



Cite this: *Phys. Chem. Chem. Phys.*,
2021, 23, 21690

Lanthanide-dependent coordination interactions in lanmodulin: a 2D IR and molecular dynamics simulations study†

Stephanie Liu,^a Emily R. Featherston,^b Joseph A. Cotruvo Jr. ^{*b} and
Carlos R. Baiz ^{*a}

The biological importance of lanthanides, and the early lanthanides (La³⁺–Nd³⁺) in particular, has only recently been recognized, and the structural principles underlying selective binding of lanthanide ions in biology are not yet well established. Lanmodulin (LanM) is a novel protein that displays unprecedented affinity and selectivity for lanthanides over most other metal ions, with an uncommon preference for the early lanthanides. Its utilization of EF-hand motifs to bind lanthanides, rather than the Ca²⁺ typically recognized by these motifs in other proteins, has led it to be used as a model system to understand selective lanthanide recognition. Two-dimensional infrared (2D IR) spectroscopy combined with molecular dynamics simulations were used to investigate LanM's selectivity mechanisms by characterizing local binding site geometries upon coordination of early and late lanthanides as well as calcium. These studies focused on the protein's uniquely conserved proline residues in the second position of each EF-hand binding loop. We found that these prolines constrain the EF-hands for strong coordination of early lanthanides. Substitution of this proline results in a more flexible binding site to accommodate a larger range of ions but also results in less compact coordination geometries and greater disorder within the binding site. Finally, we identify the conserved glycine in the sixth position of each EF-hand as a mediator of local binding site conformation and global secondary structure. Uncovering fundamental structure–function relationships in LanM informs the development of synthetic biology technologies targeting lanthanides in industrial applications.

Received 6th August 2021,
Accepted 18th September 2021

DOI: 10.1039/d1cp03628a

rsc.li/pccp

Introduction

In the last decade, certain members of the lanthanide series have been revealed to play unexpected, yet critical biological roles in the growth of methylotrophic (and some non-methylotrophic) bacteria.^{1–3} Beginning with the characterization of the first lanthanide-dependent enzyme, a methanol dehydrogenase,^{4–6} recent work has identified the first selective lanthanide-binding protein, lanmodulin (LanM),^{7,8} as well as a system for lanthanide acquisition.^{9–12} Understanding how nature has evolved to use lanthanide elements – in particular, the light or “early” lanthanides (La–Nd) that organisms tend to prefer – is not only essential to uncovering the roles of lanthanides in biochemistry, but it could also lead to new

biomolecule-based approaches for lanthanide separations and recycling in industrial applications.¹³

LanM is of particular interest due to its ability to selectively bind lanthanides *in vivo*,⁷ along with its utility in isolating lanthanides from other metals in industrial waste streams.¹⁴ LanM is a unique EF-hand-containing protein (Fig. 1A) with picomolar affinity and high selectivity for the lanthanide series, even over more abundant and available ions with similar ionic radii such as Ca²⁺.⁷ EF-hands consist of an entering helix, a ~12-residue metal-binding loop, and an exiting helix.¹⁵ These metal-binding motifs are typically associated with Ca²⁺ binding proteins, such as calmodulin (CaM), the ubiquitous Ca²⁺ sensor protein.¹⁶ Most EF-hands demonstrate a marginal preference for lanthanides over Ca²⁺,^{17–19} due to their higher charge. This feature has frequently been exploited in the laboratory as a method for the investigation of calcium binding.^{20–23} In the case of LanM, there is a remarkable 100-million fold selectivity for lanthanide binding over Ca²⁺ and other divalent metal ions.⁷ This makes LanM an ideal model for understanding the relationship between EF-hand geometry and ion selectivity, as well as a scaffold for developing synthetic biological

^a Department of Chemistry, University of Texas at Austin, Austin, TX 78712, USA.
E-mail: cbaiz@cm.utexas.edu

^b Department of Chemistry, The Pennsylvania State University, University Park,
PA 16802, USA. E-mail: juc96@psu.edu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cp03628a

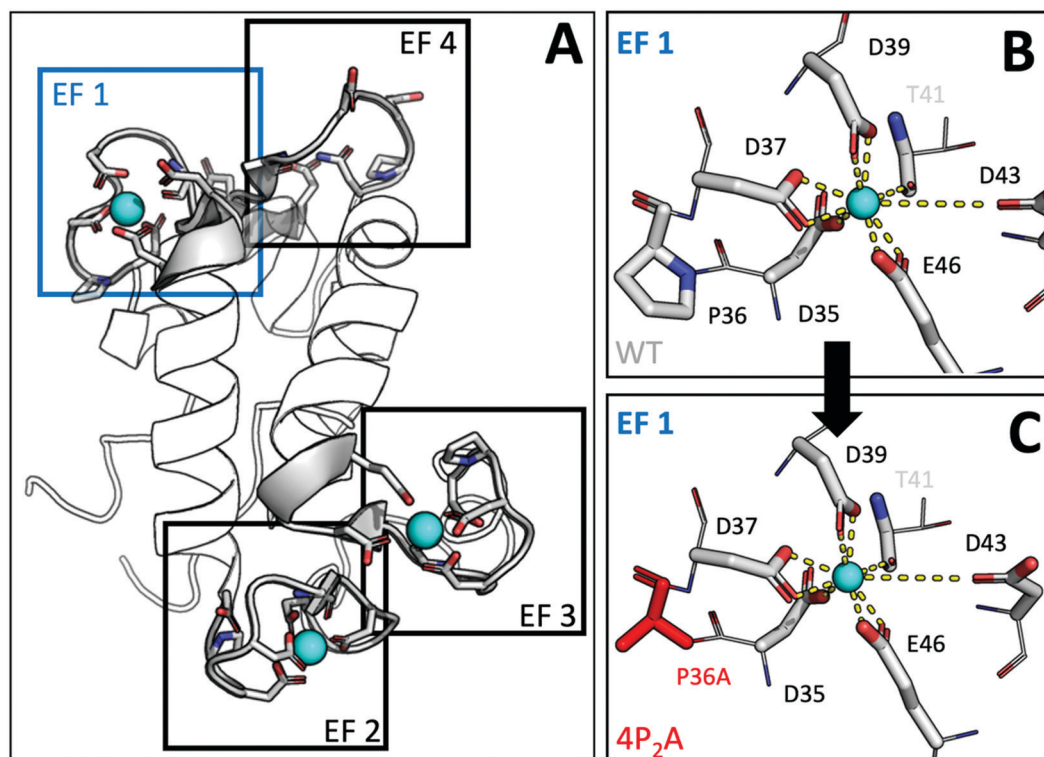


Fig. 1 Cartoon representations of LanM bound to Y^{3+} . (A) NMR solution structure of LanM (PDB: 6MI5), with each of the four EF-hand binding sites highlighted in blue (EF 1) and black (EF 2–4). The fourth EF-hand remains uncoordinated in this structure. (B) Structure of WT EF1, as modeled in the NMR structure. The coordinated ion is shown in cyan, and coordinating oxygens are connected by dashed yellow lines. Residues with coordinating carboxylate sidechains are labeled in black. Threonine residues that coordinate with a backbone carbonyl are labeled in gray, as they are not included in our analysis. The coordinating residue sequence is the same across all four EF-hands, with the exception of EF4 which has an Asp in the first position. (C) Generated structure of $4P_2A$ EF1 variant used as the starting point for MD simulations. The Pro $\rightarrow$ Ala mutation was produced in Pymol and is highlighted in red. This Pro $\rightarrow$ Ala mutation was also applied to the other three EF hands to determine the sequence-dependent geometry.

approaches for isolating lanthanides from other metals.^{13,14} In particular, the structural principles underlying the significant affinity difference are not understood. Characterizing these principles deepens our understanding of lanthanide selectivity in biology and lays a foundation for designing LanM analogs with tailored ion-specific affinities. There are a few relevant observations that serve as a starting point for rationalizing the measured affinities: LanM undergoes a significant conformational change from disordered to ordered upon binding of three Ln^{3+} ions, to its EF-hands 1, 2, and 3 (the fourth EF-hand has very low lanthanide affinity). In the presence of Ca^{2+} , the protein's conformational change is only partial and does not occur until nearly millimolar Ca^{2+} concentrations.⁷ The NMR solution structure of the Y^{3+} -bound LanM⁸ suggests that LanM's metal sites may bind with a higher coordination number (8 or 9, depending on the site) compared to typical EF-hand/ Ca^{2+} interactions (7 coordinate).¹⁵ However, this method was unable to provide atomic-level detail of these binding sites, limiting the specific conclusions that can be drawn about lanthanide selectivity.

