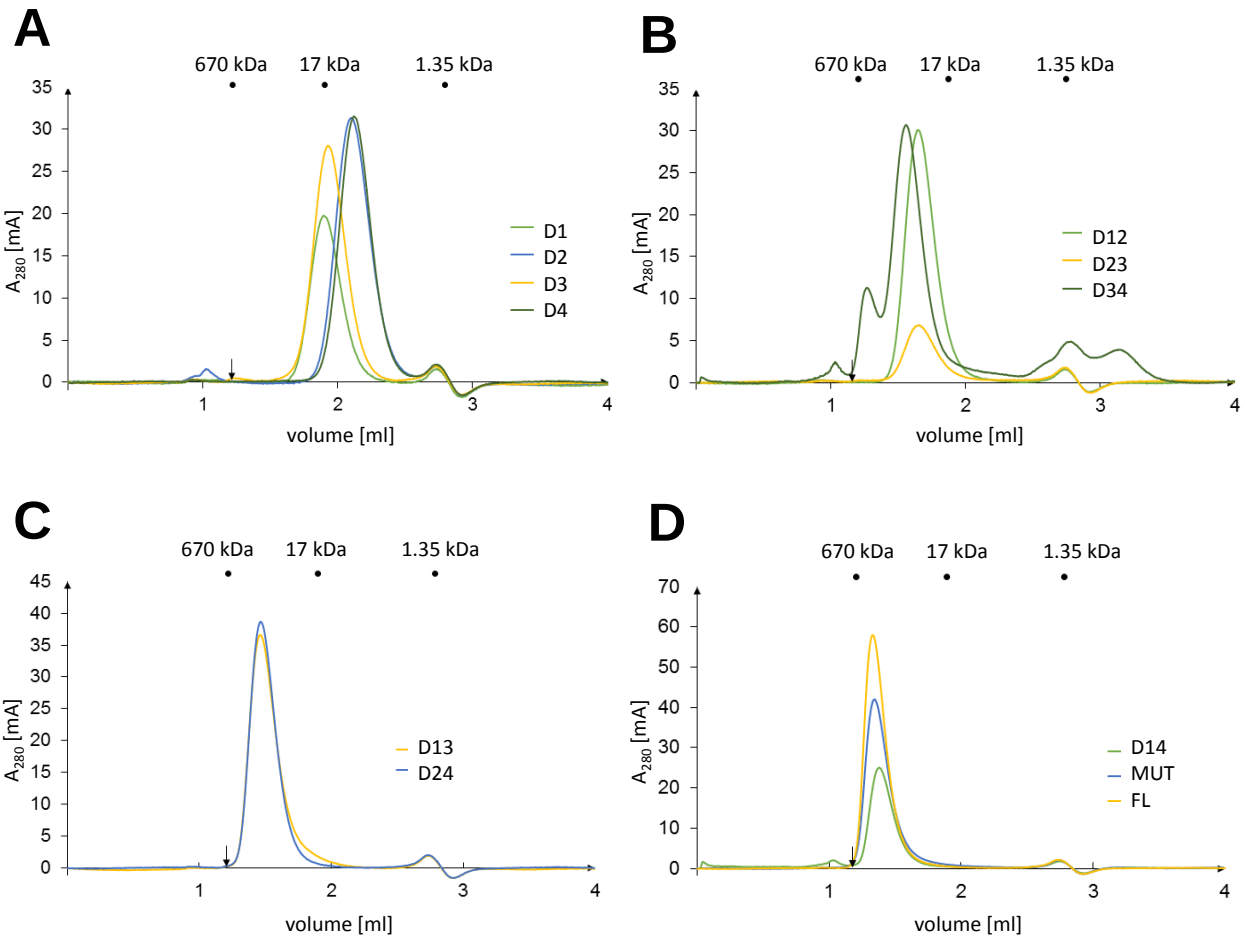
Table S1. Overview of used PDZK1 constructs and their binding affinities to peptides. Related to figure 1.

