Supporting Information

Calcium Binding to the Innate Immune Protein Human Calprotectin Revealed by Integrated Mass Spectrometry

Jagat Adhikari[1], Jules R. Stephan[2], Don L. Rempel[1], Elizabeth M. Nolan[2]* and Michael L. Gross[1]*
[1]Department of Chemistry, Washington University at St Louis, St Louis, MO 63130, USA
[2]Department of Chemistry, Massachusetts Institute of Technology, Boston, MA 02139, USA
*Corresponding Authors: Elizabeth M. Nolan and Michael L. Gross
Phone: 617-452-2495 (EMN) and 314-935-4814 (MLG); Email: lnolan@mit.edu and mgross@wustl.edu

This Supporting Information section includes:

Supporting Discussion of PLIMSTEX and "Sharp-break" Data

Additional data for the HDX kinetics of CP-Ser performed at 4 μM protein concentration and the plots from unchanged regions of CP-Ser in PLIMSTEX "sharp-break"titrations done at 40 μM protein concentration are provided in Figures S1-S3. These data complement the results in the main text and are presented to represent conformational changes occurring throughout the protein, not only on the Ca^{2+} binding sites.

Supporting Discussion of Mathematical Modeling

Modeling of 4 μM Protein Titration Data (PLIMSTEX): To make clear the modeling and its implementation for this study, we describe its evolution from modeling of PLIMSTEX data taken at 4 μM protein concentration to a final extended model of the 4 μM PLIMSTEX, the 15 μM native-MS and the 40 μM "sharp-break" PLIMSTEX results with a reduced model. The foundation of this modeling strategy in a general context is presented in a recent book chapter.[1] All modeling to which the analyticl Ca^{2+} concentration was added was done assuming there was initially some Ca^{2+} present in the sample, corresponding to 2.4 μM detected in representative buffer/D_2O mixture by ICP-MS.

The modeling involves the nonlinear least squares (NLLS) regression analysis of the 4 μM data obtained by titration, using an approach described previously[1], to test models for Ca^{2+} binding and to obtain binding constants.[1] This approach was designed to address the question of Ca^{2+} binding and its relation to tetramerization. Our working assumptions were that the information supplied by PLIMSTEX pertained to Ca^{2+} binding or to Ca^{2+} binding accompanied by tetramerization. Thus, we chose different mathematical models to fit the PLIMSTEX data only. The first model was a 1:4 sequential binding model for the four Ca^{2+}-binding events per heterodimer with no tetramerization event included (Figure 4 of the main text).

Examples of classical Ca^{2+}-binding proteins such as calmodulin show cooperativity in Ca^{2+}-binding within a domain. To accommodate tetramerization and cooperativity, we introduced a combination of tetramerization into four additional mathematical models (Models A-D) with cooperativity also included in Models C and D. These models include tetramerization occurring upon binding of either the first Ca^{2+} ion (model A) or of the second Ca^{2+} ion (model B). In the third model, tetramer formation occurs after the first Ca^{2+}-binding event followed by cooperative binding of two more Ca^{2+} ions (model C). In the fourth model, tetramerization occurs after binding the first two Ca^{2+} ions followed by cooperative binding of a fifth Ca^{2+} ion to the heterotetramer (model D). The absence of K_3 in model B, of K_2 and K_3 in model C, and K_3 and K_4 in model D are consistent with the assumptions of the models. Each model was used to calculate the Ca^{2+}-binding affinities, and all four models demonstrated good fits as judged by visual inspection and the calculated residual RMS and R^2 values (Table S1 and Figure S4).

Table S1. Fitted association constants from models A and B.[a]

Model	A		B	
	Value (M^{-1})	Model Term	Value (M^{-1})	Model Term
K_1	6.2 x 10^5	$K_1[Ca^{2+}][CP]$	3.1 x 10^5	$K_1[Ca^{2+}][CP]$
K_2	4.5 x 10^4	$K_2K_1^2[Ca^{2+}]^2[CP]^2$	6.0 x 10^5	$K_2K_1[Ca^{2+}]^2[CP]$
K_3	5.2 x 10^5	$K_3K_2K_1^2[Ca^{2+}]^3[CP]^2$	-	
K_4	5.6 x 10^5	$K_4K_3K_2K_1^2[Ca^{2+}]^4[CP]^2$	1.1 x 10^5	$K_4K_2^2K_1^2[Ca^{2+}]^4[CP]^2$
K_5	2.8 x 10^4	$K_5K_4K_3K_2K_1^2[Ca^{2+}]^5[CP]^2$	1.9 x 10^5	$K_5K_4K_2^2K_1^2[Ca^{2+}]^5[CP]^2$
K_6	1.9 x 10^4	$K_6K_5K_4K_3K_2K_1^2[Ca^{2+}]^6[CP]^2$	1.2 x 10^4	$K_6K_5K_4K_2^2K_1^2[Ca^{2+}]^6[CP]^2$
K_7	1.6 x 10^5	$K_7K_6K_5K_4K_3K_2K_1^2[Ca^{2+}]^7[CP]^2$	1.7 x 10^4	$K_7K_6K_5K_4K_2^2K_1^2[Ca^{2+}]^7[CP]^2$
K_8	2.4 x 10^5	$K_8K_7K_6K_5K_4K_3K_2K_1^2[Ca^{2+}]^8[CP]^2$	1.0 x 10^2	$K_8K_7K_6K_5K_4K_2^2K_1^2[Ca^{2+}]^8[CP]^2$
resRMS	0.98		1.32	
R^2	0.997		0.995	

[a] The association constants were determined by fitting each model to the data presented in Figure S4. They are for Ca^{2+} binding except that K_2 for Model A and K_4 for Model B are tetramerization constants.

Table S1 (cont.). Fitted association constants for models C and D.[b]

Model	C		D	
	Value (M^{-1})	Model Term	Value (M^{-1})	Model Term
K_1	7.2 x 10^5	$K_1[Ca^{2+}][CP]$	1.4 x 10^9	$K_1[Ca^{2+}][CP]$
K_2	-		1.6 x 10^6	$K_2K_1[Ca^{2+}]^2[CP]$
K_3	-		-	
K_4	5.5 x 10^{17}	$K_4K_1^2[Ca^{2+}]^4[CP]^2$	-	
K_5	1.3 x 10^6	$K_5K_4K_1^2[Ca^{2+}]^5[CP]^2$	3.3 x 10^{12}	$K_5K_2^2K_1^2[Ca^{2+}]^5[CP]^2$
K_6	2.0 x 10^4	$K_6K_5K_4K_1^2[Ca^{2+}]^6[CP]^2$	6.9 x 10^6	$K_6K_5K_2^2K_1^2[Ca^{2+}]^6[CP]^2$
K_7	2.0 x 10^4	$K_7K_6K_5K_4K_1^2[Ca^{2+}]^7[CP]^2$	1.1 x 10^6	$K_7K_6K_5K_2^2K_1^2[Ca^{2+}]^7[CP]^2$
K_8	8.4 x 10^2	$K_8K_7K_6K_5K_4K_1^2[Ca^{2+}]^8[CP]^2$	6.6 x 10^4	$K_8K_7K_6K_5K_2^2K_1^2[Ca^{2+}]^8[CP]^2$
resRMS	1.09		0.72	
R^2	0.996		0.998	

[b] The association constants were determined by fitting each model to the data presented in Figure S4. They are for Ca^{2+} binding except that K_4 (M^{-3}) for Model C is the tetramerization and two-Ca^{2+} "cooperative" binding constant and K_5 (M^{-2}) for Model D is the tetramerization and one-Ca^{2+} "cooperative" binding constant.

Despite the substantial effort to model cooperativity described above, we found no significant improvement in the fits to the PLIMSTEX data compared to the ones obtained without cooperativity, and this outcome, as evidenced by the statistical parameters for different models (Table S1), shows no strong preference for any of the models. For human calprotectin (hCP), self-association of heterodimers to form the heterotetramer in the presence of Ca^{2+} occurs, indicating that this self-association event must be considered in the modeling.

