JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Contrast-Matched Small-Angle X-ray Scattering from a Heavy-Atom-Labeled Protein in Structure Determination: Application to a Lead-Substituted Calmodulin–Peptide Complex

Alexander Grishaev, Nicholas J. Anthis, and G. Marius Clore*

Laboratories of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

Supporting Information

ABSTRACT: The information content in 1-D solution X-ray scattering profiles is generally restricted to low-resolution shape and size information that, on its own, cannot lead to unique 3-D structures of biological macromolecules comparable to all-atom models derived from X-ray crystallography or NMR spectroscopy. Here we show that contrast-matched X-ray scattering data collected on a protein incorporating specific heavy-atom labels in 65% aqueous sucrose buffer can dramatically enhance the power of conventional small- and wide-angle X-ray scattering (SAXS/WAXS) measurements. Under contrast-matching conditions the protein is effectively invisible and the main contribution to the X-ray scattering intensity arises from the heavy atoms, allowing direct extraction of pairwise distances between them. In combination with conventional aqueous SAXS/WAXS data, supplemented by NMR-derived residual dipolar couplings (RDCs) measured in a weakly aligning medium, we show that it is possible to position protein domains relative to one another within a precision of 1 Å. We demonstrate this approach with respect to the determination of domain positions in a complex between calmodulin, in which the four Ca^{2+} ions have been substituted by Pb^{2+}, and a target peptide. The uniqueness of the resulting solution is established by an exhaustive search over all models compatible with the experimental data, and could not have been achieved using aqueous SAXS and RDC data alone. Moreover, we show that the correct structural solution can be recovered using only contrast-matched SAXS and aqueous SAXS/WAXS data.

Small- and wide-angle X-ray scattering (SAXS/WAXS) in solution yield 1-D profiles that are determined by the pairwise distances between all atoms in a molecule.[1] Because of the convoluted nature of SAXS/WAXS, it is not possible, outside of the low-q range, to directly relate features of the scattering profiles to a particular structure. Further, it is generally not feasible to derive unique 3-D structures from 1-D curves, as many models may be compatible with a given scattering profile. On the other hand, direct refinement against SAXS/WAXS data in combination with other constraints, such as those from NMR data, can be extremely powerful.[2] Here we investigate the utility of scattering arising from a few heavy atoms, such as Pb^{2+}, under contrast-matched conditions at which the protein is rendered effectively invisible at low scattering angles by using a 65% aqueous sucrose solution as a solvent.[3] Contrast-matched scattering profiles arise from only a small number of heavy atoms and should therefore effectively constrain the distances between them. We demonstrate the utility of contrast-matched SAXS using, as an example, a calmodulin (CaM)–myosin light-chain kinase (MLCK) peptide complex in which the four coordinated Ca^{2+} ions have been substituted by Pb^{2+}. We show that the combination of contrast-matched SAXS, aqueous SAXS/WAXS, and NMR-derived residual dipolar couplings (RDCs)[4] is sufficient to define the positions of the two domains of CaM within a precision of 1 Å, and that the contrast-matched SAXS data are critical for discriminating among several models that are fully consistent with aqueous SAXS/WAXS and RDC data.

Examination of Ca^{2+}-loaded CaM–peptide complexes solved by NMR and crystallography reveals considerable variability in the relative positions of the N- and C-terminal domains (Figure 1). While some of this variability arises from differences in target peptide sequences leading to distinct binding modes,[5] even unique binding modes exhibit substantial structural variation. Further, none of these structures fit the SAXS/WAXS data within experimental error (Figures S4 and S5 and Table S1), and only in a handful of cases are the fits of the RDC data (measured in a

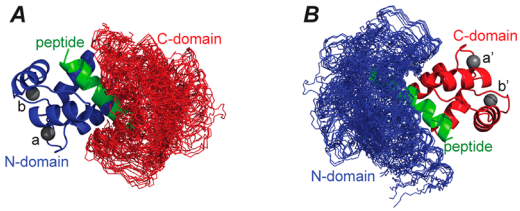


Figure 1. Distribution of the relative positions of the N- and C-domains of CaM in complexes with target peptides. Alignments on the N- (blue) and C- (red) domains are shown in (A) and (B), respectively, with the aligned domain represented by a ribbon of a single structure, and the non-aligned domain by backbone traces of the crystal structures listed in Table S1. The Ca^{2+} atoms of the N- and C-domains are shown in panels A and B, respectively, as gray spheres. The peptide from one structure (1MXE[9]) is shown as a green ribbon to illustrate how the two domains clamp the peptide.

Received: June 29, 2012
Published: August 21, 2012

dx.doi.org/10.1021/ja306359z | J. Am. Chem. Soc. 2012, 134, 14686–14689

weakly aligning medium of phage pf1[6]) to the whole complex comparable to those for the individual domains (Table S1).

Substitution of Ca^{2+} by Pb^{2+} does not significantly perturb the structure of CaM (as judged by crystallography,[7] NMR, and SAXS/WAXS), its function,[8] or its binding affinity for the MLCK peptide (see SI). Similarly, the structure of the CaM−MLCK complex appears unperturbed by the addition of sucrose as judged by NMR (see SI).

Rather than refine directly against all available SAXS/WAXS and NMR data by simulated annealing, starting from a limited set of initial coordinates,[2] we chose to evaluate the agreement with the experimental data by exhaustive sampling all possible, stereochemically feasible, relative geometries of the two domains of CaM. The latter approach guarantees that the best-fitting solution corresponds to the global minimum of the target function and is not simply one of many possible solutions, thereby removing the issue degeneracy which is universally recognized as one of the main pitfalls of biomolecular SAXS.

Fits of the RDC data to the individual domains of CaM from a large array of crystal structures indicate that the most accurate representation of the N- (residues 1−75) and C- (residues 82−148) domains in solution is that from the 1.7 Å resolution structure of a CaM−peptide complex from CaM-dependent protein kinase I (PDB code 1MXE)[9] with RDC *R*-factors of 17% (Table S1). Thus we used the N- and C-domain coordinates from the 1MXE structure throughout this study.

The rotational and translational parameters (three of each) describing the position of the two domains of CaM were systematically sampled in 2° and 1 Å steps, respectively, to generate an initial set of $\sim 4.5 \times 10^{11}$ geometries with a spatial resolution of ~1 Å. Structures were removed that (i) had steric clashes <2.5 Å between backbone or Cβ atoms of the two domains; (ii) exhibited an increase of >10% in the RDC *R*-factor (equivalent to ~10° orientational uncertainty) relative to the minimum value obtained by rigid-body minimization of the relative orientations; or (iii) had a radius of gyration (R_{gyr}) outside a range of 13−19 Å. [The experimental R_{gyr} obtained from the aqueous SAXS data via Guinier and $P(r)$ analyses is 17.8 ± 0.4 Å for the CaM−MLCK complex with Ca^{2+} and 18.3 ± 0.4 Å for the Pb^{2+}-substituted complex.] For relative domain geometries that passed the latter requirements, rotation and translation of the MLCK peptide taken from the solution structure of the CaM−MLCK complex (2BBM)[10] were obtained by minimization of a target function comprising the intermolecular NOE distance restraints from the 2BBM deposition and a repulsion term between the heavy atoms of the two CaM domains and MLCK to prevent atomic overlap. Additional geometries were then excluded in instances with steric clashes <2.5 Å between the heavy atoms of MLCK and the two domains of CaM or with NOE violations >3 Å. Finally the linker (residues 76−81) was built from a PDB database of 6250 non-redundant protein chains comprising 2.2×10^6 residues. All contiguous eight-residue stretches that did not contain either Gly or Pro (consistent with the composition of the CaM linker) were selected. The terminal residues of these eight-residue fragments were best-fitted to the backbone atoms of Lys75 of the N-domain and Glu82 of the C-domain for each CaM domain geometry. All linkers exhibiting backbone rms differences <1 Å relative to the coordinates of Lys75 and Glu82 were retained. Backbone linker geometries that did not exhibit steric clashes <2.5 Å with the two CaM domains and the MLCK peptide were processed further, and the best-fitting six-residue backbone segment corresponding to Met76−Ser81 was decorated with the appropriate side chains using residue-specific rotamers,[11] avoiding steric clashes.

The calculated SAXS/WAXS curves for the resulting 75 000 models of the CaM−MLCK complex were best-fitted to the experimental data (recorded on beamlines 12-IDC and 12-IDB, APS) using the AXES formalism which makes use of explicit water molecules to model the solvent boundary layer.[12] The scattering intensity from the Pb-substituted sample in 65% sucrose was predicted via the Debye formula from the coordinates of the Pb sites. Agreement of the aqueous SAXS/WAXS and contrast-matched SAXS data sets with structural models was evaluated using χ^2 statistics.

As is clear from Figure 2, the discriminating power of the Pb-substituted contrast-matched SAXS data is essential for selecting between candidate solutions. Three clusters of structures fit equally well to the aqueous SAXS/WAXS data (Figures 2A and 3A), and the RDCs (Figure 3C) within a few percentage points of their minimum values (Table 1). These clusters, depicted in blue (Cluster I), green (Cluster II), and red (Cluster III) in Figures 2−4 also satisfy the NOE restraints between CaM and the bound MLCK peptide (Figure 2B), indicating that all three capture the placement of the peptide within experimental error. The only observable that discriminates between the three clusters is the Pb-substituted contrast-matched SAXS data, and only one cluster, namely Cluster I, satisfies the contrast-matched SAXS data within experimental error (Figures 2A and 3B and Table 1).

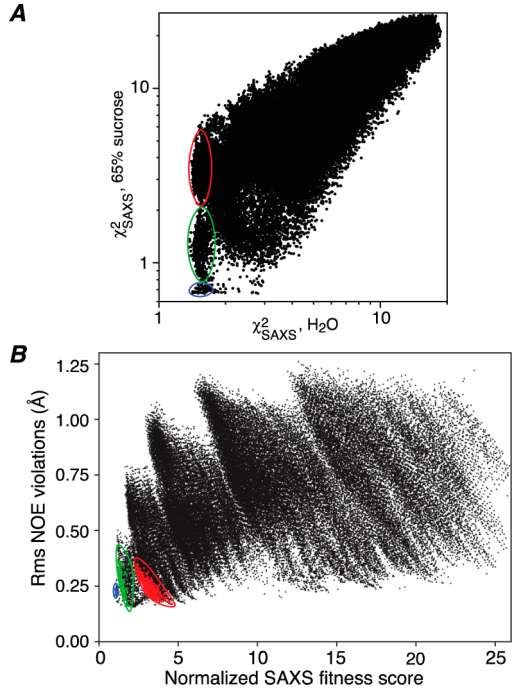


Figure 2. Agreement with SAXS data. (A) Correlation of the quality of fit of the aqueous SAXS/WAXS data with that from the contrast matched SAXS data for the Pb-substituted CaM−MLCK complex in 65% (w/v) sucrose for the 75 000 structural models generated as described in the text. (B) Correlation of the normalized SAXS fitness score with the intermolecular NOE distance restraint violations between CaM and the MLCK peptide. The three clusters discriminated by the contrast-matched SAXS data that fit equally well to the aqueous SAXS/WAXS, RDC, and NOE data are highlighted by the blue, green, and red ovals. The normalized SAXS fitness score is given by $\{[(\chi^2_{SAXS,water}/\chi^2_{SAXS,water,min})^2 + (\chi^2_{SAXS,sucrose}/\chi^2_{SAXS,sucrose,min})^2]/2\}^{1/2}$.

dx.doi.org/10.1021/ja306359z | *J. Am. Chem. Soc.* 2012, 134, 14686−14689

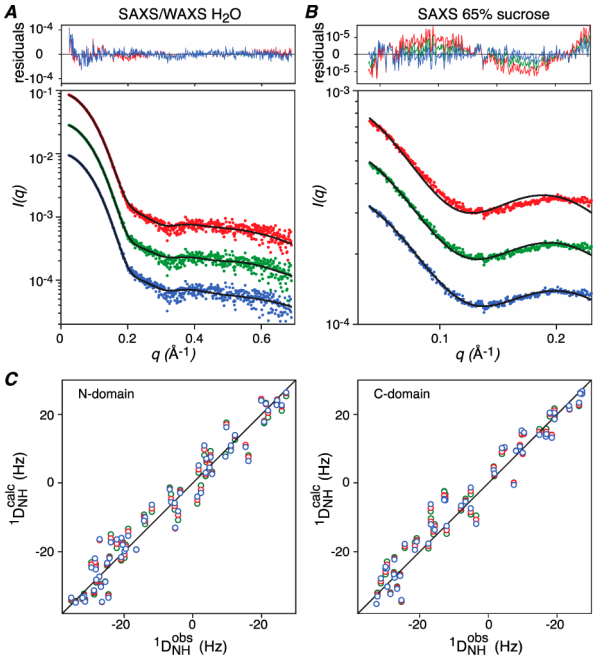


Figure 3. Agreement between observed and calculated SAXS and RDC data for the structures closest to the mean of each cluster: (A) SAXS/WAXS in water, (B) contrast-matched SAXS in 65% (w/v) sucrose, and (C) RDCs for the N- (left panel) and C- (right panel) domains. Cluster I is depicted in blue, cluster II in green, and cluster III in red.

