LanM Peptides – Unravelling the Binding Properties of the EF-Hand Loop Sequences Stripped from the Structural Corset

Sophie M. Gutenthaler^[a], Satoru Tsushima^[b,c], Robin Steudtner^[b], Manuel Gailer^[a], Anja Hoffmann-Röder^[a], Björn Drobot^{,[b]}, Lena J. Daumann^{*,[a]}*

[a] Department of Chemistry Ludwig-Maximilians-University Munich Butenandtstraße 5-13, 81377 München (Germany)

[b] Institute of Resource Ecology Helmholtz-Zentrum Dresden-Rossendorf e.V. Bautzner Landstraße 400, 01328 Dresden (Germany)

[c] World Research Hub Initiative (WRHI), Institute of Innovative Research, Tokyo Institute of Technology, Meguro 152-8550, Tokyo (Japan)

Keywords: Lanmodulin, metal-binding peptides, lanthanide-binding peptides, actinide-binding peptides, EF-Hand loop-based peptides

Abstract: Since the discovery of the biological relevance of lanthanides (Lns) for methylotrophic bacteria in the last decade, the field has seen a steady rise in discoveries of bacteria using Lns. The major role of lanthanides here is in the active sites of enzymes: methanol dehydrogenases. Additionally, lanthanide binding proteins have also been identified. One such protein is lanmodulin (LanM) and, with a remarkable selectivity for Lns over Ca(II) and affinities in the picomolar range, it makes an attractive target to address challenges in lanthanide separation. Why LanM has such a high selectivity is currently not entirely understood, both the specific amino acid sequences of the EF-hand loops, together with cooperativity effects have been suggested. Consequently, we decided to remove the effect of cooperativity by focusing on the amino acid level. Thus, we synthesized all four 12-amino acid EF-Hand loop peptides of LanM using solid phase peptide synthesis and investigated their affinity for Lns (Eu(III), Tb(III)), the actinide Cm(III) and Ca(II). Using isothermal titration calorimetry and time resolved laser fluorescence spectroscopy combined with parallel factor analysis, we show that in the absence of cooperativity the short EF-Hand loop peptides have all similar affinities for lanthanides and that these are all in the micromolar range. Furthermore, calcium was shown not to bind to the peptides which was verified with circular dichroism spectroscopy. This technique also revealed that the peptides undergo a change to a more ordered state when lanthanides are added. These experimental observations were further supported by molecular dynamics simulations. Lastly, we put Eu(III) and Cm(III) in direct competition using TRLFS. Remarkably, a slightly higher affinity for the actinide, as was also observed for LanM, was found. Our results demonstrate that the picomolar affinities in LanM are largely an effect of pre-structuring in the full protein and therefore reduction of flexibility in combination with cooperative effects, and that all EF-Hand loops possess similar affinities when detached from the protein backbone, albeit still retaining the high selectivity for lanthanides and actinides over calcium.

1 Introduction

Lanthanides (Lns), the group of 15 elements at the centre of the periodic table, belong, together with Sc and Y, to the rare earth elements (REEs) and play a key role in our modern society.^[1] These elements are essential in phosphors (e.g. LEDs, screens), strong permanent magnets (e.g. hard-drives, phones, wind turbines), rechargeable batteries and are even found in medical applications (e.g. Gd(III) in MRI contrast agents).^[2,3] Luckily, lanthanides are not as rare as the misleading name REE indicates, however, despite their widespread occurrence – apart of the radioactive promethium – around the world, Lns mostly co-appear in ores at low concentrations.^[4] This, and the chemical similarity within the series (prevalent oxidation state +III and similar ionic radii), makes the mining, separation and purification of these elements extremely cumbersome.^[1,4] Additionally, the procurement of Lns is also extremely harmful to the environment. The commonly used mining process involves harsh acidic or alkaline conditions leading to groundwater pollution by mobilising heavy metals from the ores as well as producing large amounts of radioactive waste (Lns often co-occur with actinides such as thorium or uranium).^[2,3,5,6] This negative environmental impact is in strong conflict with the necessity of these elements for sustainable energy production. Therefore, greener solutions to satisfy the high demand for Lns are urgently needed. One suitable, more sustainable Ln-source, is believed to be found in End-of-Life (EoL) products.^[7] However, similarly to the naturally occurring ores, Lns are used in combinations with other elements and at low concentrations in those EoL products. This highlights that for sustainable Ln acquisition, highly selective and effective methods are needed. As nature has always been an inspirational source for scientists, it comes in handy that Lns were found to be biologically essential less than 10 years ago with the finding of Op den Camp and co-workers who first reported a Ln-dependent methanotrophic bacterium in 2014, opening up a completely new research field.^[8] Nakagawa and co-workers had demonstrated earlier in 2012 that *M. extorquens* AM1 (short: AM1) can express a previously uncharacterised methanol dehydrogenase (MDH) upon addition of La(III) to the growth medium.^[9] Along with SolV, this organism belongs to the ones best studied for its lanthanide metabolism.^[10] In particular, *M. extorquens* AM1 is already strongly linked with bio-inspired REE-recycling strategies as it can, for example, directly leach Lns from magnet swarf.^[11] Furthermore, AM1 has also been evolved to tackle the rising gadolinium levels in waterways, stemming from the widespread use of MRI contrast agents.^[12]