We then conducted native-spray MS experiments to find evidence to support models A and B and the inclusion of cooperative binding of Ca^{2+} to the heterodimer (models C and D). The approach was to examine the pattern of Ca^{2+} adducts to the heterodimer in the native MS. For example, if model A were correct, the native MS would show a heterodimer with no binding of a second, third, or fourth Ca^{2+} ion because the next step is tetramerization. The tetramer would then be seen with binding to at least two Ca^{2+} ions. If model C were correct, the native MS would show also show the absence of a second, third, and fourth Ca^{2+} ion (as in model A) but the tetramer would have at least four Ca^{2+} ions. The native MS data did not show this or related behavior consistent with the four models that incorporate cooperativity. At this point in our efforts, we removed cooperativity from the modeling and looked for a means to incorporate additional data, particularly at protein concentrations higher than 4 μM at which the PLIMSTEX data were gathered.

Expanding the Data for Modeling

The outcomes of the four models discussed above indicated that we need to expand our data by working at higher protein concentration. To address this need, we (i) modeled the PLIMSTEX Ca^{2+} titration data from the four peptides representing the binding sites (Figure 3, main text) and the titration data from native MS (Figure 6, main text), (ii) performed model reduction by removing terms that do not contribute significantly to the overall outcome (e.g., tetramer with 1, 2, 3, 4 Ca^{2+}), and (iii) expanded the data further by considering not only the PLIMSTEX and the native MS titration data but also the "sharp break" PLIMSTEX data (Figure 5, main text). This design allows us to consider data from three concentration regimes of the protein: 4 μM for PLIMSTEX titration, 15 μM for native MS, and 40 μM for the "sharp-break" PLIMSTEX, thus moving to higher concentrations than for PLIMSTEX and improving the opportunity to look more closely at tetramerization. For our first effort in conducting step (i), we considered whether the native MS data, taken in the gas phase of the mass spectrometer, were compromised by tetramerization occurring upon solvent evaporation in the ESI process and attendant increases in droplet compression and concentration.

1. Probability for one molecule of CP-Ser in nanospray droplet: To use properly the native-MS data, we ruled out tetramerization caused by solvent evaporation from the charged droplets. Proper treatment would require that

there is no more than one heterodimer or one heterotetramer per nanospray droplet in the 15-μM range of protein concentrations that were used in the native MS titration.[2] The following considerations were made:

- Analytical concentration of the CP-Ser heterodimer, C = 15 μM
- Radius of a droplet in a nanospray, r = 10 nM
- Volume of a droplet in nanospray, $V = \frac{4\pi r^3}{3} = 4.19 * 10^{-21}$ L
- Amount of CP-Ser in a droplet, m = concentration(C)* Volume (V) = $6.28 * 10^{-26}$ moles.
- Total number of molecules in a droplet, $N = 3.78 * 10^{-2}$

Therefore, the probability of having one molecule of CP-Ser heterodimer in a 10-nm droplet volume with an analytical concentration of 15 μM of CP-Ser heterodimer is 0.0378, and an upper bound on the fraction of two or more unassociated heterodimers in the droplet being counted in the heterotetramer signal is at most 4%, justifying our premise that the overall signals for heterodimer and heterotetramer in native MS represent those in solution.

The overall count of heterodimer signals to overall count of heterotetramer signals in native MS is representative of speciation of the protein in the solution phase. For each Ca^{2+} concentration investigated in the native-MS titration, the signals were first corrected to account for heterotetramer composition (2 x heterodimer). The signals for each of the heterodimer and the heterotetramer species were then combined and expressed as a fraction of the total protein. For each titration point, we performed this procedure on the dominant charge state of the heterodimer and heterotetramer. The signals for each species were divided by the number of charge states to adjust for the nature of the instrument detection, namely charge displacement as is done for measurements made with an Orbitrap mass spectrometer–see Experimental). Thus, the calculated values from the native-MS experiment can be used to construct PLIMSTEX-like peptide data. The results are included in the modeling as a peptide that can be viewed as a "pseudo peptide". For the native-MS-derived pseudo peptide, the maximum signal, D_o, was set at 100% and the change in extent of HDX signal, ΔD, was set to 100 (i.e., starting at 100% and decreasing to 0%). The processed native-MS data, viewed as for the pseudo peptide, were incorporated with the data from other measurements as discussed below.

2. Modeling PLIMSTEX and native MS data: The 4 μM PLIMSTEX data provide no definitive features that signal the appearance of the first Ca^{2+}-bound heterotetramer. The data trends in Figure S3 for the peptides containing at least parts of the the four EF hands are remarkably the same in character, making it difficult to compare the first and second Ca^{2+} binding events in each heterodimer and to consider the tetramerization event if or when it occurs at higher protein concentration.

The first expanded model (Model E) considered the 4 μM PLIMSTEX data and the results from the native MS but not the "sharp-break" PLIMSTEX data. Because there is a lack of evidence for cooperativity (*vide supra*) and the PLIMSTEX data do not represent the whole binding scenario, a symmetry argument was invoked, and the first two Ca^{2+}-binding events were modeled as being independent and having the same affinity in the heterodimer. For the binding of the next two Ca^{2+} ions to the heterodimer, which are expected to occur with relatively low affinities, negative cooperativity is possible, but it is mathematically indistinguishable from independent binding–see reference[3], p 46, discussion following equation 2.41. The construction of the overall Adair constants in the heterodimer modeling uses the reported *K* values as microscopic *K* values.

Specifically, for Model E, we utilized four independent Ca^{2+}-binding sites, two high-affinity and two low-affinity sites in the heterodimer. That treatment yielded nine Adair constants (i.e., a tetramerization constant and eight Ca^{2+}-binding constants (Figure S5 for data and Table S2 for equilibrium constants)) calculated for the heterotetramer. For the Ca^{2+}-bound species with k-Ca^{2+}ions (k represents the number of Ca^{2+} ions bound to CP-Ser), the k-Ca^{2+} heterodimer species maximize when the 2k-Ca^{2+} heterotetramer species do in the model (see Figure S6B for illustration). Because it is difficult to distinguish the contribution of the k-Ca^{2+} heterodimer species from the contribution of the 2k-Ca^{2+} heterotetramer species in the PLIMSTEX data, the ΔD_i (the difference in HDX of the apo and holo protein–see reference 1) for each k-Ca^{2+} and 2k-Ca^{2+} were locked together in the model.

Table S2. Experimental association and Adair constants from model fit (Model E).[a]

Heterodimer Association Constants	Value (M^{-1})
K_1	2.4×10^{6}
K_2	2.4×10^{6}
K_3	6.2×10^{-2}
K_4	6.9×10^{0}
Overall Heterotetramer Adair Constants	**Value (unit)**
β_0	2.7×10^{3} (M^{-1})
β_1	2.6×10^{10} (M^{-2})
β_2	9.0×10^{16} (M^{-3})
β_3	1.5×10^{23} (M^{-4})
β_4	9.4×10^{28} (M^{-5})
β_5	4.2×10^{27} (M^{-6})
β_6	1.2×10^{28} (M^{-7})
β_7	1.5×10^{29} (M^{-8})
β_8	1.9×10^{45} (M^{-9})

[a] The association and Adair constants were determined from the data presented in Figure S5. The heterodimer association constants are for Ca^{2+} binding.

The two high-affinity sites (K_1 and K_2) are fixed to be the same and solve as 2.4×10^{6} M^{-1} (Table S2). The third site (K_3) is of very low affinity and solves as 0.062 M^{-1}. The fourth site (K_4) is of low affinity and solves as 6.9 M^{-1}. In the case of heterotetramer species, dimerization of apo heterodimer, β_0, solves as 2.7×10^{3} M^{-1}. The first four Ca^{2+} species, β_1-β_4, are modeled as independent binding sites with the same high affinity as for the heterodimer as explained above. The fifth Ca^{2+} species is modeled with overall Adair binding constant by itself to allow the possibility that this species can be zero, and the β_5 solves as 4.2×10^{27} M^{-6}. For the 6-8 calcium species, overall Adair binding constants were constructed sequentially; these solve as $\beta_6 = 1.2 \times 10^{28}$ M^{-7}, $\beta_7 = 1.5 \times 10^{29}$ M^{-8}, and $\beta_8 = 1.9 \times 10^{45}$ M^{-9}, respectively (Table S2).

