Electronic Supplementary Material (ESI) for Chemical Science.


# *Supporting Information*

# *Identification of novel functional mini-receptors by combinatorial screening of split-WW domains*

Hermann Neitz,[+,c] Niels Benjamin Paul,[+,d] Florian Häge, [a,e] Christina Lindner,[a,e] Roman Graebner, [a,e] Michael Kovermann*[,b], and Franziska Thomas*[,a,e]

a Institute of Organic Chemistry, Heidelberg University, Im Neuenheimer Feld 270, 69120 Heidelberg, Germany

b Department of Chemistry, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany

c Institute of Organic Chemistry, University of Würzburg, Am Hubland, 97074 Würzburg, Germany

d Institute of Organic and Biomolecular Chemistry, University of Göttingen, Tammannstr. 2, 37077 Göttingen, Germany

e Centre for Advanced Materials, Heidelberg University, Im Neuenheimer Feld 225, 69120 Heidelberg, Germany

+ These authors contributed equally.

*To whom correspondence should be addressed: franziska.thomas@oci.uni-heidelberg.de, michael.kovermann@uni-konstanz.de.

## Table of Content

# 1. Experimental

**General.** Fmoc-protected amino acids were purchased from *GL Biochem LTD* (Shanghai, China) and isotopically labelled Fmoc-protected amino acids were acquired from CortecNet Europe (Voisins-Le-Bretonneux, France). DIC and OxymaPure were obtained from *Iris Biotech GmbH* (Marktredwitz, Germany). The H-Rink Amide-ChemMatrix® resin was acquired from *Sigma Aldrich* (Taufkirchen, Germany). DMF used for peptide synthesis was supplied by *Fisher Scientific* (Schwerte, Germany) and was of peptide grade quality. Acetonitrile used for HPLC was supplied by *Fisher Scientific* (Schwerte, Germany) with HPLC grade quality. Water used for HPLC and reactions was obtained by purifying deionized water with the purification device *Arium Mini* from *Sartorius*. All other reagents were supplied by *Sigma Aldrich* (Taufkirchen, Germany), *Thermo Fisher Scientific* (Langenselbold, Germany), *VWR International* (Darmstadt, Germany) and *Carl Roth* (Karlsruhe, Germany). All reagents were of synthesis grade quality and were used as supplied. Unless otherwise stated, biophysical measurements were performed in phosphate buffered saline (PBS, 8.2 mM $Na_2HPO_4$, 1.8 mM $K_2HPO_4$, 137 mM NaCl, 2.7 mM KCl, pH 7.4). Peptide concentrations were determined by UV-absorbance at 280 nm ($\varepsilon_{280}$(Trp) = 5540 $mol^{-1}$ $cm^{-1}$, $\varepsilon_{280}$(Tyr) = 1480 $mol^{-1}$ $cm^{-1}$) or at 214 nm in the absence of Tyr and Trp [1,2]

**Peptide Synthesis.** Peptide amides were synthesized on a H-Rink Amide-ChemMatrix® resin (loading density: 0.40-0.60 mmol $g^{-1}$) and peptide acids on a HMBC-ChemMatrix® resin, which was preloaded with glycine ($B_{CC}$-$^{s}$hPin1$_{WW}$-L1) or histidine ($B_{CC}$-$^{s}$hPin1$_{WW}$-L2). General synthesis scale was 0.1 mmol, and automated synthesis was performed on a *Liberty Blue CEM* microwave-assisted peptide synthesizer. The synthesis was conducted *via* a standard Fmoc/*t*Bu-protocol using DIC and Oxyma Pure as coupling reagents and a solution of piperidine in DMF (1:4 (v/v)) for Fmoc-removal. Where required, *N*-acetylation of the peptides was carried out by equilibrating the peptide resin with 5 mL of acetic acid anhydride / pyridine (1:9 (v/v)) for 5 min at room temperature. Acidic cleavage from the resin was achieved by treatment with a mixture of trifluoroacetic acid (TFA) / triisopropylsilane / water (90:5:5 (v/v/v), 3 h). The resin was washed with additional TFA (5 mL), and the combined TFA fractions were concentrated to a third of the initial volume under a flow of nitrogen. The crude peptide was then precipitated from cold diethylether (40 mL) and isolated by centrifugation and decantation of the supernatant. The precipitate was washed twice with ice-cold diethylether and subsequently dissolved in 5 mL of a 1:1 (v/v) mixture of acetonitrile and water and then freeze-dried to give a fine white solid.

The CTD heptapeptide was synthesized by manual microwave-assisted solid-phase peptide synthesis using a *CEM Discover* microwave. The temperature was monitored by a fiber optic temperature probe. The synthesis was performed in 0.05 mmol scale on a H-Rink Amide-ChemMatrix® resin (loading density: 0.40-0.60 mmol $g^{-1}$), which was swollen in DMF for 1 h prior to synthesis. Fmoc-deprotection was performed at room temperature for 5 min and another 15 min using a solution of piperidine in DMF (1:4 (v/v)). After each deprotection step, the peptide resin was washed with DMF (5 x), $CH_2Cl_2$ (5 x) and DMF (5 x). Coupling of Fmoc-protected amino acid building blocks was performed with solutions at 0.2 M (1.0 eq) concentration using HBTU (0.45 M, 4.5 eq) as coupling reagent and DIPEA (2.0 M, 10 eq) as base additive. Whereas standard Fmoc amino acids were coupled for 5 min at 25 W and a maximum temperature of 75 °C, coupling time for Fmoc-(*O*)-(benzylphospho)-serine and all following couplings was extended to 15 min at 25 W and slightly reduced temperature of 72 °C. After each coupling step, the peptide resin was washed with DMF (5 x), $CH_2Cl_2$ (5 xs) and DMF (5 x). *N*-Acetylation was carried out by equilibrating the peptide resin with 5 mL of acetic acid anhydride / pyridine (1:9 (v/v)). The resin was washed in DMF (5 x),

$CH_2Cl_2$ (5 x), DMF ( 5 x) and another $CH_2Cl_2$ (10 x) and dried *in vacuo*. The cleavage from the resin and the global deprotection was performed as described above.

**Peptide purification.** Peptides were purified by semi-preparative reversed-phase HPLC using a *JASCO* chromatography system (pumps *PU-2080 Plus*, degasser *DG-2080-53*, detector *MD-2010 Plus*) at a flow rate of 3 mL·min$^{-1}$, and a *Nucleodur* 100-5-C18 ec, (250 mm by 10 mm, 5 µm) reversed-phase column from *Macherey-Nagel*, which was equilibrated to 50 °C. Linear gradients of water and acetonitrile (buffer A: water, 0.1 % TFA, buffer B: acetonitrile, 0.1 % TFA) were used for purification. Chromatograms were monitored at 220 nm and 280 nm wavelengths.

The following gradients were used for purification:

gradient 1: 10-20% B, 30 min (CTD peptide)

gradient 2: 10-50% B, 30 min (WW-1-8, WW-2-8, WW-2-10)

gradient 3: 10-60% B, 30 min (hPin1$_{WW}$(6-20); hPin1$_{WW}$(21-39))

gradient 4: 20-40% B, 30 min (hPin1$_{WW}$)

gradient 5: 20-50% B, 30 min ($B_{CC}$, $B_{CC}$-WW-2, $B_{CC}$-WW-3, $B_{CC}$-WW-4)

gradient 6: 20-70% B, 30 min ($A_{CC}$-WW-6 to $A_{CC}$-WW-10)

gradient 7: 25-45% B, 30 min ($B_{CC}$-hPin1$_{WW}$-L1; $B_{CC}$-hPin1$_{WW}$-L2)

gradient 8: 30-50% B, 30 min (hPin1$_{WW}$-L1-$A_{CC}$; hPin1$_{WW}$-L2-$A_{CC}$)

gradient 9: 30-60% B, 30 min ($A_{CC}$)

**Peptide characterization.** The peptides were characterized by mass spectrometry on a *Bruker Autoflex Speed* MALDI-TOF mass spectrometer operating in positive-ion reflector mode. (matrix: dihydroxybenzoic acid (DHB) saturated solution in acetonitrile/water (1:1 (v/v), external calibration). Analytical HPLC measurements were performed on a *VWR Hitachi Primaide* chromatography system and a *VDSPher PUR* 100 C18-SE (5 µm, 250 mm by 4.6 mm) reversed-phase column from *Optilab* at a flow rate of 1 mL·min$^{-1}$. For peptide characterization linear gradients of water and acetonitrile (buffer A: water, 0.1 % TFA, buffer B: acetonitrile, 0.1 % TFA) were used. Chromatograms were monitored at 220 nm and 280 nm wavelengths.

The following gradients were used for analysis:

gradient 1: 5-40% B, 30 min (CTD peptide)

gradient 2: 10-50% B, 30 min ($B_{CC}$, WW-1-8, WW-2-8, WW-2-10)

gradient 3: 10-60% B, 30 min (hPin1$_{WW}$(6-20), hPin1$_{WW}$(21-39))

gradient 4: 20-40% B, 30 min (hPin1$_{WW}$)

gradient 5: 20-50% B, 30 min ($B_{CC}$-WW-2 to $B_{CC}$-WW-4)

gradient 6: 20-70% B, 30 min ($A_{CC}$-WW-6 to $A_{CC}$-WW-10)

gradient 7: 20-80% B, 30 min ($A_{CC}$, $B_{CC}$-hPin1$_{WW}$-L1, $B_{CC}$-hPin1$_{WW}$-L2, hPin1$_{WW}$-L1-$A_{CC}$, hPin1$_{WW}$-L2-$A_{CC}$)

**Circular Dichroism Spectroscopy.** CD spectra and CD thermal denaturation profiles were recorded on a *JASCO J-1500* CD spectrometer, which was equipped with a *JASCO PTC510* temperature measuring unit. CD spectra were recorded in a range of 190-260 nm at 40 µM peptide concentration in PBS at 20 °C in

1 mm quartz cuvettes from *Starna* at 100 nm · min$^{-1}$ scanning speed. The spectra were averaged from 5 individual measurements. All spectra were corrected by blank CD signals and the mean residue ellipticity MRE (deg·cm$^2$·dmol$^{-1}$·res$^{-1}$) was calculated according to equation 1:

$$MRE = \frac{100 \cdot \theta}{l \cdot c \cdot n} \quad (1)$$

with $\theta$ being the ellipticity, $l$ the optical path length, $c$ the peptide concentration (mmol · L$^{-1}$) and $n$ (res) the number of residues.

CD thermal denaturation experiments were performed by heating from 5 to 95 °C at a heat rate of 60 °C · h$^{-1}$. The CD thermal denaturation profiles were recorded at 208 nm (coiled-coil peptides) or 227 nm (hPin1$_{WW}$) at 1 °C intervals (1 nm bandwidth, 16 s response time). All thermal denaturation profiles were depicted as change of fraction folded. Therefore, the change of ellipticity was transformed to the fraction folded $\alpha$ using equation 2:

$$\alpha = \frac{\theta - (\theta_u + m_u T)}{\theta_f + m_f T - (\theta_u + m_u T)} \quad (2)$$

with $\theta$ being the ellipticity, $\theta_u$ the ellipticity of the unfolded peptide, $\theta_f$ the ellipticity of the fully folded peptide, $m_u$ the slope of the post-transition, $m_f$ the slope of the pre-transition and $T$ the temperature in Kelvin.

**Analysis of CD data.** CD spectra were analyzed for the fraction of secondary structure elements on the DichroWeb Webserver using the CDSSTR algorithm with peptide set 4 (range 190-240 nm).[3,4]

To determine the midpoints of the thermal denaturation curves ($T_m$), two-state transition models were assumed.[5] The thermal denaturation curves of hPin1$_{WW}$ in the absence and presence of CTD ligand were fitted to a model for the unfolding of a monomer using the following equations:

$$\Delta G_{Di} = \Delta H_m \left(1 - \frac{T}{T_m}\right) - \Delta C_{p,m} \left((T_m - T) + T ln \left(\frac{T}{T_m}\right)\right) \quad (3)$$

$$K = \exp\left(-\frac{\Delta G_{Di}}{RT}\right) \quad (4)$$

$$\alpha = \frac{K}{(1 + K)} \quad (5)$$

$$\theta = \alpha(\theta_f - \theta_u) + \theta_u \quad (6)$$

All other thermal denaturation curves were fitted to a model of a two-state transition between a folded dimer and unfolded monomer using the following equations and assuming that $\Delta C_{p,m}$ = 0:[5]

$$\Delta G_{Di} = \Delta H_m \left(1 - \frac{T}{T_m}\right) + \ln(c) \quad (7)$$

$$K = \exp\left(-\frac{\Delta G_{Di}}{RT}\right) \quad (4)$$

$$\alpha = \frac{(4cK + 1) - \sqrt{(8cK + 1)}}{(4cK)} \quad (8)$$

$$\theta = \alpha(\theta_f - \theta_u) + \theta_u \quad (6)$$

Determination of $\Delta H_m$ and $\Delta C_{p,m}$ of the two-state transitions of the dimers to the respective monomers were performed using a model of a two-state transition between a folded dimer and unfolded monomers:[5]

$$\Delta G_{Di} = \Delta H_m + \Delta C_{p,m}(T - T_m) - T\left(\frac{\Delta H_m}{T_m} + \left(\Delta C_{p,m} ln\left(\frac{T}{T_m}\right)\right)\right) \quad (9)$$

$$K = \exp\left(-\frac{\Delta G_{Di}}{RT}\right) \quad (4)$$

$$\alpha = \frac{(8Kc^2 + c) - \sqrt{((-8Kc^2 - c)^2 - 4(4Kc^2)^2}}{(8Kc^2)} \quad (10)$$

$$\theta = \alpha(\theta_f - \theta_u) + \theta_u \quad (6)$$

with $\Delta G_{Di}$ being the free enthalpy of folding, $\Delta H_m$ the enthalpy of folding, $\Delta C_{p,m}$ the change of the heat capacity upon folding, $K$ the equilibrium constant of folding and $c$ the peptide concentration. $T_m$ is the concentration independent melting temperature at $K = 1$. $T$ is the temperature in Kelvin.

