Supplement: Conservation of specificity in two low-specificity proteins

Lucas C. Wheeler,[1,2] Jeremy A. Anderson[1,2], Anneliese J. Morrison[1,2], Caitlyn E. Wong[1,2], Michael J. Harms[1,2,*]

1.Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR, USA

2. Institute of Molecular Biology, University of Oregon, Eugene, OR, USA

* harms@uoregon.edu

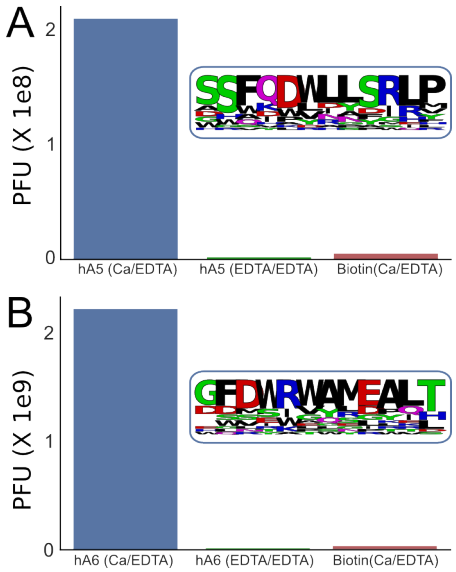

Fig S1. Randomer phage enrichment is dependent on Ca^{2+} and protein. Bar graphs show the plaque forming units (PFU) for phage solutions after the third round of enrichment for screens using hA5 (A) or hA6 (B). For each round of panning, we incubated phage with biotinylated protein, pulled down bound phage via a streptavidin plate, and finally eluted the phage from the protein with an elution buffer. To verify that binding occurred in a Ca^{2+}-dependent manner, we compared Ca^{+}-loading/$EDTA$-elution to $EDTA$-loading/$EDTA$-elution. We also performed a Ca^{2+}-loading/$EDTA$-elution experiment using biotin alone. Insets show sequence logos (WebLogo) generated from 20 plaque sequences from each $Ca^{2+}/EDTA$ panning experiment. The most frequent residue at each position was used to generate the A5cons and A6cons peptides.