Binding affinities obtained from Model E and the associated species fraction plots revealed high coefficients (large ΔDs, the differences in D uptake between apo and holo) for the low-order species with little contribution to the actual fitting of the data (Table S2 and Figure S6). The species with the largest contributions demonstrate their prominence in the species fraction plot (Figure S6A). The low-order species reveal themselves only when the species plots are adjusted to their maximum (Figure S6B) and are otherwise not noticeable (Figure S6A).

3. Model reduction: Based on these observations from Model E, we introduced several simplifications to the modeling scheme by:

(a) Setting lower bounds on the overall Adair constants and utilizing coupled *ΔDs* in data fitting where the k-Ca^{2+} heterodimer species maximize with the 2k-Ca^{2+} heterotetramer species. For low-abundant species, changes in the species' overall Adair binding constants will have only a small effect on the equilibrium. As a consequence, there is uncertainty in the overall Adair binding constants for these species. This results in uncertainty of the low-abundant species' *ΔD*. The deuteration of each species may be at most 100%. This fact was used to constrain the *ΔDs*, which sets lower bounds on the overall Adair binding constants for low-abundant species.

(b) Incorporating the major contributors (heterodimer 1-4 Ca^{2+} species, and heterotetramer 8 Ca^{2+} species) and leaving out the low-order species (small numbers of Ca^{2+}) for the heterotetramer, which do not contribute significantly to the equilibria. The latter will reduce the number of unknowns in the modeling. Justification is provided by the native MS where no small-order species are found associated with the heterotetramer (Figure 6, main text).

(c) Constraining model choices by understanding that species fractions add up to one. For most of the titration range (up to about 20 equivalents of Ca^{2+} for the native MS), the heterodimer species fractions

dominate (Figures S5 and S6A). The drop in deuteration shown in the PLIMSTEX data (Figure S5) coincides with the drop in the 1-Ca^{2+} heterodimer species as the earliest possibility and occurs at approximately 2 equivalents of Ca^{2+} (Figure S4). The last possibility for the rise in heterotetramer species is formation of the 8-Ca^{2+} heterotetramer species. The rise in this species necessarily coincides with the rise in the 4-Ca^{2+} heterodimer. There is a gap between these two points in the titration that must be filled with a 2-Ca^{2+} heterodimer species or a 3-Ca^{2+} heterodimer species or both. The detailed model described above suggests that the 2-Ca^{2+}species fills this gap (Figure S5).

4. Modeling PLIMSTEX and native-MS data with three reduced models: We then turned to reduced modeling that includes a heterotetramer signal arising with six Ca^{2+} as a possibility. The best reduced model incorporated a 2-Ca^{2+} heterodimer species filling the "gap" with species between heterodimer and heterotetramer, as is illustrated by the species fraction plot for the PLIMSTEX and native-MS data (Figure S7).

Is the 3-Ca^{2+} heterodimer or 4-Ca^{2+} heterodimer competent to form the heterotetramer? Three different models for the heterodimer species with either 4-Ca^{2+} heterodimer or 3-Ca^{2+} heterodimer → heterotetramer model were investigated to determine whether the 3-Ca^{2+} heterodimer or the 4-Ca^{2+} heterodimer is competent to form the heterotetramer. The statistical parameters from these model fits were compared (Table S3, Figure S8). First, we looked at the fits for the exclusively 3-Ca^{2+} heterodimer to heterotetramer model (Model F). The comparison of native spray data to its fitted model shows an overall RMS, "resRMS" of 3.56. Next, we looked at tetramerization starting at the 3-Ca^{2+} heterodimer (Model G). Heterotetramer species at 7 and 8 Ca^{2+} were also included in this model. These species could be from binding of one Ca^{2+} or two Ca^{2+} to the 6-Ca^{2+} heterotetramer. The comparison of native spray data to its fitted model resulted in the "resRMS" of 2.62. The fit is better than the model that postulates that the tetramer arises from the 3-Ca^{2+} heterodimer and ends at 6-Ca^{2+}.

Table S3. Comparison of statistical parameters for dimer-to-tetramer models with 3 or 4 Ca^{2+}: Reduced modeling using protein titration data at 4 and 15 μM.

Model #	4 μM PLIMSTEX RMS	15 μM Native MS RMS	resRMS	Parameters Searched
F[a]	1.81	12.03	3.56	3
G[b]	1.78	7.63	2.61	6
H[c]	1.72	5.31	2.15	4

[a] Independent Dimer 123 Ca^{2+} →Tetramer 6 Ca^{2+}.
[b] Independent Dimer 123 Ca^{2+}, Decouple 1 and 2 → Tetramer 678 Ca^{2+}.
[c] Independent Dimer 1234 Ca^{2+} → Tetramer 8 Ca^{2+}.

Finally, we fit the data to the 4-Ca^{2+} heterodimer to heterotetramer mode (Model H). The comparison of native spray data to its fitted model resulted "resRMS" of 2.15. This model presented the best fit thus far, and the 4-Ca^{2+} heterodimer to 8-Ca^{2+} heterotetramer was used in fitting the PLIMSTEX and native-MS data and gave the binding affinities and tetramerization constant (Figure S9, Table S4) and to generate fractional species plots as a function of Ca^{2+} concentration (Figure S7).

Table S4. Calculated dissociation constants for 4 and 15 μM data (Figure S9) analyzed by using reduced Model H and the data presented in Figure S9.[a]

Dissociation Constants[a]	Value (M)
K_1	7.7×10^{-7}
K_2	7.7×10^{-7}
K_3	2.6×10^{-3}
K_4	1.9×10^{-2}
$K_{tetramer}$	5.6×10^{-12} ($\beta = 2.3 \times 10^{44}$ M^{-9})

[a] K_1-K_4 refer to binding 1-4 Ca^{2+} to the heterodimer; $K_{tetramer}$ is the tetramerization constant.

5. Modeling PLIMSTEX, native MS, and "sharp-break" data with reduced model: Normally, we view "sharp-break" PLIMSTEX data as reporters on stoichiometry rather than restraints on modeling. These data, however, are also generated in a titration. Thus, their use in the modeling with the titrations at 4 μM M (PLIMSTEX) and 15 μM M (native MS) is plausible and can lead to an increase in accuracy.

These results obtained by modeling the three titration data sets (4, 15, and 40 μM) make use of the reduced model. The outcome of the modeling (Models I-K), including the equilibrium constants, are discussd in the main text, and the statistical parameters (Table S5), fractional species plots obtained at the three concentrations (Figure S10), and a comparison of several reduced model fits to native-MS data (Figure S11) are provided here. The resRMS indicates the outcome is slightly worse when the 40 μM data are included in the modeling, and the native-MS RMS has increased probably because the 15 μM data are defined by the fewest number of points.

Table S5. Comparison of statistical parameters for dimer-to-tetramer Models I-J with 3 or 4 Ca^{2+}: Reduced modeling using protein titration data at 4, 15, and 40 μM

Model #	4 μM PLIMSTEX RMS	15 μM Native MS RMS	40 μM Sharp-Break RMS	resRMS	Parameters Searched
I[a]	2.30	10.93	1.43	2.89	3
J[b]	1.80	10.09	2.87	3.02	6
K[c]	1.60	9.26	1.34	2.34	4

[a] Independent Dimer 123Ca^{2+} →Tetramer 6 Ca^{2+}.
[b] Independent Dimer 123 Ca^{2+}, Decouple 1 and 2 →Tetramer 678 Ca^{2+}.
[c] Independent Dimer 1234 Ca^{2+} → Tetramer 8 Ca^{2+}.

6. Modeling PLIMSTEX, native MS, and "sharp-break" data with reviewer suggested equilibrium models: One of the Reviewers suggested using other models, recognizing that the heterodimer and heterotetramer species are coupled. Further, the reviewer suggested using the overall Adair binding constants of the species as the searched parameters. Use of these constants in the following effort removes the redundancy that comes with the use of stepwise macroscopic binding constants and the several pathways that lead to the same product.