Table 1. Parameters of the Three Clusters Selected Based on $\chi^2_{SAXS,\ water}$ Values <1.5 and Grouped According to $\chi^2_{SAXS,\ sucrose}$[a]

	Cluster I	Cluster II	Cluster III
$\chi^2_{SAXS,\ sucrose}$	0.69 ± 0.02	1.25 ± 0.22	3.42 ± 0.53
$\chi^2_{SAXS,\ water}$	1.46 ± 0.03	1.46 ± 0.03	1.46 ± 0.03
RDC *R*-factor	0.18 ± 0.00	0.19 ± 0.00	0.18 ± 0.01
D_a^{NH} (Hz)	−17.6 ± 0.1	−17.6 ± 0.2	−17.7 ± 0.1
η	0.35 ± 0.01	0.35 ± 0.01	0.35 ± 0.01
$r_{Pb(a)-Pb(a')}$ (Å)	30.9 ± 0.2	31.9 ± 0.3	33.3 ± 0.3
$r_{Pb(a)-Pb(b')}$ (Å)	34.3 ± 0.2	35.6 ± 0.9	37.4 ± 0.4
$r_{Pb(b)-Pb(a')}$ (Å)	36.9 ± 0.2	37.5 ± 0.9	38.2 ± 0.3
$r_{Pb(b)-Pb(b')}$ (Å)	39.7 ± 0.2	40.7 ± 0.3	41.8 ± 0.2

[a]RDC *R*-factor is given by $[\langle (D_{obs} - D_{calc})^2\rangle / 2\langle D_{obs}^2\rangle]^{1/2}$, where D_{obs} and D_{calc} are the observed and calculated RDCs, respectively. D_a^{NH} and η are the magnitude of the principal component of the alignment tensor and the rhombicity, respectively. There are 115 RDCs (61 and 54 for the N- and C-domains, respectively). The correspondence of the Pb^{2+} atoms within the structure is shown in Figure 4.

A structural comparison of the three clusters is shown in Figure 4, and the rms differences within and between the clusters are provided in Table 2. For Cluster I, the relative position of the two CaM domains is defined with a precision of 1 Å. When best-fitted to the N-domain, the backbone rms displacements of the C-domains of Clusters II and III relative to Cluster I are much larger, 1.8 and 5.4 Å, respectively, well outside the coordinate precision of Cluster I. Structural differences between the three clusters reflect systematic lengthening of the Pb−Pb distances from Cluster I to Clusters II and III (Figure 4 and Table 1).

Given the discriminating power of the Pb-substituted contrast-matched SAXS data, can such data be directly used to accurately extract Pb−Pb distances? To assess this we carried out fits to the contrast-matched SAXS data using a Monte Carlo-generated random sampling of 4-atom geometries to represent the Pb sites

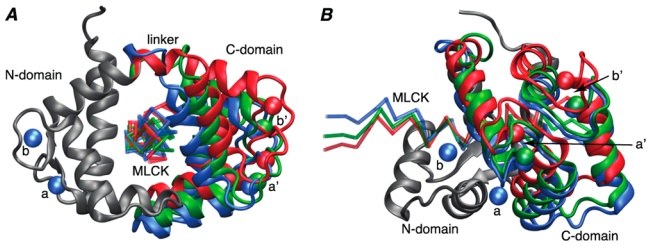


Figure 4. Structural comparison of Clusters I, II, and III. The N-domain (gray) of the three clusters is superimposed, illustrating the relative displacements of the C-domain of Clusters I (blue), II (green), and III (red). The backbone of CaM is displayed as a ribbon diagram, the backbone of the MLCK peptide is shown as a Cα trace, and the positions of the Pb^{2+} ions are displayed as spheres. Only Cluster I accounts for the Pb-substituted contrast-matched SAXS data. The backbone rms displacement of the C-domain of Cluster I relative to the 1MXE coordinates[9] is 2.5 Å.

Table 2. Backbone Root-Mean-Square Difference between the C-Terminal Domains of Clusters I, II, and III, When Best-Fitted to the N-Domain[a]

	backbone rms difference (Å)		
Cluster	I	II	III
I	1.1	1.8	5.4
II		2.4	3.9
III			1.4

[a]Diagonal elements are the average rms difference within a cluster to the structure closest to the mean coordinate positions of the cluster, and off-diagonal elements are rms differences between the structures closest to the means of the respective clusters.

Table 3. Pb−Pb Distances Determined Directly from the Pb-Substituted Contrast-Matched SAXS Data

no.[a]	Pb−Pb (Å)
4[b]	31.5 ± 1.0, 33.8 ± 0.9, 36.5 ± 1.1, 39.6 ± 1.1
5[c]	11.1 ± 2.9, 30.2 ± 2.3, 33.9 ± 1.3, 36.3 ± 1.3, 40.6 ± 1.5
6	6.4 ± 3.6, 14.7 ± 3.1, 29.9 ± 1.9, 32.3 ± 1.9, 38.1 ± 1.8, 41.1 ± 1.3

[a]Number of unknown distances. [b]Only the four interdomain Pb−Pb distances are varied, and the two intradomain Pb−Pb distances are fixed to their known value of 11.7 Å. [c]The two intradomain Pb−Pb distances are varied but constrained to be equal.

(see SI). We consider three cases: four variable interdomain Pb−Pb distances with the two intradomain distances fixed to the values measured from the domain coordinates (11.7 Å); five variables comprising the four interdomain Pb−Pb distances with the two intradomain distances constrained to have the same value; and six variables in which all four interdomain and two intradomain Pb−Pb distances are allowed to vary. For the four-variable case, the average interdomain Pb−Pb distances derived from the contrast-matched SAXS data have uncertainties of only ~1 Å (Table 3) and are in excellent agreement with the corresponding Pb−Pb distances in Cluster I (cf. Table 1). The five-variable case results in larger uncertainties, but agreement with the Cluster I Pb−Pb distances is still excellent and within the uncertainties of the distance estimates. Finally, the six-variable case results in reasonable agreement with the longer, interdomain Pb−Pb distances from the all-atom model but exhibits marginal agreement with the correct intradomain Pb−Pb separations. These calculations therefore indicate that up to five (from a total of six) Pb−Pb distances can be extracted both

dx.doi.org/10.1021/ja306359z | J. Am. Chem. Soc. 2012, 134, 14686−14689

precisely and accurately from a single curve of the contrast-matched SAXS data arising from four Pb labels. Accelerated deterioration of the uncertainties of the shorter separations is due to increased uncertainty at wider angles and the limited q_{max} of the fitted data, as the impact of the shorter Pb−Pb distances is increasingly felt at higher q.

In light of the above results, we sought to investigate whether the contrast-matched SAXS and aqueous SAXS/WAXS data alone are sufficient to arrive at the correct solution without recourse to filtering by RDCs and intermolecular NOE distance violations. Using the same grid procedure, structures were filtered by the fits to the X-ray scattering data using a cutoff of $\chi^2_{SAXS,water} < 1.69$ and $\chi^2_{SAXS,sucrose} < 0.72$, the absence of steric clashes, an R_{gyr} of 13−19 Å, and the ability to form a linkage between the N- and C-domains. The results in Figure 5 indicate that the cluster with the lowest normalized SAXS fitness score yields solutions with a coordinate accuracy of better than 1 Å by reference to Cluster I.

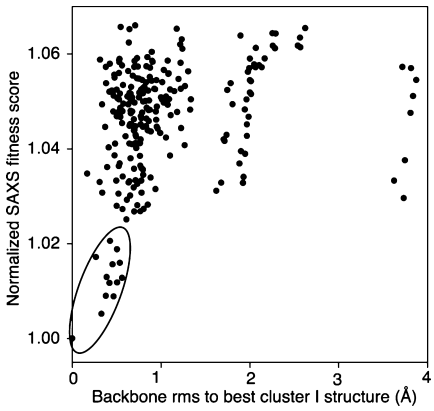


Figure 5. Accuracy of best-fitting models ($\chi^2_{SAXS,water} < 1.69$ and $\chi^2_{SAXS,sucrose} < 0.72$) obtained from the grid search procedure using aqueous and contrast-matched SAXS data as the only experimental restraints. The normalized SAXS fitness score is defined in the legend to Figure 2. The best Cluster I structure is the structure closest to the mean coordinate positions of the Cluster I structures.

While this work capitalizes on the ability of CaM to specifically bind Pb^{2+} in place of Ca^{2+}, applications of this approach can be readily extended beyond metal-binding proteins by incorporating heavy-atom ions such as Pb^{2+} or Hg^{2+} into EDTA moieties conjugated via disulfide bonds to engineered surface cysteines as routinely done in NMR paramagnetic relaxation enhancement studies.[13] Although the EDTA−metal moiety samples a large region of conformational space, the metal−metal separations measured by contrast-matched SAXS are simple linear averages of all the conformations present in solution and each metal atom can therefore be represented by a single average position.

Although, in principle, similar information has been obtained from small-angle neutron scattering (SANS) of ^{240}Pu,[14] or X-ray scattering of DNA with attached gold nanoclusters,[15] the present approach does not suffer from the lower signal-to-noise of SANS and the use of a radioactive isotope, or the necessity to decompose the observed data into individual scattering functions from measurements on a series of samples.

Finally, the results obtained without recourse to filtering by NMR data suggest that accurate "triangulation-driven" assembly of multi-component protein architectures, based only on a combination of aqueous and contrast-matched/heavy-atom-labeled SAXS data, is feasible providing the structures of the individual subunits are known.

■ ASSOCIATED CONTENT

Supporting Information

Experimental procedures, data tables, figures, and complete ref 1b. This material is available free of charge via the Internet at http://pubs.acs.org. Coordinates of the best-fitting Cluster I model and experimental data (PDB 2LV6) have been deposited in the Protein Data Bank (http://www.rcsb.org).

■ AUTHOR INFORMATION

Corresponding Author

mariusc@mail.nih.gov

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank A. Bax for useful discussions, and S. Seifert and X. Zuo for assistance with SAXS data collection. This work was supported by the intramural program of NIDDK and the Intramural AIDS Antiviral Program of the Office of the Director of the NIH (G.M.C.). The shared beamline resource (PUP-77 agreement between NCI, NIH, and Argonne National Laboratory) at the APS is acknowledged.