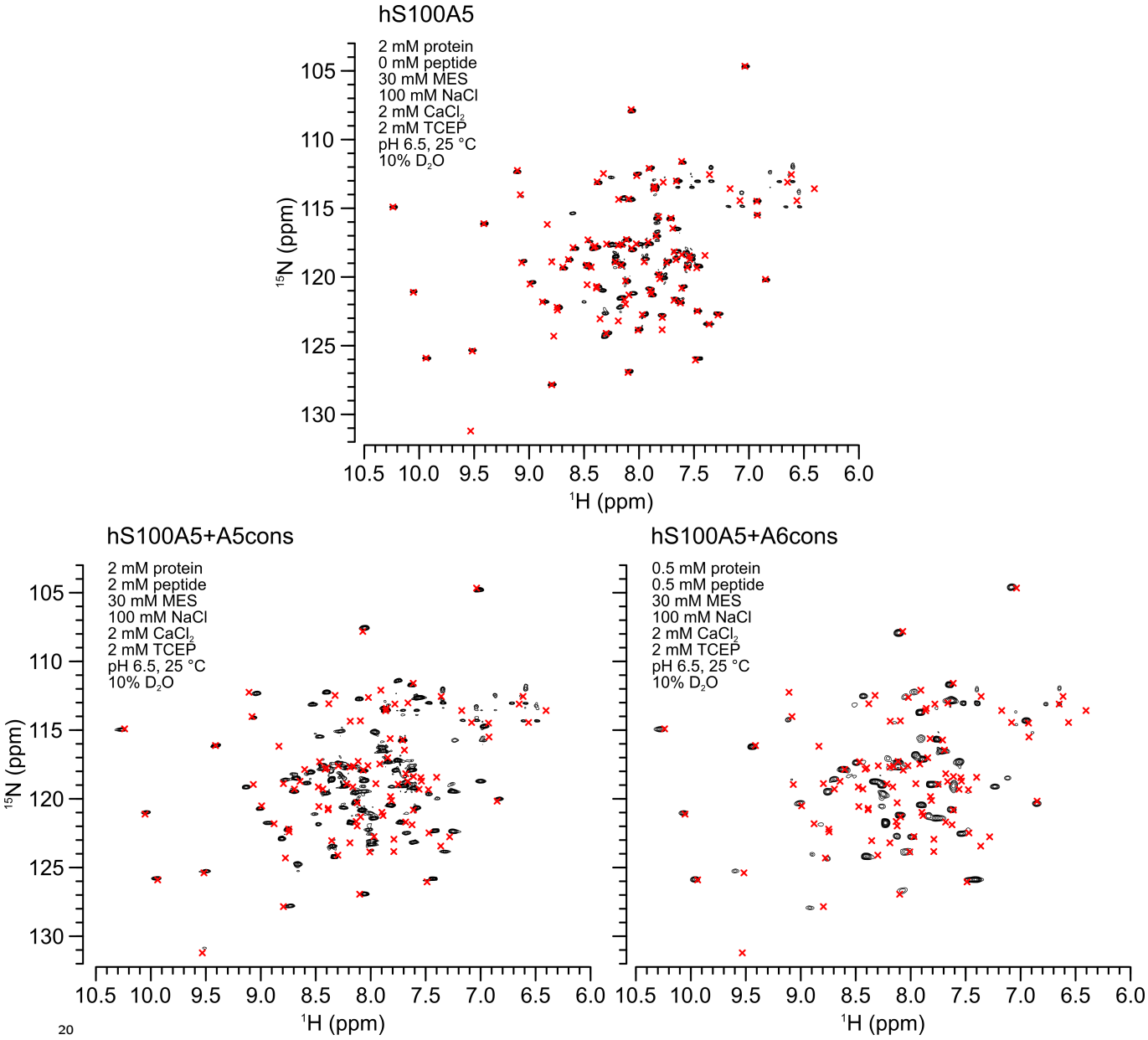

Fig S2. Changes in HSQC spectra of hA5 with peptide binding. 2D $^1H-^{15}N$ TROSY-HSQC NMR spectra for hA5 alone or in the presence of saturating A5cons or A6cons. Buffer conditions are on each plot. Assignments (red ×) are from (1).