In the past decade, it has been firmly established that Ln-using bacteria incorporate Lns in the active site of quinone-dependent MDH,^[13,14] an enzyme class formerly only associated with calcium.^[15–18] As calcium and Lns have similar ionic radii, chemists have used Lns as substitutes for calcium to investigate Ca-binding proteins and enzymes by using the excellent spectroscopic properties of these metals long before it was discovered that Lns are biological relevant.^[19,20] With that at mind, it is not surprising that the first discovered naturally occurring Ln-binding protein lanmodulin (LanM) isolated from *M. extorquens* AM1 belongs to the class of EF-Hand proteins, which are again associated with binding strongly to calcium. The EF-Hand motif refers to a structural helix-loop-helix unit in which two alpha-helices are linked by a short 12-amino acid calcium(II)-binding loop.^[21] LanM has four EF-Hands with metal-binding loops of which three have been reported to have a picomolar affinity for Lns and the fourth with lower, micromolar affinity, although newer studies have also suggested that this site does not bind lanthanides under the usually used conditions.^[22–24] The affinity for Lns has been demonstrated to be 100-million-fold higher for Lns than for calcium.^[22] In addition, it was recently shown that LanM not only has a high affinity for Lns, but also for actinides (Ans).^[25–27] Interestingly, all conducted studies observe a higher affinity of LanM for Ans than for the Lns.^[25–27]

Due to the aforementioned similarities between calcium and Lns, calcium-binding proteins such as calmodulin (CaM) have been used in the past as template for specific Ln-binding peptides.^[28–31] Particularly noteworthy are lanthanide-binding tags (LBTs) which have been developed starting from EF-Hand binding loop sequences as tool for analysing protein structures and protein-protein interactions by incorporating those short Ln-binding amino acid sequences (the LBTs) into the protein structure which enables analysis *via* methods involving luminescence, NMR spectroscopy and others.^[19,32–37] Additionally, Uranyl-binding peptides, based on EF-Hand sequences of CaM, are known.^[38–40]

Now, with the discovery of LanM, a protein, designed by nature to bind Lns and a high affinity for Ans, can be used directly for the development of REE-recycling and Ans separation methods^[24,26,27,41,42], or as foundation to understand nature's design rules. With this inspiration, new *de novo* proteins^[43], peptides^[44] and organic ligands can be envisioned. An example for a bio-inspired approach towards Ln separation would be the use of pyrroloquinoline quinone (PQQ) which is the cofactor of the aforementioned Ln-MDHs and was demonstrated to be capable of separating Lns by precipitation in aqueous solution.^[45] A key-step in using natural ligand-design as a foundation to develop potent Ln-binding ligands that are suitable, for example, for recycling of EoL products and the separation of Lns and Ans, is understanding why a used design element is beneficial. For CaM numerous extensive studies and systematic analyses, which amino acid combinations in which positions are most common in CaM and related proteins were conducted.^[21,46] Interestingly, features that decrease Ca(II) affinity while increasing the affinity towards Lns or other trivalent ions have been already reported in the literature.^[47–49] Remarkably, some of these can be found in lanmodulin. For example, aspartic acid can be found in the ninth position in all four EF-Hands of lanmodulin^[22] and the presence of an acidic amino acid in this specific position was already shown to be beneficial in order to decrease the Ca(II) selectivity and increase the affinity for trivalent ions.^[47–49] However, aspartic acid in position nine is not highly uncommon in canonical EF-loops as this position is occupied by Asp in almost a third of CaMs and CaM-like proteins.^[21] More unusual is the presence of proline in position two (again in all four EF-Hands) which is why it was suggested by Cotruvo and co-workers^[22] that this might be one reason for lanmodulin's unique features. This was already investigated on the protein level *via* molecular dynamics simulations and 2D IR.^[50] The authors found when exchanging the proline in position two against an alanine that the selectivity for the

early lanthanides decreases. This was attributed to the resulting higher flexibility of the binding site which can bind to a broader range of ions, highlighting the proposed role of proline.

Interestingly, this is also supported by a study published in 2004^[51] in which 36 amino acid long helix-loop-helix peptides derived from a Ca(II)-binding EF-Hand protein were investigated, revealing that by replacing the proline by glycine the affinity for Ca(II) can be further increased, demonstrating that the presence of proline decreases the selectivity for Ca(II) and increases the selectivity for trivalent ions. The lower affinity (or no affinity in more recent studies) of EF-Hand 4 in LanM is suggested to be mainly due to the presence of an asparagine residue in position one (EF1-EF3, D at position 1) which is normally occupied by an aspartic acid residue in EF-Hand proteins.^[21] One possibility to take a closer look at the differences between LanM's four binding loops and unravelling the affinity and differences fully detached from cooperativity and other effects observed in proteins, is to take the protein apart and look at it on the peptide level.

Therefore, this study brings us one step closer to understand LanM high affinity for Lns – in the absence of protein pre-structuring and cooperativity – by investigating the EF-Hand binding loops as four isolated 12-amino acid peptides. We use time-resolved laser-induced fluorescence spectroscopy (TRLFS), circular dichroism (CD) spectroscopy, isothermal titration calorimetry (ITC) and molecular dynamic (MD) simulations to investigate LanM's EF-Hand peptides for their binding properties with terbium, europium, curium and calcium with a main focus on europium, to gain a basic understanding of mother nature's design rules. Eu(III), Tb(III) and Cm(III) were selected as representatives for trivalent Lns and Ans based on their excellent fluorescence properties.

2 Results and Discussion

Synthesis, Purification and Determination of the Target Peptide Concentration

In this study a total of four 12-amino acid peptides based on the four metal-binding EF-Hand loops (numbered 1 to 4) of the naturally occurring Ln-binding protein LanM were synthesised *via* automated microwave-assisted solid phase peptide synthesis (SPPS) using *Fmoc* conditions. For all peptides a preloaded Wang-resin that gives a C-terminal carboxylic acid was used. The C-terminal carboxylic acid was preferred over a terminal amide function, as an additional carboxylic group was expected to be beneficial for binding lanthanides. As all four sequences contain at least one amino acid combination (e.g. Gly-Asp) which is prone to aspartimide formation,^[52–55] 5% formic acid was added during the *Fmoc* deprotection step to minimise the formation of unwanted side-products as recommended in the literature^[56]. With these conditions, all four sequences were successfully synthesised as determined by mass spectrometry and purified by preparative RP-HPLC (Chapter 2, SI). As no sequence contained a tryptophan side chain or an aromatic residue in general, the absorption at 280 nm could not be used to determine the net peptide concentration. Therefore, the net peptide concentration was determined by using the Pierce™ quantitative fluorometric peptide assay. Together with the peptide purity determined by analytical HPLC and the theoretical molecular weight, the target peptide concentration was obtained. The determined concentration was later corrected with a correction factor N determined by ITC (Section 4.2, SI), as the analysis suggested a slight underestimation of the real peptide concentration by using the method described above.

