Chem Soc Rev

REVIEW ARTICLE

View Article Online
View Journal



Cite this: DOI: 10.1039/d4cs01148a

# Lanthanide(III)-binding peptides and proteins: coordination properties and applications

Enrico Falcone, *† Emilie Mathieu *† and Christelle Hureau

Lanthanides play a crucial role in modern medicine and technology as well as in the metabolism of methylotrophic bacteria. In this context, the research on lanthanide-binding peptides and proteins is an active and rapidly developing field. This comprehensive and critical review focuses on the structural, thermodynamic (affinity and selectivity) and kinetic parameters governing the interaction of $Ln^{3+}$ ions with different peptides and proteins, including both naturally occurring and *de novo*-designed scaffolds. It thus provides guidelines and future directions for the rational design of Ln-binding peptides and proteins with suitable features for the main applications explored to date, including luminescent sensing, magnetic resonance imaging, Ln separation and recovery and Ln-based (photo)-catalysis.

Received 21st December 2024

DOI: 10.1039/d4cs01148a

rsc.li/chem-soc-rev



## 1. Introduction

### 1.1 Context and scope

Lanthanides (Ln) are f-block elements known for their unique magnetism and luminescence properties. They are used in a wide variety of fields: as contrast agents in medicine, in laser technology, in superconductors and in permanent magnets.[1] Rare earth elements (REE), which include Ln, lanthanum (La), yttrium (Y) and scandium (Sc), have been identified as strategic resources by the European Union (EU).[2] Indeed, many advanced technologies depend on them, and China remains the main exporter of these resources, supplying over 98% of REE importation to the EU. Of note, the mining and extraction of REE have a deleterious impact on the environment.[3]

Even before the emergence of interest in Ln biochemistry due to the discovery of the first Ln-utilizing bacteria in the 2010s,[4–7] the design of $Ln^{3+}$-binding peptides was an intensive research field aimed at exploiting the unique physical properties of $Ln^{3+}$ ions for applications in sensing or imaging.[8]

The discovery of Ln-utilizing bacteria has been accompanied by the identification of natural $Ln^{3+}$-binding proteins and $Ln^{3+}$–enzymes, paving the way for a better description of the lanthanome, which corresponds to the proteins involved in $Ln^{3+}$-uptake, trafficking and utilisation.[9–12] These findings have inspired a large community of biologists and chemists, resulting in a proliferation of studies on the rational design of Ln-binding peptides and proteins, and their applications in biomedicine,[13] biohydrometallurgy,[14,15] and catalysis.[16–21]

Previous reviews in the field focused on the rational design of lanthanide-binding peptides,[8] and a general presentation of the coordination of f-block elements to bio-relevant ligands, including peptides and proteins.[22] Here, we first discuss some critical parameters of $Ln^{3+}$ aqueous chemistry and the design of peptides and proteins with a high affinity for $Ln^{3+}$ ions. Then, we provide guidelines on how selectivity against metal ions ($Ca^{2+}$, actinides ($An^{3+}$), d-block metal ions) or among $Ln^{3+}$ can be achieved, and on how to include kinetic considerations in the description of these systems. Last, we highlight recent applications for sensing, MRI, Ln separation and catalysis. In this review, we focus on $Ln^{3+}$-binding peptides and proteins using natural and/or unnatural amino acids, excluding peptides and proteins functionalized with macrocycles such as DOTA, some of which were reviewed in ref. 23. This choice was driven by the focus on the interplay between the structure and the coordination properties (affinity, selectivity and kinetics) of $Ln^{3+}$-binding peptides and proteins scaffold, which do not apply to macrocycles. Nevertheless, small chelators will be occasionally discussed to introduce key features and concepts (*e.g.* selectivity) useful for Ln-binding peptides and proteins, and peptides with appended macrocycles will be briefly mentioned among the applications, as they represent useful models for the design of luminescent and MRI probes.

### 1.2 Chemistry of Ln ions in aqueous solution

The most stable oxidation state for Ln in aqueous solution is +3 (electronic configuration $[Xe]4f^n$, with $n$ = 0–14 from $La^{3+}$ to $Lu^{3+}$). Other oxidation states are also accessible for some Ln ions in aqueous solution, such as +2 for Eu and +4 for Ce, due to the enhanced stability associated with a half-filled 4f subshell ($[Xe]4f^7$ for $Eu^{2+}$) and with closed shell ($[Xe]4f^0$ for $Ce^{4+}$) configuration (Fig. 1A). Of note, in organic solvents, the entire

*LCC-CNRS, Université de Toulouse, CNRS, Toulouse, France.*
*E-mail: enrico.falcone@lcc-toulouse.fr, emilie.mathieu@lcc-toulouse.fr*
† These authors contributed equally to this work.

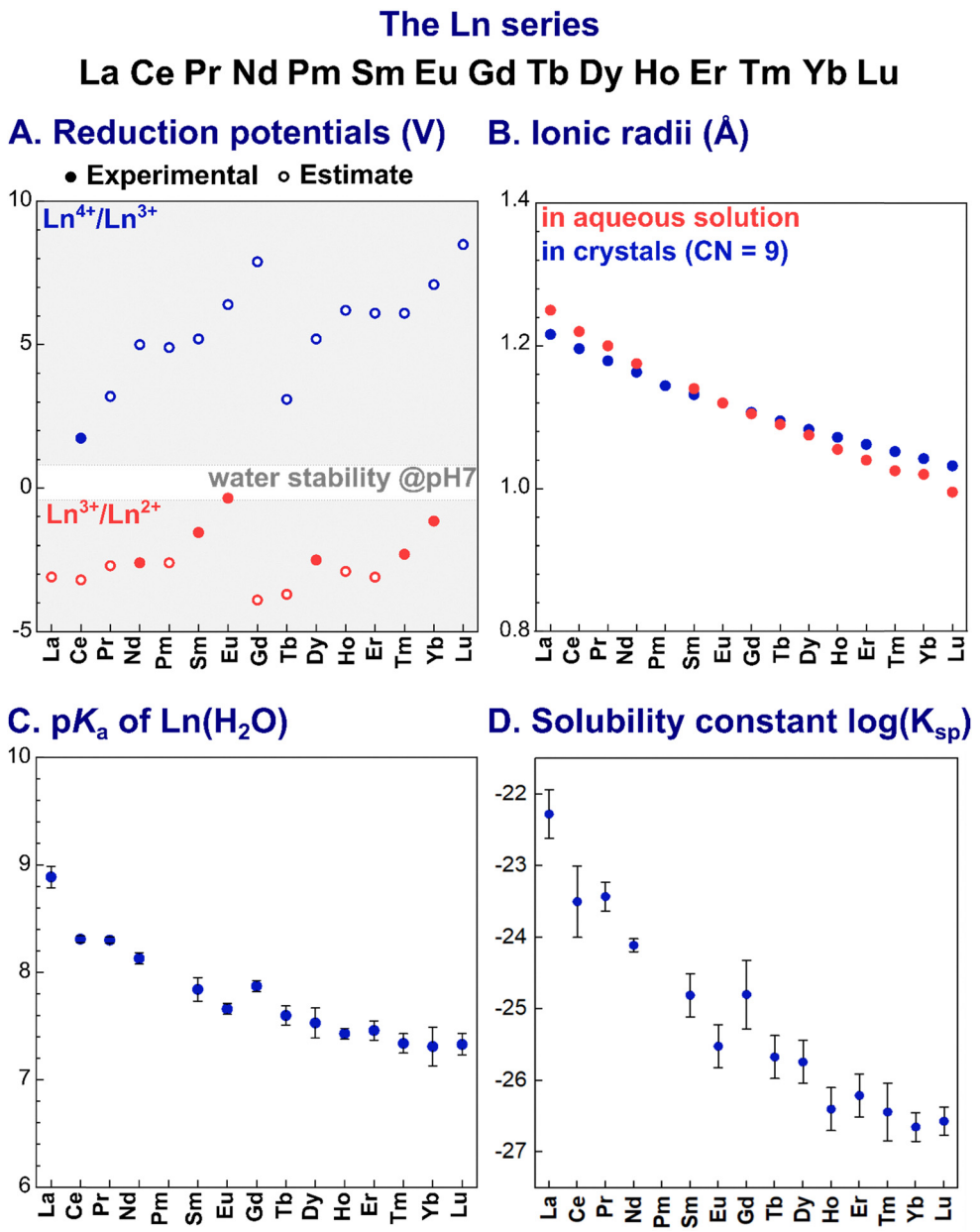


Fig. 1 Trends of key parameters along the Ln series: (A) reduction potentials;[26] (B) ionic radii in aqueous solution[27] and in crystals (CN = 9);[28] (C) acidity constants ($pK_a$) for $Ln^{3+}$-bound water molecule; and (D) solubility constants ($\log K_{sp}$).[29] The values plotted are reported in Table 14.

series of Ln (except for the radioactive Pm) could be stabilised in the +2 oxidation state, while for the +4 oxidation state only $Ce^{4+}$, $Tb^{4+}$ and $Pr^{4+}$ complexes were isolated.[24]

$Ln^{3+}$ ions are hard Lewis acids. Their Lewis acidity slightly increases along the series as a consequence of the decrease of the ionic radius, known as lanthanide contraction (Fig. 1B). Noteworthy, $Ln^{3+}$ ions have a similar size to $Ca^{2+}$ ($\sim$1.0 Å), but are larger than d-block metal ions. As a consequence, $Ln^{3+}$ have higher coordination numbers (CN) than d-block elements.

Due to their hard character, according to Pearson's HSAB principle,[25] $Ln^{3+}$ ions prefer hard negatively charged ligands (*e.g.* carboxylate and phosphate). As a result of the core nature of the f orbitals, the nature of the $Ln^{3+}$–ligand bond is mainly ionic, and coordination geometries are hence mostly dictated by electrostatic and steric repulsions between the ligands.

In water, the larger $Ln^{3+}$ ions ($La^{3+}$–$Nd^{3+}$) are coordinated by 9 water molecules and adopt a tri-capped trigonal prismatic coordination geometry, whereas smaller $Ln^{3+}$ ($Gd^{3+}$–$Lu^{3+}$) are bound to 8 water molecules in a square-antiprismatic arrangement. For $Pm^{3+}$ to $Eu^{3+}$, $[Ln(H_2O)_8]^{3+}$ and $[Ln(H_2O)_9]^{3+}$ are in equilibrium.[30]

Importantly, as $Ln^{3+}$ aqua ions display Brønsted acidity in aqueous solution, $Ln^{3+}$ speciation depends on the pH, with the acidity constant $pK_a$ (eqn (1)) decreasing along the series

 

(Fig. 1C):

$$[Ln(H_2O)_n]^{3+} + H_2O \rightleftarrows [Ln(H_2O)_{n-1}(OH)]^{2+} + H_3O^+$$

$$K_a = \frac{[Ln(H_2O)_{n-1}(OH)][H_3O^+]}{[Ln(H_2O)_n]} \quad (1)$$

The stepwise deprotonation of bound water molecules leads to the formation of poorly soluble $Ln^{3+}$ hydroxides, $Ln(OH)_3$. In particular, the solubility (expressed by the solubility product $K_{sp}$, eqn (2)) decreases along the Ln series (Fig. 1D).

$$Ln(OH)_3\ (s) \rightleftarrows Ln^{3+}(aq) + 3HO^- \qquad K_{sp} = [Ln^{3+}][HO^-]^3 \quad (2)$$

In light of the trends described above, it appears that $Ln^{3+}$ share some properties: they are all predominantly found in their +3 redox state and have similar Lewis acidities due to their similar ionic radii. At a more detailed level, $Ln^{3+}$ display differences in their chemical properties: along the series, $Ln^{3+}$ Lewis acidity increases, which can modulate their reactivity, the $pK_a$ of the $Ln^{3+}$-bound water molecule decreases by approximately 2 units (Fig. 1C) and the solubility of $Ln(OH)_3$ drops from almost 5 order of magnitude (Fig. 1D); moreover Ln ions have different redox properties (Fig. 1A), which have been capitalized on for catalysis (see 6.4 Catalysis).

## 2. Scaffolds structure & $Ln^{3+}$ coordination

### 2.1 Generalities

**2.1.1 Type of scaffolds.** The binding of $Ln^{3+}$ ions to peptides and proteins has been accomplished either by using existing proteins with adequate coordination environment (*e.g.* 2.2 EF-hand peptides and proteins and 2.5.1 Naturally occurring proteins) or by the rational design of a $Ln^{3+}$-coordination site using natural and/or unnatural amino-acids (*e.g.* 2.3 Short peptides and 2.4 Peptide scaffolds of intermediate size). These scaffolds range from sequences of a few amino acids (AAs) to proteins, and can be divided into two main categories (Fig. 2): (i) those relying on an "EF-hand-like" motif for Ln-binding (2.2 EF-hand peptides and proteins); and (ii) other types of scaffolds.

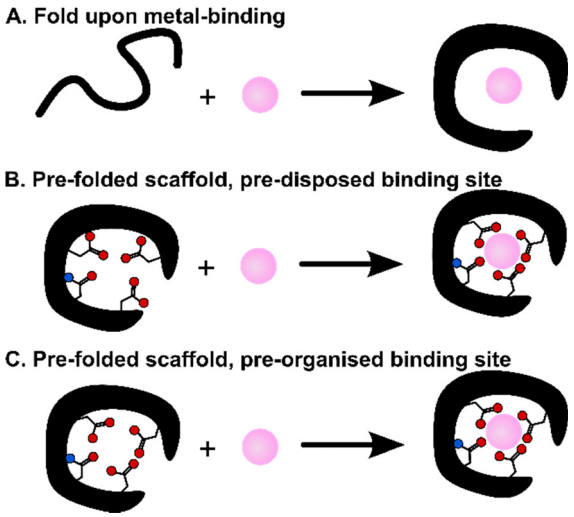


Fig. 3 Scaffolds that fold upon metal-binding (A), or are folded in the apo-form with a pre-disposed (B) or a pre-organized (C) binding site.

These scaffolds can be also differentiated based on their structure in the apo-form, *i.e.* in the absence of metal ions (Fig. 3). Whereas some are disordered in the apo-form and fold upon metal binding, others are pre-folded even in the absence of metal ions. In this latter case, the binding site can be (i) pre-disposed, *i.e.* the coordinating residues are placed at appropriate positions for metal-binding, or (ii) pre-organised, *i.e.* the coordinating residues are not only at an appropriate position but also with a good orientation that requires minimal reorganisation upon metal-binding.[31] To our knowledge, there is no example of a pre-organised metal binding site in $Ln^{3+}$-binding peptides and proteins, and hence, in the following, we will only describe unfolded scaffolds, or folded ones with a predisposed binding site. In order to discuss the scaffold structure and the $Ln^{3+}$-coordination, we grouped the scaffolds into four sub-categories: (i) scaffolds relying on the well-described EF-hand motifs, which are found in short peptides such as lanthanide-binding tags (LBT), in helix-loop-helix (HLH) peptides, and naturally-occurring proteins such as calmodulin and lanmodulin; (ii) short sequences of ~10 amino acids; (iii) peptides of intermediate size, spanning from Ln fingers (LF) to coiled coils (CC); (iv) $Ln^{3+}$-binding proteins either naturally occurring or

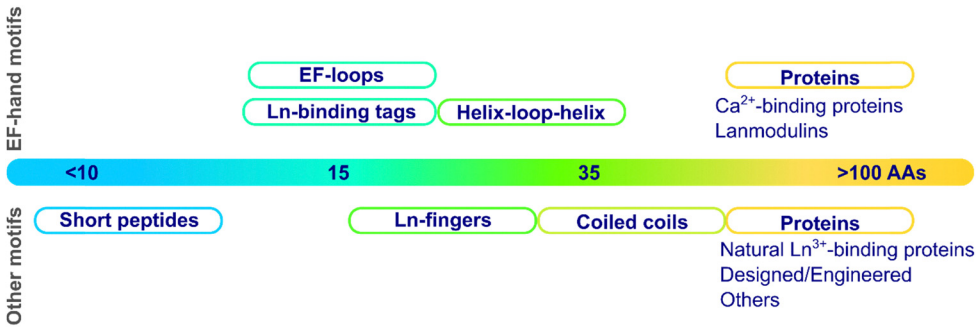


Fig. 2 Categories of scaffolds by size (number of amino acids).

engineered, in which the $Ln^{3+}$ ion is not coordinated in an EF-hand motif.

**2.1.2 Characterization of $Ln^{3+}$ coordination sphere and scaffold stability.** Several techniques give access to information on the $Ln^{3+}$ coordination sphere and on the stability of the apo-peptide and Ln–peptides, which are shortly presented below.

*$Ln^{3+}$ coordination sphere.* X-ray crystallography is the most accurate way to determine the $Ln^{3+}$ coordination sphere, especially the CN and the mode of binding of certain amino acids (*e.g.* mono- or bi-dentate for carboxylates in Asp and Glu residues). However, it requires the formation of crystals, which are often difficult to obtain, in particular when working with peptides.

Nevertheless, information on the mono- or bi-dentate binding mode of carboxylates can also be obtained in solution through FT-IR measurements.[32,33] Moreover, the CN may be also investigated in solution through X-ray absorption spectroscopy, specifically EXAFS (extended X-ray absorption fine structure),[34–37] which has not been applied to $Ln^{3+}$-binding peptides and proteins so far.

In addition, insights into the $Ln^{3+}$ coordination sphere can be gained by luminescence measurements. $Eu^{3+}$-hypersensitive transitions can report on the geometry of the coordination environment.[38] The hydration number $q$, which corresponds to the number of $Ln^{3+}$-bound water molecules, is also easily determined by measuring $Ln^{3+}$ luminescence lifetimes (Ln = $Eu^{3+}$, $Tb^{3+}$, $Yb^{3+}$) for a given complex in $H_2O$ and $D_2O$ and using empirical relationships,[39,40] which have been established based on a library of complexes with $q$ ranging from 0 to 6. This is the most solid method and should be favoured compared to other empirical equations that rely only on lifetimes measured in water. Importantly, when working with peptides and proteins it may be difficult to work in $H_2O$-free conditions. Instead, measurements can be performed for different $H_2O:D_2O$ ratios in order to extrapolate the value in 100% $D_2O$.[41]

Information beyond the number of $Ln^{3+}$-bound water molecules can be gained from NMR measurements, by taking advantage of the magnetic properties of lanthanides.[42–46] Comparison of spectra obtained with diamagnetic $Ln^{3+}$ (La, Lu) and paramagnetic ones (all the other) can be used to identify residues in the first coordination sphere or further away from the $Ln^{3+}$-binding site. For well-folded scaffolds, paramagnetic Ln can even help to refine the protein solution structure. This has been accomplished either by incorporating $Ln^{3+}$ in an intrinsic metal-binding site, such as in $Ca^{2+}$-binding proteins,[47,48] or by introducing Ln-binding tags (LBT, see Section 2.2.2) in the sequence of the targeted protein.[46]

*Scaffolds stability.* Information on scaffold stability can be gained from denaturation experiments, either using a temperature ramp or chemical denaturants, followed by circular dichroism. By comparison to a ''native'' scaffold, these measurements can inform on the impact of engineering a $Ln^{3+}$-binding site into the scaffold and the range of temperatures for which the peptide/protein is folded. It is also helpful to quantitatively assess the stabilisation induced by rational optimisations of the scaffold and by $Ln^{3+}$ binding.

### 2.2 EF-hand motifs

EF-hand motifs represent the largest and best-known family of $Ln^{3+}$-binding sites in peptides/proteins. They are found in most $Ca^{2+}$-binding proteins, including calmodulin (CaM), parvalbumin and troponin C (TnC), and in a recently discovered selective Ln-binding protein, lanmodulin. The EF-hand consists of two α-helices (almost perpendicular to each other, a topology that reminds the spread thumb and forefinger of human hands) linked by a 12-residues loop, which is responsible for $Ca^{2+}$ and $Ln^{3+}$ binding through the amino acids in positions 1, 3, 5, 7, 9 and 12 (Fig. 4).[49–53] In the following paragraphs, EF-hand proteins and peptides are described starting from long-known $Ca^{2+}$-binding proteins (Section 2.2.1), which have inspired the design of optimized $Ln^{3+}$-binding peptides (see 2.2.2 LBTs and 2.2.3 HLHs), and finishing by a recently discovered family of $Ln^{3+}$-selective proteins (2.2.4 Lanmodulins).

**2.2.1 $Ca^{2+}$-binding proteins.** In most EF-hand proteins, $Ca^{2+}$ is coordinated by 7 oxygen ligands with a pentagonal bipyramidal geometry: residues in positions 1, 3, and 5 serve as monodentate ligands, while the $Glu_{12}$ residue binds always in a bidentate fashion (Fig. 4 and entry 1 in Table 1). The amino acid in position 7 (Thr in most natural scaffolds) binds instead with its backbone carbonyl. Finally, the 9th residue, which is most commonly (~30%) Asp, does not bind directly to the metal ion, but it forms a hydrogen bond with a water molecule that completes the coordination sphere.

Since $Ln^{3+}$ ions have similar ionic radius to $Ca^{2+}$, the luminescent properties of some $Ln^{3+}$ ions, mostly $Tb^{3+}$ and $Eu^{3+}$, have been exploited to probe the metal binding sites in $Ca^{2+}$-binding proteins,[54–57] showing that the same coordination sphere with CN = 7 and $q$ = 1 is observed for $Ln^{3+}$ ions bound to $Ca^{2+}$–proteins such as TnC (Fig. 4 and entry 1 in Table 1).[52,58,59]

**2.2.2 Lanthanide-binding tags (LBT).** Following the seminal work of Szabo and coworkers,[66] the group of Imperiali has widely exploited the $Ln^{3+}$-binding ability of EF-hand loops found in $Ca^{2+}$-binding proteins to develop so-called lanthanide-binding tags (LBT) for various applications (see 6. Applications).[60,67–71]

LBTs are 15- to 20-residue-long peptides that were designed by elongating the core 12-residue EF-loop at both termini with apolar amino acids that stabilize the $Ln^{3+}$-bound conformation *via* hydrophobic interactions. The primary sequence of LBTs was optimized to feature high affinity for $Tb^{3+}$ (see 3.2 EF-hand motifs),[72,73] and to control the hydration number for specific applications. For instance, LBT featuring a $Ln^{3+}$-bound water molecule (entries 2–4, in Table 1) were explored for applications as MRI contrast agents (see 6.2 MRI). The crystal structure of $Gd^{3+}$ bound to a double LBT-ubiquitin fusion construct showed a CN = 7, which is as low as that observed in $Ca^{2+}$-binding proteins.[60]

As $Ln^{3+}$-bound water molecules quench its luminescence emission (see 6.1 Luminescent tags and probes), the

 

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| D/N | X | D/N | X | D/N | X | T | X | D/E | X | X | D/E |

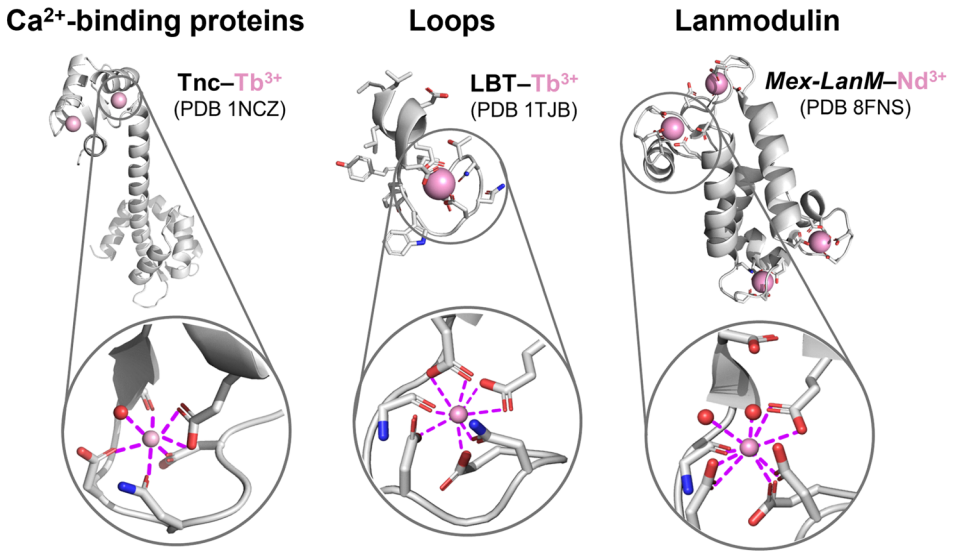


Fig. 4 EF-hand peptides and proteins. Top: EF-loop sequence showing the most common $Ln^{3+}$-binding amino acids (in bold and highlighted in pink) in each position. Bottom: Protein structures and $Ln^{3+}$ coordination spheres in: left, troponin C (TnC) bound to two $Tb^{3+}$ ions (PDB 1NCZ); middle, a lanthanide-binding tag (LBT) bound to $Tb^{3+}$ (PDB 1TJB); right, *Mex*-LanM bound to four $Nd^{3+}$ ions (PDB 8FNS). Figures were generated using Pymol.

coordination sphere was further optimized to afford efficient luminescent LBT by replacing the hydrogen-bonded $Asp_9$ with bidentate $Glu_9$. Thus, $Tb^{3+}$–LBT complexes with $q = 0$ and CN = 8 were obtained (Fig. 4 and entries 9–11 in Table 1).[64] Moreover, in order to sensitize $Tb^{3+}$ emission a Trp residue serving as the antenna was introduced in the 7th position of the EF-hand motif (see 6.1 Luminescent tags and probes).[66,74]

**2.2.3 Chimeric helix-loop-helix (HLH) motifs.** The group of Franklin designed chimeric helix-loop-helix (HLH) motifs by engineering an EF-hand loop in a transcription factor helix-turn-helix DNA-binding domain, which were investigated as artificial nucleases (6.4 Catalysis) and DNA-responsive MRI probes (6.2. MRI).[16,75–80] The NMR solution structure of the $La^{3+}$–P3 complex revealed a coordination site similar to natural EF-hand loops. However, $Asp_9$ points away from the metal ion, leaving an open face that could easily accommodate two inner-sphere water molecules, which is in agreement with luminescence data recorded on the $Eu^{3+}$–P3 complex (entry 5 in Table 1).

**2.2.4 Lanmodulins.** In 2018, a novel family of EF-hand proteins, called lanmodulins (LanM), was discovered in $Ln^{3+}$-utilising bacteria.[81] Compared to previously known $Ca^{2+}$-binding proteins, LanM shows considerably higher selectivity for $Ln^{3+}$ *vs.* $Ca^{2+}$ (see 4. Selectivity). Recently, X-ray crystal structures of two LanM proteins from *M. extorquens* AM1 (*Mex*-LanM, Fig. 4) and *H. quercus* (*Hans*-LanM) were reported, thus facilitating a thorough comparison of their binding sites with those found in $Ca^{2+}$-binding proteins.[62] Of note, while CaM is folded in the apo-state and undergoes a conformational change upon metal-binding,[82] LanM is disordered and only upon binding four $Ln^{3+}$ ions, it forms a compact three-helix bundle fold that is unique among EF-hand proteins.[83] The $Ln^{3+}$ coordination spheres in LanMs feature higher CN (9 or 10) than $Ln^{3+}$ bound to $Ca^{2+}$-binding loops and LBT (CN = 7–8). In *Mex*-LanM, this stems from a monodentate-to-bidentate switch of $Asp_5$ (entries 1 *vs.* 6 in Table 1) and the presence of an additional water molecule.