**NMR spectroscopy.** All one-dimensional proton NMR spectra acquired in this study have been measured in 0.1 mM peptide concentration in PBS, pH = 7.4 using a *Bruker Avance III* 600 MHz spectrometer equipped with a CP-TCI probehead operating at $T$ = 283 K.

Two-dimensional heteronuclear $^{1}H$-$^{15}N$ HSQC NMR spectra of selectively $^{15}N$ labeled peptide samples and of hPin1$_{WW}$ (natural abundance of $^{15}N$ nuclei) have been measured at a concentration of 1 mM in PBS buffer at pH = 7.4 using a *Bruker Avance Neo* 800 MHz spectrometer equipped with a CP-TCI probehead operating at $T$ = 283 K.

NMR diffusion spectra have been acquired for hPin1$_{WW}$, CC-hPin1$_{WW}$-L1, B$_{CC}$-hPin1$_{WW}$-L1 and hPin1$_{WW}$-L1-A$_{CC}$ at the proton dimension by using a pulsed field bipolar gradient stimulated echo experiments at $T$ = 298 K. For each diffusion profile, 21 different gradient strengths $G$ were used for 6 ms along the z axis followed by a 100 ms recovery delay. The diffusion was allowed to proceed for 80 ms. The calibration of $G$ was performed by a standard protocol. For error estimation, four different gradient strengths were repeated (relative gradient strengths of 1, 10, 40, and 70 %). The measured proton NMR spectra were integrated using chemical shifts ranging between 0 and 4 ppm and the following equation has been applied for data fitting:

$$I(G) = I_0 \cdot e^{(-G^2\gamma^2\delta^2 2D(\Delta - \frac{\delta}{3}))} \quad (11)$$

where $\gamma$ is the gyromagnetic ratio, $\delta$ is the gradient length, $\Delta$ is the diffusion time and $D$ is the calculated diffusion coefficient.[6]

**NMR titration experiments.** Unlabeled heptapeptide YSPTpSPS has been stepwise added to selectively $^{15}$N labeled CC-hPin1$_{WW}$-L1. A series of two-dimensional heteronuclear $^{1}$H-$^{15}$N HSQC NMR spectra has been then acquired (Bruker Avance Neo 800 MHz, CP-TCI probehead, $T$ = 283 K) to monitor the potential structural impact CTD peptide has on $^{s}$hPin1$_{WW}$-L1. Changes in chemical shifts, $\Delta\omega$, have been computed according to:

$$\Delta\omega = ((\Delta^1 H)^2 + (1/25)((\Delta^{15} N)^2))^{0.5} \quad (12)$$

where $\Delta^1 H$ is the change in proton and $\Delta^{15} N$ is the change in nitrogen dimension.

The dissociation constant, $K_D$, characterizing the interaction between CTD peptide and CC-hPin1$_{WW}$-L1 has been determined by using:

$$\Delta\omega^{obs} = \Delta\omega^{max}\left(\frac{n[P]_t + [L]_t + K_D - ((n[P]_t + [L]_t + K_D)^2 - 4n[P]_t[L]_t)^{0.5}}{2n[P]_t}\right) \quad (13)$$

where $\Delta\omega^{obs}$ represents the change in chemical shift per point of titration, $\Delta\omega^{max}$ the maximum of the change in chemical shift, $n$ the stoichiometry of binding, $[P]_t$ the entire concentration of CC-hPin1$_{WW}$-L1 and $[L]_t$ the entire concentration of CTD peptide. Two-dimensional $^{1}$H-$^{15}$N HSQC NMR spectra have been acquired using following concentrations, $c$, of CC-hPin1$_{WW}$-L1 and stoichiometric ratios, $n$, regarding CTD peptide: $c$ = 50 $\mu$M, $n$ = 0; $c$ = 42 $\mu$M, $n$ = 0.67; $c$ = 38 $\mu$M, $n$ = 2; $c$ = 32 $\mu$M, $n$ = 5.

**Fluorescence spectroscopy.** Ligand binding was studied by recording the changes in intrinsic tryptophan fluorescence in hPin1$_{WW}$ and CC-hPin1$_{WW}$-L1 and their derivatives with increasing concentrations of the respective ligands using a *BMG Labtech* microplate reader.

*Binding studies hPin1$_{WW}$ and CC-hPin1$_{WW}$.* Binding studies were performed in a 96 well-plate format using black non-binding well plates with F-bottom from *Greiner Bio-One*. WW domain-CTD ligand mixtures were prepared directly in the plate with a constant WW-domain concentration of 2 µM and varying ligand concentrations from 0.5-500 µM and with a total volume of 100 µL. The stock solutions used for sample preparation were concentrated tenfold. All stock solutions were prepared in protein low binding tubes, and protein low binding tips were used for pipetting. The ligand stock solutions were prepared as a 1:1 dilution series, starting with a concentration of 5 mM. The peptide mixtures were left to equilibrate at 4 °C overnight. Intrinsic tryptophan fluorescence was recorded at 350 nm using the fluorescence intensity endpoint mode and an excitation filter of $\lambda_{ex}$ = 295 nm with 20 flashes per well and top-to-bottom reading. The focal height was 4.4 mm and the gain was 1800. Each concentration was measured as sextuplicate. Saturation binding curves were plotted from the change of the background-corrected fluorescence signals. To determine the $K_d$ values, the obtained saturation binding curves were fitted to a one-site binding model using the following equation:

$$\Delta F_{350} = \frac{\Delta F_{350}^{max} \cdot c_{Ligand}}{K_d + c_{Ligand}} \quad (14)$$

with $\Delta F_{350}$ being the change in intrinsinsic fluorescence, $\Delta F_{350}{}^{max}$ the maximum fluorescence change at saturation, $c_{Ligand}$ the concentration of the ligand (here CTD) and $K_d$ the dissociation constant. To enable

background correction over the entire range of CTD-ligand titration, a ligand-to-buffer titration was performed as sextuplicate and subtracted from the sample plates.

*Library screening.* Library screening was performed in a 96 well-plate format using black non-binding well plates with F-bottom from *Greiner Bio-One*. WW domain-ligand mixtures (ligands: ATP, cAMP, phophocholine, inositol hexaphosphate) were prepared directly in the plate with a constant WW-domain concentration of 1 µM and varying ligand concentrations from 0.5-2 µM and with a total volume of 100 µL to estimate differences in the dissociation binding constants of a split-WW-domain variant and the respective ligands. The stock solutions used for sample preparation were concentrated tenfold. All stock solutions were prepared in protein low binding tubes and protein low binding tips were used for pipetting. The peptide mixtures were left to equilibrate at 20 °C for one hour. Intrinsic tryptophan fluorescence spectra were recorded from 335 - 400 nm with 20 flashes per well and top-to-bottom reading. Excitation was performed at $\lambda_{ex}$ = 295 nm using an excitation filter (295 nm). The focal height was 4.1 mm and the gain 2200. Each concentration was measured as at least triplicate. Saturation binding curves were plotted from the change of the background-corrected fluorescence signals at fitted to a one-site binding model using Equation 14. To enable background correction over the entire range of the ligand titration, a ligand-to-buffer titration was performed. In case of IP6 no change in background fluorescence was observed.

*Binding studies WW-1-8 and WW-2-10.* Binding studies were performed in a 96 well-plate format using black non-binding well plates with F-bottom from *Greiner Bio-One*. WW domain-ligand mixtures were prepared directly in the plate with a constant WW-domain concentration of 2 µM and varying ligand concentrations from 0.5-500 µM and with a total volume of 100 µL. The stock solutions used for sample preparation were concentrated tenfold. All stock solutions were prepared in protein low binding tubes, and protein low binding tips were used for pipetting. The ligand stock solutions were prepared as an 1:1 dilution series, starting with a concentration of 5 mM. The peptide mixtures were left to equilibrate at 4 °C overnight. Intrinsic tryptophan fluorescence was recorded at 350 nm using the fluorescence intensity endpoint mode and an excitation filter of $\lambda_{ex}$ = 295 nm with 20 flashes per well and top-to-bottom reading. The focal height was 4.4 mm and the gain was 1800. Each concentration was measured as octuplicate. Saturation binding curves were plotted from the change of the background-corrected fluorescence signals and fitted to a one-site binding model using Equation 14. To enable background correction over the entire range of ligand titration, a ligand-to-buffer titration was performed as octuplicate and subtracted from the sample plates.

**Computational modelling.** To model the binding of the phosphate ligands to WW-2-10, a computational model of WW-2-10 was produced based on a $hPin1_{WW}$ NMR structure (pdb 1i6c), in which the respective amino acid residues were mutated and which was relaxed three times using the Rosetta Relax application.[7,8] *AutoDockVina* was used for molecular docking calculations.[9] Prior to computational docking, the .pdb file of WW-2-10, which included all hydrogens, no water molecules and no other ligands, as obtained by the Rosetta Relax application, was used to prepare the receptor .pdbqt file in *AutoDockTools* (ADT) 1.5.7. The entire receptor was assumed to be rigid for the following molecular docking. The ligands were drawn in ChemDraw Professional 20, opened in Chem3D 20 and energy minimized using the MM2 force field. The 3D structure of the respective ligands was saved as pdb file and prepared as ligand in ADT 1.5.7 with rotatable bonds being defined. Molecular docking was performed covering the entire WW domain with an exhaustiveness of 200 and five repetitions to obtain five different seeds. Ligand-WW-domain complexes were visualized and figures created using PyMol 2.4.1. The ligands used for the docking studies were ATP (quadruple negatively charged), cAMP (ionized), phosphorylcholine (negatively charged), and IP6 (sixfold negatively charged).

## 2. Additional Tables and Figures

**Table S1.** Overview of peptide sequences containing calculated and experimentally determined molar masses and retention times (amino acids highlighted in red are isotopically labelled; G1-4: Gradients used in this study).

| | **Sequence** | **$[M+H]^+$ (calc.)** | **$[M+H]^+$ (found)** | **$t_r$ (min)** |
|---|---|---|---|---|
| $B_{CC}$ | Ac-GAQLKKKLQANKKELAQLKKKLQALKKKLAQG-$NH_2$ | 3613.3 | 3613.2 | 26.1 (G2) |
| $A_{CC}$ | Ac-GAQLEKELQALEKKLAQLEWENQALEKELAQG-$NH_2$ | 3676.9 | 3677.3 | 22.9 (G7) |
| $hPin1_{WW}$ | H-KLPPGWEKRMSRSSGRVYYFNHITNASQWERPSG-OH | 4021.9 | 4022.3 | 13.4 (G4) |
| CTD | Ac-YSPTpSPS-$NH_2$ | 857.3[a] | 857.4[a] | 16.8 (G1) |
| $B_{CC}$-$hPin1_{WW}$-L1 | Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSRSSG-OH | 5197.1 | 5197.9 | 16.5 (G7) |
| $hPin1_{WW}$-L1-$A_{CC}$ | H-RVYYFNHITNASQWERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5769.0 | 5769.3 | 20.2 (G7) |
| $B_{CC}$-$hPin1_{WW}$-R-$^{13}C_6$, $^{15}N_4$ | Ac-GAQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMS**R**SSG-OH | 5264.1 | 5264.7 | 16.6 (G7) |
| $hPin1_{WW}$-L1-$A_{CC}$-F-$^{15}N$ | H-RVYY**F**NHITNASQWERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5770.0 | 5770.3 | 20.3 (G7) |
| $B_{CC}$-$hPin1_{WW}$-L2 | Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSRSSGRVYYFNH-OH | 6176.6 | 6177.9 | 16.5 (G7) |
| $hPin1_{WW}$-L2-$A_{CC}$ | H-ITNASQWERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 4789.5 | 4789.9 | 20.5 (G7) |
| $hPin1_{WW}$(6-20) | H-KLPPGWEKRMSRSSG-OH | 1715.9 | 1716.0 | 22.3 (G3) |
| $hPin1_{WW}$(21-39) | H-RVYYFNHITNASQWERPSG-OH | 2325.1 | 2325.4 | 24.8 (G3) |

a) mass spectrometry of the CTD peptide was performed in linear negative mode.

