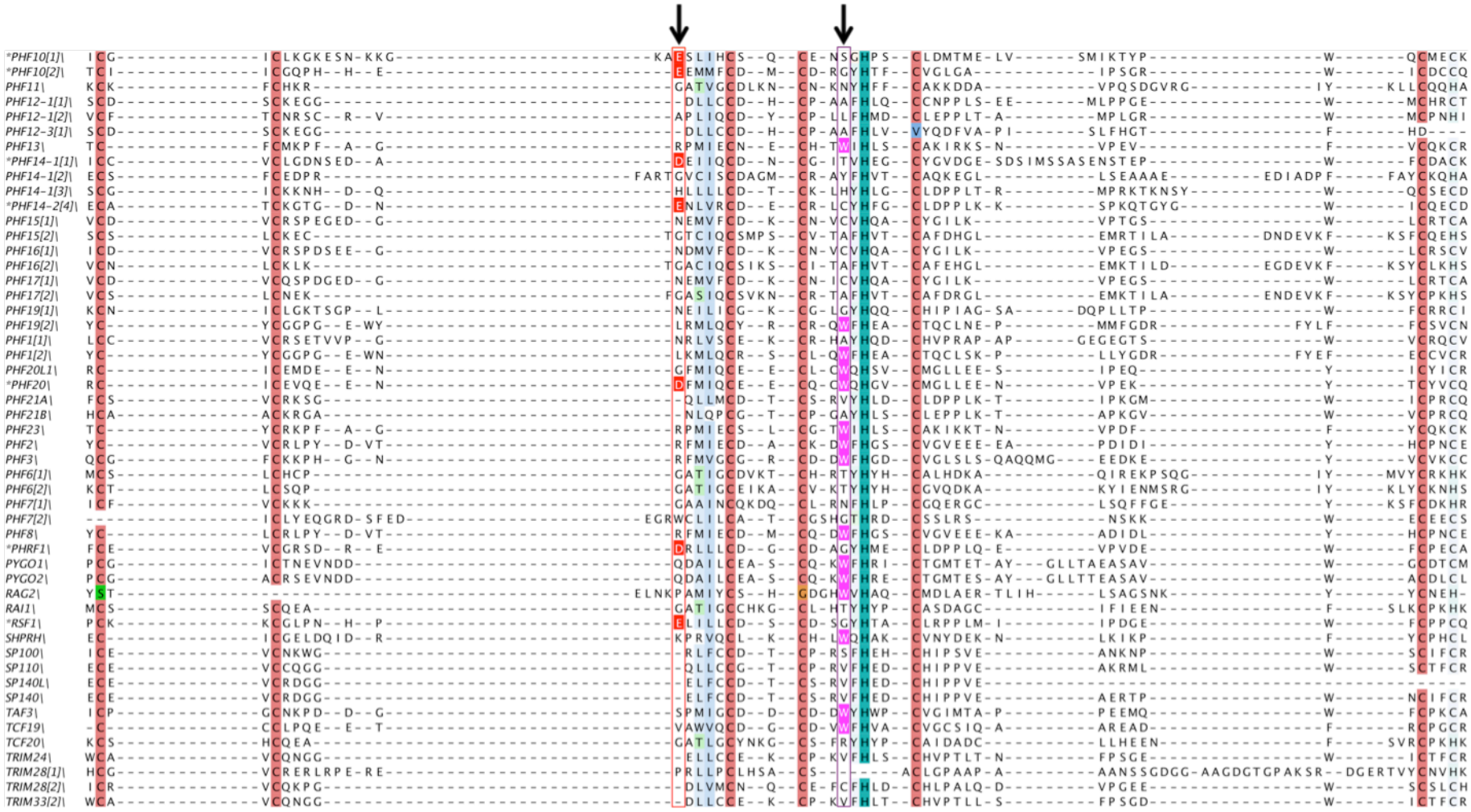


Figure S4. (A) Sequence alignment of all the PHD fingers whose structure was solved in complex with an H3 N-terminal tail peptide. Residues corresponding to E1689 of BAZ2A and residues corresponding to the absolutely conserved Trp of PHD fingers that recognize methylated-K4 are highlighted through the alignment with a red and a black box

respectively. PHD fingers in a blue box induce the H3 N-terminal peptide to adopt a helical folded-back conformation, in a green box induce H3 N-terminal peptide to adopt a bent conformation and in a magenta box bind H3 N-terminal peptide into an extended conformation **(B)** Structures of PHD fingers (shown in gray) that have an acidic residue (shown in red) in the position corresponding to E1689 of BAZ2A and that induce the H3 N-terminal peptide (shown in green) to adopt an helical folded-back conformation. **(C)** Structures of PHD fingers that induce the H3 N-terminal peptide to bend. In contrast to the recognition of helical folded-back H3 tail, the bent conformation of H3 bound to PHD appears to be stabilized by a different set of interactions. The interactions between PHD finger and H3 N-terminal peptide that stabilize the bent conformation are highlighted. Dotted lines represent hydrogen bonds and double brackets hydrophobic interactions. **(D)** Representative structures of PHD fingers that bind H3 N-terminal peptide in an extended conformation.


AIRE[1]\
AIRE[2]\
ASH1L\
*ASH2L\
ASXL1\
ASXL2\
ASXL3\
*BAZ1A\
*BAZ1B\
*BAZ2A\
*BAZ2B\
BPTF[1]\
BPTF[2]\
BRD1[1]\
BRD1[2]\
BRPF1[1]\
BRPF1[2]\
BRPF3[1]\
BRPF3[2]\
CHD3[1]\
CHD3[2]\
CHD4[1]\
CHD4[2]\
CHD5[1]\
CHD5[2]\
*CREBBP\
CXXC1\
DIDO1\
*DPF1-1[1]\
*DPF1-1[2]\
*DPF1-2[2]\
*DPF2[1]\
*DPF2[2]\
*DPF3[1]\
*DPF3[2]\
*EP300\
FBXL19\
ING1\
ING2\
ING3\
ING4\
ING5\
INTS12\
JHDM1D\
KDM2A\
KDM2B\
KDM4A[1]\
KDM4A[2]\
KDM4B[1]\
KDM4B[2]\
KDM4C[1]\





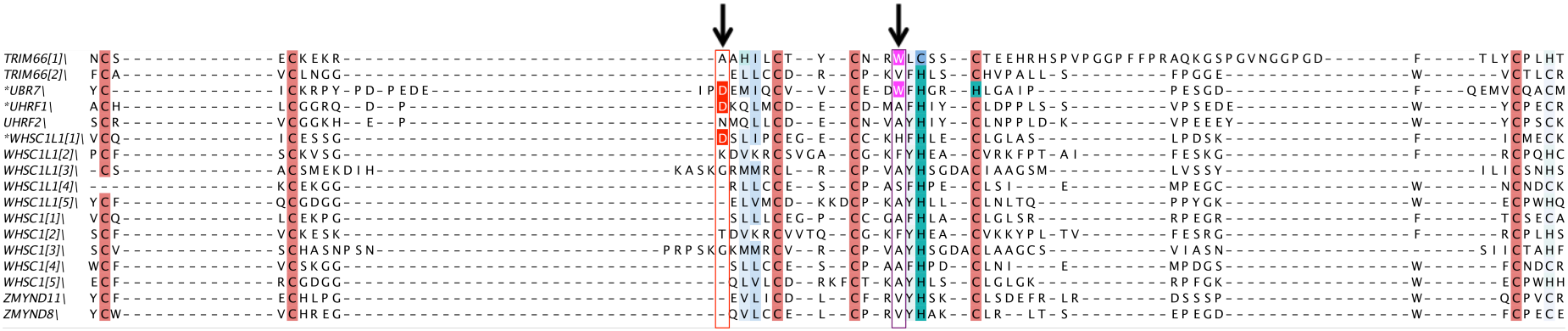


Figure S5. Sequence alignment of human PHD fingers (Structural Genomics Consortium database, at http:// www.thesgc.or). The column corresponding to E1689 of BAZ2A is highlighted through the alignment with a red box and a black arrow, and Asp or Glu residues in this column are colored in red. The column corresponding to the absolutely conserved Trp in PHD fingers that recognize methylated-K4 is highlighted through the alignment with a magenta box and a black arrow, and Trp residues in this column are colored in magenta. Sequences that have an Asp or a Glu in the position corresponding to E1689 of BAZ2A are marked with an asterisk (*).