The construction of these models, as suggested by the Reviewer, begins with the enumeration of the species in the coupled equilibrium. There are five heterodimer species beginning with the apo heterodimer and including the four species formed by the stepwise binding of four calcium ions. There are five heterotetramer species beginning with the apo heterotetramer followed by those species formed by the stepwise binding of two calcium ions at a time. The Reviewer suggested that the apo heterotetramer and the two-calcium heterotetramer concentrations can be regarded to remain low as indicated by the native-spray MS so their effect on the equilibrium need not be considered. The Reviewer made a similar suggestion for the four-calcium heterodimer species. Taken together, we thus have seven species with six overall Adair binding constants in the description for the coupled equilibrium. Optionally, the Reviewer suggested that a tetramer species with seven calciums bound may be included in the equilibrium model, giving two suggested models.

Our experience with modeling this shows that the model parameter space has two distinguishable regions in which solutions are found. The first region is characterized by low-binding affinities for the first calcium bindings to the heterodimer. A value of 5×10^4 M^{-1} for the first calcium binding event is representative, but we regard this value to be too small. The second region is characterized by values typically greater than 1×10^6 M^{-1}. Solutions in this region are thought to represent better the actual properties of the equilibrium. Because searches that vary the overall Adair binding constants can find a minimum in either region, the searches discussed below were deliberately initiated in the second region. In addition, the overall Adair binding constants for the first search trial were deliberately biased so that all species fractions were clearly and graphically visible, allowing the search either to enhance them or to push them down to optimize the fit. All of the searches are fits to the PLIMSTEX, native MS and "sharp-break" data with the outcomes tabulated in Table S6.

Columns 2 and 3 of Table S6 show results that are recalculated with model K (*vide supra*) except that the initial analytical Ca^{2+} concentration for the native-MS experiment is set to zero, not 2.4 μM as before, so that comparisons can be made within Table S6. The overall Adair binding constants that are reported in the second column are computed from the site-specific microscopic association constants. The relative propagated error remains as that of the association constants because these are the parameters searched in this case. The eight-calcium heterotetramer species fraction values maximize at 0.816 for the model, for example, fit to the native-MS data whereas the four-calcium species fraction maximizes at a relatively small value of 1.53×10^{-4}.

Columns 4 and 5 of Table S6 give the results for a model (Model L) very much like Model K in all respects except that the overall Adair binding constants are the searched parameters. The removal of the constraints imposed by the assumptions of the microscopic structure result in only slightly different answers except for the overall Adair binding constant for the fourth calcium binding to the heterodimer. That value is now five orders of magnitude smaller than the value obtained from Model K. The relative propagated errors for this case are for the overall Adair binding constants. This case establishes the basis for comparison with the first of the Reviewer suggested models (Model M and Model N), as discussed next.

The first Reviewer-suggested model (Model M) was based on the apo, first, second and third calcium bindings to the heterodimer and the fourth, sixth and eighth calcium binding to the heterotetramer (results in columns 6 and 7 of Table S6). Those overall Adair binding constants are remarkably similar except that the relative propagated errors for the fourth- and sixth-calcium tetramer species are unusually high (column 7 of Table S6), providing little confidence in the numbers. This outcome is likely due to the fact that whereas the eight-calcium heterotetramer species is well represented, the four- and six-calcium species are not. We observed this with the model fitted to the native-MS data. The species fraction plot for the fits show that although the eight-calcium species fraction attains a maximum value of 0.816, the four-calcium species fraction attains a maximum value of 1.05×10^{-9} and the six-calcium species fraction attains a maximum of 2.83×10^{-7}. These fractions show that the four- and six-calcium heterotetramer species are insignificant, having been pushed down by the search.

The second Reviewer-suggested model (Model N, columns 8 and 9 of Table S6) considers the case when the seventh-calcium heterotetramer species is added to Model M. There are four difficulties with the solution to Model N. The first is that for peptide 27-34 of the 4 μM PLIMSTEX data the search result is up against the physical limit of 100% deuteration for the Dimer-1Ca species. The second is that for peptide 22-36 of the 4 μM PLIMSTEX data, the search result is up against physical limit of 0% deuteration limit for the Tetramer-7Ca species. The third is that the fitted model curve for peptide 22-36 of the 4 μM PLIMSTEX data has a run of residuals of the same sign for the last 12 data points. The fourth is that the fitted model curve for peptide 27-39 of the 40 μM "sharp-break" data has a run of residuals of the same sign for the last five data points. We conclude that Model N is not compatible with the data, at least in our hands. The calculated relative propagated errors are small and not thought to be valid because the search has encountered hard limits (e.g., the %D content has reached 100%). These outcomes indicate that there are too many parameters to allow the model to be successful.

Table S6. Calculated overall Adair binding constant values from experimental data and model fits.

1	2	3	4	5	6	7	8	9
Adair Constant	**Model K**[a] Dimer 1234 Tetramer 8 Adair Expressed	**Relative Propagated Error**[b] **for underlying K's**	**Model L**[a] Dimer 1234 Tetramer 8, Adair Search	**Relative Propagated Error**[b]	**Model M**[a] Dimer 123 Tetramer 468 Adair Search	**Relative Propagated Error**[b]	**Model N**[a] Dimer 123 Tetramer 4678, Adair Search	**Relative Propagated Error (see text)**
β_{11}	2.1×10^{6}	0.08	1.6×10^{6}	0.16	1.6×10^{6}	0.04	2.5×10^{6}	0.05
β_{21}	1.2×10^{12}	0.08	8.7×10^{11}	0.20	8.7×10^{11}	0.06	3.4×10^{12}	0.03
β_{31}	2.4×10^{16}	0.06	1.8×10^{16}	0.21	1.8×10^{16}	0.08	8.6×10^{18}	0.03
β_{41}	4.9×10^{16}	0.01	4.2×10^{11}	0.14				
β_{42}					5.2×10^{19}	190	7.4×10^{19}	0.17
β_{62}					1.3×10^{31}	41	7.5×10^{31}	0.15
β_{72}							3.2×10^{45}	0.09
β_{82}	2.8×10^{45}	0.006	1.6×10^{45}	0.41	1.6×10^{45}	0.15	2.1×10^{50}	0.06
4 μM RMS	1.6		1.6		1.6		1.2	
40 μM RMS	1.3		1.3		1.3		1.0	
15 μM RMS	9.2		9.2		9.2		11.4	
RMS Overall	2.3		2.3		2.3		2.5	
Number of Search Variables	4		5		6		7	

[a]Association constants were determined from the data presented in Figure 7. Initial Ca^{2+} concentration is 2.4 μM except for native-spray MS for which $[Ca^{2+}]$ is taken to be zero because extensive buffer exchange was required for this experiment and D_2O, a possible source of contaminant Ca^{2+}, was no longer used.

[b]First order approximation of the uncertainty in the fit parameters as propagated through the least squares fitting from the standard deviations of the replicate HDX measurements.

References

1. Zhang, Y., Rempel, D.L., Gross, M.L., Hydrogen exchange mass spectrometry for the analysis of ligand binding and protein aggregation. In *Hydrogen exchange mass spectrometry of proteins*, Weis, D. D., Ed. John Wiley & Sons: 2016; pp 185-207.
2. Juraschek, R.; Dulcks, T.; Karas, M., Nanoelectrospray--more than just a minimized-flow electrospray ionization source. *J Am Soc Mass Spectrom* **1999,** *10* (4), 300-8.
3. Wyman, J. G., S.J., *Binding and Linkage - Functional Chemistry of Biological Macromolecules*. University Science Books: Mill Valley, California, 1990.