**Table S2.** Overview of peptide sequences of the peptide library containing calculated and experimentally determined molar masses and retention times (Peptides WW-2, WW-3, WW-4 are derived from $B_{CC}$-$^{s}$hPin1$_{WW}$-L1 ; peptides WW-6, WW-7, WW-8, WW-9, and WW-10 are derived from $^{s}$hPin1$_{WW}$-L1-$A_{CC}$. WW-2-10 is the ATP-binding WW domain identified by screening of peptide combinations from $B_{CC}$-$^{s}$hPin1$_{WW}$-L1 and $^{s}$hPin1$_{WW}$-L1-$A_{CC}$ derived peptides. G1-4: Gradients used in this study).

| | **Sequence** | **$[M+H]^+$ (calc.)** | **$[M+H]^+$ (found)** | **$t_r$ (min)** |
|---|---|---|---|---|
| $B_{CC}$-WW-2 | Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSRSRG-OH | 5266.2 | 5264.5 | 20.5 (G5) |
| $B_{CC}$-WW-3 | Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKKMSKSKG-OH | 5182.1 | 5180.8 | 20.1 (G5) |
| $B_{CC}$-WW-4 | Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSHTSG-OH | 5192.1 | 5190.1 | 20.7 (G5) |
| $A_{CC}$-WW-6 | H-RVYYFNHITNAEQWERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5811.0 | 5809.5 | 20.2 (G6) |
| $A_{CC}$-WW-7 | H-RVYYFNHITNASQKERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5711.0 | 5710.5 | 19.5 (G6) |
| $A_{CC}$-WW-8 | H-RVYYSNHITNASQSERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5609.9 | 5609.7 | 19.1 (G6) |
| $A_{CC}$-WW-9 | H-RVYYFNHITNASQYERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5746.0 | 5745.1 | 20.0 (G6) |
| $A_{CC}$-WW-10 | H-RVYYHNHITNAEQSERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-$NH_2$ | 5701.0 | 5700.4 | 19.5 (G6) |
| WW-1-8 | H-KLPPGWEKRMSRSSGRVYYSNHITNASQSERPSG-OH | 3862.9 | 3862.8 | 18.5 (G2) |
| WW-2-8 | H-KLPPGWEKRMSRSRGRVYYSNHITNASQSERPSG-OH | 3932.0 | 3931.0 | 18.4 (G2) |
| WW-2-10 | H-KLPPGWEKRMSRSRGRVYYHNHITNAEQSERPSG-OH | 4023.0 | 4022.8 | 20.6 (G2) |

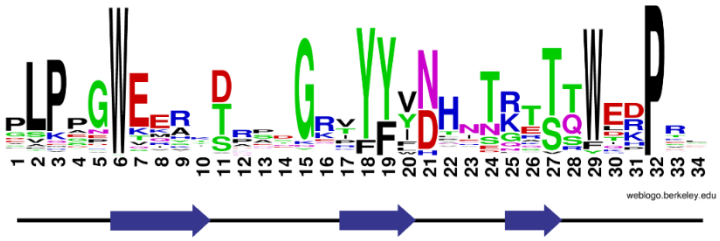


**Figure S1**. Sequence logo (top) of a sequence alignment of 85 WW domain sequences.[9] The height of a stacking symbol at a sequence position represents the degree of residual conservation. Schematic representation of secondary structure motifs (bottom). Black lines represent disordered and loop regions; blue arrows represent β-sheet regions.

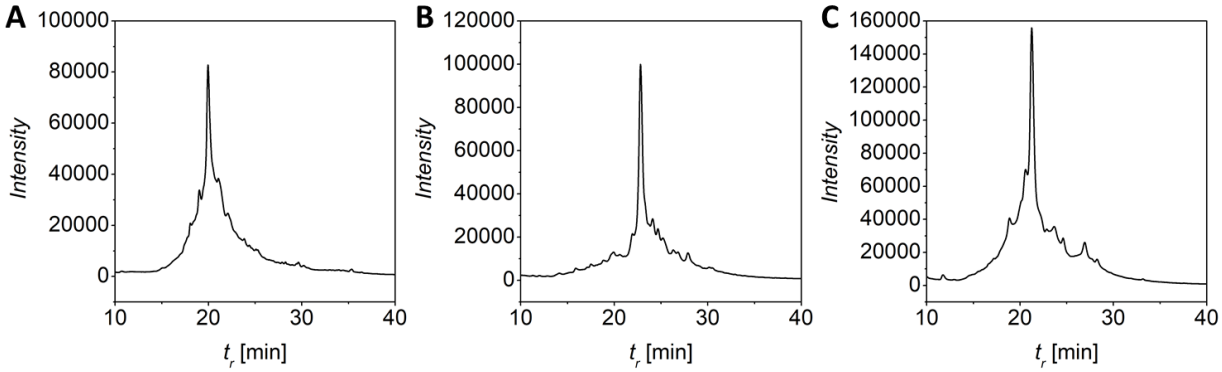


**Figure S2**. HPLC traces of crude hPin1$_{WW}$ (A), hPin1$_{WW}$-L1-A$_{CC}$ (B), and B$_{CC}$-hPin1$_{WW}$-L1 (C) at 220 nm. Gradient: 10-70% B, 30 min; Column: Nucleodur 100-5-C18 ec, (250 mm by 4.6 mm, 5 µm).

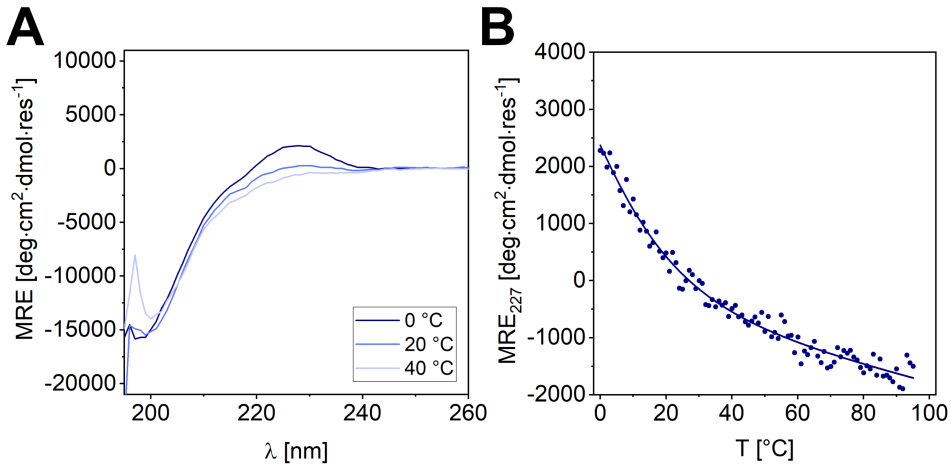


**Figure S3**. CD spectroscopic data of an 1:1-mixture of hPin1$_{WW}$(6-20) and hPin1$_{WW}$(21-29). A) CD spectra acquired at 0 °C, 20 °C, and 40 °C; B) Thermal denaturation profile at 227 nm. Data was fitted using model of a two-state transition between a folded dimer and unfolded monomer and assuming that $\Delta C_{p,m} = 0$ (equations 4, 6, 7, 8).

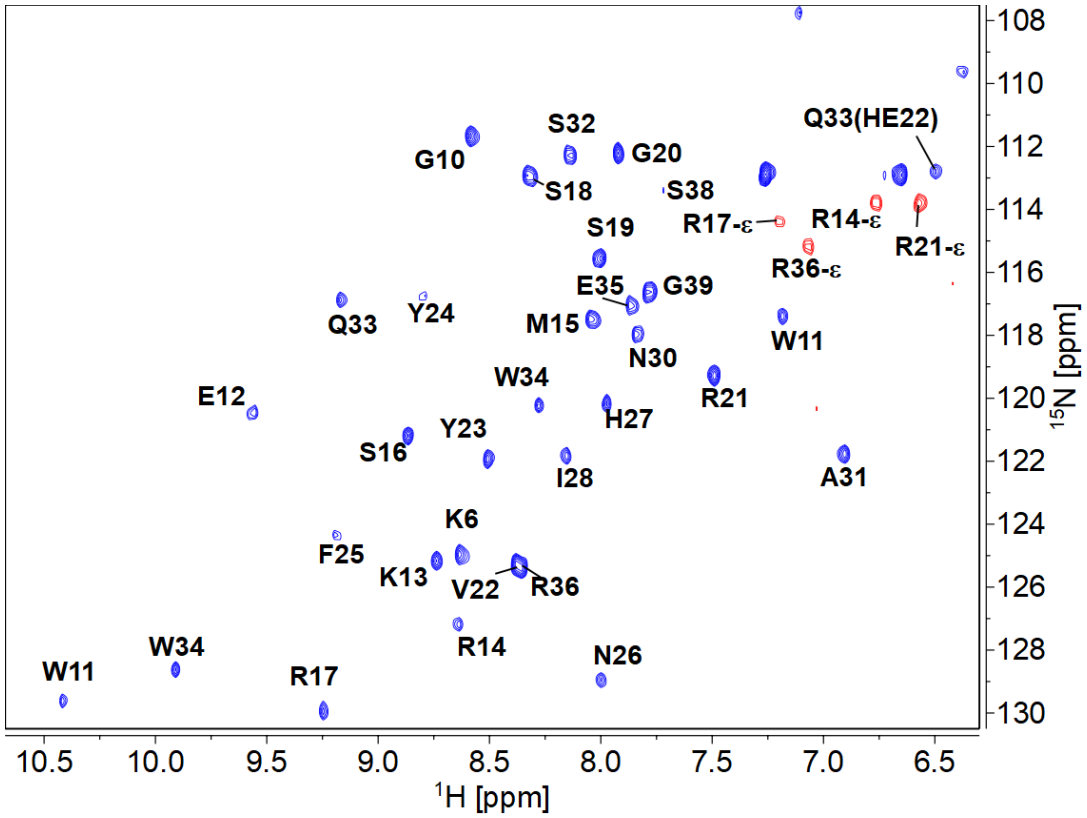


**Figure S4**. 2D $^{1}$H-$^{15}$N HSQC NMR spectrum of the amide region of non-isotopically labeled hPin1$_{WW}$ ($c$ = 1 mM), acquired at $T$ = 283 K, $B_0$ = 18.8 T. The signals of R14-ε, -R17-ε, R21-ε and R36-ε are folded into the spectral range shown and thus appear negative (colored in red). Assignment of cross-peaks is shown using the one letter code of amino acids following the position in the primary sequence of hPin1$_{WW}$ and is based on literature.[10–12]

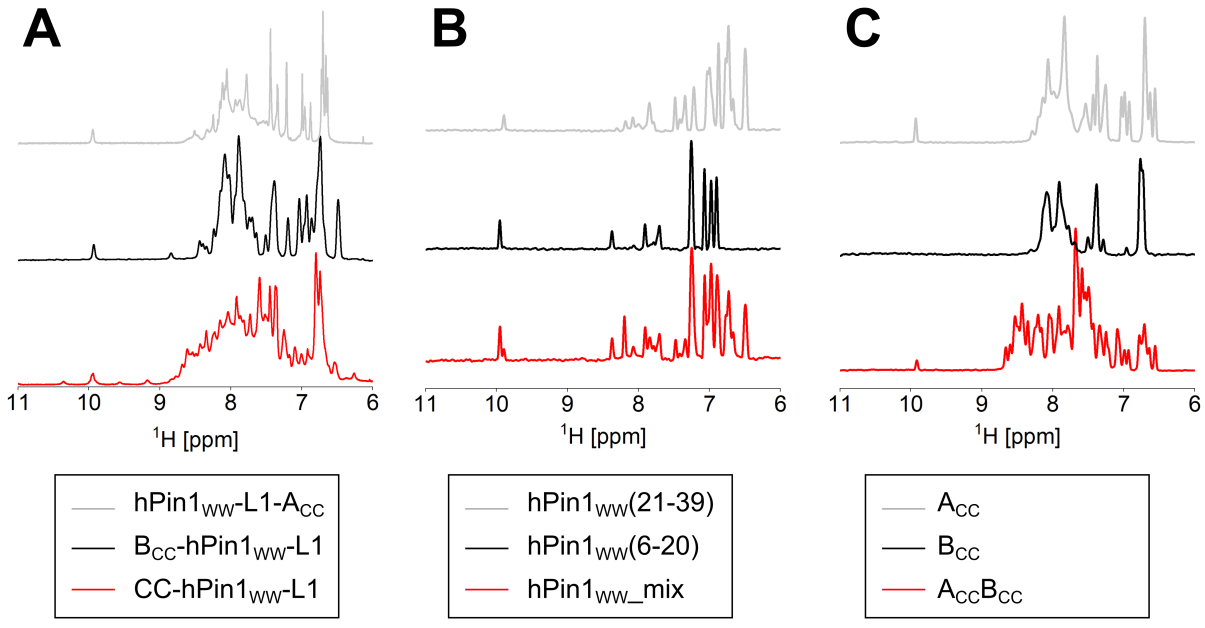


**Figure S5**. 1D proton NMR spectra of individual fragments and corresponding one-to-one mixtures of A) CC-hPin1$_{WW}$-L1, B) hPin1$_{WW}$(6-20)/hPin1$_{WW}$(21-39), and C) A$_{CC}$B$_{CC}$ . Note: A$_{CC}$B$_{CC}$ contains a tryptophan residue in A$_{CC}$ that is not present in the A$_{CC}$-segment of hPin1$_{WW}$-L1-A$_{CC}$.