Metal-binding Studies

LanM possesses an unparalleled binding affinity for lanthanides (and actinides). Out of the four EF-Hands it is now suggested that only two of them have high affinity (EF2 and EF3), while EF1 has lower affinity and under the conditions usually reported, EF4 is proposed to not bind lanthanides.^[24] Both, the high affinity for lanthanides as well as the high selectivity over calcium(II) have been proposed to be a combination of cooperativity and the amino

acid sequence. Here, we investigated the binding affinities of all four EF-Hands on the amino acid level using several methods to unravel the molecular and thermodynamic basis for LanM's unusual selectivity.

Eu(III) and Ca(II) binding to the EF-Hand loop peptides

The investigations were started with ITC experiments which were performed with Eu(III) and Ca(II). The obtained thermogram and integrated heat plot of the EF1/Eu(III) titration experiment are shown as an example in Figure 2 on the left. The data was first analysed by solely considering the formation of a 1:1 peptide to Eu(III) complex when fitting binding isotherms. However, a slight increase of enthalpy after the second titration step as well as a significant offset between the obtained fit and the raw data with increasing Eu(III) concentration (which were both reproducible when repeating the experiments and systematic for all peptides), suggested the formation of more than one peptide/Eu(III) complex (a speciation of the allocated species is shown in Figure 3 on the left). The slight increase at the beginning of the titration experiment might be explained with the formation of a 2:1 peptide to Eu(III) complex. However, due to the used set-up and most probably the instability of this potential complex, it was not possible to fully thermodynamically characterise this complex which is why the 2:1 complex was neglected during further analysis and this second titration step excluded from the final analysis (compare Section 4.3 of the SI and Figure S4). The mentioned offset could be explained with the formation of a 1:2 peptide to Eu(III) complex after reaching an excess of Eu(III). This complex can be fully thermodynamically characterised and together with the 1:1 complex the obtained ITC data can be fully explained. The reaction enthalpy (ΔH) for the formation of the 1:1 complex is slightly higher than for the 1:2 complexes. In general, for all four peptides very similar values ($\Delta H_{1:1}=20-24$ kJ/mol; $\Delta H_{2:1}=21-19$ kJ/mol for EF1-EF3) were obtained. Solely for the 1:2 complex of EF4 a lower reaction enthalpy can be observed (13 kJ/mol). For the 1:1 complexes the Gibbs energy (ΔG) is around -29 kJ/mol and for the 1:2 complexes around -20 kJ/mol, showing that for both the complex formation is thermodynamically favourable which suggests, together with the observed endothermic behaviour, an entropy-driven binding event. Not solely the metal binding process, but also the stripping of the first and second hydration shells of the Eu(III) aquo ion (increase in entropy),^[49] as well as the presumably dehydration of the peptide (increase in entropy) in addition to structural changes upon metal-binding contribute to the final thermogram. As the complex formation seems to be for all peptides similarly exergonic, the lower enthalpy of the 2:1 Eu(III) to EF4 complex goes hand in hand with a lower entropic component (ΔH , ΔG , $-T\Delta S$ for both species shown in Table S4). In addition to the thermodynamic variables, the determined complex binding affinities (K_D) again show very similar values for all four peptides. For the formation of a 1:1 complex the Eu(III) affinity is in the lower micromolar range (around $8.5 \mu\text{M}$) while the affinity for the formation of a 1:2 complex is significantly lower (see Table 1). Under the tested conditions no significant differences between the four peptides can be determined. This is interesting as at least for EF4 – the proposed low-affinity site in the parent protein LanM^[22] – a lower affinity was expected. However, on the amino acid level no differences between the four EF-Hands could be determined suggesting that the different binding affinities for the four binding-loops in the parent protein are dominated by cooperative effects.

Figure 2 Right: ITC of EF1 binding to Eu(III). Top: Thermogram obtained by performing a titration of 1.5 mM Eu(III) to 100 μM EF1 at pH 6.6 (10 mM MOPSO, 100 mM KCl) and the corresponding background measurement (Eu(III) to buffer). Bottom: Integrated heat and best fit with a 2-component model. Datapoints which are displayed as circles were excluded from the analysis. Left: TRLFS titration experiment of Eu (0-122.5 μM) to EF1 (12 μM) at pH 6.6 (10 mM MOPSO, 100 mM KCl), λ_{ex} (Eu) = 394 nm. Top: Deconvoluted spectra. Bottom: Lifetimes of the three observed species.

In contrast to the ITC experiments with Eu(III), the addition of Ca(II) to the peptide solutions showed under the tested conditions no complex formation, even though the added Ca(II)-concentration was five times the amount used in the Eu(III) set-up. The obtained thermograms hardly deviated from the background titrations (Figure S10 and S11). This clearly shows that *via* ITC under the tested conditions no Ca(II)-binding to the peptides could be observed. However, an affinity in the molar range cannot be fully ruled out with this technique, as ITC is better suited for systems with an affinity in the millimolar to nanomolar range. One could conclude, based on this first comparison between Ca(II) and Eu(III), that the high selectivity for Lns over Ca(II) for which the parent protein LanM is known in the literature,^[22] is preserved in the short amino acid sequences of the binding-loops. However, a significantly lower Ca(II) affinity of short EF-Hand based peptides is to be expected, based on the available literature on Ca(II)-binding protein-inspired small peptides. It was shown multiple times that the Ca(II) affinity is significantly reduced in such peptides and strongly depends on the length of the amino acid chain. For most of the reported 12 AA peptides no Ca(II) affinity or solely a low affinity in the millimolar range is reported while longer peptides (starting at around 20 AA) show to have affinities in the higher micromolar range.^[57–60] One reason for the absence of even a millimolar affinity between Ca(II) and the LanM peptides could be the proline residue at position two. It has already been shown in 2004 that the presence of a proline residue significantly reduces the affinity for Ca(II) in protein-based peptides.^[51]

Figure 3 Chemical species distribution observed in Eu(III) to EF1 titration experiments (corresponding ITC and TRLFS data shown in Figure 2).