Supporting Figures

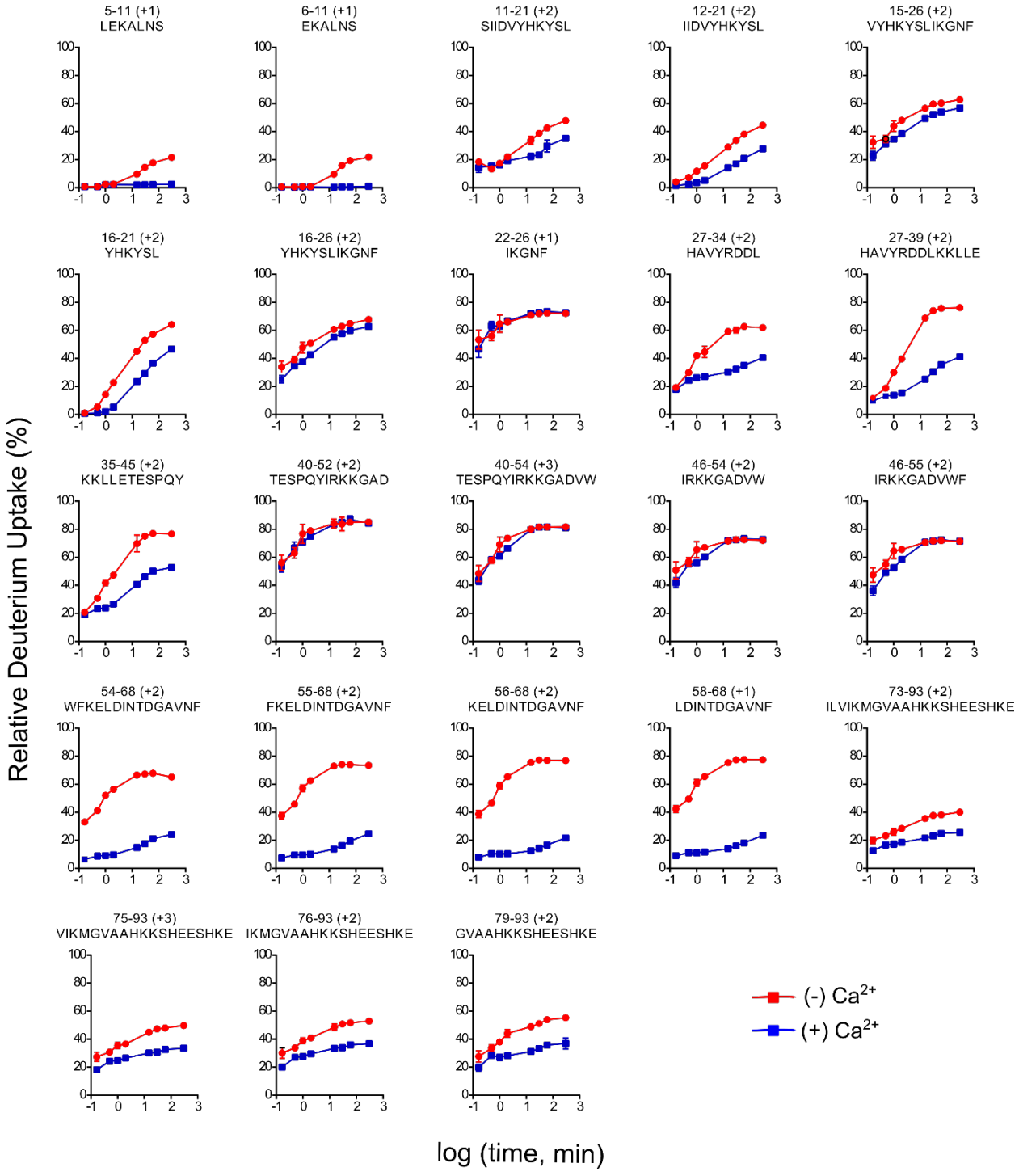

Figure S1: HDX-MS identifies regions from the Ca(II)-binding sites and tetramerization interface in CP-Ser in a differential HDX experiment in the presence or absence of Ca^{2+} for the S100A8 subunit. Relative deuterium uptake curves for all representative peptides covering entire S100A8 subunit of CP-Ser. Comparison of deuterium incorporation of CP-Ser alone (red) with the deuterium incorporation of the CP-Ser in the presence of Ca^{2+} (blue).

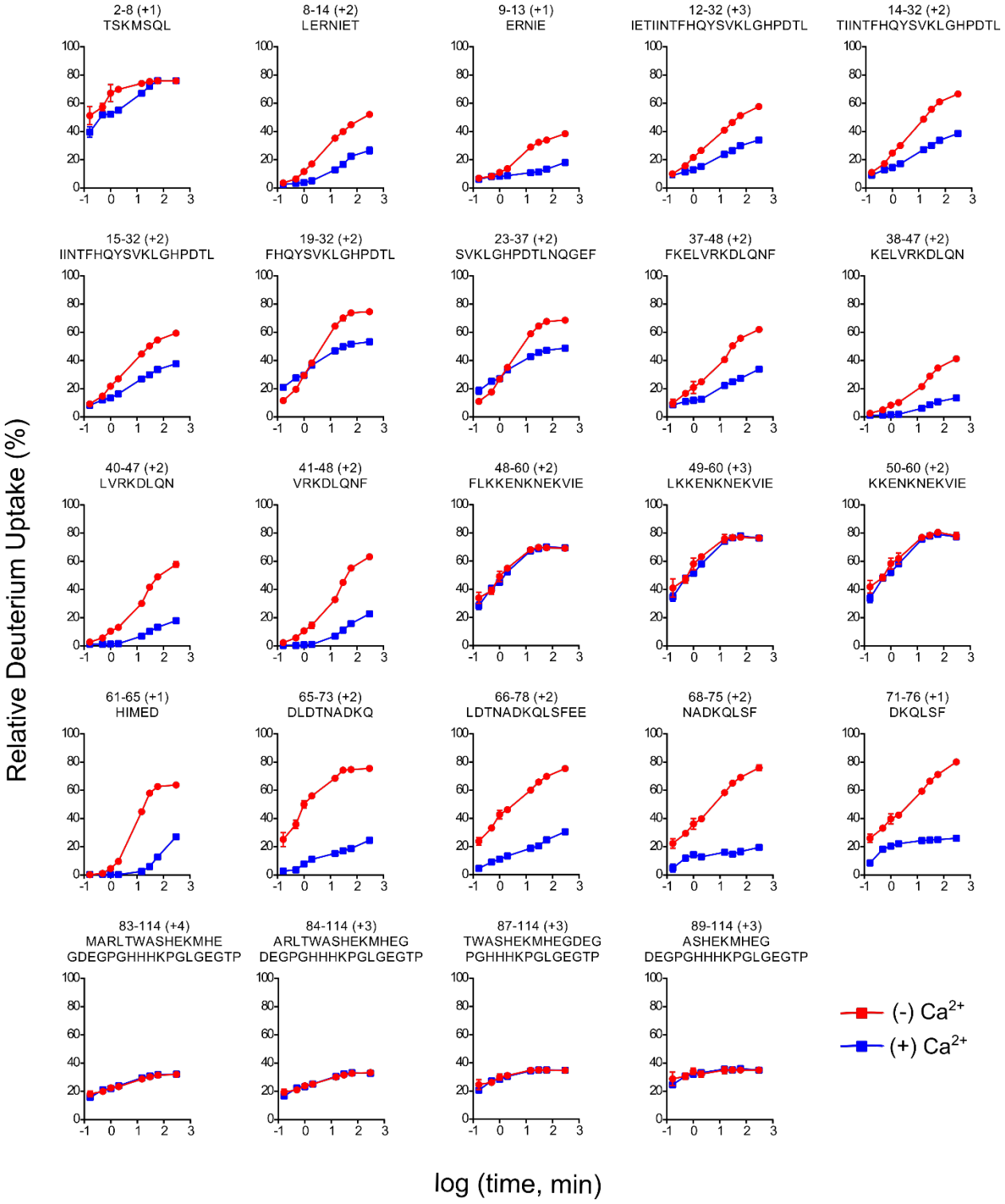

Figure S2: HDX-MS identifies regions from the Ca^{2+}-binding sites in CP-Ser in a differential HDX experiment in the presence or absence of Ca^{2+} for the S100A9 subunit. Relative deuterium uptake curves for all representative peptides covering entire S100A9 subunit of CP-Ser. Comparison of deuterium incorporation of CP-Ser alone (red) with the deuterium incorporation of the CP-Ser in the presence of Ca^{2+} (blue).