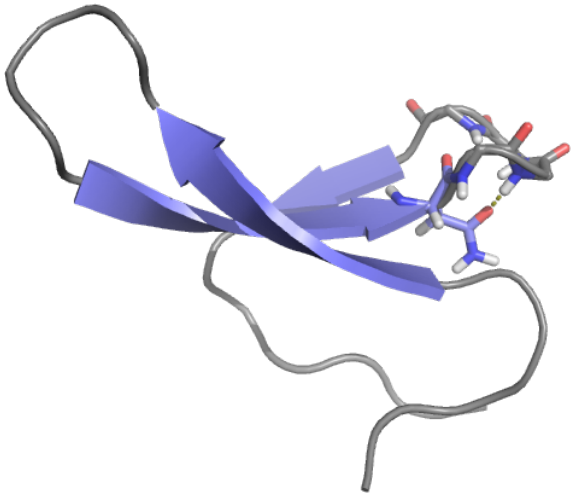

**Figure S6**. Cartoon representation of the NMR structure of hPin1$_{WW}$ (pdb 2m8i) highlighting the hydrogen bond between N26 and the amide proton of I28 (dashed line) that stabilizes the entire fold of this peptide.[13]

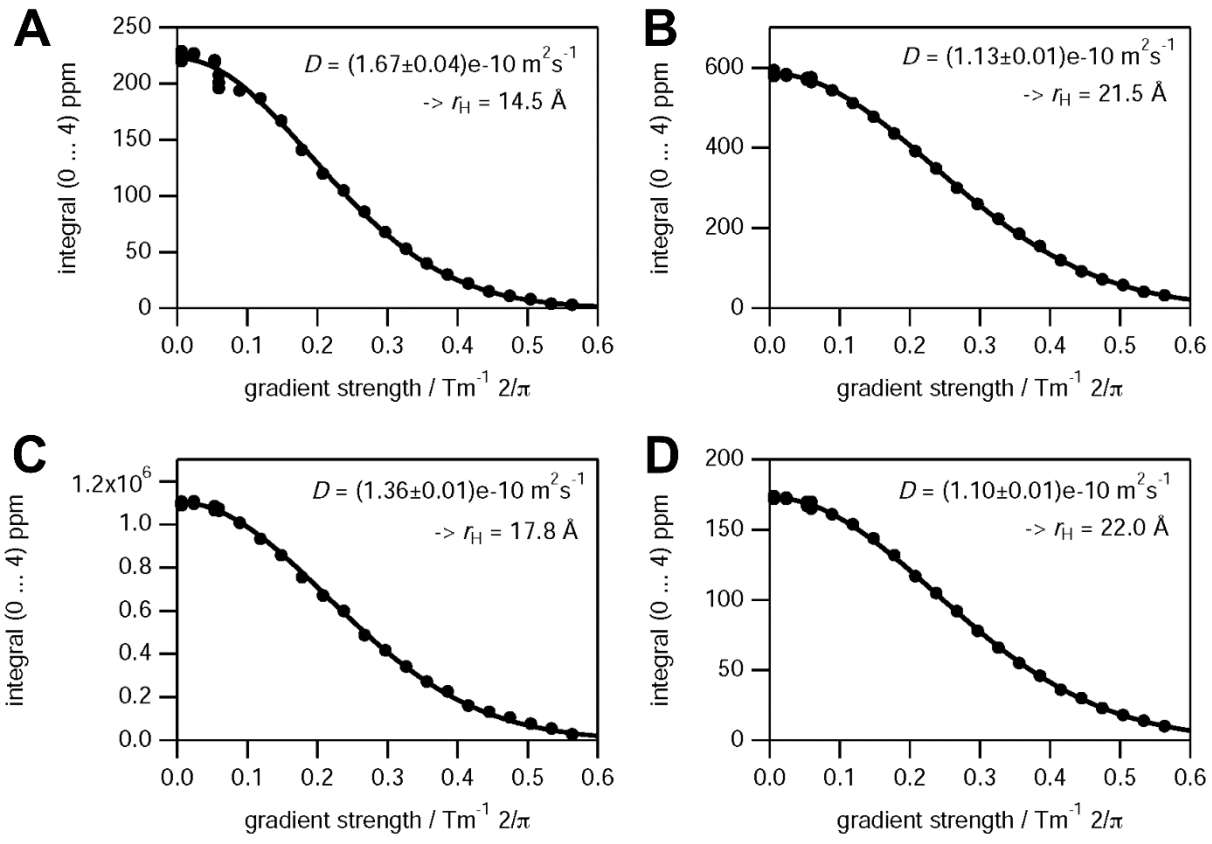


**Figure S7**. NMR diffusion profiles obtained for A) hPin1$_{WW}$, B) B$_{CC}$-hPin1$_{WW}$-L1, C) hPin1$_{WW}$-L1-A$_{CC}$, and D) CC-hPin1$_{WW}$-L1.

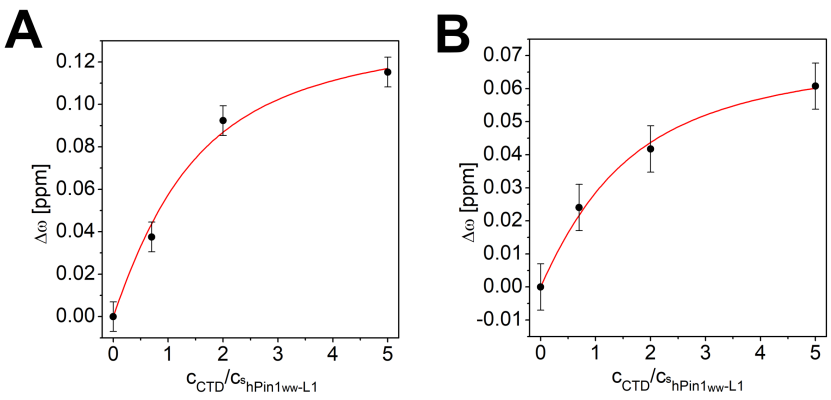


**Figure S8**. Saturation binding of CTD obtained from chemical shift perturbation in R17 (A, $K_d$ = 45 ± 20 µM) and R17-ε (B, $K_d$ = 50 ± 15 µM). All dissociation constants, $K_d$, characterizing binding were obtained from nonlinear least square fit applying saturation binding curves (see SI Eq. 13)

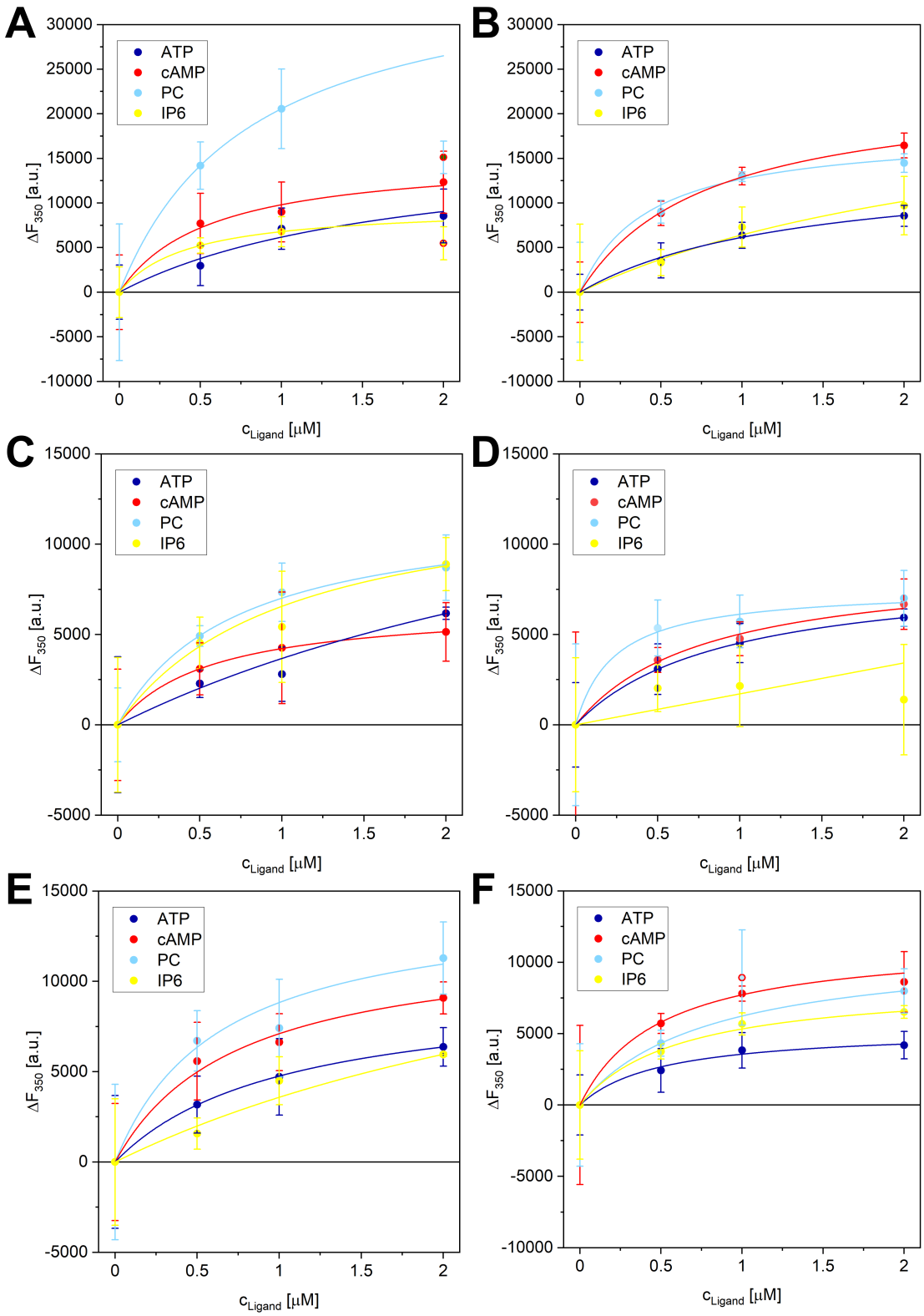


**Figure S9**. Change in tryptophan fluorescence observed for the split WW domain library upon ligand binding. A) Ligand binding to CC-WW-1-5, B) to CC-WW-1-6, C) to CC-WW-1-7, D) to CC-WW-1-8, E) to CC-WW-1-9, and F) to CC-WW-1-10. Data was fitted to an one-site binding model (Eq. 14).

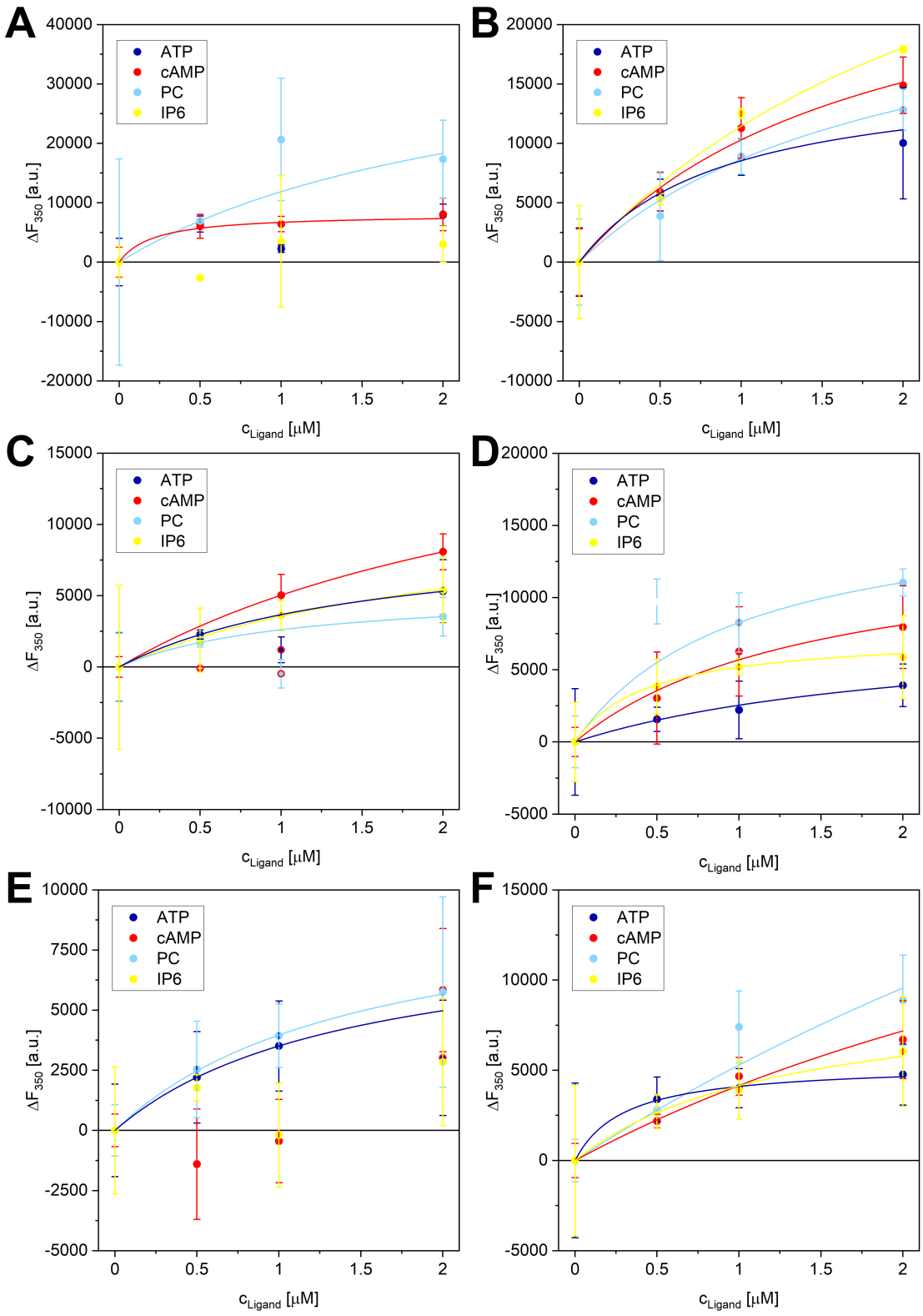


**Figure S10**. Change in tryptophan fluorescence observed for the split WW domain library upon ligand binding. Ligand binding A) to CC-WW-2-5, B) to CC-WW-2-6, C) to CC-WW-2-7, D) to CC-WW-2-8, E) to CC-WW-2-9, and F) to CC-WW-2-10. Data was fitted to an one-site binding model (Eq. 14).