■ REFERENCES

(1) (a) Koch, M. H.; Vachette, P.; Svergun, D. I. *Q. Rev. Biophys.* **2003**, *36*, 147. (b) Hura, G. L.; et al. *Nature Methods* **2009**, *6*, 606.

(2) (a) Grishaev, A.; Wu, J.; Trewhella, J.; Bax, A. *J. Am. Chem. Soc.* **2005**, *127*, 16621. (b) Schwieters, C. D.; Clore, G. M. *Biochemistry* **2007**, *46*, 1152. (c) Grishaev, A.; Tugarinov, V.; Kay, L. E.; Trewhella, J.; Bax, A. *J. Biomol. NMR* **2008**, *40*, 95. (d) Schwieters, C. D.; Suh, J. Y.; Grishaev, A.; Ghirlando, R.; Takayama, Y.; Clore, G. M. *J. Am. Chem. Soc.* **2010**, *132*, 13026. (e) Takayama, Y.; Schwieters, C. D.; Grishaev, A.; Ghirlando, R.; Clore, G. M. *J. Am. Chem. Soc.* **2011**, *133*, 424.

(3) Lipfert, J.; Doniach, S. *Annu. Rev. Biophys. Biophys. Chem.* **2007**, *36*, 307.

(4) Bax, A.; Grishaev, A. *Curr. Opin. Struct. Biol.* **2005**, *15*, 563.

(5) (a) Hoeflich, K. P.; Ikura, M. *Cell* **2002**, *108*, 739. (b) Maximciuc, A. A.; Putkey, J. A.; Shamoo, Y.; Mackenzie, K. R. *Structure* **2006**, *14*, 1547.

(6) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571.

(7) Kursula, P.; Majava, V. *Acta Crystallogr.* **2007**, *F63*, 653.

(8) Fullmer, C. S.; Edelstein, S.; Wasserman, R. H. *J. Biol. Chem.* **1985**, *260*, 6816.

(9) Clapperton, J. A.; Martin, S. R.; Smerdon, S. J.; Gamblin, S. J.; Bayley, P. M. *Biochemistry* **2002**, *41*, 14669.

(10) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632.

(11) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr. *Proteins* **2009**, *77*, 778.

(12) Grishaev, A.; Guo, L.; Irving, T.; Bax, A. *J. Am. Chem. Soc.* **2010**, *132*, 15484.

(13) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108.

(14) Seeger, P. A.; Rokop, S. E.; Palmer, P. D.; Henderson, S. J.; Hobart, D. E.; Trewhella, J. *J. Am. Chem. Soc.* **1997**, *119*, 5118.

(15) Mathew-Fenn, R. S.; Das, R.; Silverman, J. A.; Walker, P.; Harbury, P. A. B. *PloS ONE* **2008**, *3*, e3229.

Supplementary Information

Contrast-Matched Small Angle X-ray Scattering from a Heavy Atom-Labeled Protein in Structure Determination: Application to a Lead-Substituted Calmodulin-Peptide Complex

Alexander Grishaev, Nicholas J. Anthis and G. Marius Clore*

Laboratories of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Experimental Procedures

Protein production and sample preparation

The 148-residue human calmodulin (CaM) protein was expressed and purified as described previously.[S1] NMR experiments were performed on CaM in a buffer consisting of 5% D_2O/95% H_2O, 100 mM KCl, 0.1x Roche Complete Protease Inhibitor, and 25 mM HEPES, pH 6.5. Depending on the specific experiment $CaCl_2$ or $PbCl_2$ was also added. SAXS experiments were performed under similar conditions, in a buffer consisting of 100% H_2O, 150 mM KCl, 1 mM TCEP, 0.1x Roche Complete Protease Inhibitor, and 25 mM HEPES, in the presence or absence of 65 % (w/v) sucrose. $CaCl_2$ (3 mM), $PbCl_2$ (3 μM), or EGTA (1 mM) was included in each buffer. For the Ca^{2+}- or Pb^{2+}-bound CaM samples, four molar equivalents of Ca^{2+} or Pb^{2+} were first added to CaM, and then CaM was exchanged into these conditions using a HiPrep 26/10 Desalting Column (GE Life Sciences) or dialysis with a 3.5 kDa molecular weight cutoff membrane (Pierce) (e.g. a SAXS sample containing 300 μM $4Pb^{2+}$-CaM would contain 1200 μM Pb^{2+} bound to CaM, and a small excess—3 μM—free in solution). Full removal of Ca^{2+}, for studies on $4Pb^{2+}$-CaM, required first equilibrating CaM with a chelating buffer (e.g. containing at least 1 mM EGTA), and then equilibrating into a chelator-free buffer before the addition of $PbCl_2$. Samples were concentrated to 0.3 mM (5 mg/mL)-0.6 mM (10 mg/mL) with Amicon Ultra Centrifugal Filter Units (3 kDa molecular weight cutoff). For peptide-bound CaM ($4Ca^{2+}$-CaM-MLCK or $4Pb^{2+}$-CaM-MLCK), CaM was mixed with 1.2 equivalents of target peptide (skeletal muscle myosin light chain kinase [skMLCK] M13 peptide; commercially synthesized by Anaspec; KRRWKKNFIAVSAANRFKKISSSGAL) in the appropriate buffer and then concentrated. For SAXS experiments, these samples were further dialyzed into the appropriate buffer.

NMR spectroscopy

NMR experiments were performed on uniformly $^{2}H/^{13}C/^{15}N$-labeled CaM. Data were processed with NMRPipe[S2] and analyzed using the program Sparky (www.cgl.ucsf.edu/home/sparky). Data were acquired using transverse relaxation optimized (TROSY) 2D ^{1}H-^{15}N HSQC spectra, acquired at 27°C on a Bruker 600 MHz spectrometer equipped with a triple resonance *z*-gradient cryoprobe. $^{1}D_{NH}$ RDCs were measured on samples comprising 0.3 mM $^{2}H/^{13}C/^{15}N$-labeled CaM in the presence or absence of the MLCK peptide by taking the difference in one-bond ^{1}H-^{15}N splittings between aligned (15 mg/mL phage pf1[S3]) and isotropic media using the ARTSY method.[S4]

The addition of up to 4 molar equivalents of Ca^{2+} or Pb^{2+} causes similar changes in the $^{1}H_N/^{15}N$ chemical shifts of CaM (Fig. S1). However, in contrast to Ca^{2+}, addition of Pb^{2+} over and above 4 equivalents does lead to some further spectral changes (Fig. S1B) due to the presence of lower affinity Pb^{2+} sites on CaM outside of the 4 Ca^{2+}-binding sites.[S5] For this reason, whereas

SAXS experiments on Ca^{2+}-CaM were conducted in the presence of a large molar excess of Ca^{2+} (3 mM), it was important to conduct SAXS experiments on Pb^{2+}-CaM with only a very small molar excess of Pb^{2+} (3 μM) to ensure that all Pb^{2+} present was bound exclusively to the 4 Ca^{2+}-binding sites.

Since the presence of the high concentrations of sucrose (65% w/v) used in the contrast-matched SAXS experiments could potentially perturb the structure of CaM, we also verified that sucrose at a concentration of 50% (w/v) did not result in any significant perturbation in $^{1}H_N/^{15}N$ chemical shifts (Fig. S2). (The high viscosity at 65% w/v sucrose and concomitant line broadening owing to the accompanying increase in the rotational correlation time precluded recording a ^{1}H-^{15}N correlation spectrum at a sucrose concentration above 50% w/v).

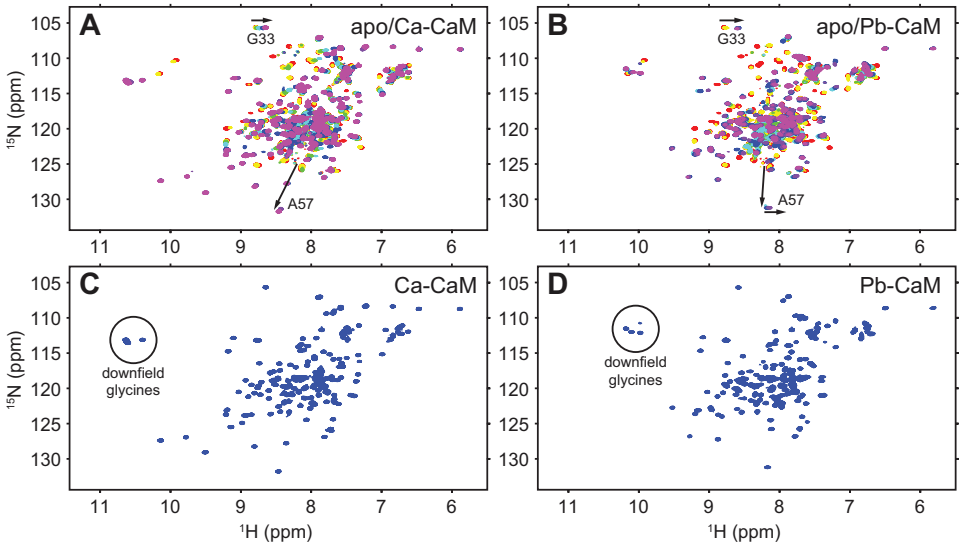


Fig. S1. CaM binding to Ca^{2+} and Pb^{2+} observed by NMR. (A) ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled CaM (300 μM) titrated with Ca^{2+}. Apo CaM (red) is shown overlaid with spectra of CaM in the presence of 1 (yellow), 2 (green), 3 (cyan), 4 (blue), 5 (purple), and 6 (magenta) molar equivalents of Ca^{2+}. (B) As in panel A, but with Pb^{2+} instead of Ca^{2+}. The two titrations are broadly similar, although Pb^{2+} binding to CaM occurs in the slow-intermediate exchange regime, compared to the fast-intermediate exchange for Ca^{2+} binding, indicative of a higher affinity of CaM for Pb^{2+} over Ca^{2+}. Peaks for Gly33 are highlighted in panels A and B, showing the same shifts upon binding Ca^{2+} and Pb^{2+}. Whereas Ca^{2+}-loaded CaM approaches saturation at 4 Ca^{2+} equivalents (and only achieves full saturation in the presence of an excess of Ca^{2+} due to lower affinity binding for the N-domain), CaM appears saturated at 4 equivalents of Pb^{2+}. Above 4 equivalents of Pb^{2+}, however, the spectrum undergoes additional changes, with peaks losing intensity and shifting in a different direction, due to additional lower affinity binding sites for Pb^{2+} outside of the 4 Ca^{2+}-binding sites.[S5] The cross-peak for Ala57, a residue located both within Ca^{2+}-binding loop a (see Fig. 4) and near a non-Ca^{2+} Pb^{2+}-binding site seen in the crystal structure[S5], illustrates the fact that the non-Ca^{2+} sites are only affected by Pb^{2+} at concentrations above those required to saturate the four Ca^{2+}-binding sites. Thus, whereas SAXS experiments on Ca-CaM were conducted in the presence of a large excess of Ca^{2+} (3 mM), it was important to conduct experiments on Pb-CaM with only a very small excess of Pb^{2+} (3 μM) to achieve binding at only the 4 Ca^{2+}-binding sites. Panels (C) and (D) show Ca-CaM (3 mM molar excess of Ca^{2+}) and Pb-CaM (3 μM molar excess of Pb^{2+}) under the conditions used in SAXS experiments. Four glycine residues (Gly25, Gly61, Gly98, Gly134; one from each Ca^{2+}-binding loop) show characteristic downfield shifts in the ^{1}H dimension upon binding either Ca^{2+} or Pb^{2+}.