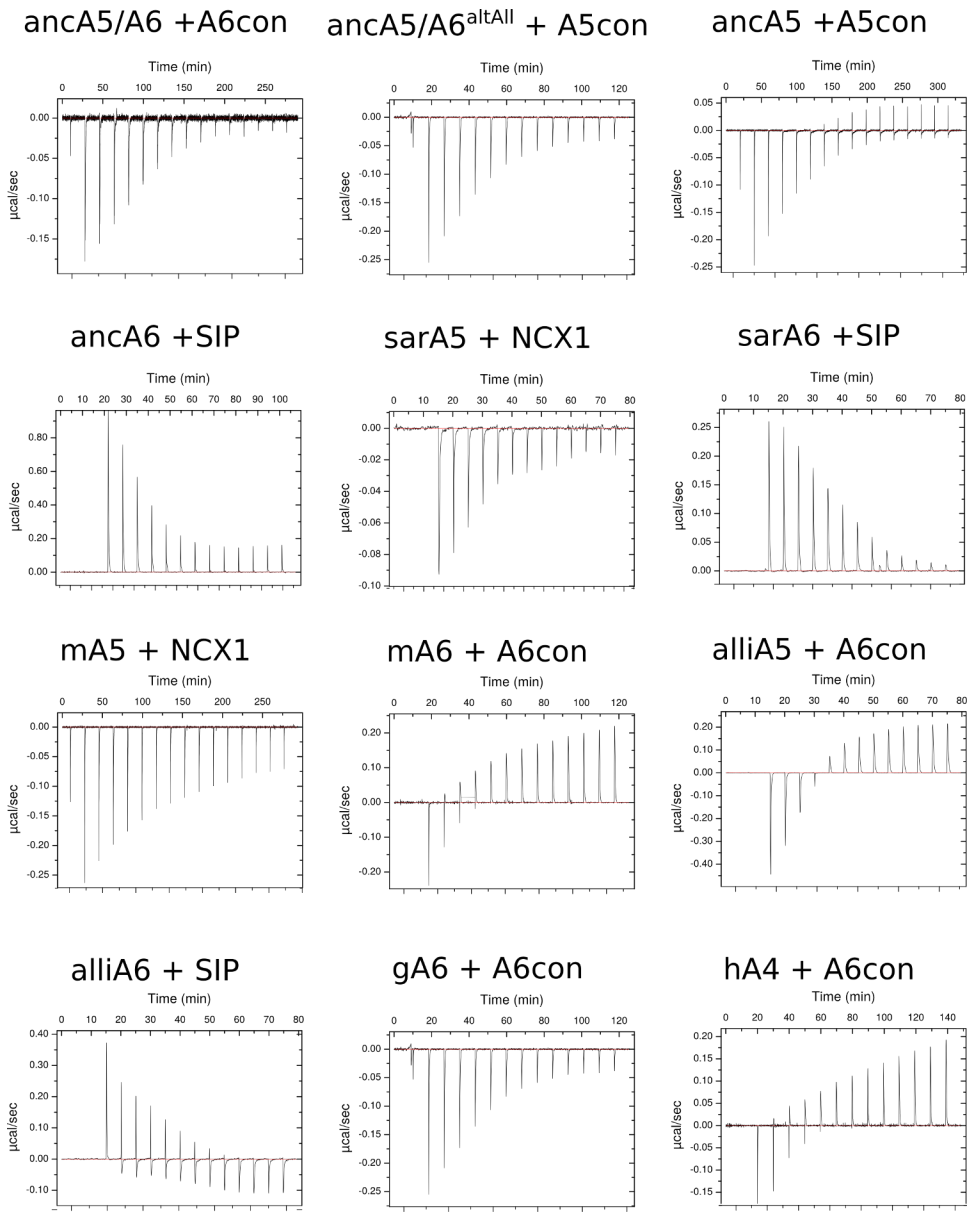

Fig S3. Representative ITC data traces for each ancestor and ortholog. ITC traces show baseline-corrected titration of various peptides onto S100 proteins in the presence of 2 mM Ca^{2+}. All experiments were done with ≈ 100 μM protein in 25 mM TES, 100 mM NaCl, 1 mM $TCEP$ at pH 7.4, 25 °C.