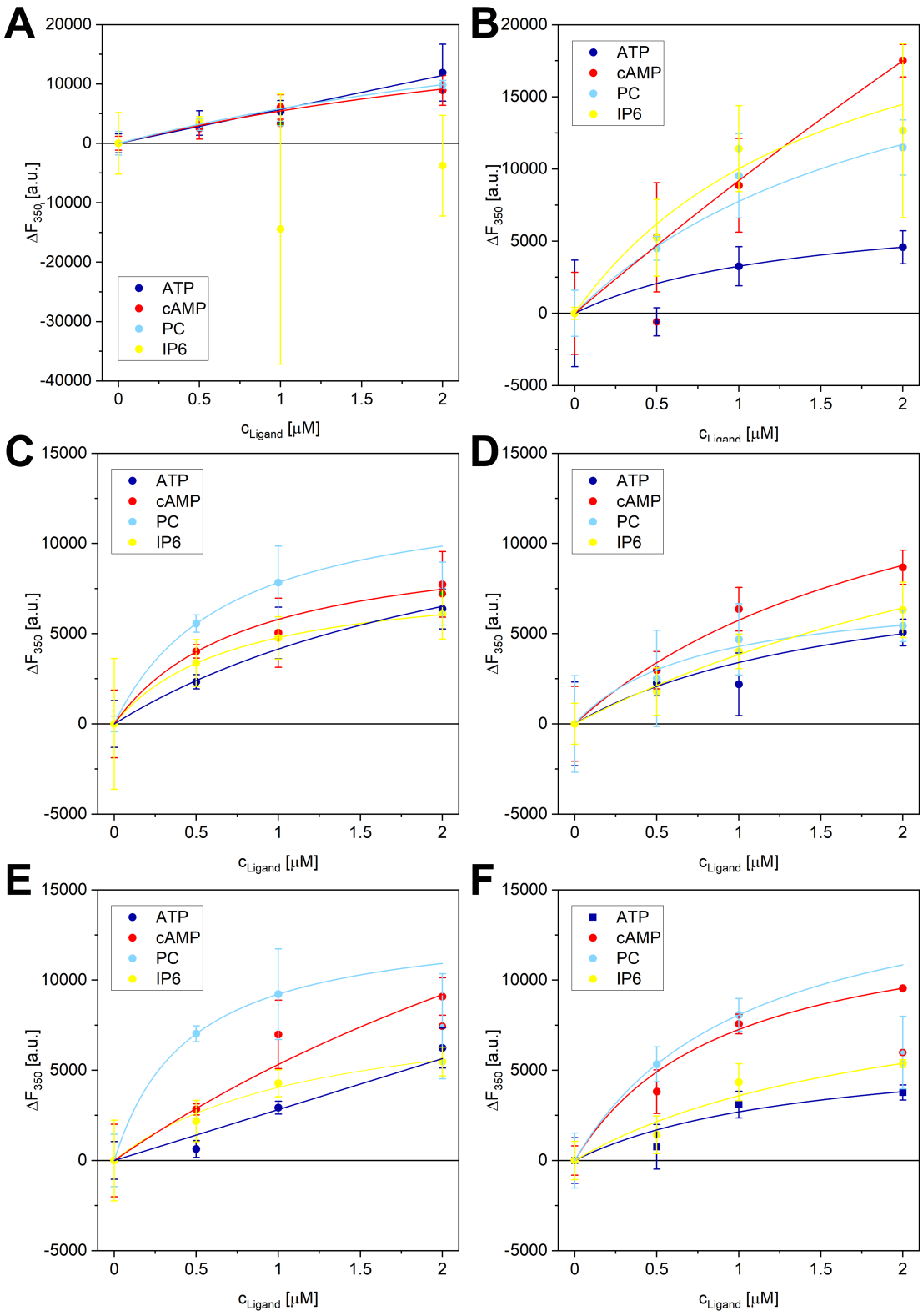


**Figure S11**. Change in tryptophan fluorescence observed for the split WW domain library upon ligand binding. Ligand binding A) to CC-WW-3-5, B) to CC-WW-3-6, C) to CC-WW-3-7, D) to CC-WW-3-8, E) to CC-WW-3-9, and F) to CC-WW-3-10. Data was fitted to an one-site binding model (Eq. 14).

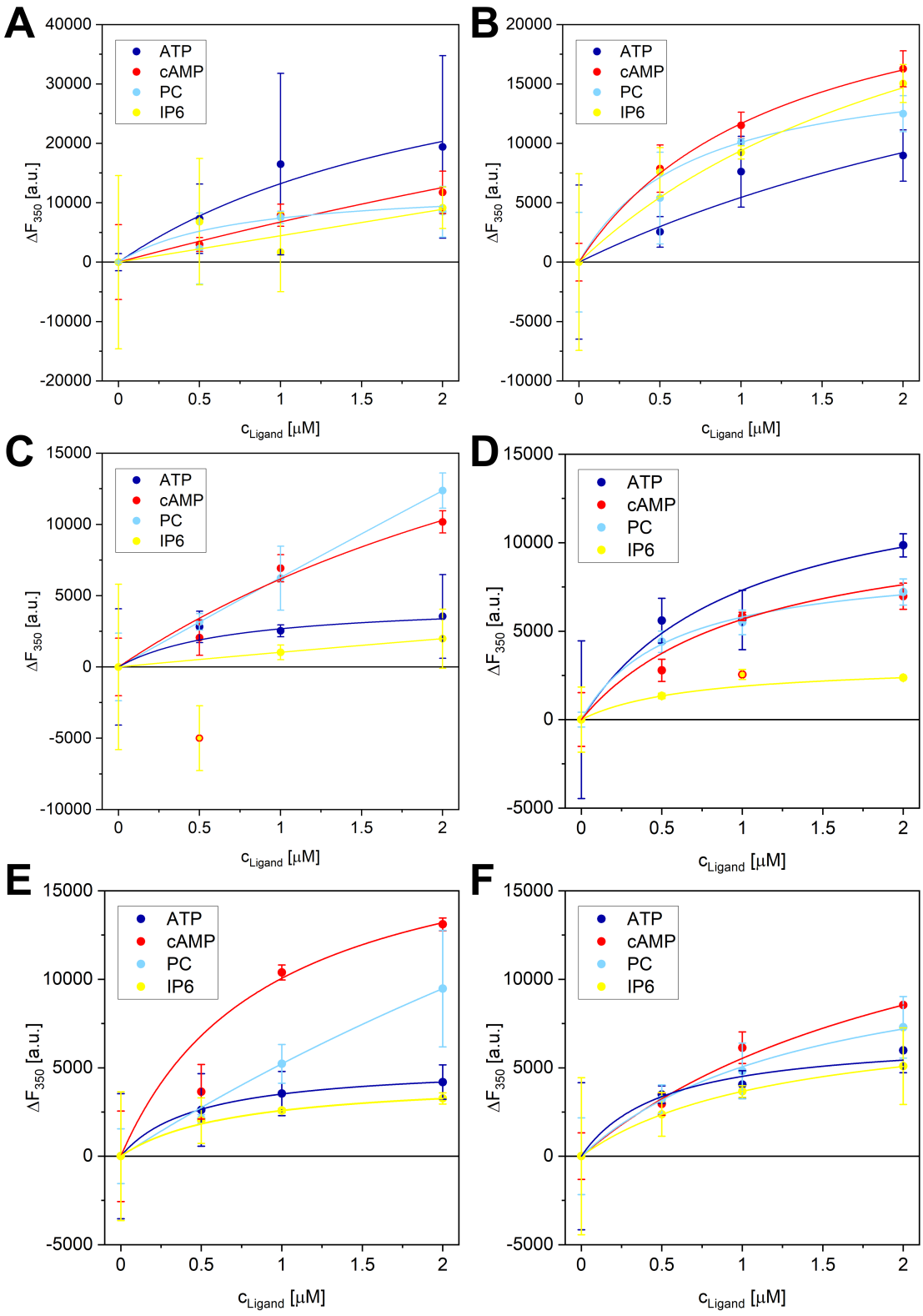


**Figure S12**. Change in tryptophan fluorescence observed for the split WW domain library upon ligand binding. Ligand binding A) to CC-WW-4-5, B) to CC-WW-4-6, C) to CC-WW-4-7, D) to CC-WW-4-8, E) to CC-WW-4-9, and F) to CC-WW-4-10. Data was fitted to an one-site binding model (Eq. 14).

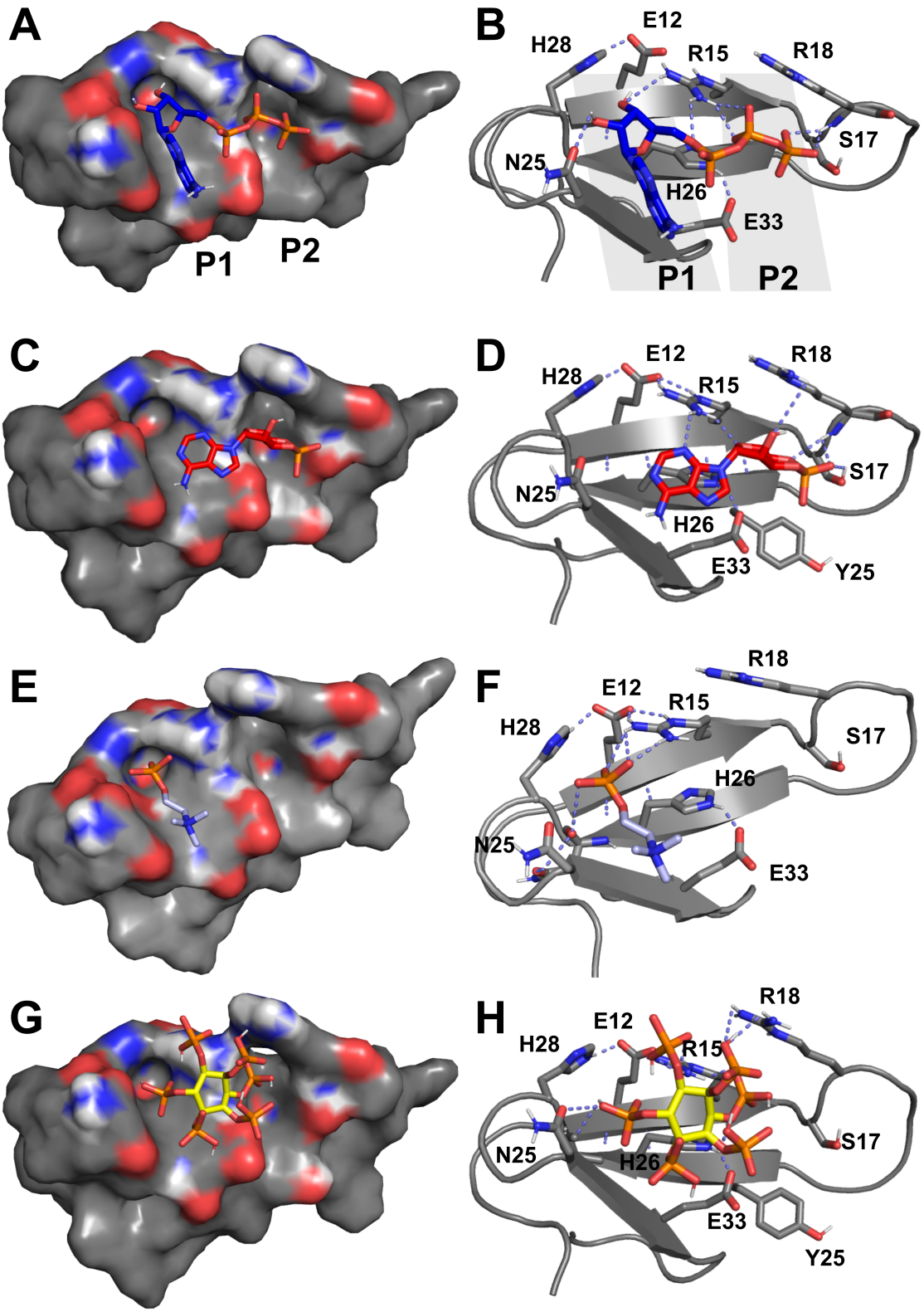


**Figure S13**. Computational models of WW-2-10 binding to phosphate esters. Surface representations of WW-2-10 binding to A) ATP, C) cAMP, E) phosphorylcholine, and G) inositol hexaphosphate. B, D, F, H) Cartoon representations of WW-2-10 binding to B) ATP, D) cAMP, F) phosphorylcholine, and H) inositol hexaphosphate. The polar contacts are shown as light blue dashed lines. P1 indicates binding pocket 1 and P2 indicates binding pocket 2.