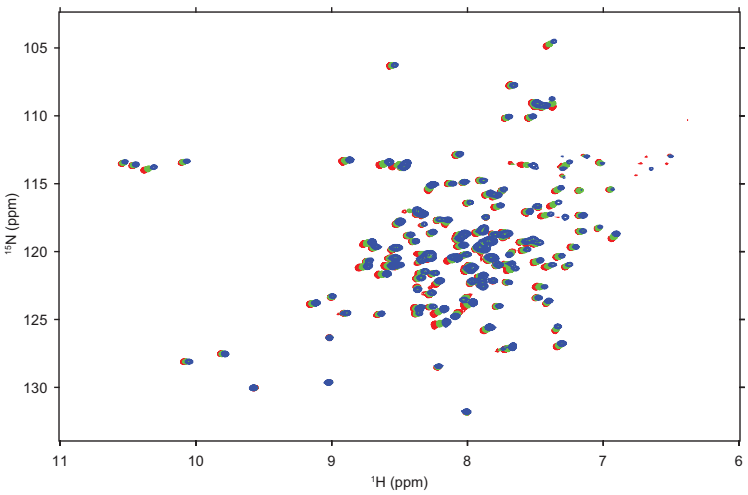


Fig. S2. The CaM-MLCK complex is unperturbed by the addition of sucrose. ^{1}H-^{15}N HSQC spectra of Ca^{2+}-loaded CaM bound to the MLCK peptide, in the presence of 0% (red), 25% (green), and 50% (blue) sucrose (w/v). Spectra were acquired at 50°C on a sample of 250 µM ^{2}H/^{13}C/^{15}N-labeled CaM. The very small shifts observed in the spectra upon the addition of sucrose (< 0.1 ppm in the ^{1}H dimension) are due to solvent effects and not to any changes in the structure of CaM.

Fluorescence experiments

Fluorescence experiments were carried out at 27°C using a Jobin Ybon FluoroMax-3 fluorometer equipped with a Peltier temperature control unit. The intrinsic fluorescence intensity of the single tryptophan residue in the MLCK was monitored with excitation at 280 nm and emission readings at 357 nm (free peptide) and 331 nm (bound to CaM). Measurements were acquired on 8 nM MLCK in the presence of 0-20 nM of either $4Ca^{2+}$-CaM or $4Pb^{2+}$-CaM. Experiments were performed in the same buffers used for NMR experiments, using either an 8 mM $CaCl_2$ or 3 µM $PbCl_2$ molar excess over CaM. Individual readings were taken for 2 sec. and 6-10 individual readings were taken for each concentration of CaM. Binding was monitored by subtracting the fluorescence intensity at 357 nm (corresponding to peak fluorescence emission for the free peptide) from that at 331 nm (corresponding to peak fluorescence emission for CaM-bound peptide). A linear function was then subtracted from the data to correct for baseline intensity and the fluorescence intensity of the two tyrosine residues in CaM (which increases linearly with the concentration of CaM). K_D values were determined by fitting the normalized linearly corrected fluorescence signal to the following equation:

$$F = \{[L]+[U]+K_D - \sqrt{([L]+[U]+K_D)^2 - 4[L][U]}\} / 2[L]$$

where F is the fraction of MLCK bound to CaM, and $[L]$ and $[U]$ are the total concentrations of MLCK peptide and CaM, respectively. Data were fit using OriginPro 8.

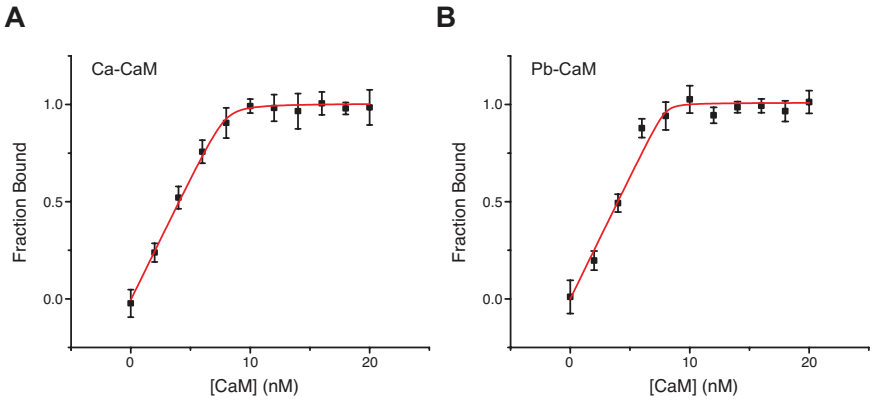


Fig. S3. Binding of the MLCK peptide to $4Ca^{2+}$-CaM and Pb^{2+}-CaM monitored by fluorescence spectroscopy. Fluorescence measurements indicate that (A) $4Ca^{2+}$-CaM and (B) $4Pb^{2+}$-CaM (B) both bind to MLCK with a very high affinity. Intrinsic tryptophan fluorescence of the MLCK peptide was measured for a range of CaM concentration, as described in the experimental procedures. Each point is the average of 6-10 individual measurements, with the standard deviation indicated by error bars. A best-fit line for a simple binding isotherm has been calculated for each titration, giving K_D values of 50±50 pM and 10±30 pM for $4Ca^{2+}$-CaM and $4Pb^{2+}$-CaM, respectively. The affinity of the MLCK peptide for both Ca^{2+} and Pb^{2+}-loaded CaM, however, is too high to be accurately determined from this data, so a more conservative estimate would be that for both interactions the K_D is <<1 nM. Thus, these data indicate that Pb^{2+}-loaded CaM binds to MLCK similarly to Ca^{2+}-loaded CaM.

Acquisition and analysis of SAXS data

Preliminary optimization of SAXS contrast-matching conditions for Pb-substituted CaM-MLCK complex was carried out with scattering data collected on an Anton Paar SAXSess instrument using the Kratky camera geometry, line-shaped 15 mm beam at Cu K^{α} wavelength (8 keV) and protein solutions in sucrose-containing buffers at 5 mg/mL protein concentration. Final solution X-ray scattering data were acquired at the Beam Lines 12-IDC and 12-IDB at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). For Pb-substituted CaM-MLCK samples, data collection was done at station 12-IDC using a Gold CCD detector positioned 3.5 m away from the sample capillary. Incident radiation with an energy of 18 keV resulted in an observable q-range of 0.01 - 0.23 A^{-1}. Data in aqueous buffer were collected at APS station 12-IDB using 12 keV incident radiation, 3m sample/detector distance and an off-center Pilatus 2M detector, resulting in an observable q-range from ~0.01 $Å^{-1}$ to ~0.7 $Å^{-1}$. Q-axis mapping was done using scattering from the silver behenate standards. A total of 20 sequential data frames was recorded for each data collection with the sample temperature set to 25°C. To prevent radiation damage, 120 µL volumes of samples and buffers were oscillating during data collection using a flow-though setup. Individual data frames were masked, corrected for the detector sensitivity and solid angles per each pixel, radially integrated and normalized by the corresponding incident beam intensities and transmissions. The final one-dimensional scattering profiles and their uncertainties were calculated as means and mean uncertainties over the 20 individual frames. The buffer data were then subtracted from the samples. To evaluate the magnitude of a possible structure factor, data were collected at protein concentrations ranging from 1.5 to 5.0 mg/mL for aqueous samples and from 5.0 to 9.0 mg/mL for samples in 65% sucrose buffer. At these concentrations the data proved to be indistinguishable for q >0.02 $Å^{-1}$ in aqueous buffers and >0.035 $Å^{-1}$ in buffers with 65% (w/v) sucrose. The absence of aggregation in the aqueous samples was verified by extrapolation of the scattering intensities to the zero angle,

by reference to the sample concentrations measured by UV absorption and the $I(0)$ values obtained for well-behaved standards of with known concentrations (hen egg white lysozyme and horse heart cytochrome c).

The coordinates of the individual N- (residues 1-75) and C- (residues 82-148) domains of $4Ca^{2+}$-CaM (taken from PDB 1MXE) were rotated into the frames of their alignment tensors SVD-fitted to the $4Ca^{2+}$-CaM-MLCK RDCs so that the z-axes of the molecular frames coincided with the SVD-fitted C_2–proximal axes for both domains. The domains were positioned at the origin of the coordinate system and the orientations of the C-domain relative to the N-domain were systematically sampled with a step size of ~2° using an 11,000-vector Fibonacci grid of order N=20 for rotations relative to the z axis and 2° rotation steps for the orthogonal dimension. To avoid the well-known inhomogeneity of the Fibonacci grid in the vicinity of the z-axis, the vector grid was rotated by 90° around the x-axis prior to usage. The C- domain was then uniformly translated in all three dimensions on a 1 Å cubic grid. Such orientational and translational grids are internally consistent since a 2° rotation of the C-domain about its center of mass produces a ~1 Å r.m.s.d. change in its coordinates. This procedure generated an initial set of $\sim 4.5 \times 10^{11}$ orientational/translational configurations, from which $\sim 4 \times 10^{9}$ passed the RDC fit criterion and were subsequently filtered to ~75,000 configurations by the clash avoidance, R_{gyr}, NOE and linkage criteria, as described in the main text.

Utility of SAXS contrast variation data

Unlike neutron scattering, where contrast variation is conveniently achieved by changing the $^1H/^2H$ composition of the aqueous solvent, contrast-matching a protein is more challenging with SAXS as the electron density of the solvent has to be increased by ~30% to make a protein of approximately constant electron density invisible for X-ray scattering relative to the buffer. Of the two possible choices of buffer additives – heavy alkali/halide salts or carbohydrates, the latter are preferable since they do not result in a dramatic increase of X-ray absorption. Out of all carbohydrates, sucrose is used most frequently due to its very high solubility in water. Since such a high concentration of sucrose could potentially affect the structure of the studied protein, we verified the absence of any sucrose-induced changes in the Ca^{2+}-loaded CaM-MLCK structure relative to that in aqueous buffer by analysis of ^1H-^{15}N HSQC spectra, which revealed no significant changes in $^1H_N/^{15}N$ chemical shifts (Fig. S2).

The above observation is in line with other reports in the literature demonstrating minimal structural effects on tightly folded proteins by concentrations of sucrose up to 2 M (~ 68% (w/v); in excess of the concentrations used in this study).[S7,S8] Sucrose does stabilize folded proteins, due to the exclusion of sucrose from the protein surface,[S8] and high concentrations encourage the compaction of more open and unfolded conformations.[S7] Thus, a rigid structure, such as that of the CaM-MLCK complex should be largely unaffected by high concentrations of sucrose, as our NMR data indicate. It is also worth noting that buffers containing significant amounts of hydroxyl-containing small molecules, such as glycerol, have long been used in macromolecular crystallography without adverse effects. For example, it has been shown that co-crystallization of lysozyme with trehalose, sorbitol or sucrose does not significantly affect either the protein structure or its first hydration layer.[S8] Moreover, di-saccharides such as trehalose have been recognized as stress response factors that help maintain protein structure when organisms are subjected to extreme environments.[S9]

The absence of effects of sucrose on the structure of the CaM/MLCK complex is further supported by the results of the exhaustive conformational search against the two types of SAXS data and the RDC and NOE NMR data. The general correlation of the quality of fit indicators in Fig. 2 of the main text and the very low discrepancy between observed and calculated data by the best-fitting Cluster I models indicate that the Pb-substituted contrast-matched scattering data is fully compatible with the rest of the experimental restraints. In addition, the agreement between the Pb/Pb separations derived from the contrast-matched Pb-substituted data alone with the corresponding distances in the best-fitting Cluster I further supports this observation (see Tables 1

and 3 of the main text). A more flexible or open system, such as peptide-free CaM, may necessitate additional validation, but is beyond the scope of the current study.