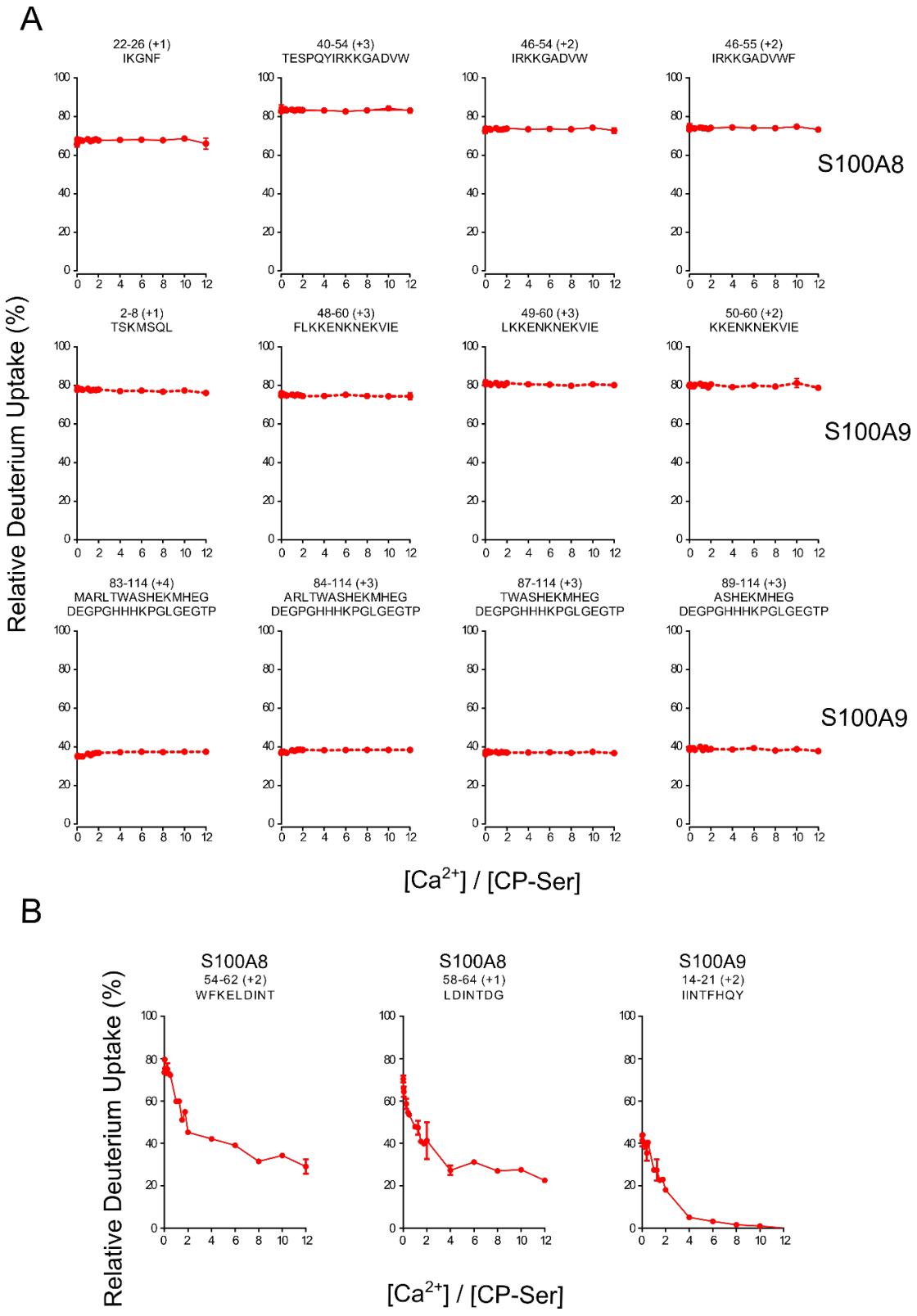

Figure S3. Sharp-break PLIMSTEX data from regions in CP-Ser that show different behavior than 4:1 sharp break. (A) Peptides that show no difference in deuterium uptake upon titration with increasing concentrations of Ca^{2+} from S100A8 subunit, top panel, and S100A9 subunit, bottom two panels, respectively. (B) Peptides that show break at 2:1 or 3:1 Ca^{2+}:CP-Ser stoichiometry.

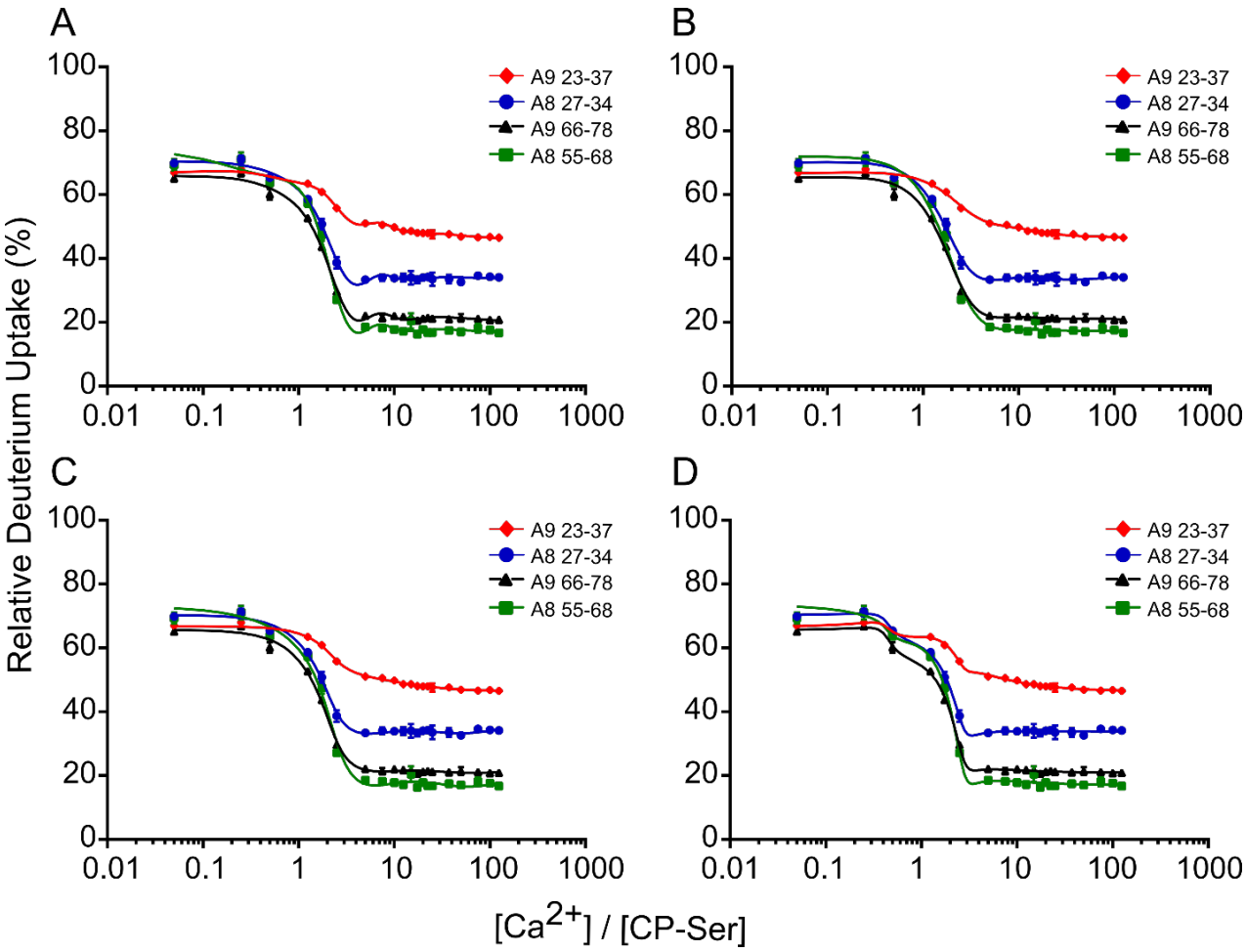

Figure S4. PLIMSTEX data for peptides from four different binding sites (shown above with different legends and residue numbers listed) fitted using the NLLS algorithm to 1:4 sequential binding model with tetramerization occurring at different Ca^{2+}-binding events included in the parameter. The fitting from the models with tetramerization occurring at (A) first Ca^{2+}-binding, (B) second Ca^{2+} -binding, and with cooperative binding considered in (C) for conditions as in (A), and (D) for conditions as in (B) are shown, respectively.