The LanM EF-hand sequence is unique in several regards. Perhaps most distinctive is that each of LanM's four EF-hands contains a proline in the second position of the binding loop, which is the first time this has been observed out of the

hundreds of characterized EF-hand proteins.⁷ This notable deviation from the highly conserved standard EF-hand motif²⁴ is thought to be a significant contributor to the lanthanide selectivity of LanM. This hypothesized function was initially investigated by constructing a variant of LanM with these four prolines mutated to alanine ($4P_2A$, Fig. 1B and C). Whereas $4P_2A$ -LanM shows the same conformational response with similar apparent K_d values as the WT protein for La^{3+} and Nd^{3+} ions, this substitution both tightens the apparent K_d for Ca^{2+} binding and allows Ca^{2+} to induce the same, full conformational response as to Ln^{3+} ions.⁸ These studies supported the proposed role of these prolines in lanthanide selectivity, but leave unanswered how the proline impacts metal ion coordination within the EF-hand.

Vibrational spectroscopy directly accesses protein structure and conformational dynamics, including binding geometries, and is an established technique for investigating ion-dependent changes in coordination structure.^{22,25,26} Specifically, the EF-hand glutamic and aspartic acid carboxylate vibrations, together with those of the backbone and side-chain amide carbonyls, are inherently spectroscopically sensitive to the ion radius, charge, and coordination geometry. The carbonyl stretching frequencies shift in response to changes in the local electrostatic environment.^{27–29} Furthermore, EF-hands natively contain both

monodentate and bidentate carboxylate coordination geometries, the populations of which will shift depending on the charge and size of the coordinating ion. For example, these methods have revealed that Ln^{3+} -binding perturbs CaM's binding site from its native Ca^{2+} -binding geometry, with a decrease in bidentate coordination and an increase in structural disorder.²² Simulations of CaM bound to Eu^{3+} and Cm^{3+} also point towards increase in disorder, albeit with the opposite trend in bidentate coordination.³⁰ This contrast embodies the challenges of simulation and emphasizes the necessity of one to one comparisons between simulation and experiment. The accurate simulation of highly charged ions like trivalent lanthanides is a non-trivial aspect of this approach. Here we use Ln^{3+} force field parameters developed by Migliorati and coworkers³¹ to simulate LanM complexed to three ionic conditions—divalent calcium, an early lanthanide represented by La^{3+} , and a late lanthanide represented by Lu^{3+} . These parameters were recently used to elucidate the structural characteristics, particularly water's role, underlying a lanthanide binding tag (LBT) variant's binding selectivity for certain lanthanides.³²

In the present study, we focus on sequence and metal ion-dependent differences in LanM by mapping the coordination geometries as reflected in IR lineshapes. We examine the biologically relevant early Ln^{3+} ions that LanM binds with the highest affinity (e.g., La^{3+} , Pr^{3+} , and Nd^{3+}), and the late Ln^{3+} ions that LanM also recognizes selectively but with weaker affinity (e.g., Tb^{3+} and Lu^{3+}), and Ca^{2+} , the typical substrate of EF-hands. We compare metal binding to WT LanM and to the 4P₂A variant, which is more Ca^{2+} responsive than the WT protein. These vibrational spectroscopy measurements are interpreted through simulations, which produce an atomistic representation of the ion size and charge effects on coordination. This approach directly links coordination ensembles with the high lanthanide affinities of LanM relative to well-studied EF-hand proteins, revealing ion-dependent structural characteristics underlying LanM's remarkable selectivity.

Methods

Preparation of LanM for spectroscopic measurements

Lanmodulin (LanM), both WT and the 4P₂A variant, were expressed and purified according to an established protocol, with Chelex treatment and exchange into 20 mM MOPS, pH 7.0.⁷ Purified LanM was frozen in liquid N_2 and lyophilized. The dried protein was deuterated by reconstituting in pure D_2O and incubating at room temperature for two hours before being frozen and lyophilized again. The deuterated LanM/MOPS lyophilisate was then dissolved in D_2O and pH-adjusted with DCl (Sigma-Aldrich, 99 atom% D) and NaOD (Sigma-Aldrich, 99 atom% D) to a pH reading of 6.5 uncorrected for deuterium effects. Protein buffer solutions were then combined with solutions of anhydrous metal chlorides (CaCl_2 , LaCl_3 , PrCl_3 , NdCl_3 , TbCl_3 , and LuCl_3) in 50 mM deuterated MOPS/ D_2O to give a final solution of 1.5 mM LanM and 5 mM metal ion to ensure saturation of LanM's high-affinity binding sites, even in

the case of Ca^{2+} ($K_{\text{d,app}} = 710 \mu\text{M}$ for WT Ca^{2+} -LanM).⁷ This final solution was once again adjusted to an uncorrected pH reading of 6.5.

Two-dimensional infrared (2D IR) spectroscopy

2D IR spectra were collected with a custom-built 2D IR spectrometer described previously.²² Briefly, the protein vibrational modes are excited by a pair of pump pulses resonant with the carboxylate asymmetric stretch, which are then probed by a third probe pulse. The excitation frequency is obtained by numerical Fourier transformation of the time delay between the two pump pulses, while the detection frequency axis is generated by measuring the probe and signal in the frequency domain using a grating-based spectrometer. Plotting the amplitude of the 2D IR signal for each pump frequency gives a spectrum analogous to absorption spectra. This pump-slice-amplitude (PSA) analysis method is described elsewhere.³³ All 2D IR spectra were collected at a fixed pump-probe waiting time of 150 fs. Pump and probe polarizations were set to the perpendicular geometry. Spectra were averaged over 5 million shots, or ~ 1.5 hours of acquisition. Pump scatter was minimized through phase cycling using a pulse shaper.

Computational methods

Simulations were carried out using CHARMM36 force field³⁴ and the TIP3P water model³⁵ as implemented in the GROMACS 2019.4 package.³⁶ Simulations were carried out starting with the Y^{3+} -bound NMR structure (PDB: 6MI5) which was energy minimized and equilibrated in a box containing ~ 6000 water molecules. Sodium ions were added to maintain charge neutrality. The structure of the 4P₂A-LanM variant was generated using the Pymol mutagenesis function starting from the same initial NMR configuration. Three MD potentials for different metal ion-bound forms were simulated in complex with LanM—a calcium ion using the CHARMM36 parameters for Ca^{2+} , as well as representative early and a late lanthanide ion generated using parameters for La^{3+} ($\sigma_{\text{LnO}} = 3.310$, $\epsilon_{\text{LnO}} = 0.4308$) and Lu^{3+} ($\sigma_{\text{LnO}} = 2.997$, $\epsilon_{\text{LnO}} = 0.4308$) developed by Migliorati *et al.*³¹ The simulation box was equilibrated for 10 ns to solvate the protein before a 50 ns NPT production at 300 K and 1 bar using the Nose-Hoover and Parrinello-Rahman thermostat and barostat respectively.^{37,38} Equilibration and convergence were confirmed *via* comparison of protein-ion distance distributions in the first and second halves of the production trajectory.