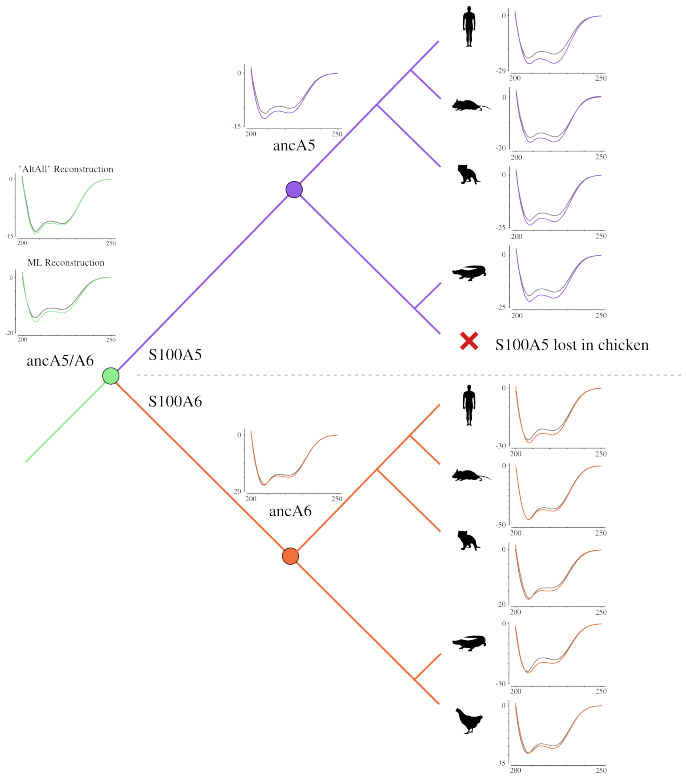

Fig S4. Far UV CD spectra are diagnostic for the S100A5 and S100A6 clades. CD spectra are mapped onto a diagram of the S100A5-S100A6 clade. Curves are spectra of apo (gray) and Ca^{2+}–bound (orange/purple) proteins. The S100A5 proteins (purple) are characterized by a deep alpha-helical signal at 222nm that substantially increases in response to binding of Ca^{2+}. S100A6 proteins (orange) show comparatively minimal response and maintain a deeper peak at 208nm. These patterns hold for the ancestors at the base of each clade. The spectra of ancA5/A6 and the ancA5/A6 altAll version (both shown in green) resemble that of an extant S100A6, indicating that the large Ca^{2+}-driven conformational change seen in the extant S100A5s is a derived feature of this lineage.

Table S1. Accession numbers of S100 proteins used to build the multiple sequence alignment.

paralog	accession	species
A1	F1R758	*Danio rerio*
A1	A5WW32	*Danio rerio*
A1	H2TQM5	*Takifugu rubripes*
A1	H2ST19	*Takifugu rubripes*
A1	H2L492	*Oryzias latipes*
A1	H2M1B8	*Oryzias latipes*
A1	G3NKS0	*Gasterosteus aculeatus*
A1	G3PEI0	*Gasterosteus aculeatus*
A2	P29034	*Homo sapiens*
A2	F6Q7Q8	*Ornithorhynchus anatinus*
A2	P10462	*Bos taurus*
A2	G3W672	*Sarcophilus harrisii*
A2	JH205580.1	*Pelodiscus sinensis*
A3	P33764	*Homo sapiens*
A3	P62818	*Mus musculus*
A3	A4FUH7	*Bos taurus*
A3	G3W5T7	*Sarcophilus harrisii*
A3	F6SL13	*Monodelphis domestica*
A3	F6Q7S6	*Ornithorhynchus anatinus*
A3	JH205580.1	*Pelodiscus sinensis*
A4	P35466	*Bos saurus*
A4	predicted*	*Crocodylus porosus*
A4	P26447	*Homo sapiens*
A4	H0Z1G5	*Taeniopygia guttata*
A4	P07091	*Mus musculus*
A4	F6SKU1	*Monodelphis domestica*
A4	F6Q7T6	*Ornithorhynchus anatinus*
A4	XP_015743713.1	*Python bivittatus*
A4	JH205580.1	*Pelodiscus sinensis*
A4	G3W5H2	*Sarcophilus harrisii*
A4	H9H0S2	*Meleagris gallopavo*
A5	P33763	*Homo sapiens*
A5	P63084	*Mus musculus*
A5	E1B8S0	*Bos taurus*
A5	G3W581	*Sarcophilus harrisii*
A5	XP_019412310.1	*Crocodylus porosus*
A5	JH205580.1	*Pelodiscus sinensis*
A6	P06703	*Homo sapiens*
A6	P14069	*Mus musculus*
A6	F6SKR4	*Monodelphis domesitica*
A6	F6R394	*Ornithorhynchus anatinus*
A6	G3W4S8	*Sarcophilus harrisii*
A6	H9H0S3	*Meleagris gallopavo*
A6	XP_019412316.1	*Crocodylus porosus*
A6	Q98953	*Gallus gallus*
A6	EOB07085.1	*Anas platyrhynchos*
A6	XP_015284753.1	*Gekko japonicus*
A6	XP_007429160.1	*Python bivittatus*
A6	JH205580.1	*Pelodiscus sinensis*

* Unannotated BLAST hit from the Crocodylus porosus genome (GB2012 version). **Table S2.**

Binding of 12-mer phage display peptides does not depend on solubilizing flanks. List of phage display consensus peptides used in the study. The sequences of flank variants of A5cons and A6cons are shown. Flanks are indicated by lower-case letters. The third column shows dissociation constants for peptides binding to hA5 with 95% credibility regions from Bayesian fits of one ITC dataset per variant. Flank variants bind with similar K_D.