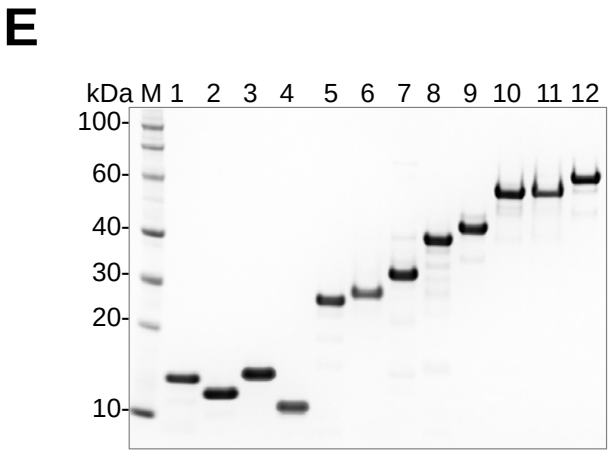
PDZK1 constructs	Amino acids	MW [kDa]	K_D [µM] PEPT2-CT-FITC	K_D [µM] NHE3-CT-FITC	K_D [µM] PDZK1-CT-FITC	EC [$cm^{-1} \cdot M^{-1}$]
D1	7-116	15.0	83 ± 8	3.0 ± 0.2	413 ± 13	4470
D2	133-219	12.3	91 ± 8	n.b.*	>500	5960
D3	239-324	12.0	82 ± 10	>50	n.b.*	4470
D4	374-460	12.1	39 ± 2	n.b. *	314 ± 30	5960
D12	7-219	26.4	49 ± 3	n/d**	n/d**	14440
D23	133-324	23.8	32 ± 2	n/d**	n/d**	8940
D34	239-460	26.8	23 ± 2	n/d**	204 ± 10	13410
D13	7-324	37.9	24 ± 1	n/d**	n/d**	17420
D24	133-460	38.7	12.7 ± 0.5	n/d**	n/d**	17880
D14	7-460	52.7	4.1 ± 0.1	10 ± 1	143 ± 21	26360
MUT	7-460	52.2	15.3 ± 0.7	17 ± 2	117 ± 42	24995
FL	1-519	59.8	4.9 ± 0.2	19 ± 3	286 ± 36	27850

* n.b., no binding

** n/d., not determined

Table S1 legend. List of PDZK1 constructs and affinities. Details of the PDZK1 constructs: Residues ranges for the different constructs, molecular weights (MW) including affinity tags and calculated molar extinction coefficients (EC). These columns are followed by dissociation constants of the PDZK1 constructs with the PEPT2-CT-FITC, NHE3-CT-FITC and PDZK1-CT peptides (C-terminal peptides from these proteins). Values were obtained by fluorescence polarization experiments. K_D's were calculated from three replicate experiments, where the uncertainty corresponds to the standard error of the fit.





Lane	PDZK1 domain	MW [kDa]
1	D1	15.0
2	D2	12.3
3	D3	9.5
4	D4	9.6
5	D12	26.4
6	D23	23.8
7	D34	26.8
8	D13	37.9
9	D24	38.7
10	D14	52.7
11	MUT	52.2
12	FL	59.8

Figure S1. Quality control of PDZK1 constructs by analytical gel filtration and SDS PAGE. Related to figure 1. Chromatograms show samples analyzed on a Superdex 75 (SD75) 5/150 home-packed column using the 1260 Infinity Bio-inert high-performance liquid chromatography system (Agilent Technologies). (A) Gel-filtration chromatogram of single domains, (B) double domains, (C) triple domains and (D) tetra domain constructs. The void volume of the column is indicated by a black arrow and elution volumes of

protein standards by a dots. (E) Representative SDS-PAGE gel showing the purity of the samples. For reference, the SDS gel was run together with a molecular weight (MW) marker (M) (molecular weights for reference proteins are indicated in kDa). Lanes 1-4: single domain constructs, lanes 5-7: double domain construct, lanes 8-9: triple domain constructs and lanes 10-12: tetra domain constructs. MWs of constructs are shown in the table on the right.