Table 1 K_D values obtained by Eu(III) to EF titration experiments *via* ITC and TRLFS. The ITC values are averaged from two independent experiments using different concentrations; the deviation between both experiments is given as error. The given error for the TRLFS data is the standard deviation. Conditions: 25 °C, pH 6.6 in 10 mM MOPSO buffer with an ionic strength of 100 mM KCl.

EF-Eu(III) complex	K_D (μ M) measured by ITC		K_D (μ M) measured by TRLFS	
	1:1	1:2	1:1	1:2
EF1	8.55 ± 0.17	325 ± 93.0	4.39 ± 2.21	30.5 ± 12.7
EF2	8.61 ± 0.04	309 ± 49.5	4.44 ± 0.82	36.9 ± 3.17
EF3	8.02 ± 1.23	511 ± 169	3.05 ± 2.83	20.5 ± 10.7
EF4	8.85 ± 0.16	456 ± 229	2.42 ± 1.94	50.5 ± 19.3

The observations made by ITC were confirmed and investigated in depth by TRLFS titration experiments. One advantage of this method is that very low concentrations can be used, perfectly suited to investigate affinities in the micromolar or even nanomolar range while using only small sample amounts. Furthermore, the number of bound water molecules can be estimated, giving first hints towards how many amino acid residues are involved in the metal binding as well as structural insights. This method has already been applied successfully for the investigation of CaM and LanM in the literature.^[25,61]

In general, the metal-binding information is obtained due to spectral deconvolution by parallel factor analysis (PARAFAC), which is a robust analytic tool for determining unique explanatory factors directly. The comparison between the ITC and TRLFS experiments is shown in Figure 3 and Table 1.

First, Eu(III) to peptide titration experiments were performed by adding an Eu(III) solution to the respective peptide in solution. The analysis of this data set strongly supported the formation of two different peptide-Eu(III) complexes as suggested by ITC. Furthermore, the determined K_D values (see Table 1) for the 1:1 complex are in the same order of magnitude. Differences between the K_D values can be explained with the different concentration range used for the two methods and general experimental error. However, no hint for the formation of a 2:1 (peptide to Eu(III)) complex could be observed which can be explained with the fact that the complex would be expected to be only formed at very low Eu(III) concentrations in comparison to the respective peptide. This makes it difficult to detect this species in this experimental set-up in which the read-out, Eu(III), is present at very low concentrations during the first titration steps. To verify a 2:1 complex formation, further investigations would be needed. One way to avoid

such issues arising from the low Eu(III) fluorescence signal intensity, is a complementary experimental set-up in which the respective peptide is added in increasing amounts to a metal-solution whose concentration is kept constant (the peptide solution contains the same metal-concentration). A titration experiment with Eu(III) using this set-up confirms the formation of two different peptide/Eu(III) complexes in the experiment described above, as in this set-up solely the 1:1 complex was observed, and the spectral deconvolution showed the strong consistency of the measured data due to the matching measured lifetimes for the observed species (see Figure 7). In both Eu(III) data sets, lifetimes around 180 μ s are observed for the 1:1 complexes (compare Table S5 Table 2). Furthermore, an experiment confirming the observations made by ITC for the Ca(II) binding affinity was conducted using TRLFS. For this, Eu(III) was added to a sample containing EF1 and a high excess of Ca(II). The analysis showed no significant differences to the Eu(III) to EF1 data-set despite the high concentration of Ca(II) present, fully supporting the observation made *via* ITC. The spectral deconvolution yielded again the presence of three species, the free Eu(III) aquo ion, a 1:1 and a 1:2 peptide/Eu(III) complex each with lifetimes matching the values obtained earlier within the experimental error (compare Table S5, Table S6 and Table 2).

In addition, the influence of Eu(III) and Ca(II) on the structure of the four peptides was investigated *via* CD spectroscopy. Titration experiments in which the buffer and peptide concentration was kept constant while the metal concentration was increased. The obtained spectra are shown in Figure 4. In the absence of Ca(II)/Eu(III) the CD spectrum shows for all peptides a distinct minimum around 200 nm and a slight shoulder at 222 nm suggesting a random-coil conformation with very little secondary (if at all) structure elements.^[62] The addition of Eu(III) in increasing amounts leads to a slight decrease of (molar) ellipticity at 222 nm and a slight increase of (molar) ellipticity at around 200 nm. Isodichroic points are visible around 210 nm. The decrease of (molar) ellipticity at 222 nm is associated with an increase of α -helicity caused by metal-binding in the literature.^[28,29,31,63,64] The observed changes appear to increase within the peptide series from EF1 to EF4. The first two EF-Hand peptides as well as EF3 and EF4 show a very similar change in the CD upon Eu(III) addition. EF4 shows slightly more secondary structure content at an excess of Eu(III) in comparison to e.g. EF1 when looking at the (molar) ellipticity at around 220 nm and 222 nm. However, no shift of the minimum at 220 nm towards 210 nm which would be typical for larger peptides/proteins with a high α -helical content.^[62] When normalising the change of ellipticity at 222 nm (Figure S5, right) in respect to the Eu(III) amount, a very similar progression for all four peptides can be observed (Figure S5), suggesting that the peptides undergo structural changes to a similar degree at the same Eu(III) equivalents, supporting the observed similar binding affinities. In contrast to Eu(III), the addition of Ca(II) does not trigger any conformational change for any of the peptides. This again strongly supports the conclusions drawn from the ITC and TRLFS experiments: At the tested conditions the LanM peptides do not bind Ca(II) to a significant degree.