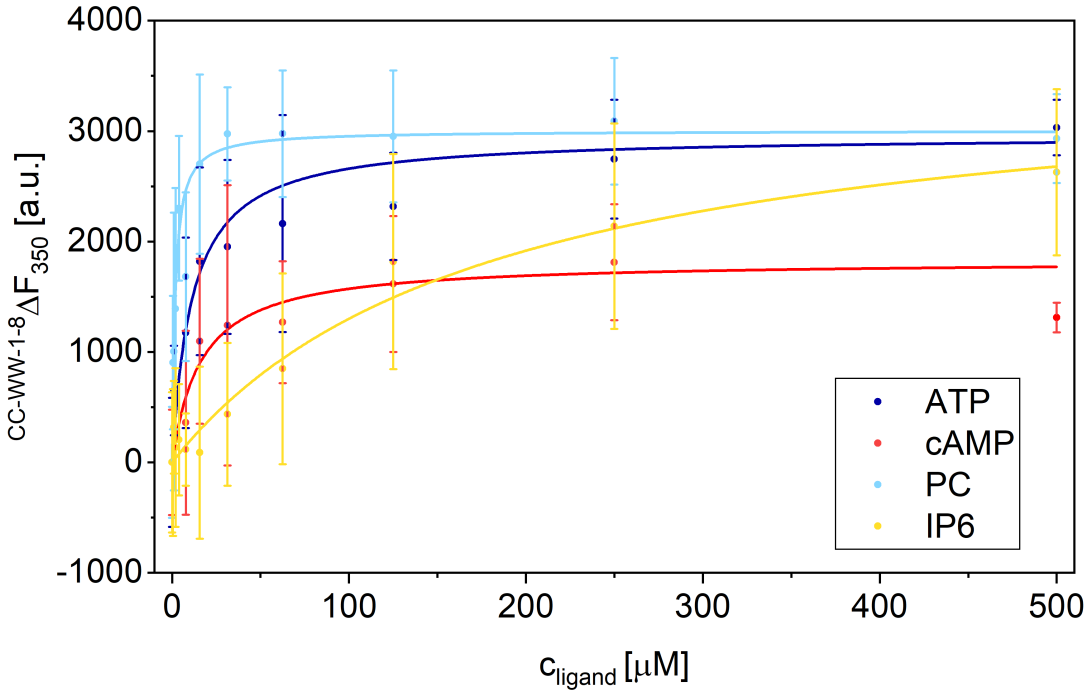


**Figure S14**. Intrinsic tryptophan fluorescence titrations of ATP, cAMP, phosphorylcholine, and IP6 to CC-WW-1-8 lead to the following $K_d$-values: $K_d$(ATP) = 11.4 ± 4.6 µM, $K_d$(cAMP) = 16.3 ± 0.7 µM, $K_d$(PC) = 1.7 ± 0.4 µM, $K_d$(IP6) = 180 ± 61 µM.

# 3. Analytical HPLC and MALDI-TOF mass spectrometry

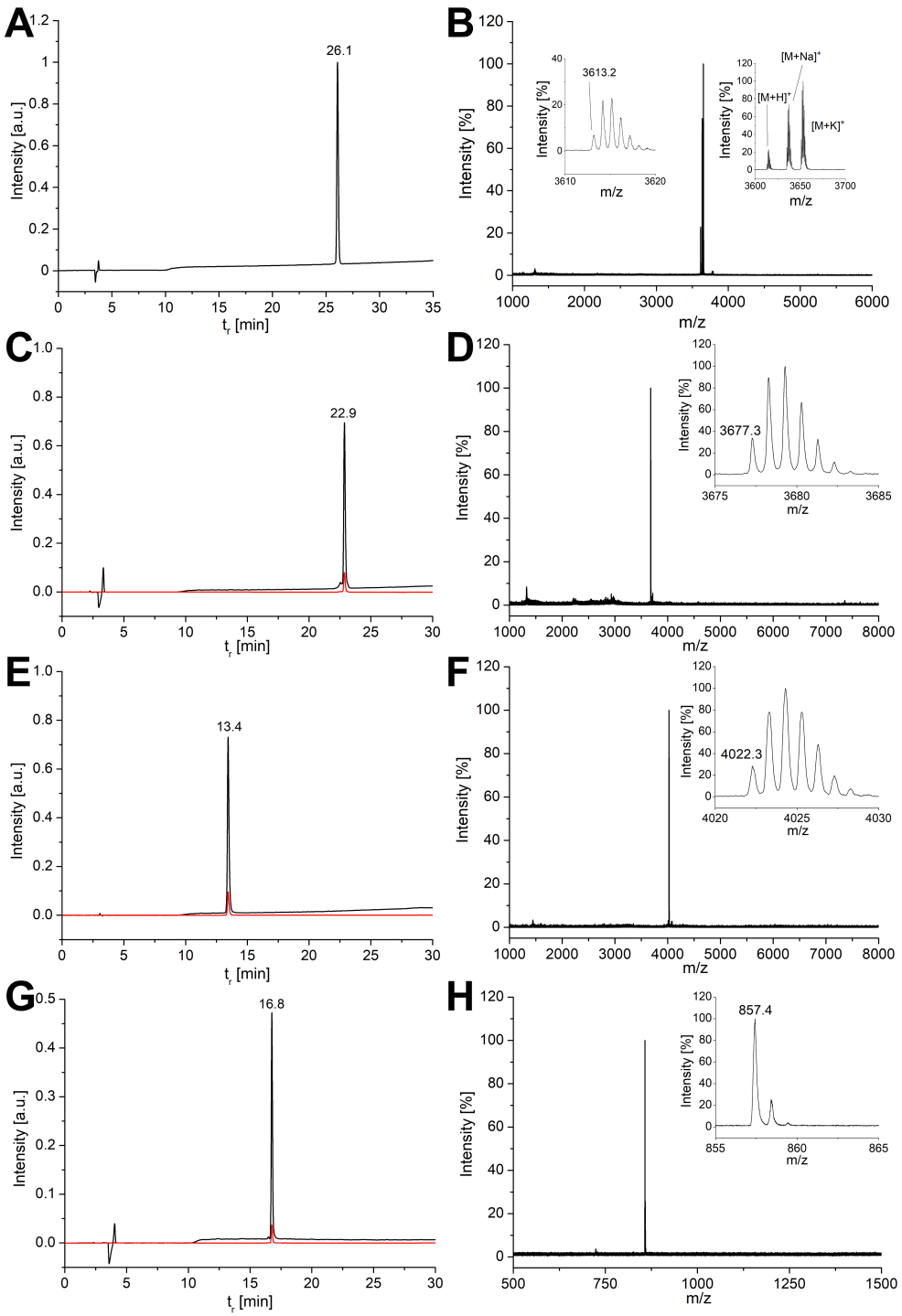


**Figure S15.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of $B_{CC}$ (A), $A_{CC}$ (C), hPin1$_{WW}$ (E), and CTD (G); MALDI-TOF mass spectra of $B_{CC}$ (B), $A_{CC}$ (D), hPin1$_{WW}$ (F), and CDT (H, mass was recorded in linear negative mode). Note: $B_{CC}$ does not contain chromophores. Therefore only the HPLC trace at 220 nm is shown.

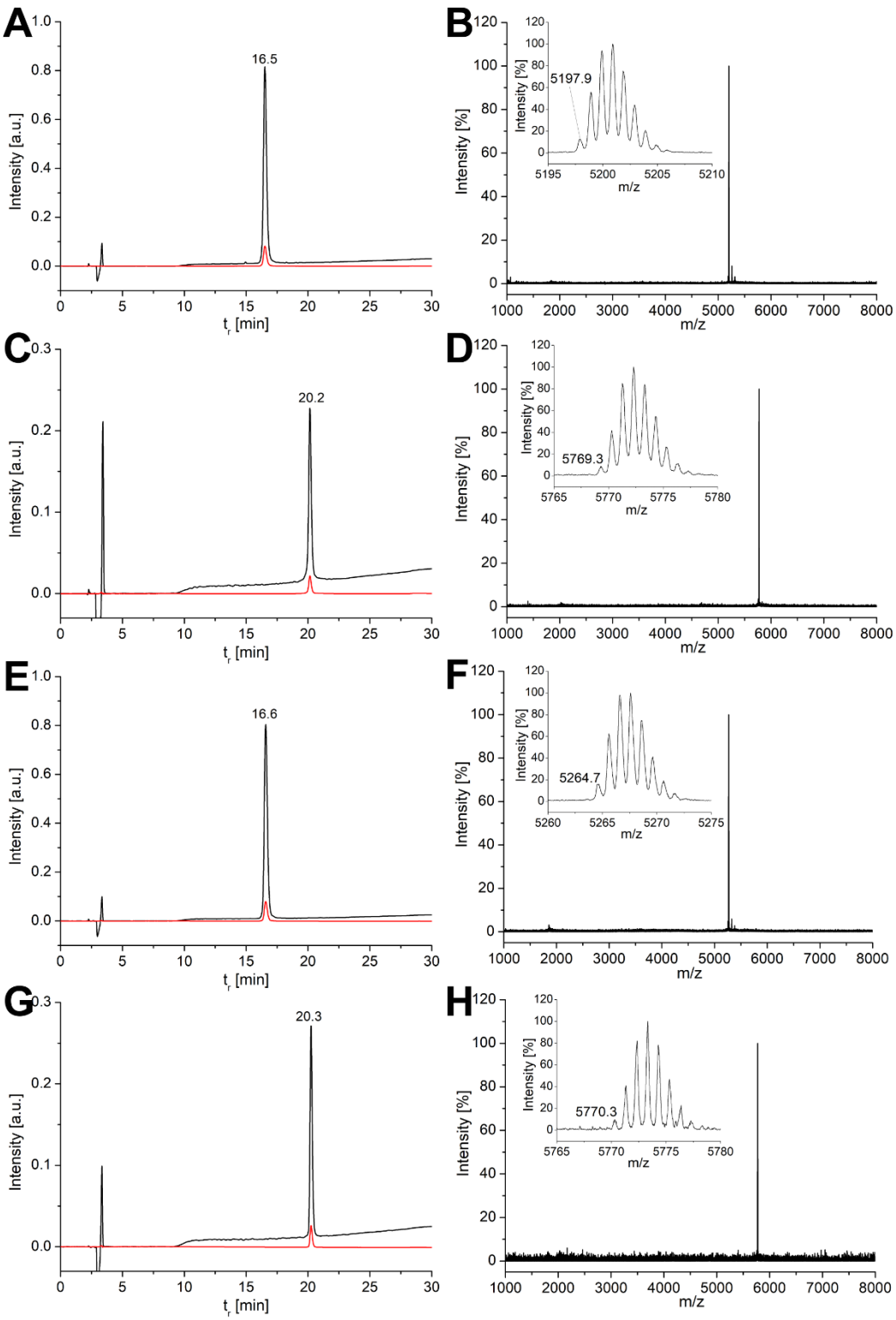


**Figure S16.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of $B_{CC}$-hPin1$_{WW}$-L1 (A), hPin1$_{WW}$-L1-$A_{CC}$ (C), $B_{CC}$-hPin1$_{WW}$-R-$^{13}C_6$, $^{15}N_4$ (E), and hPin1$_{WW}$-L1-$A_{CC}$-F-$^{15}$N (G); MALDI-TOF mass spectra of $B_{CC}$-hPin1$_{WW}$-L1 (B), hPin1$_{WW}$-L1-$A_{CC}$ (D), $B_{CC}$-hPin1$_{WW}$-R-$^{13}C_6$, $^{15}N_4$ (F), and hPin1$_{WW}$-L1-$A_{CC}$-F-$^{15}$N (H).

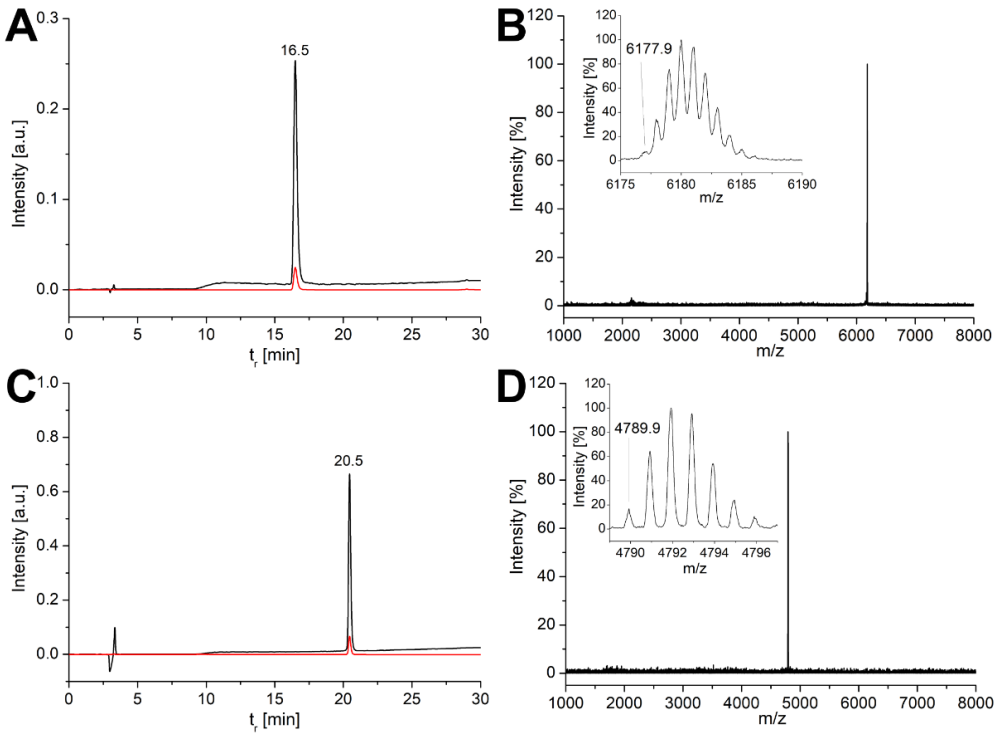


**Figure S17.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of $B_{CC}$-hPin1$_{WW}$-L2 (A), hPin1$_{WW}$-L2-$A_{CC}$ (C); MALDI-TOF mass spectra of $B_{CC}$-hPin1$_{WW}$-L2 (B), hPin1$_{WW}$-L2-$A_{CC}$ (D).