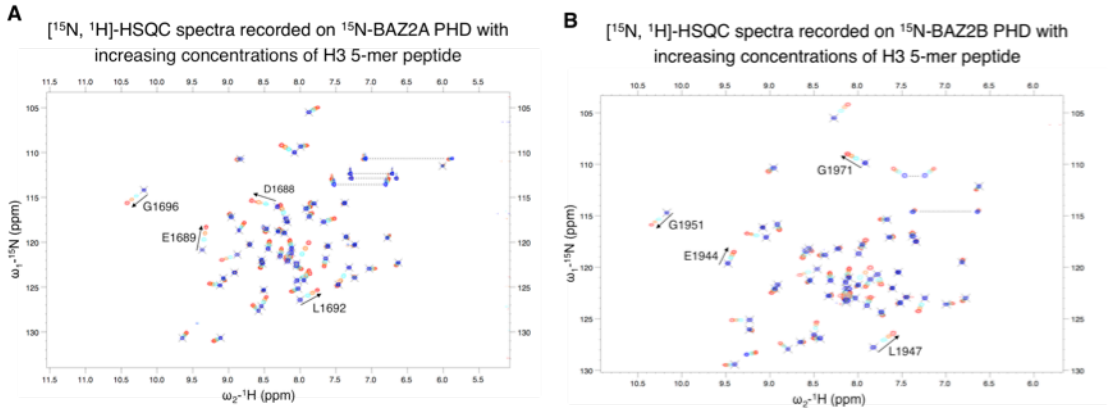
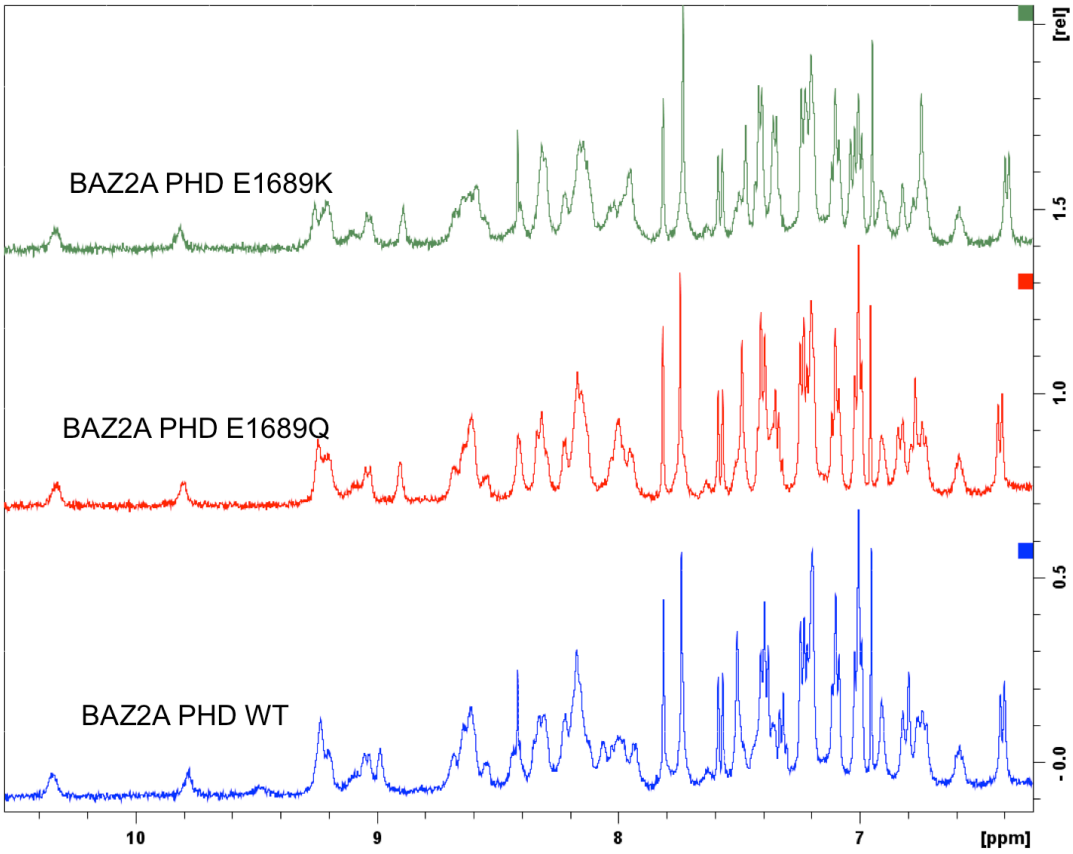
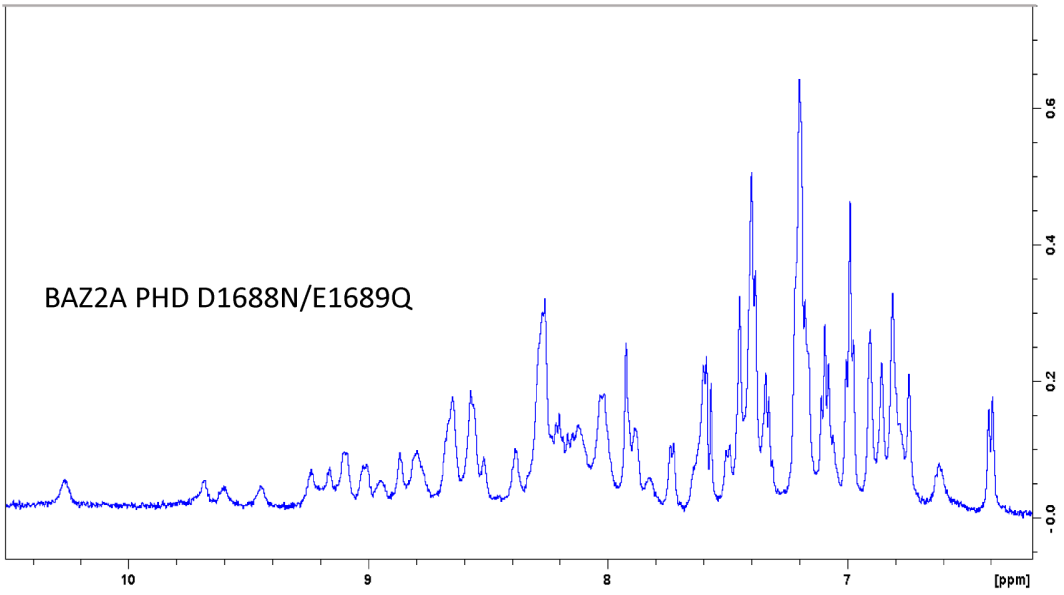


Figure S6. Characterization of the interaction between the H3 N-terminal tail and BAZ2A PHD and BAZ2B PHD in solution by NMR. (A and B) Overlay of [^{15}N, ^{1}H]-HSQC spectra recorded on ^{15}N-BAZ2A PHD **(A)** and on ^{15}N-BAZ2B PHD **(B)** with increasing concentrations of H3 5-mer peptide (ARTKQ). Spectra were recorded at the following protein:peptide molar ratios: 1:0 (blue spectra), 1:2 (cyan), 1:4 (orange) and 1:8 (red). For a set of peaks the direction of the shifts are indicated with black arrows. All the backbone amide protons of BAZ2A PHD and BAZ2B PHD were assigned except for the first serine residue of BAZ2A PHD (BMRB deposition numbers: 26754 and 25988, for BAZ2A PHD and BAZ2B PHD, respectively).


BAZ2A PHD E1689K
BAZ2A PHD E1689Q
BAZ2A PHD WT
[rel]
1.5
1.0
0.5
- 0.0
10
9
8
7
[ppm]


BAZ2A PHD D1688N/E1689Q
0.6
0.4
0.2
- 0.0
10
9
8
7
[ppm]

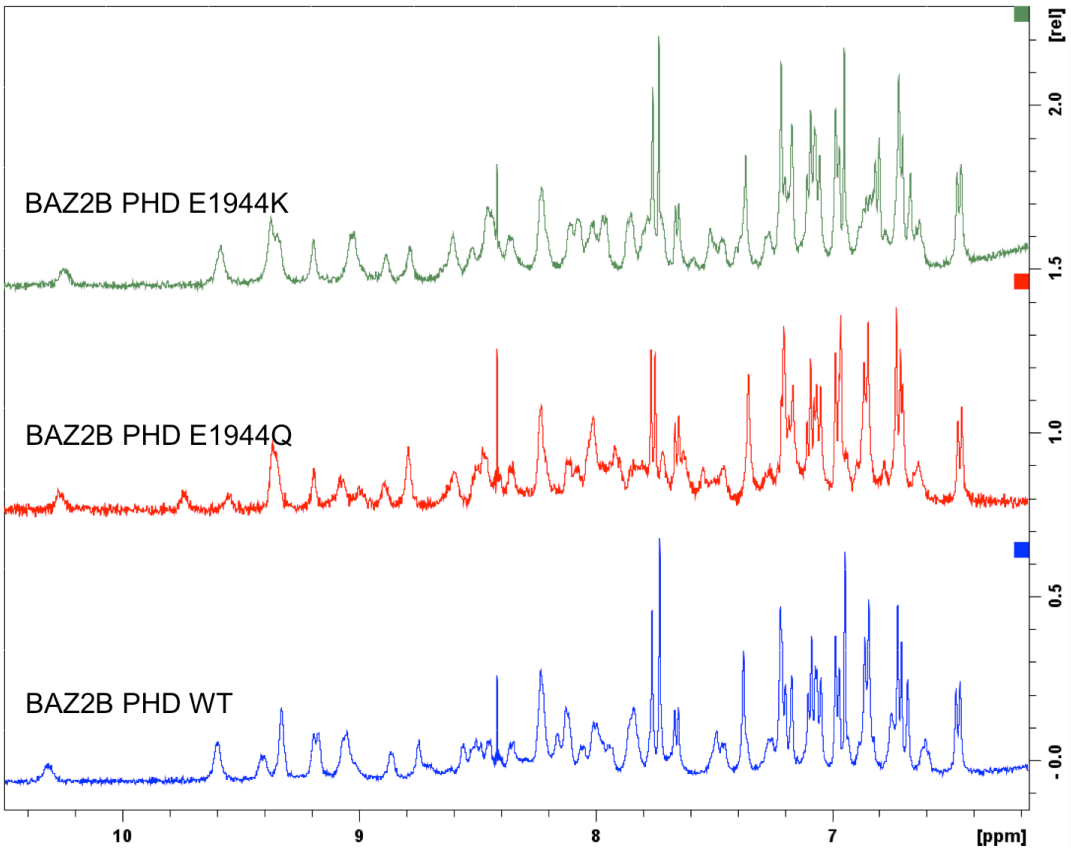


Figure S7. ^{1}H 1D NMR spectra were recorded on WT and mutants BAZ2A and BAZ2B PHD fingers and all the samples show a significant dispersion of the signals from the backbone NH groups suggesting that the proteins are folded. The spectra were recorded on samples at a concentration of 60 μM in a buffer containing 10 mM Na_2HPO_4 pH 6.0, 50 mM NaCl, 1 mM DTT and 0.02% w/v NaN_3 and 10% v/v D_2O.