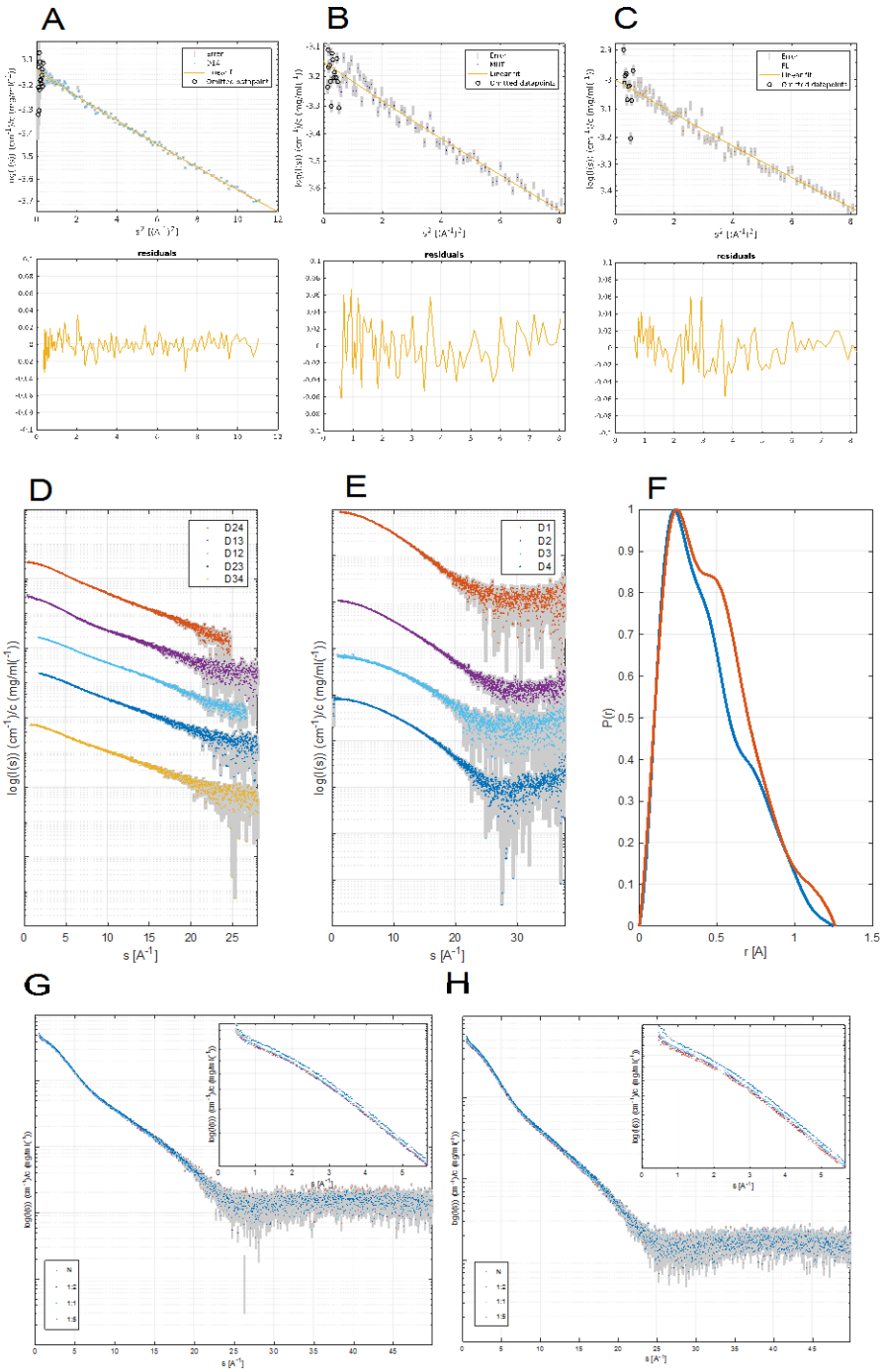

A
B
C
residuals
residuals
residuals
D
D24
D13
D12
D23
D34
log(I(s)) (cm⁻¹)/c (mg/ml(⁻¹))
s [A⁻¹]
E
D1
D2
D3
D4
log(I(s)) (cm⁻¹)/c (mg/ml(⁻¹))
s [A⁻¹]
F
P(r)
r [A]
G
H

Figure S2. Supplementary SAXS data. Related to figure 2-5 and 7. (A,B,C) Guinier plots of Log(s) vs s^2 in the low angle region for the three four domain constructs. (A) D14. (B) MUT, and (C) FL. The residuals from the linear fit are shown below each plot. (D,E) SAXS profiles of single, double and triple domains shown on a log(I(s)) vs s normalized by concentration scale. The profiles are offset for clarity. (D) Double and triple domain constructs: D12, D23, D34, D13, and D24. (E) Single domain constructs: D1, D2, D3, and D4. (F) P(r) profile for triple domain constructs. There is a clear difference in shape of the *P(r)* between D13 and D24. D24 has a slower decay indicative of an extended protein. D13 has a second peak, indicating regularity in inter-domain distances. (G,H) SAXS profile for titration studies of D14 with inset graph showing the low-s region. (G) D14 alone and with PEPT2-CT at three different protein to peptide ratios. PEPT2-CT binds to D14 without inducing structural changes. (H) D14 alone and with NHE3-CT at three different protein to peptide ratios. Similarly to PEPT2-CT, NHE3-CT binds without changing the overall shape of D14. (N – no peptide, 1:2 two-fold excess peptide, 1:1 equal ratio of peptide and protein, 1:5, five-fold excess of peptide)