Results and discussion

2D IR spectroscopy

2D IR spectroscopy is a useful technique to investigate protein structure and dynamics and has been applied to the study of metal-binding proteins such as CaM.^{39–43} Briefly, a 2D IR spectrum is represented along excitation and detection frequency axes. Lineshapes contain rich structural information.⁴⁴ Here we analyze spectra by extracting the

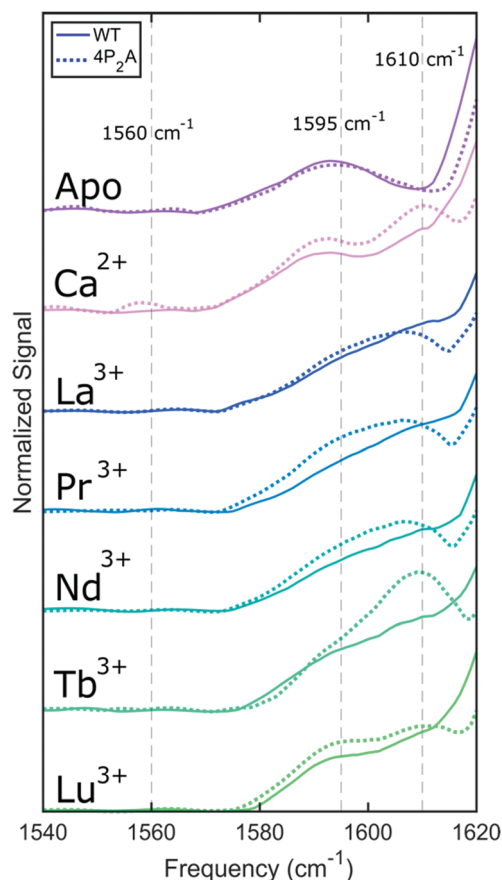


Fig. 2 Pump slice amplitude spectra of WT (solid line) and 4P₂A LanM (dashed line) complexed to different metal ions. Vertical dashed lines indicate observed features: a bidentate feature at 1560 cm⁻¹, unbound carboxylate centered around 1595 cm⁻¹, and monodentate carboxylate near 1610 cm⁻¹. Spectra are arranged in decreasing ionic radius from upper to lower, except for Ca²⁺, the only divalent and non-lanthanide metal ion investigated.

excitation-axis contribution to the spectra using a pump-slice method. Pump-slice spectra (Fig. 2) were extracted from experimental 2D IR spectra of wild-type (WT) LanM and the (4P₂A) variant complexed with 3 equivalents of a representative series of early and late lanthanide ions, as well as Ca²⁺ (Fig. S1 and S2, ESI†). These spectra are analogous to IR absorption spectra and can be interpreted similarly.³³ Given the few binding site carbonyls in each protein, and the concentration range used in protein studies, obtaining adequate signal-to-noise ratio and proper background subtraction remains a challenge for IR absorption measurements, making 2D IR a useful alternative.^{43,45,46} In combination with computational methods, 2D IR tracks the metal ion-dependent changes in coordination geometries and conformational disorder.

Fig. 2 shows spectra of WT and 4P₂A LanM metal complexes, revealing marked differences in the carboxylate asymmetric stretch region, 1540–1620 cm⁻¹, both between different ions and between the WT and 4P₂A variants. Broadly, the WT protein bound to the early lanthanides (La³⁺, Pr³⁺, Nd³⁺) exhibits comparable line shapes that contrast against the spectra of

the protein bound to Ca²⁺ and late lanthanides, Tb³⁺ and Lu³⁺. The spectra of the metal-bound 4P₂A variant differ from those of the WT but similarly fall into the groups of early Ln³⁺ versus late Ln³⁺ and Ca²⁺. The apo spectra for both WT and the 4P₂A variant are similar, showing a single band centered at around 1595 cm⁻¹ that is assigned to an unbound carboxylate asymmetric stretch. This feature can be observed in all of the spectra presented in the figure but is most pronounced for the late Ln³⁺ ions and Ca²⁺ with the WT protein, and generally becomes enhanced in the 4P₂A variant. EF-hand 4, which remains predominantly unbound due to its lower metal ion-affinity, may contribute to this unbound feature.⁸ Metal-bound WT and 4P₂A LanM produce a second feature present in the 1610 cm⁻¹ region, which is assigned to monodentate carboxylates.^{22,47} Another notable feature in the Ca²⁺ bound 4P₂A spectrum is a low amplitude peak centered around 1558 cm⁻¹, a feature that has been observed in the natively Ca²⁺-responsive EF-hand protein, CaM.²² This feature is indicative of a bidentate carboxylate configuration.^{48,49} Its presence in the 4P₂A but not the WT spectrum suggests that the proline produces Ca²⁺ binding configurations that are different from CaM or other EF-hand proteins. Previous studies have shown the bidentate feature to shift to lower frequencies (~1535 cm⁻¹) upon binding lanthanides instead of Ca²⁺ due to their increased charge.²² This low-frequency feature is not visible in the Ln³⁺-LanM spectra, which indicates that the bidentate carboxylate frequency in LanM may appear even lower, outside of the measured spectral range, and is therefore obscured on the low-frequency side by the OD bend vibration of water.^{50,51} This interpretation is further supported by MD simulations, described below.

Fitting the lineshapes to a combination of Gaussian functions semi-quantitatively describes the spectral features (Fig. S3, ESI†). Peak areas report the populations (*i.e.*, number of residues) of the different coordination geometries, whereas peak frequencies report binding strength (*i.e.*, strength of the electrostatic potential at the C=O positions). Table 1 shows the trend in the extracted center peak frequencies of the monodentate carboxylate asymmetric stretch. In the WT protein, decreasing lanthanide ionic radius is accompanied by a red shift of the monodentate feature. A higher center peak frequency has been correlated to stronger electrostatic interaction between ion and monodentate coordinating residues. For bidentate configurations, on the other hand, stronger interactions are accompanied by a shift towards lower frequencies.^{22,47} As applied to LanM, this correlation may suggest that the metal-ligand interactions are optimal for the early Ln³⁺ ions

Table 1 Center peak frequencies of the monodentate carboxylate population for WT and 4P₂A LanM from Gaussian fitting (Fig. S3)

Ion	WT (cm ⁻¹)	4P ₂ A (cm ⁻¹)
Ca ²⁺	1609.8	1609.7
La ³⁺	1612.4	1609.4
Pr ³⁺	1611.3	1609.6
Nd ³⁺	1611.4	1609.0
Tb ³⁺	1610.0	1609.2
Lu ³⁺	1609.2	1610.8

but become weaker across the series, as well as in response to Ca^{2+} binding, perhaps due to the backbone's inability to adopt favorable configurations given the smaller radii of the later Ln^{3+} ions. Conversely, in the case of the 4P₂A variant, frequencies are relatively independent of metal ion and the monodentate frequencies are generally red-shifted compared to WT. Therefore, although the 4P₂A mutation does not strongly affect lanthanide affinity and the native metal-dependent global conformational response, at least among the early Ln^{3+} ions,⁷ the coordination environment of these lanthanides is perturbed relative to the WT. The Ca^{2+} -bound LanM monodentate peak frequency does not change with the 4P₂A mutation, suggesting that the Ca^{2+} binding configurations are similar, at least with respect to monodentate coordination. Together, these results support the hypothesis that proline residues in LanM's EF-hands play an important role in creating a coordination environment uniquely suited for the early Ln^{3+} ions,⁷ which are biologically relevant.^{5,10,13,52}

The fitting of the carboxylate features in Fig. S3 (ESI[†]) also enables analysis of the ratio between the isolated monodentate peak area to the sum of all carboxylate peaks, as a semi-quantitative metric of the relative populations of the monodentate carboxylates (Table 2). The monodentate populations range between 44% and 80%, with most ions being around 50%, with Ca^{2+} and Tb^{3+} lying on the outer edges. This metric is also useful as a benchmark for MD simulations, as described below, where populations of unbound, monodentate, and bidentate carboxylates are extracted from a structural analysis of the trajectories. The ratios calculated from simulation are reasonably similar to our experimentally calculated ratios, showing that MD simulations capture the structural ensembles observed in the experiment (Tables 2 and 3).