Compared to *Mex*-LanM, no $Ln^{3+}$-bound water molecules are found in *Hans*-LanM binding sites due to replacement of $Asp_9$ with $Glu_9$ (entries 6–8 *vs.* 12–14 in Table 1), as already observed for LBT (entries 9 *vs.* 2 in Table 1). In *Hans*-LanM, the increased CN relative to $Ca^{2+}$-binding proteins results from the presence of a bidentate $Asp_3$ (entries 12 and 14) rather than monodentate $Asn_3$ (entries 1, 9–11) and a monodentate-to-bidentate shift of $Asp_5$ (entries 12 and 14 *vs.* 9–11 in Table 1). Furthermore, a decrease in the CN is observed when $La^{3+}$ is replaced by the smaller $Dy^{3+}$ stemming from a bidentate-to-monodentate switch of $Glu_9$ (entries 12 *vs.* 14 in Table 1).[62]

### 2.3 Short peptides

**2.3.1 Cyclic decapeptide (PA).** Following the works of Franklin and Imperiali on the design of $Ln^{3+}$-binding peptides inspired by the $Ca^{2+}$ coordination site in EF-hand motifs, Delangle and coworkers explored a different approach to improve the affinity of small peptides for $Ln^{3+}$. First, they selected a regioselectively addressable functionalized templates (RAFT) type of cyclic decapeptide (noted PA) in order to design a $Ln^{3+}$-binding site made of two Asp and two Glu residues (Fig. 5).[41] The

**Table 1** Structural features of $Ln^{3+}$ bound to EF-hand motifs. EF-loop patterns report the amino acids found in positions 1, 3, 5, 9 and 12 (note that the residue in position 7, not shown, always binds $Ln^{3+}$ *via* a backbone carbonyl donor) and their binding mode (m = monodentate, b = bidentate, hb = hydrogen-bonded to a coordinating water molecule, nb = non-bound). Coordination (CN) and hydration (*q*) numbers are derived from X-ray crystal structures (for which PDB code is reported) or luminescence data obtained in solution using $Tb^{3+}$ or $Eu^{3+}$

| Entry | EF-loop pattern | | | | | *q* | CN | $Ln^{3+}$ | EF-hand motifs | Peptide/Protein | PDB code | Ref. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 3 | 5 | 9 | 12 | | | | | | | |
| 1 | D (m) | N (m) | D (m) | D (hb) | E (b) | 1 | 7 | Tb | **$Ca^{2+}$-protein** | TnC | 1NCZ | 58 |
| 2 | D | N | D | D | E | 1 | n/a | Tb | **LBT** | m-sSE3 | | 60 |
| 3 | D (m) | D (nb) | D (b) | D (hb) | E (b) | 1 | 7 | Gd | | xq-dSE3-ubiquitin | 3VDZ | 60 |
| 4 | D | D | D | D | E | 1 | n/a | Tb | | q-dSE3-ubiquitin | | 60 |
| 5 | D | D | D | D | E | 2 | n/a | Eu | **HLH** | P3 (HLH) | | 61 |
| 6 | D (m) | D (m) | D (b) | D (hb) | E (b) | 2 | 9 | Nd | **LanM** | *Mex*-LanM (EF1-3) | 8FNS | 62 |
| 7 | D | D | D | D | E | 2 | n/a | Tb | | *Mex*-LanM T41W | | 63 |
| 8 | N (m) | D (b) | D (b) | D (hb) | E (b) | 2 | 10 | Nd | | *Mex*-LanM (EF4) | 8FNS | 62 |
| 9 | D (m) | N (m) | D (m) | E (b) | E (b) | 0 | 8 | Tb | **LBT** | LBT12 | 1TJB | 64 |
| 10 | D | N | D | E | E | 0 | n/a | Tb | | sSE3 | | 60 |
| 11 | D (m) | N (m) | D (m) | E (b) | E (b) | 0 | 8 | Tb | | ILB1β-S1 | 3LTQ | 65 |
| 12 | N (m) | D (b) | D (b) | E (b) | E (b) | 0 | 10 | La | **LanM** | *Hans*-LanM (EF1-3) | 8DQ2 | 62 |
| 13 | N | D | D | E | E | 0 | n/a | Eu | | | | 62 |
| 14 | N (m) | D (b) | D (b) | E (m) | E (b) | 0 | 9 | Dy | | | 8FNR | 62 |

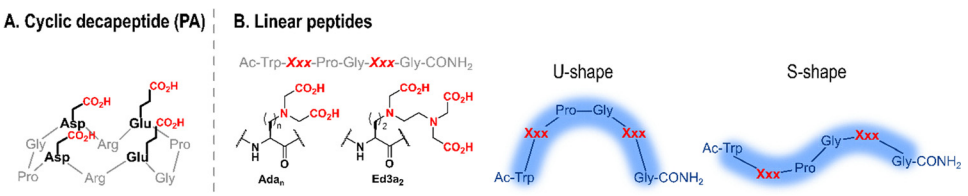


**Fig. 5** Schematic structures of short cyclic (A) and linear (B) Ln-binding peptides. In (B) the structures of unnatural amino acids Xxx = $Ada_n$ and $Ed3a_2$, and shapes adopted by linear peptides are shown. Adapted from ref. 8.

apo-peptide adopted an antiparallel β-sheet structure with a predisposed Ln-binding site at pH 3 with the side chains of the four coordinating residues on the same face of the cycle. Upon increasing the pH to 7, coulombic repulsions due to the deprotonation of the carboxylates led to an apparently disordered apo-peptide. Nevertheless, the addition of $Ln^{3+}$ ($La^{3+}$, $Tb^{3+}$) compensated for this repulsion. The authors showed the presence of two water molecules in $Tb^{3+}$ first coordination sphere ($q$ = 2). By comparison with earlier work by Franklin[16,61,76] and Imperiali,[64] they proposed that the $Ln^{3+}$ coordination sphere includes the four carboxylates (monodentate or bidentate) and two water molecules.

**2.3.2 Linear peptides.** Next, Delangle and coworkers investigated the use of unnatural amino acids in short linear hexapeptides for improving the affinity for $Ln^{3+}$ (see 3. Thermodynamic stability). All the peptides have the sequence Ac–W–*X*–P–G–*X*–G–$CONH_2$, where *X* is either $Ada_n$ with an aminodiacetate side chain ($n$ = 1–3), or $Ed3a_2$ with an ethylenediamine triacetate side chain (Fig. 5), and Pro-Gly is a β-turn-inducer motif.[84–87] These peptides are disordered in the apo-form and fold upon metal binding. The speciation, coordination and solution structure of Ln–peptides were analysed in details and the authors showed that it was impacted by (i) side chain length; (ii) position of the Ln-binding residues in the sequence; and (iii) the type of residues ($Ada_n$ or $Ed3a_2$, Table 2).

All the peptides formed 1 : 1 ligand : metal complexes below 1 equivalent of $Ln^{3+}$ relative to the peptide, however, in the presence of more equivalents of $Ln^{3+}$ the formation of multimetallic species was also observed (Table 2).

Changes in $Ada_n$ side chain length impacted its coordination mode to $Ln^{3+}$ (Table 2). $Ada_2$ and $Ada_3$ bind $Ln^{3+}$ in a tridentate manner,[85] whereas $Ada_1$ is tri- or tetra-dentate depending on its position in the sequence (fifth or second, respectively).[84] The additional chelating group was proposed to be the backbone carbonyl of the $Ada_1$ main chain, which can form an additional chelate ring. This change of coordination mode impacted the number of water molecules in the first coordination sphere of $Tb^{3+}$, which dropped from 3 ($P^{22}$, $P^{21}$ and $P^{33}$) to 1 and 0 in peptides where $Ada_1$ is tetradentate ($P^{11}$ and $P^{12}$, respectively). It also impacted the structure of $Ln^{3+}$–peptide complexes in solution: in the presence of $Ln^{3+}$, $P^{22}$ and $P^{21}$ adopted a U-shape conformation, likely further stabilised by the formation of a type-II β-turn,[85] while coordination of $Ada_1$ backbone carbonyl to $Ln^{3+}$ did not allow for the formation of a β-turn and resulted instead in an S-shape conformation of $P^{11}$ and $P^{12}$ (Fig. 5).[84]

Building upon this knowledge, Delangle and co-workers combined the pentadentate $Ed3a_2$ unnatural amino acid with $Ada_2$ either by placing $Ed3a_2$ in the second ($P^{HD2}$) or fifth ($P^{HD5}$) position in the sequence (Table 2).[87] Again, the speciation of the complexes depended on the position of the $Ln^{3+}$-binding amino acids in the sequence. For both peptides, the pentadentate $Ed3a_2$ excluded water molecules from the $Tb^{3+}$ coordination sphere ($q$ = 0).

More recently, short $Ln^{3+}$-binding peptides rich in Asp or Glu residues were reported. Veliscek-Carolan and co-workers studied peptides made of two to four Glu residues, either with an L- or D-stereochemistry, and coupled to a naphthalene antenna.[88] Simoni and co-workers studied the structure–affinity relationships of three Asp-rich pentapeptides (ADPDA, DPDPD, DGDGD) with actinide ions and $Eu^{3+}$ and they predicted that the $Eu^{3+}$ coordination sphere contained two carboxylates, one backbone carbonyl, and 4–5 water molecules.[89]

### 2.4 Peptide scaffolds of intermediate size

In the late 1980s, chemists began to dissect the complexity of proteins' fold and function, with the ultimate goal of being able to predict the primary sequence–structure–activity relationship.[90] This led to the development of *de novo* proteins, which refers to proteins designed from first principles. The design of α-helices was soon mastered, and the study of α-helical coiled coil made of the assembly of several α-helices has helped unravel design principles to obtain well-folded coiled coil assemblies. Since then, *de novo* proteins have been implemented with metal-binding site(s) and functions such as catalytic properties or electron transfer capacities.[91] Regarding Ln-binding peptides, a large amount of work has been dedicated to the design of Ln-binding coiled coils (2.4.1 Coiled coils),[92–102] and a seminal work described the design of Ln-fingers[103] (2.4.2 Ln fingers).

**2.4.1 Coiled coils.** The sequences of α-helices in coiled coils are constituted by a repetition of a heptad of amino acids (*abcdefg*) (Fig. 6 and 8).[93] Their assembly is driven by hydrophobic interactions of core amino acids in positions *a* and *d*. Residues in positions *e* and *g* are used to provide stabilizing electrostatic interactions, which also control the number of helices that assemble as well as their orientations (parallel or anti-parallel). Last, amino acids in position *f* are chosen for their propensity to induce α-helix folding, whereas those in positions *b* and *c* are used to improve solubility.

Three strategies were employed to insert an $Ln^{3+}$-binding site in such scaffolds: (i) in two-stranded coiled coils (2SCC,

Table 2 Linear peptides speciation, coordination mode, structure and hydration number. The peptides have the sequence Ac–W–*X*–P–G–*X*–G–$CONH_2$, where *X* is either $Ada_n$ with an aminodiacetate side chain ($n$ = 1–3) or $Ed3a_2$

| Peptide | Speciation | $Ln^{3+}$-binding AAs | Coordination mode | Structure | $q$ | Ref. |
|---|---|---|---|---|---|---|
| $P^{11}$ | 1 : 1 and multimetallic species | $Ada_1$ | Tri- and tetra-dentate | S-shape | 1 | 84 |
| $P^{22}$ | 1 : 1 | $Ada_2$ | Tridentate | U-shape | 3 | 85 |
| $P^{33}$ | 1 : 1 and multimetallic species | $Ada_3$ | Tridentate | n.d. | 3 | 85 |
| $P^{12}$ | 1 : 1 | $Ada_1$ and $Ada_2$ | $Ada_1$: tetradentate | S-shape | 0 | 86 |
| $P^{21}$ | 1 : 1 and multimetallic species | $Ada_2$ and $Ada_1$ | $Ada_1$: tridentate | U-shape | 3 | 86 |
| $P^{HD2}$ | 1 : 1 | $Ada_2$ and $Ed3a_2$ | $Ed3a_2$: pentadentate | n.d. | 0 | 87 |
| $P^{HD5}$ | 1 : 1 and multimetallic species | $Ed3a_2$ and $Ada_2$ | $Ed3a_2$: pentadentate | n.d. | 0 | 87 |

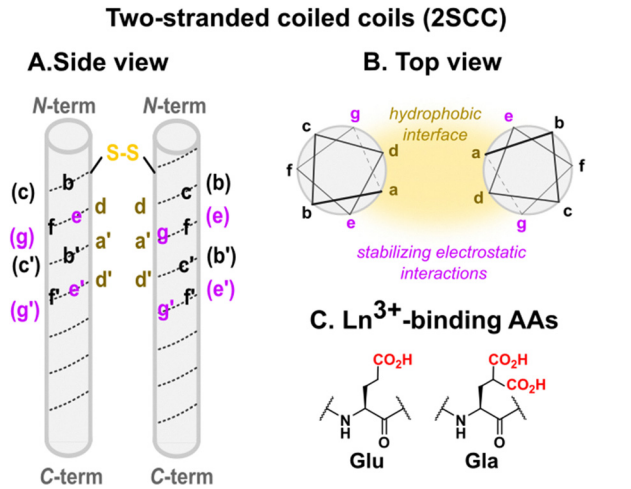


Fig. 6 Two-stranded coiled coils scaffolds designed by Hodges and coworkers: (A) side view, (B) top view, (C) Ln-binding sites are placed at the helical interface using Glu or Gla AAs in *e* and *g* positions. In (B) stabilising non-covalent interactions (hydrophobic, electrostatic) responsible for the folding and assembly of coiled coils are highlighted.

Fig. 6 and Table 3), positions *e* and *g* were used to introduce chelating moieties and place the $Ln^{3+}$-binding site at the helical interface; in three-stranded coiled coils (3SCC, Fig. 8), the binding site was (ii) either placed at the N-terminal (Table 4) or (iii) buried in the hydrophobic core by replacing core amino acids in position *a* and *d* by $Ln^{3+}$-binding residues (Table 5).

*Two-stranded coiled coils (2SCC).* Pioneering works on $Ln^{3+}$-binding 2SCCs were done by Hodges and coworkers.[94,95] They used as a starting point a *de novo* α-helical peptide made of the repetition of 5 heptads (35 amino acids, Table 3). The formation of a disulphide bridge ensured the obtention of a parallel 2SCC (Fig. 6). The native sequence was modified to include two ($E_2$) or three ($E_3$) Glu residues in *e* or *g* positions to form two $Ln^{3+}$-binding sites per coiled coil.[94] The stability of the apo-peptides in their oxidized form was quantitatively investigated by urea denaturation followed by circular dichroism (Fig. 7). The destabilisation induced by the mutation of Gln to Glu was the greatest for the $E_2(15,20)_{ox}$ construct, in which Glu in positions 15 and 20 generates the highest interchain repulsion. The introduction of a third Glu residue in the sequence ($E_3(13,15,20)$ and $E_3(15,20,22)$) further destabilised the coiled coil.

A stabilising effect of $La^{3+}$ on peptide unfolding was noted for $E_2(15,20)$ and $E_3$ constructs, in which the Glu residues are well positioned to bind $La^{3+}$ at the interhelical interface, the resulting interhelical bridges being responsible for the stability enhancement. The higher stabilisation measured for the $E_3$ sites may also indicate that the third Glu could participate in binding.

In addition, the authors evaluated the stability of the reduced peptides. These were less stable than their oxidised counterparts, highlighting the importance of the disulphide bridge for the stability of the scaffold. The addition of $LaCl_3$ resulted in an important stabilisation, which was attributed not only to $La^{3+}$-binding but also to a $La^{3+}$-driven dimerization. Hence, the most stable $Ln^{3+}$-binding constructs were achieved thanks to the presence of a disulphide bridge to form exclusively a 2SCC, and to $Ln^{3+}$-binding residues at the helical interface.

In a following work, Hodges and co-workers turned to γ-carboxyglutamic acid (Gla, Fig. 6), with two carboxylate functions, in order to (i) enhance the $Ln^{3+}$-affinity of the coiled coil thanks to the higher denticity of Gla *vs.* Glu, and (ii) design a system that folds upon metal-binding as the higher electrostatic

Table 3 Two-stranded coiled coil sequences designed by Hodges and coworkers

| Name | Peptide sequence | Ref. |
|---|---|---|
| | g abcdefg abcdefg abcdefg abcdefg abcdef | |
| N | Ac-Q CGALQKQ VGALQKQ VGALQKQ VGALQKQ VGALQK-$CONH_2$ | 94 |
| $E_2$(13,22) | Ac-Q CGALQKQ VGALEKQ VGALQKE VGALQKQ VGALQK-$CONH_2$ | |
| $E_2$(20,22) | Ac-Q CGALQKQ VGALQKQ VGALEKE VGALQKQ VGALQK-$CONH_2$ | |
| $E_2$(15,20) | Ac-Q CGALQKQ VGALQKE VGALEKQ VGALQKQ VGALQK-$CONH_2$ | |
| $E_3$(13,15,20) | Ac-Q CGALQKQ VGALEKE VGALEKQ VGALQKQ VGALQK-$CONH_2$ | |
| $E_3$(15,20,22) | Ac-Q CGALQKQ VGALQKE VGALEKE VGALQKQ VGALQK-$CONH_2$ | |
| $Gla_2$(15,20)N | Ac-Q CGALQKQ VGALQKX VGALXKQ NGALQKQ VGALQK-$CONH_2$<br>X = Gla | 95 |

Table 4 Three-stranded coiled coil sequences designed by Kashiwada and Ashkenazy

| Name | Peptide sequence | Ref. |
|---|---|---|
| | *efg abcdefg abcdefg abcdefg abcdefg* | |
| Pep1 | YGG EEK IAAIEKK IAAIEEK IAAIEKK IAAIEEK GGY | 96 |
| Pep3 | YGG EEK IAAIEKK IAAAEEK XAAIEKK IAAIEEK GGY<br>X = Gla | |
| | abcdefg abcdefg abcdefg abcdefg | |
| HPO-1 | HPO-β-A IKQIEDK IEEILSK IYSIENE IARIKKL | 97 |

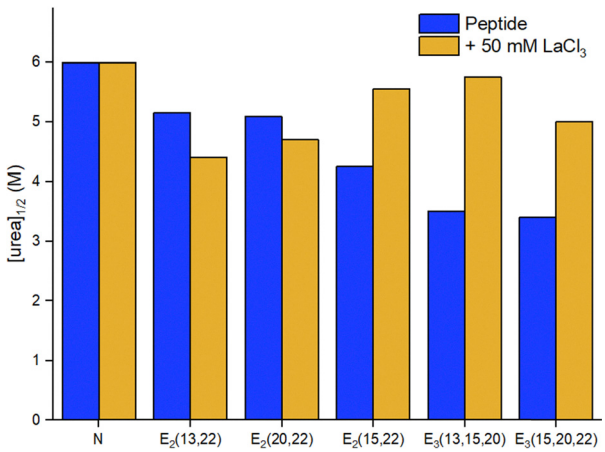


Fig. 7 Urea denaturation of 2SCC at pH 7 followed by circular dichroism.[94] Peptide sequences are given in Table 3. Experiments performed at 20 °C in 50 mM TRIS, 100 mM KCl, [pep] = 70–140 μM, without or with 50 mM $LaCl_3$. $[urea]_{1/2}$ corresponds to the concentration at which 50% of the peptide is unfolded.

repulsion induced by Gla was expected to lead an unfolded apo-peptide.[95] In order to further destabilise the apo-peptides, they mutated a Val residue (V) in position 23 by an Asn (N) (Table 3). The resulting peptide in its oxidised form, $Gla_2$(15,20)N, was fully unfolded at 20 °C pH 7. The peptide folded upon the addition of 0.5 mM $LaCl_3$, the amount of coiled coil formation increasing from ~3 to 100% to match the one of the native Asn.

*Three-stranded coiled coils (3SCC).* A few years later, Kashiwada and coworkers used the same Gla residue for $Ln^{3+}$-binding, but inserted it in a three-stranded coiled coil (3SCC) scaffold.[96] In their design, the Gla residue was placed in an *a* layer of Pep1 (Table 4) so that the $Ln^{3+}$-binding site would be buried in the hydrophobic core rather than placed at the helical interface. Due to the insertion of the Gla residue in the sequences, the resulting peptide Pep3 was unfolded at 20 °C, pH 7. Circular dichroism experiments showed that the addition of 1 equivalent of $Ln^{3+}$ ($Ce^{3+}$, $Eu^{3+}$, $Tb^{3+}$) per trimer drove the assembly and folding of the coiled coil.

Ashkenazy and coworkers investigated an alternative strategy, by placing the $Ln^{3+}$-binding site at the N-terminal of the 3SCC (Fig. 8 and Table 4), which has the advantage of not perturbing the folding and assembly of 3SCC.[97] They coupled a hydroxyphenol oxazoline moiety (HPO) to the peptide through a triazole linker. HPO acted both as a chelating moiety and as a sensitizing antenna.

The group of Peacock contributed to the design of $Ln^{3+}$-binding 3SCCs and provided a detailed understanding of the influence of $Ln^{3+}$-binding site position, $Ln^{3+}$-binding residues, second sphere effects, and scaffold size on $Ln^{3+}$-affinity, selectivity, and hydration number (Table 5).[98–102,104] They also reported the design of a heterobimetallic 3SCC bearing an $Ln^{3+}$-binding site and a Hg-binding site.[92]

The first peptide designed by the Peacock group, MB1-2, was a parallel 3SCC made of the repetition of 5 heptads and with an $(Asn)_3(Asp)_3$ metal binding site with Asn residues in a *d* layer, and Asp residues in the following *a* layer (Table 5).[98] The peptide also contains Trp residues to sensitize $Tb^{3+}$ emission (see 6.1 Luminescent tags and probes).

The positioning of the $Ln^{3+}$-binding site within the hydrophobic core of the 3SCC destabilised the peptide, which was unfolded in the apo-state but assembled and folded upon $Ln^{3+}$-binding (Fig. 9). The hydrophobic environment and the adequate coordination sphere ($(Asn)_3(Asp)_3$) provided a suitable CN to the $Ln^{3+}$ ion, preventing water from binding directly to the metal ion ($q$ = 0).

Building on this work, the group of Peacock investigated how the position of the binding site in the 3SCC scaffold and the type of binding site ($(Asn)_3(Asp)_3$ *vs.* $(Asp)_3$) would impact $Ln^{3+}$ coordination and hydration number.[99] With this aim, they extended the MB1 series containing an $(Asn)_3(Asp)_3$ binding site with peptides MB1-1, MB1-3 and MB1-4, and designed the CS1 series containing an $(Asp)_3$ binding site with peptides CS1-1, CS1-2 and CS1-4 (Table 5). In the MB1 series, the position of the metal-binding site impacted differently the stability of the apo-peptides (Fig. 9). Whereas positioning the metal-binding

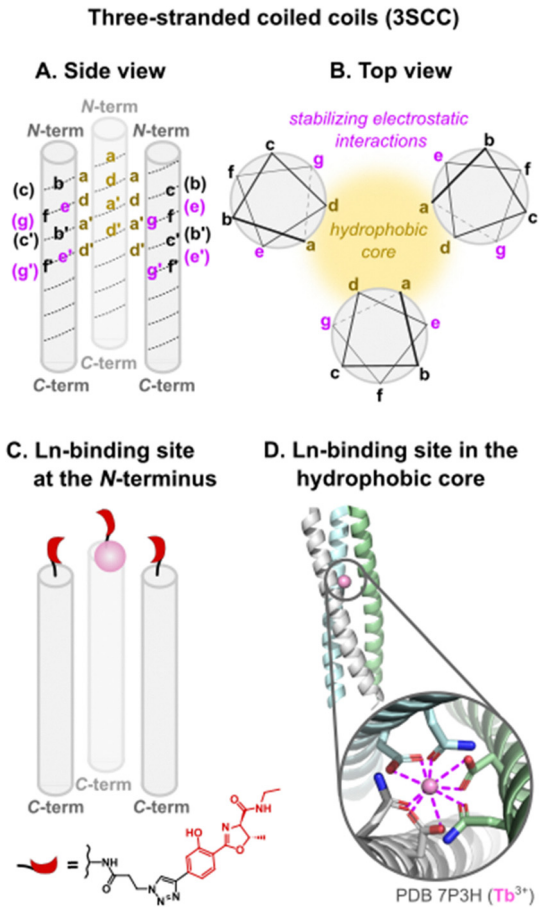


Fig. 8 Three-stranded coiled coils scaffolds: (A) side view, (B) top view, (C) Ln-binding site placed at the N-terminal with a hydroxyphenol oxazoline (HPO) ligand, (D) Ln-binding site in the hydrophobic core (*a* and *d* positions). In (B) stabilising non-covalent interactions responsible for the folding and assembly of coiled coils are highlighted.

Table 5 Three-stranded coiled coil sequences designed by Peacock and coworkers

| Name | Peptide sequence | $q$ | Ref. |
|---|---|---|---|
| | abcdefg abcdefg abcdefg abcdefg abcdefg | | |
| MB1-1 | Ac-G IAANEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 3 | 99,100 |
| MB1-2 | Ac-G IAAIEQK IAANEWK DAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 0 | 98,99 |
| MB1-3 | Ac-G IAAIEQK IAAIEQK IAANEWK DAAIEQK IAAIEQK G-$CONH_2$ | 0 | 99 |
| MB1-4 | Ac-G IAAIEQK IAAIEQK IAAIEQK IAANEWK DAAIEQK G-$CONH_2$ | 2 | 99 |
| MB1-1(2*A*) | Ac-G *A*AANEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 3 | 100 |
| MB1-1(2*F*) | Ac-G *F*AANEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 2 | |
| MB1-1(2*W*) | Ac-G *W*AANEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 0 | |
| MB1-1(2*Y*) | Ac-G *Y*AANEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 1 | |
| MB1-4(37W) | Ac-G IAAIEQK IAAIEQK IAAIEQK IAANE*W*K DAAIEQK *W*G-$CONH_2$ | 0 | |
| MB1-1S | Ac-G IAANEWK DAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | nd | 102 |
| MB1-1L | Ac-G IAANEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 4 | |
| CS1-1 | Ac-G IAAIEWK DAAIEQK IAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | 4 | 99 |
| CS1-2 | Ac-G IAAIEQK IAAIEWK DAAIEQK IAAIEQK IAAIEQK G-$CONH_2$ | nd | |
| CS1-4 | Ac-G IAAIEQK IAAIEQK IAAIEQK IAAIEWK DAAIEQK G-$CONH_2$ | nd | |
| HC02 | Ac-E WEAIEKK IAANESK DQAIEKK IQAIEKK IEAIEHG-$CONH_2$ | 0 | 101 |

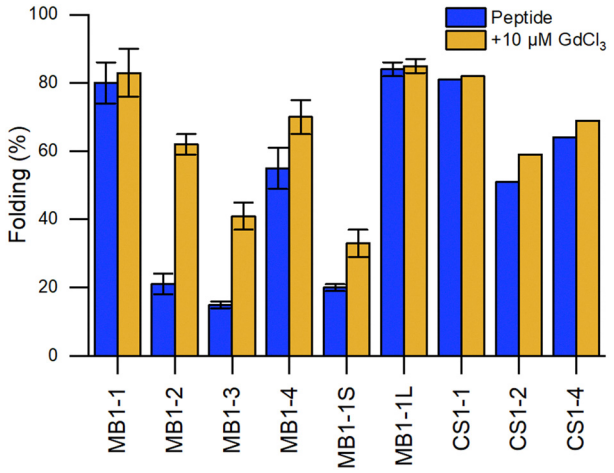


Fig. 9 Comparison of the percentage of folding of 3SCC.[99,102] Peptide sequences are given in Table 5. Experiments were performed in 5 mM HEPES, pH 7, 293 K, [pep] = 30 μM, without or with 10 μM $GdCl_3$. The percentage folding was calculated based on: $\%_{folded} = \frac{[\theta]_{222nm} - [\theta]_{coil}}{[\theta]_{max} - [\theta]_{coil}} \times 100$, with $[\theta]_{222nm}$ the molar ellipticity at 222 nm, $[\theta]_{coil}$ the molar ellipticity for a random coil, and $[\theta]_{max} = -42\,500 \times \left(1 - \left(\frac{3}{n}\right)\right)$ the theoretical maximum molar ellipticity, where $n$ is the number of residues in the sequence.[105]

site in the middle of the scaffold destabilized the apo-peptides (21% and 15% folded for MB1-2 and MB1-3, respectively), positioning at the C-terminal was better tolerated (55% folded), and when placed close to the N-terminal the apo-peptide seemed almost unaffected (80% folded). For the CS1 series, in which there is no top $(Asn)_3$ layer, the apo-peptides were always more stable than their MB1 analogues (Fig. 9). As was previously observed for MB1-2,[98] $Gd^{3+}$ stabilised the 3SCC assembly and led to the formation of 1 : 3 $Ln^{3+}$ : peptide complex. However, the authors noted that for CS1-2 and CS1-4, there was only a small improvement of folding upon $Tb^{3+}$-binding compared to their MB analogues (Fig. 9), and only a small increase in $Tb^{3+}$-emission which suggested that these peptides may bind $Tb^{3+}$ through non-specific interactions.

Depending on the position of the binding site, the number of $Tb^{3+}$-bound water molecules varied from 4 (CS1-1) and 3 (MB1-1) for the more solvent-exposed sites at the N-terminal, to 2 (MB1-4) for the site at the C-terminal, to 0 (MB1-2 and MB1-3) for sites more buried into the scaffold (Table 5). These results demonstrated that it is possible to control the number of $Ln^{3+}$-bound water molecules by controlling the position of the metal binding site in the 3SCC. The more solvent-exposed sites at the N- and C-termini had the higher hydration number, which suggested a change in Asp coordination mode (monodentate *vs.* bidentate) or that not all Asp and Asn were involved in $Ln^{3+}$-coordination. On the other hand, the sites that were more buried within the hydrophobic core of the 3SCC were shielded from the solvent. This is consistent with the $(Asn)_3(Asp)_3$ site fulfilling the preference of $Ln^{3+}$ for a high coordination number. The recent publication of the crystal structure of HC02 (PDB 7P3H), a 3SCC similar to MB1-2, confirmed that $Tb^{3+}$ is nona-coordinated by three Asn (monodentate) and three Asp (bidentate) as shown in Fig. 8.[101]

In later works, Peacock and coworkers investigated both the effect of the peptide length on the overall stability of the assembly[102] and of second-sphere effects on the hydration number (Table 5).[100] Whereas the shorter MB1-1S was too destabilised to bind $GdCl_3$, there was no marked difference in peptide folding or hydration number between MB1-1 and the longer MB1-1L (Fig. 9), and only a qualitative enhancement of stability for MB1-1L compared to MB1-1. One explanation proposed was that the binding site is too far (4–5 nm) from the sixth heptad to detect any positive effect on folding, $Ln^{3+}$ affinity and hydration number.

In order to investigate 2nd sphere effects on the hydration number, core Ile (I) residues placed in the *a* layer above the metal-binding site of MB1-1 were mutated to Ala (A), Phe (F), Tyr (Y) or Trp (W) (MB1-1(2*X*) series, Table 5). Tuning of the steric hindrance provided control of the hydration number of $Tb^{3+}$ that varied from $q$ = 3 (MB1-1, MB1-1(2A)), to 2 (MB1-1(2F)), to 1 (MB1-1(2Y)), to 0 (MB1-1(2W)). This strategy was also successfully applied at the C-terminal for MB1-4 site, which was mutated to MB1-4(37W) in order to shield $Tb^{3+}$ from the solvent, which decreased the hydration number from 2 to 0.