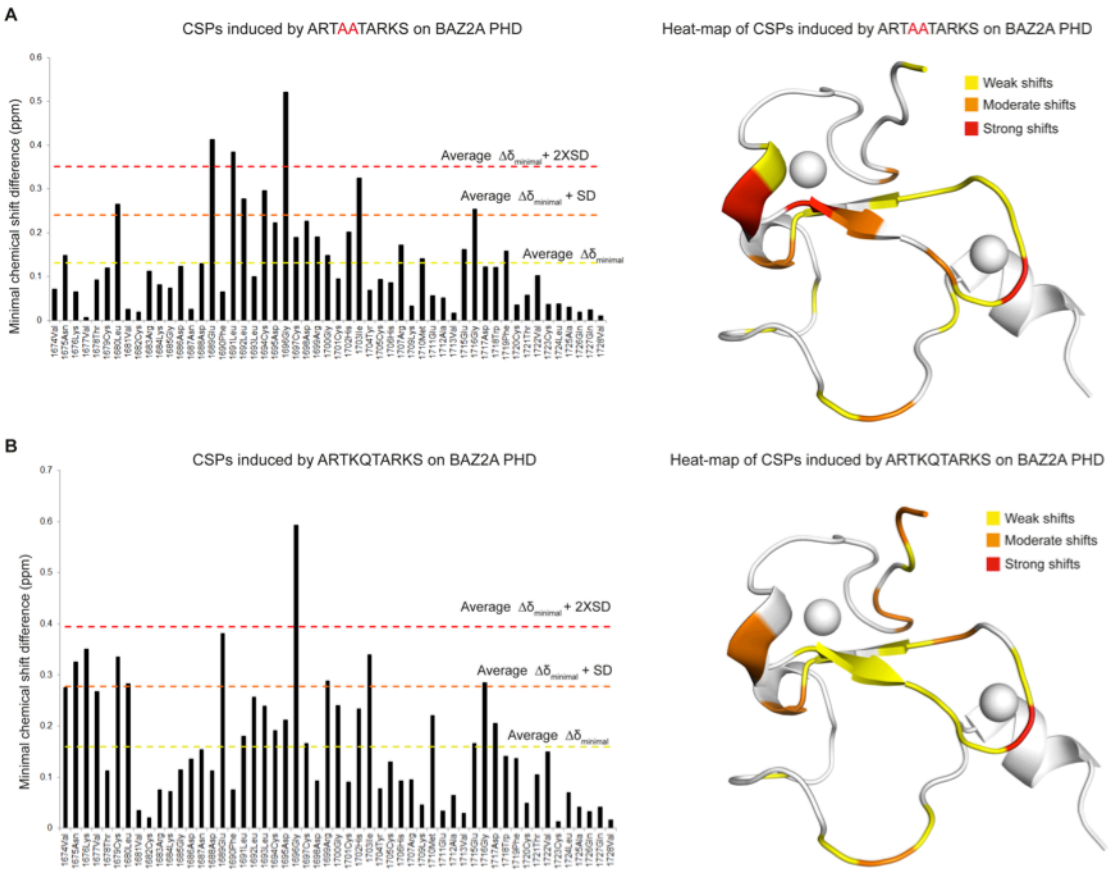


Figure S8. CSPs induced by H3 10-mer wild-type (ARTKQTARKS) and double-Ala mutant (ARTAATARKS) peptides on BAZ2A PHD. Consistent with a tighter binding affinity, most the shifts induced by H3 10-mer double-Ala mutant peptide (ARTAATRAKS) are in the slow exchange regime on the NMR timescale and peaks shifts are difficult to follow. To allow unbiased analysis of the double-Ala mutant peptide CSP data we used the "minimal-shift approach" (material and methods). The minimal shifts found were plotted against BAZ2A PHD sequences, clustered into weak, moderate and strong shifts (as described in material and method) and used to generate a heat-map representative of the peptide binding site **(A)**. To allow direct comparison the same approached was used to analyze the CSPs induced by H3 10-mer wild-type (ARTKQTARKS) **(B)**.

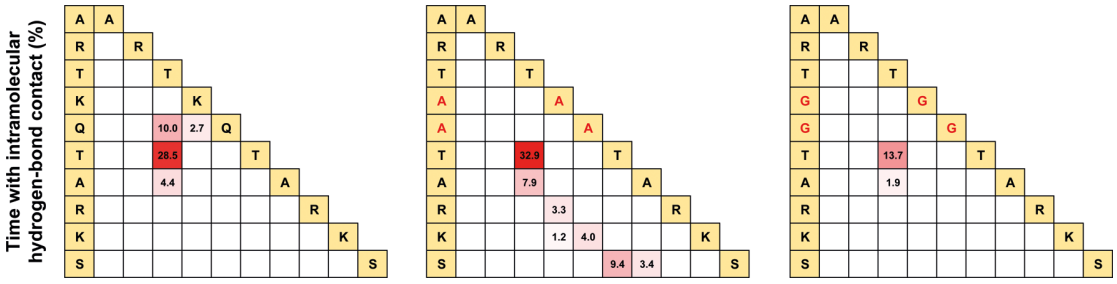


Figure S9. Intramolecular hydrogen-bond contacts within the 10-mer wild-type (ARTKQTARKS), double-Ala (ARTAATARKS) and double-Gly (ARTGGTARKS) mutant peptides in complex with BAZ2A PHD occurring during the last 60 ns of MD simulations, reported as the median percentage of time with contact out of 4 replicas. Median intramolecular hydrogen bond times below 1% have been omitted for clarity. In 10-merAA, a clear "i to i+4" pattern characteristic of α-helices is found.