Our strategy of isolating a contribution from the metal sites within a macromolecule is in line with earlier work in which the distance probability distribution functions from gold nanoclusters attached to DNA were extracted from X-ray solution scattering data.[S10] Since the latter data, however, were collected in aqueous buffer, extraction of the gold/gold cross-term required multi-atom gold labels and data decomposition into three contributions (gold/gold, gold/DNA, DNA/DNA). In our case single-atom Pb^{2+} labels are much less intrusive and extraction of the metal/metal correlation function is straightforward when the X-ray scattering data are acquired at the contrast matching conditions.

Our strategy is also related to an earlier neutron resonance scattering study on ^{240}Pu-substituted CaM, which enabled extraction of the Pu/Pu site separation within the individual CaM domains.[S11] While contrast-matched SAXS from heavy atom labels permits only approximate cancellation of the contribution from the protein within a limited q-range, neutron resonant scattering of ^{240}Pu potentially ensures a more accurate isolation of the Pu/Pu cross-term. However, collection of the Pb-substituted contrast-matched X-ray scattering data is much less challenging and can done in a routine fashion since it does not require production and handling of the rare, highly radioactive ^{240}Pu isotope and usage of very short wavelength neutrons at $\lambda = 0.28$ Å to achieve resonance scattering conditions.

The current work capitalizes on the ability of CaM to specifically bind Pb^{2+} in place of Ca^{2+}. CaM binding to Pb^{2+} is a well-established phenomenon, and Pb^{2+} activates CaM and subsequently CaM target proteins similarly to Ca^{2+}.[S12,S13] Consistent with the latter observations, we observe very similar MLCK peptide binding to Ca^{2+}- and Pb^{2+}-bound CaM (Fig. S3). The primary difficulty encountered here in substituting Pb^{2+} for Ca^{2+} is the presence of additional non-specific Pb^{2+} binding sites outside of the four Ca^{2+} sites, as observed previously by crystallography.[S5] The NMR titration data, however, clearly demonstrate that the four Ca^{2+} sites have a much higher affinity for Pb^{2+} than the non-specific sites, and therefore under low concentrations of excess Pb^{2+} only the four specific binding sites are occupied (Fig. S1). In addition, the procedure used to analyze the contrast-matched SAXS data, by design, minimizes the impact of low-occupancy sites if such sites were present. Rather than extract the Pb/Pb $P(r)$ distribution from the scattering data, we restrict the fitted model to contain exactly 4 Pb atoms. The quality of the fit of the contrast-matched data in Fig. 3 of the main text indicates that this model is sufficient to describe the observed scattering intensities within experimental error. If additional binding sites were occupied to any significant extent, the Pb/Pb distances extracted via Monte Carlo sampling would not have been consistent with the Pb/Pb distances in the structure of the complex derived from the exhaustive conformational search against the aqueous and contrast-matched SAXS data and the NMR data. The excellent agreement between these values (within less than 1Å in the 4- and 5-variable cases; see Tables 1 and 3 of the main text) illustrates this point.

Although not performed in the current study, the methodology presented here can be readily extended to other proteins by incorporating heavy atom ions, such as Pb^{2+} or Hg^{2+}, into EDTA moieties conjugated via disulfide bonds to engineered surface cysteines. Such tagging is employed routinely in paramagnetic relaxation enhancement (PRE) studies (e.g. EDTA-Mn^{2+}).[S14] The heavy metal ions Pb^{2+} and Hg^{2+} have a higher affinity for EDTA than Mn^{2+} (or Mg^{2+} or Ca^{2+}), and pre-loading the tag with the heavy ion would have the added advantage of avoiding working with free heavy metal ions in the presence of protein. Although the EDTA moiety is connected to the protein by a flexible linker and hence the heavy atom ion can potentially sample a substantial region of conformational space, the fact that the metal-metal separations measured by SAXS are simple linear averages of all conformations present in solution means that each metal atom can be represented by a single average position. As a result analysis the contrast-matched SAXS data in these circumstances is in principle significantly more straightforward than the quantitative interpretation of PRE data arising from conjugated EDTA-Mn^{2+}, which generally requires an ensemble approach to take into account the $\langle r^{-6} \rangle$ averaging of the PRE.[S14]

Analysis of Pb-Pb distances from contrast-matched SAXS data

Fits of the contrast-matched Pb-substituted SAXS data were also performed using a random sampling of 4-atom geometries to represent the Pb sites within CaM. For this purpose, 10^6 random 4-atom geometries were generated using 4, 5, and 6 independent variables. The coordinates of the 4 atoms were represented as (0, 0, 0), (0, 0, r_{intra1}), ($r_1\sin(\theta_1)$, 0, $r_{intra1}+r_1\cos(\theta_1)$), ($r_1\sin(\theta_1)+r_{intra2}\sin(\theta_2)\cos(\phi_2)$, $r_{intra2}\sin(\theta_2)\sin(\phi_2)$, $r_{intra1}+r_1\cos(\theta_1)+r_{intra2}\cos(\theta_2)$). For the 4-variable case, r_1, θ_1, θ_2, and ϕ_2 are the sampled variables and $r_{intra1} = r_{intra2} = 11.7$ Å is the fixed intra-domain Pb-Pb distance. In the 5-variable case, $r_{intra1} = r_{intra2}$ becomes an additional variable, and in the 6-variable case, r_{intra1} and r_{intra2} are allowed to vary independently. The intervals of random uniform sampling for the variables are [0, π] for θ_1 and θ_2; [0, 2π] for ϕ_2; [0, 50Å] for r_1; and [0, 25 Å] for r_{intra1} and r_{intra2}. For each configuartion generated, the fit of the experimental contrast-matched SAXS data was performed as described in the text. For the fits that exhibited $\chi^2 < 0.81$, averages and standard deviations for the conformational variables were accumulated, with a success ratio of $\sim 10^{-3}$.

Comments on χ^2 values of fits to aqueous SAXS/WAXS data for Clusters I, II and III

It should be noted that the χ^2 values of the fits to the aqueous SAXS/WAXS data for all three clusters is ~1.5 which appears to be due to a small systematic discrepancy between q = 0.3 to 0.37 Å^{-1} (Fig. 3A, main text). This could be due to inadequacies of prediction of SAXS/WAXS curves from structural coordinates or unknown systematic errors in the SAXS/WAXS data, but more likely reflects small residual errors of the structural models. This is supported by the observation that none of the sampled structures display a $\chi^2_{SAXS,water}$ value below 1.37, and the finding of very small but possibly systematic differences in rhombicities when the RDC data are fitted to the N- and C-domains separately (Table S2). Both these observations could be accounted for by the existence of an ensemble of closely related structures that describe the Ca^{2+}-loaded CaM-MLCK complex. However, building such an ensemble is outside the scope of this paper since it would require not only a distribution of the interdomain geometries but also an ensemble representation of the individual domains of CaM to account for the previously noted possibility of residual scissor-type motions within the EF-hands.[S15] In addition, small differences in the rhombicities reported in Table S2 could also be due to the limited number and precision of the fitted RDCs, consistent with the intrinsically lower accuracy of the fitted rhombicities compared to the alignment tensor magnitudes.

The CaM-MLCK complex appears rigid based on all available experimental data.

Our procedure for determining the structure of the CaM-MLCK complex uses single-model fits, and therefore assumes a rigid architecture for the complex. This assumption is justified by the excellent agreement between the magnitudes and rhombicities of the alignment tensors fitted separately to the RDCs of the N- and C- domains of CaM (Table S2), as well as the appearance of commonly used scattering-based indicators of flexibility[S16] such as Kratky and Kratky-Porod plots (Fig. S7). Molecular volumes for the Ca^{2+}-loaded CaM-MLCK complex extracted from SAXS data in H_2O via the Porod formalism range from 24,700Å^3 to 25,900Å^3 and correspond to specific densities of 1.27-1.33 g/cm^3, also consistent with a folded rigid protein.[S17]

Table S1. Quality of the fits of the RDC and SAXS/WAXS data to previously determined structures of CaM-peptide complexes deposited in the RCSB Protein Databank. The entries with the best agreement to the experimental data are shown in bold typeface.

PDB accession Code	Resolution (Å)	RDC R-factor[a]			$\chi^2_{SAXS,water}$	$\chi^2_{SAXS,sucrose}$
		N-domain (6-71)	C-domain (83-144)	N+C (6-144)		
1J7O[b]	n.a. (NMR)	**0.19**	n,a.[b]	n.a.[b]	n.a.[b]	n.a[b]
1J7P[b]	n.a. (NMR)	n.a.[b]	**0.19**	n.a.[b]	n.a.[b]	n.a[b]
2BBM[c]	n.a. (NMR)	0.48	0.31	0.45	4.70	1.20
1CDL:A	2.0	0.26	0.24	0.27	2.06	1.08
1CDL:B		0.23	**0.18**	**0.21**	2.21	**0.96**
1CDL:C		0.30	0.22	0.27	2.23	**0.79**
1CDL:D		0.28	0.28	0.32	2.13	**0.88**
1CDM	2.0	0.22	0.22	0.31	n.a.[d]	n.a.[d]
1CM1	2.0	0.21	0.22	0.30	2.09	1.23
1CM4:1	2.0	0.34	0.28	0.41	**1.61**	**0.99**
1CM4:2		0.26	0.27	0.32	2.02	1.51
1CM4:3		0.24	0.30	0.33	2.09	2.72
1CM4:4		0.30	0.25	0.35	3.11	1.08
1MXE:A	1.7	**0.17**	**0.17**	**0.21**	**1.78**	1.29
1MXE:B		**0.18**	**0.18**	**0.20**	**1.73**	1.50
1N1W:A	2.05	0.25	0.21	0.33	2.32	1.13
1N1W:C		0.32	**0.17**	0.33	2.16	1.23
1N1W:E		0.28	0.23	0.33	2.83	1.77
1N1W:G		0.27	0.20	0.29	2.13	1.77
1QS7:A	1.8	**0.17**	**0.19**	**0.19**	2.24	**0.85**
1QS7:C		**0.18**	**0.18**	**0.20**	2.30	**0.80**
1QTX	1.65	**0.20**	**0.19**	**0.20**	2.26	1.06
1VRK	1.9	**0.16**	0.26	0.24	3.50	1.20
1YR5	1.7	0.25	**0.16**	0.32	2.40	3.71
1ZUZ	1.91	0.22	0.19	0.28	**1.69**	2.82
2BCX	2.0	**0.19**	0.33	0.50	17.6	10.9
2BE6;A	2.0	**0.19**	**0.17**	0.24	2.92	2.07
2BE6:B		**0.18**	0.24	**0.21**	9.50	2.70
2BE6:C		0.23	**0.19**	0.26	**1.68**	1.60
2F3Y	1.45	**0.18**	**0.17**	**0.22**	1.70	1.42
2FOT	2.45	0.29	0.22	0.33	4.86	3.42
2HQW	1.9	**0.20**	**0.18**	0.37	3.31	**0.79**
2O5G	1.08	0.20	0.25	0.29	n,a.[c]	n.a.[c]
2O60	1.55	0.23	0.31	0.29	**1.61**	1.38
2VAY	1.94	**0.20**	0.22	0.30	**1.75**	1.10
3BXL	2.3	0.22	0.23	0.27	2.16	1.28
3DVE	2.35	0.21	0.22	0.38	2.11	**0.85**
3DVJ	2.8	0.24	0.20	0.40	2.17	**0.76**
3DVK	2.3	0.25	0.20	0.40	3.08	4.96
3DVM	2.6	0.33	**0.17**	0.44	3.85	4.13
3EWT	2.4	**0.19**	**0.19**	0.33	5.22	**0.89**
3EWV	2.6	0.25	0.23	0.40	5.13	**0.71**
3G43	2.1	**0.18**	0.24	0.27	**1.79**	1.30
3GN4:B	2.7	0.60	0.63	2.21	17.3	37.3
3GN4:D		0.36	0.24	0.32	1.98	3.01
3GN4:F		0.36	0.30	0.35	2.18	2.70
3GN4:H		0.56	0.61	2.01	20.1	3.01
3GP2	1.46	0.65	**0.19**	0.58	2.68	**0.81**
3OXQ:A	2.55	0.27	0.24	0.34	2.03	1.33
3OXQ:C		0.27	0.22	0.34	2.51	1.53

[a]Fits were carried out using an alignment tensor fitted to the structural coordinates by singular value decomposition (SVD).