The results obtained by Peacock and coworkers highlight that with the same set of coordinating AAs, the number of $Ln^{3+}$-bound water molecules can be controlled either by controlling the position of the $Ln^{3+}$-binding site within the scaffold or by tuning second sphere residues thanks to steric effects. The replacement of core amino acids by $Ln^{3+}$-binding residues resulted in the destabilisation of peptide folding and assembly that depended on the position of the site within the scaffold and could be in part compensated for by $Ln^{3+}$-binding.

**2.4.2 Ln fingers.** A seminal work by Zondlo and coworkers described the rational design of lanthanide fingers (LF, Fig. 10 and Table 6).[103] Their goal was to obtain a peptide that adopts the same fold as the naturally occurring zinc fingers, whose structure is constituted of an α-helix placed above a β-hairpin.

The $(His)_2(Cys)_2$ $Zn^{2+}$-binding site of Zn-fingers was replaced by an $(Asp)_2(Glu)_2$ binding site, more suited for $Ln^{3+}$-binding. The LF scaffold was redesigned to favour the formation of secondary structures, leading to LF1. Modifications included notably the use of α-helix or β-sheet inducer amino acids and of

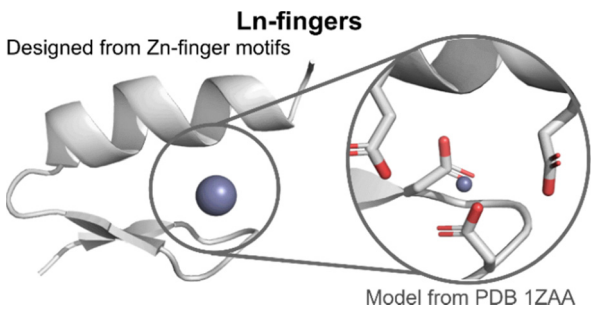


Fig. 10 Lanthanide fingers. A model of the hypothetic structure and binding site of Ln-fingers was generated from a closely related zinc finger structure (PDB 1ZAA) using Pymol. Zn is represented as a grey sphere.

Table 6 Ln fingers sequences designed by Zondlo and coworkers. $^{D}$P = D-proline

| Name | Peptide sequence |
|---|---|
| | β-hairpin α-helix |
| LF1 | AYTD$^{D}$PSDTKTFSQSSALAKELKAEAG |
| LF2 | AYTD$^{D}$PSDTKTFS**TPEQ**LAKELKAEAG |
| LF3 | AYTD$^{D}$PSDTKTFSTPEQLAKELL**QEHGI** |
| LF4 | AYTD$^{D}$PSD__TFSTPEQLAKELLQEHGI |
| LF5 | AYTD **TI**D__TFSTPEQLAKELLQEHGI |
| LF6 | AYTD TID__T**W**STPEQLAKELLQEHGI |
| LF7 | AYTD$^{D}$PS**N**__TFSTPEQLAKELLQEHGI |

a type II′ β-turn ($^{D}$Pro-Ser) to favour the formation of the β-hairpin. The addition of an N-cap (LF2) and a C-cap (LF3), which are motifs favouring the formation of α-helices,[106–108] improved the folding. The extent of stabilisation of peptides fold following these modifications was qualitatively assessed by comparing CD spectra obtained for apo-peptides. The influence of the size of the binding site on $Ln^{3+}$-binding was investigated by opening up the site thanks to the deletion of two AAs in the sequence (LF4), as well as the type of $Ln^{3+}$-coordinating residues (LF7). Genetically encodable LFs were designed by replacing the type II′ β-turn ($^{D}$Pro-Ser) with Thr-Ile (LF5 and LF6).

A detailed NMR analysis of LF4 without and with $Lu^{3+}$ showed rearrangements throughout the peptide upon $Lu^{3+}$-binding, which suggested that, as for Zn-fingers, Ln-fingers fold upon metal binding. For all the sequences, $Ln^{3+}$ binding improved the folding of the peptides and the data obtained were consistent with a 1:1 peptide:$Ln^{3+}$ ratio. Hydration numbers were not determined, but $H_2O$ molecules are likely to participate in $Ln^{3+}$ coordination sphere in addition to the $Ln^{3+}$-binding AAs.

### 2.5 Other Ln-binding protein scaffolds

#### 2.5.1 Naturally occurring proteins

*Ln-ADH.* The discovery of Ln-utilizing bacteria has sparked a great interest in the identification of the lanthanome.[9–12] The first Ln-binding protein discovered in the 2010s was an enzyme: a pyrroloquinoline quinone (PQQ)-dependent methanol dehydrogenases (MDH) encoded by the gene *xoxF*, whose active site also contains an $Ln^{3+}$ taking part in the catalysis.[4–7] A characteristic of this enzyme is that it is directly involved in the metabolism of the organisms that produce it, mostly methylotrophic bacteria, which are hence the sole organisms known so far for which $Ln^{3+}$ can be essential elements.[9] Following this discovery, several Ln-MDH were isolated and their crystal structures were reported with either $La^{3+}$,[109,110] $Ce^{3+}$,[7] $Nd^{3+}$,[111] or $Eu^{3+}$ [112] in the active site (Fig. 11). In addition, the family of Ln-dependent enzymes was extended to ethanol dehydrogenases (EDH) following the identification of ExaF[113] and PedH[114] in methylotrophs and non-methylotrophs.[115] The crystal structure obtained for a $Pr^{3+}$-EDH showed that the active site of this enzyme shared a lot of similarities with the one of Ln-MDH, including conserved residues involved in $Ln^{3+}$- and PQQ-binding.[116] Due to these similarities, Ln-MDH and Ln-EDH are often grouped under the term Ln-alcohol dehydrogenases (Ln-ADH).

The active site of Ln-ADH is closely related to other PQQ-dependent alcohol dehydrogenases, whose active site contains a $Ca^{2+}$ ion, but which are encoded by different genes (*mxaF* for Ca-MDH, *exaA* and *pedE* for Ca-EDH).[9] The main difference between the active sites of Ca- and Ln-ADH is the presence of an additional Asp residue in the $Ln^{3+}$ coordination sphere. Crystal structures showed that in Ln-ADH, the $Ln^{3+}$ ion is nona-coordinated by two Asp (one monodentate, one bidentate), a Glu (bidentate), an Asn (monodentate), and the PQQ cofactor (tridentate), as shown in Fig. 11. For some structures, an additional ligand (PEG,[7] MeOH[111]) is also found coordinated to the $Ln^{3+}$, which suggests that there is a free site for substrate coordination to the metal ion, giving a total coordination number of 10.

It is not clear how the mature Ln-ADHs are formed, especially how they acquire their two cofactors, PQQ and $Ln^{3+}$, and how these impact the stability of the scaffold. Based on the obtention of crystal structures with only $Ln^{3+}$ (and not PQQ) in the active site,[110,111] it has been proposed that $Ln^{3+}$ could be loaded independently of PQQ. However, it is also possible that PQQ leaks out of the enzyme pocket during the crystallization

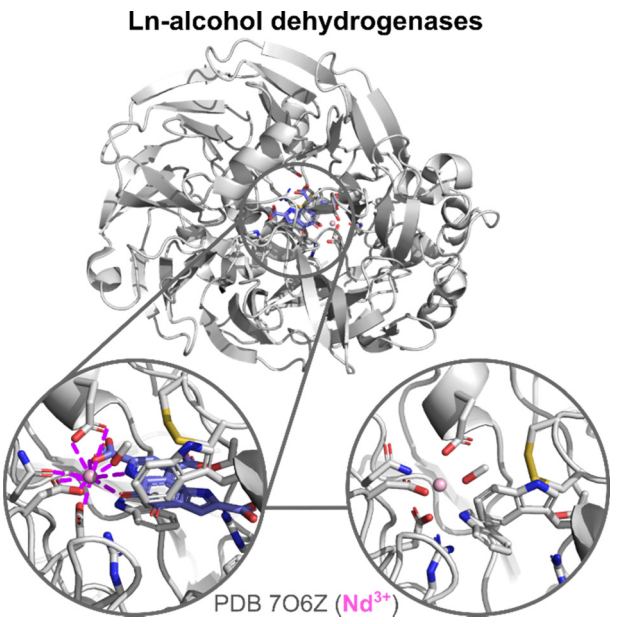


Fig. 11 Ln-ADH structure and active site. Zoom on the active site and Ln-coordination sphere of crystals obtained with the PQQ cofactor (blue) in the active site (left) and without it (right). In addition to Ln-binding AAs and PQQ, a methanol molecule is also bound to $Nd^{3+}$.

steps. The comparison of the holo-MDH (with Ln and PQQ) and Ln-only-MDH (no PQQ) shows nearly identical structures (Fig. 11).[110,111] Ln is coordinated by the same residues in both cases. Residues in interaction with PQQ show minimal rearrangement in the holo-MDH relative to the Ln-only MDH, which suggests that the MDH active site may be predisposed for PQQ binding.

On the other hand, ADH enzymes can also be obtained with PQQ-only in their active site. Indeed, several papers report on the purification of Ln-ADH with 0.4 to 1 equivalent of Ln per protein.[6,7,109–113,117,118] Moreover, Daumann and coworkers demonstrated that upon incubation of partially metallated MDH with $Ln^{3+}$ the activity of the enzyme could be improved, which suggests the *in situ* reconstitution of a functional enzyme in which PQQ was already present.[112,119] Martinez-Gomez and coworkers also observed the partial metallation of a Nd-MDH, without change in the PQQ content of the enzyme.[117] However, with their conditions, the reconstitution of a holo-enzyme by incubation with $Nd^{3+}$ was unsuccessful.

Working on a parent enzyme, the PQQ- and $Ca^{2+}$-dependent soluble glucose dehydrogenase (sGDH), Stines-Chaumeil, Limoges and coworkers demonstrated that reconstitution of the holo-enzyme could follow either a path where PQQ is added first to the active site or one where $Ca^{2+}$ is bound first.[120] Their results evidenced that the kinetic of reconstitution depended on the order of addition of the two cofactors. The reconstitution was fast when the protein was metallated with $Ca^{2+}$ before the binding of PQQ, and slow if PQQ binding happened first. Although little is known about the mechanism and kinetics of Ln-ADH reconstitution, the two reconstitution paths proposed for sGDH could also be considered for Ln-ADH.

The data on Ln-ADH stability are scarce. Some data show that Ln-MDHs are sensitive to temperature and denature at temperatures higher than 50 °C.[111,112] However, the exact value at which denaturation occurs may be influenced, among other factors, by the organism Ln-MDH is purified from, as the enzyme obtained from the extremophile *M. fumariolicum* SolV has an optimum temperature for catalytic activity at 60 °C.[7] In addition, the thermal stability of Ln-MDH could be Ln-sensitive, as suggested by Nakagawa and coworkers who compared the thermal stability of a La- and a Nd-MDH and found that the latter denatured at lower temperatures although it is not clear what could cause such a difference.[118]

*Lanpepsy (LanP).* Following the discovery of ADH, other $Ln^{3+}$-binding proteins were identified: lanmodulins (LanM, see above), lanpepsy (LanP), and landiscernin (LanD also referred to as LutD, see below).[121] LanP is composed of two PepSY domain.[122] Studies on deletion strains suggested that LanP function is not directly linked to $Ln^{3+}$ sensing or uptake. The 3D structure predicted by AlphaFold suggested that the folded protein forms a cavity with several negatively charged residues. Competitive titration experiments determined that LanP could bind up to 3–4 $Ln^{3+}$ ions, which was confirmed by ICP-MS, although a higher number of sites was determined by ITC titration (~6). The exact binding sites and coordination environments of $Ln^{3+}$ in LanP still need to be uncovered.

*Landiscernin (LanD).* A role for LanD in the acquisition of $Ln^{3+}$ ions was first suggested in 2019.[123,124] Following these preliminary works, the protein was isolated and crystallized by Cotruvo and coworkers.[125] It forms a three-helix bundle stabilised by a disulphide bridge and hydrophobic interactions at its core. Several negatively charged residues are found at its surface, close to the N-terminal. Crystal structures obtained with 0.5 equivalents of $Ln^{3+}$ (La, Ce, Eu, and Ho) per protein showed the formation of a $Ln^{3+}$-binding site at the interface between two proteins made of three Glu residues from each protein (Glu70, Glu73, Glu75), which except for this coordination site do not interact with each other (Fig. 12). Interestingly, as the $Ln^{3+}$ ionic radius decreases, the coordination number decreases from 9 ($La^{3+}$, $Ce^{3+}$) to 8 ($Eu^{3+}$, $Ho^{3+}$), due to the loss of a $Ln^{3+}$-bound water molecule for the latter (Fig. 12).

In solution, the protein was shown to be monomeric in the apo-state with a dissociation constant for the dimer of $K_{dimer}$ = 610 μM but was differently influenced by metal ions. Whereas $La^{3+}$, $Ce^{3+}$ and $Nd^{3+}$ seemed to favour the dimerization of the protein ($K_{dimer}$ = 117, 200 and 253 μM, respectively), this was not the case with $Eu^{3+}$ and $Ho^{3+}$ ($K_{dimer}$ = 1400 and 700 μM, respectively). Based on the crystal structures obtained, the authors proposed that as the ionic radius of the metal ion decreases, electrostatic repulsions due to the presence of

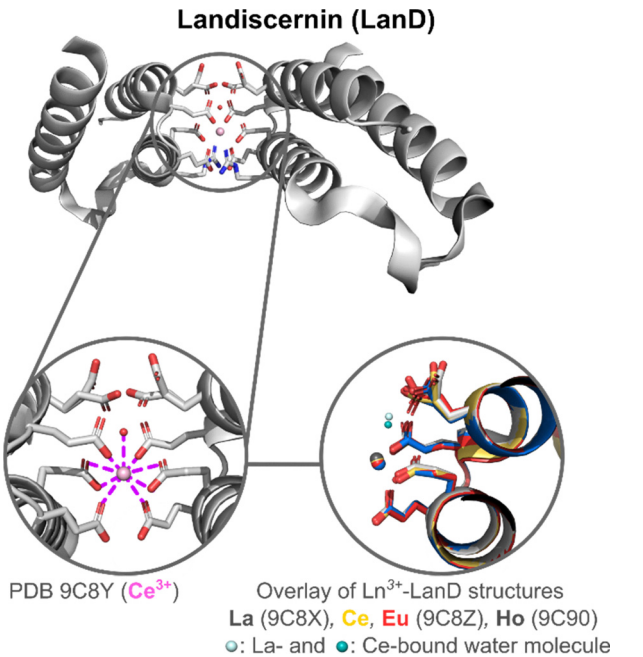


Fig. 12 Landiscernin (LanD) structure and active site. Zoom on the $Ln^{3+}$ coordination sphere (left). Overlay of $Ln^{3+}$–LanD structures ($Ln^{3+}$ = La, Ce, Eu, Ho) was generated with Pymol (right). A $Ln^{3+}$-bound water molecule is only present in the structure of $La^{3+}$–LanD and $Ce^{3+}$–LanD. Two glutamate residues adopt different orientations in the crystal structures and their rotamers are represented. They do not participate in $Ln^{3+}$-binding but are important for dimer formation.

negative charges on the surface of the proteins increase and disfavour the dimerization.

### 2.5.2 Proteins designed/engineered for Ln-binding

*PqqT.* The discovery of Ln-utilising bacteria has also inspired the field of artificial enzyme design, with two lead examples recently published. Olshansky and coworkers aimed at designing a Ln-ADH mimic, which could serve as a model to study the mechanism of the enzymes.[20] Analysis of the X-ray structure of PqqT, a PQQ-binding protein, showed that several residues near PQQ may be predisposed for Ln-binding, including $Glu_{111}$, $Asp_{141}$ and $Asn_{143}$ (Fig. 13). In addition, a lysine ($Lys_{142}$) interacts through electrostatic interactions with PQQ with its side chain positioned in the PQQ pocket where $Ln^{3+}$ binds. Mutation of $Lys_{142}$ ($K_{142}$) to an Asp afforded the mutant PqqT–$K_{142}$D that binds 1 PQQ (model structure in Fig. 13) and 1 $Ln^{3+}$ ion. In the engineered active site, PQQ was required for the binding of $La^{3+}$ and hence participated in its coordination sphere. In addition to $Asp_{142}$, other residues ($Glu_{111}$, $Asp_{141}$, $Asn_{143}$) may also participate in Ln-binding, which still needs to be confirmed. The mutant containing $La^{3+}$ and PQQ, noted $La^{3+}$–PQQ⊂PqqT–$K_{142}$D, was able to oxidize benzyl alcohol under aqueous conditions (see 6.4 Catalysis).

*TIM barrel.* The second example of an Ln artificial enzyme came from Zeymer and coworkers. Starting from a computationally designed *de novo* TIM barrel scaffold,[126] they aimed to design an enzyme for $Ce^{4+}$ photoredox catalysis (see 6.4 Catalysis).[19]

The initial scaffold was designed by combining a TIM domain made of eight parallel β-strands surrounded by eight α-helices, and a ferredoxin (FD) insert domain, either with a $(His)_2(Glu)_2$ or with a $(Glu)_4$ binding site.[126] The stability of the scaffolds with these two binding sites was assessed by thermal and chemical denaturation. The two scaffolds showed a high stability with temperature ($T_m$ > 95 °C), and unfolded in two steps with increasing concentration of guanidinium chloride, each event corresponding to the unfolding of one domain (TIM and FD). X-ray structure of the apo-protein scaffold showed a pre-disposed $(Glu)_4$ binding site. The hydration number of $Tb^{3+}$ was determined using a third mutant with a Trp for $Tb^{3+}$ sensitisation. In this mutant, $Tb^{3+}$ was bound by one water molecule. X-ray structure showed a CN = 9 for $Tb^{3+}$, with the four Glu acting as bidentate ligands (Fig. 14).

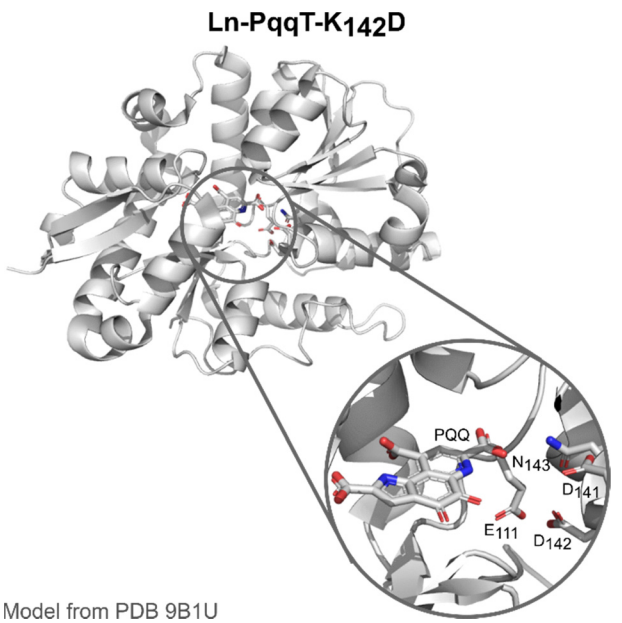


Fig. 13 Ln–PqqT–$K_{142}$D mutant bound to PQQ and its possible $Ln^{3+}$-binding site (zoom). Model structure was generated from the crystallographic structure of PQQ-bound PqqT (PDB 9B1U) using the mutagenesis tool in PyMOL.

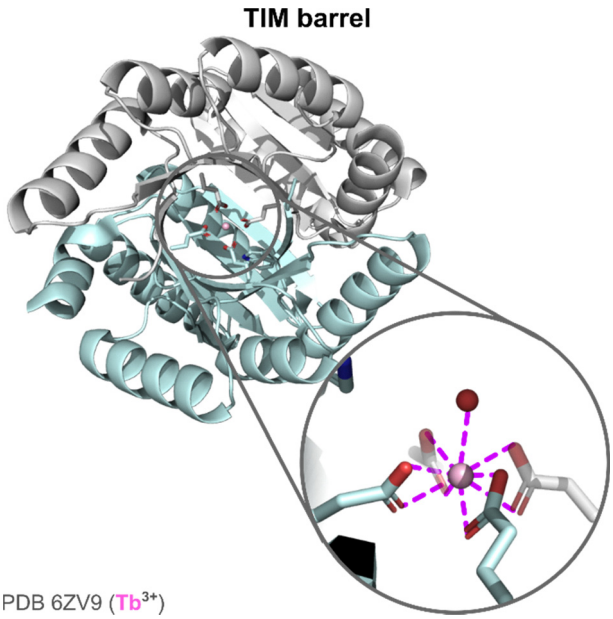


Fig. 14 TIM barrel structure and active site. Zoom on Ln-coordination sphere showing the presence of a $Ln^{3+}$-bound water molecule.

### 2.5.3 Ln-binding to other metalloproteins for specific applications.

Investigation of Ln-binding was also conducted with other metalloproteins (RTX, PTE-R18) to exploit these scaffolds for specific applications.

*RTX.* RTX corresponds to the block V of the repeats-in-toxins (RTX) domain of adenylate cyclase.[127–129] It is a $Ca^{2+}$-binding protein that can bind up to eight $Ca^{2+}$ ions. In the apo-state, the protein is intrinsically disordered and folds upon metal binding. The folded protein adopts a parallel β-helix structure. $Ca^{2+}$-binding sites are located between turns of successive layers and are composed of one Asp residue and other donor atoms from the main chain to provide CN = 7. Recently, the capacity of RTX to bind $Ln^{3+}$ was investigated more thoroughly for potential applications in $Ln^{3+}$ recycling (see 6.3 Ln recovery and separation).[129] Circular dichroism experiments demonstrated that $Ln^{3+}$ induced the folding of the protein. However, it also suggested that the protein might adopt a different structure relative to the $Ca^{2+}$-induced folding, and the exact Ln-coordination sphere in this protein is still unknown.

*PTE-R18.* PTE-R18 is a mutant from a $Zn^{2+}$-dependent phosphotriesterase enzyme (PTE) designed for the hydrolysis of

2-naphthyl hexanoate. Jackson and coworkers investigated the use of this scaffold for $Ln^{3+}$-binding to design an Ln–enzyme and exploit the Lewis acid properties of $Ln^{3+}$ (see Catalysis).[21] Crystal structures obtained by co-incubation of the apo-protein and $Ln^{3+}$ (Eu, Gd) showed the positioning of the $Ln^{3+}$ ion in the $Zn^{2+}$-binding site coordinated by two His, one Asp (bidentate), a water molecule and a chloride ion (CN = 6). This low CN is surprising given the preference of $Ln^{3+}$ for a higher coordination number, and could explain the low occupancy ($\sim$20%) of the metal-site by $Ln^{3+}$.

### 2.6 Discussion and guidelines

A large variety of scaffolds were reported for $Ln^{3+}$-binding, which range from peptide sequences of a few AAs to proteins, that either fold upon metal-binding or that are pre-folded. For unfolded scaffolds or pre-folded scaffolds of intermediate size, additional stabilising interactions beyond the binding site are required to obtain a stable assembly, which in turn may improve Ln affinity (see Thermodynamic stability). In particular, the design of pre-folded scaffolds with a predisposed $Ln^{3+}$-binding site using short or intermediate-size peptides requires finding a balance between the stability of the structure, which relies mostly on hydrophobic interactions, and the introduction of metal-binding residues that can destabilise the scaffold. With this aim, two strategies were employed, sometimes in combination: (i) to put the metal binding site at a position that ensures a minimal perturbation of the scaffold (*e.g.* on the surface *vs.* inside the hydrophobic core of coiled coils, or at different positions within the hydrophobic core of coiled coils); (ii) to select amino acids known for their propensity to favour the formation of specific secondary structures. Overall, only a few studies investigated in detail the stability of their constructs and the impact of $Ln^{3+}$-binding site incorporation and $Ln^{3+}$ binding. These data could be helpful to guide the optimisation of scaffolds and validate the strategies followed.

All $Ln^{3+}$ coordination sites reported meet $Ln^{3+}$ preference for a high coordination number (CN 8–10) and hard Lewis bases. This includes either natural (Asp, Glu, Gla) or unnatural ($Ada_n$, $Ed3a_2$) negatively charged AAs and other AAs with O-donor atoms either in their side chains (Asn, Gln) or main chain (carbonyl of peptide bond). When such high CN cannot be achieved solely by the peptide or protein residues, water molecules are found coordinated to $Ln^{3+}$.

Control over the hydration number is important for applications in imaging (see 6. Applications). A few general principles can be drawn upon the analysis of the CN and $q$ of EF-hand motifs with distinct amino acid patterns (identity and position of residues 1, 3, 5, 9 and 12), which constitute the larger database on this topic. First, in EF-hand the presence of $Glu_9$ (rather than $Asp_9$) warrants a coordination sphere devoid of water molecules ($q$ = 0), which, as mentioned above, can be sought to optimize luminescent emission (see 6.1 Luminescent tags and probes). Instead, $Asp_9$ is never directly involved in $Ln^{3+}$-binding, but instead hydrogen-bonded to an $Ln^{3+}$-bound water molecule. Furthermore, higher CN (9–10) are observed in LanM, which are proteins evolved by Nature to bind $Ln^{3+}$ in physiological conditions, relative to $Ca^{2+}$-binding proteins or engineered LBTs (CN = 7–8). This does not stem merely from a change in the pattern of coordinating amino acids, but rather from second-sphere factors including a fine-tuning of backbone conformation and orientation of amino acid side chains. From the analysis of the data obtained with non-EF-hand scaffolds, two additional general principles can be underlined. First, the denticity of Ln-binding residues impacts $q$, as is seen with short peptides (*e.g.* $Ada_n$ *vs.* $Ed3a_2$) and coiled coils (*e.g.* Glu *vs.* Gla). Second, the environment in which the Ln-binding site is placed participates in the tuning of $q$, as evidenced in 3SCC, for which $Ln^{3+}$-hydration state depends on the position of $Ln^{3+}$-binding site in different regions (*e.g.* the more solvent-exposed termini *vs.* the buried hydrophobic core of the coiled coils), and on second-sphere effects (*e.g.* tuning of steric hindrance of second sphere residues). Thus, to summarize, control over the hydration number can be achieved by tuning three key parameters: (i) the denticity of Ln-binding residues and the length of their side chain; (ii) the position of the binding site in the scaffold; and (iii) the steric hindrance of second sphere residues. In addition, works on EF-hand motifs underline the influence of a fourth key parameter, more difficult to control, which is the spatial positioning of AAs side chains that allows AAs to bind in a mono- or bi-dentate fashion, and so participate in the fine-tuning of $Ln^{3+}$ hydration number.

## 3. Thermodynamic stability

### 3.1 Generalities

**3.1.1 Stability constants.** The thermodynamic stability, or affinity, of metal–ligand complexes $ML_n$ can be expressed by the overall stability constant $\beta_n$, which is the product of stepwise stability constants $K_n$ (eqn (3)):

$$\beta_n = K_1 K_2 \dots K_n = \frac{[M(L)_n]}{[M][L]^n} \quad (3)$$

The measurement of such a pH-independent formation constant is accessible through, for instance, potentiometric measurements of both apo-ligand and metal–ligand complex $pK_a$s, which represent the most accurate method to determine metal-binding affinities. However, for practical reasons (*e.g.* poor stability over a wide pH range and a large number of protonable sites), potentiometric titrations can be rarely performed on long peptides and proteins. Moreover, $K$ and $\beta$ values have poor significance under biological conditions, as they correspond to the affinity of a fully deprotonated ligand (*i.e.* at very high pH).

Nevertheless, conditional stability constants, $^cK$, at a given pH can be derived from the stability constant $K$ taking into account the protonation constants ($K_{a,n}$) of the ligand (eqn (4)).

$$^cK = \frac{K}{\left(1 + K_{a,1}[H^+] + K_{a,1}K_{a,2}[H^+]^2 + \dots K_{a,1}K_{a,2}K_{a,n}[H^+]^n\right)} \quad (4)$$

Conditional stability constants can be also determined through either direct metal–ligand titrations or competition experiments

 

using a second chelator of known affinity in the conditions used.[130,131] In this case, an apparent stability constant, $K_{app}$, is measured in the presence of the competitor in a given buffer, from which $^{c}K$ can be derived by taking into account the concentrations ($C$) and the conditional stability constants of the competitor (comp) and the buffer (eqn (5)).

$$^{c}K = K_{app}(1 + {}^{c}K_{comp}C_{comp} + {}^{c}K_{buffer}C_{buffer}) \quad (5)$$

The stability constants of peptide and protein scaffolds for $Ln^{3+}$ ions found in the literature and reported in this review have not been corrected for the buffer contribution, and hence they are apparent ($K_{app}$). This is also due to the lack of reliable stability constants for $Ln^{3+}$–buffer systems (see Section 3.1.2). In addition, it is noteworthy that the buffer concentration has a considerable impact on the $K_{app}$ of the complex under investigation: for instance, based on eqn (5), in the presence of a buffer with $^{c}K_{buffer} = 10^3$, the $K_{app}$ will be ~10-fold or ~100-fold lower than $^{c}K$ when the buffer concentration is 10 or 100 mM, respectively.

Another important point to keep in mind is that when working with peptides and proteins, the techniques used to determine the apparent stability constant can monitor the metal binding and/or the induced folding of the scaffold. For instance, CD (circular dichroism) can only inform indirectly on metal binding, since this technique measures metal-induced changes in the scaffold folding. Variations of chemical shifts in NMR spectroscopy are due to both metal-binding and subsequent conformational changes; similarly, ITC (isothermal titration calorimetry) measures the heat exchanged for all events happening during the titration, including conformational changes and metal-binding. To some extent, luminescence titration is also dependent on folding and metal binding. $Ln^{3+}$–luminescence increases upon $Ln^{3+}$-binding due to the substitution of water molecules. Thus, in general, $Ln^{3+}$–luminescence is also sensitive to induced local folding events (*i.e.* the formation of the binding site), but not to conformational changes that are more distant.