Figure 4 CD spectra of Eu(III) (top) and Ca(II) (bottom) titration experiments with all four EF-Hand peptides at pH 6.6 (10 mM MOPSO, 100 mM KCl).

The experimental observations made for Eu(III) were further supported by molecular dynamics simulations. Calculations of the four EF-Hand loop peptides were performed using two different sets of initial structures. A first set of MD simulations was performed by taking the local structures of the EF-Hand loop from the NMR structure of LanM^[23] (PDB 6mi5) and removing the residual protein. As the metal-binding loop of EF-Hand 4 is not occupied with a metal, in the structure for the peptide EF4 the NMR structure of EF-Hand loop 1 was taken and the sequence “mutated”. These calculations converged to structures (Table S9) in which Eu(III) is fully coordinated by amino acid residues and there is virtually no water except for EF4 in which two water molecules are coordinated to Eu(III). These calculations show the high potential of the amino acid sequence for Ln(III) coordination, but from our experimental observations it became clear that these calculations do not match the experimental evidence. From the experiments we observed micromolar binding affinity for Eu(III) with about 5 water molecules completing the Eu(III) coordination sites when in a 1:1 complex (calculated from the luminescence lifetimes, see Table 2 and Table S5) as well as very little secondary structure elements in the absence of metal which increases upon Eu(III)-binding (Figure 4). Therefore, a more realistic set of simulations was performed in the presence and absence of Eu(III) without any initial restraints. Each peptide was initially prepared as a “string” having the sequences of the corresponding EF-Hand loop and was allowed to freely wrap around Eu(III). This second set of simulations converged to a totally different set of structures than the first simulation experiment, thereby better matching the experimental data. Eu(III) is coordinated over less amino acid residues and on average about four to five water molecules are coordinated to Eu(III) (Table S10). The coordination of EF1 to Eu(III) based on the simulations is shown in Figure 5. Even though the association of the peptides to the metal is only through 2 to 4 residues (in contrast to 5 to 6 residues in structures obtained when starting from the crystal structure) the peptide structure is

well-preserved during the simulation in comparison to those without metals (Figure 6). In the absence of Eu(III) the peptide flexibility is very high and retains ill-defined secondary structures and only after the addition of Eu(III) upon metal-binding, a more organised structure is obtained, which is in good agreement with the observations made by CD spectroscopy (compare Figures 4 and 6). This observation is also further confirmed through DSSP analysis of the MD trajectories (Figure S18). Very recently, a new set of 12-6-4 Lennard-Jones-type parameters has been developed for better parametrization of the interaction between metals and specific ligating groups and tested on the EF-Hand loop EF1 of LanM and CaM.^[65] These parameters will be used in future investigations to refine the modelling.

Figure 5 Representative ball-and-stick drawing of Eu(III)-bound EF-Hand 1 from MD trajectory. Grey ribbon depicts the peptide backbone (orange: Eu, blue: N, black: C, red: O, white: H).

Figure 6 Superimposed MD snapshots of metal-free (upper panels) and Eu(III)-bound (lower panels, Eu in orange) bound to EF-Hands (grey ribbon) for every 1 ns of the 100 ns MD trajectory (waters excluded from display).

Eu(III) binding compared with Tb(III) and Cm(III) binding

Figure 7 TRLFS data: Spectral deconvolution (top) and lifetimes (bottom) for the peptide to metal titration experiments (Eu/Tb pH 6.6; Cm 6.0¹; 10 mM MOPSO, 100 mM KCl); light blue= full spectrum obtained by the used model, dark blue= raw spectrum; λ_{ex} (Eu)= 394 nm, λ_{ex} (Tb)= 220 nm, λ_{ex} (Cm) = 396 nm.

In addition to the extensive metal-binding studies using Eu(III) as an example for the Lns, this study was further completed by including another Ln, terbium. Furthermore, the binding properties of the EF-Hand loop peptides were tested with the actinide curium to see whether the even higher affinity of Ans over Lns which was observed for lanmodulin^[25–27] is also preserved in the short peptides. The comparison between the different metals with EF1 as an example can be seen in Figure 7. In contrast to Eu(III), Tb(III) does not show hypersensitivity like the $^5D_0 \rightarrow ^7F_0$ transition of Eu(III). A change in the coordination sphere causes only minor changes in the shape of the emission spectrum. Therefore, a direct readout of the Tb(III) interaction is complicated, and usually its complexation is only traced by changes in its lifetime. Here we can show that the minor changes in the emission spectra combined with the changes in the luminescence lifetimes can be used to decompose TRLFS Tb(III) titration experiments using PARAFAC thereby demonstrating its suitability for the methods described here. The experiment with Tb(III) further confirmed the Eu(III) experiments by showing the formation of one 1:1 complex. The obtained K_D values are as well in the lower micromolar range suggesting a very similar binding affinity for Tb(III) as well as Eu(III) (see Table 2). Overall, the Tb(III) titrations strongly support the observations already made with Eu(III).