MD simulations

Simulations produce an atomistic interpretation of the structural ensembles observed in the experiments, allowing for a more detailed residue-level analysis. This is a frequently utilized approach for investigating both protein structure and metal binding.^{53–59} Developing accurate models for lanthanide ions remains an active area of research.^{60–65} Lanthanide ion interactions with the hard ligands typical of protein lanthanide binding sites are largely driven by electrostatics.^{64,66} We simulated three ionic conditions: Ca^{2+} , La^{3+} , and Lu^{3+} , to cover the range of ionic radii of the lanthanide series (Fig. 3B).⁶⁷ The Y^{3+} -bound NMR structure⁸ serves as the starting structure for our simulations. This method provides a tractable single-parameter model to directly capture the interplay between ionic radii and binding site structures. On the basis of prior biochemical and structural studies, the fourth EF-hand (Fig. 1A) does not

Table 3 Ligand coordination populations (percent) extracted from MD trajectories. Categorized based on distance histograms shown in Fig. 3A and Fig. S4

	Bidentate		Unassigned		Monodentate		Unbound	
	WT	4P ₂ A	WT	4P ₂ A	WT	4P ₂ A	WT	4P ₂ A
Ca^{2+}	23.8	19.7	1.94	0.90	37.6	45.1	36.6	34.3
La^{3+}	46.4	58.9	0.63	0.12	21.2	17.9	31.8	23.0
Lu^{3+}	52.1	49.4	0.11	0.66	30.0	26.5	17.8	23.5

coordinate lanthanides with high affinity;^{7,8} therefore, we focus on the carboxylate ligands of the first three binding sites. Within this structural analysis, each ligand is classified as monodentate, bidentate, intermediate, or unbound, using a two-step classification protocol. First, the α -carbon to ion distance of each ligand was calculated, with a distance larger than 0.4 nm being classified as “unbound,” selected by inspection of the histograms (Fig. S4, ESI[†]). Then, for each bound ligand, the difference in oxygen-to-ion distance was used to classify each ligand as monodentate (>0.1 nm), bidentate (0–0.05 nm), or unassigned (0.05–0.1 nm). These cut-offs were selected by examining the histograms shown in Fig. 3A and are consistent with values previously reported in the literature.⁶⁸ The resulting populations are shown in Table 3. The populations of unassigned carboxylates is very small $<2\%$.

This geometry-based coordination analysis leads to several observations: 1. The bidentate population is significantly higher for the lanthanides relative to Ca^{2+} (Table 3). 2. The monodentate population does not greatly change between LanM variants. The 4P₂A variant bound to Ca^{2+} shows the highest monodentate population. 3. Finally, WT LanM shows the highest percentage of unbound ligands when coordinated to Ca^{2+} . This could suggest that water molecules within the binding site may also contribute to the total coordination. For example, simulations of Ca^{2+} -bound WT LanM suggest a significant number of water molecules in EF2, discussed below. While 4P₂A has been shown to conformationally respond to Ca^{2+} in the same way that WT LanM responds to Ln^{3+} , its coordination geometry is still distinct from that of Ln^{3+} – WT LanM. This suggests that, in addition to coordinating ligand geometries, the orientations of non-coordinating residues or the protein backbone may also play a role.

Simulation allows access to experimentally inaccessible parameters by analysis of the MD trajectories. For example, we can characterize additional binding geometries such as the bidentate configuration *via* simulation. This vibrational mode appears at a lower frequency range relative to unbound and monodentate carboxylates, which complicates analysis due to the presence of overlapping water vibrational modes. However, through MD, we extract an estimate of this population. Furthermore, since simulations produce residue-level information, the specific contributions of individual residues (e.g., prolines) towards binding site geometry can be analyzed. This residue-level information is an important strength of simulations, as these local interactions are challenging to probe experimentally.

Table 2 Relative populations of monodentate ligands to total observable peak area (percent total)

Protein	Ca^{2+}	La^{3+}	Pr^{3+}	Nd^{3+}	Tb^{3+}	Lu^{3+}
WT	44	63	74	61	59	60
4P ₂ A	51	52	50	61	80	48

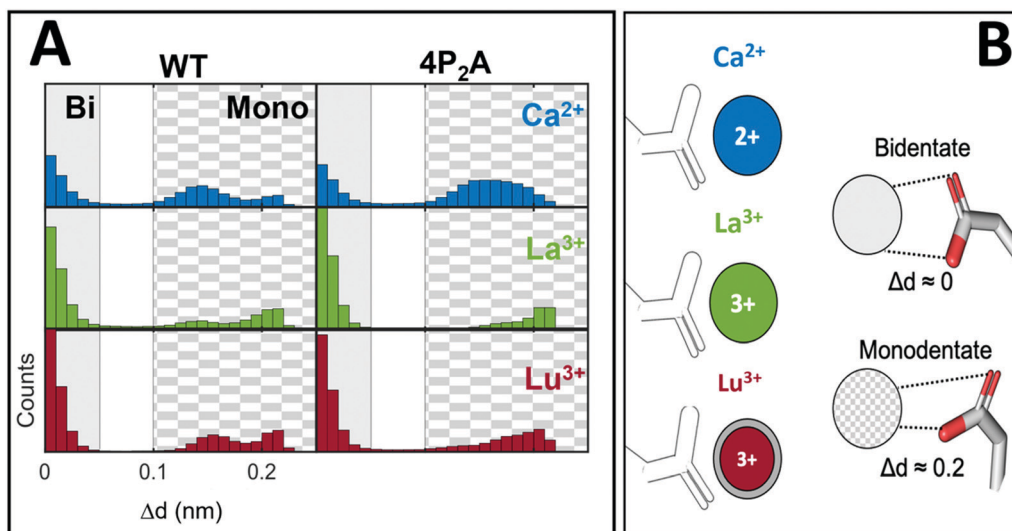


Fig. 3 (A) Histograms of the oxygen to metal ion distance difference (Δd) used as the order parameter to categorize the binding coordination of simulated WT and 4P₂A LanM. This histogram contains aggregated data for all three metal-binding EF-hands, normalized to the total number of frames in the trajectory. Solid gray background indicates the region categorized as bidentate binding, checkered background indicates monodentate binding. (B) Cartoon illustrations of the categories used in Fig. A, illustrating the three ionic conditions (Ca²⁺ in blue, La³⁺ in green, and Lu³⁺ in red); along with bidentate binding (solid gray, $\Delta d \approx 0$ nm) and monodentate binding (checkered, $\Delta d \approx 0.2$ nm).