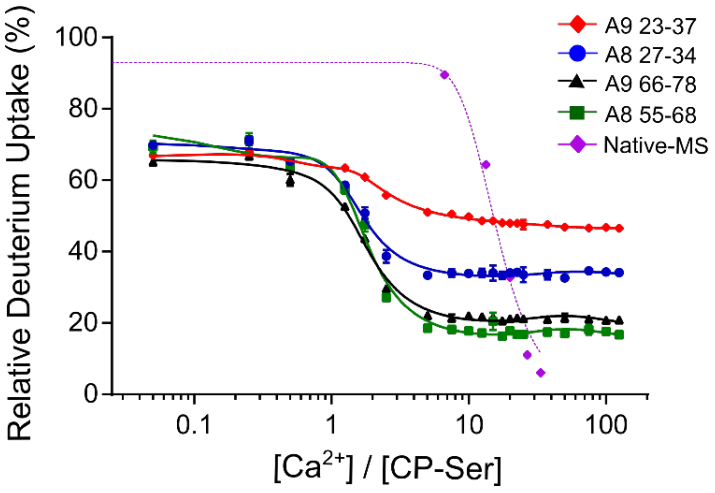

Figure S5. PLIMSTEX and native-MS data fitted to the detailed model E described in the text.

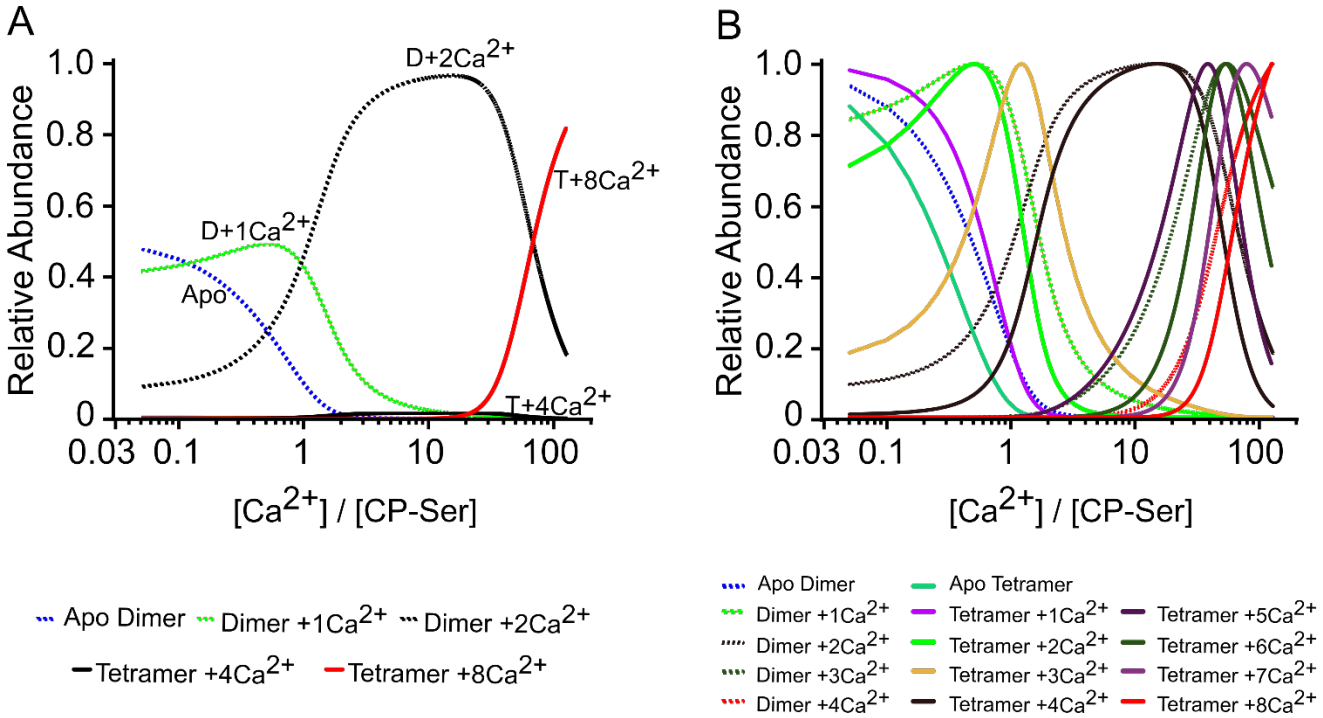

Figure S6. Fractional species plots of detailed model E shown as (A) un-normalized fractional species plots and (B) normalized fractional species plots. The relative abundance of the Ca^{2+}-bound CP-Ser species are shown as a function of the ratio of the Ca^{2+} and CP-Ser concentrations.

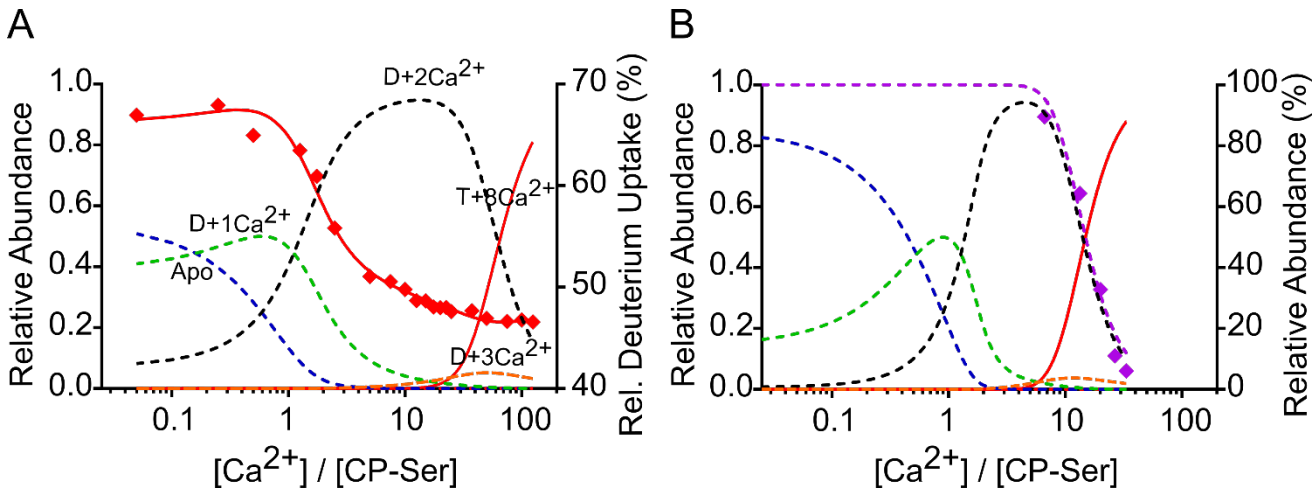

Figure S7. Fractional species plots for the reduced model H described in the text with (A) PLIMSTEX data and (B) native-MS data demonstrating that the second Ca^{2+} species fills the gap between the hetetodimer and heterotetramer signal for the PLIMSTEX data. The PLIMSTEX data are represented by the second peptide (red) (A), and native-MS data (magenta) (B). Relative abundance of each of the species as function of the [Ca^{2+}] for the heterodimer (dotted lines) and heterotetramer (solid lines) species are represented as follows: Apo (blue), dimer with 1 Ca^{2+} (green), dimer with 2 Ca^{2+} (black), dimer with 3 Ca^{2+} (orange), and tetramer with 8 Ca^{2+} (red).