Peptide Name	Amino Acid Sequence	$K_D(\mu M)$
A5cons (variant 1)	`rshsSSFQDWLLSRLPgggsae`	$4.9 \leq 6.1 \leq 7.8$
A5cons (variant 2)	`----SSFQDWLLSRLP-ggsae`	$1.1 \leq 2.8 \leq 7.9$
A5cons (variant 3)	`rshsSSFQDWLLSRLP------`	$7.2 \leq 9.6 \leq 13.1$
A6cons (variant 1)	`rshsGFDWRWGMEALTgggsae`	$0.3 \leq 0.9 \leq 2.4$
A6cons (variant 2)	`----GFDWRWGMEALT-ggsae`	$1.5 \leq 2.5 \leq 4.0$

Table S3. Thermodynamic parameters for binding of the peptide rshs-GFDWRWAMEALTggsae (A6cons) to S100A5 and S100A6 proteins. Species abbreviations are "alli" (alligator), "gal" (chicken), "sar" (tasmanian devil), "m" (mouse), and "h" (human). Fit parameters, with standard deviation from fits, for to the data shown schematically in Fig 4A. Parameters are for a single-site binding model. "NA" indicates that there was no detectable binding. We floated the fraction competent parameter to capture uncertainty in peptide and protein concentration, particularly given the low extinction coefficients of S100A5 and S100A6. If an experiment was done at both 10 and 25 $°C$, the parameters correspond to the 10 $°C$ experiment.

protein	K_A (M^{-1})	$\Delta H°$ $(kcal/mol)$	$fx\ comp.$	$num\ reps$	T $(°C)$
ancA5/A6	$8.30e5 \pm 1.9e5$	-12.20 ± 1.2	0.70 ± 0.02	2	25
altAll	$1.10e5 \pm 5.2e4$	-8.70 ± 0.6	0.90 ± 0.07	2	25
ancA5	$7.70e5 \pm 2.1e5$	-3.30 ± 0.8	0.80 ± 0.07	2	25
alliA5	$4.40e5 \pm 6.8e4$	-10.60 ± 0.7	0.70 ± 0.01	2	25
sarA5	$2.50e5 \pm 1.6e5$	-5.90 ± 2.5	0.90 ± 0.13	2	25
mA5	$2.10e5 \pm 5.4e4$	-11.70 ± 2.7	1.10 ± 0.05	2	25
hA5	$4.10e5 \pm 9.8e4$	-8.50 ± 1.5	1.00 ± 0.02	2	25
ancA6	$2.80e5 \pm 1.9e5$	-6.40 ± 2.7	1.10 ± 0.14	2	25
alliA6	$9.50e4 \pm 4.7e4$	-10.40 ± 3.7	0.60 ± 0.09	2	25
gA6	$4.20e5 \pm 2.0e5$	-8.10 ± 2.0	0.70 ± 0.06	2	25
sarA6	$1.40e5 \pm 6.7e4$	-6.20 ± 1.9	0.80 ± 0.10	2	25
mA6	$2.60e5 \pm 1.2e5$	-6.40 ± 1.5	0.60 ± 0.05	2	25
hA6	$2.00e5 \pm 4.8e4$	9.60 ± 1.4	0.80 ± 0.02	2	25
hA4	$2.80e6 \pm 6.5e6$	-1.80 ± 0.5	0.60 ± 0.04	2	25

Table S4. Thermodynamic parameters for binding of the peptide rshsSS-FQDWLLSRLPgggsae (A5cons) to S100A5 and S100A6 proteins. Species abbreviations are "alli" (alligator), "gal" (chicken), "sar" (tasmanian devil), "m" (mouse), and "h" (human). Fit parameters, with standard deviation from fits, for to the data shown schematically in Fig 4A. Parameters are for a single-site binding model. "NA" indicates that there was no detectable binding. We floated the fraction competent parameter to capture uncertainty in peptide and protein concentration, particularly given the low extinction coefficients of S100A5 and S100A6. If an experiment was done at both 10 and 25 $°C$, the parameters correspond to the 10 $°C$ experiment.