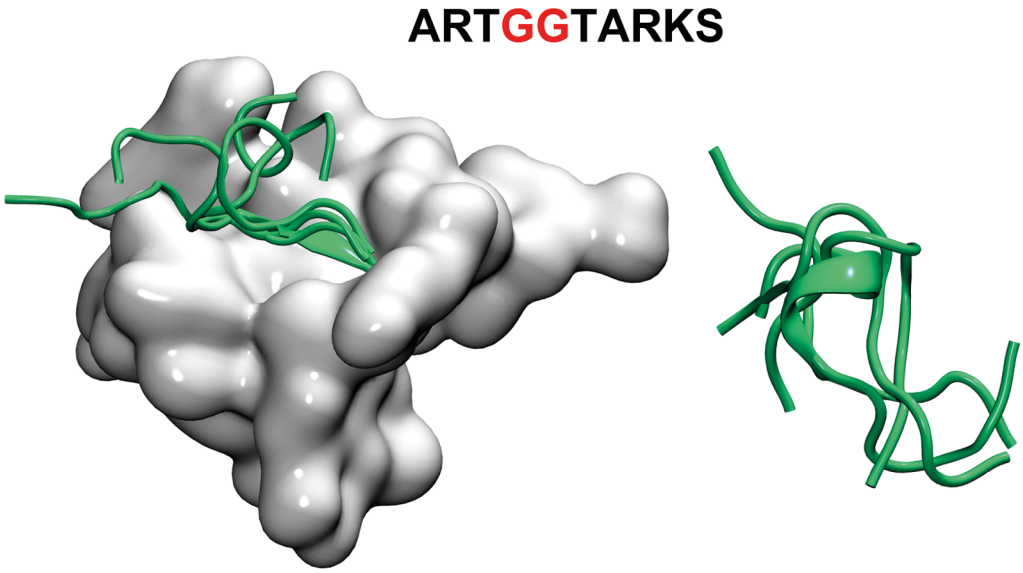
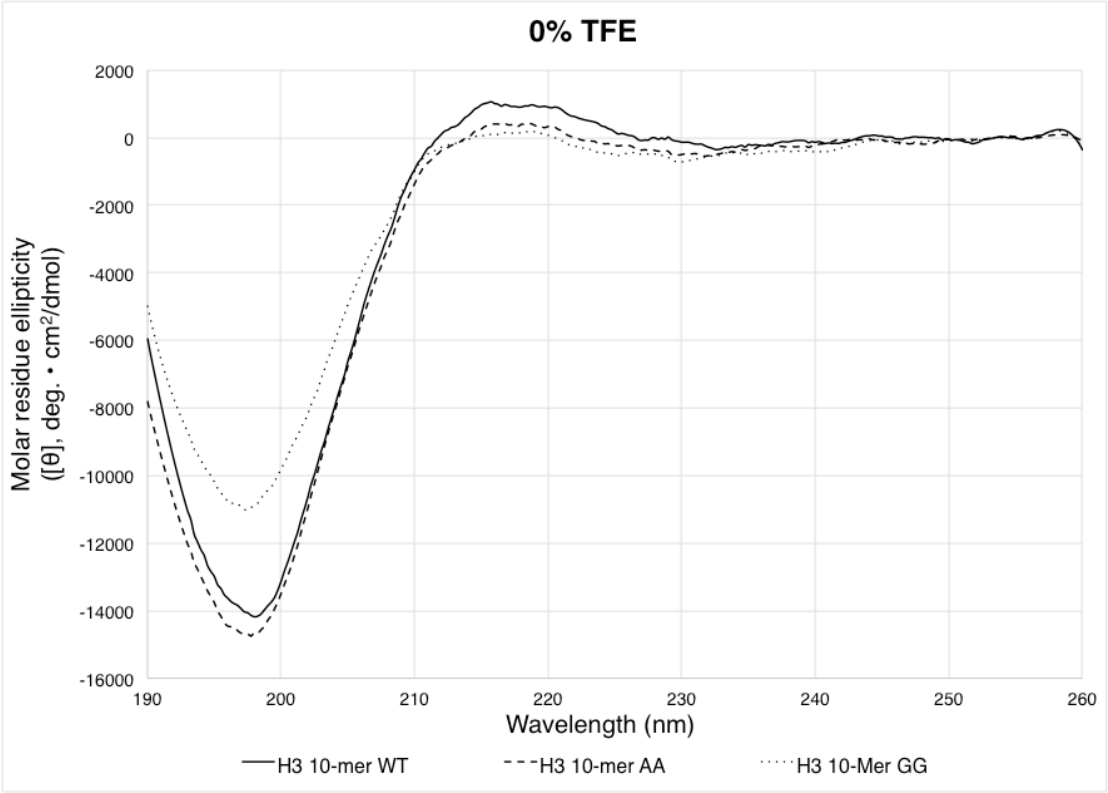
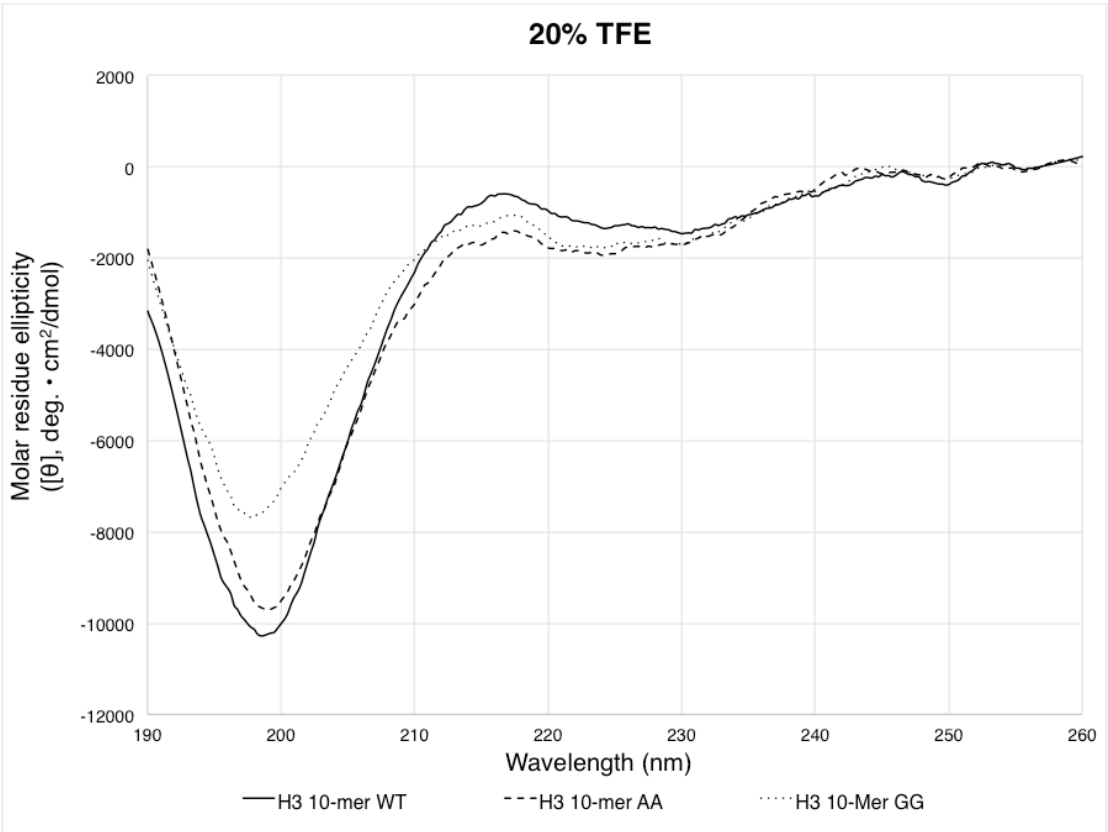
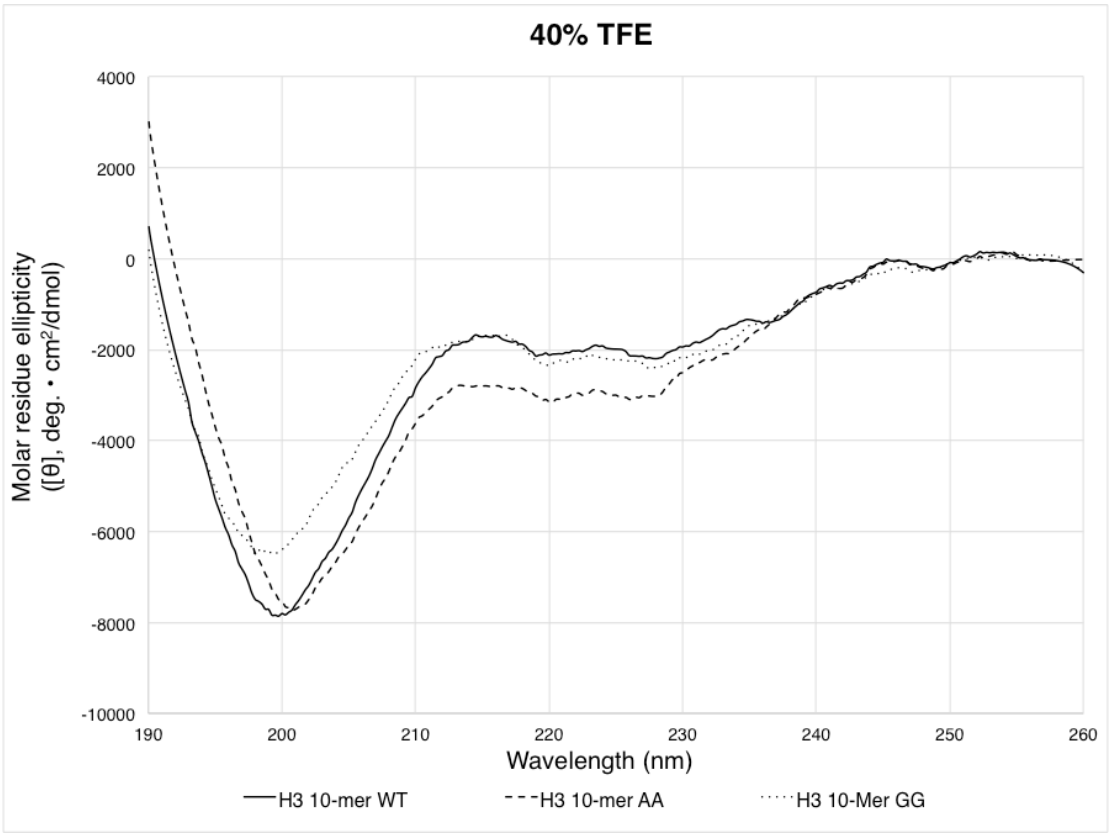
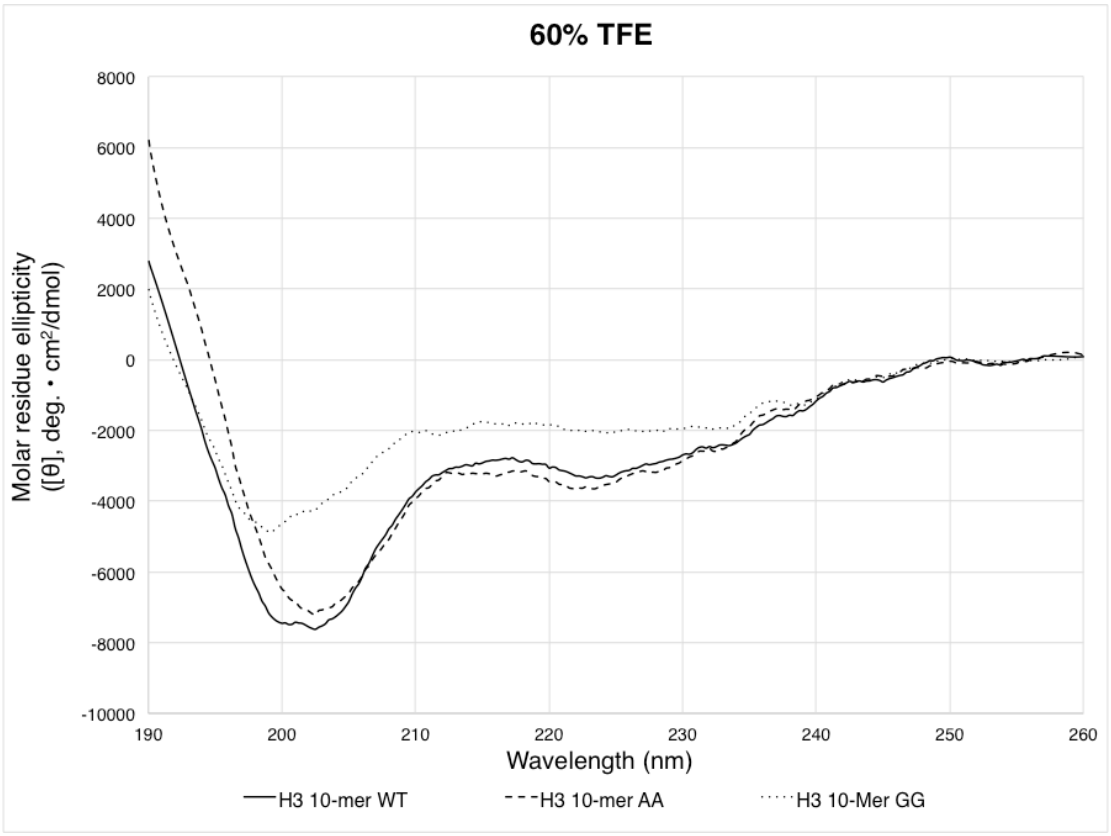