The analogously conducted experiment with Cm(III) again pointed out the similarity of the four different peptides as overall binding affinities in the same range were obtained (Table 2). A comparison of Cm(III) and Eu(III) luminescence lifetimes after interaction with the full protein^[25] and isolated EF-Hand loop peptides further supports the current conclusions from experiment and simulation. The luminescence lifetimes were significantly shorter for the EF-Hand loop peptides compared to the full protein (180 μs vs 370 μs ^[25] for Eu(III) and 140 μs vs 200 μs ^[25] for Cm(III)). This indicates more water and less peptide functionalities in the first coordination sphere of the metal centre. The analysis suggested that the affinity for Cm(III) is slightly higher than for both investigated lanthanides (2–3 μM). In addition to the fitted affinity the calculated water molecules based on the luminescence lifetime also show on average one water molecule less which is coordinated to Cm(III) when compared to Eu(III) again indicating

¹ pH needed to be adjusted as at a pH of 6.6 solely the free Cm(III)-aquo ion was observed.

a tighter binding. This nicely fits into the observations made in the literature for LanM which all showed that the naturally Ln-binding protein favours actinides.^[25–27] Our data supports the assumption that this preference is preserved even in these short peptides.

Table 2 K_D values and lifetimes (τ) obtained by EF to metal TRLFS titration experiments. The number of coordinated water molecules (q) was calculated for Eu(III) and Cm(III) by using the equations established by Horrocks^[66,67] and Kimura^[67], respectively.

EF-Eu(III) 1:1 complex			EF-Tb(III) 1:1 complex		EF-Cm(III) 1:1 complex	
	K_D (μM)	τ (μs) q_{water}	K_D (μM)	τ (μs)	K_D (μM)	τ (μs) q_{water}
EF1	5.50 ± 0.57	186.2 ± 3.7 5.1	8.47 ± 0.325	450 ± 18.0	2.28 ± 0.104	143.1 ± 19 3.7
EF2	6.52 ± 0.48	181.8 ± 8.6 5.3	8.24 ± 1.31	469 ± 5.30	2.86 ± 0.084	147.7 ± 18 3.5
EF3	7.42 ± 1.04	182.1 ± 2.2 5.3	8.46 ± 0.864	502 ± 11.0	2.77 ± 0.229	137.4 ± 15 3.9
EF4	7.56 ± 1.49	186.3 ± 4.3 5.1	7.85 ± 1.23	502 ± 4.00	2.34 ± 0.088	142.7 ± 3.0 3.7

Binding competition: Lanthanide versus Actinide

To confirm the higher affinity for Cm(III) over Tb(III) and Eu(III) found in TRLFS experiments, a “fair” competition experiment was envisioned in which the peptides were added in increasing amounts to a solution containing the same concentration of Cm(III) and Eu(III). To the best of our knowledge, this is the first time that such an experiment in which both metals were simultaneously probed was performed. One difficulty in such a set-up is the higher sensitivity of Cm(III) (higher quantum yield) in comparison to Eu(III). To optimise the weighting of the detected signal towards Eu(III), the excitation wavelength of Eu(III) was chosen for both metals. This was important, as both the Cm(III) as well as the Eu(III) emissions superimpose which makes the data deconvolution extra challenging. Nevertheless, we were able to analyse the measured data by using PARAFAC and obtained single component emission spectra, corresponding lifetimes and the speciation. The analysis showed the expected formation of the 1:1 peptide/Eu(III) and peptide/Cm(III) complexes as well as the presence of the respective free aquo ions (see Figure 8). In this direct setting, we were able to confirm the slightly higher affinity for Cm(III) over Eu(III).

In the future, modifications of the natural EF-Hand loop peptides towards higher selectivity for Ans over Lns can be envisioned. It was already shown in the literature^[68] for LBTs that by substituting the aspartic acid in position 5 with a cysteine or a 2-methylenepyridine group the selectivity for Ans over Lns can be increase by a factor of 10. Although the affinity of the short peptides is significantly reduced when compared to the full protein LanM, remarkable affinities for such short peptides were reached. For the full protein calmodulin – a natural Ca(II)-binding protein which also has a higher affinity for Lns/Ans than Ca(II) – similar average K_D values are reported in the literature^[61] ($6.6 \mu\text{M}$ for Eu(III)) as for the isolated EF-Hand loop peptides of LanM. This indicates the great potential of these sequences in *f*-element accumulation processes as required by e.g. efficient recycling strategies.

Figure 8 TRLFS data of the titration experiment of EF1 to Cm(III)/Eu(III), $\lambda_{\text{ex}}(\text{Eu/Cm}) = 393.9 \text{ nm}$. Spectral deconvolution of the different species (top), lifetimes (middle) and speciation (bottom).

3 Conclusions

In this study an extremely consistent picture of the Ln and An-binding capabilities of the LanM-based 12amino acid peptides (EF-Hand loops) could be drawn. We were able to show that the Ln-affinity, the preference for actinides and the low affinity for Ca(II) is preserved even in these short peptide sequences. Although the high affinity of LanM, which has been reported^[22] to be in the picomolar range is not observed for its EF-Hand loop peptides, their affinity is remarkably high, especially considering them not being modified or optimised. These short peptides even reach affinities in the micromolar range, similar to the average affinity of the four calmodulin binding sites which is known to also have a higher affinity for Eu(III)/Cm(III) over Ca(II)^[61]. Furthermore, the investigation without cooperative and pre-structuring effects of the full protein, enabled us to study the impact of the amino acid sequence on the Ln-affinity. We show that all four peptides behave very similarly, strongly supporting the assumption that the major differences in Ln-affinity observed in the literature^[22,24] for the different EF-Hands is mainly due to cooperativity and hydrogen bonding. Furthermore, we demonstrate that these peptides using the sequences optimised by nature to bind Lns are a good starting point for the development of selective and highly affine peptides which could find applications in medicine, recycling and An/Ln separation, all of which are still urgent issues of our times. One major difference between the metal-binding loops in the full protein and the isolated peptides is the

flexibility. Therefore, for future work a restriction of the flexibility could be envisioned leading to higher affinities and enhanced selectivity.