[b]1J7O and 1J7P are the NMR coordinates of the N- and C-domains, respectively, of Ca^{2+}-loaded CaM in the absence of peptide.[S15] Hence the X-ray scattering data cannot be fitted since the two domains of CaM, in the absence of ligand, tumble essentially independently of one another in solution, and therefore their relative orientation could not be determined.

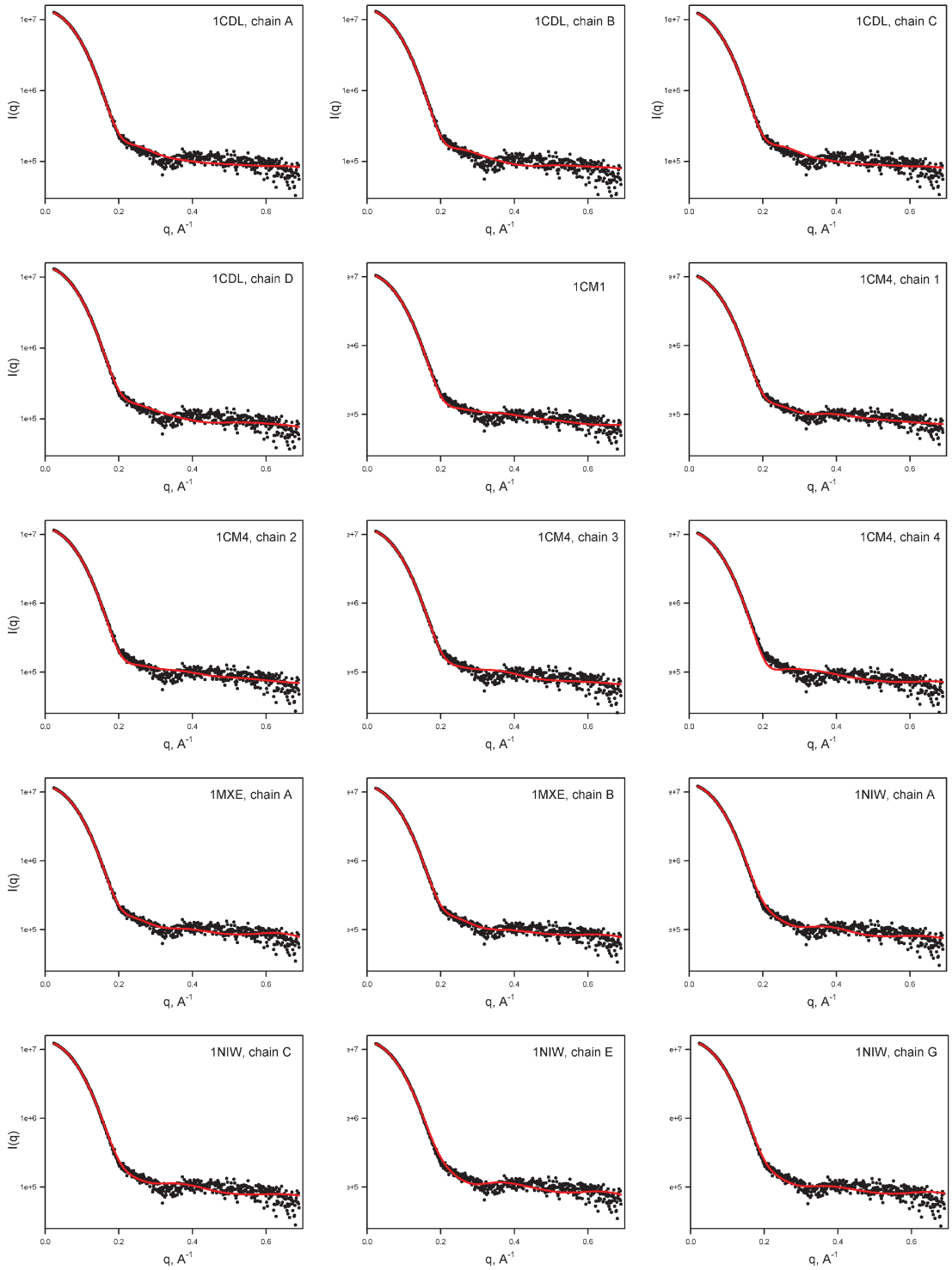
[c]NMR coordinates of Ca^{2+}-loaded CaM-MLCK complex.[S18]

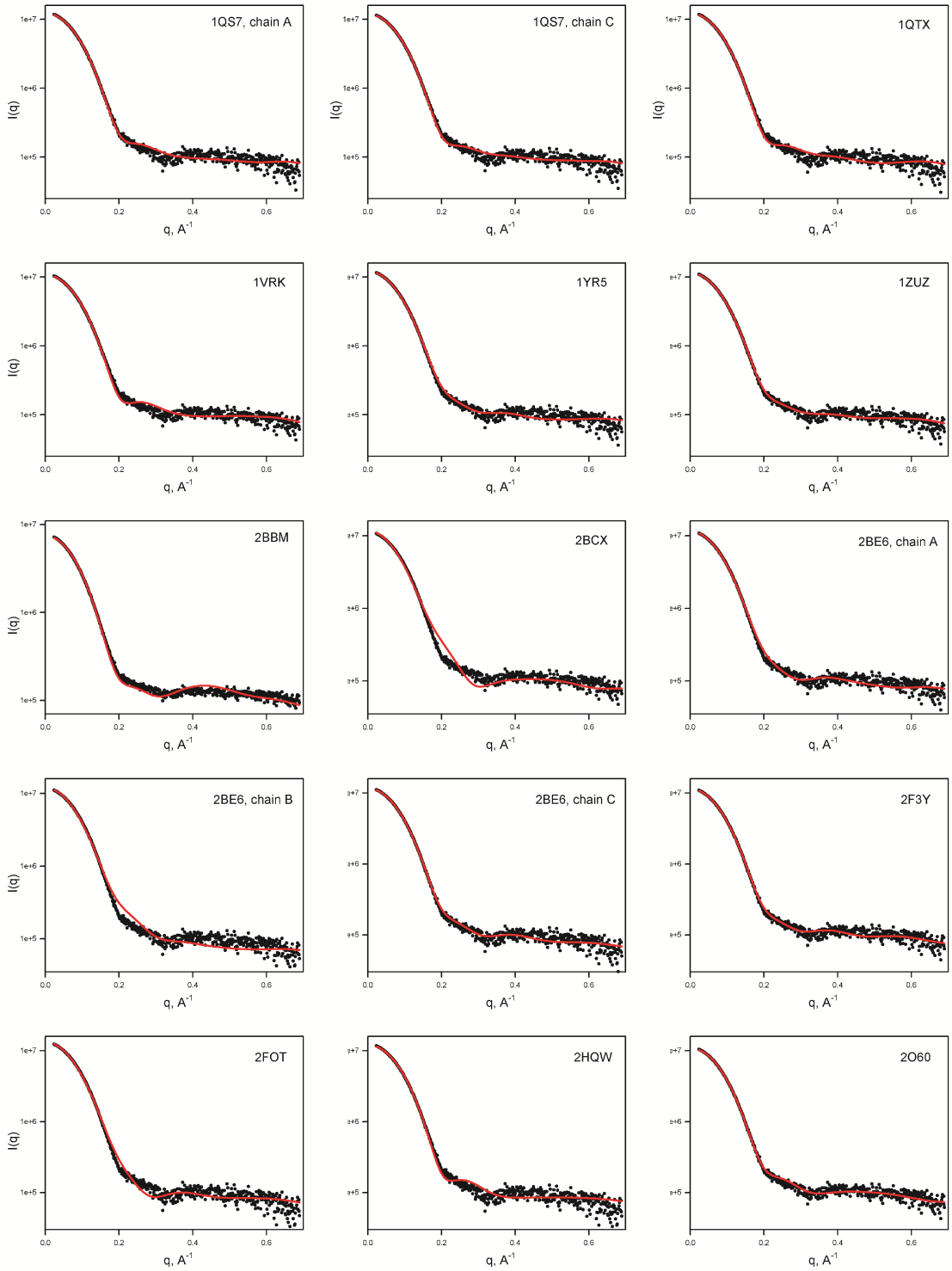
[d]X-ray scattering data cannot be fitted accurately from the coordinates of these Ca^{2+}-loaded CaM-peptide complexes as a significant fraction of residues are missing from the PDB depositions.

Table S2 Alignment tensor parameters obtained by SVD fitting of the $^{1}D_{NH}$ RDC data recorded on the Ca^{2+}-loaded CaM-MLCK complex to the structures of the individual N and C domains of CaM. The magnitude (D_a^{NH}) and rhombicities (η) of the alignment tensors are provided for several best-fitting models.

Model	N-domain (6-71)		C-domain (83-144)	
	D_a^{NH} (Hz)	η	D_a^{NH} (Hz)	η
1MXE	-17.3	0.32	-17.6	0.39
1QS7	-17.4	0.35	-17.2	0.45
2BE6	-17.3	0.33	-17.4	0.38

Fig. S4 Fits to the SAXS/WAXS data collected on the Ca^{2+}-loaded CaM-MLCK complex in aqueous buffer for the crystal structures of the CaM-peptide complexes listed in Table S1 (3 pages). The slight variation in the level of the high-q data in the plots is due to the fact that measured data are fitted to the scattering intensities predicted from each structural model as $I_{measured}(q) = I_{sample}(q) - \alpha I_{buffer}(q) + const$, where both α and $const$ are optimized in each fit between the measured and calculated data, as proposed previously.[S19]




1QS7, chain A
1QS7, chain C
1QTX
1VRK
1YR5
1ZUZ
2BBM
2BCX
2BE6, chain A
2BE6, chain B
2BE6, chain C
2F3Y
2FOT
2HQW
2O60
I(q)
q, A^{-1}