Proline is not part of the EF-hand sequence in other characterized EF-hand proteins, but it is present in the second position in each of *M. extorquens* LanM's four EF-hands. The conservation of Pro residues at this position in LanM and its homologs, along with the enhanced conformational response to Ca²⁺ in 4P₂A suggests a significant role of these residues in LanM's unique selectivity, though the details are not well

understood.⁷ The 4P₂A variant tests proline's role in LanM's binding by substituting an alanine at proline's position within each of the four EF-hands. Proline's constrained backbone would be expected to reduce binding site flexibility, which may favor orientations that preferentially bind lanthanides.¹³ Fig. 4 shows a 2nd position (Pro for WT, Ala for 4P₂A) specific Ramachandran plot, extracted from MD trajectories, for each of

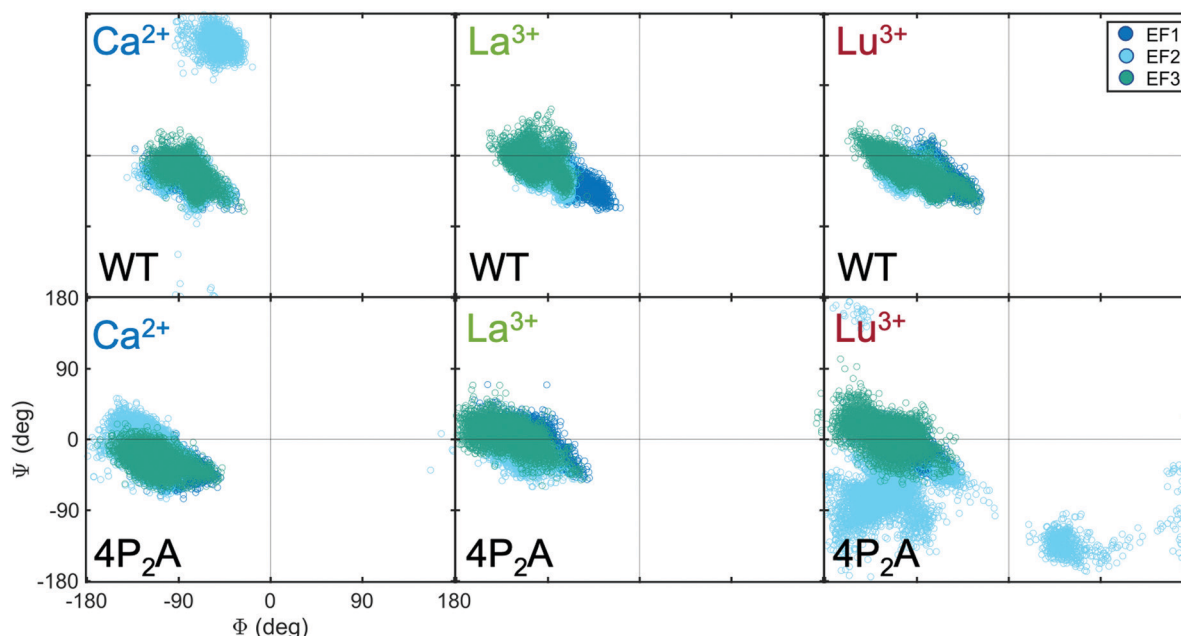


Fig. 4 Ramachandran plots of WT LanM's proline (upper row) and 4P₂A's alanine (lower row) in the second position of each EF hand binding loop, derived from the MD trajectories for Ca²⁺, La³⁺, and Lu³⁺ ions as indicated. The distributions of the second residue of each metal binding EF hand are shown in blue, cyan, and green respectively. EF2 shown in cyan experiences a significant deviation in distribution in a metal- and variant-dependent fashion, supporting its potential significance in LanM's metal recognition.

LanM's three metal-binding EF-hands, EF1, EF2, and EF3. This illustrates the Φ , Ψ angular distributions of these residues, representing a measure of the conformational range and disorder at this position. As seen by comparing the upper and lower rows of Fig. 4, the 4P₂A alanine displays a larger distribution along the Φ coordinate compared to WT's proline. This is expected due to the substitution of the rigid pyrrolidine ring for the flexible alanine.⁶⁹ Across the WT simulations (Fig. 4, upper row), there are relatively minor metal-specific shifts in the proline's Φ angles, with the exception of EF2 bound to Ca²⁺. In that case, importantly, Pro displays two separate constrained distributions—one clustered with the other EF hand Pro distributions, while the other is centered around $-90/180^\circ$. This perturbation is not seen in the 4P₂A variant case. Previously, 4P₂A was observed to have enhanced conformational response to Ca²⁺, similar to the WT variant's response to Ln³⁺.⁷ In these results, we can see that the Ala mutation causes EF2's second position residue in the binding loop to adopt a conformation closer to that seen in Ln³⁺-WT interactions. Furthermore, in the Lu³⁺-bound case, Ala in EF2 in the 4P₂A variant is also strongly perturbed, displaying a wide range of orientations. This indicates increased disorder in the Lu³⁺-EF2 binding interaction in 4P₂A, relative to WT LanM. These results help provide a molecular interpretation for the differences in IR lineshapes between WT and 4P₂A LanM (Fig. 2). In particular, the 2D IR lineshapes are similar across the series for WT, whereas 4P₂A displays a greater diversity. This suggests that 4P₂A may display increased structural flexibility relative to WT.

Further insight can be extracted by analyzing conformational ensembles of other residues within the binding site. The glycine in the sixth position of the EF-hand is highly

conserved in biology and plays a significant role in orienting the EF-hand loop upon metal coordination, with Φ, Ψ angles around $60/20^\circ$ in other canonical Ca²⁺-binding EF-hand proteins.^{24,70} In WT and 4P₂A LanM bound to La³⁺ and Lu³⁺ ions, the Φ, Ψ angles for the sixth position Gly residues cluster in this region, although in the case of EF2 in 4P₂A Lu³⁺ the range of Ψ angles is slightly broader (Fig. 5). However, in the presence of Ca²⁺, EF2 in the WT and 4P₂A proteins exhibits substantial structural differences at the Gly residue. In the WT, the EF2 Gly residue experiences a wide range of dihedral angles, suggesting a disordered and unstable orientation. As just two of the three metal-binding EF-hands appear to undergo a conformational change to Ca²⁺ in WT LanM, and only at high Ca²⁺ concentrations,⁷ the disorder in EF2 observed in our simulations may be related to this experimental observation. In 4P₂A, this distribution becomes increasingly constrained and similar to La³⁺-bound WT (Fig. 5). Indeed, experimentally, the Ca²⁺-bound 4P₂A variant exhibits the full conformational response of the Ln³⁺-bound WT protein.⁷ Therefore, the computational results provide insight into LanM's experimentally observed conformational response, suggesting that EF2 may play an especially important role in Ln³⁺/Ca²⁺ selectivity. Not only does the 4P₂A variant result in altered orientations locally at the position of the mutation, but it also impacts the orientation of neighboring residues within the binding site.

Finally, it is important to note that in our simulations we have not explicitly included nor excluded waters within the binding sites in the starting conformations bulk waters are able to enter the binding sites and interact with the ions during the simulation. The presence of coordinating water molecules could not be probed in the initial NMR structure

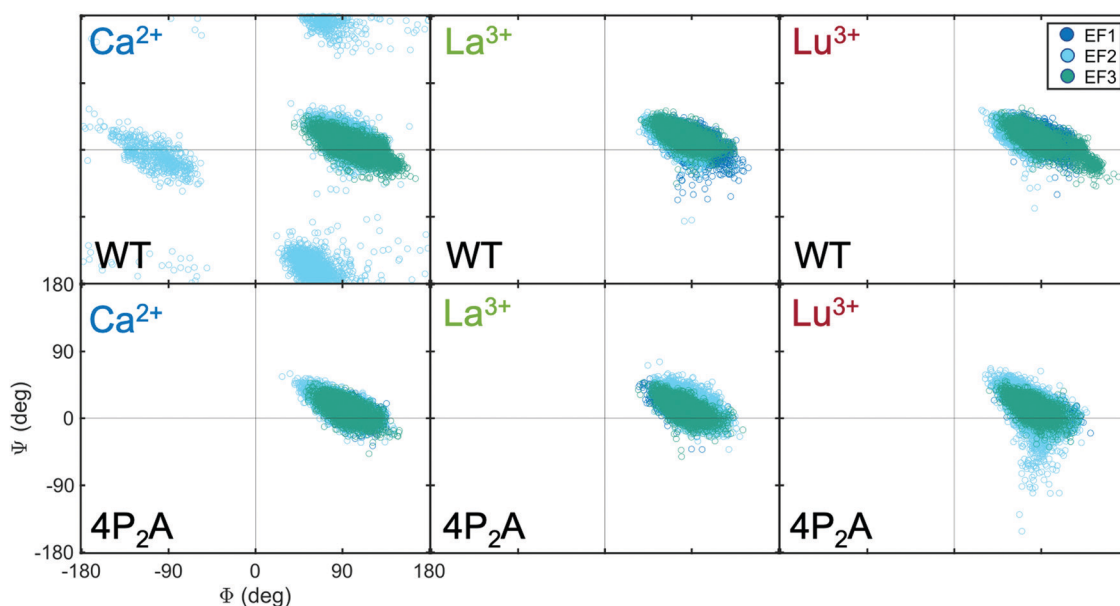


Fig. 5 Ramachandran plots of WT LanM's (upper row) and 4P₂A's (lower row) Gly6 over the course of each simulation bound to Ca²⁺, La³⁺ and Lu³⁺ as indicated. The distributions of each individual binding site are shown in blue, cyan, and green. The WT shows distinct orientations, especially in EF2 of the Ca²⁺-bound case (top left). This changes in the 4P₂A case to become much more like the biologically relevant early Ln³⁺, which display similar distributions in both the WT and 4P₂A case.