Table S2. SAXS parameters for the individual PDZK1 domains. Related to figure 1-3.

	D1	**D2**	**D3**	**D4**
Structural Parameters				
I(0) (cm^{-1}) [from P(r)]	0.01	0.01	0.007	0.008
R_g (Å) [from P(r)]	19	19	16	17
I(0) (cm^{-1}) [from Guinier]	0.009	0.01	0.007	0.008
R_g (Å) [from Guinier]	18.7	18.5	15	16.5
d_{max} (Å)	74	72	70	65
Dry volume calculated from seq. ($Å^3$)	18205	14914	11508	11850
Guinier region	52-321	38-322	36-403	37-360
Molecular-mass determination				
Partial specific volume (cm^3g^{-1})	0.729	0.731	0.728	0.74
Contrast	2.969	2.922	2.981	2.81
Molecular mass from (from *I(0)*) (kDA)	12.4	13.8	9.7	11
Molecular mass from Standard P. (kDA)	13	15	10.4	11.9
Calculated monomeric Mr from sequence (kDA)	15.0	12.3	12.0	9.6
Sample Details				
Organism	*Homo sapiens*			
Source	*E. coli;* BL21(DE3)			
Injection volume (microliters)	30			
Concentration range (mg/mL)	1.15-9.26	1.5-12.26	0.5-4.46	0.8-11.8
Solvent (Buffer)	20mM HEPES, 150 Mm NaCl, 0.5 mM TCEP			

Figure S3. The ConSurf platform was used to calculate the sequence conservation and solvent accessibility for PDZK1-FL (Berezin *et al.*, 2004). Related to figure 1. Note that the residue numbers are all shifted by +23 compared to Figure 1 because of the N-terminal affinity tag. e - exposed residue according to the neural-network algorithm. b – buried residue according to neural network algorithm. f – predicted functional residue (highly conserved and exposed). s – predicted structural residue (highly conserved and buried). The ConSurf results were generated with default settings. The linker regions are sparsely conserved, however, there are some buried residues, most notably in the beginning of the linker between D3-D4 (349-358). The conservation score for the residues are overall higher in the tail region (484-542) compared to the linker. In particular residues 483-496 and 529-540. Many of these aforementioned residues are predicted to be potentially of functional interest and a smaller number are also predicted to

be structurally important. Taken together, this could suggest that the tail region may adapt a structured conformation.