It is important to underscore that $K_{app}$ values should be compared only if they were determined in the same conditions (metal and ligand concentration, buffer, pH, *etc.*) and through the same technique, and refers to complexes with the same metal–ligand stoichiometry. For a more accurate comparison of complexes with different metal–ligand stoichiometry or different metal/ligand concentrations, the use of pM (eqn (6)) would be recommended.[132,133]

$$pM = -\log[M] = \log\left(^{c}K\frac{[L]^n}{[ML_n]}\right) \quad (6)$$

In the following, the thermodynamic stability of different $Ln^{3+}$-binding peptide and protein scaffolds is discussed. The reported apparent stability constants have been determined with a variety of methods and conditions, preventing rigorous comparisons but allowing discussion and speculation about trends and binding mechanisms.

**3.1.2 $Ln^{3+}$ coordination by pH buffers.** As mentioned in the former paragraph, knowledge of the stability constant of $Ln^{3+}$ complexes with pH buffers would be beneficial to take their contribution into account and derive buffer-independent binding constants. Only a few data are available in the literature on the affinity of the most common buffer for $Ln^{3+}$ (Table 7). Azab and coworkers studied the thermodynamic stability and speciation of $Ln^{3+}$–buffers systems using potentiometry.[134–136] They suggested the formation of 1 : 1, 1 : 2 $Ln^{3+}$–buffer complexes as well as [Ln(buffer)$(OH)_n$] species. The stability constants ($\log\beta$) reported for 1 : 1 $Ln^{3+}$–buffer complexes range between 3 and 4, but contrasting and hence no reliable values have been reported for the same $Ln^{3+}$–buffer system.

More recently, the interaction of some buffers with $Eu^{3+}$ was investigated employing NMR and time-resolved laser-induced fluorescence spectroscopy (TRLFS).[137] However, it is not clear how the values found in this study were determined. Overall, there is little agreement on the stability constants of $Ln^{3+}$–buffers systems, for which reliable values need to be determined. Notwithstanding, given the range of their stability constants, buffers could compete with peptides and proteins with low affinity for $Ln^{3+}$ ions. Therefore, the choice of the pH buffer and its concentration should be considered carefully.

Furthermore, buffer molecules could replace loosely bound water in the coordination sphere of $Ln^{3+}$ complexes, leading to the formation of ternary species, which has been investigated only in a few cases. For instance, Delangle and coworkers compared the number of $Ln^{3+}$-bound water molecules (hydration number, $q$) of a $Tb^{3+}$-cyclic decapeptide complex in HEPES buffer (10 mM, pH 7) and in the absence of buffer at pH 6.5.[41] In both cases, they obtained similar $Tb^{3+}$ luminescence lifetimes and hydration numbers ($\tau_{H_2O}$ = 1.65 ms, $\tau_{D_2O}$ = 6.99 ms, and $q$ = 2.0 in HEPES buffer; $\tau_{H_2O}$ = 1.68 ms, $\tau_{D_2O}$ = 6.28 ms and $q$ = 1.9 in absence of buffer), suggesting the absence of buffer molecules in the first coordination sphere of $Tb^{3+}$.

**3.1.3 $Ln^{3+}$ coordination by peptides and proteins.** The affinity of peptides and protein scaffolds for $Ln^{3+}$ ions spans over a very wide range ($\log K_{app}$ ~ 2–13; Fig. 15). Table 8 gives some examples of $\log K_{app}$ for $Tb^{3+}$ around neutral pH. Most of the scaffolds, regardless of their origin (natural or artificial) and size, have an affinity in the micro- to nanomolar range ($\log K_{app}$ = 5–9). The highest affinity is exhibited by large

Table 7 Stability constant ($\log\beta$) of 1 : 1 $Ln^{3+}$–buffers systems

| | Buffer | | | | | |
|---|---|---|---|---|---|---|
| $Ln^{3+}$ | MOPSO | MES | PIPES | HEPES | POPSO | MOPS |
| $La^{3+}$ | 3.34[a] | 3.40[a] | n.d. | 3.33[a] | n.d. | n.d. |
| $Ce^{3+}$ | 3.31[a] | 3.36[a] | n.d. | 3.40[a] | n.d. | n.d. |
| $Pr^{3+}$ | 3.36[a] | 3.39[a] | 4.11[b] | 3.44[a] | n.d. | n.d. |
| | 4.18[b] | 4.26[b] | | | | |
| $Eu^{3+}$ | 3.39[a] | 3.38[a] | 4.22[b] | 3.43[a] | 3.28[c] | 3.68[d] |
| | 4.24[b] | 4.27[b] | 2.51[c] | 3.04[c] | 3.68[d] | |
| | | 2.57[d] | 3.04[d] | 3.4[d] | | |
| $Gd^{3+}$ | 3.27[b] | 3.8[b] | 3.57[b] | n.d. | n.d. | n.d. |
| $Dy^{3+}$ | 4.09[b] | 4.19[b] | 4.03[b] | n.d. | n.d. | n.d. |

[a] Ref. 134. [b] Ref. 135. [c] Ref. 136. [d] Values have been calculated by correcting the $\log\beta$ reported in ref. 137 for the buffer $pK_a$s ($\log\beta$ + $pK_a$): MES, 6.27; PIPES, 7.14; HEPES, 7.56; MOPS, 7.18.[138]

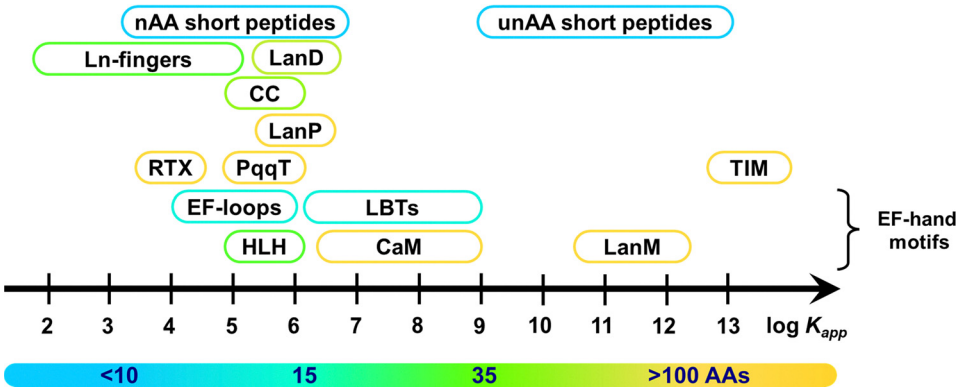


Fig. 15 Affinity range of different $Ln^{3+}$-binding peptide and protein scaffolds around neutral pH (6–7.5). Colour code is based on the scaffold size (as in Fig. 2). Abbreviations: nAA, natural amino acids; unAA, unnatural amino acids; CC, coiled coils; LBT: Ln-binding tag; HLH, helix-loop-helix.

Table 8 Apparent stability constants of different $Tb^{3+}$ complexes with peptide and protein scaffolds

| Scaffold (peptide/protein) | log $K_{app}$ ($Tb^{3+}$) | Ref. |
|---|---|---|
| RTX | 4.6[a] | 129 |
| Ln-fingers (LF4) | 4.9[b] | 103 |
| EF-hand (LBT1) | 5.1[c] | 139 |
| 3SCC (MB1-1) | 5.3[d] | 99 |
| HLH (P3W) | 5.4[e] | 80 |
| Cyclic decapeptide (PA) | 6.3[f] | 41 |
| CaM (bovine, site I) | 8.8[g] | 82 |
| LBT–protein conjugate (ILB1β-L3) | 8.9[h] | 65 |
| LanM (*Mex*-LanM) | 11.1[i] | 63 |
| unAA short peptides ($P^{HD2}$) | 12.7[j] | 87 |
| TIM barrel (TFD-EE N6W) | 13.1[k] | 126 |

[a] Bis-Tris 20 mM pH 6, NaCl 25 mM. [b] HEPES 10 mM pH 7.5, NaCl 10 mM. [c] HEPES 10 mM pH 7, NaCl 100 mM. [d] HEPES 10 mM pH 7. [e] HEPES 10 mM pH 6.9. [f] HEPES 10 mM pH 7. [g] PIPES 25 mM pH 6.8, KCl 100 mM. [h] HEPES 10 mM pH 7, NaCl 100 mM. [i] MOPS 30 mM pH 7.2, KCl 100 mM. [j] HEPES 10 mM pH 7, NaCl 100 mM. [k] HEPES 25 mM pH 7.5, NaCl 100 mM.

scaffolds, such as the natural LanM[81] and the *de novo* designed TIM barrel protein (TFD-EE N6W),[126] and by short peptides (6 amino acids) featuring multidentate unnatural amino acids.[87] In the following paragraphs, several factors influencing the affinity of each scaffold will be described and discussed.

### 3.2 EF-hand motifs

Thanks to the large variety of natural and engineered EF-hand motifs, it is possible to evaluate the effect of several structural factors influencing their affinity for $Ln^{3+}$ ions.

**3.2.1 Influence of the protein scaffold.** All naturally-occurring 12-residue EF-hand motifs isolated from their either $Ca^{2+}$- or $Ln^{3+}$-binding native protein scaffolds show similar affinities in the high micromolar range (log $K_{app} \approx 5$). Natural $Ca^{2+}$-binding proteins and LBT–protein conjugates bind $Ln^{3+}$ ions with at best log $K_{app} \approx$ 8–9 at pH $\sim$ 7. LanMs display the highest affinity (log $K_{app} \approx$ 10–12 at pH $\sim$ 7) among naturally-occurring $Ln^{3+}$-binding scaffolds. The structural determinants of such enhanced affinity are not clear. Based on the current literature,[62,81,83] it seems that it results from second-sphere effects, including different loop backbone conformation and hydrogen bonding networks that ensure an optimal orientation of coordinating amino acids within the loop.

The loss of 5 orders of magnitude in $Tb^{3+}$ affinity when EF1-3 loops are isolated from their native *Mex*-LanM scaffold ($\Delta$log $K_{app} \approx 5$, see entries 1 and 2, Table 9) highlights the momentous impact of a suitably folded, but not necessarily predisposed (note the LanM is disordered in the apo-state) scaffold on the affinity.

The important role of the core helical bundle in LanM has been further underscored by the recent study of a *Mex*-LanM fragment encompassing the highest-affinity and cooperative EF2-3 domains devoid of flanking helices, which showed much weaker affinity (log $K_{app} < 5$) than the whole protein.[140] Besides, another recent study has explored the impact of replacing EF-loops in CaM with those found in *Mex*-LanM.[141] Interestingly, this chimaera (LanM-GCaMP) showed a weak conformational response to $Ln^{3+}$ (from the micro- to the milli-molar range, depending on the $Ln^{3+}$). Nevertheless, the authors were able to obtain a modified CaM scaffold (LanTERN) with a conformational transition similar to LanM (log $K_{app}$ = 10.3 for $La^{3+}$) by introducing key mutations in CaM loops that are considered responsible for enhanced $Ln^{3+}$ selectivity in LanM (*i.e.* $Pro_2$ and $Asp_9$, see Section 3.2). Altogether, this evidence suggests that to further boost the $Ln^{3+}$-binding affinity of EF-hand proteins, the loop sequence, including non-coordinating amino acids, and the protein folding must be optimised simultaneously.

**3.2.2 Influence of primary sequence and post-translational modifications.** The primary sequences of LBT have been artificially evolved mainly to optimize $Tb^{3+}$ luminescence and affinity. Hence, they represent an outstanding dataset to derive structure–affinity relationships. In particular, the impact of specific modification can be quantified by comparing structurally analogous peptides (Table 9). For instance, the incorporation of hydrophobic amino acids (Phe, Tyr, Ile) at both N- and C-terminal to constrain the loop contributes to almost 1 order of magnitude of affinity (entries 4 *vs.* 3 in Table 9).[72] A similar gain in affinity (about 1.5 order of magnitude) was accomplished by

Table 9 Apparent stability constants of $Tb^{3+}$ complexes with distinct EF-hand motifs

| Entry | Sequence | Peptide/Protein | log $K_{app}$ | Ref. |
|---|---|---|---|---|
| 1 | DPDKDGTIDLKE<br>DPDKDGTLDAKE<br>DPDNDGTLDKKE<br>NPDNDGTIDARE | Isolated EF1-4 from Mex-LanM | 5.1 [a] | 143 |
| 2 | DPDKDGTIDLKE<br>DPDKDGTLDAKE<br>DPDNDGTLDKKE | LanM EF1-3 | 10.7 [b] | 63 |
| 3 | GDYNADGWIEFEEL | LBT precursor | 5.1 [c] | 139 |
| 4 | YIDTNNDGWIDGDELLA | m-sSE3 (LBT) | 5.9 [d] | 60 |
| 5 | YIDTDNDGWIDGDELLA | q-sSE3 (LBT) | 7.5 [d] | 60 |
| 6 | YIDTNNDGWIEGDELLA | sSE3 (LBT) | 7.4 [c] | 68 |
| 7 | GDYNKDGWYEELEL | LBT1 | 5.1 [c] | 139 |
| 8 | c(ACAGDYNKDGWYEELECAA) | LBTC2 | 6.6 [c] | 139 |
| 9 | …YIDTNNDGWIEGDEL… | ILB1β-S1 (LBT-protein) | 8.1 [c] | 65 |
| 10 | …YIDTNNDGWIEGDEL… | ILB1β-L3 (LBT-protein) | 8.9 [c] | 65 |
| 11 | …DKDGDGTIDERE… | P3W (HLH) | 5.4 [e] | 80 |

[a] MOPSO 10 mM pH 6.6, KCl 100 mM
[b] MOPS 30 mM pH 7.2, KCl 100 mM
[c] HEPES 10 mM pH 7, NaCl 100 mM
[d] HEPES 20 mM pH 7.5, NaCl 100 mM
[e] HEPES 10 mM pH 6.9

the cyclization of the EF-loop termini through the formation of disulfide bonds between Cys suitably located upstream and downstream of the loop (entries 8 *vs.* 7 in Table 9).[139,142]

Furthermore, as already mentioned for natural proteins, the insertion of LBT within a protein scaffold increases the affinity up to $\log K_{app} \approx 8$–9 (entries 8–9 in Table 9). Of note, the rigidity of the regions flanking the loops is also crucial to achieving such an enhancement of the affinity, which is not accomplished, for instance, in the chimeric helix-loop-helix domains ($\log K_{app} \approx 5$–6) likely due to the high flexibility of the flanking helices (entry 11 in Table 9).[80]

The pattern of coordinating amino acids in the EF loop also influences the affinity. Notably, the affinity increases of more than 1 order of magnitude (entries 5–6 *vs.* 4 in Table 9) when (i) $Asn_3$ is mutated to Asp (entries 5 *vs.* 4 in Table 9), likely due to the more negative charge of Asp relative to Asn, and (ii) $Asp_9$ is replaced by Glu (entries 6 *vs.* 4 in Table 9), which serves as bidentate ligands displacing the $Ln^{3+}$-bound water hydrogen-bonded to $Asp_9$.

In order to develop kinase/phosphatase- and nitration-responsive probes (see 6.1.5 Probe design and selected applications), the group of Zondlo has also widely explored EF-hand motifs featuring phosphorylated and nitrated amino acids, including phosphotyrosine (pTyr), phosphoserine (pSer), phosphothreonine (pThr) and 3-nitrotyrosine (nTyr). These studies have shown that $Glu_9$ can be replaced by pSer, pTyr and nTyr with minor impact ($\Delta\log K_{app} < \pm 1$) of the EF-loop affinity (entries 1–5 in Table 10).[144–146] A similar effect was also observed when Asp residues in positions 1, 3 or 5 were replaced by Cys sulfinic acid (entries 7–9 in Table 10).[147] Remarkably, an affinity gain of about 1 order of magnitude was reported when $Asp_9$ was replaced by pThr (entry 6 in Table 10).[148] Interestingly, a similar affinity gain was also observed for C-terminally truncated 9-residues loops bearing $pThr_9$ or $pSer_9$ (entries 10–12 in Table 10).[148] Unfortunately, the lack of structural studies prevents a detailed description of the $Ln^{3+}$ coordination sphere in these modified EF-hand motifs. Noteworthy, the impact of phosphorylated or nitrated amino acids in natural EF-hand proteins has not been investigated yet.

With respect to $Ln^{3+}$ binding to phosphorylated peptides, it is also worth pointing out that a micromolar affinity for $Tb^{3+}$ ($\log K_{app} = 6.5$ at pH 7) has been reported for a fragment of the

Table 10 Apparent stability constants of $Tb^{3+}$ complexes with different EF-hand motifs bearing modified amino acids (in red). pSer = phosphoserine; pTyr = phosphotyrosine; nTyr = 3-nitrotyrosine; pThr = phosphothreonine

| Entry | Sequence | Peptide/Protein | Modified amino acid | log $^{c}K$ | Ref. |
|---|---|---|---|---|---|
| 1 | DKNADGWIDAAEK | EF-E12 | n/a | 4.1 [a] | 144 |
| 2 | DKNADGWIDAASLK | phospho-pKID-PKC | $pSer_9$ | 3.7 [a] | 145 |
| 3 | DKNADGYIDAAYK | phospho-pKID-pTyr12 | $pTyr_9$ | 3.7 [a] | 144 |
| 4 | DKDADGWISPAYAK | pNO2-Y(OPO3)12 | $pTyr_9$ | 4.4 [a] | 146 |
| 5 | DKDADGWISPAYAK | pNO2-Y(NO2)12 | $nTyr_9$ | 4.8 [a] | 146 |
| 6 | DLNADGWITAATAK | phospho-pKID-Min-pThr9 | $pThr_9$ | 4.9 [b] | 148 |
| 7 | CKDADGWISPAEAK | $pSO_2$-C1-$SO_2^-$ | $sCys_1$ | 5.0 [a] | 147 |
| 8 | DKCADGWISPAEAK | $pSO_2$-C3-$SO_2^-$ | $sCys_3$ | 3.8 [a] | 147 |
| 9 | DKDACGWISPAEAK | $pSO_2$-C5-$SO_2^-$ | $sCys_5$ | 3.9 [a] | 147 |
| 10 | DLNADGWIT | phospho-pKID-Min2 | $pThr_9$ | 5.0 [b] | 148 |
| 11 | DLNADGWIS | phospho-pKID-Min1 | $pSer_9$ | 4.8 [b] | 148 |
| 12 | DKDADGWIS | phospho-pKID-Min4 | $pSer_9$ | 5.2 [b] | 148 |

[a] HEPES 5 mM pH 7.8, NaCl 100 mM
[b] HEPES 10 mM pH 7.5, NaCl 100 mM, $MgCl_2$ 2 mM

protein α-synuclein (119–132) with an EF-hand-like amino acid pattern and bearing a pTyr residue ($Asp_1$, $Asp_3$, $Glu_5$, $pTyr_7$, $Glu_{12}$).[149]

### 3.3 Other scaffolds

**3.3.1 Short and intermediate-size scaffolds.** In short and intermediate-size peptides, three main parameters influenced the affinity for $Ln^{3+}$: (i) the set of coordinating amino acids; (ii) their denticity and (iii) scaffold stability.

Among the short peptides reported in the literature (cyclic or linear), the one with the best affinity was obtained when using unnatural amino acids for $Ln^{3+}$-binding ($P^{HD2}$ with $Ada_2$ and $Ed3a_2$, $\log K_{app} = 12.7$, Fig. 5).[87] Compared to Glu and Asp, these unnatural amino acids have a higher denticity, which is more favourable for $Ln^{3+}$ binding. Furthermore, the peptide backbone designed to adopt a β-turn-fold upon metal binding provides additional stabilising interactions that make the peptide a better ligand than aminodiacetate groups separated by long alkyl chains ($n > 2$), but not as good as ligands such as EDTA for which the formation of 5- and 6-membered chelate rings result in a large stabilising chelate effect.[85]

The influence of the coordination sphere on the apparent affinity constants is evidenced for LF scaffolds, in which modification of the coordination sphere from $(Asp)_2(Glu)_2$ (LF4) to $AspAsn(Glu)_2$ (LF7) decreased the apparent affinity constant for $Eu^{3+}$ from $\log K_{app} = 4.6$ to $\log K_{app} = 4.0$. It is also apparent in the 3SCC designed by Peacock and coworkers when comparing the affinity of MB1-1 and CS1-1 for $Tb^{3+}$ ($\log K_{app} = 5.3$ *vs.* 4.6, respectively).[99] The additional $(Asn)_3$ layer in MB1-1 contributes to the slightly higher affinity observed relative to CS1-1. However, a change of hydration number for a given coordination sphere does not impact the apparent affinity constant, as can be seen in the 3SCC MB1-*n* series ($n = 1$–4), which have comparable affinities for $Tb^{3+}$ ($\log K_{app} = 5.2$–5.5) and hydration number ranging from $q = 0$ (MB1-2 and MB1-3) to $q = 3$ (MB1-1).[99]

The influence of denticity on Ln-affinity is well illustrated by the 2SCC described by Hodges when using Gla instead of Glu.[95] The affinity of $E_2(15,20)_{ox}$ has a $\log K_{app} \approx 2$ for $La^{3+}$, whereas the one of $Gla_2(15,20)N_{ox}$ has a $\log K_{app} = 6.2$ for $La^{3+}$, which can be linked to the higher denticity of the latter.

Finally, the link between scaffold stability and Ln-affinity was investigated in the design of the Ln-finger series. Changes in the sequence to improve the formation of secondary structures resulted in improved apparent affinity constants from $\log K_{app} < 2$ for $Eu^{3+}$ of LF1, to $\log K_{app} = 3.7$ of LF3.[103] Increasing the size of the binding site by deletion of two AAs (LF4) gave the best apparent affinity constant for $Eu^{3+}$ with $\log K_{app} = 4.6$. In the MB1 series shifting the position of the binding site within the 3SCC scaffold impacted the stability of the assembly (see 2.4.1 Coiled coils).[99] However, this did not significantly impact the apparent affinity constants for $Tb^{3+}$, which were all in the range $\log K_{app} = 5.2$–5.5. The addition of an additional heptad in MB1-1L also did not impact the apparent affinity constant for $Tb^{3+}$ ($\log K_{app} = 5.0$).[102]

**3.3.2 Proteins.** Proteins without EF-hand motifs cover a wide range of affinities for $Ln^{3+}$ with $\log K_{app}$ ranging from 4 to 13. This large range of affinities can be explained by $Ln^{3+}$-binding sites that are either not fully optimised for $Ln^{3+}$-binding (*e.g.* in $Ca^{2+}$–proteins), or found in natural proteins whose biological function may require affinity in an intermediate range (*e.g.* LanP, LanD).

The $Ca^{2+}$-binding protein RTX has an apparent affinity constant for $Ln^{3+}$ in the range $\log K_{app}$ = 4.1–4.6 at pH 6, and may bind 4 to 7 $Ln^{3+}$ per protein.[129] Note that these values were obtained for a fusion protein, in which RTX is fused with two fluorescent proteins for detection of folding upon metal binding by FRET. The affinity of LanP for $Ln^{3+}$ is $\log K_{app}$ = 6.0 and this protein binds 3 to 4 $Ln^{3+}$.[122] It was identified in Ln-utilizing organisms and due to its $Ln^{3+}$-binding ability, it was hypothesized that LanP could be of importance in Ln-dependent methylotrophy. Similarly, the apparent affinity constant of the monomeric protein LanD for $Ln^{3+}$ ranges from $\log K_{app}(La^{3+})$ = 5.7 to $\log K_{app}(Nd^{3+})$ = 6.6.[125] Because LanD and LanM are part of the same gene cluster involved in $Ln^{3+}$ transport and utilisation, it was proposed that LanD could be responsible for transferring $Ln^{3+}$ to LanM. This hypothesis was supported by the metalation of the modified fluorescent LaMP1[123] in the presence of $La^{3+}$–LanD, and by ITC titration experiments that showed a specific 1 : 1 interaction between apo-*Mex*-LanD and apo-*Mex*-LanM with a dissociation constant of 4 μM of physiological relevance. As of now, little is known about the coordination environment of $Ln^{3+}$ in these proteins, except for LanD, and the stability of these scaffolds. This does not allow the identification of specific features that could explain the affinity range observed.

When the protein scaffold and $Ln^{3+}$-binding site are optimised, high affinity can be obtained by design, as is the case for the computationally designed TIM barrel that displayed the highest affinity for $Ln^{3+}$.[126,150] The affinity of the mutant with a $(Glu)_4$ binding site and Trp for $Tb^{3+}$-sensitisation (TFD-EE N6W) was first suggested to be $\log K_{app} \approx 15$ by competitive titration with EGTA.[126] However, when the same experiments were reproduced with a higher concentration of $Tb^{3+}$–EGTA relative to TFD-EE N6W and longer equilibration times, the apparent affinity constant for $Tb^{3+}$ was rather found to be $\log K_{app}$ = 13.1.[150] Of note, this stressed the importance of taking into account kinetic considerations (see 5. Kinetic stability), since when protein metalation, or metal ion exchange with a given competitor, is slow, reaching the thermodynamic equilibrium may take days as it is the case for this example. Nevertheless, it also shows that in a rigid scaffold ($T_m > 95$ °C) with a predisposed binding site suitable for $Ln^{3+}$-binding, high affinity can be achieved by rational design.

Ln-ADH and its mimics stand apart since Ln-binding depends not only on the protein but also on the presence of the cofactor PQQ. Early work on Ln-ADH extrapolated apparent affinity constant based on activity assays and predicted an affinity in the low μM range for $Ln^{3+}$ in the presence of PQQ ($\log K_{app} \sim 6$).[112,119] More recently, Zeymer and coworkers studied the affinity of a PedH mutant for $Ln^{3+}$ in the absence of PQQ in the active site.[19] They determined an apparent affinity constant $\log K_{app}$ = 7.0 for $Tb^{3+}$ by luminescence titration, and $\log K_{app}$ = 6.2 for $Ce^{3+}$ based on competition with $Tb^{3+}$. Similarly, for the engineered PqqT scaffold, a $\log K_{app}$ = 6.2 was obtained for the mutant PqqT–$K_{142}$D containing PQQ (noted PQQ⊂PqqT–$K_{142}$D).[20] This represents a 10-fold improvement compared to the wild-type scaffold (PQQ⊂PqqT) and to the PQQ⊂PqqT–$K_{142}$A mutant. The presence of PQQ inside the protein was required for $Ln^{3+}$-binding, as demonstrated by controlled experiments performed with the apo-scaffolds (PqqT, PqqT–$K_{142}$A, and PqqT–$K_{142}$D).

### 3.4 Effect of the pH

Ln recovery requires ligands with high affinity and selectivity for $Ln^{3+}$ ions at pH ranging from pH < 1 to pH = 5–6 depending on the source (see 6. Applications).[151,152] The capacity to bind $Ln^{3+}$ ions in acidic conditions is an overlooked feature of peptides and proteins, mainly because (i) protein folding is generally sensitive to acidic pH and (ii) typical $Ln^{3+}$-binding residues (Asp, Glu) become protonated below pH 3–4, weakening $Ln^{3+}$ binding.

Nonetheless, some works investigated the pH sensitivity of their scaffolds. 3SCC loose their $Ln^{3+}$-binding ability below pH ∼ 4–5.[92] CaM and LBTs start releasing $Ln^{3+}$ below pH 6 and show no binding below pH 4.[153,154] Similarly to small chelators such as EDTA and DTPA, LanMs are able to bind $Ln^{3+}$ ions down to pH ∼ 2.5, retaining approximately nanomolar affinity at pH 4 ($\log K_{app} \approx$ 8–9) and 5 ($\log K_{app} \approx$ 9–10).[155] The reason behind the higher pH stability of LanM relative to CaM and LBTs is not understood. Interestingly, the RTX scaffold has been recently reported to retain partial $Ln^{3+}$-binding below pH 2. The protein could recover up to 20% of $Nd^{3+}$ from a synthetic NdFeB magnet solution containing Nd, Dy, Fe and Co at pH < 1.[129]

On the other end of the pH scale, studies on 3SCC scaffolds also showed that such moderate-affinity ligands undergo partial de-complexation of $Ln^{3+}$ at basic pH (> ∼8.5), likely due to competing formation of lanthanide hydroxides.[92]

More systematic studies will be necessary to understand the different pH-dependence of $Ln^{3+}$ binding to different peptide and protein scaffolds.

### 3.5 Discussion and guidelines

Among natural peptide and protein scaffolds, LanM stands out for its $\log K_{app} \approx$ 10–12. Although the structural factors for this remarkably high affinity have not been fully elucidated yet, the constraint imposed by the global protein scaffold, even if not pre-folded, is paramount. Indeed, isolated EF-loops with the same potential coordinating residues are much weaker ligands.