protein	K_A (M^{-1})	$\Delta H°$ $(kcal/mol)$	$fx\ comp.$	$num\ reps$	T $(°C)$
ancA5/A6	$9.30e4 \pm 3.0e4$	-5.20 ± 1.6	1.40 ± 0.07	2	25
altAll	$4.70e4 \pm 2.2e4$	-3.90 ± 1.3	1.30 ± 0.19	2	25
ancA5	$1.30e5 \pm 3.6e4$	-6.90 ± 1.3	0.90 ± 0.05	2	10, 25
alliA5	$2.30e4 \pm 3.8e3$	13.80 ± 2.4	1.10 ± 0.07	2	10, 25
sarA5	$2.10e5 \pm 1.5e5$	-4.80 ± 1.9	0.70 ± 0.1	2	25
mA5	$4.70e4 \pm 1.9e4$	-6.90 ± 2.1	0.60 ± 0.08	2	25
hA5	$3.60e5 \pm 2.1e5$	-5.70 ± 1.7	0.80 ± 0.06	2	25
ancA6	NA	NA	NA	2	10, 25
alliA6	NA	NA	NA	2	25
gA6	NA	NA	NA	2	25
sarA6	NA	NA	NA	2	10, 25
mA6	NA	NA	NA	2	25
hA6	NA	NA	NA	2	25
hA4	$1.70e4 \pm 5.1e3$	-4.10 ± 0.8	0.90 ± 0.3	2	25

Table S5. Thermodynamic parameters for binding of the peptide RRLL-FYKYVYKR (NCX1) to S100A5 and S100A6 proteins. Species abbreviations are "alli" (alligator), "gal" (chicken), "sar" (tasmanian devil), "m" (mouse), and "h" (human). Fit parameters, with standard deviation from fits, for to the data shown schematically in Fig 4A. Parameters are for a single-site binding model. "NA" indicates that there was no detectable binding. We floated the fraction competent parameter to capture uncertainty in peptide and protein concentration, particularly given the low extinction coefficients of S100A5 and S100A6. If an experiment was done at both 10 and 25 $°C$, the parameters correspond to the 10 $°C$ experiment. (*) Data from ancA5 binding to NCX1 were difficult to fit. The binding curves for this interaction had shallow curvature and did not appear to reach baseline saturation even in with higher titrant/titrate molar ratio, leading to the high fraction competent.

protein	K_A (M^{-1})	$\Delta H°$ $(kcal/mol)$	*fx comp.*	*num reps*	T $(°C)$
ancA5/A6	$3.3e4 \pm 7.6e3$	-1.70 ± 0.3	0.60 ± 0.05	2	25
altAll	$2.3e4 \pm 8.3e3$	-3.80 ± 1.2	0.60 ± 0.05	2	25
ancA5*	$1.98e5 \pm 1.7e5$	-0.68 ± 0.4	2.90 ± 0.20	2	10
alliA5	$5.80e3 \pm 1.6e3$	-7.20 ± 2.0	0.90 ± 0.26	2	10, 25
sarA5	$2.50e4 \pm 1.7e4$	-2.80 ± 1.3	0.70 ± 0.20	2	25
mA5	$1.20e5 \pm 1.7e5$	-1.30 ± 0.5	0.80 ± 0.20	2	25
hA5	$5.50e4 \pm 1.3e4$	-3.60 ± 0.9	1.40 ± 0.10	2	25
ancA6	NA	NA	NA	2	10, 25
alliA6	$4.60e4 \pm 3.3e4$	-2.50 ± 0.2	0.70 ± 0.15	2	10, 25
gA6	$1.10e5 \pm 1.7e4$	3.40 ± 0.6	1.70 ± 0.05	2	25
sarA6	$1.30e4 \pm 5.8e3$	-4.30 ± 1.8	0.90 ± 0.30	2	25
mA6	NA	NA	NA	2	25
hA6	NA	NA	NA	2	25
hA4	NA	NA	NA	2	25