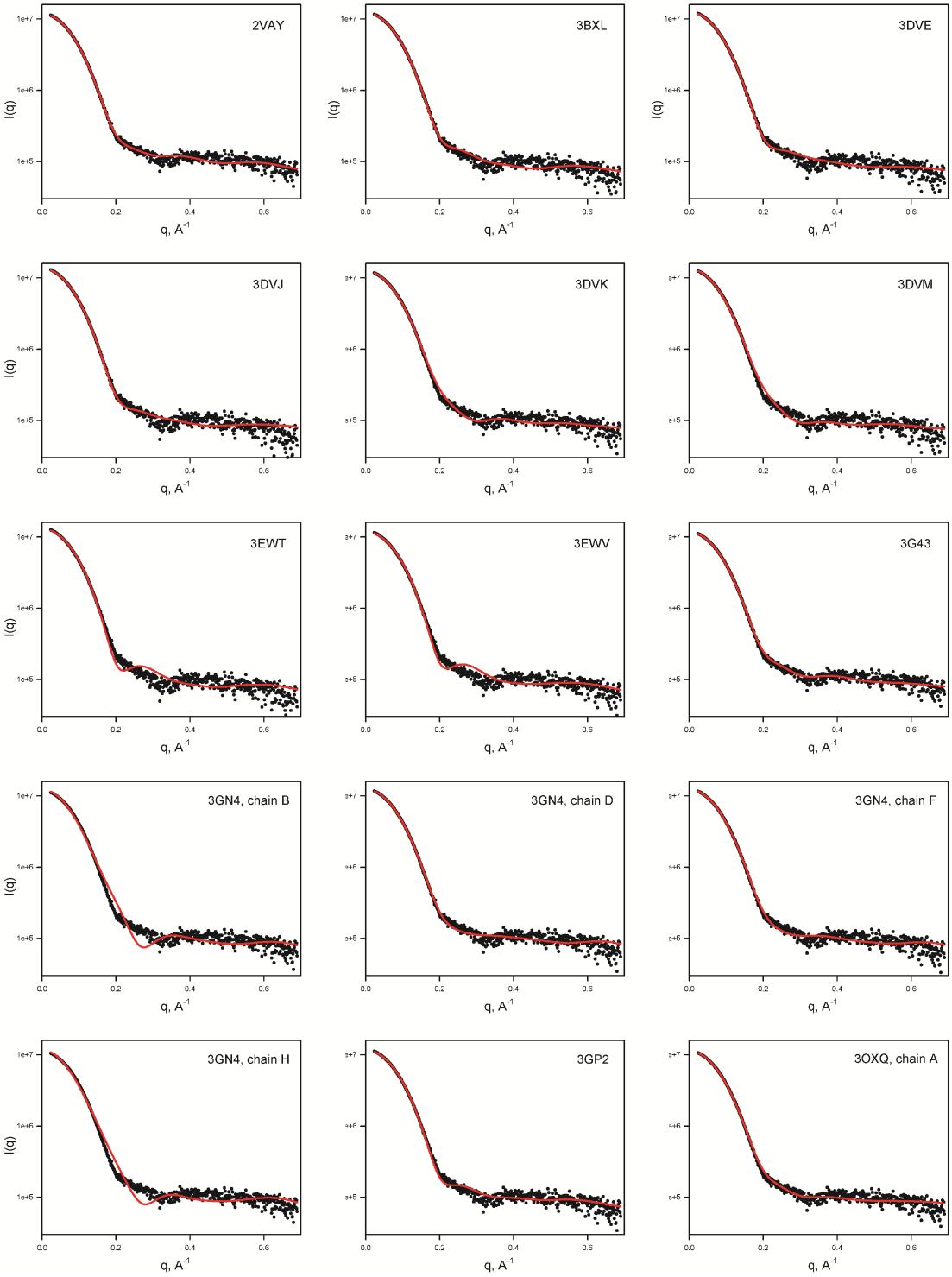
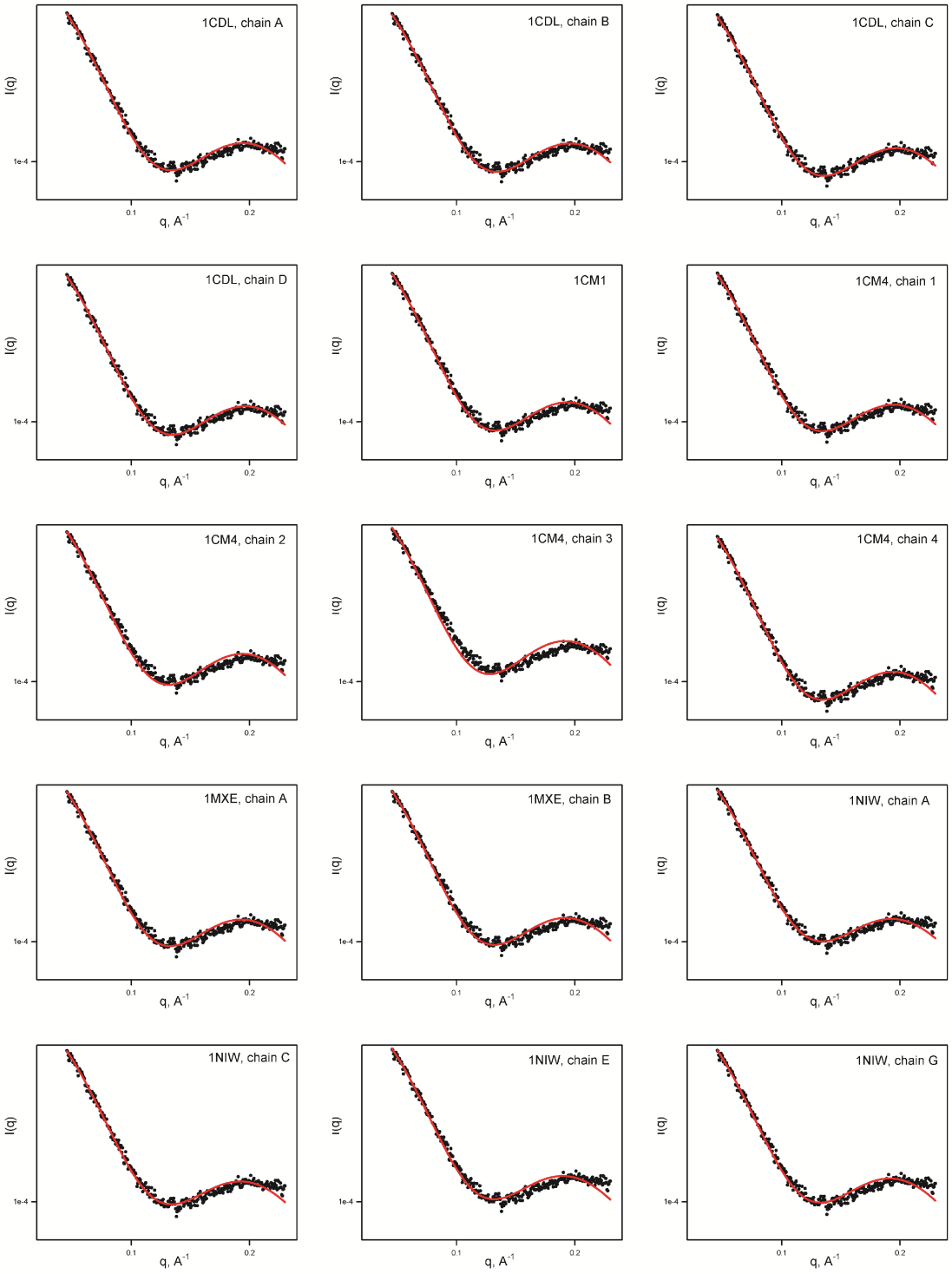
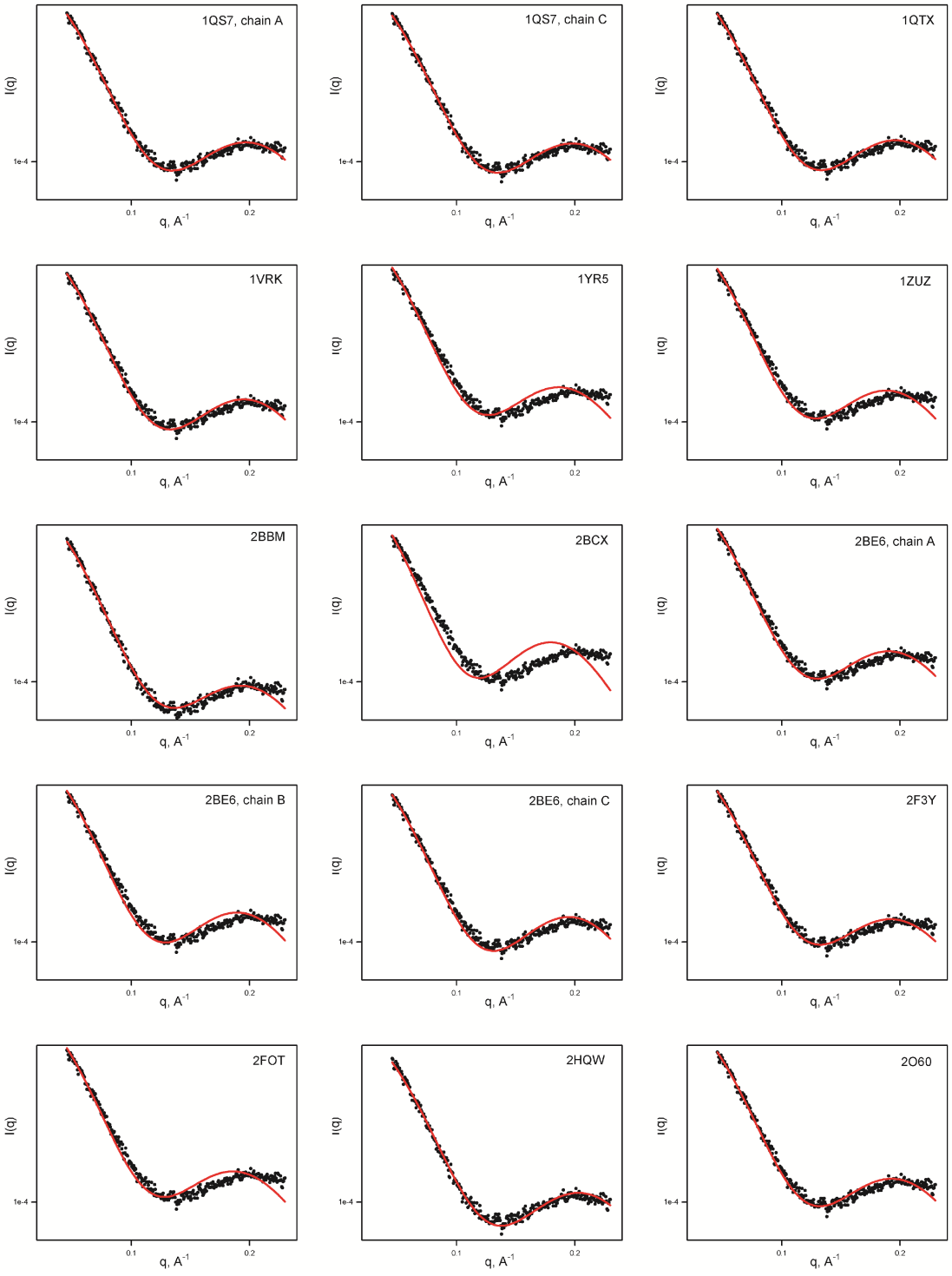
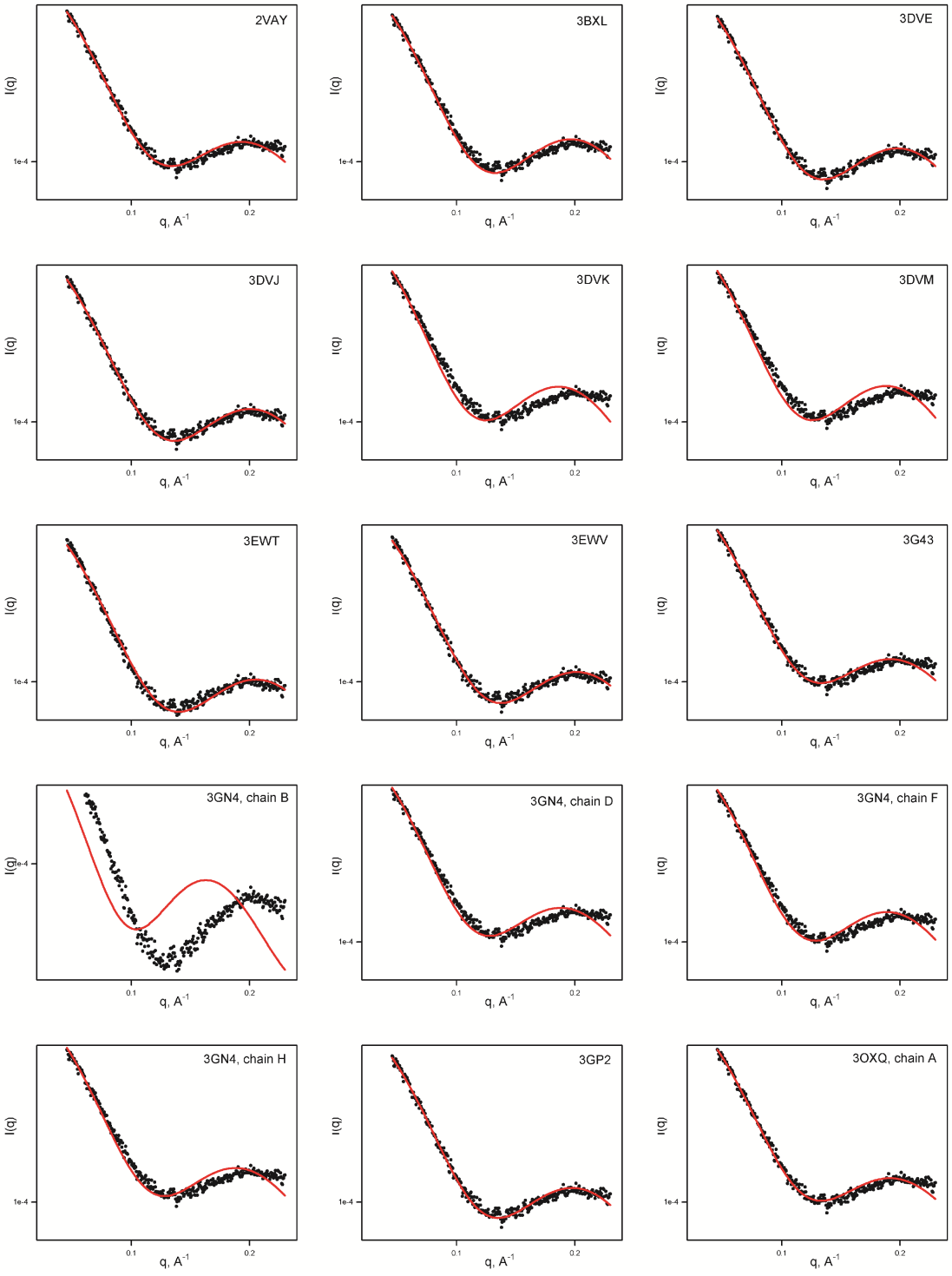