Figure S10. Superposed cartoon representation of the last frame of four MD replicas of ARTGGTARKS in complex with BAZ2A PHD (left) and in aqueous solution (right).


0% TFE
2000
0
-2000
-4000
-6000
-8000
-10000
-12000
-14000
-16000
190
200
210
220
230
240
250
260
Molar residue ellipticity ([θ], deg. • cm²/dmol)
Wavelength (nm)
H3 10-mer WT
H3 10-mer AA
H3 10-Mer GG


20% TFE
2000
0
-2000
-4000
-6000
-8000
-10000
-12000
190
200
210
220
230
240
250
260
Molar residue ellipticity ([θ], deg. • cm²/dmol)
Wavelength (nm)
H3 10-mer WT
H3 10-mer AA
H3 10-Mer GG


40% TFE
Molar residue ellipticity ([θ], deg. • cm²/dmol)
4000
2000
0
-2000
-4000
-6000
-8000
-10000
190
200
210
220
230
240
250
260
Wavelength (nm)
H3 10-mer WT
H3 10-mer AA
H3 10-Mer GG


60% TFE
Molar residue ellipticity ([θ], deg. • cm²/dmol)
8000
6000
4000
2000
0
-2000
-4000
-6000
-8000
-10000
190
200
210
220
230
240
250
260
Wavelength (nm)
H3 10-mer WT
H3 10-mer AA
H3 10-Mer GG

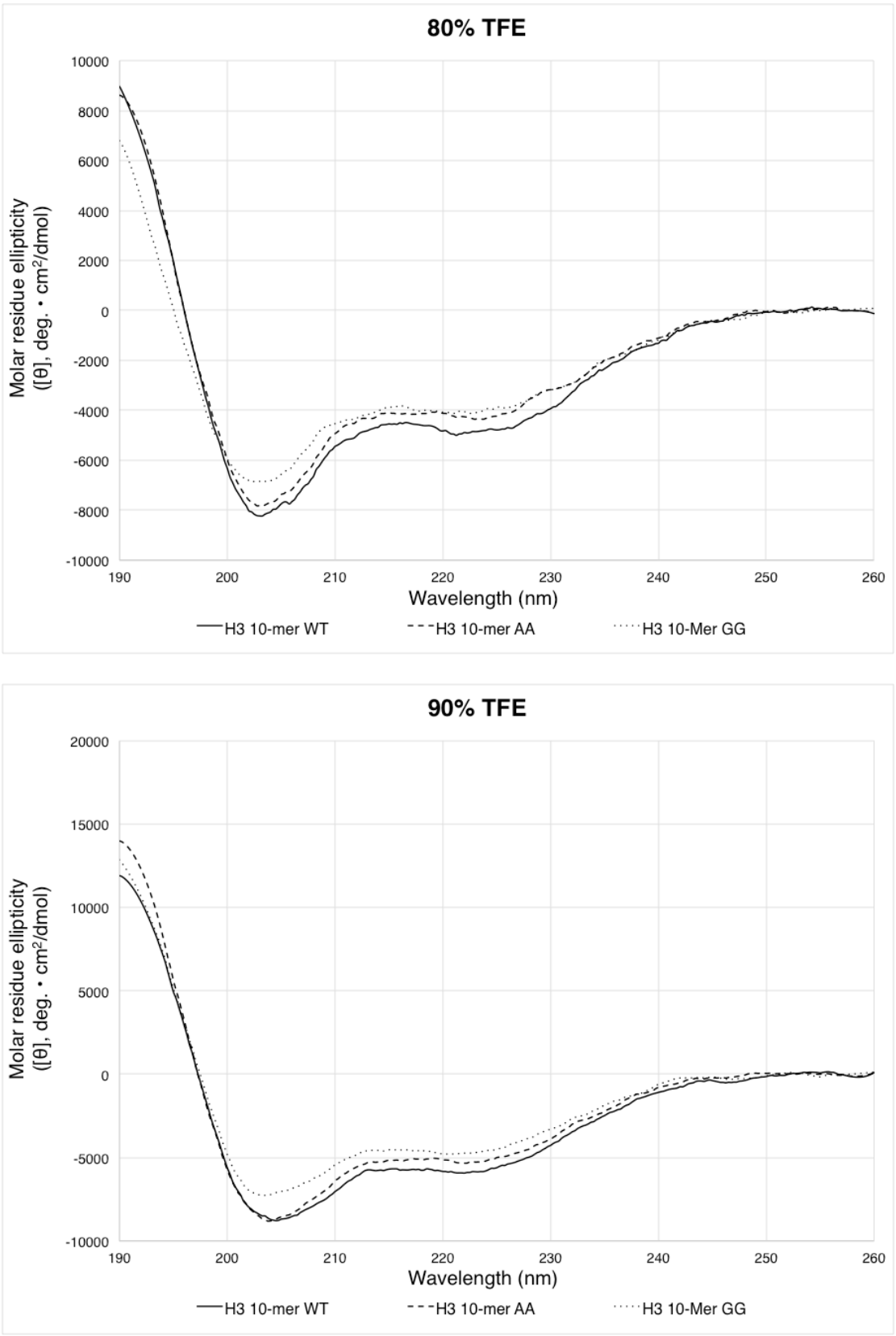


Figure S11. CD spectra recorded on samples of the H3 10-mer WT (ARTKQTARKS), H3 10-mer AA (ARTAATARKS) and H3 10-mer GG (ARTGGTARKS) peptides at different TFE concentrations (v/v).