Table S6. Thermodynamic parameters for binding of the peptide SEGLMNVLKKIYEDG (SIP) to S100A5 and S100A6 proteins. Species abbreviations are "alli" (alligator), "gal" (chicken), "sar" (tasmanian devil), "m" (mouse), and "h" (human). Fit parameters, with standard deviation from fits, for to the data shown schematically in Fig 4A. Parameters are for a single-site binding model. "NA" indicates that there was no detectable binding. We floated the fraction competent parameter to capture uncertainty in peptide and protein concentration, particularly given the low extinction coefficients of S100A5 and S100A6. If an experiment was done at both 10 and 25 $°C$, the parameters correspond to the 10 $°C$ experiment.

protein	K_A (M^{-1})	$\Delta H°$ $(kcal/mol)$	$fx\ comp.$	$num\ reps$	T $(°C)$
ancA5/A6	$1.30e4 \pm 1.7e3$	-8.10 ± 0.9	1.50 ± 0.01	2	25
altAll	$2.40e4 \pm 1.2e4$	-1.50 ± 0.5	1.30 ± 0.20	2	25
ancA5	NA	NA	NA	2	25
alliA5	NA	NA	NA	2	10, 25
sarA5	NA	NA	NA	2	25
mA5	NA	NA	NA	2	25
hA5	NA	NA	NA	2	25
ancA6	$3.90e4 \pm 3.0e2$	3.80 ± 0.3	1.20 ± 0.02	2	10, 25
alliA6	$3.00e4 \pm 9.3e3$	3.20 ± 0.7	1.40 ± 0.09	2	25
gA6	$5.80e4 \pm 8.9e3$	4.00 ± 0.4	1.90 ± 0.04	2	25
sarA6	$3.30e5 \pm 1.5e5$	0.90 ± 0.1	2.00 ± 0.02	2	25
mA6	$3.90e5 \pm 2.8e5$	0.50 ± 0.3	1.80 ± 0.02	2	10, 25
hA6	$3.80e4 \pm 5.7e3$	2.90 ± 0.3	1.50 ± 0.03	2	15, 25
hA4	NA	NA	NA	2	25

Table S7. Thermodynamic parameters for binding of the A5cons and SIP peptides to hA5 ancestral reversion mutants. Table entries show 95% credibility region from the posterior distribution of each parameter. Parameters are for a single-site binding model. "NA" parameters indicate that there was no detectable binding. All experiments were done at 25 $^{\circ}C$.

protein	peptide	K_A ($\times 10^5$ M^{-1})	ΔH° ($kcal/mol$)	fx *comp.*
hA5	A5cons	$1.3 \leq 3.6 \leq 8.9$	$-9.5 \leq -5.7 \leq -3.0$	$0.71 \leq 0.84 \leq 0.97$
hA5	SIP	NA	NA	NA
hA5/E2a	A5cons	$1.2 \leq 1.3 \leq 1.4$	$-5.2 \leq -5.1 \leq -4.88$	$0.87 \leq 0.90 \leq 0.93$
hA5/E2a	SIP	NA	NA	NA
L44i	A5cons	$1.6 \leq 1.7 \leq 1.9$	$-5.2 \leq -5.1 \leq -4.88$	$0.87 \leq 0.90 \leq 0.93$
L44i	SIP	NA	NA	NA
D54k	A5cons	$3.2 \leq 3.5 \leq 3.7$	$-5.1 \leq -5.0 \leq -4.9$	$1.15 \leq 1.17 \leq 1.19$
D54k	SIP	NA	NA	NA
M78a	A5cons	$1.0 \leq 1.1 \leq 1.2$	$-3.5 \leq -3.3 \leq -3.1$	$1.24 \leq 1.28 \leq 1.32$
M78a	SIP	NA	NA	NA
A83m	A5cons	$2.6 \leq 2.8 \leq 3.0$	$-5.5 \leq -5.5 \leq -5.3$	$1.52 \leq 1.53 \leq 1.55$
A83m	SIP	$0.4 \leq 0.6 \leq 1.0$	$-0.8 \leq -0.6 \leq -0.4$	$0.99 \leq 1.22 \leq 1.54$

References

[1] Bertini I, Gupta SD, Hu X, Karavelas T, Luchinat C, Parigi G, et al. Solution Structure and Dynamics of S100A5 in the Apo and Ca2+-Bound States. JBIC Journal of Biological Inorganic Chemistry. 2009 Sep;14(7):1097–1107.