Table 4 Populations of water molecules within 0.3 nm of the metal ion within each binding site (cutoff determined by ion-water RDF (Fig. S7))

	EF 1		EF 2		EF 3	
	WT	4P ₂ A	WT	4P ₂ A	WT	4P ₂ A
Ca ²⁺	1.00	1.00	4.94	2.00	1.99	0.00
La ³⁺	2.00	1.00	3.00	3.95	0.00	0.00
Lu ³⁺	1.00	1.00	1.00	3.00	0.03	0.00

determination,⁸ but other canonical EF-hands such as CaM are known to coordinate metal-ions with water molecules.^{30,71,72} Our simulations of LanM show that there are EF-hand dependent populations of water near the binding sites which could directly coordinate the metal ion, replace carboxylates, and stabilize these displaced side-chains through hydrogen bonding (Table 4). These waters contribute to the total population of oxygen atoms near the metal ion (Table S1, ESI[†]), which reproduce the expected lower coordination number of Ca²⁺ (~7–8) and the higher coordination number for La³⁺ and Lu³⁺ (~9–10). EF2, which was most sensitive to Ca²⁺ vs. lanthanide coordination, as well as substitution of its proline residue, has the largest population of waters when coordinated to Ca²⁺ and lanthanides. On the other hand, EF3 has zero nearby water molecules for all ion and sequence cases except for Ca²⁺ bound WT LanM, and solvent near EF1 is less sensitive to metal and WT/4P₂A. While the presence of water molecules in the binding site is yet to be confirmed experimentally, these simulation results indicate the significant role that water molecules may play in LanM's metal ion recognition, much like in other EF-hand proteins^{30,70} and in the lanthanide binding tag.³²

Conclusions

This spectroscopic study provides detailed information regarding the set of structural and dynamic characteristics underlying LanM's unique affinity for lanthanide ions. We used a combination of 2D IR spectroscopy and MD simulations to characterize structural differences at the metal-binding sites of LanM and its 4P₂A variant, associated with binding of Ln³⁺ and Ca²⁺ ions. We found that Ca²⁺, La³⁺, and Lu³⁺ ions elicit different conformations of both metal-coordinating carboxylates and non-coordinating residues.

The spectroscopic measurements presented here show the protein's complex binding response, where small differences in ion size determine coordination geometries and are likely the determinants of binding affinities. Biologically relevant early lanthanides induce stronger monodentate coordination, whereas late Ln³⁺ ions and Ca²⁺ are associated with more disordered carboxylate conformations, which may contribute to the observed affinities of the protein for these metal ions. It further illustrates the interplay of metal binding and global conformational response—conformational subtleties at the local binding site can influence the coupling of large-scale secondary structure changes and metal binding. Simulations provide more atomistic interpretations to describe the

differences in WT and 4P₂A LanM's binding behavior. The proline at the second position in the EF-hand binding loop impacts the orientation of structurally important residues, such as the glycine in the sixth position. Our results also suggest that EF2 is particularly sensitive to the identity of the metal ion, such that it may play a pivotal role in the selectivity of the protein. Together, this analysis supports the hypothesis that proline constrains the binding site into an orientation optimized for biological function. In the context of prior biochemical data, the nuanced relationship between the proline and glycine rigidity and flexibility offers potential explanations for the vastly differing responsiveness of the WT protein and variant to lanthanides and Ca²⁺, in contrast to other EF-hand proteins.

These findings lay a foundation for a more detailed interrogation of this unique metal sensing protein and of lanthanide recognition in biology more generally. Our simulations provide an atomistic view of the fast timescale fluctuations within the binding site that would be inaccessible by crystallographic methods. Future efforts may be directed towards exploring global LanM ion-dependent conformational landscapes. Finally, these conformational studies provide a further structural basis for exploring and engineering LanM for metal extraction and separation applications.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by the Welch Foundation (F-1891) and the National Institutes of Health R35GM133359 (to C. R. B.), as well as the Pennsylvania State University and the National Science Foundation (CHE-1945015, to J. A. C.). Simulations were carried out using the computational resources of the Texas Advanced Computing Center (TACC).

Notes and references

- 1 L. Chistoserdova, Lanthanides: New life metals?, *World J. Microbiol. Biotechnol.*, 2016, **32**, 138.
- 2 L. J. Daumann, Essential and Ubiquitous: The Emergence of Lanthanide Metallobiochemistry, *Angew. Chem., Int. Ed.*, 2019, **58**, 12795–12802.
- 3 E. R. Featherston and J. A. Cotruvo, The biochemistry of lanthanide acquisition, trafficking, and utilization, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2021, **1868**, 118864.
- 4 Y. Hibi, K. Asai, H. Arafuka, M. Hamajima, T. Iwama and K. Kawai, Molecular structure of La³⁺-induced methanol dehydrogenase-like protein in *Methylobacterium radiotolerans*, *J. Biosci. Bioeng.*, 2011, **111**, 547–549.
- 5 A. Pol, T. R. M. Barends, A. Dietl, A. F. Khadem, J. Eygensteyn, M. S. M. Jetten and H. J. M. Op den Camp,