All rationally designed scaffolds with natural amino acids (*i.e.* short peptides, LFs and CCs), except for the TIM barrel ($\log K_{app}(Tb^{3+})$ = 13.1), have lower or at best comparable, affinity than natural scaffolds. As was noted earlier, these small to intermediate scaffolds are sensitive to the insertion of a metal-binding site, which destabilises the structure and in turn, may decrease $Ln^{3+}$-affinity. Peptides tend to be flexible and explore a large conformational space. $Ln^{3+}$-binding may improve the folding of the peptide; however, it is generally not sufficient

by itself to compensate for the entropic cost of restricting the peptide conformational space. Optimisations of the peptide sequence to improve the formation of secondary structures can help fold the peptide, and hence, increase the affinity for $Ln^{3+}$. This is observed qualitatively by the successive optimisations performed on Ln-fingers. It is also the case for the *de novo* designed TIM barrel, whose packing was optimised to rigidify the structure, and which displays a high stability ($T_m > 95$ °C) and affinity for $Ln^{3+}$. Similarly, even though LanMs are not pre-folded, their affinity for $Ln^{3+}$ is coupled to the formation of a thermally stable ($T_m > 95$ °C) folded protein. Nonetheless, there seem to be limitations to this, as increasing the size and stability of 3SCC (MB1-1 *vs.* MB1-1L) did not result in an impact on $Ln^{3+}$-affinity.[102] This could be due to the position of the binding site at the N-terminal, whereas the elongation was at the C-terminal. Interestingly, the correlation between the stability of the scaffold and metal-binding affinity has been observed for other metal-binding peptides. Studies performed on β-sheet WW-domain-like metal binding mini-proteins showed an improved affinity for $Zn^{2+}$, $Cu^{2+}$ and $Ni^{2+}$ by improving the stability of the scaffold.[156] Moreover, a study on Hg- and Cd-binding coiled coils demonstrated a direct correlation between the stability of the coiled coil scaffold and metal ion affinity.[157] Altogether, these data validate the strategy of optimising the scaffold structure in order to improve $Ln^{3+}$-affinity.

An alternative strategy is to use unnatural amino acids. Indeed, affinities comparable to that of LanM or TIM barrel were achieved in very short peptides (6 AAs) by combining $Ada_n$ and $Ed3a_2$ ($P^{HD2}$, $\log K_{app}(Tb^{3+}) = 12.7$). Even in this case, the primary sequence played an important part since residues that favour the formation of a β-turn upon metal-binding were chosen to bring additional stabilising interactions.

# 4. Selectivity

The selectivity of peptide and protein scaffolds for $Ln^{3+}$ *vs.* other metal ions is a key aspect for their applications in both environmental and biological samples, as well as for understanding Ln biochemistry. In particular, the selectivity among $Ln^{3+}$ and against actinides ($An^{3+}$) is of interest for $Ln^{3+}$ mining and separation, while the selectivity for $Ln^{3+}$ over $Ca^{2+}$ and other physiological cations (*e.g.* $Zn^{2+}$ and $Cu^{2+}$) is particularly relevant in the biological context.

The selectivity between two different metal ions can be quantified by calculating the selectivity factor, $\Delta\log K$ (or $\Delta\log\beta$), which corresponds to the difference between the affinities of a certain ligand for distinct metal ions. In the following, we will express the $Ln^{3+}$-selectivity, $\Delta\log K^{Ln}$ (or $\Delta\log\beta^{Ln}$) as the difference between the lowest and the highest $\log K^{Ln}$ along the series (eqn (7)).

$$\Delta\log K^{Ln} = \log K^{Ln}_{max} - \log K^{Ln}_{min} \quad (7)$$

Similarly, the selectivity for $Ln^{3+}$ *vs.* $Ca^{2+}$, $\Delta\log K^{Ca}$ ($\Delta\log\beta^{Ca}$), is expressed as the difference between the $\log K^{Ln}$ for a certain $Ln^{3+}$ and $\log K^{Ca}$ (eqn (8)).

$$\Delta\log K^{Ca} = \log K^{Ln} - \log K^{Ca} \quad (8)$$

## 4.1 Selectivity among $Ln^{3+}$

Prior to peptides and proteins, it is useful to discuss the selectivity of small chelators (Fig. 16 and 17A), which can be classified into four types.[158–161]

Most ligands, including EDTA (Fig. 16), show increasing stability constants across the Ln series (type I, black dots, Fig. 17A), commensurate with the increase of $Ln^{3+}$ Lewis acidity. For other ligands, such as DTPA and DOTA (Fig. 16), the affinity increases along the first part of the series, reaches a maximum and then remains constant or even decreases for late $Ln^{3+}$ ions (type II, red dots, Fig. 17A). This is the result of the ligand steric hindrance, which impairs to accommodate the smaller $Ln^{3+}$ ions.[158,159,162]

An uncommon behaviour is shown by a few ligands, including Macropa (Fig. 16), for which a reverse-size selectivity is observed due to a fairly rigid scaffold better suited for larger than smaller $Ln^{3+}$ ions (type III, blue dots, Fig. 17A).[160] Finally, a biphasic selectivity trend (type IV, green dots, Fig. 17A) has been recently reported with ligands such as Macrotripa (Fig. 16) that are better suited for both early and late $Ln^{3+}$ due to a switch between a 10-coordinated conformation accommodating larger $Ln^{3+}$ ions and an 8-coordinated conformation stabilizing smaller $Ln^{3+}$ ions.[161,163]

Overall, the selectivity trend among $Ln^{3+}$ ions results from a compromise between electronic and steric effects, whose

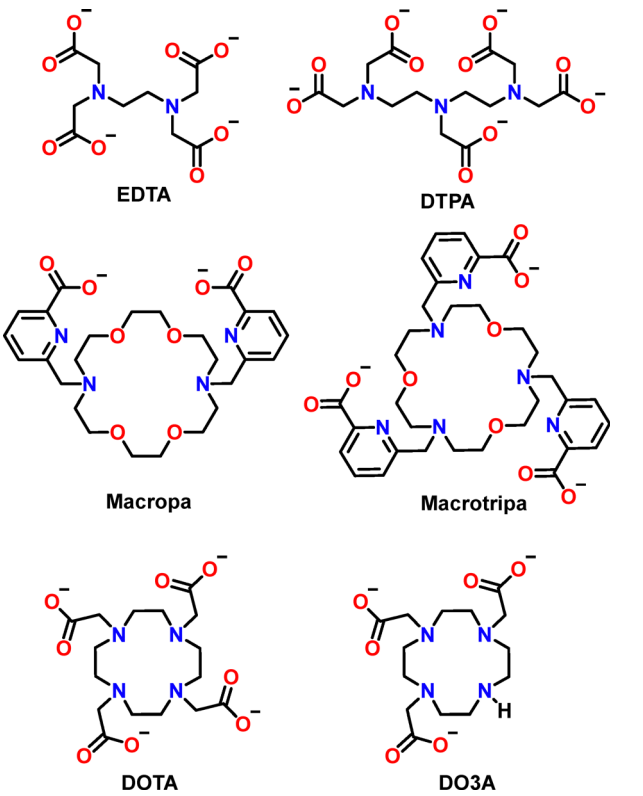


Fig. 16 Chemical structure of the $Ln^{3+}$ chelators mentioned in the text.

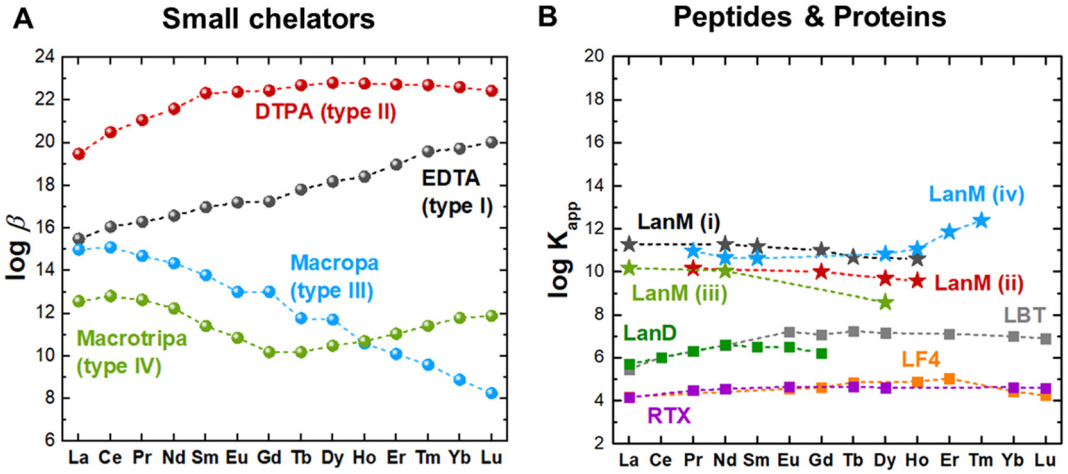


Fig. 17 Selectivity of different ligands across the $Ln^{3+}$ series. (A) Small chelators and (B) peptides and protein scaffolds. Note that log $\beta$ and log $K_{app}$ are reported in (A) and (B), respectively. Values for LanM are referred to (i) *Mex*-LanM at pH 7.2 measured by CD spectroscopy; (ii) *Mex*-LanM at pH 5 measured by CD spectroscopy; (iii) *Hans*-LanM at pH 5 measured by CD spectroscopy; (iv) *Mex*-LanM at pH 5 measured by UV-vis spectroscopy. The values plotted are reported in Table 15.

balance varies depending on the properties of the ligand, such as its denticity and rigidity. Moreover, the behaviour of these small chelators points out a common misconception: $Ln^{3+}$ are not all the same, especially when bound to a chelator. Indeed, a remarkable difference can be observed in the stability constants of certain ligands for distinct $Ln^{3+}$ ions: for instance, EDTA and Macropa show selectivity for $Lu^{3+}$ *vs.* $La^{3+}$ ($\Delta\log\beta^{Ln}$) of ~5 and ~7 orders of magnitude, respectively.

The $Ln^{3+}$-selectivity of only a few peptides and protein scaffolds has been reported (Fig. 17B). LBTs, Ln-fingers, RTX and LanD display a type II selectivity (Fig. 17B). The affinity of LBTs increases by almost 2 orders of magnitude from $La^{3+}$ up to $Ln^{3+}$ in the middle of the series ($Eu^{3+}$–$Tb^{3+}$) and then slightly decreases towards the end of the series.[64] Since LBTs were derived from $Ca^{2+}$-binding loops, their preference for middle-sized $Ln^{3+}$ ions (1.04–1.07 Å, CN = 8 in LBTs) could be explained by the closer similarity to $Ca^{2+}$ ionic radius (1.06 Å, CN = 7 in EF loops).

Ishida and coworkers employed molecular dynamics simulations, as well as ITC and NMR measurements, in order to elucidate the selectivity trend of LBTs.[164] The authors found that one or two water molecules can be accommodated in the coordination sphere of large $Ln^{3+}$ ions ($La^{3+}$–$Nd^{3+}$); this weakens the binding between the $Ln^{3+}$ and $Asn_5$ and enhances the flexibility of the complex, resulting in reduced affinity. For $Ln^{3+}$ from $Sm^{3+}$ to $Lu^{3+}$, water binding is rarely observed, correlating with the higher affinity for these smaller ions.

For LF4, the affinity increases by nearly 1 order of magnitude from $La^{3+}$ to $Er^{3+}$, then decreases back towards $Lu^{3+}$, which was found to have the same affinity as $La^{3+}$.[103] A similar trend is shown by RTX and LanD proteins, but in this case the $\Delta\log K^{Ln}$ measured is much lower (<1).[125,129]

The relative selectivity of 3SCC among $Ln^{3+}$ has been also reported based on luminescence measurements of $Tb^{3+}$ displacement by other $Ln^{3+}$ ions.[101] A bell-shaped selectivity trend was found for several 3SCC scaffolds that differ for the location of the $Ln^{3+}$-binding site along the helices (MB1 series, Fig. 18 and Table 5). In particular, no significant discrimination was observed among medium-sized $Ln^{3+}$ ions ($Nd^{3+}$–$Tb^{3+}$) for all scaffolds. For competing ions smaller than $Tb^{3+}$, size-dependent discrimination was observed with scaffolds where the binding site is located around the centre or the C-terminal of the coiled coil. The higher promiscuity of the N-terminal

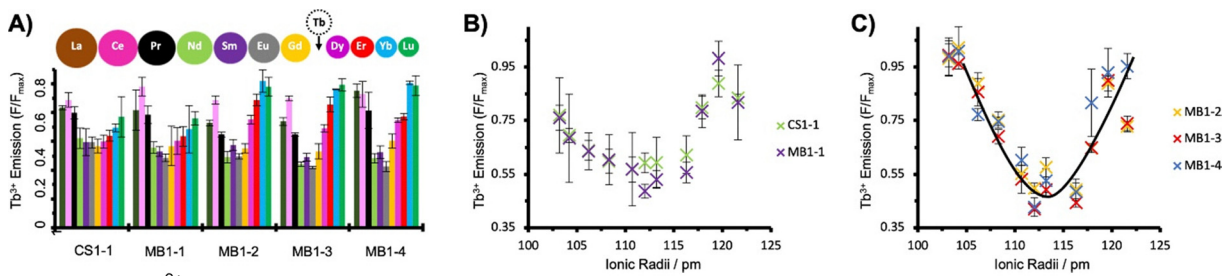


Fig. 18 $Tb^{3+}$–displacement experiments followed by luminescence in three-stranded coiled coils showing a bell-shape selectivity within the $Ln^{3+}$ series (MB1 series, Table 5). (A) Luminescence displacement, (B) Comparison of CS1-1 and MB1-1, (C) Comparison of MB1-2, MB1-3, and MB1-4. Reproduced from ref. 101.

binding site was attributed by the authors to its greater flexibility, which hence allows it to better accommodate also smaller $Ln^{3+}$ ions.[101]

LanP[122] and PqqT[20] do not display significant selectivity among early $Ln^{3+}$ ions ($La^{3+}$–$Gd^{3+}$; late $Ln^{3+}$ were not studied). The TIM barrel TDF-EE N6W binds $Eu^{3+}$, $Gd^{3+}$, and $Tb^{3+}$ with comparable affinity, while ~10-fold weaker binding was observed with $Ce^{3+}$.[126]

Among $Ln^{3+}$-binding peptides and proteins, LanM represents a peculiar and controversial case. In most reports by the groups of Cotruvo and Daumann, an unusual preference of LanM for larger ($La^{3+}$–$Eu^{3+}$) over smaller $Ln^{3+}$ ions has been underscored based on CD and luminescence measurements,[81,123,165] which are responsive to both $Ln^{3+}$ binding and $Ln^{3+}$-induced conformational change (Fig. 17B, LanM(i) and LanM(ii)). It must be noted again, that similarly to RTX and 3SCC, *Mex*-LanM shows a very modest $Ln^{3+}$-discrimination ($\Delta\log K^{Ln} \sim 0.5$) relative to small chelators and LBTs (Fig. 17B). An enhanced selectivity ($\Delta\log K^{Ln} \approx 1.6$) for early ($La^{3+}$, $Nd^{3+}$) *vs.* late $Ln^{3+}$ ($Dy^{3+}$) was found for *Hans*-LanM (Fig. 17B, LanM(iii)), thanks to a $Ln^{3+}$ size-dependent dimerization.[62]

Curiously, the opposite trend, *i.e.* an affinity increase across the end of the $Ln^{3+}$ series, was observed when the intrinsic $Ln^{3+}$-affinity of *Mex*-LanM was determined *via* UV-vis-NIR spectrophotometric competition experiments (Fig. 17B, LanM(iv)).[155] As suggested by the authors,[166] this could highlight a decoupling between $Ln^{3+}$-binding and conformational change in LanM, which warrants further investigations, and underscores the importance of the method and the conditions chosen to determine and compare affinity values. Nonetheless, higher retention of early *vs.* late $Ln^{3+}$ ions was observed for immobilized LanM upon pH-induced desorption.[62,167,168] Such (at least apparent) inconsistency between the affinity values measured *via* different techniques and in different conditions is worth further systematic investigations.

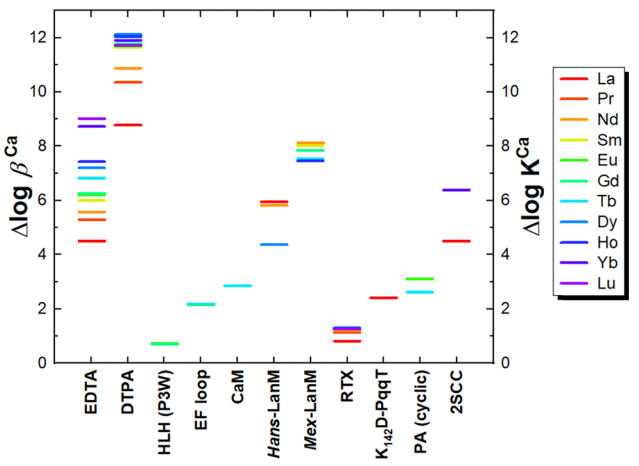


Fig. 19 Selectivity of small chelators and some peptide/protein scaffolds for $Ln^{3+}$ over $Ca^{2+}$. $\Delta\log\beta$ and $\Delta\log K_{app}$ are reported for small chelators and peptide/protein scaffolds, respectively. Values plotted were calculated from values reported in Table 16.

Finally, it is also worth mentioning that LanM showed higher affinity for the rare-earth element $Sc^{3+}$ relative to $Ln^{3+}$ ions (3-fold higher relative to $Nd^{3+}$),[167] commensurate with its smaller ionic radius and higher Lewis acidity.

### 4.2 Selectivity over $Ca^{2+}$

Peptide and protein scaffolds generally show selectivity for $Ln^{3+}$ relative to $Ca^{2+}$ (and $Mg^{2+}$) which is classically attributed to their higher charge and coordination numbers (Fig. 19). In addition, as $Ln^{3+}$ have higher CN than $Ca^{2+}$ in water, $Ln^{3+}$-binding releases a higher number of water molecules, resulting in a more favourable entropic change.

Among the EF-hand motifs, LanMs show a remarkably higher conformational selectivity against $Ca^{2+}$ ($\Delta\log K^{Ca} \approx 8$ for *Mex*-LanM)[81,155] relative to both natural and engineered EF loops ($\Delta\log K^{Ca} \approx 1$–3).[59,82,169] In this respect, it is worth noting that $Ca^{2+}$ seems to bind to LanM with $\log K^{Ca} = 5.5$ without inducing a conformational change, which occurs upon further $Ca^{2+}$ binding events with $\log K^{Ca} \approx 3$.[81]

The $Pro_2$ residue found in all EF loops of *Mex*-LanM is crucial for the conformational selectivity of LanM for $Ln^{3+}$ over $Ca^{2+}$. Notably, mutation of all four $Pro_2$ residues to Ala reduces the selectivity of *Mex*-LanM for $Ln^{3+}$ against $Ca^{2+}$ by strengthening the conformational response of $Ca^{2+}$-binding to LanM of more than 2 order of magnitudes.[81] Moreover, $Pro_2$ also had a key role in the $Ln^{3+}$-specific conformational response of LanTERN chimera already discussed above (see 3.2.1 Influence of the protein scaffold).[141] However, it is noteworthy that high selectivity over $Ca^{2+}$ ($\Delta\log K_a^{Ca} \approx 4$–6) is also observed in *Hans*-LanM, in which a $Pro_2$ residue is found only in one of the four EF-hand loops. Hence, it is likely that the unique selectivity of LanM for $Ln^{3+}$ *vs.* $Ca^{2+}$ stems from minor differences in the second coordination sphere and backbone conformation.[62]

The main difference between Ca-dependent and Ln-dependent alcohol dehydrogenases is the presence of an additional Asp in the coordination sphere of $Ln^{3+}$. This residue appears essential for $Ln^{3+}$-coordination and selectivity relative to $Ca^{2+}$. Indeed, mutation of this AA to Ala in MDH and EDH enzymes led to the production of enzymes metallated only with $Ca^{2+}$.[110] However, the mutation did not lead to catalytically competent Ca-enzymes, which underscores that differences between Ca-dependent and Ln-dependent ADH beyond the metal coordination sphere are important for catalysis.

Among the *de novo* designed $Ln^{3+}$-binding proteins, 2SCC scaffolds bearing non-canonical multidentate carboxyglutamic (Gla, Fig. 6) residues show a selectivity over $Ca^{2+}$ comparable to that of small chelators and LanMs.[95]

### 4.3 Selectivity over $An^{3+}$

The groups of Deblonde and Daumann have explored the capacity of *Mex*-LanM to bind $An^{3+}$ ions.[165,170–172] LanM exhibits a higher affinity for $An^{3+}$ over $Ln^{3+}$ ions and $Y^{3+}$. In particular, a ~5-fold stronger affinity was found for $Am^{3+}$ and $Cm^{3+}$ relative to $Ln^{3+}$ of similar size ($Pr^{3+}$, $Nd^{3+}$, $Sm^{3+}$),[170] and for $Ac^{3+}$ over $Y^{3+}$.[171] This trend is similar to that of small

chelators such as EDTA and DTPA, which show slightly higher selectivity for $Am^{3+}$ over $Nd^{3+}$ ($\Delta \log \beta \sim 1.5$).[173]

For small complexes, the non-negligible covalency of $An^{3+}$–ligand bonds has been generally considered accountable for the higher stability of $An^{3+}$-complexes over their $Ln^{3+}$-analogues, but this remains to be ascertained for $An^{3+}$–protein complexes.[170]

An improved selectivity for $An^{3+}$ over $Ln^{3+}$ ions at low pH ($\sim$3) was obtained by Deblonde and coworkers *via Mex*-LanM engineering with softer Asn ligands in place of native $Asp_9$ in each loop (LanM $3D_9N$).[174]

A slightly higher affinity for $Cm^{3+}$ over $Eu^{3+}$ has also been reported for the EF-hand loops isolated from *Mex*-LanM.[143] $Am^{3+}$ binding to an LBT was shown to be $\sim$6-fold stronger than $Nd^{3+}$ and comparable to smaller $Ln^{3+}$ ($Tb^{3+}$, $Eu^{3+}$). A higher ($\sim$10-fold) selectivity for $Am^{3+}$ over $Nd^{3+}$ could be obtained by a modified LBT bearing a $Cys_5$ residue functionalized with a soft 2-methylene-pyridine group in place of $Asp_5$.[175]

Daumann and coworkers also demonstrated that $An^{3+}$ ions such as $Am^{3+}$ and $Cm^{3+}$ can replace early $Ln^{3+}$ ions in the catalytic activity of the $Ln^{3+}$-dependent MDH enzyme and ensure the growth of $Ln^{3+}$-dependent bacteria in the absence of $Ln^{3+}$ ions.[176]

### 4.4 Selectivity over d-block metal ions

Despite its relevance for biomedical and bio-metallurgic applications, the selectivity of $Ln^{3+}$-binding peptides and proteins over d-block metal ions is a largely neglected aspect. Generally, a higher selectivity for $Ln^{3+}$ over d-block metal ions is obtained by leveraging their preference towards O- (carboxylates, carbonyls) over N-donors (*e.g.* histidines), and higher coordination numbers (8–10 *vs.* 4–6). For instance, LanP displayed no binding to $Fe^{2+}$, $Fe^{3+}$, or $Zn^{2+}$.[122] A 3SCC scaffold (MB1-2) engineered with Asn and Asp residues for $Ln^{3+}$ binding, shows $\sim$6-fold higher affinity for $Tb^{3+}$ over $Cu^{2+}$.[98,177] A remarkably high selectivity for $Nd^{3+}$ ions over $Cu^{2+}$ ($\Delta \log K \sim 6$) and $Zn^{2+}$ ($\Delta \log K \sim 8$) has been reported for *Mex*-LanM at pH 5.[155] On the contrary, EDTA shows higher selectivity for $Cu^{2+}$ over $Nd^{3+}$ (*i.e.* 100-fold higher affinity for $Cu^{2+}$ over $Nd^{2+}$) and no selectivity between $Nd^{3+}$ and $Zn^{2+}$, while DTPA has no selectivity between $Nd^{3+}$ and $Cu^{2+}$, but a $\sim$1000-fold higher affinity for $Nd^{3+}$ relative to $Zn^{2+}$.[155,173] It is worth noting that EDTA has a higher affinity for $Cu^{2+}$ ($\log \beta^{Cu} = 18.8$) than most $Ln^{3+}$ ions ($La^{3+}$–$Ho^{3+}$) and DTPA has a higher affinity for $Cu^{2+}$ ($\log \beta^{Cu} = 21.4$) than the early $Ln^{3+}$ ($La^{3+}$–$Pr^{3+}$). Altogether, these data point out that peptides and proteins benefit from a higher selectivity for the whole $Ln^{3+}$ series against d-block metal ions relative to small chelators.

## 5. Kinetic stability

Beyond thermodynamic parameters, kinetic aspects are also important for a complete characterization of $Ln^{3+}$-binding peptides and proteins, especially for further applications, such as MRI or Ln recovery (see 6. Applications).

### 5.1 Lability and inertness: exchange reaction

Due to the shielding by 5s and 5p orbitals, 4f orbitals do not undergo strong ligand field stabilisation, making $Ln^{3+}$ ions kinetically labile.

The lability/inertness is witnessed by the rate of the self-exchange reaction. Such a reaction can proceed according to two extreme pathways: associative or dissociative, corresponding to an increase or a decrease in the coordination number (CN) of the intermediate species, respectively. In turn, this depends on the CN of the starting complexes: for instance, $Ln^{3+}$ complexes with low CN (CN $\leq$ 8, $Gd^{3+}$–$Lu^{3+}$) are expected to proceed through an associative pathway with a nine-coordinate intermediate, while those with high CN (CN $\geq$ 9, $La^{3+}$–$Eu^{3+}$) by a dissociative pathway with an eight-coordinated intermediate (Fig. 20A).[30] The rates of self-exchange for a certain $Ln^{3+}$ ion depend on the relative stabilities of the eight- and nine-coordinated species for that $Ln^{3+}$.

Recently, Peacock and coworkers measured the activation parameters for water exchange in the $Gd^{3+}$ complex with a 3SCC scaffold (MB1-1, $q = 3$) and proposed an associative mechanism.[102] Furthermore, key insight can be deduced from the comparative NMR study of $La^{3+}$ *vs.* $Lu^{3+}$ binding to the LBT3 peptide.[164] In particular, $Ln^{3+}$ binding to LBT3 showed a fast

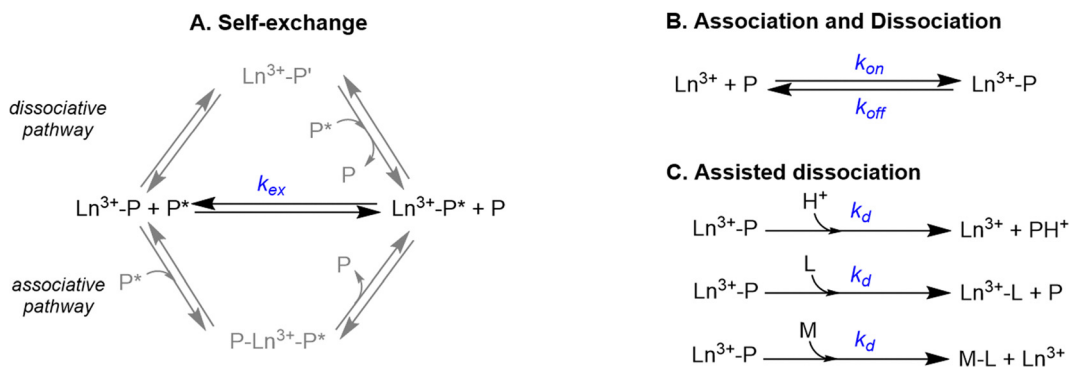


Fig. 20 (A) Self-exchange reaction characterized by the rate constant $k_{ex}$ and related pathways. P and P* are two molecules of the same peptide. In the case of the dissociative pathway, P′ corresponds to peptide P which has some of its binding residues de-coordinated from the Ln. Hence, $Ln^{3+}$ has a lower CN when bound to peptide P′ than P. (B) Association and dissociation reactions with their respective rate constants $k_{on}$ and $k_{off}$. (C) Simplified scheme of proton-, ligand- and metal-assisted dissociation pathways (rate constants $k_d$).

and slow exchange regime for $La^{3+}$ and $Lu^{3+}$, respectively,[164] which may be explained through the different structural features of the complexes. Indeed, the $La^{3+}$–LBT3 complex has CN = 9, with $q$ = 1 and is structurally flexible, hence the de-coordination of the weakly-bound (relative to the peptide ligands) $H_2O$ molecule can promote the self-exchange reaction through a dissociative pathway. Instead, the $Lu^{3+}$–LBT3 complex has CN = 8, with $q$ = 0, and poor solvent accessibility, hindering the possibility of undergoing an associative mechanism.

As a small number of data is available, it is difficult to generalize whether the exchange reaction proceeds *via* a dissociative or an associative mechanism in other $Ln^{3+}$-binding peptides and proteins, and which parameters would influence the preference between one or the other pathways.

Furthermore, the self-exchange rate $k_{ex}$ can be easily inferred by NMR experiments. For most unfolded or poorly folded peptides, a fast exchange regime characterized by a broadening and shift of the signals was reported at sub-stoichiometric $Ln^{3+}$ : P ratio[76,164,169,178] except for a 2SCC scaffold[95] and an HLH scaffold[18] for which a slow exchange regime was observed. This is in line with the intrinsic lability of $Ln^{3+}$ ions and the absence of constraints induced by the peptide scaffolds.

### 5.2 Association and dissociation rates

Beyond the exchange rate $k_{ex}$, key parameters to be defined are the association and dissociation rates, noted $k_{on}$ and $k_{off}$, respectively. These two parameters are related to the thermodynamic stability constants $K$ according to eqn (9):

$$K = \frac{k_{on}}{k_{off}} \quad (9)$$

The $k_{on}$ value can be measured experimentally by monitoring the appearance of spectroscopic signatures corresponding to the $Ln^{3+}$ complex when formed within the right time scale (milliseconds to hours). Note that such $k_{on}$ values reflect both $Ln^{3+}$ binding to the protein and the induced folding unless the protein is pre-folded in the apo-state. Qualitative estimation of the $k_{on}$ can be deduced for Ln-ADH[19] and TDF-EE N6W,[126,130] for which a long incubation time (hours) was required to load the protein with $Ln^{3+}$, indicating a slow association rate.