2VAY
3BXL
3DVE
3DVJ
3DVK
3DVM
3EWT
3EWV
3G43
3GN4, chain B
3GN4, chain D
3GN4, chain F
3GN4, chain H
3GP2
3OXQ, chain A
I(q)
q, A-1

Fig. S5. Fits to the contrast-matched SAXS data collected on the Pb^{2+}-loaded CaM-MLCK complex in 65% (w/v) sucrose buffer for the crystal structures of CaM-peptide complexes listed in Table S1 (3 pages). The slight variation in the level of the higher-q data in the plots is due to the fact that measured data are fitted to the scattering intensities predicted from each structural model as $I_{measured}(q) = I_{sample}(q) - \alpha I_{buffer}(q) + const$, where both α and $const$ are optimized in each fit between the measured and predicted data, as proposed previously.[S19]




1QS7, chain A
1QS7, chain C
1QTX
1VRK
1YR5
1ZUZ
2BBM
2BCX
2BE6, chain A
2BE6, chain B
2BE6, chain C
2F3Y
2FOT
2HQW
2O60
I(q)
1e-4
0.1
0.2
q, A-1


2VAY
3BXL
3DVE
3DVJ
3DVK
3DVM
3EWT
3EWV
3G43
3GN4, chain B
3GN4, chain D
3GN4, chain F
3GN4, chain H
3GP2
3OXQ, chain A
I(q)
1e-4
0.1
0.2
q, A-1

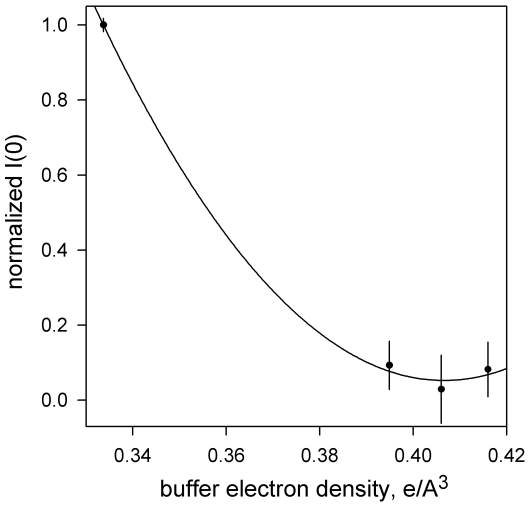


Fig. S6. Optimization of the contrast-matching conditions for the CaM-MLCK complex based on the variation of the zero-angle scattering intensity as a function of sucrose concentration in the buffer. Zero-angle scattering intensities are extracted via Guinier fits to the SAXS data for 5 mg/mL CaM-MLCK samples at 25°C with 0%, 55%, 65% and 75% w/v sucrose buffers. Corresponding electron densities of the buffers are displayed on the *x* axis. The *y* axis shows *I*(0) values normalized to the value in pure H_2O. Average and standard deviations are shown for the 4 measurement points. A parabolic fit is displayed as a solid line.

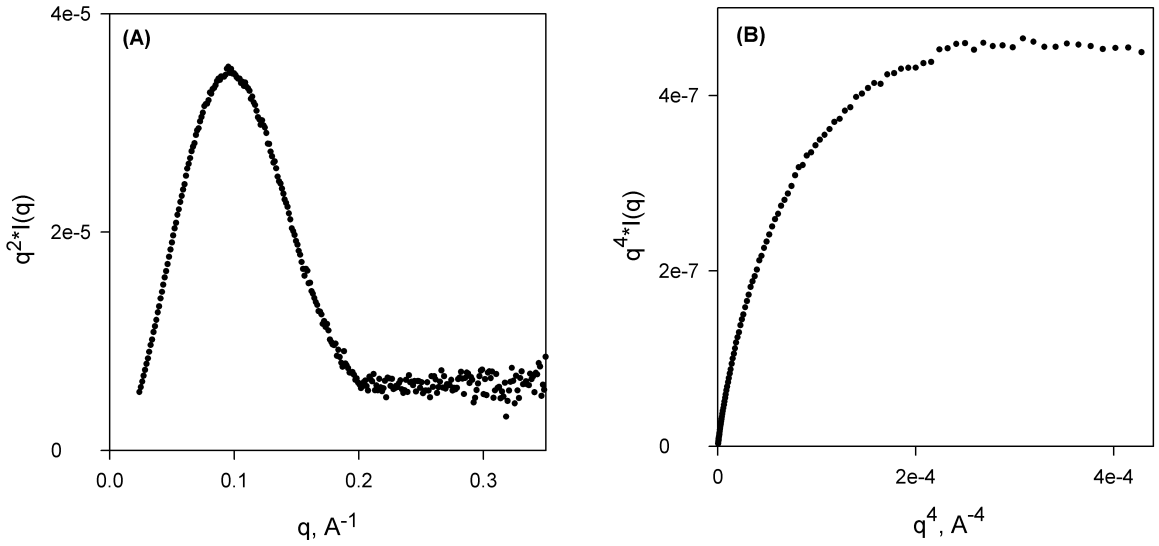


Fig. S7. Kratky (A) and Kratky-Porod (B) plots for the Ca^{2+}-loaded CaM-MLCK complex in H_2O at a concentration of 5mg/mL. The presence of a plateau at higher *q* in both plots is consistent with a rigid structure for the complex.

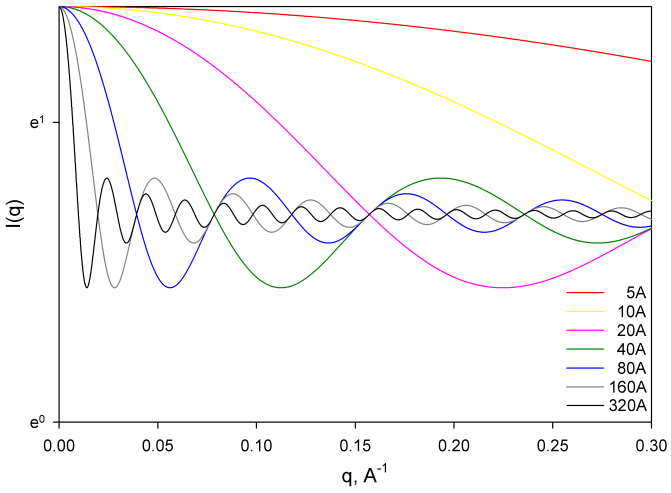


Fig. S8. The effect of variation of the distance between two scatterers on the resulting scattering profile. Simulated profiles are shown for two Pb scatterers at the protein contrast-matching conditions. The distance between the scatterers is varied between 5Å and 320 Å. The periodicity of the oscillations allows extraction of the Pb-Pb distances from the scattering data. The ability to extract short distances is limited by the q range within which the contrast-matching conditions are satisfied (q_{max} = 0.2-0.3 Å^{-1}). Accurate extraction of the long Pb-Pb distances is limited by the signal-to-noise of the experimental data.

Complete citation for ref. 1c of the main text:

Hura, G. L.; Menon, A. L., Hammel, M., Rambo, R. P.; Poole, F. L.; Tsutakawa, S. E.; Jenney, F. E.; Classen, S.; Frankel, K. A.; Hopkins, R. C.; Yang, S. J.; Scott, J. W.; Dillard, B. D.; Adams, M. W.; Tainer, J. A. *Nature Methods* **2009**, *6*, 606-612.

Supplementary References

(S1) Anthis, N. J.; Doucleff, M.; Clore, G. M. *J Am Chem Soc* **2011**, *133*, 18966.
(S2) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(S3) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J Am Chem Soc* **1998**, *120*, 10571.
(S4) Fitzkee, N. C.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 65.
(S5) Kursula, P.; Majava, V. *Acta Cryst.* **2007**, *F63*, 653.
(S6) Cioni, P.; Bramanti, E.; Stambini, G. B. *Biophys. J.* **2005**, *88*, 4213.
(S7) Lee, J. C.; Timasheff, S. N. *J. Biol. Chem.* **1981**, *256*, 7193.
(S8) Datta, S.; Biswal,B. K..; VijayaN, M. *Acta Cryst.* **2001**, *D57*, 1614.
(S9) Jain, N. K.; Roy, I. *Protein Sci.* **2009**, *18*, 24.
(S10) Mathew-Fenn, R. S.; Das, R.; Silverman, J. A.; Walker, P. A.; Harbury, P. A. *PLoS One* **2008**, *3*, e3229.
(S11) Seeger, P. A.; Rokop, S. E.; Palmer, P. D.; Henderson, S. J.; Hobart, D. E.; Trewhella, J. *J Am Chem Soc* **1997**, *119*, 5118.
(S12) Chao, S. -H.; Suzuki, Y.; Zysk, J. R.; Ceung, W. Y. *Mol,. Pharm.* **1984**, *26*, 75/
(S13) Chao, S. -H.; Bu, C. H., Ceung, W. Y. *Arch. Toxicol.* **1995**, *69*, 197.
(S14) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108.

(S15) Chou, J. J.; Li, S.; Klee, C. B.; Bax, A. *Nature Struct. Biol.* **2001**, *8*, 990.
(S16) Rambo, R. P.; Tainer, J. A. *Curr. Opin. Struct. Biol*. **2010**, *20*, 128.
(S17) Rambo, R. P.; Tainer, J. A. *Biopolymers*, **2011**, *95*, 559.
(S18) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632.
(S19) Grishaev, A.; Guo, L.; Irving, T.; Bax, A. *J Am Chem Soc* **2010**, *132*, 13026.