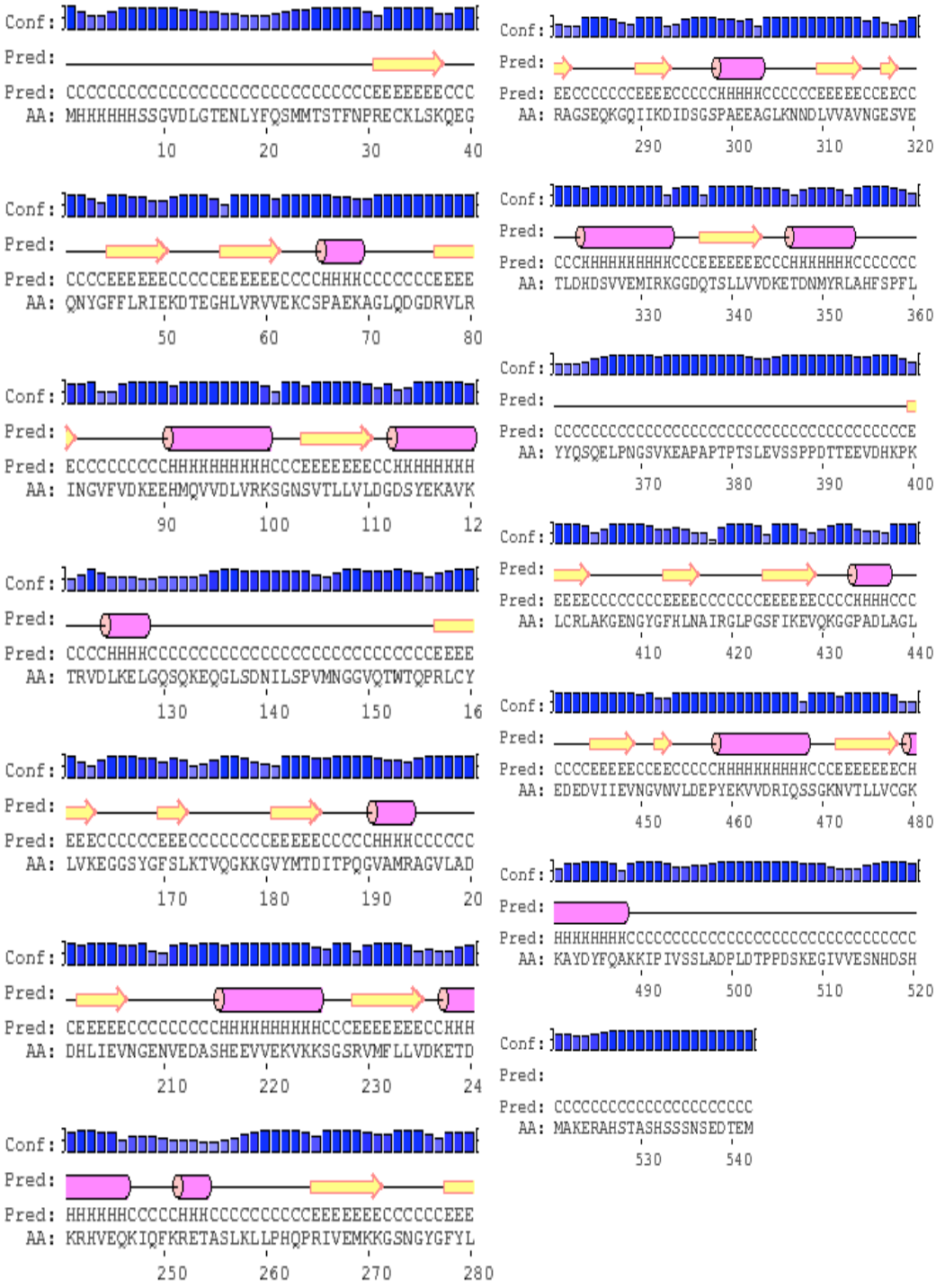


Figure S4. PSIPRED platform (http://bioinf.cs.ucl.ac.uk/psipred/) was used for secondary structure prediction (Buchan *et al.*, 2013) of PDZK1-FL. Related to figure 1. The linker regions are predicted to be mostly coils (non-secondary structure). In contrast, there is a high propensity for a helical structure for the residues DNMYRLAH (348:356), which overlap with the mutated residues

in MUT (351:359). Residue numbers are all shifted by +23 compared to Figure 1 because of the N-terminal affinity tag.