The $k_{off}$ can be then calculated from the $k_{on}$ and stability constant $K$ values, using eqn (9). Indeed, the $k_{off}$ cannot be directly determined experimentally since the release of the $Ln^{3+}$ ion from a polydentate ligand, such as peptides and proteins, to form the fully solvated species is unlikely. Nevertheless, the complete dissociation of a $Ln^{3+}$ from a polydentate ligand can be assisted by (i) protonation of the ligands (proton-assisted dissociation), (ii) capture by another ligand(s) (ligand-assisted dissociation), which can also be the buffer, (iii) trans-metalation by another metal ion (metal-assisted dissociation), or any combination of these three processes. It is worth noting that in these processes, the otherwise disfavoured $Ln^{3+}$ release occurs through the alteration of the system thermodynamics, *e.g.* by weakening the ligand affinity (acid-assisted dissociation) or by inducing metal-transfer/exchange reaction with competing ligand or cations. Thus, although the concept of complex inertness is often equated with that of "assisted dissociation" by language abuse, they refer to intrinsically distinct processes.

Ligand-assisted dissociation rate constants ($k_d$), which differs from $k_{off}$, were reported by Falke and coworkers for the complex of $Tb^{3+}$ with an EF-hand protein (galactose-binding protein, GBP) using excess EDTA as the competing ligand.[179] Interestingly, by measuring the dissociation rates of several mutants, these authors showed that the 9th residue of the EF-hand motif acts as a gateway that tunes the kinetics of $Ln^{3+}$ dissociation. Notably, the $Glu_9$ mutant displayed the slowest dissociation rate ($\sim$0.006 $s^{-1}$) compared to $Asp_9$ ($\sim$0.002 $s^{-1}$) and $Asn_9$ ($\sim$1 $s^{-1}$).

The chelator-free $k_{off}$ of $Tb^{3+}$-bound to *Mex*-LanM ($\sim$0.02–0.05 $s^{-1}$) was estimated by extrapolation from $k_{obs}$ values obtained by stopped-flow spectrofluorimetry in the presence of different concentrations of EGTA.[63] However, this approach is questionable, since the $k_{off}$ value relative to a dissociative mechanism is inferred from ligand-assisted dissociation rates, for which a different, namely associative, mechanism is at play. Moreover, it is worth noting that, as mentioned for $k_{on}$ values, dissociation rates may also take into account the contribution of protein unfolding induced by $Ln^{3+}$ release.

Daumann and coworkers monitored over time the metal exchange reactions between a pre-formed $Eu^{3+}$–LanM complex and the other $Ln^{3+}$ ions. Interestingly, equilibration times varied from minutes for the late $Ln^{3+}$ to hours for the early $Ln^{3+}$, correlating with the size and Lewis acidity of the competing ions.[165]

### 5.3 Discussion and guidelines

The kinetic aspects of $Ln^{3+}$ coordination by peptides and proteins have been largely neglected and overlooked even though they are important in applications for imaging and Ln recovery. NMR is the method of choice to get a better picture of the dynamics of the system. A simple $^1H$ NMR titration experiment can help gain qualitative insights into the exchange rates using the diamagnetic $La^{3+}$ and $Lu^{3+}$, and thus determine the exchange regime on the NMR timescale. Comparison with paramagnetic $Ln^{3+}$ can bring additional information on $Ln^{3+}$ coordination site, and even peptides/protein solution structure.

Although it is difficult to predict the kinetic behaviour, based on the results and concepts described above, it is anticipated that the spontaneous dissociation of $Ln^{3+}$ ions from polydentate peptides and proteins is negligible, especially in the case of buried sites within well-folded scaffolds. Notwithstanding, $Ln^{3+}$ dissociation can be triggered upon acidification or the addition of competing ligands or cations. In the case of ligand-assisted dissociation, the presence of weakly bound water molecules in the $Ln^{3+}$ coordination sphere is expected to promote dissociation since its replacement facilitates the formation of ternary intermediate complexes with the competing ligand.

## 6. Applications

$Ln^{3+}$-binding peptides and proteins have been employed for a number of applications spanning from luminescence sensing

 

Table 11 Non-exhaustive overview of applications reported for the main types of $Ln^{3+}$-binding peptide and protein scaffolds

| | Scaffold | Luminescence imaging/sensing | MRI | Metal ion separation | Catalysis (reaction) |
|---|---|---|---|---|---|
| EF-hand motifs | LBT | X | X | X | |
| | HLH | | X | | X (hydrolysis) |
| | LanM | X | X | X | |
| Other Ln-binding sites | Cyclic decapeptide | | X | | |
| | 3SCC | | X | | |
| | LanD | | | X | |
| | RTX | | | X | |
| | TIM barrel | | | | X (alcohols photooxidation) |
| | ADH or ADH mimics | | | | X (alcohols oxidation) |

and magnetic resonance imaging to metal ion separation and catalysis (Table 11). In the following paragraphs, applications in these four major fields will be discussed with respect to the peculiar properties of each type of $Ln^{3+}$–peptide and protein scaffold.

### 6.1 Luminescent tags and probes

**6.1.1 $Ln^{3+}$ luminescence.** Most $Ln^{3+}$ ions display luminescent emission with several advantageous features relative to organic fluorophores and have hence gained interest for applications in optical bioassays, bioimaging and luminescence-guided surgery, as reviewed in ref. 180–185. In particular, they benefit from narrow emission bands and long luminescence lifetimes (μs–ms) at relatively invariant wavelengths, in a range spanning from the visible to the near-infrared (NIR) depending on each $Ln^{3+}$, and they are not sensitive to photobleaching. Nevertheless, due to the low extinction coefficients ($\varepsilon$ = 1–10 $M^{-1}$ $cm^{-1}$) of Laporte-forbidden f–f transitions, the direct excitation of $Ln^{3+}$ emission is quite inefficient.[30,186,187] This downside is generally circumvented through the so-called antenna effect (Fig. 21).[30,186,187] This is the indirect sensitization of $Ln^{3+}$ *via* (i) the excitation of a suited chromophore – the antenna – to an excited singlet state ($S_1$), (ii) the energy transfer to an excited triplet state ($T_1$) *via* intersystem crossing, (iii) the energy transfer to an excited state of $Ln^{3+}$, from which emission takes place. Even if it is common to observe sensitisation from $T_1$, the $S_1$ state can also be the sensitising level.[188] The use of an antenna also warrants larger pseudo-Stokes shifts (difference between the antenna absorption and $Ln^{3+}$ emission wavelengths) than accessible with organic fluorophores.

The efficiency of the energy transfer from the antenna to the $Ln^{3+}$ depends on (i) the energy gap between an antenna triplet state and the $Ln^{3+}$ excited acceptor state, which depends on the antenna–$Ln^{3+}$ couple, and (ii) the antenna–$Ln^{3+}$ distance. Several non-radiative pathways compete with $Ln^{3+}$ luminescence:[189] (i) the antenna fluorescence and phosphorescence emissions, (ii) quenching of the antenna $T_1$ state by $^3O_2$, (iii) back energy transfer from $Ln^{3+}$ excited state to the antenna triplet state when the energy levels are close, (iv) photo-induced electron transfer (PeT) from the excited antenna to the $Ln^{3+}$, to which the most reducible $Ln^{3+}$ ($Eu^{3+}$, $Yb^{3+}$, $Sm^{3+}$) are particularly sensitive,[190] (v) non-radiative vibrational de-excitation of $Ln^{3+}$ excited state by high-energy oscillators such as O–H, N–H and C–H groups.[39,40] As a result, $Ln^{3+}$-bound $H_2O$ molecules quench $Ln^{3+}$ emission. Thus, the type and position of the antenna as well as the hydration number are key parameters influencing the luminescence of $Ln^{3+}$ complexes, including those with peptide and protein ligands.

**6.1.2 Type of antenna.** Early studies using $Tb^{3+}$ luminescence to investigate $Ca^{2+}$-binding sites in proteins[54,191] have established that the amino acid tryptophan (Trp) is a suitable antenna for this $Ln^{3+}$, also showing its superiority relative to Phe and Tyr. Since then, Trp has been by far the most common antenna incorporated in peptide and protein scaffolds to

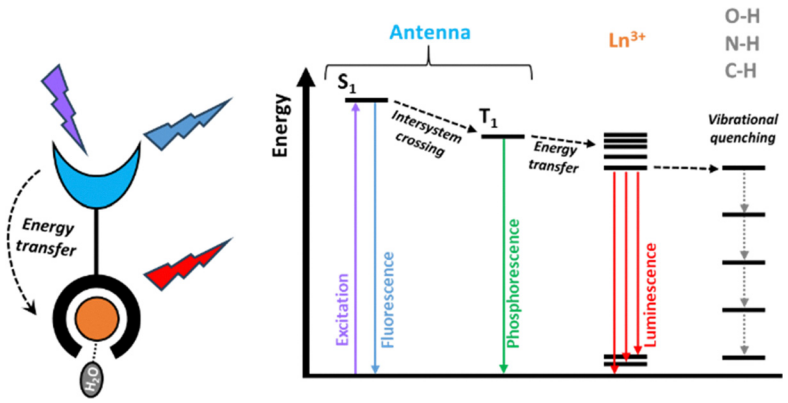


Fig. 21 Scheme of the "antenna" effect: upon irradiation, a chromophore called "antenna" (blue, which can be part of or conjugated to a Ln-chelator, black) is excited to a singlet state ($S_1$); then, the energy is transferred to an excited triplet state ($T_1$) *via* intersystem crossing and finally transferred to an excited level of the $Ln^{3+}$ (orange) from which emission takes place. Vibrational de-excitation by high-energy oscillators such as O–H groups is also depicted.

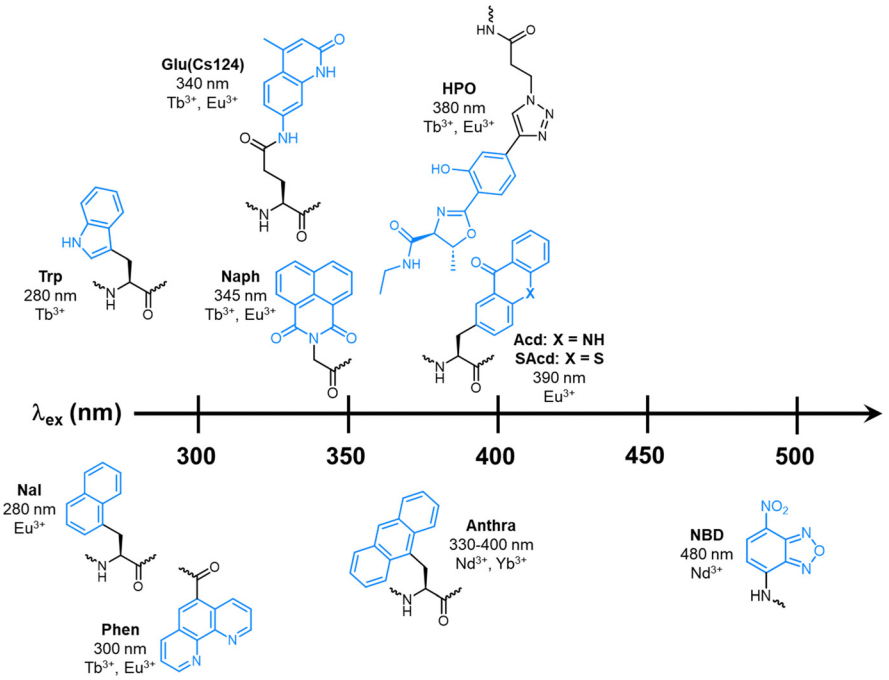


Fig. 22 Antenna amino acids found in $Ln^{3+}$-binding peptide scaffolds discussed in this review (top) and in peptides with appended polydentate $Ln^{3+}$-chelators (bottom). The excitation wavelength and the $Ln^{3+}$ ions that each antenna is able to sensitize are indicated. The antenna moieties are highlighted in blue. Abbreviations: Trp, tryptophan; Glu, glutamate; Cs124, carbostyryl-124; Naph, 1,8-naphthalimide; HPO, hydroxyphenol oxazoline; Acd, acridone; Nal, naphthyl-alanine; Phen, phenanthroline; Anthra, anthryl-alanine; NBD, nitrobenzodiazole.

sensitize $Tb^{3+}$ luminescence. Indeed, Trp ($\lambda_{ex}$ = 280–290 nm) is characterized by a triplet state located 3700 $cm^{-1}$ above the emissive $^5D_4$ state of $Tb^{3+}$, which makes it an efficient antenna for this $Ln^{3+}$ ion.[192] Trp could also be suitable for $Eu^{3+}$-sensitisation; however, PeT from Trp quenches $Eu^{3+}$-emission. PeT has also been described by Horrocks between Trp and $Yb^{3+}$ in the EF-hand protein parvalbumin.[193] While for $Eu^{3+}$ this process competes with $Eu^{3+}$-emission, it was hypothesized that PeT could be in part responsible for the population of $Yb^{3+}$ excited state and its subsequent emission.

Despite its straightforward incorporation in peptide and protein sequences, Trp is not an ideal antenna for applications in bioimaging, as its excitation wavelength in the UV spectral region challenges applications in samples with high absorption and fluorescence background, such as biological media.[194]

To overcome such limitations, different unnatural antenna amino acids have been introduced. Gunnlaugsson and coworkers incorporated at the N-terminal of an EF-loop isolated from parvalbumin a 1,8-naphthalimide chromophore ($\lambda_{ex}$ = 345 nm) able to sensitise $Tb^{3+}$ and $Eu^{3+}$ emission (Fig. 22).[169,178] The HPO ligand introduced into 3SCCs scaffolds (see 2.4.1 Coiled coils) was able to sensitize $Tb^{3+}$ and $Eu^{3+}$ upon excitation at 380 nm. Imperiali and co-workers have developed LBTs with unnatural amino acids featuring a carbostyryl-124 (Cs124) or an acridone (Acd) antenna able to sensitize $Tb^{3+}$ (Cs124) and $Eu^{3+}$ (Cs124 and Acd) at longer wavelengths, namely ~340 and 390 nm, respectively (Fig. 22).[195] In particular, the Cs124 antenna was able to sensitise $Tb^{3+}$ better than Trp, while Acd sensitised $Eu^{3+}$ better than Cs124 (Fig. 22).

Lately, Acd and its sulfur-substituted SAcd (Fig. 22) were incorporated into *Mex*-LanM *via* genetic code expansion, achieving long-lived sensitized $Eu^{3+}$ emission upon excitation at 390 nm.[196] Interestingly, SAcd showed to be a more efficient antenna than Acd due to a prolonged triplet state lifetime.

These unnatural antenna amino acids were not able to sensitise the emission from other $Ln^{3+}$ ions including $Nd^{3+}$, $Er^{3+}$, $Tm^{3+}$ and $Yb^{3+}$, which are NIR emitters. Actually, no $Ln^{3+}$-binding peptide or protein scaffolds have been reported bearing suitable antenna amino acids for NIR-emitting $Ln^{3+}$ ions, which are particularly sensitive to quenching by O–H, N–H and even C–H oscillators. Thus, in addition to a suitable antenna, the design of Ln–peptides and Ln–proteins as NIR emitters will also require avoiding O–H oscillators in their first coordination sphere, and likely also N–H oscillators. To the best of our knowledge, these antenna amino acids are the sole used in $Ln^{3+}$-binding peptide and protein scaffolds.[8] Other antennae were described in peptide and proteins functionalised with polydentate $Ln^{3+}$-chelators such as EDTA and DOTA as reviewed in part in ref. 23. In these cases, the antenna has been introduced through either (i) ligand derivatization, often with conjugated pyridine or picolinate groups,[197–199] or (ii) unnaturall amino acids (Fig. 22), including naphthalene (for $Eu^{3+}$, $\lambda_{ex}$ = 280 nm),[200] NBD (for $Nd^{3+}$, $\lambda_{ex}$ = 480 nm),[201,202] anthracene (for $Nd^{3+}$ and $Yb^{3+}$, $\lambda_{ex}$ = 330–400 nm)[203] and phenanthroline (for $Tb^{3+}$ and $Eu^{3+}$, $\lambda_{ex}$ = 300 nm).[204] Similarly to HPO, phenanthroline served as both $Ln^{3+}$ ligand and antenna. Such an $Ln^{3+}$-coordinating antenna is generally preferable because its proximity to the ion enhances the energy transfer efficiency.

**6.1.3 Position of the antenna.** As the $Ln^{3+}$–antenna distance is critical to achieve efficient sensitization, antennas are introduced in close proximity to the metal-binding site. For instance, in 3SCC scaffolds (Table 5) Trp has been mostly introduced within the two $Ln^{3+}$ binding Asn and Asp residues of each strand (–NEWKD–).[98]

Szabo and coworkers have thoroughly investigated how the position of aromatic amino acids (Trp, Phe or Tyr) within the EF-hand motif sequence influences $Tb^{3+}$ emission.[66] In particular, they compared the excitation and emission spectra of $Tb^{3+}$ bound to EF-hand peptides containing aromatic amino acids in positions 2, 4, 7 or 10. Very interestingly, they found that, regardless of which amino acid occupied positions 2, 4 and 10, the amino acid in the 7th position is responsible for $Tb^{3+}$ sensitization due to its closer distance to the ion ($\sim$5 Å). For instance, despite Phe and Tyr having a lower extinction coefficient than Trp, $Tb^{3+}$ was mainly sensitized by either Phe or Tyr in the 7th position even when Trp was simultaneously present in the 2nd position. Moreover, the authors reported a favourable impact of Tyr in the 2$^{nd}$ and 4$^{th}$, but not 10$^{th}$, position on the $Tb^{3+}$ emission in peptides having Trp in the 7$^{th}$ position. In particular, the results suggested that this occurs thanks to an energy transfer between excited Tyr and Trp, which in turn sensitises $Tb^{3+}$ emission. Based on these studies, $Trp_7$ has been incorporated as the $Tb^{3+}$ sensitizer in many EF-hand motif scaffolds, including LBTs[72,73] and *Mex*-LanM mutants.[63] The 7th loop position has also been used to incorporate unnatural Cs124 and Acd antenna amino acids in LBTs[195] and *Mex*-LanM (in the latter case after screening of 42 incorporation sites).[196] Zondlo and coworkers also showed that $Tb^{3+}$ emission can be sensitized using Trp at the 8$^{th}$ position of EF-loop-like peptides.[148]

**6.1.4 Hydration number.** As already mentioned, $Ln^{3+}$-bound water molecules quench $Ln^{3+}$–luminescence, which is used to determine $Ln^{3+}$ hydration number. This is why, for instance, the coordination sphere of luminescent $Tb^{3+}$–LBT complexes was optimized to achieve $q$ = 0 (see 2.2.2 Lanthanide-binding tags). The effect of the hydration number is also proven by the correlation between the emission intensity of $Ln^{3+}$ bound to different 3SCC scaffolds (MB1-*n* series, $n$ = 1–4) and their $q$ values (0, 2 or 3) (see 2.4.1 Coiled coils).[99]

**6.1.5 Probe design and selected applications.** A simple application of emissive $Ln^{3+}$-complexes is as luminescent tags attached to a (bio)molecule or receptor. In this regard, LBTs have been developed as encodable luminescent tags that can be easily conjugated to proteins *via* recombinant expression and applied, for instance, to the study of protein–protein interactions by luminescence resonance energy transfer (LRET).[139,205] Recently, LanM-Acd mutants conjugated with Affibodies and complexed with $Eu^{3+}$ were used as a secondary antibody reagent for immunofluorescence staining.[196]

Besides, peptide and protein scaffolds with suitable antennae may serve as luminescent probes for $Ln^{3+}$ ions. For instance, a *Mex*-LanM Trp mutant (T90W) has been successfully employed to measure $Tb^{3+}$ in acid mine drainage at very low pH ($\sim$3),[63] while LanM-Acd mutants were applied for $Eu^{3+}$ detection in cell culture.[196]

Alternatively, Walker and coworkers reported on the design of fluorescent proteins with a high concentration of negative charges at their surface for the detection of $Ln^{3+}$ (Ln = Sm, Eu, Tb, Dy, Tm, Yb) by LRET in the micro- to millimolar concentration range.[206] Instead of exciting an antenna, it is the $Ln^{3+}$ that is excited and transfers its energy to the fluorescent proteins, which resulted in a twofold higher fluorescence intensity and longer emission lifetimes.

As reviewed in ref. 23 and 180, $Ln^{3+}$-based responsive probes, whose luminescent signal changes upon the interaction with a given target, can be conceived based on the above-mentioned factors influencing $Ln^{3+}$ luminescence, such as (i) chemical modifications of the antenna, (ii) variation of the $Ln^{3+}$–antenna distance, for instance through conformational changes induced upon target-binding or *via* proteolytic cleavage, and (iii) modification of the $Ln^{3+}$ coordination sphere and hydration number.

These principles have been used to develop several responsive probes based on peptides with appended $Ln^{3+}$–chelator complexes.[23] For instance, Vazquez and coworkers designed a peptide probe that folds upon RNA binding, inducing the coordination of a phenanthroline antenna to an appended $Tb^{3+}$–EDTA complex.[204] Similarly, Sénèque and coworkers developed $Zn^{2+}$-[201] and RNA-responsive[207] peptide probes based on conformational changes that decrease the distance between the antenna and a $Ln^{3+}$–DOTA complex attached to the peptide. Exploiting the distinct emission bands of $Ln^{3+}$ couples such as $Tb^{3+}$–$Eu^{3+}$ in the visible and $Nd^{3+}$–$Yb^{3+}$ in the NIR, the group of Sénèque also developed ratiometric probes (*i.e.* based on the ratio between two signals) introducing two suitably positioned $Ln^{3+}$–chelator complexes.[200,203] In order to synthesize such regio-selectively hetero-metallated complexes, they used native chemical ligation to assemble two peptides functionalized with different $Ln^{3+}$ complexes. Of note, this strategy relies on the kinetic inertness of the $Ln^{3+}$-complexes grafted on the peptides and has never been applied so far on peptide scaffolds with ''intrinsic'' $Ln^{3+}$–binding sites.

Lately, LanM-Acd mutants have been developed as protease sensors based on the impairment of the antenna effect upon protein cleavage. To this end, protease recognitions sequences were incorporated in suited position with the LanM scaffold.[196]

Pazos and coworkers developed peptides with appended $Tb^{3+}$–DO3A complex (Fig. 15) responsive to (i) Ser phosphorylation, based on a decrease of the hydration number,[208] and (ii) to Tyr nitration, relying on the quenching of $Tb^{3+}$ emission by 3-nitro-Tyr.[209] Quenching of $Tb^{3+}$ emission was also exploited in the design of $Cu^{2+}$-selective peptide probes.[194,210]

Zondlo and coworkers also developed $Tb^{3+}$ complexes with EF-hand-like peptides (Table 10) responsive to (de)phosphorylation of Tyr, Ser and Thr,[144,145,148] Tyr nitration[146] and Cys oxidation.[147,211] In particular, Cys, Tyr, Ser or Thr residues were introduced in place of Asp or Glu residues within EF-loops (Table 10), drastically reducing $Tb^{3+}$-binding ($\log K_{app} < 4$), while their derivatization with phosphate or sulphinyl groups restored a $\log K_{app} \approx 4$–5. Hence, in this case, the response was mainly due to a change in the metallopeptide affinity upon amino acid modification.

Table 12 Properties of $Gd^{3+}$-complexes relevant for MRI applications

| Scaffold | Peptide/protein | MW (kDa) | $q^a$ | $\tau_R$ (ns) | $\tau_m{}^b$ (ns) | $r_1$ ($mM^{-1}$ $s^{-1}$) | $r_2$ ($mM^{-1}$ $s^{-1}$) | Magnetic field (T) | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| DOTA | n/a | ~0.5 | 1 | ~0.08 | 244 | 4.3 | 5.7 | 7 | 215 |
| | | | | n/a | n/a | 4.7 | n/a | 0.47 | 216 |
| DTPA | n/a | ~0.5 | 1 | ~0.06 | 300 | n/a | n/a | n/a | 215 |
| | | | | n/a | n/a | 4.7 | n/a | 0.47 | 216 |
| | | | | n/a | n/a | 3.5 | n/a | 1.4 | 77 |
| | | | | n/a | n/a | 5.4 | 8 | 1.5 | 217 |
| | | | | n/a | n/a | 4.2 | 6.8 | 3 | 217 |
| | | | | n/a | n/a | 5.1 | 9.4 | 7 | 218 |
| HLH | P3W | ~4 | 2 | n/a | n/a | 16.2 | n/a | 0.47 | 77 |
| | | | | | | 21.2 | n/a | 1.4 | |
| LBT | m-sSE3 | ~2 | 1 | n/a | n/a | 5.5 | n/a | 11.75 | 60 |
| | sSE3 | | 0 | n/a | n/a | 1.2 | n/a | 11.75 | 60 |
| dLBT | dSE3 | ~4 | 0.08 | n/a | n/a | 5.9 | n/a | 11.75 | 60 |
| dLBT–protein | dSE3-Ubiquitin | ~12 | 0.08 | n/a | n/a | 2.3 | n/a | 11.75 | 60 |
| | q-dSE3-Ubiquitin | | 1 | n/a | n/a | 4.2 | n/a | 11.75 | 60 |
| | xq-dSE3-Ubiquitin | | 0; 1 | n/a | n/a | 5.0 | n/a | 11.75 | 60 |
| Cyclic decapeptide | PA | ~1.3 | 2 | 0.4 | n/a | ~30 | ~40 | 4.7 | 41 and 219 |
| | | | | | | ~20 | ~35 | 9.4 | |
| | | | | | | ~20 | ~35 | 11.75 | |
| 3SCC | MB1-1 | ~12 | 3.1 | 7 | 1.56 | 10 | 89.3 | 7 | 102 |
| | | | | | | 64.3 | 87.9 | 1 | |
| | MB1-1 (2W) | | 0 | n/a | n/a | 3.9 | 24.2 | 7 | 100 |
| | MB1-2 | | 0 | n/a | n/a | 4.2 | 21.3 | 7 | 99 and 100 |
| | MB1-3 | | 0 | n/a | n/a | 4 | 20.9 | 7 | 99 |
| | MB1-4 | | 2 | n/a | n/a | 7.5 | 37.9 | 7 | 99 and 100 |
| | MB1-1L | ~15 | 3.7 | 10 | n/a | 10.9 | 81.8 | 7 | 102 |
| | | | | | | 67.4 | 96.5 | 1 | |
| Proteins | CA1.CA2 | ~12 | 2 | 9.1 | n/a | 117 | 129 | 1.5 | 217 |
| | | | | | | 48 | 88 | 3 | |
| | | | | | | 6 | 50 | 9.4 | |
| | ProCA32 | ~12 | 0.5 | n/a | n/a | 33.4 | 44.6 | 1.4 | 218 |
| | | | | | | 21.9 | 56.9 | 4.7 | |
| | | | | | | 18.9 | 48.6 | 7 | |
| | LanM | ~12 | 2 | n/a | n/a | 12.1 | 17 | 3 | 220 |
| | LanND | ~12 | 2 | n/a | n/a | 13.2 | 25.1 | 3 | |
| | | | | | | 6.9 | 31.2 | 7 | |

[a] Values obtained by the measurement of $Tb^{3+}$ or $Eu^{3+}$ luminescence lifetimes. [b] Values measured at 298 K.

### 6.2 Magnetic resonance imaging

Clinical contrast agents (CAs) for magnetic resonance imaging (MRI), reviewed by Caravan and coworkers in ref. 212, are mostly $Gd^{3+}$-complexes with polyaminocarboxylate chelators, such as DOTA and DTPA (Fig. 16). $Gd^{3+}$-based CAs (GBCAs) enhance image contrast by shortening the longitudinal ($T_1$) and transverse ($T_2$) relaxation time of proximal water protons, to an extent that depends on their longitudinal ($r_1$) and transverse ($r_2$) relaxivity, as defined by eqn (10):

$$\frac{1}{T_i^{obs}} = \frac{1}{T_i^0} + r_i[Gd] \quad i = 1, 2 \qquad (10)$$

where $1/T_i^0$ is the inherent relaxation rate of the tissue and $1/T_i^{obs}$ is the relaxation rate in the presence of the GBCA. Of note, the relaxivity ($r$) depends on the external magnetic field strength (Table 12). Clinical MRI scanners currently use low magnetic fields (~0.5–3 T), while higher fields (up to ~11.7 T) are used for research purposes. Nevertheless, in order to improve signal-to-noise ratio and spatial resolution, clinical MRI is progressively moving towards higher field strength.[213]

In order to increase MRI sensitivity, reduce the injected GBCA dose and hence limit toxic side-effects, GBCAs with higher relaxivity are currently sought.[212] The approaches developed to achieve this goal have been reviewed in ref. 214. The main molecular factors affecting the inner-sphere relaxivity of GBCAs are the hydration number $q$, the mean residency time $\tau_m$ of Gd-bound water and the rotational correlation time $\tau_R$ (Fig. 23).[212] Hereafter, these factors are discussed with respect to the $Gd^{3+}$ complexes with peptides and proteins that have been investigated as GBCAs (Table 12).