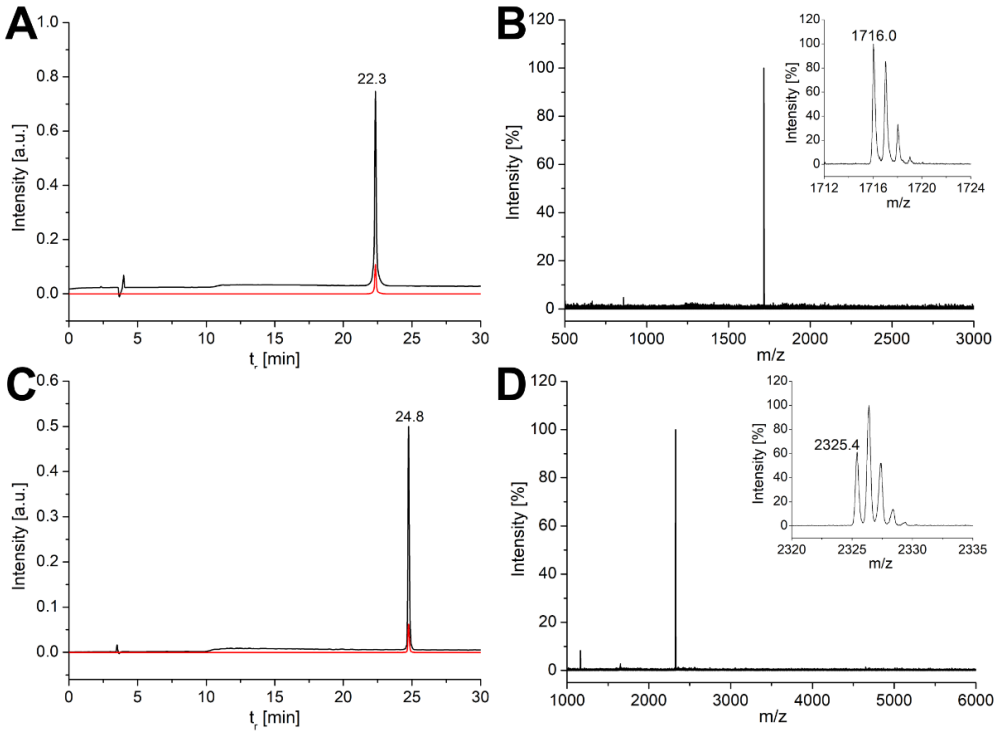


**Figure S18.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of hPin1$_{WW}$(6-20) (A), hPin1$_{WW}$(21-39) (C); MALDI-TOF mass spectra of hPin1$_{WW}$(6-20) (B), hPin1$_{WW}$(21-39) (D).

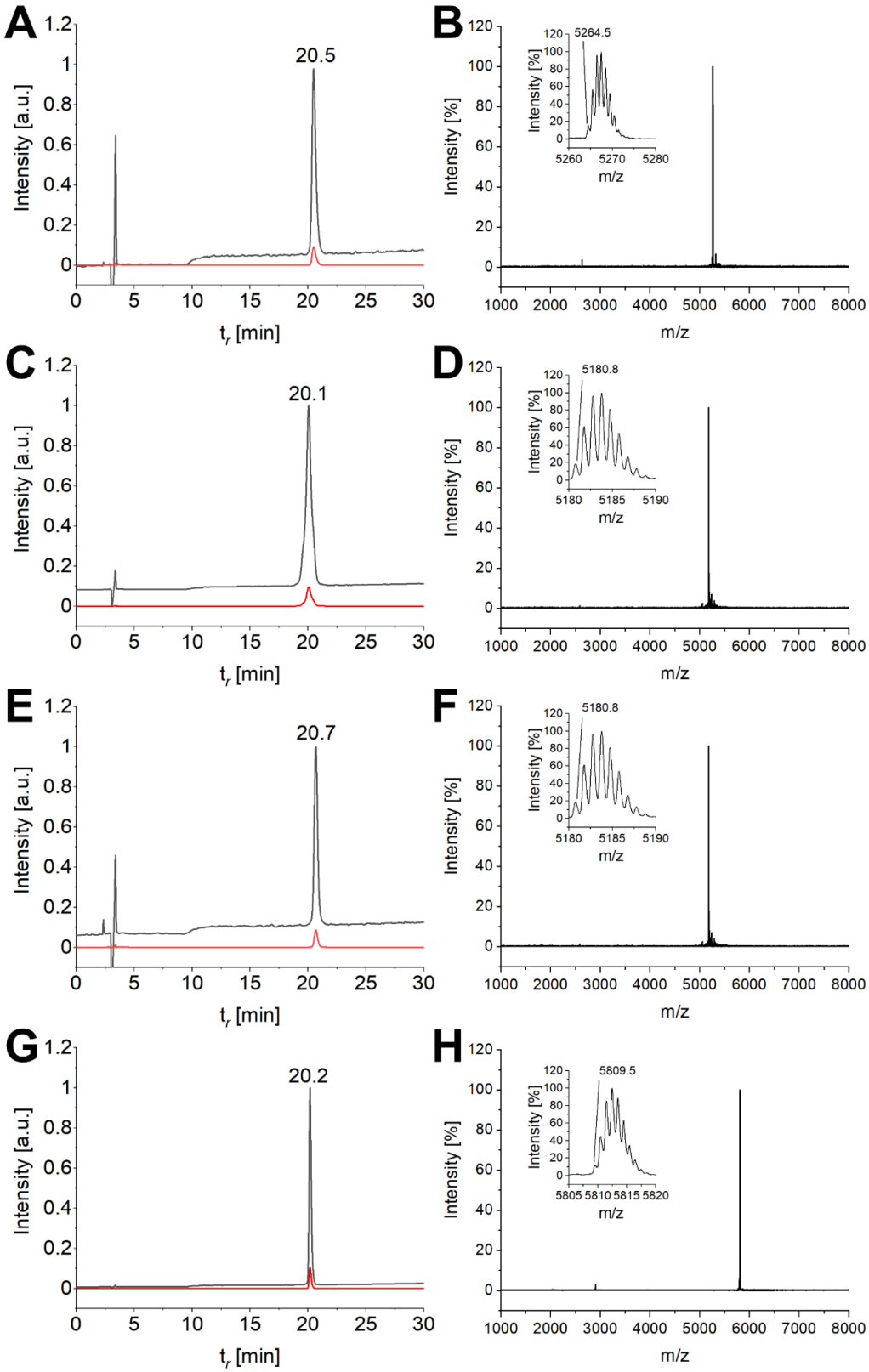


**Figure S19.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of $B_{CC}$-WW-2 (A), $B_{CC}$-WW-3 (C), $B_{CC}$-WW-4 (E), and $A_{CC}$-WW-6 (G); MALDI-TOF mass spectra of $B_{CC}$-WW-2 (B), $B_{CC}$-WW-3 (D), $B_{CC}$-WW-4 (F), and $A_{CC}$-WW-6.

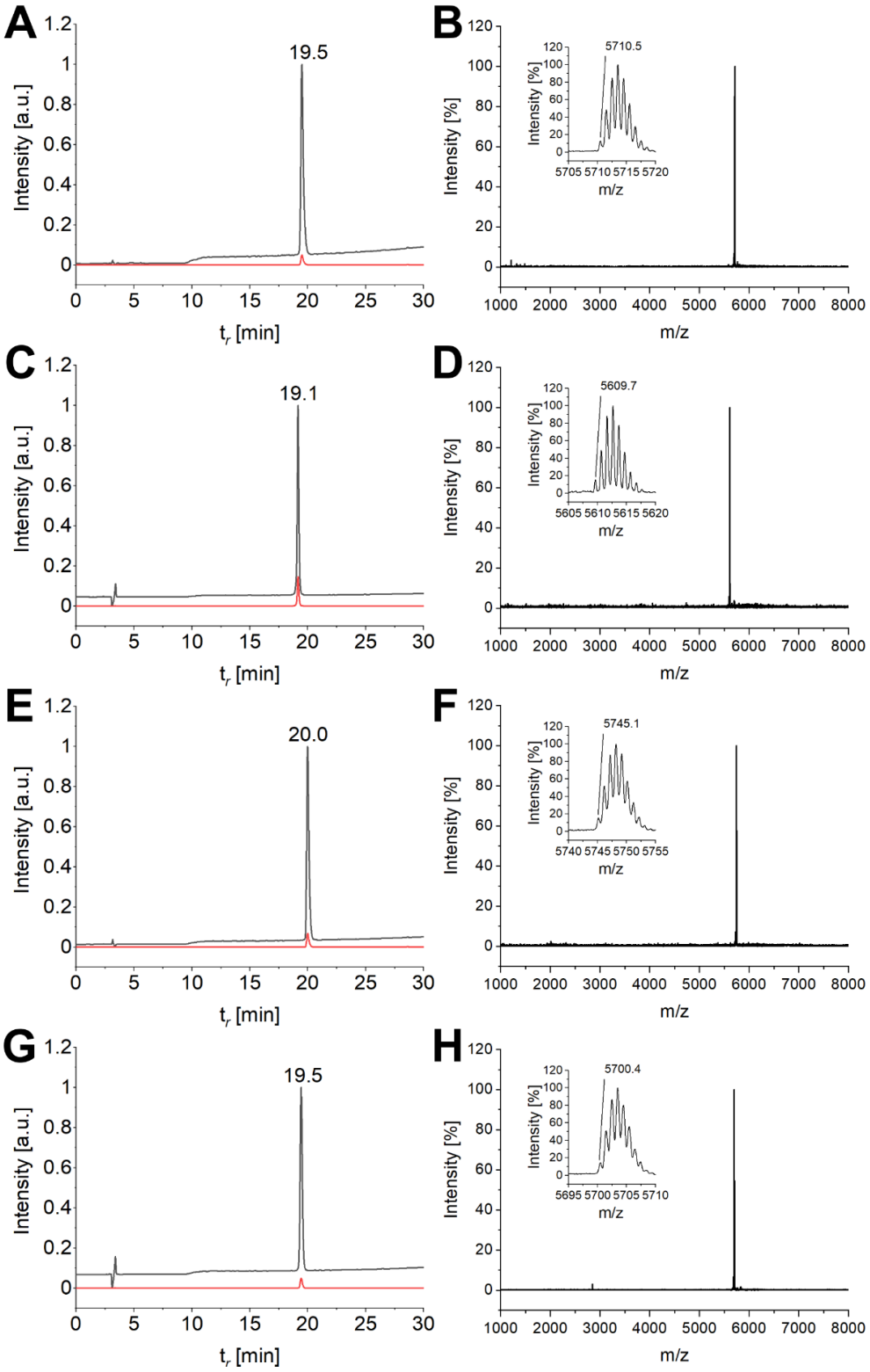


**Figure S20.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of $A_{CC}$-WW-7 (A), $A_{CC}$-WW-8 (C), $A_{CC}$-WW-9 (E), and $A_{CC}$-WW-10 (G); MALDI-TOF mass spectra of $A_{CC}$-WW-7 (B), $A_{CC}$-WW-8 (D), $A_{CC}$-WW-9 (F), and $A_{CC}$-WW-10.