4 Acknowledgements

SMG thanks the Studienstiftung des Deutschen Volkes for funding, Manuel Gebauer for his support as Hiwi with the peptide purifications and Falk Lehmann for conducting spontaneous MS measurements at the HZDR. LJD acknowledges the ERC Starting Grant Lanthanophor (945846) and the Klaus Tschira Boost Fund. Furthermore, the authors wish to thank the TU Dresden for computation time as all MD calculations were performed at the Centre for Information Services and High-Performance Computing (ZIH) at the Technische Universität Dresden. Additionally, the authors wish to thank Prof. Peter Klüfers for access to the CD Instrument.

5 References

- [1] T. Cheisson, E. J. Schelter, *Science* **2019**, 363, 489–493.
- [2] V. Balaram, *Geosci. Front.* **2019**, 10, 1285–1303.
- [3] Z. Weng, S. M. Jowitt, G. M. Mudd, N. Haque, *Econ. Geol.* **2015**, 110, 1925–1952.
- [4] J. H. L. Voncken, *The Rare Earth Elements: An Introduction*, Springer International Publishing, **2015**.
- [5] N. Haque, A. Hughes, S. Lim, C. Vernon, *Resources* **2014**, 3, 614–635.
- [6] S. H. Ali, *Resources* **2014**, 3, 123–134.
- [7] S. M. Jowitt, T. T. Werner, Z. Weng, G. M. Mudd, *Curr. Opin. Green Sustain. Chem.* **2018**, 13, 1–7.
- [8] A. Pol, T. R. M. Barends, A. Dietl, A. F. Khadem, J. Eygensteyn, M. S. M. Jetten, H. J. M. Op den Camp, *Environ. Microbiol.* **2014**, 16, 255–264.
- [9] T. Nakagawa, R. Mitsui, A. Tani, K. Sasa, S. Tashiro, T. Iwama, T. Hayakawa, K. Kawai, *PLOS ONE* **2012**, 7, e50480.
- [10] A. M. Ochsner, L. Hemmerle, T. Vonderach, R. Nüssli, M. Bortfeld-Miller, B. Hattendorf, J. A. Vorholt, *Mol. Microbiol.* **2019**, 111, 1152–1166.
- [11] N. C. Martinez-Gomez, H. N. Vu, E. Skovran, *Inorg. Chem.* **2016**, 55, 10083–10089.
- [12] N. M. Good, H. Lee, E. R. Hawker, A. A. Gilad, N. C. Martinez-Gomez, *bioRxiv* **2021**, 2021.06.12.448192.
- [13] J. T. Keltjens, A. Pol, J. Reimann, H. J. M. Op den Camp, *Appl. Microbiol. Biotechnol.* **2014**, 98, 6163–6183.
- [14] L. J. Daumann, *Angew. Chem. Int. Ed.* **2019**, 58, 12795–12802.
- [15] O. Adachi, K. Matsushita, E. Shinagawa, M. Ameyama, *Agric. Biol. Chem.* **1990**, 54, 3123–3129.
- [16] G. E. Cozier, C. Anthony, *Biochem. J.* **1995**, 312, 679–685.
- [17] M. Ghosh, C. Anthony, K. Harlos, M. G. Goodwin, C. Blake, *Structure* **1995**, 3, 177–187.
- [18] C. Anthony, P. Williams, *Biochim Biophys Acta.* **2003**, 1647, 18–23.
- [19] K. N. Allen, B. Imperiali, *Curr. Opin. Chem. Biol.* **2010**, 14, 247–254.
- [20] S. Lim, S. J. Franklin, *Cell. Mol. Life Sci.* **2004**, 61, 2184–2188.
- [21] J. L. Gifford, M. P. Walsh, H. J. Vogel, *Biochem. J.* **2007**, 405, 199–221.
- [22] J. A. Cotruvo, Jr., E. R. Featherston, J. A. Mattocks, J. V. Ho, T. N. Laremore, *J. Am. Chem. Soc.* **2018**, 140, 15056–15061.
- [23] E. C. Cook, E. R. Featherston, S. A. Showalter, J. A. Cotruvo, Jr., *Biochemistry* **2019**, 58, 120–125.
- [24] E. R. Featherston, E. J. Issertell, J. A. Cotruvo, *J. Am. Chem. Soc.* **2021**, 143, 14287–14299.
- [25] H. Singer, B. Drobot, C. Zeymer, R. Steudtner, L. J. Daumann, *Chem. Sci.* **2021**, 15581–15587.