**6.2.1 Hydration number.** Inner-sphere relaxivity is directly proportional to the hydration number $q$, and indeed commercial GBCAs feature one or two water molecules in the coordination sphere. To meet this requisite, Imperiali and coworkers developed LBTs with $q$ = 1 (m-sSE3 and q-dSE3-Ubiquitin, Table 12), which showed similar relaxivity relative to commercial GBCAs at high magnetic field (11.75 T).[60] As already mentioned above (2.2.2 Lanthanide-binding tags), the increase of $q$ from 0 to 1 was accomplished by mutating $Glu_9$ to $Asp_9$, and it was accompanied by a nearly 5-fold increase in longitudinal relaxivity (Table 12).

Peacock and co-workers have investigated the relaxivity of $Gd^{3+}$–3SCC scaffolds, showing a significant dependence on the

 

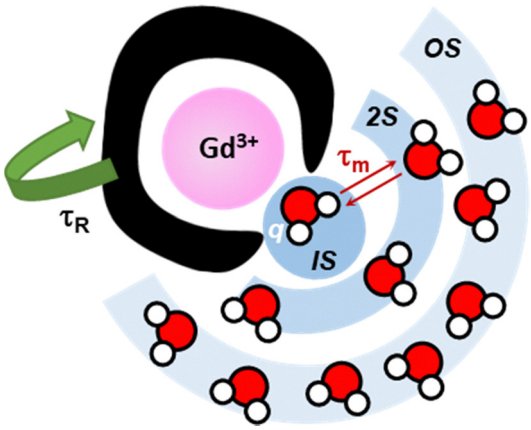


Fig. 23 Scheme of a GBCA and graphical description of the parameters that influence its relaxivity, including inner-sphere (IS) contributions (hydration number, $q$; water residency time, $\tau_m$; rotational correlation time, $\tau_R$) as well as second (2S) and outer (OS) sphere effects.

location of the $Ln^{3+}$-binding site along the coiled coil and the second sphere layer, which both affect the hydration number (see 2.4.1 Coiled coils).[98–100,102,104] Remarkably, scaffolds with high hydration numbers, such as MB1-1 ($q$ = 3), MB1-4 and the recently reported LanM derivative LanND ($q$ = 2), display significantly higher relaxivity compared to clinical GBCAs, especially at low clinical field (1 T, Table 12).

**6.2.2 Rotational correlation time.** The rotational correlation time, $\tau_R$, reflects the rate of molecular tumbling. Of note, the impact of $\tau_R$ on the relaxivity depends on the strength of the magnetic field, as illustrated in Fig. 24 by the nuclear magnetic relaxation dispersion (NMRD) profiles (*i.e.* relaxivity profiles as a function of the magnetic field strength) for different $\tau_R$ values (A) short ~0.1 ns; B) intermediate; C) long, ~10 ns).[212] In general, the relaxivity increases with increasing $\tau_R$ at low magnetic fields (< ~1 T in Fig. 24), whereas at higher fields, the relaxivity increases with increasing $\tau_R$ up to a threshold beyond which the relaxivity decreases upon further $\tau_R$ increase.

To improve the relaxivity, a $\tau_R$ increase can obtained by increasing the molecular weight and the rigidity of GBCAs.[214]

For instance, the $Gd^{3+}$ complex with the rigid cyclodecapeptide scaffold PA (see 2.3.1 Cyclic decapeptide) showed remarkably high relaxivity (20–40 $mM^{-1}$ $s^{-1}$, Table 12) at high magnetic fields (4.7–11.75 T) thanks to a medium-range molecular weight (~1.3 kDa) implying an optimal $\tau_R \approx 0.4$ ns.[41]

Besides, Yang and co-workers have developed protein-based CAs by engineering protein scaffolds with $Gd^{3+}$-binding sites, achieving much higher relaxivity relative to small GBCAs at both low and high fields (Table 12).[217,218] In CA1.CD2 (~12 kDa), where a *de novo* designed rigid $Gd^{3+}$-binding site was engineered into a compact CD2 (a cell adhesion protein) domain, the enhanced relaxivity relative to small GBCAs is mainly the result of a higher hydration number ($q$ = 2) and a remarkably longer $\tau_R$ (~9 ns) relative to commercial GBCAs.[217] It is worth noting that the local rigidity of the $Gd^{3+}$ complex, and not only the molecular weight, is crucial to achieving high relaxivity. As an example, Yang and coworkers also showed that $Gd^{3+}$ bound to a CD2 domain engineered with a flexibly conjugated EF-loop from CaM showed a relaxivity comparable to that of small commercial GBCAs (~3.4 $mM^{-1}$ $s^{-1}$ at 3 T).[217] Analogously, the mere conjugation of double LBT (dLBT) on a protein scaffold (ubiquitin) did not improve their relaxivity (Table 12).[60]

A long $\tau_R$ (7–10 ns) was also reported for $Gd^{3+}$–3SCC complexes;[102] however, it is noteworthy that lengthening the coiled coil from five (MB1-1, ~12 kDa) to six heptads (MB1-1L, ~15 kDa) and the consequent $\tau_R$ increase had negligible impact on the relaxivity (Table 12), suggesting that the tumbling rate is not a limiting factor for such scaffolds.

**6.2.3 Water residency time.** The relaxivity has roughly a bell-shaped dependence on the water residency time ($\tau_m$), as the water exchange must be neither too fast nor too slow to maximize the relaxation of bulk water molecules. Similarly to the case of $\tau_R$, the impact of $\tau_m$ on the relaxivity also depends on the strength of the magnetic field: at the higher fields, the optimal relaxivity can be attained by a broad range of $\tau_m$, while at low fields the optimal relaxivity can be obtained only for a narrow range of $\tau_m$.[212] Of note,

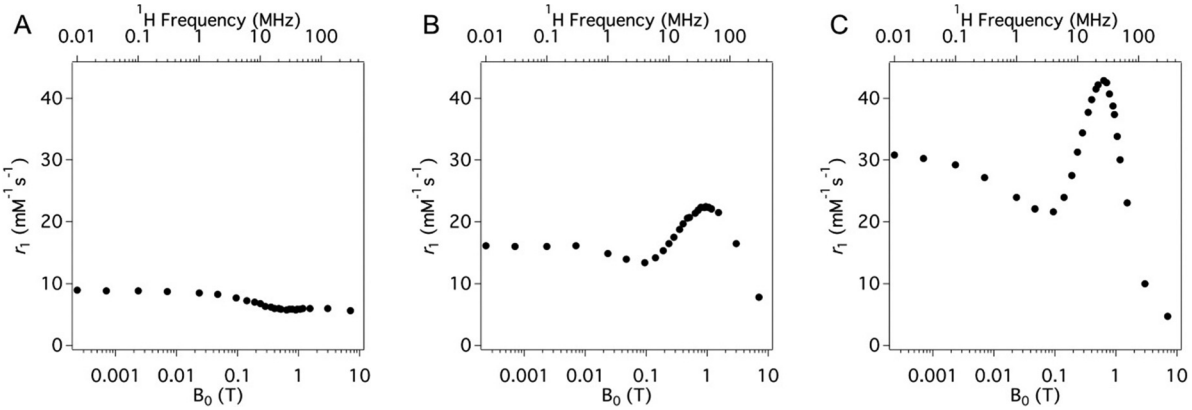


Fig. 24 Experimental $^1$H NMRD profiles showing the dependence of the relaxivity (per $Gd^{3+}$ ion) as a function of the magnetic field strength for Gd-DTPA derivatives with short (A, ~0.1 ns), intermediate (B), and long (C, ~10 ns) rotational correlation times. Reprinted with permission from ref. 212. Copyright 2019 American Chemical Society.

the impact of $\tau_m$ on the relaxivity tends to be negligible for small molecules with fast tumbling, but it can be important for complexes with slow tumbling such as bigger peptides and proteins.[212] Peacock and coworkers have recently demonstrated that a short water residency time is the predominant factor by which $Gd^{3+}$–3SCC complexes achieve high relaxivity.[102,104] In particular, the authors measured an extremely fast water exchange rate ($\tau_m$ = 1.56 ns), comparable to that of free $Gd^{3+}$ in water ($\tau_m$ = 1.02 ns) and much faster relative to that of small GBCAs ($\tau_m \approx$ 240–300 ns, Table 12).[102]

**6.2.4 Outer-sphere relaxivity.** Besides the inner-sphere (IS) relaxivity, second- (2S) and outer-sphere (OS) contributions exist involving hydrogen-bonded water or exchangeable protons in the second coordination sphere, and water molecules that diffuse close to the $Gd^{3+}$ centre, respectively. Remarkably, such OS effects account for about 40% of the overall relaxivity of commercial GBCAs.[212] Moreover, note that the relaxivity of $Gd^{3+}$-complexes with no inner-sphere water molecules ($q$ = 0) basically stems from 2S/OS contributions.

This is the case, for instance, of some $Gd^{3+}$–3SCC scaffolds (*e.g.* MB1-2 and MB1-3) that showed a relaxivity comparable to that of commercial GBCAs at high field (7 T), despite having $q$ = 0 (Table 11). Indeed, the authors hypothesized that this could be mainly related to an OS mechanism involving hydrogen bonding and proton exchange networks between the peptide surface and bulk water.[99]

Delangle and coworkers measured the IS and OS relaxivities of the $Gd^{3+}$–PA complex and indirectly determined the 2S contribution, which accounts for about 25% of the total relaxivity.[41,219] By means of MD simulations and the fit of the experimental data with appropriate models, the authors provided a molecular interpretation of the 2S contribution, identifying three 2S water molecules with a high residence time ($\approx$1 ns).

Important 2S and OS contributions were also evidenced for the protein-based GBCAs reported by Yang and coworkers.[218]

**6.2.5 Discussion and perspectives.** Although an exhaustive comparison among the magnetic properties of different $Ln^{3+}$-binding peptide and protein scaffolds is prevented by the different conditions in which relaxivity data were collected, the highest relaxivity is observed (i) at low fields (0.47–3 T), for large and compact protein scaffolds with long $\tau_R$ ($\sim$10 ns), such as some 3SCCs and CA1.CA2, and (ii) at high fields (7–11.75 T), for small rigid scaffolds with intermediate $\tau_R$ ($\sim$0.5–2 ns) and important 2S/OS contributions, such as PA. This trend, which is consistent with the dependence of relaxivity on $\tau_R$ and the magnetic field strength (Fig. 22), points to the fact that, while large protein scaffolds could be suitable for current clinical (low-field) MRI settings, smaller scaffolds will meet the future need (see above) for GCBAs applicable at higher field. The analysis of the relaxivity data reported indicates that GBCAs based on peptide and protein scaffolds can benefit from higher relaxivity than current clinical GBCAs. This is due to their longer $\tau_R$ (related to higher molecular weight and/or rigidity) as well as important second- and outer-sphere contributions mediated by the peptide surface (*e.g.* hydrogen bonding networks). It is also worth noting that protein-based GBCAs such as ProCA32 and LanND have enhanced molecular relaxivity ($\sim$40–50 $mM^{-1}$ $s^{-1}$ at 7 T), as they bind two and four $Gd^{3+}$ ions, respectively.

However, most peptide and protein scaffolds display a lower affinity for $Gd^{3+}$ ($\log K_{app} \leq \sim$13, see 3. Thermodynamic stability) than currently used GBCAs ($\log {}^{c}K \approx$ 15–19 at pH 7.4[212]) and their kinetic inertness has been poorly explored, challenging their *in vivo* applications. Nevertheless, some protein-based GBCAs, including ProCA32 and LanND-Gd benefit from higher selectivity against physiological metal ions (*e.g.* $Ca^{2+}$, $Cu^{2+}$, $Zn^{2+}$) than small chelators (see 4.4 Selectivity over d-block metal ions), and showed good biocompatibility, resulting as a promising alternative as next-generation GBCAs. Indeed, ProCA32 showed a half-life of $\sim$4 hours in the blood plasma of mice, which increased to $\sim$10 hours upon PEGylation. At 14 days postinjection of ProCA32, a 3- to 10-fold lower amount of $Gd^{3+}$ was detected in different organs compared to the administration of Gd bound to a DTPA derivative.[221] Hence, ProCA32 has been conjugated with different targeting moieties, including PSMA (prostate-specific membrane-associated antigen), CXCR4 (chemokine receptor 4) and a collagen-targeting peptide, to selectively target prostate cancer, liver metastasis or lung fibrosis.[222–226] Similarly to clinical GBCAs, LanND concentration in mice blood dropped below detection level within 1 hour, and renal clearance was observed within 3 hours after LanND injection. Moreover, LanND did not show cellular and neural toxicity and did not elicit an evident immune response. Based on these favourable properties, LanND allowed the high-resolution visualization of brain vessels and the monitoring of kidney dysfunction in mice.[220]

Beyond the research of safer and more sensitive GBCAs, there is a growing interest in the development of responsive GBCAs, which have been reviewed in ref. 227–231. For instance, GBCAs responsive to the pH, molecular targets (cations, anions, metabolites, nucleic


Fig. 25 Overview of the most developed Industrial uses of Ln: rechargeable batteries, permanent magnets, lightings, catalysis and refinery, and medicine. Pm is a radioactive element and does not occur in Nature.

 

acids), biomarkers, or enzyme activity can be designed based on a change of the relaxivity, that is on the variation of hydration number ($q$-based) or the tumbling rate ($\tau_R$-based), or both.

In the early 2000s, the HLH scaffold P3W (see 2.2.3 Chimeric helix-loop-helix motifs) was the first, and currently unique, $Ln^{3+}$-binding peptide investigated as a responsive GBCA. $Gd^{3+}$–P3W showed a 4- to 6-fold higher relaxivity than $Gd^{3+}$–DTPA at low clinical magnetic fields (Table 11), as a result of higher hydration number ($q$ = 2) and slower tumbling rate, whose further decrease upon DNA binding yielded a ~100% relaxivity increase.[77] Lately, Sénèque and coworkers have also developed $Zn^{2+}$-responsive peptides with appended $Gd^{3+}$–DOTA/DO3A derivatives that show a $Zn^{2+}$-dependent relaxivity increase based on a change of the system rigidity and/or hydration number upon $Zn^{2+}$ binding.[232,233] These examples illustrate how responsive peptide/protein-based GBCAs could be designed.

### 6.3 Ln recovery and separation using Ln-binding peptides and proteins

**6.3.1 Challenges in REE separation.** Rare earth elements, which include Ln, Y, and Sc, have become indispensable for medicine (Gd, Lu), and green energies as they are essential elements of permanent magnets and rechargeable batteries (La, Ce, Pr, Nd, Sm, Gd, Dy), lightings (La, Eu, Gd, Tb, Y), and catalysis and refinery (La, Ce) as reviewed in ref. 234 and 235 (Fig. 25). Even Er, Tm and Yb, which are not used on a large scale, find industrial applications. Er is used in silica-glass fibres for communications and in laser for medical applications; Tm is also used in laser for laser surgery, $^{170}$Tm radionuclide is used in portable X-ray sources, and so is $^{169}$Yb.[235]

The REE market is dominated by China, which supplies 98% of REE used in the EU.[2] Due to the expected increasing demand for REE in the coming years and to the lack of high concentrated REE mining sites in Europe, REEs have been identified as critical elements by the EU.[2,235] Thus, there is a strong interest in developing REE recovery technologies from secondary sources, such as the enormous amounts of waste from electrical and electronic equipment generated each year (WEEE 1–30% wt REE).[234] Additionally, sources with low concentrations of REE such as phosphogypsum (0.01–2% wt REE), red mud (500–2500 ppm REE), mine tailings (3–5% wt REE), coal ashes (0.1–1% wt REE), or acid mine drainages (200–900 μg $L^{-1}$ REE) are available in such amounts that the total content of REE they contained could be worth exploiting. The industrial and economic feasibility of the valorisation of such wastes was discussed in ref. 236.

Beyond economic and geopolitical aspects, it is urgent to develop new and more sustainable REE recovery technologies to limit damages caused by mining and reduce the greenhouse gas emissions and the detrimental impact on the environment of actual processes that rely heavily on pyrometallurgy (high energy consumption, release of toxic gases, volatile metals and dust) and hydrometallurgy (use of strong acids such as HCl, $HNO_3$, or $H_2SO_4$; solvents such as kerosene or hexane; large amount of liquid wastes and sludges; risk of heavy metal pollution, including by radionuclides).[3] The separation of REE is an additional challenge, which is performed by multiple separation cycles involving liquid/liquid extraction, or precipitation, and relying on the use of organic ligands.[14,237]

In this context, biometallurgy is emerging as a sustainable alternative that could reduce the energy consumption, waste production, and volumes of strong acids and solvents used, even though it has not yet been translated at the industrial scale.[14,238–243] This approach relies on the use of microorganisms or biomaterials for the (selective) recovery of critical metal ions and is particularly suited for low-grade REE feedstocks with low REE concentrations.

Several bio-inspired strategies have been investigated for the recovery of REE from secondary sources which were discussed in recent reviews: *(i) non-modified microorganisms:* microorganisms were used either for biosorption of REE on their surface, or for their (active) bioaccumulation as reviewed in ref. 239–243. Notably, the Ln-dependent methylotrophic strain *M. fumariolicum* SolV was recently described for its ability to recover REE from various low-grade sources;[244] *(ii) display of REE-binding motifs:* micro-organisms were also modified to display $Ln^{3+}$-binding peptides on their surface in order to improve REE biosorption, which was recently reviewed.[15,242] Examples include LBT displayed on phages,[245,246] or *E. coli*,[154,247–251] and LanM displayed on *Y. lipolytica* and *E. coli*;[252,253] *(iii) Ln-binding peptides and proteins:* $Ln^{3+}$-binding peptides and proteins were also used directly for recycling (Table 13) as reviewed in ref. 15, 241, 242 and 254 and discussed in more detail below.

**6.3.2 REE recovery strategies.** The development of recovery strategies based on Ln-binding peptides and proteins has been done using EF-hand scaffolds, as well as other types of scaffolds.

Joshi and coworkers used LBT to develop filters for REE recovery from dilute acid wash.[255] They genetically engineered a Curli fibre produced in *E. coli* biofilm and known to form amyloid fibres extracellularly, to fuse it with two or four repeats of an LBT sequence. Following a similar strategy, Kaplan and co-workers developed filters based on silk nanofibrils (SNF) functionalized with LBT peptides (Fig. 26), either by chemical coupling or by genetically fusing LBT to a silk-elastin-like protein (SELPS).[256]

Following the discovery of LanM, their potential for the recovery of REE (including $Y^{3+}$ and $Sc^{3+}$) was investigated mostly with *Mex*-LanM, and more recently with *Hans*-LanM (Table 13). The strategies employed relied on size-exclusion filtration, on LanM-functionalised agarose beads,[257] on functionalized magnetic nanoparticles (MNP-LanM, Fig. 27),[260] on polymeric membranes functionalized with LanM-derived peptides,[261] and on liquid–liquid phase separation made by fusing LanM to a thermo-responsive genetically encoded polypeptide RELP.[258]

Cotruvo and coworkers investigated the use of LanD for large Ln-separation.[125] They hypothesized that by combining mutations of LanD to lower the affinity of its monomer for $Ln^{3+}$ (E75Q) and favour the dimer formation by removing repulsive electrostatic interactions (E78A), they could obtain a dimer that preferentially binds $Pr^{3+}$ and $Nd^{3+}$ over $La^{3+}$ and $Ce^{3+}$.

Table 13 Strategies based on Ln-binding peptides and proteins for REE recovery. Details are given regarding the scaffolds used, the pH range in which $REE^{3+}$-binding was performed, the composition and pH of the solutions from which REE were recovered, the sorption/desorption strategies employed, and if intra-REE selectivity was achieved. n.d.: not determined; n/a: not applicable

| Scaffold | pH range | $REE^{3+}$ Competing metal ions (source, pH) | Sorption (S) Desorption (D) | Intra-REE selectivity | Ref. |
|---|---|---|---|---|---|
| LBT | >5 | Ln<br>Al, Ca, Cu, Fe, Ni<br>(mine tailings, pH 7) | S: filtration on functionalised fibres<br>D: $HNO_3$, pH 2 | Yes | 255 |
| | ~4 | Tb<br>Ca or Cu or Fe or Zn<br>— | S: filtration on functionalised fibres<br>D: HCl, pH 1.7 | n.d. | 256 |
| RTX | >0.5 | Nd, Dy<br>Fe, Co<br>(NdFeB magnet solution, pH 6) | S: filtration on 3 kDa cutoff ultrafilter<br>D: n/a | n.d. | 129 |
| LanM | >2.5 | Sc, Y, Ln<br>Li, Be, Na, Mg, Ca, Sr, Al, Si, Mn, Fe, Co, Ni, Cu, Zn, U<br>(lignite leachate, pH 3.7) | S: filtration on 3 kDa cutoff ultrafilter (*Mex*-LanM)<br>D: n/a | no | 155 |
| | >2.4 | Sc, Y, Ln<br>Li, Na, K, Rb, Mg, Ca, Sr, Ba, Al, Si, Pb, V, Mn, Fe, Co, Ni, Cu, Zn, Se, Th, U<br>(coal fly ash leachate, pH 5) | S: filtration on functionalized agarose beads (*Mex*-LanM)<br>D: pH steps (HCl, pH 2.3, then pH 1.5) | Yes | 257 |
| | — | Sc, Y, Ln<br>—<br>(coal combustion, mine tailings, pH 3) | S: filtration on functionalized agarose beads (*Mex*-LanM)<br>D: two-cycle process: 1. Malonate, then pH ramp (2.3-2.1-1.9), 2. Citrate, then pH 1.5 | Yes | 167 |
| | >2 | Y, Ln<br>Mg, Na, Ca<br>(coal fly ash leachate, pH 4.5) | S: liquid–liquid phase separation (*Mex*-LanM)<br>D: phosphate-citrate buffer, pH 2.2 | Yes | 258 and 259 |
| | >3 | Sc, Y, Ln<br>Li, Na, K, Rb, Mg, Ca, Ba, Al, Si, Pb, V, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Cs<br>(coal fly ash leachate, pH 5) | S: functionalised magnetic nanoparticles (*Mex*-LanM)<br>D: HCl, pH 1.7 | n.d. | 260 |
| | — | Dy, Nd<br>—<br>(pH 5) | S: filtration on functionalized agarose beads (*Hans*-LanM, *Hans*-LanM-R100K)<br>D: 1. Malonate, 2. HCl, pH 1.5 | Yes | 62 |
| LanD | n.d. | La, Ce, Pr, Nd<br>—<br>(pH 6) | S: filtration on 10 kDa cutoff ultrafilter (LanD-E75Q/E78A)<br>D: n/a | Yes | 125 |

Still with the aim of developing a method for REE recovery, Banta and coworkers proposed to use the RTX protein, which can bind up to 8 lanthanides at pH 5.5.[129]

**6.3.3 pH sensitivity.** As stressed above (see 3.4 Effect of the pH), the impact of pH on $Ln^{3+}$-binding is generally an overlooked feature. Nonetheless, because applications for recovery require working with acidic REE sources, the pH sensitivity of some of the peptides/proteins used or their derived materials has been investigated.

LBT-based fibres displayed a similar pH-sensitivity than observed in solution.[255] The fibres could bind $Tb^{3+}$ down to pH 5, and their $Tb^{3+}$-sorption capacity was decreased to the level of non-functionalized fibres at pH 3. Park and Cotruvo showed that *Mex*-LanM functionalised agarose beads retained their ability to bind REE at pH 3,[257] which was consistent with *Mex*-LanM capacity to bind $Ln^{3+}$ down to pH 2.5 in solution.[155] RTX is the protein that seems to be the more tolerant to pH as it retains some $Ln^{3+}$-binding capacity at pH < 1.5. However, the proteins precipitate at low pH, which complicated the analysis.

Although the different scaffolds described to date present various pH-sensitivity, most of them were assayed for their ability to bind Ln on slightly acidic solutions (pH 4–7, Table 12) corresponding to REE sources obtained from mine tailings, or leachates. Desorption of REE was only investigated for LBT- and LanM-based materials and could be achieved with acidic solutions (HCl, $HNO_3$) at pH < 2. This is a consequence of the protonation of $Ln^{3+}$-binding amino acids (*e.g.* Asp, Glu), which results in a lower affinity for $Ln^{3+}$.

 

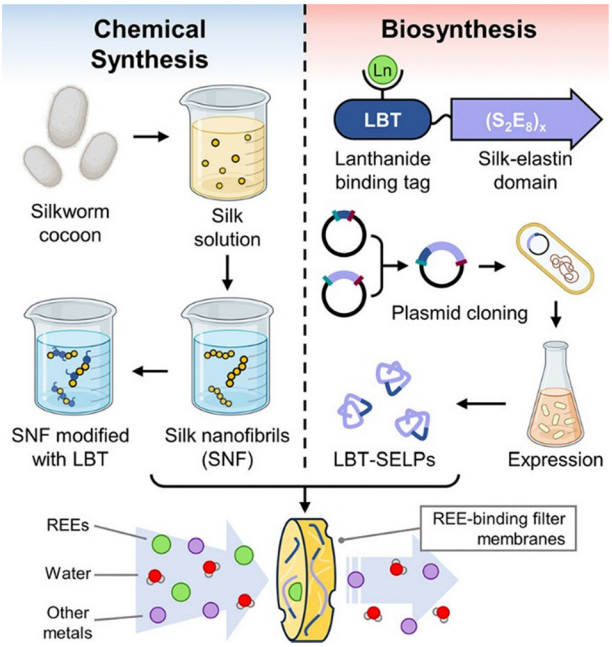


Fig. 26 Overview of the two approaches used to fabricate biobased filter membranes. Reprinted with permission from ref. 256. Copyright 2024 American Chemical Society.

### 6.3.4 Selectivity

*REE vs. non-REE.* Working with low-grade, secondary sources of REE that have a low content in REE (WEEE, phosphogypsum, bauxite, mine tailings, coal ashes, mine drainages) requires a good selectivity for REE *vs.* non-REE in order to minimize the purification steps. Usual competing metal ions consist notably of alkaline-earth (Ca, Mg), and d-block (Mn, Fe, Co, Ni, Cu, Zn) elements (Table 12). The ability of the different scaffolds or derived materials to retain selectively REE *vs.* non-REE has been assessed in most studies.

For the LBT-based fibres described by Joshi and coworkers,[255] REE sorption fell in the presence of 10 to 100 times more concentrated non-REE metal ions (Al, Cu, Fe, Ni), whereas Ca only had a small impact. Similarly, Kaplan and coworkers reported that the $Tb^{3+}$-binding capacity of silk nanofibrils functionalised with LBT was decreased in the presence of 10 to 100 times more concentrated Cu or Fe, but not by Ca and Zn.[256] LBT has a selectivity against $Ca^{2+}$ of $\Delta\log K^{Ca} \approx 2$–3 (see 4. Selectivity), which appears sufficient to compensate for the difference in concentrations. However, from the observations made in these two studies, it is anticipated that LBT selectivity against d-block elements is quite low.

LanMs display a very high selectivity against Ca ($\Delta\log K^{Ca} \approx 8$ for *Mex*-LanM,[81,155] $\Delta\log K_a^{Ca} \approx 4$–6 for *Hans*-LanM[62]) and d-block elements ($\Delta\log K^{Cu} \approx 6$ and $\Delta\log K^{Zn} \approx 8$ for *Mex*-LanM[155]) (see 4. Selectivity). This can explain the impressive ability of LanM to selectively recover REE among other metal ions in complex solutions (Table 12) with high recovery yields and purity when using size-exclusion separation,[155] functionalised agarose beads,[257] or liquid–liquid phase separation.[258,259] Moreover, Gao, Wei and coworkers demonstrated that even with concentrations of non-REE 10 to 1000 times higher than of REE, LanM-functionalized magnetic nanoparticles could selectively bind REE over competing metal ions and achieve a REE purity of 31% wt (compared to the 0.03% wt of the initial solution) corresponding to a 967-fold improvement.[260]

RTX is interesting due to its ability to bind several Ln per protein over a broad pH range.[129] However, one limitation to its use is the precipitation of the protein at low pH. To overcome precipitation issues, Banta and coworkers coupled the RTX protein to a maltose-binding protein (MBP). Working at pH 6, they showed that RTX-MBP could selectively recover up to ~85% of $Nd^{3+}$ and $Dy^{3+}$ in a solution containing also $Fe^{2+}$ and $Co^{2+}$, which suggests a good selectivity of this modified protein for Ln over d-block metal ions.

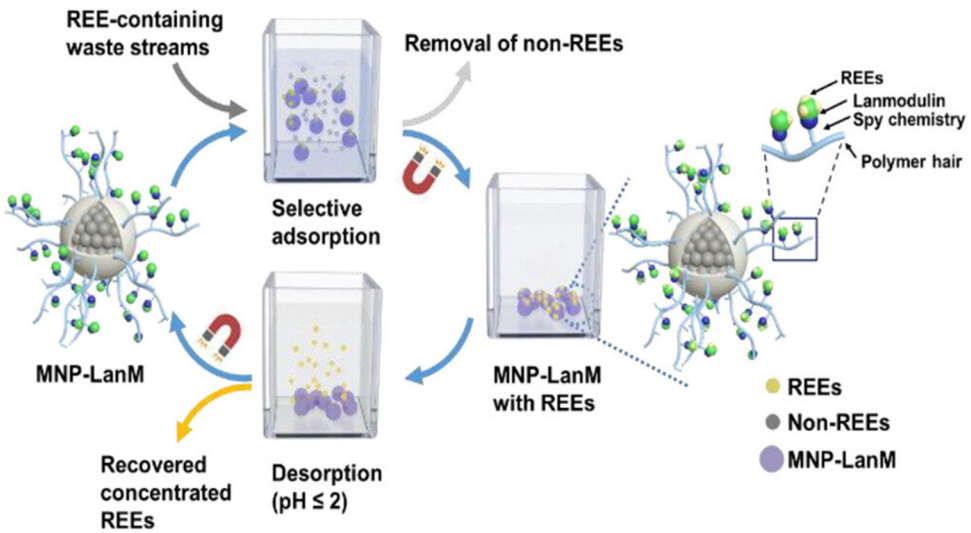


Fig. 27 Use of LanM-functionalized magnetic nanoparticles for REE recovery.[258,259] Reprinted with permission from ref. 260. Copyright 2023 American Chemical Society.