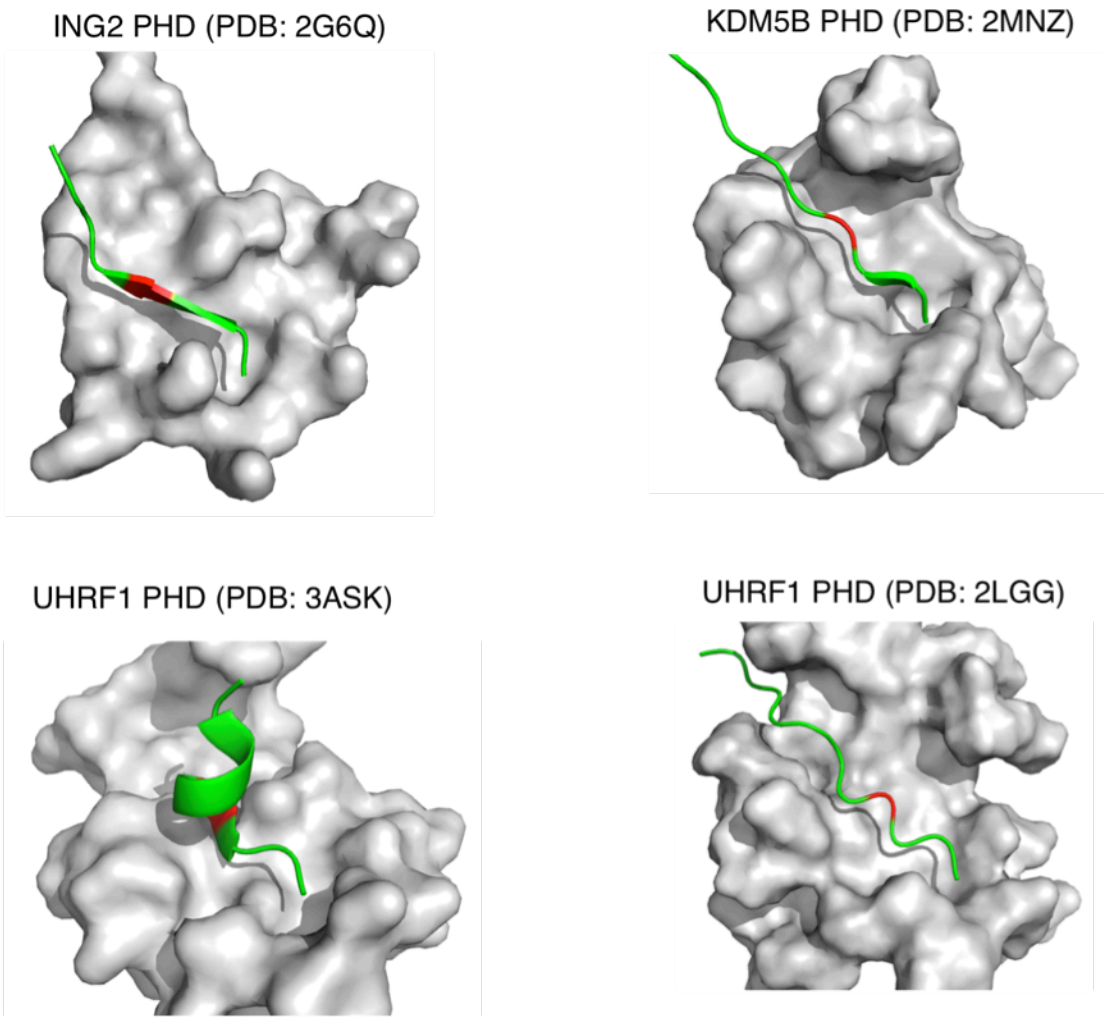
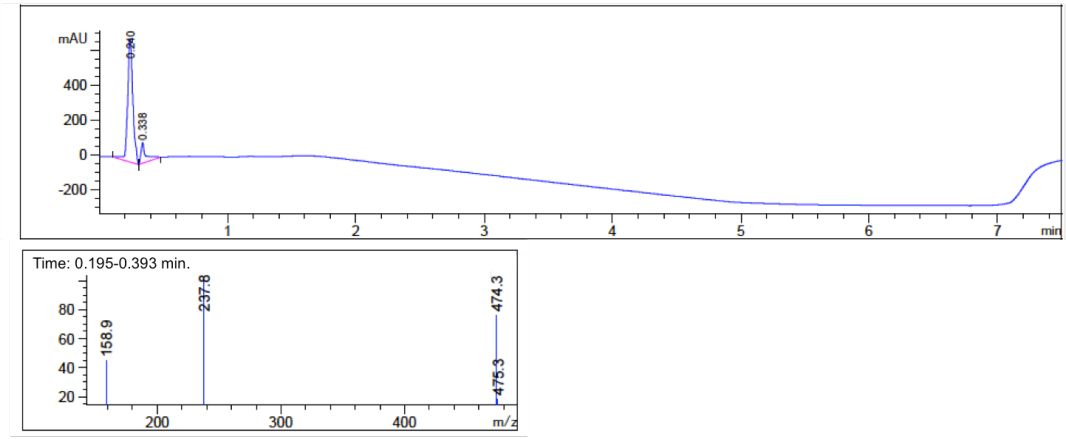
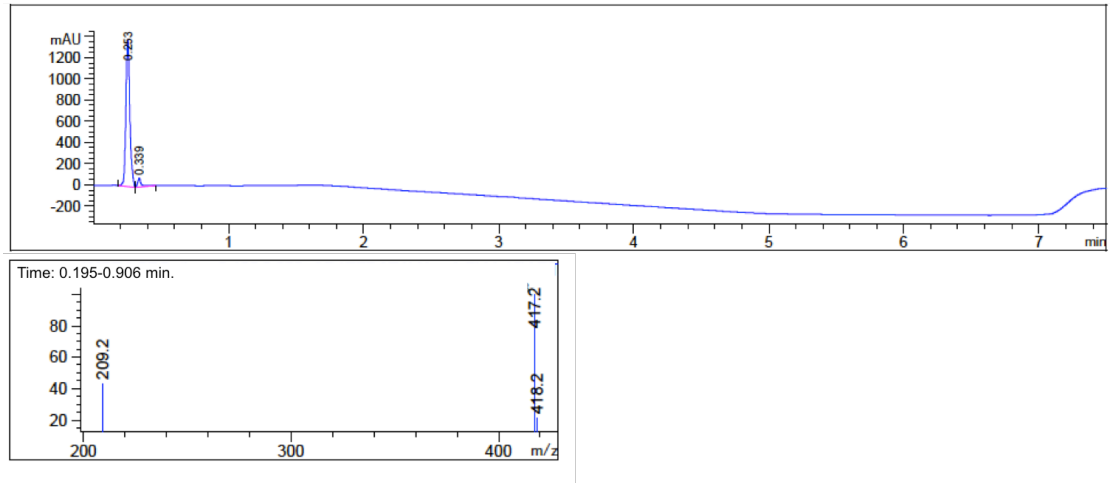


Figure S12. Structures of the PHD fingers of ING2, KDM5B and UHRF1 (shown in grey and surface representation) in complex with H3 N-terminal peptide (shown in green and cartoon representation). The K4 of H3 peptide is colored in red.

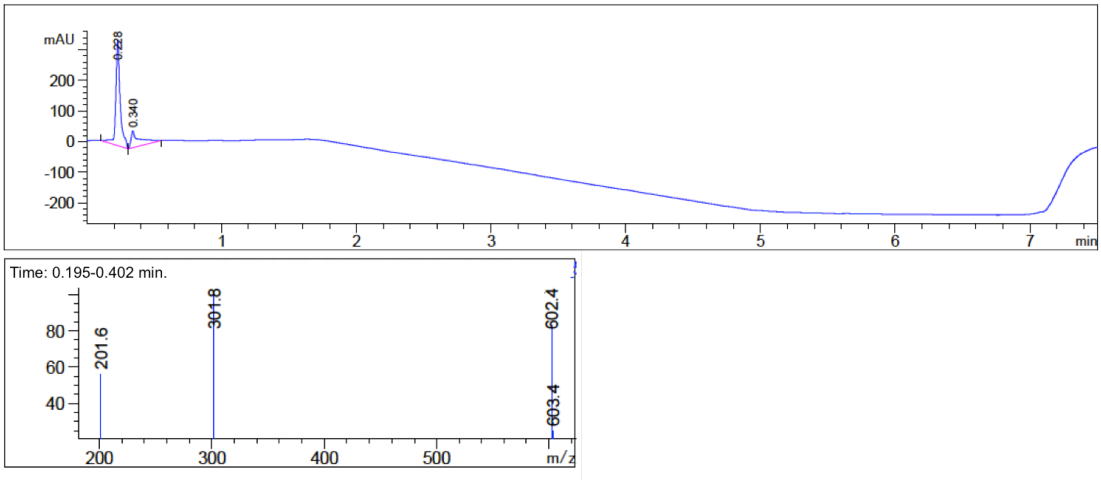
ARTK (MW: 475.6 Da)



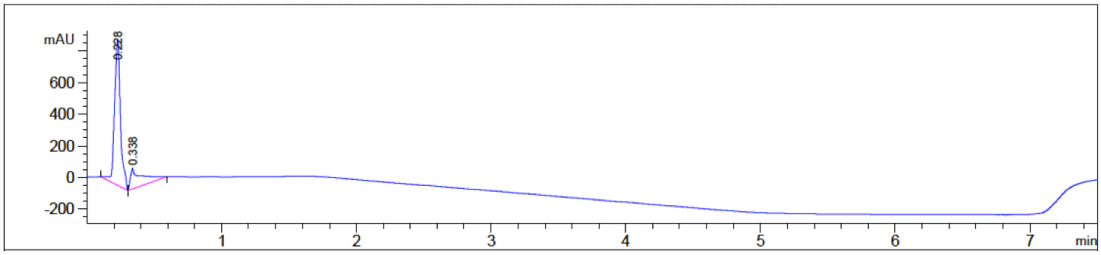
ARTA (MW: 417.5 Da)

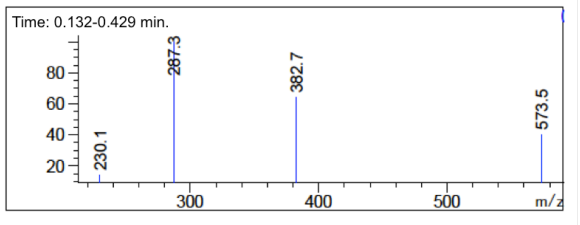


ARTKQ (MW: 602.7 Da)

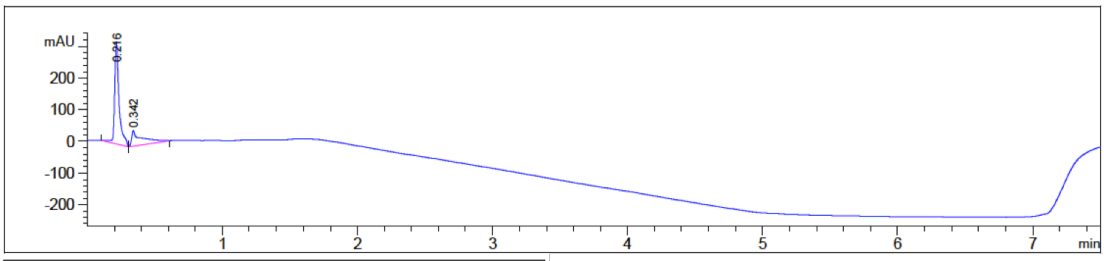


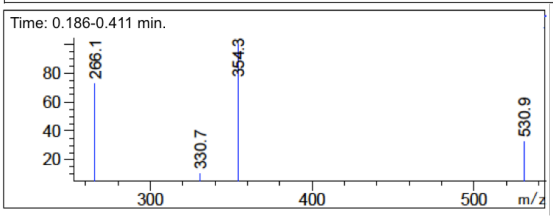
ARTKQTARKS (MW: 1146.3 Da)