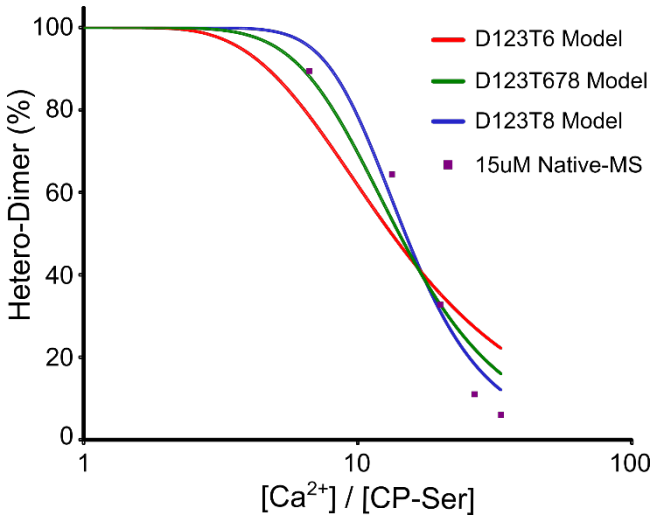

Figure S8. Comparison of reduced model F, G, and H fits to native-MS data. The native spray data is shown in black squares and the fits are represented by the curves. Model #1 (red line) represents the fit with an independent dimer with 1, 2, and 3 Ca^{2+} and tetramer with 6 Ca^{2+} bound; model #2 (green line) is described by an independent 1, 2, and 3 Ca^{2+}-bound dimer with decoupled binding of 1 and 2 Ca^{2+} in the dimer and tetramer with 6, 7, and 8 Ca^{2+}; and model #3 (blue line) represents the fit for an independent dimer 1, 2, 3, and 4 Ca^{2+} and tetramer with 8 Ca^{2+} bound to CP-Ser.

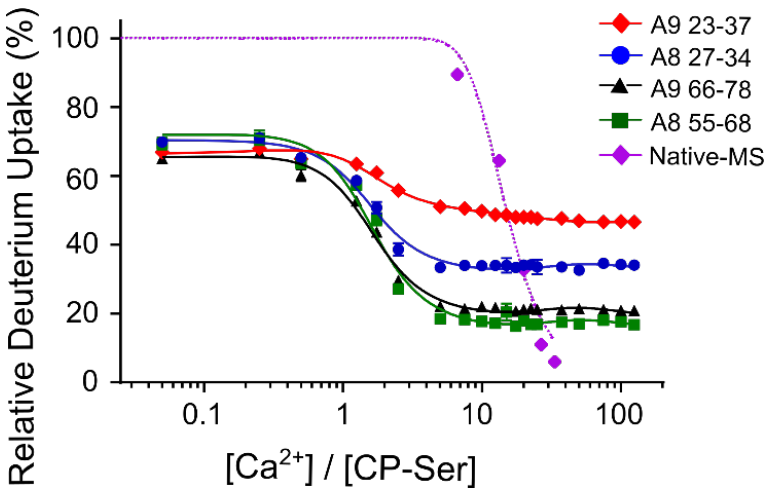

Figure S9. PLIMSTEX and native-MS data fitted together with NLLS to model H incorporating four independent Ca^{2+} (1 - 4 Ca^{2+}) in the heterodimer and 8 Ca^{2+} in the heterotetramer. PLIMSTEX data for the four peptides from the binding sites are S100A9(C3S) 23-37 (+2) (red diamonds), S100A8(C42S) 27-34 (+2) (blue circles), S100A9(C3S) 66-78 (+2) (black triangles), S100A8(C42S) 55-68 (+2) (green squares), and native-MS data (magenta diamonds), respectively, are shown.

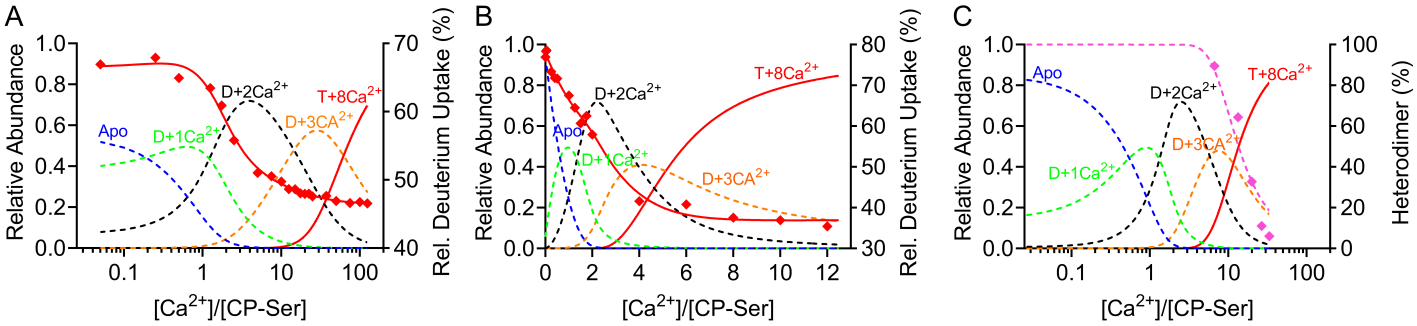

Figure S10. Fractional species plots for the reduced model K described in the text with (A) PLIMSTEX data, (B) sharp-break PLIMSTEX data, and (C) native-MS data demonstrating that the second and third Ca^{2+} species fill the gap between the hetetodimer and heterotetramer signal for the PLIMSTEX data. The PLIMSTEX data are represented by the A9 23-37 (red) peptide (A), the A8 27-39 (red) peptide (B), and native-MS data (purple) (C). Relative abundance of each of the species as function of the [Ca^{2+}] for the heterodimer (dotted lines) and heterotetramer (solid lines) species are represented as follows: Apo (blue), dimer with 1 Ca^{2+} (green), dimer with 2 Ca^{2+} (black), dimer with 3 Ca^{2+} (orange), and tetramer with 8 Ca^{2+} (red).

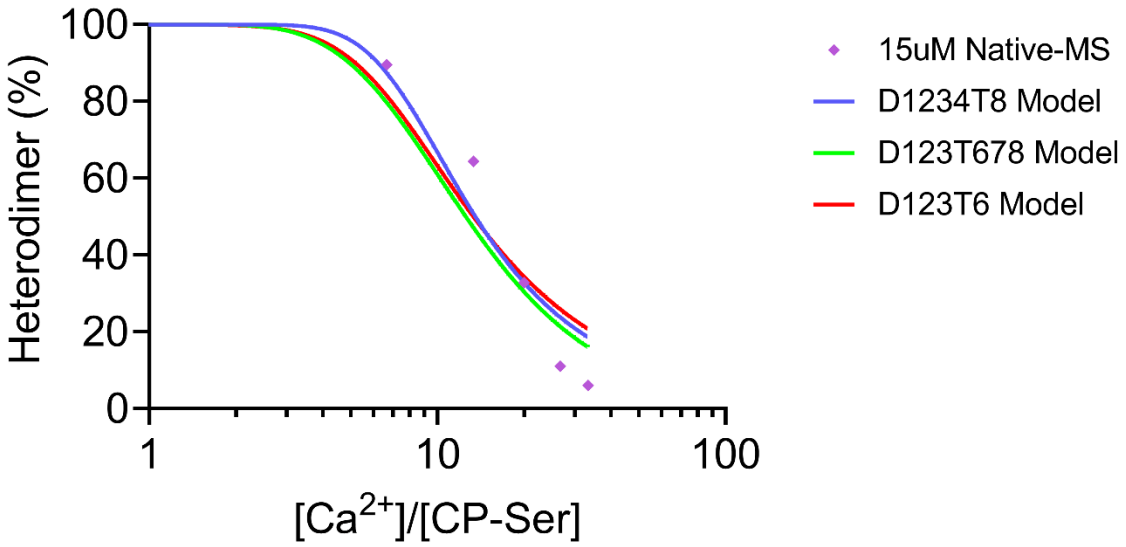

Figure S11. Comparison of reduced model fits to native-MS data. The native spray data is shown in purple diamonds and the fits are represented by the curves. Model I (red line) represents the fit with an independent dimer with 1, 2, and 3 Ca^{2+} and tetramer with 6 Ca^{2+} bound; model J (green line) is described by an independent 1, 2, and 3 Ca^{2+}-bound dimer with decoupled binding of 1 and 2 Ca^{2+} in the dimer and tetramer with 6, 7, and 8 Ca^{2+}; and model K (blue line) represents the fit for an independent dimer 1, 2, 3, and 4 Ca^{2+} and tetramer with 8 Ca^{2+} bound to CP-Ser.