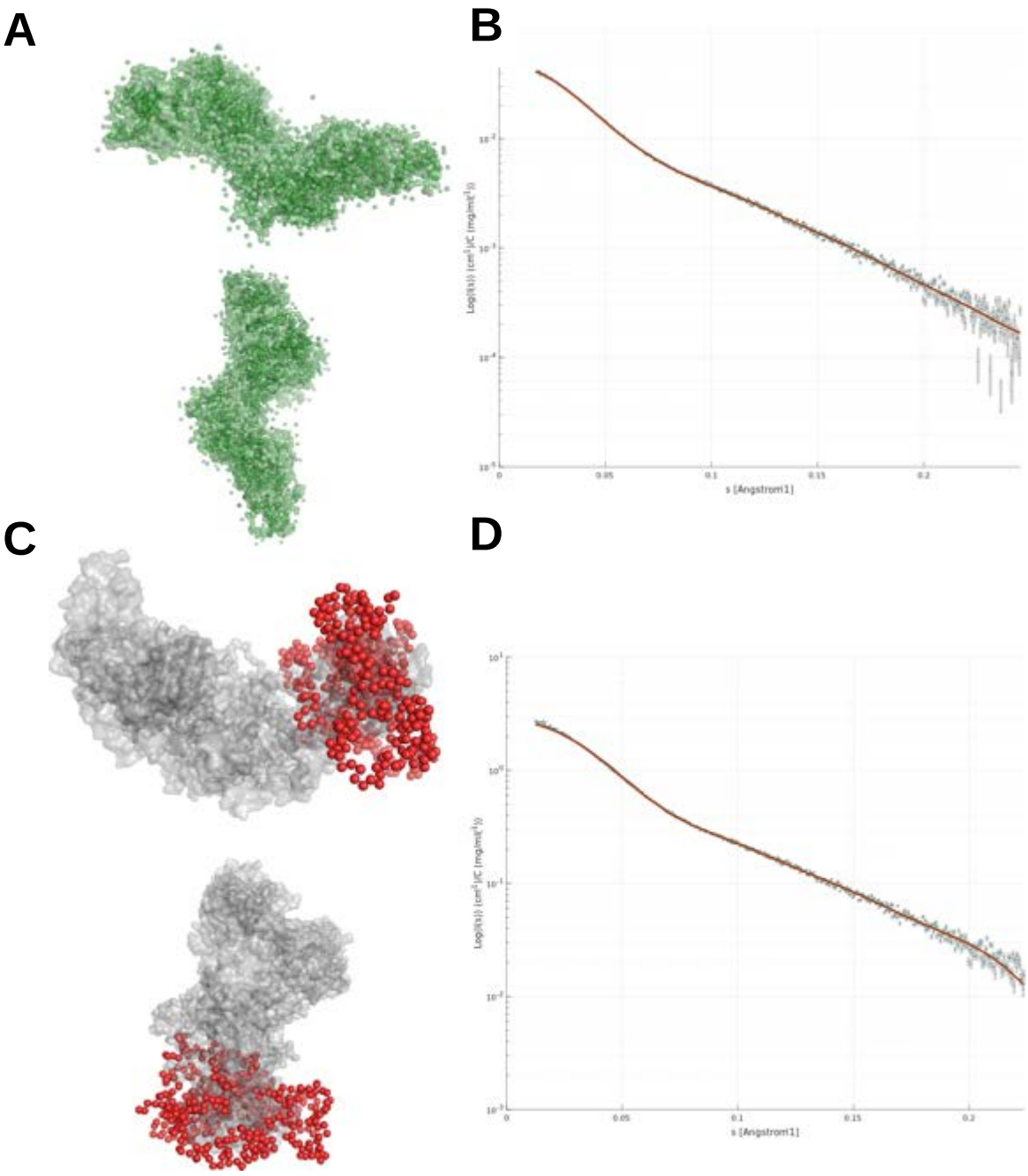


Figure S5. Ab-initio and Rigid body model of PDZK1-FL protein. Related to figure 2 and 3. (A) Nine superimposed GASBOR Ab-initio models (B) Representative fit of a GASBOR model to the SAXS profile, χ^2=0.51. (C) 10 superimposed BUNCH models. Tail region shown in red. (D) Representative fit from one run, χ^2 =0.67.