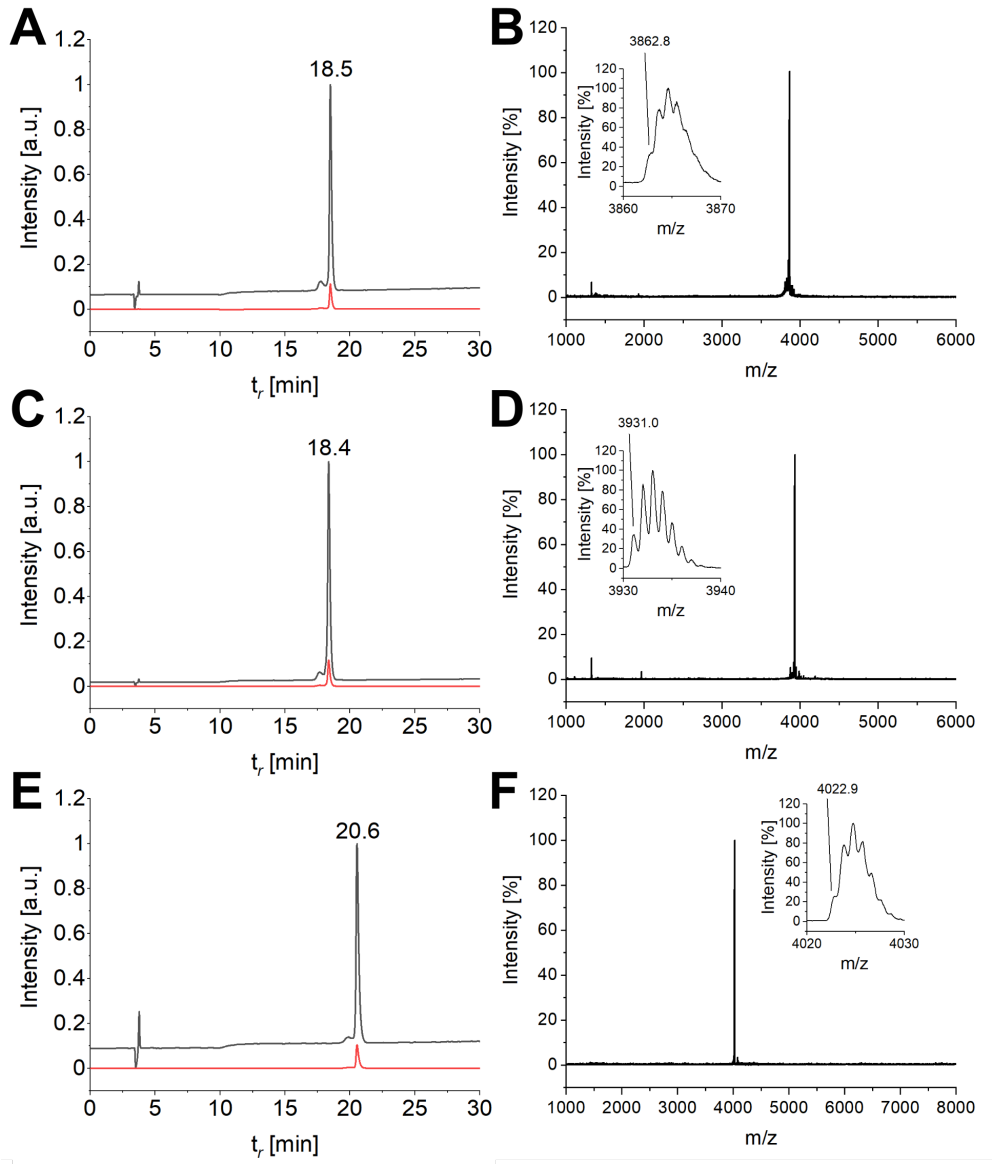


**Figure S21.** HPLC traces at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of synthesized peptides. HPLC traces of WW-1-8 (A),WW-2-8 (C), and WW-2-10 (E); MALDI-TOF mass spectra of WW-1-8 (B), WW-2-8 (D), and WW-2-10.

# 4. Analysis of CD thermal denaturation profiles

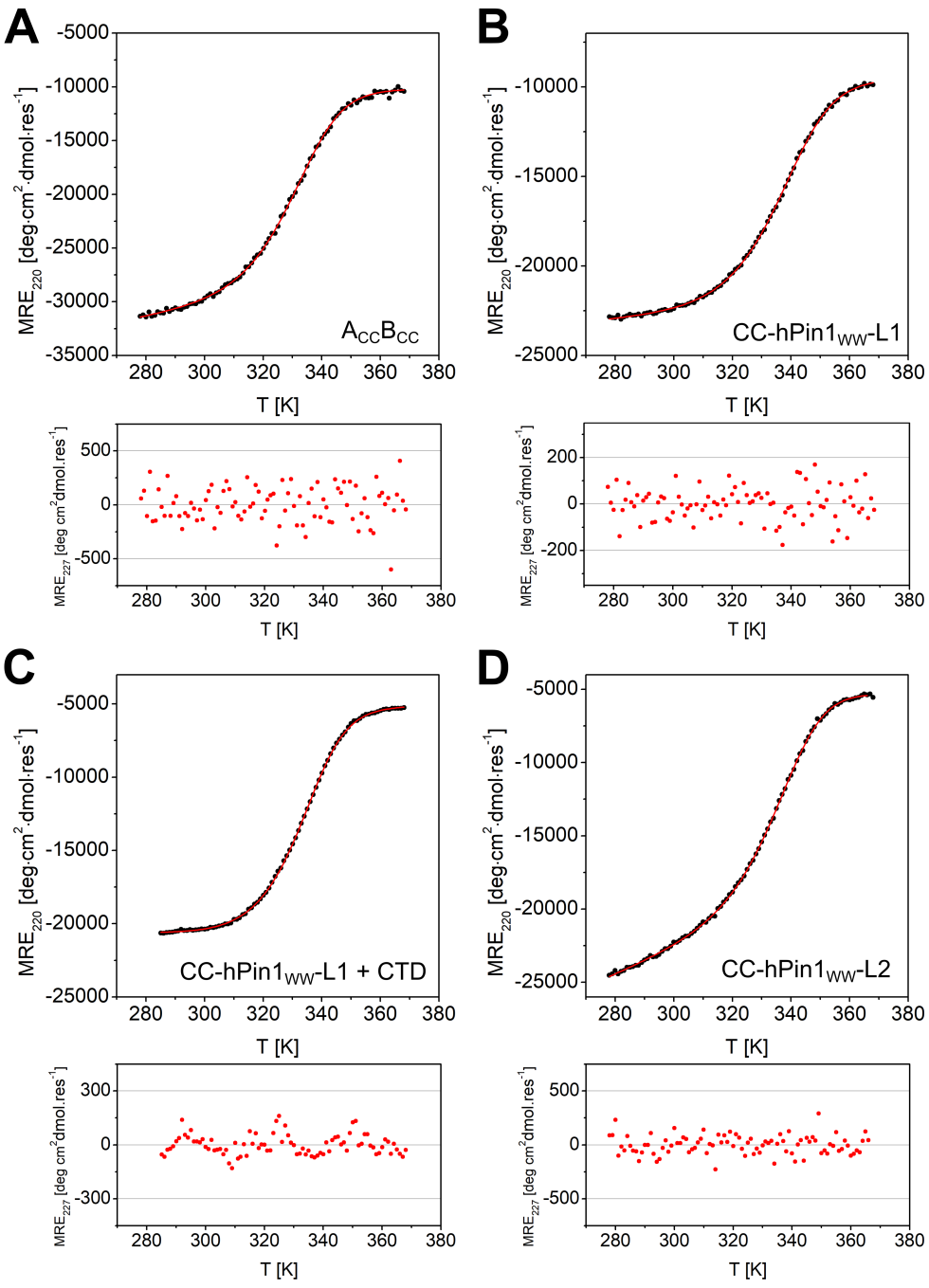


**Figure S22.** Analysis of thermal denaturation profiles (top) and residuals (bottom) to determine the melting temperatures of $A_{CC}B_{CC}$ (A), CC-hPin1$_{WW}$-L1 (B), CC-hPin1$_{WW}$-L1 in the presence of CTD ligand (C), and CC-hPin1$_{WW}$-L2 (D). A model of a two-state transition between a folded dimer and an unfolded monomer assuming $\Delta C_{p,m} = 0$ was applied for data fitting (SI equations 4, 6, 7, 8). Residuals are shown below the respective denaturation profiles.

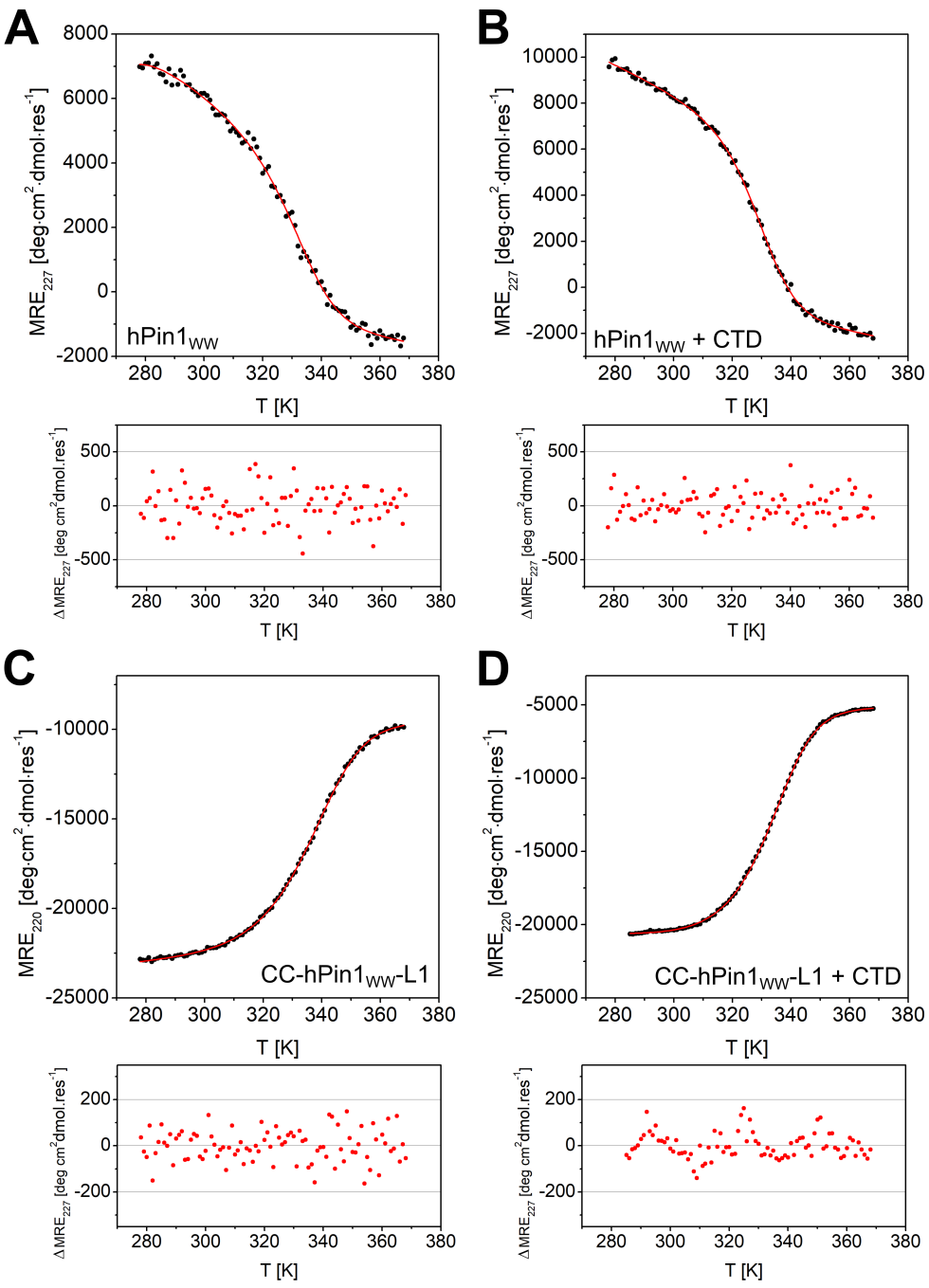


**Figure S23.** Analysis of the thermal denaturation profiles (top) and residuals (bottom) to determine thermodynamic parameters of $hPin1_{WW}$ (A), $hPin1_{WW}$ in the presence of CTD ligand (B), CC-$hPin1_{WW}$-L1 (C), and CC-$hPin1_{WW}$-L1 in the presence of CTD ligand (D). A model for the unfolding of a monomer (A and B, SI equations 3 to 6) or a model of a two-state transition between a folded dimer and an unfolded monomer was applied for data fitting (C and D, SI equations 4, 6, 9, 10). Residuals are shown below the respective denaturation profiles.

## 5. References


1 H. Mach, C. R. Middaugh and R. V. Lewis, *Anal. Biochem.*, 1992, **200**, 74–80.

2 B. J. H. Kuipers and H. Gruppen, *J. Agric. Food Chem.*, 2007, **55**, 5445–5451.

3 L. Whitmore and B. A. Wallace, *Biopolymers*, 2008, **89**, 392–400.

4 L. Whitmore and B. A. Wallace, *Nucleic Acids Res.*, 2004, **32**, 668–673.

5 N. J. Greenfield, *Nat. Protoc.*, 2006, **1**, 2527–2535.

6 J. A. Jones, D. K. Wilkins, L. J. Smith and C. M. Dobson, *J. Biomol. NMR*, 1997, **10**, 199–203.

7 F. Khatib, S. Cooper, M. D. Tyka, K. Xu, I. Makedon, Z. Popović, D. Baker and F. Players, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 18949–18953.

8 J. B. Maguire, H. K. Haddox, D. Strickland, S. F. Halabiya, B. Coventry, J. R. Griffin, S. V. S. R. K. Pulavarti, M. Cummins, D. F. Thieker, E. Klavins, T. Szyperski, F. DiMaio, D. Baker and B. Kuhlman, *Proteins Struct. Funct. Bioinforma.*, 2021, **89**, 436–449.

9 S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell and A. J. Olson, *Nat. Protoc.*, 2016, **11**, 905–919.

10 J. A. Kowalski, K. Liu and J. W. Kelly, *Biopolymers*, 2002, **63**, 111–121.

11 K. Kurakula, D. S. Koenis, M. A. Herzik, Y. Liu, J. W. Craft, P. B. van Loenen, M. Vos, M. K. Tran, H. H. Versteeg, M. J. T. H. Goumans, W. Ruf, C. J. M. de Vries and M. Şen, *Haematologica*, 2018, **103**, 1073–1082.

12 C. Schelhorn, P. Martín-Malpartida, D. Sunõl and M. J. Macias, *Sci. Rep.*, 2015, **5**, 1–12.

13 L. M. Luh, R. Hänsel, F. Löhr, D. K. Kirchner, K. Krauskopf, S. Pitzius, B. Schäfer, P. Tufar, I. Corbeski, P. Güntert and V. Dötsch, *J. Am. Chem. Soc.*, 2013, **135**, 13796–13803.