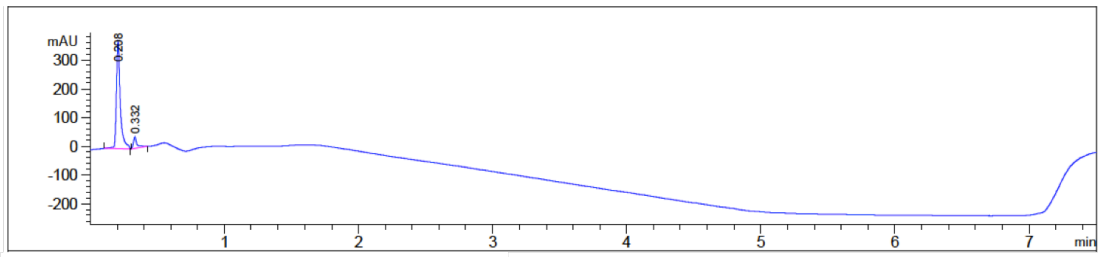


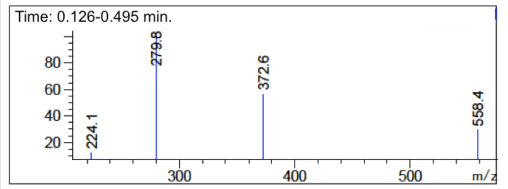
AATKQTARKS (MW: 1061.2 Da)



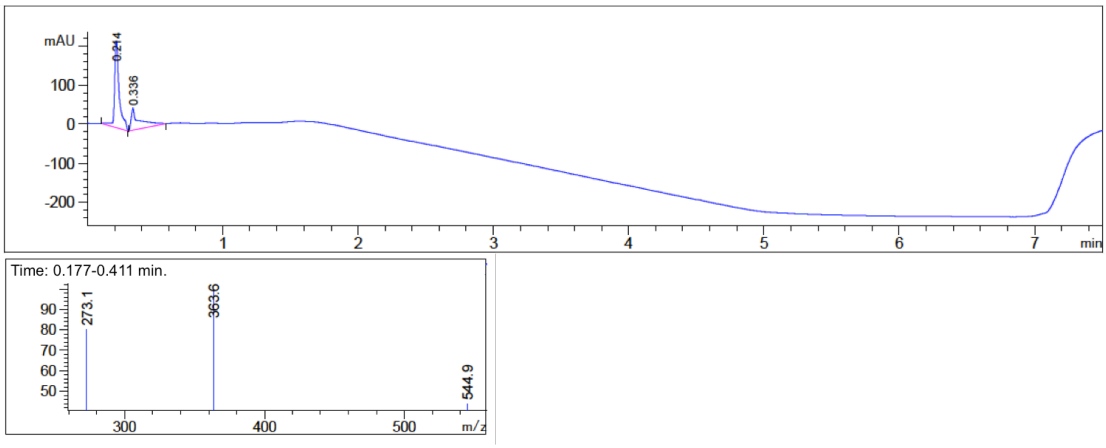


ARAKQTARKS (MW: 1116.3 Da)

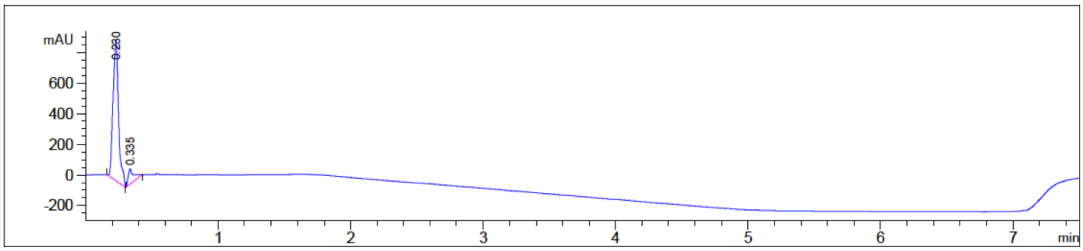


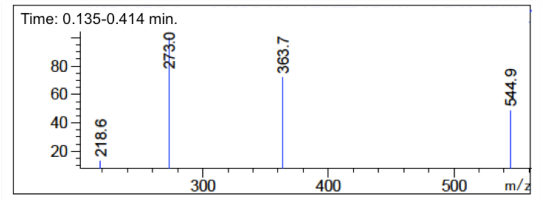


ARTAQTARKS (MW: 1089.2 Da)

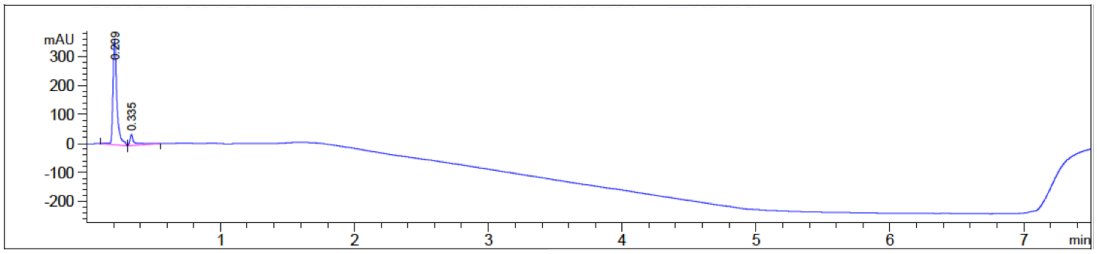


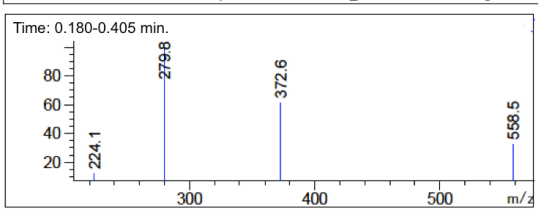
ARTKATARKS (MW: 1089.2 Da)





ARTKQAARKS (MW: 1116.2 Da)





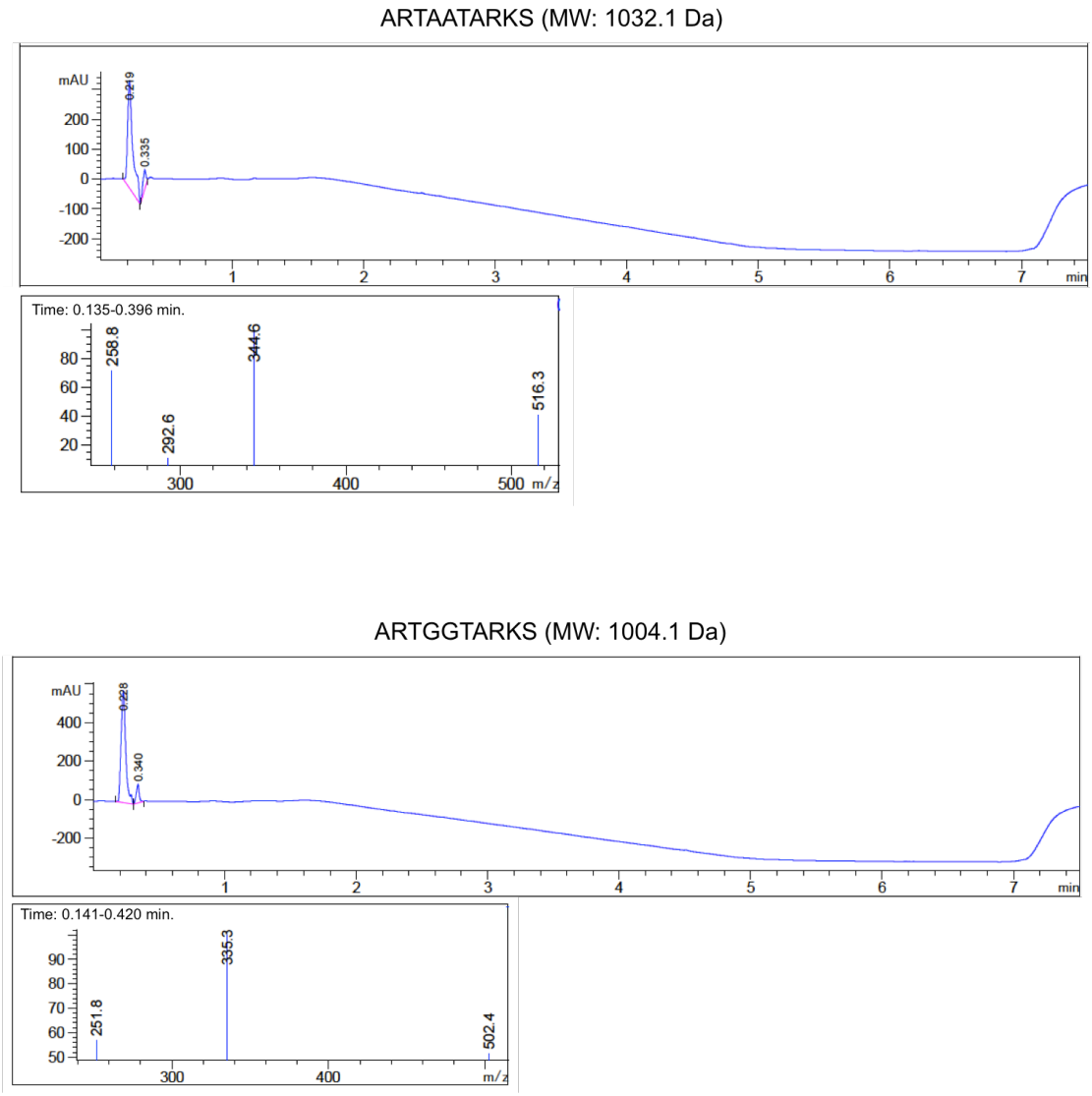


Figure S13. Liquid chromatography mass spectrometry (LCMS) analysis of H3 N-terminal peptides. For each peptide is reported the expected molecular weight (MW), the chromatogram monitoring absorption at 210 nm (upper panel) and the MS spectrum (lower panel).