- Rare earth metals are essential for methanotrophic life in volcanic mudpots, *Environ. Microbiol.*, 2014, **16**, 255–264.
- 6 T. Nakagawa, R. Mitsui, A. Tani, K. Sasa, S. Tashiro, T. Iwama, T. Hayakawa and K. Kawai, A Catalytic Role of XoxF1 as La^{3+} -Dependent Methanol Dehydrogenase in *Methylobacterium extorquens* Strain AM1, *PLoS One*, 2012, **7**, e50480.
 - 7 J. A. Cotruvo, E. R. Featherston, J. A. Mattocks, J. V. Ho and T. N. Laremore, Lanmodulin: A Highly Selective Lanthanide-Binding Protein from a Lanthanide-Utilizing Bacterium, *J. Am. Chem. Soc.*, 2018, **140**, 15056–15061.
 - 8 E. C. Cook, E. R. Featherston, S. A. Showalter and J. A. Cotruvo, Structural Basis for Rare Earth Element Recognition by *Methylobacterium extorquens* Lanmodulin, *Biochemistry*, 2019, **58**, 120–125.
 - 9 A. M. Ochsner, L. Hemmerle, T. Vonderach, R. Nüssli, M. Bortfeld-Miller, B. Hattendorf and J. A. Vorholt, Use of rare-earth elements in the phyllosphere colonizer *Methylobacterium extorquens* PA1, *Mol. Microbiol.*, 2019, **111**, 1152–1166.
 - 10 J. A. Mattocks, J. V. Ho and J. A. Cotruvo, A Selective, Protein-Based Fluorescent Sensor with Picomolar Affinity for Rare Earth Elements, *J. Am. Chem. Soc.*, 2019, **141**, 2857–2861.
 - 11 P. Roszczenko-Jasińska, H. N. Vu, G. A. Subuyuj, R. V. Crisostomo, J. Cai, N. F. Lien, E. J. Clippard, E. M. Ayala, R. T. Ngo, F. Yarza, J. P. Wingett, C. Raghuraman, C. A. Hoeber, N. C. Martinez-Gomez and E. Skovran, Gene products and processes contributing to lanthanide homeostasis and methanol metabolism in *Methylobacterium extorquens* AM1, *Sci. Rep.*, 2020, **10**, 12663.
 - 12 M. Wehrmann, C. Berthelot, P. Billard and J. Klebensberger, Rare Earth Element (REE)-Dependent Growth of *Pseudomonas putida* KT2440 Relies on the ABC-Transporter PedA1A2BC and Is Influenced by Iron Availability, *Front. Microbiol.*, 2019, **10**, 2494.
 - 13 J. A. Cotruvo, The Chemistry of Lanthanides in Biology: Recent Discoveries, Emerging Principles, and Technological Applications, *ACS Cent. Sci.*, 2019, **5**, 1496–1506.
 - 14 G. J.-P. Deblonde, J. A. Mattocks, D. M. Park, D. W. Reed, J. A. Cotruvo and Y. Jiao, Selective and Efficient Biomacromolecular Extraction of Rare-Earth Elements using Lanmodulin, *Inorg. Chem.*, 2020, **59**, 11855–11867.
 - 15 H. Kawasaki, S. Nakayama and R. H. Kretsinger, Classification and evolution of EF-hand proteins, *Biomaterials*, 1998, **11**, 277–295.
 - 16 D. Chin and A. R. Means, Calmodulin: A prototypical calcium sensor, *Trends Cell Biol.*, 2000, **10**, 322–328.
 - 17 E. E. Snyder, B. W. Buoscio and J. J. Falke, Calcium(II) site specificity: effect of size and charge on metal ion binding to an EF-hand-like site, *Biochemistry*, 1990, **29**, 3937–3943.
 - 18 C.-L. A. Wang, R. R. Aquaron, P. C. Leavis and J. Gergely, Metal-Binding Properties of Calmodulin, *Eur. J. Biochem.*, 1982, **124**, 7–12.
 - 19 K. N. Allen and B. Imperiali, Lanthanide-tagged proteins—an illuminating partnership, *Curr. Opin. Chem. Biol.*, 2010, **14**, 247–254.
 - 20 H. G. Brittain, F. S. Richardson and R. B. Martin, Terbium(III) emission as a probe of calcium(II) binding sites in proteins, *J. Am. Chem. Soc.*, 1976, **98**, 8255–8260.
 - 21 M. C. Kilhoffer, J. G. Demaille and D. Gerard, Terbium as luminescent probe of calmodulin calcium-binding sites; domains I and II contain the high-affinity sites, *FEBS Lett.*, 1980, **116**, 269–272.
 - 22 S. C. Edington, A. Gonzalez, T. R. Middendorf, D. B. Halling, R. W. Aldrich and C. R. Baiz, Coordination to lanthanide ions distorts binding site conformation in calmodulin, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, 201722042.
 - 23 P. Mulqueen, J. M. Tingey and W. D. Horrocks, Characterization of lanthanide(III) ion binding to calmodulin using luminescence spectroscopy, *Biochemistry*, 1985, **24**, 6639–6645.
 - 24 D. B. Halling, B. J. Liebeskind, A. W. Hall and R. W. Aldrich, Conserved properties of individual Ca^{2+} -binding sites in calmodulin, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E1216–E1225.
 - 25 M. Nara, M. Tasumi, M. Tanokura, T. Hiraoki, M. Yazawa and A. Tsutsumi, Infrared studies of interaction between metal ions and Ca^{2+} -binding proteins Marker bands for identifying the types of coordination of the side-chain COO^- groups to metal ions in pike parvalbumin ($\text{pI} = 4.10$), *FEBS Lett.*, 1994, **349**, 84–88.
 - 26 M. Nara, M. Tanokura, T. Yamamoto and M. Tasumi, A comparative study of the binding effects of Mg^{2+} , Ca^{2+} , Sr^{2+} , and Cd^{2+} on calmodulin by Fourier-transform infrared spectroscopy, *Biospectroscopy*, 1995, **1**, 47–54.
 - 27 M. Nara, H. Morii and M. Tanokura, Coordination to divalent cations by calcium-binding proteins studied by FTIR spectroscopy, *Biochim. Biophys. Acta, Biomembr.*, 2013, **1828**, 2319–2327.
 - 28 F. Yumoto, M. Nara, H. Kagi, W. Iwasaki, T. Ojima, K. Nishita, K. Nagata and M. Tanokura, Coordination structures of Ca^{2+} and Mg^{2+} in Akazara scallop troponin C in solution. FTIR spectroscopy of side-chain COO^- groups, *Eur. J. Biochem.*, 2001, **268**, 6284–6290.
 - 29 J. W. DePalma, P. J. Kelleher, L. C. Tavares and M. A. Johnson, Coordination-Dependent Spectroscopic Signatures of Divalent Metal Ion Binding to Carboxylate Head Groups: H^2 - and He-Tagged Vibrational Spectra of $\text{M}^{2+} + \text{RCO}_2^-$ ($\text{M} = \text{Mg}$ and Ca , $\text{R} = -\text{CD}_3$, $-\text{CD}_2\text{CD}_3$) Complexes, *J. Phys. Chem. Lett.*, 2017, **8**, 484–488.
 - 30 B. Drobot, M. Schmidt, Y. Mochizuki, T. Abe, K. Okuwaki, F. Brullfert, S. Falke, S. A. Samsonov, Y. Komeiji, C. Betzel, T. Stumpf, J. Raff and S. Tsushima, $\text{Cm}^{3+}/\text{Eu}^{3+}$ induced structural, mechanistic and functional implications for calmodulin, *Phys. Chem. Chem. Phys.*, 2019, **21**, 21213–21222.
 - 31 V. Migliorati, A. Serva, F. M. Terenzio and P. D'Angelo, Development of Lennard-Jones and Buckingham Potentials for Lanthanoid Ions in Water, *Inorg. Chem.*, 2017, **56**, 6214–6224.
 - 32 T. Hatanaka, N. Kikkawa, A. Matsugami, Y. Hosokawa, F. Hayashi and N. Ishida, The origins of binding specificity