- [26] G. J.-P. Deblonde, J. A. Mattocks, H. Wang, E. M. Gale, A. B. Kersting, M. Zavarin, J. A. Cotruvo, *J. Am. Chem. Soc.* **2021**, *143*, 15769–15783.
- [27] G. J.-P. Deblonde, J. A. Mattocks, Z. Dong, P. T. Wooddy, J. A. Cotruvo, M. Zavarin, *Sci. Adv.* **2021**, *7*, eabk0273.
- [28] J. T. Welch, W. R. Kearney, S. J. Franklin, *Proc. Natl. Acad. Sci.* **2003**, *100*, 3725–3730.
- [29] M. Xu, Z. Su, J. N. Renner, *Pept. Sci.* **2019**, *111*, e24133.
- [30] J. Gariepy, B. D. Sykes, R. S. Hodges, *Biochemistry* **1983**, *22*, 1765–1772.
- [31] J. Wójcik, J. Góral, K. Pawłowski, A. Bierzynski, *Biochemistry* **1997**, *36*, 680–687.
- [32] M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen, B. Imperiali, *Angew. Chem.* **2004**, *116*, 3768–3771.
- [33] L. J. Martin, B. R. Sculimbrene, M. Nitz, B. Imperiali, *QSAR Comb. Sci.* **2005**, *24*, 1149–1157.
- [34] K. J. Franz, M. Nitz, B. Imperiali, *ChemBioChem* **2003**, *4*, 265–271.
- [35] M. Nitz, K. J. Franz, R. L. Maglathlin, B. Imperiali, *ChemBioChem* **2003**, *4*, 272–276.
- [36] B. R. Sculimbrene, B. Imperiali, *J. Am. Chem. Soc.* **2006**, *128*, 7346–7352.
- [37] T. Hatanaka, N. Kikkawa, A. Matsugami, Y. Hosokawa, F. Hayashi, N. Ishida, *Sci. Rep.* **2020**, *10*, 19468.
- [38] L. Le Clainche, C. Vita, *Env. Chem Lett* **2006**, *4*, 45–49.
- [39] S. Sauge-Merle, F. Brulfert, R. Pardoux, P. L. Solari, D. Lemaire, S. Safi, P. Guilbaud, E. Simoni, M. L. Merroun, C. Berthomieu, *Chem. Eur. J.* **2017**, *23*, 15505–15517.
- [40] R. Pardoux, S. Sauge-Merle, D. Lemaire, P. Delangle, L. Guilloreau, J.-M. Adriano, C. Berthomieu, *PLOS ONE* **2012**, *7*, e41922.
- [41] G. J.-P. Deblonde, J. A. Mattocks, D. M. Park, D. W. Reed, J. A. Cotruvo, Y. Jiao, *Inorg. Chem.* **2020**, *59*, 11855–11867.
- [42] Z. Dong, J. A. Mattocks, G. J.-P. Deblonde, D. Hu, Y. Jiao, J. A. Cotruvo, D. M. Park, *ACS Cent. Sci.* **2021**, *7*, 1798–1808.
- [43] S. J. Caldwell, I. C. Haydon, N. Piperidou, P.-S. Huang, M. J. Bick, H. S. Sjöström, D. Hilvert, D. Baker, C. Zeymer, *Proc. Natl. Acad. Sci.* **2020**, *117*, 30362–30369.
- [44] L. N. Slope, O. J. Daubney, H. Campbell, S. A. White, A. Peacock, *Angew. Chem. Int. Ed.* **2021**, *60*, 14473–24477.
- [45] H. Lumpe, A. Menke, C. Haisch, P. Mayer, A. Kabelitz, K. V. Yussenko, A. G. Buzanich, T. Block, R. Pöttgen, F. Emmerling, L. J. Daumann, *Chem. – Eur. J.* **2020**, *26*, 10133–10139.
- [46] D. B. Halling, B. J. Liebeskind, A. W. Hall, R. W. Aldrich, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E1216–E1225.
- [47] S. K. Drake, K. L. Lee, J. J. Falke, *Biochemistry* **1996**, *35*, 6697–6705.
- [48] S. K. Drake, J. J. Falke, *Biochemistry* **1996**, *35*, 1753–1760.
- [49] R. C. Hapak, P. J. Lammers, W. A. Palmisano, E. R. Birnbaum, M. T. Henzl, *J. Biol. Chem.* **1989**, *264*, 18751–18760.
- [50] S. Liu, E. R. Featherston, J. A. Cotruvo, C. R. Baiz, *Phys. Chem. Chem. Phys.* **2021**, 21690–21700.
- [51] D. Muralidhar, M. Kunjachen Jobby, A. Jeromin, J. Roder, F. Thomas, Y. Sharma, *Peptides* **2004**, *25*, 909–917.
- [52] R. Dölling, M. Beyermann, J. Haenel, F. Kernchen, E. Krause, P. Franke, M. Brudel, M. Bienert, *J. Chem. Soc. Chem. Commun.* **1994**, 853–854.
- [53] R. Behrendt, P. White, J. Offer, *J. Pept. Sci.* **2016**, *22*, 4–27.
- [54] J. L. Lauer, C. G. Fields, G. B. Fields, *Lett. Pept. Sci.* **1995**, *1*, 197–205.
- [55] D. Samson, D. Rentsch, M. Minuth, T. Meier, G. Loidl, *J. Pept. Sci.* **2019**, *25*, e3193.
- [56] T. Michels, R. Dölling, U. Haberkorn, W. Mier, *Org. Lett.* **2012**, *14*, 5218–5221.

- [57] R. E. Reid, D. M. Clare, R. S. Hodges, *J. Biol. Chem.* **1980**, 255, 3642–3646.
- [58] S. Yokokawa, S. Tsubuki, H. Ito, H. Kasai, *Chem. Lett.* **1989**, 18, 1627–1630.
- [59] P. Kanellis, J. Yang, H. C. Cheung, R. E. Lenkinski, *Arch. Biochem. Biophys.* **1983**, 220, 530–540.
- [60] N. A. Malik, G. M. Anantharamaiah, A. Gawish, H. C. Cheung, *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1987**, 911, 221–230.
- [61] B. Drobot, M. Schmidt, Y. Mochizuki, T. Abe, K. Okuwaki, F. Brulfert, S. Falke, S. A. Samsonov, Y. Komeiji, C. Betzel, T. Stumpf, J. Raff, S. Tsushima, *Phys Chem Chem Phys* **2019**, 21, 21213–21222.
- [62] W. C. Johnson, *Annu. Rev. Biophys. Biophys. Chem.* **1988**, 145–166.
- [63] R. E. Reid, J. Gariépy, A. K. Saund, R. S. Hodges, *J. Biol. Chem.* **1981**, 256, 2742–2751.
- [64] G. Borin, P. Ruzza, M. Rossi, A. Calderan, F. Marchiori, E. Peggion, *Biopolymers* **1989**, 28, 353–369.
- [65] P. Kantakevičius, C. Mathiah, L. Johannissen, S. Hay, **2021**, DOI 10.33774/chemrxiv-2021-l3g6h.
- [66] W. DeW. Horrocks, D. R. Sudnick, *J. Am. Chem. Soc.* **1979**, 101, 334–340.
- [67] T. Kimura, G. R. Choppin, *J. Alloys Compd.* **1994**, 213–214, 313–317.
- [68] S. Özçubukçu, K. Mandal, S. Wegner, M. P. Jensen, C. He, *Inorg. Chem.* **2011**, 50, 7937–7939.