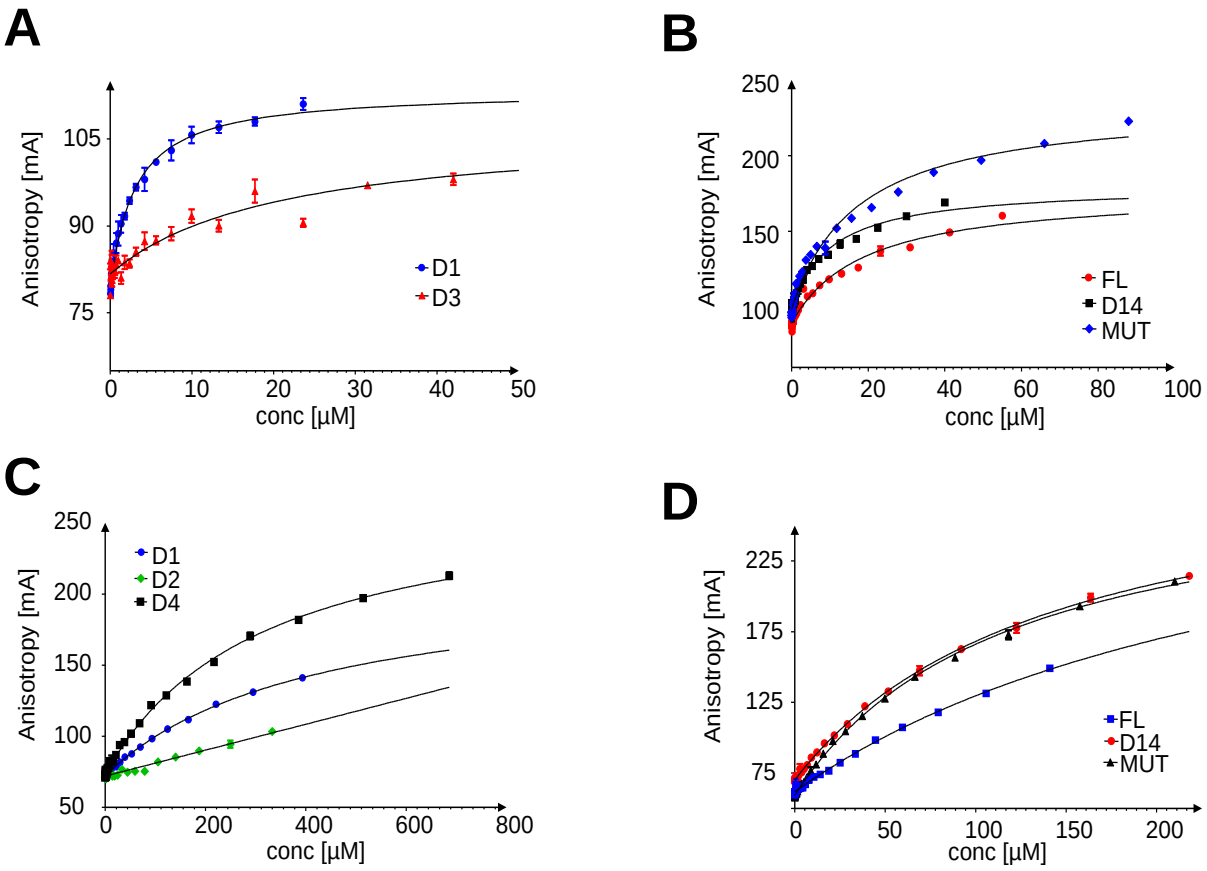


Figure S6. Binding of PDZK1 domains to NHE3-CT-FITC peptide and PDZK1-CT-FITC peptide as monitored by fluorescence polarization. Related to figure 7. The graphs represent fluorescence anisotropy as a function of protein concentration**.** (A) Binding of PDZK1 single domains to the NHE3-CT-FITC peptide. Data points are shown in blue for D1 and in red for D3. (B) Binding of PDZK1 multiple domain constructs to the NHE3-CT-FITC peptide. Data points are marked in red for FL, black for D14, and blue for MUT. (C) Binding of PDZK1 single domains to the PDZK1-CT-FITC peptide. Data points are marked in blue for D1, green for D2, and black for D4. (D) Binding of PDZK1 multiple domains to the PDZK1-CT-FITC peptide. Data points are marked with a red square for FL, red dot for D14, and black triangle for MUT. All data points were recorded in triplicates and error bars indicate the standard deviation of these measurements. The K_D values calculated from three replicate experiments are listed in Table S1.