H3 10-mer WT								
TFE % (v/v)	**Result**	**Helix1**	**Helix2**	**Strand1**	**Strand2**	**Turns**	**Unordered**	**Total**
0%	1	0	0.103	0.194	0.111	0.223	0.368	0.999
	2	0	0.061	0.226	0.128	0.214	0.372	1.001
20%	1	0	0.104	0.189	0.116	0.245	0.346	1
	2	0.004	0.067	0.236	0.128	0.22	0.344	0.999
40%	1	0	0.095	0.209	0.118	0.239	0.338	0.999
	2	0.007	0.068	0.247	0.128	0.218	0.332	1
60%	1	0.018	0.109	0.196	0.115	0.238	0.323	0.999
	2	0.026	0.085	0.231	0.123	0.222	0.314	1.001
80%	1	0.068	0.109	0.177	0.106	0.224	0.315	0.999
	2	0.058	0.094	0.189	0.111	0.231	0.317	1
90%	1	0.093	0.117	0.161	0.102	0.222	0.306	1.001
	2	0.08	0.106	0.169	0.105	0.229	0.312	1.001

Table S1. Estimation of the secondary structure content of the H3 10-mer WT (ARTKQTARKS) peptide from deconvolution of CD spectra acquired at different TFE concentrations. Helix1 is the content of regular α-helix; Helix2 of distorted α-helix; Strand1 of regular β-sheet and Strand2 of distorted β-sheet. Result 1 reports the secondary structure content found using the closest matching solution during deconvolution and Result 2 the average of all matching solutions.

H3 10-mer AA								
TFE % (V/V)	**Result**	**Helix1**	**Helix2**	**Strand1**	**Strand2**	**Turns**	**Unordered**	**Total**
0%	1	0	0.103	0.184	0.114	0.231	0.368	1
	2	0.002	0.062	0.202	0.123	0.221	0.39	1
20%	1	0	0.091	0.199	0.118	0.239	0.353	1
	2	0.005	0.064	0.229	0.125	0.221	0.356	1
40%	1	0.014	0.095	0.208	0.112	0.23	0.341	1
	2	0.02	0.074	0.233	0.12	0.219	0.335	1.001
60%	1	0.034	0.098	0.199	0.113	0.228	0.327	0.999
	2	0.035	0.085	0.237	0.122	0.214	0.308	1.001
80%	1	0.059	0.104	0.188	0.111	0.227	0.311	1
	2	0.048	0.091	0.204	0.118	0.231	0.309	1.001
90%	1	0.094	0.106	0.175	0.104	0.215	0.306	1
	2	0.069	0.097	0.182	0.112	0.224	0.314	0.998

Table S2. Estimation of the secondary structure content of the H3 10-mer AA (ARTAATARKS) peptide from deconvolution of CD spectra acquired at different TFE concentrations. Helix1 is the content of regular α-helix; Helix2 of distorted α-helix; Strand1 of regular β-sheet and Strand2 of distorted β-sheet. Result 1 reports the secondary structure content found using the closest matching solution during deconvolution and Result 2 the average of all matching solutions.

H3 10-mer GG								
TFE % (V/V)	**Result**	**Helix1**	**Helix2**	**Strand1**	**Strand2**	**Turns**	**Unordered**	**Total**
0%	1	0	0.092	0.204	0.119	0.227	0.357	0.999
	2	0.001	0.057	0.234	0.131	0.216	0.36	0.999
20%	1	0	0.089	0.208	0.121	0.235	0.347	1
	2	0.005	0.059	0.24	0.13	0.219	0.348	1.001
40%	1	0	0.093	0.215	0.119	0.237	0.337	1.001
	2	0.006	0.064	0.245	0.127	0.218	0.34	1
60%	1	0.004	0.081	0.223	0.126	0.236	0.329	0.999
	2	0.01	0.061	0.264	0.136	0.217	0.312	1
80%	1	0.043	0.094	0.199	0.114	0.227	0.322	0.999
	2	0.041	0.083	0.224	0.12	0.222	0.31	1
90%	1	0.074	0.093	0.192	0.11	0.216	0.316	1.001
	2	0.051	0.083	0.205	0.119	0.226	0.317	1.001

Table S3. Estimation of the secondary structure content of the H3 10-mer GG (ARTGGTARKS) peptide from deconvolution of CD spectra acquired at different TFE concentrations. Helix1 is the content of regular α-helix; Helix2 of distorted α-helix; Strand1 of regular β-sheet and Strand2 of distorted β-sheet. Result 1 reports the secondary structure content found using the closest matching solution during deconvolution and Result 2 the average of all matching solutions.

Primer Name	Sequence (from 5' to 3')
F_BAZ2A_PHD_E1689Q	CGCAAAGGCGATAATGATCAGTTTCTGCTGCTGTGTGAT
R_BAZ2A_PHD_E1689Q	ATCACACAGCAGCAGAAACTGATCATTATCGCCTTTGCG
F_BAZ2A_PHD_E1689K	CGCAAAGGCGATAATGATAAATTTCTGCTGCTGTGTGAT
R_BAZ2A_PHD_E1689K	ATCACACAGCAGCAGAAATTTATCATTATCGCCTTTGCG
F_BAZ2B_PHD_E1944Q	CGCAAAGGCGATAATCAGGAACTGCTGCTGCTGTGC
R_BAZ2B_PHD_E1944Q	GCACAGCAGCAGCAGTTCTGCATTATCGCCTTTGCG
F_BAZ2B_PHD_E1944K	CGCAAAGGCGATAATAAAGAACTGCTGCTGCTGTGC
R_BAZ2B_PHD_E1944K	GCACAGCAGCAGCAGTTCTTTATTATCGCCTTTGCG
F_BAZ2A_PHD_D1688N/E1689Q	CGCAAAGGCGATAATAATCAGTTTCTGCTGCTGTGTGAT
R_BAZ2A_PHD_D1688N/E1689Q	ATCACACAGCAGCAGAAACTGATTATTATCGCCTTTGCG

Supplementary Table S4. Primers used to perform site directed mutagenesis.

System	Nr. replica	Duration (ns)	Temperature (mean ± 1*σ, K)	Total energy (mean ± 1*σ, kcal/mol)	$RMSD_{BAZ2A}$ from crystal (mean ± 1*σ, Å)
BAZ2A 10-mer	1	80	299.5 ± 1.0	–54786 ± 151	1.4 ± 0.2
	2	80	299.5 ± 1.0	–54788 ± 245	1.4 ± 0.3
	3	80	299.5 ± 1.0	–54789 ± 151	1.3 ± 0.2
	4	80	299.5 ± 1.0	–54779 ± 151	1.9 ± 0.5
BAZ2A 10-mer AA	1	80	299.5 ± 1.0	–54523 ± 150	1.4 ± 0.2
	2	80	299.5 ± 1.0	–54521 ± 150	1.4 ± 0.3
	3	80	299.5 ± 1.0	–54525 ± 151	1.4 ± 0.2
	4	80	299.5 ± 1.0	–54523 ± 150	1.4 ± 0.2
BAZ2A 10-mer GG	1	80	299.5 ± 1.0	–54544 ± 150	1.5 ± 0.3
	2	80	299.5 ± 1.0	–54551 ± 150	1.4 ± 0.2
	3	80	299.5 ± 1.0	–54544 ± 150	1.5 ± 0.3
	4	80	299.5 ± 1.0	–54548 ± 150	2.2 ± 0.6
10-mer	1	80	299.7 ± 1.7	–21838 ± 94	-
	2	80	299.7 ± 1.7	–21835 ± 94	-
	3	80	299.7 ± 1.6	–21837 ± 94	-
	4	80	299.7 ± 1.7	–21837 ± 94	-
10-mer AA	1	80	299.7 ± 1.7	–20470 ± 91	-
	2	80	299.7 ± 1.7	–20473 ± 92	-
	3	80	299.7 ± 1.7	–20471 ± 91	-
	4	80	299.7 ± 1.7	–20471 ± 92	-
10-mer GG	1	80	299.7 ± 1.7	–20501 ± 92	-
	2	80	299.7 ± 1.7	–20501 ± 91	-
	3	80	299.7 ± 1.7	–20502 ± 91	-
	4	80	299.7 ± 1.7	–20502 ± 91	-

Supplementary Table S5. Convergence and stability data of MD simulations.