- of a lanthanide ion binding peptide, *Sci. Rep.*, 2020, **10**, 19468.
- 33 M. L. Valentine, Z. A. Al-Mualem and C. R. Baiz, Pump Slice Amplitudes: A Simple and Robust Method for Connecting Two-Dimensional Infrared and Fourier Transform Infrared Spectra, *J. Phys. Chem. A*, 2021, **125**(29), 6498–6504.
 - 34 J. Huang and A. D. MacKerell, CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data, *J. Comput. Chem.*, 2013, **34**, 2135–2145.
 - 35 P. Mark and L. Nilsson, Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K, *J. Phys. Chem. A*, 2001, **105**, 9954–9960.
 - 36 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*, 2015, **1–2**, 19–25.
 - 37 M. Parrinello and A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
 - 38 G. J. Martyna, M. L. Klein and M. Tuckerman, Nosé–Hoover chains: The canonical ensemble via continuous dynamics, *J. Chem. Phys.*, 1992, **97**, 2635–2643.
 - 39 L. Minnes, D. J. Shaw, B. P. Cossins, P. M. Donaldson, G. M. Greetham, M. Towrie, A. W. Parker, M. J. Baker, A. J. Henry, R. J. Taylor and N. T. Hunt, Quantifying Secondary Structure Changes in Calmodulin Using 2D-IR Spectroscopy, *Anal. Chem.*, 2017, **89**, 10898–10906.
 - 40 L. Minnes, G. M. Greetham, D. J. Shaw, I. P. Clark, R. Fritzsche, M. Towrie, A. W. Parker, A. J. Henry, R. J. Taylor and N. T. Hunt, Uncovering the Early Stages of Domain Melting in Calmodulin with Ultrafast Temperature-Jump Infrared Spectroscopy, *J. Phys. Chem. B*, 2019, **123**, 8733–8739.
 - 41 J. M. Schmidt-Engler, R. Zangl, P. Guldán, N. Morgner and J. Bredenbeck, Exploring the 2D-IR repertoire of the –SCN label to study site-resolved dynamics and solvation in the calcium sensor protein calmodulin, *Phys. Chem. Chem. Phys.*, 2020, **22**, 5463–5475.
 - 42 M. K. Petti, J. P. Lomont, M. Maj and M. T. Zanni, Two-Dimensional Spectroscopy Is Being Used to Address Core Scientific Questions in Biology and Materials Science, *J. Phys. Chem. B*, 2018, **122**, 1771–1780.
 - 43 A. Ghosh, J. S. Ostrander and M. T. Zanni, Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy, *Chem. Rev.*, 2017, **117**, 10726–10759.
 - 44 C. R. Baiz, B. Błasiak, J. Bredenbeck, M. Cho, J.-H. Choi, S. A. Corcelli, A. G. Dijkstra, C.-J. Feng, S. Garrett-Roe, N.-H. Ge, M. W. D. Hanson-Heine, J. D. Hirst, T. L. C. Jansen, K. Kwac, K. J. Kubarych, C. H. Londergan, H. Maekawa, M. Reppert, S. Saito, S. Roy, J. L. Skinner, G. Stock, J. E. Straub, M. C. Thielges, K. Tominaga, A. Tokmakoff, H. Torii, L. Wang, L. J. Webb and M. T. Zanni, Vibrational Spectroscopic Map, Vibrational Spectroscopy, and Intermolecular Interaction, *Chem. Rev.*, 2020, **120**, 7152–7218.
 - 45 C. R. Baiz, M. Reppert and A. Tokmakoff, in *Ultrafast Infrared Vibrational Spectroscopy*, ed. M. Fayer, 2013, p. 38.
 - 46 S. C. Edington, S. Liu and C. R. Baiz, in *Methods in Enzymology*, ed. J. A. Cotruvo, Academic Press, 2021, vol. 651, pp. 157–191.
 - 47 S. C. Edington and C. R. Baiz, Vibrational Relaxation in EDTA Is Ion-Dependent, *J. Phys. Chem. A*, 2018, **122**(32), 6585–6592.
 - 48 J. E. Tackett, FT-IR Characterization of Metal Acetates in Aqueous Solution, *Appl. Spectrosc.*, 1989, **43**, 483–489.
 - 49 G. B. Deacon and R. J. Phillips, Relationships between the carbon-oxygen stretching frequencies of carboxylate complexes and the type of carboxylate coordination, *Coord. Chem. Rev.*, 1980, **33**, 227–250.
 - 50 A. Barth, The infrared absorption of amino acid side chains, *Prog. Biophys. Mol. Biol.*, 2000, **74**, 141–173.
 - 51 A. Barth and C. Zscherp, What vibrations tell about proteins, *Q. Rev. Biophys.*, 2002, **35**, 369–430.
 - 52 H. N. Vu, G. A. Subuyuj, S. Vijayakumar, N. M. Good, N. C. Martinez-Gomez and E. Skovran, Lanthanide-Dependent Regulation of Methanol Oxidation Systems in *Methylobacterium extorquens* AM1 and Their Contribution to Methanol Growth, *J. Bacteriol.*, 2016, **198**, 1250–1259.
 - 53 M.-C. Bellissent-Funel, A. Hassanali, M. Havenith, R. Henchman, P. Pohl, F. Sterpone, D. van der Spoel, Y. Xu and A. E. Garcia, Water Determines the Structure and Dynamics of Proteins, *Chem. Rev.*, 2016, **116**, 7673–7697.
 - 54 G. G. Dodson, D. P. Lane and C. S. Verma, Molecular simulations of protein dynamics: new windows on mechanisms in biology, *EMBO Rep.*, 2008, **9**, 144–150.
 - 55 J. Nasica-Labouze, P. H. Nguyen, F. Sterpone, O. Berthoumieu, N.-V. Buchete, S. Coté, A. De Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. S. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang and P. Derreumaux, Amyloid β Protein and Alzheimer's Disease: When Computer Simulations Complement Experimental Studies, *Chem. Rev.*, 2015, **115**, 3518–3563.
 - 56 R. J. Xu, B. Błasiak, M. Cho, J. P. Layfield and C. H. Londergan, A Direct, Quantitative Connection between Molecular Dynamics Simulations and Vibrational Probe Line Shapes, *J. Phys. Chem. Lett.*, 2018, **9**, 2560–2567.
 - 57 M. Reppert and A. Tokmakoff, Computational Amide I 2D IR Spectroscopy as a Probe of Protein Structure and Dynamics, *Ann. Rev. Phys. Chem.*, 2016, **67**, 359–386.
 - 58 T. Hayashi, W. Zhuang and S. Mukamel, Electrostatic DFT Map for the Complete Vibrational Amide Band of NMA, *J. Phys. Chem. A*, 2005, **109**, 9747–9759.
 - 59 S. Mitra, K. Werling, E. J. Berquist, D. S. Lambrecht and S. Garrett-Roe, CH Mode Mixing Determines the Band Shape of the Carboxylate Symmetric Stretch in Apo-EDTA, Ca^{2+} -EDTA, and Mg^{2+} -EDTA, *J. Phys. Chem. A*, 2021, **125**(22), 4867–4881.
 - 60 B. Qiao, S. Skanthakumar and L. Soderholm, Comparative CHARMM and AMOEBA Simulations of Lanthanide

- Hydration Energetics and Experimental Aqueous-Solution Structures, *J. Chem. Theory Comput.*, 2018, **14**, 1781–1790.
- 61 F. E. Jorge, L. S. C. Martins and M. L. Franco, All-electron double zeta basis sets for the lanthanides: application in atomic and molecular property calculations, *Chem. Phys. Lett.*, 2016, **643**, 84–88.
 - 62 P. D'Angelo, A. Zitolo, V. Migliorati, G. Chillemi, M. Duvail, P. Vitorge, S. Abadie and R. Spezia, Revised Ionic Radii of Lanthanoid(III) Ions in Aqueous Solution, *Inorg. Chem.*, 2011, **50**, 4572–4579.
 - 63 F. Martelli, S. Abadie, J.-P. Simonin, R. Vuilleumier and R. Spezia, Lanthanoids(III) and actinoids(III) in water: diffusion coefficients and hydration enthalpies from polarizable molecular dynamics simulations, *Pure Appl. Chem.*, 2012, **85**, 237–246.
 - 64 D. C. Cantu, in *Methods in Enzymology*, ed. J. A. Cotruvo, Academic Press, 2021, vol. 651, pp. 193–233.
 - 65 P. Li and K. M. Merz, Metal Ion Modeling Using Classical Mechanics, *Chem. Rev.*, 2017, **117**, 1564–1686.
 - 66 S. Cotton, *Lanthanide and Actinide Chemistry*, John Wiley & Sons, 2013.
 - 67 R. D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 1976, **32**, 751–767.
 - 68 M. M. Harding, Geometry of metal–ligand interactions in proteins, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2001, **57**, 401–411.
 - 69 H. K. Ganguly and G. Basu, Conformational landscape of substituted prolines, *Biophys. Rev.*, 2020, **12**, 25–39.
 - 70 J. L. Gifford, M. P. Walsh and H. J. Vogel, Structures and metal-ion-binding properties of the Ca²⁺-binding helix–loop–helix EF-hand motifs, *Biochem. J.*, 2007, **405**, 199–221.
 - 71 R. Chattopadhyaya, W. E. Meador, A. R. Means and F. A. Quiocho, Calmodulin structure refined at 1.7 Å resolution, *J. Mol. Biol.*, 1992, **228**, 1177–1192.
 - 72 W. De W. Horrocks, *Methods in Enzymology*, Academic Press, 1993, vol. 226, pp. 495–538.