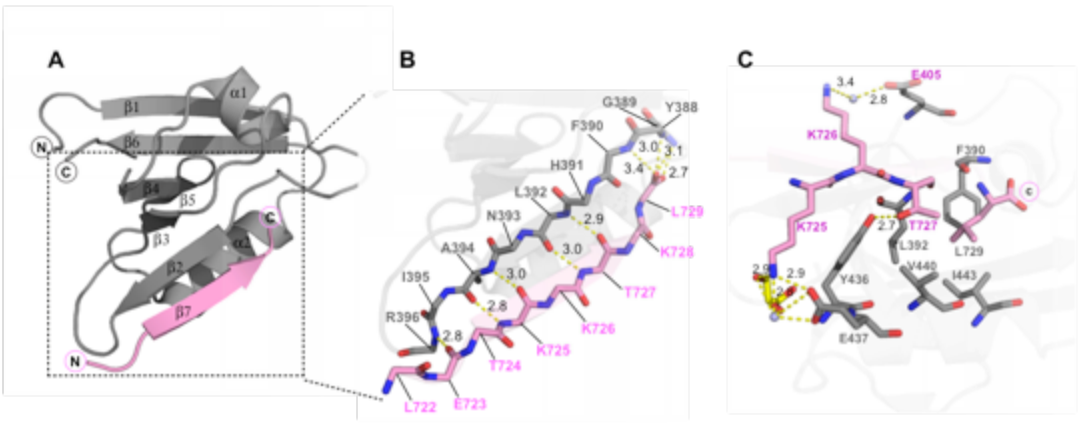


Figure S7. X-ray structure of D4 in complex with the C-terminal peptide of PEPT2. Related to figure 1 and 7. D4 has the typical PDZ domain fold (Doyle *et al.*, 1996; Von Ossowski *et al.*, 2006; Kocher *et al.*, 2010) with a six-stranded antiparallel β-sheet (β1 to β6) flanked by two α-helices (α1 and α2) similar to the mouse homologue (4R2Z.pdb). (A) Cartoon model of the derived complex showing the overall structure of D4 in grey and the C-terminal peptide of PEPT2 in pink. D4 binds PEPT2-CT in a similar fashion as described for other PDZ domains bound to target peptides (Karthikeyan, Leung and Ladias, 2001; Kocher *et al.*, 2011). Binding occurs via beta-sheet complementation, and the peptide is stabilized by an intensive network of hydrogen bonds (partly mediated by water molecules) and hydrophobic interactions, which extends the β-sheet of D4 and forms an additional β-strand (β7) that runs antiparallel to β2. N- and C-termini are labeled. (B) Zoom-in of the dashed box in (A) illustrating main chain interactions (dashed yellow lines) between D4 and the peptide. The residues involved in the formation of the β-sheet (sticks) and the length of the bonds are indicated. Oxygen and nitrogen atoms are colored in red and blue, respectively. For clarity, the side chains have been removed and the rest of the protein is shown as a transparent cartoon model. (C) Close up view of side chain interactions between D4 and PEPT2. Residues (sticks) are labeled, water molecules are colored in cyan and a glycerol molecule that was observed in

the electron density is colored in yellow (carbon atoms). Oxygen and nitrogen atoms are colored as in (B). In particular, Leu^{729} and Thr^{727} are the key residues in PEPT2 for establishing an interaction with D4.

Table S3. Crystallographic table for D4 in complex with PEPT2-CT. Related to figure 1 and 7.

Wavelength (Å)	1.07166
Resolution range (Å)	47.11 - 1.50 (1.55 - 1.50)
Space group	P 65 2 2
Unit cell dimensions	
a, b, c (Å)	54.4, 54.4, 150.76
α, β, γ (°)	90, 90, 120
Total reflections	764851 (73014)
Unique reflections	21857 (2107)
Multiplicity	35.0 (34.70)
Completeness (%)	1.00 (0.99)
Mean I/σ(I)	33.85 (3.96)
Wilson B-factor ($Å^2$)	18.86
R-merge	0.06 (0.96)
R-meas	0.06 (0.97)
CC1/2	1 (0.62)
CC*	1 (0.87)
Reflections used in refinement	21854 (2107)
Reflections used for R-free	1093 (105)
R-work	0.19 (0.38)
R-free	0.21 (0.40)
CC(work)	0.97 (0.75)
CC(free)	0.96 (0.68)
Number of non-hydrogen atoms	917
Macromolecules	787
Ligands	6
Protein residues	94
RMS (bonds) (Å)	0.007
RMS (angles) (°)	0.84
Ramachandran favored (%)	100
Ramachandran allowed (%)	0

Ramachandran outliers (%)	0
Rotamer outliers (%)	1.2
Clashscore	3.09
B-factors ($Å^2$)	
Average	36.68
Macromolecules	34.72
Ligands	43.53
Solvent	48.80
Number of TLS groups	5

Table S3 legend. Crystallographic parameters Table. Data collection and refinement statistics for the crystal structure of D4 in complex with C-terminal peptide PEPT2 are shown. Statistics for the highest-resolution shell are shown in parentheses. The final model was refined at 1.5Å resolution with R_{work} and R_{free} of 0.19 and 0.21, respectively**.** Residues 457 to 460 in addition to the side chains of residues 374, 375 and 456 from D4, and residues 720-721 in addition to the side chains of residues 722-723 from PEPT2-CT could not be modeled due to missing electron density.