*Intra-REE.* The design of Ln-binding peptides and proteins with good intra-REE selectivity would be highly appealing with the aim to reduce the number of steps required to obtain pure and isolated REE, but it is an outstanding challenge given the similarities of REE (see 1. Introduction) and the moderate intra-REE selectivity displayed so far by Ln-binding peptides and proteins (see 4. Selectivity). The LBT-fibers reported by Joshi and coworkers showed a moderate preference for the binding of $Eu^{3+}$, $Tb^{3+}$, and $Dy^{3+}$ relative to other $Ln^{3+}$.[255] Cotruvo and coworkers noted an intra-REE selectivity between heavy (HREE: $Tb^{3+}$–$Lu^{3+}$, $Y^{3+}$) and light (LREE: $La^{3+}$–$Gd^{3+}$) when using LanM-functionalised on agarose beads similar to the one observed in solution.[257]

To overcome the moderate intra-REE selectivity of Ln-binding peptides and proteins, some studies have proposed innovative strategies. The separation process based on LanM-functionalised on agarose beads was optimised by using mild competitors for Ln desorption (malonate, citrate).[167] A two-cycle process was proposed in order to recover five distinct fractions: $Sc^{3+}$, $Y^{3+}$, HREE ($Gd^{3+}$–$Lu^{3+}$), middle REE (MREE: $Pr^{3+}$–$Eu^{3+}$), and LREE ($La^{3+}$, $Ce^{3+}$). A solution reproducing metal ions concentrations found in coal combustion products or mine tailings was adsorbed onto the column. The first desorption cycle used malonate in order to desorb up to 99% of Sc with >99% purity. Next, pH desorption was performed by decreasing the pH successively to 2.3, 2.1 and 1.9, which enabled a first separation of REE into HREE and $Y^{3+}$, LREE, and MREE. These fractions were further purified using citrate desorption, which led to the recovery of HREE with 88–100% purity, >95% recovery of $Y^{3+}$ with 92% purity, recovery of MREE with ~80% purity, and recovery of LREE with 98% purity.

As an alternative strategy, Cotruvo and coworkers investigated how intra-REE selectivity could be improved by controlling the dimerization state of *Hans*-LanM to separate $Dy^{3+}$ from $Nd^{3+}$ working with a synthetic leachate solution of NdFeB magnets.[62] With *Hans*-LanM functionalised agarose beads, $Dy^{3+}$ could be desorbed from the column with malonate and obtained with 83% purity without significant desorption of $Nd^{3+}$, which was then desorbed at pH 1.5 and obtained at 99.8% purity. In a similar approach, the LanD-E75Q/E78A mutant optimised to form dimer with medium REE, was able to retain $Pr^{3+}$ and $Nd^{3+}$ better than the larger $Ce^{3+}$ and $La^{3+}$.

**6.3.5 Concluding note.** The development of biomaterials based on Ln-binding peptides and proteins is well suited to exploit secondary REE sources, which are more dilute and of lower grade. As such, it represents a complementary approach to the development of chemical strategies that aim to improve the extraction and separation processes used on primary REE sources.[14,237]

REE recovery using Ln-binding peptides and proteins requires to meet several criteria. The first is the selectivity of the peptides and proteins for REE *vs.* non-REE, and ideally also for intra-REE separation. The discovery of LanM, which has a high selectivity *vs.* non-REE has represented a breakthrough compared to LBT-based strategies that are more sensitive to high concentrations of d-block metal ions. Nonetheless, Ln-binding peptides and proteins still present a moderate intra-REE selectivity

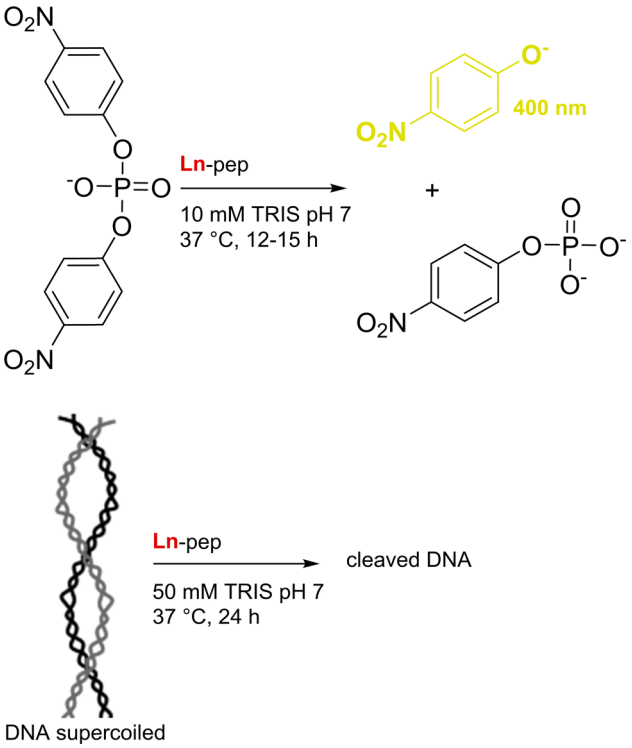


Fig. 28 Phosphate ester and DNA cleavage activities of Ln-pep (Ln = $Eu^{3+}$, $Ce^{4+}$).

(see 4. Selectivity) that will need to be improved. The second criterion is to improve the pH range in which the scaffold is stable and retains Ln-binding capacity. Last, different strategies were proposed for the sorption and recovery of REE. Scaling up to meet industrial needs will require a facile and low-cost production of the scaffold and its derived material. The strategy employed for functionalizing magnetic nanoparticles with LanM using Spy chemistry is interesting for this purpose, as it enabled the selective coupling of the protein directly from cell lysates, thus avoiding protein purification steps.[262]

Importantly, despite the considerable efforts made to develop innovative, more sustainable technologies to exploit secondary REE sources, the main challenge to their industrial implementation remains their economic viability. The economic feasibility of the biosorption process for the recovery of REE from different low-grade sources was evaluated and strongly depends on REE market prices. Only an increase of REE prices back to their level of 2011 would make actual biosorption processes profitable.[263] The fluctuating demands in REE and its impact on the development at the industrial scale of more environmentally friendly technologies have been discussed in ref. 235.

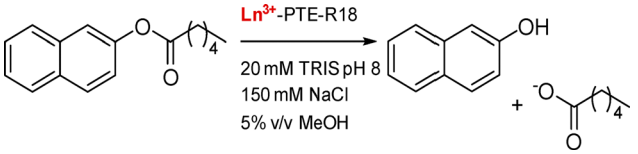


Fig. 29 Hydrolysis of 2-naphthyl hexanoate by Ln-PTE-R18.

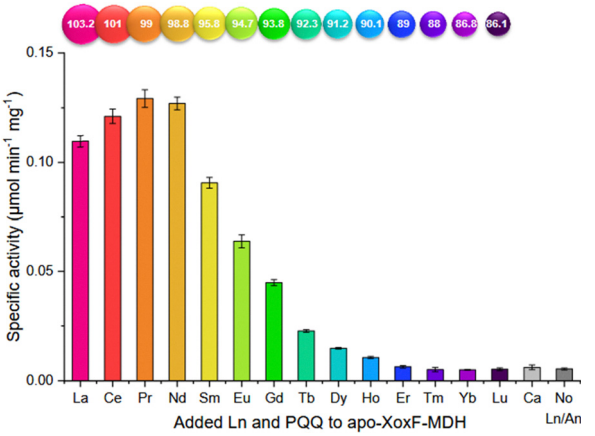


Fig. 30 Ln-dependent activity of XoxF-MDH. The enzyme was reconstituted from the apo-enzyme by incubation with PQQ and metal ion for 72 h at 4 °C. Reproduced from ref. 176 (Creative Commons Attribution Non-Commercial License 2023).

### 6.4 Catalysis

The field of Ln–proteins and Ln–peptides catalysis is emerging, and a few examples have been published. $Ln^{3+}$ Lewis acidity has been exploited for hydrolysis reactions and in alcohol dehydrogenase (ADH)-like reactions (Section 6.2.1). $Ln^{3+}$ photoredox properties have been exploited for the homolytic cleavage and oxidation of diols (Section 6.2.2).

**6.4.1 Lewis acidity.** Before the discovery of Ln–enzymes, Franklin and coworkers used $Ln^{3+}$ Lewis acidic properties for the design of Ln–peptides as chimeric nucleases.[17,18] Helix-turn-helix peptides (HTH) combined an $Ln^{3+}$-binding loop and helices from engrailed homeodomain, the latter offering the possibility to recognize a specific DNA sequence. The resulting Ln–peptides (Ln = $Eu^{3+}$, $Ce^{4+}$) displayed DNA- and phosphate esters cleavage activities (Fig. 28).[264,265]

The $Ce^{4+}$–peptides were more active than their $Eu^{3+}$ counterparts, consistent with the increased Lewis acidity and nucleophilicity of the $Ln^{3+}$-bound water molecules.[16] The effect of the pH on Ln–peptides reactivity was evaluated in the range 6.9–9.0.[266] The rate increased at pH 8.2, which is close to the p$K_a$ of $Ln^{3+}$-bound water molecule (Fig. 1), indicating that it is involved in DNA hydrolytic cleavage. In addition, the cleavage catalysed by Ln–peptides was site-selective, suggesting that DNA sequences could be recognized and cleaved preferentially.[16]

$Ln^{3+}$ Lewis acidic properties were also used for the catalysis of the hydrolysis of 2-naphthyl hexanoate (Fig. 29). Jackson and coworkers used the PTE-R18 protein, which has a $Zn^{2+}$-binding site, and showed that the demetallated protein could bind $Ln^{3+}$.[21] The activity of the resulting Ln–protein was pH-dependent, which was linked to the deprotonation of a Ln-bound water molecule.

The first and sole known Ln–enzymes, Ln-alcohol dehydrogenases (Ln-ADH), are an example of how the fine-tuning of $Ln^{3+}$ Lewis acid properties along the series may impact reactivity. In the enzyme, $Ln^{3+}$ is essential for catalysis because it tunes the redox potential of the redox cofactor, the pyrroloquinoline quinone (PQQ). Interestingly, the activity of the XoxF-methanol dehydrogenase (MDH) is strongly dependent on the $Ln^{3+}$ in the active site, the lighter and less Lewis acidic $Ln^{3+}$ giving the higher activity while no activity is detected for $Ln^{3+}$ heavier than $Er^{3+}$ (Fig. 30).[176] A similar trend has been also recently observed for an Ln-ADH from *Pseudomonas putida* KT2440 (PedH).[267] Several hypotheses have been proposed and are still under investigation to explain this dramatic impact on Ln-ADH reactivity. They consider not only PQQ redox properties, but also changes in the coordination number, or exchange rate constants of $Ln^{3+}$-bound substrate/product.

Working towards bringing new insights into Ln-MDH mechanism, Olshansky and coworkers built a structural mimic of the enzyme.[20] Starting from the PQQ-binding protein PqqT, they engineered the binding site with a single mutation ($K_{142}D$) in order to introduce a $Ln^{3+}$-binding site. They tested the ability of this structural mimic of Ln-MDH to oxidize benzyl alcohol, using a chemical trap (*O*-((perfluorophenyl)methyl)hydroxylamine (PFBHA)) to capture the benzaldehyde formed and identify it by LC-MS (Fig. 31). Only the $La^{3+}$–PQQ⊂$K_{142}$D–PqqT mutant was able to perform this reaction, and not the WT protein or $K_{142}$A mutant both lacking an $Ln^{3+}$-binding site, nor the mutants lacking $La^{3+}$, or PQQ.

The ability of ADH to oxidize alcohols has been exploited further by Bange, Klebensberger and coworkers for foreseen application in bio-based plastic production.[116] Starting from the Pr-PedH from

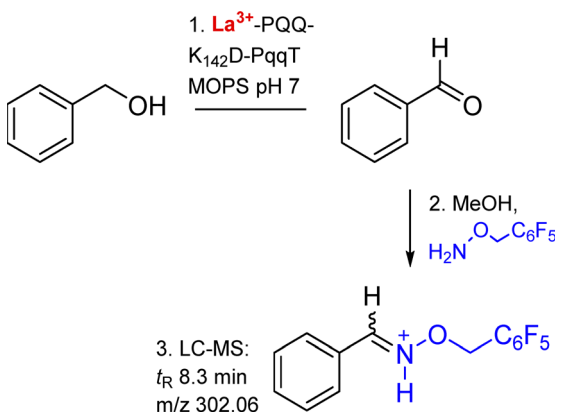


Fig. 31 Benzyl alcohol oxidation by $La^{3+}$-bound PQQ⊂$K_{142}$D–PqqT.

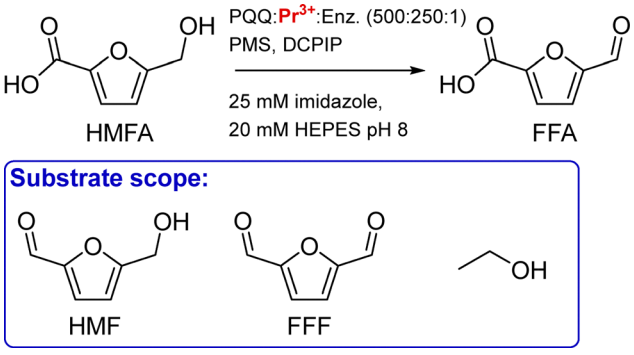


Fig. 32 Oxidation of 5-(hydroxymethyl)furoic acid (HMFA) and derivatives by mutants of PedH. FFA: 5-formylfuroic acid; HMF: 5-(hydroxymethyl)-furfural; FFF: 5-formylfurfural; PMS: phenazine methosulfate; DCPIP: 2,6-dichlorophenol indophenol.

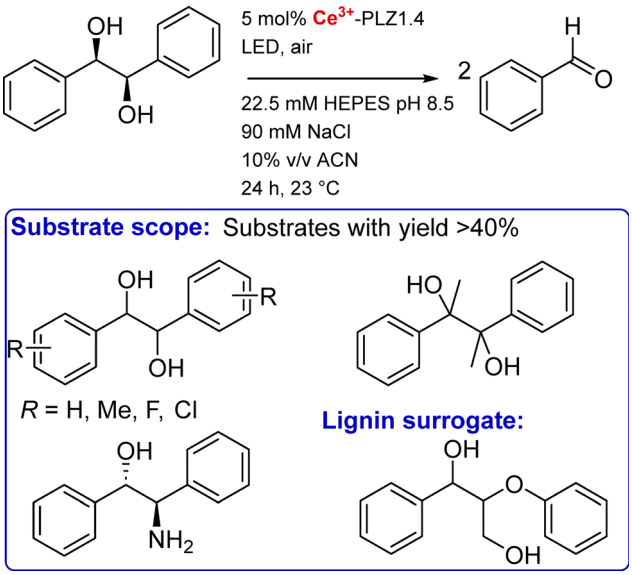


Fig. 33 Photoredox catalysis of aromatic diols homolytic cleavage by Ce-artificial enzymes.

*P. putida* KT2440, they constructed libraries of mutants of the enzyme expressed in *E. coli* and screened each colony for *in vivo* enzymatic activity towards the oxidation of HMFA (Fig. 32). This compound is an important intermediate for the obtention of furan-2,5-dicarboxylic acid (FDCA), a synthon that could serve as a plant-based alternative in the production of polymers. Out of 30 000 clones, 3 mutants were identified and their activity was confirmed by testing the purified enzymes. They shared mutations of the same residues, F412 and W561, positioned in the substrate channel and on a lid loop, respectively. These mutations enable bulkier substrates, such as HMFA, to access the active site. The mutant with the best catalytic activity for HMFA oxidation, PedH$^{F412V/W561A}$, could also successfully oxidize ethanol, HMF and FFF (Fig. 32), which opens up the route to the whole-cell biocatalytic production of FDCA.

**6.4.2 Photoredox properties.** So far, only a single example reported on the use of Ln photoredox properties in artificial enzymes for catalysis. Zeymer and co-workers designed an Ln-binding protein based on a TIM-barrel scaffold with the aim of exploiting the photocatalytic properties of Ce.[19]

The scaffold TFD-EE N6W was the starting point for the following catalysis study (Fig. 33).[19] Changes were made to the scaffold in order to introduce asymmetry in the binding site (single chain variant), to limit photodamage (removal of Trp residues), and to prevent non-specific metal binding to the protein surface (removal of clusters of negatively charged residues by mutating Asp to Asn and Glu to Gln), leading to the optimised scaffold PLZ1.4.

Upon irradiation, the artificial enzyme Ce-PLZ1.4 was able to catalyse the homolytic cleavage of homobenzoin into benzaldehyde (Fig. 33). Compared to molecular catalysts, the interest of the protein scaffold is to be able to control the stereoselectivity of the reaction. The authors showed that Ce-PLZ1.4 catalysed better the cleavage of (*R*,*R*)- and (*S*,*S*)-hydrobenzoin compared to the *meso*-isomer, and so displayed some diastereoselectivity. Investigation of the substrate scope of the enzyme also evidenced chemoselectivity, with aromatic diols being better cleaved than aliphatic ones. Last, the PLZ1.1 protein was expressed at the surface of *E. coli* cells, which were subsequently treated with $CeCl_3$ and light and were able to cleave a lignin surrogate (2-phenoxy-1-phenylpropane-1,3-diol) opening the way to photocontrolled whole-cell catalysis.

**6.4.3 Concluding note.** Lanthanides have versatile properties ranging from their Lewis acidity which offers a fine-tuning of the reactivity, to their redox properties that cover a broad range of potentials, and to their photochemical properties, which open the way to photocatalytic reactions. The similarities of Ln allow, with a single ligand, to modulate the reactivity of the catalyst simply by choosing the adequate $Ln^{3+}$ ion, which is not possible with any other family of metal ions.

To date, the scope of Ln–peptides and Ln–proteins catalysts has been limited mainly to reactions exploiting $Ln^{3+}$ Lewis acidity properties, with the exception of an original work demonstrating that photoredox catalysis is also achievable.[19] In addition, the peptide scaffold of the catalysts has been used to select the substrate, whether by using DNA recognition sequence, or by engineering the catalytic pocket,[116] or by favouring one stereoisomer over others.[19]

In the medium term, recent advances obtained with molecular catalysts should be transposable to Ln artificial enzymes. Ln-complexes have been used as Lewis acids catalysts,[268,269] in particular for C–C bond formation in aqueous media.[268] Ln-complexes are also powerful photocatalysts, that can be used for photo-reduction or photo-oxidation.[270–274] Notably, the reduction of $CO_2$, $N_2$ and $NO_x$ was achieved in mild aqueous conditions.[273,274] Similar reactions could be implemented in artificial Ln–enzymes, which operate in aqueous media and whose peptide scaffold allows modulation of the hydrophilicity of the active site. Combining the chemical properties of Ln with the modularity of the peptide scaffold will help control the reactivity of the metal centre, the selectivity of the reaction and fine-tuning of second sphere interactions, and is expected to broaden the reactivities accessible with artificial enzyme.[275]

## 7. Conclusions & perspectives

In the last decade, significant progress has been made in the study of natural Ln-dependent proteins as well as in the design and engineering of Ln-binding peptides and proteins for a variety of relevant applications. Notwithstanding, numerous challenges remain to be addressed for the advancement of this flourishing field, starting with a sound and quantitative understanding of Ln coordination by pH buffers to refine the accuracy of available data and allow reliable comparisons.

The current limitations in designing Ln-binding systems with high affinity, selectivity, and inertness underscore the complexity of controlling key factors such as the coordination sphere (*i.e.* coordination and hydration number), scaffold stability, and the influence of long-range effects.

The control over the hydration number $q$ is well understood and achievable for EF-hand motifs, and it can be tuned for other scaffolds by adjusting the denticity and positioning of the binding site within a hydrophobic environment. However, certain long-range outer-sphere factors, such as the local

conformation of the main chain and the orientation of amino acids can influence $q$ in ways that are not yet fully controllable.

The majority of peptide and protein scaffolds exhibit an Ln-binding affinity in the micro- to nanomolar range ($\log K_{app} \sim 6$–$9$), with only a limited number extending to the picomolar range ($\log K_{app} \sim 12$), which is significantly lower than the affinity achievable with small chelators. Similarly to the tuning of the hydration number $q$, long-range outer-sphere factors, which are not yet understood, strongly influence the Ln affinity, as in the case of LanM compared to its isolated EF-hand motifs. It is worth noting that this gap between the Ln-affinity of peptide/proteins and small organic chelators is not observed for d-block metal ions, for which proteins and chelators of similar affinity, up to $\log K_{app} \sim 21$, are known. The underlying reason for this remains to be elucidated. One possible explanation is that, considering that we are at the dawn of lanthanome research, the tightest Ln-binding proteins have yet to be discovered. Besides, due to the ionic nature of Ln bonding and their preference for higher CN than d-block metal ions, peptides and proteins might need more demanding constraints to bind $Ln^{3+}$ as efficiently as d-block metal ions.

Achieving higher affinities requires careful optimisation of both the Ln-binding site and scaffold stability by fine-tuning key parameters such as the type and denticity of the coordinating amino acids, as well as local and long-range structural constraints in the scaffold. To this end, computational design appears as a powerful tool to obtain a well-structured and stable protein scaffold, as is the case for the Ln-binding TIM barrel. In addition, the pH-dependency of Ln-binding by peptides and proteins has only been seldom looked at but is an important parameter for applications in REE recovery.

The selectivity relative to $Ca^{2+}$ is generally favourable due to the higher coordination number and higher charge of $Ln^{3+}$. Nevertheless, the remarkable selectivity exhibited by LanM remains to be fully elucidated. Again, this is likely attributable to the influence of outer sphere effects.

With respect to d-block elements, the selectivity of peptides and proteins for $Ln^{3+}$ often surpasses that of small chelators, a phenomenon attributable to distinct preferences for ligand types, coordination number and geometry.

In comparison to $Ln^{3+}$, peptide and protein scaffolds demonstrate a slightly higher affinity for $An^{3+}$, presumably due to enhanced covalency. However, further investigation is warranted to fully elucidate the underlying mechanisms and to boost the selectivity between $Ln^{3+}$ and $An^{3+}$. Since $An^{3+}$ can be (radioactive) contaminants of common $Ln^{3+}$ sources, this will be crucial in the context of Ln recovery and radioactive decontamination.

The intra-Ln selectivity remains an outstanding challenge due to their similar ionic radii and Lewis acidity. The intra-Ln selectivity exhibited by peptide and protein scaffolds is not yet comparable to that observed for small chelators. This discrepancy may be attributed to the greater flexibility of peptides, which consequently enables a higher degree of adaptability to $Ln^{3+}$ ions differences in ionic radius and coordination number. Further studies are required to ascertain whether size discrimination can be enhanced based on both electronic and steric considerations, for instance by modulating the rigidity of the coordination site and its position (buried or surface-exposed) within the scaffold.

The kinetic inertness of Ln coordinated to peptides and proteins is a largely neglected aspect that merits greater attention and research focus in the upcoming years, especially with respect to the development of efficient imaging agents and tools for Ln separation. In order to understand how to control the kinetic inertness of Ln–peptide and protein complexes, it is also essential to accumulate a substantial body of experimental data, which is currently lacking.

Overall, it appears that the elusive contributions beyond the first coordination sphere are of paramount importance in tuning properties like affinity, selectivity, and inertness, and must be studied further to achieve optimized designs. It is expected that emerging technologies based on machine learning and artificial intelligence will soon be beneficial in unravelling the influence of such long-range interactions, provided that solid experimental knowledge is established.[276–278] In this context, significant contributions are also expected from approaches beyond rational design, such as directed evolution.[279]

By deepening our understanding of these long-range effects and their impact on the coordination environment, we will be able to address the key challenges in the field: (i) developing high-affinity, intra-Ln selective peptides and proteins for enhanced Ln-recovery; (ii) achieving kinetically inert Ln-binding scaffolds for medical imaging and therapy; and (iii) optimizing the design of Ln-based (photo)catalysts. These advancements hold the potential to revolutionize a wide range of applications, from environmental sustainability to cutting-edge medical technologies.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this review.

## Appendices

Table 14 Values from Fig. 1: trends of key parameters along the Ln series

| $Ln^{3+}$ | $E^{3+/2+}$ (V)[26] | $E^{4+/3+}$ (V)[26] | $r$ (Å)[27] | $pK_a$[28] | $\log K_{sp}$[29] |
|---|---|---|---|---|---|
| $La^{3+}$ | −3.1 | | 1.216 | 8.89 | −22.3 |
| $Ce^{3+}$ | −3.2 | 1.7 | 1.196 | 8.31 | −23.5 |
| $Pr^{3+}$ | −2.7 | 3.4 | 1.179 | 8.3 | −23.4 |
| $Nd^{3+}$ | −2.6 | 4.6 | 1.163 | 8.13 | −24.1 |
| $Pm^{3+}$ | −2.6 | 4.9 | 1.144 | n/a | n/a |
| $Sm^{3+}$ | −1.55 | 5.2 | 1.132 | 7.84 | −24.8 |
| $Eu^{3+}$ | −0.34 | 6.4 | 1.12 | 7.66 | −25.5 |
| $Gd^{3+}$ | −3.9 | 7.9 | 1.107 | 7.87 | −24.8 |
| $Tb^{3+}$ | −3.7 | 3.3 | 1.095 | 7.6 | −25.7 |
| $Dy^{3+}$ | −2.5 | 5 | 1.083 | 7.53 | −25.7 |
| $Ho^{3+}$ | −2.9 | 6.2 | 1.072 | 7.43 | −26.4 |
| $Er^{3+}$ | −3.1 | 6.1 | 1.062 | 7.46 | −26.2 |
| $Tm^{3+}$ | −2.3 | 6.1 | 1.052 | 7.34 | −26.4 |
| $Yb^{3+}$ | −1.05 | 7.1 | 1.042 | 7.31 | −26.7 |
| $Lu^{3+}$ | | 8.5 | 1.032 | 7.33 | −26.6 |

Table 15 Values from Fig. 17: selectivity of different ligands across the $Ln^{3+}$ series

| $Ln^{3+}$ | EDTA[a] | DTPA[a] | Macropa[b] | Macrotripa[c] | LBT12[d] | LF4[e] | RTX[f] | LanD[g] | LanM[h] | LanM[i] | LanM[j] | LanM[k] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $La^{3+}$ | 15.5 | 19.5 | 15.0 | 12.6 | 5.5 | 4.2 | 4.1 | 5.7 | 11.3 | | 10.2 | |
| $Ce^{3+}$ | 16.1 | 20.5 | 15.1 | 12.8 | 6.0 | | | 6.0 | | | | |
| $Pr^{3+}$ | 16.3 | 21.1 | 14.7 | 12.7 | | | 4.5 | 6.3 | | 10.2 | | 11.0 |
| $Nd^{3+}$ | 16.6 | 21.6 | 14.4 | 12.3 | 6.6 | | 4.5 | 6.6 | 11.3 | | 10.0 | 10.7 |
| $Sm^{3+}$ | 17.0 | 22.3 | 13.8 | 11.4 | | | | 6.5 | 11.2 | | | 10.6 |
| $Eu^{3+}$ | 17.2 | 22.4 | 13.0 | 10.9 | 7.2 | 4.6 | 4.6 | 6.5 | | | | |
| $Gd^{3+}$ | 17.3 | 22.5 | 13.0 | 10.2 | 7.1 | 4.6 | | 6.2 | 11.0 | 10.0 | | |
| $Tb^{3+}$ | 17.8 | 22.7 | 11.8 | 10.2 | 7.2 | 4.9 | 4.6 | | 10.7 | | | |
| $Dy^{3+}$ | 18.2 | 22.8 | 11.7 | 10.5 | 7.2 | | 4.6 | | | 9.7 | 8.6 | 10.8 |
| $Ho^{3+}$ | 18.4 | 22.8 | 10.6 | 10.7 | | 4.9 | | | 10.6 | 9.6 | | 11.1 |
| $Er^{3+}$ | 18.9 | 22.7 | 10.1 | 11.1 | 7.1 | 5.0 | | | | | | 11.9 |
| $Tm^{3+}$ | 19.6 | 22.7 | 9.6 | 11.4 | | | | | | | | 12.4 |
| $Yb^{3+}$ | 19.7 | 22.6 | 8.9 | 11.8 | 7.0 | 4.4 | 4.6 | | | | | |
| $Lu^{3+}$ | 20.0 | 22.4 | 8.3 | 11.9 | 6.9 | 4.2 | 4.6 | | | | | |

[a] $\log \beta_{ML}$ from ref. 173. [b] $\log \beta_{ML}$ from ref. 160. [c] $\log \beta_{ML}$ from ref. 161. [d] $\log K_{app}$ in HEPES 10 mM pH 7, NaCl 100 mM, ref. 64. [e] $\log K_{app}$ in HEPES 10 mM pH 7.5, NaCl 10 mM, ref. 103. [f] $\log K_{app}$ in bis-Tris 20 mM pH 6, NaCl 25 mM, ref. 129. [g] $\log K_{app}$ in MOPS 30 mM pH 7, KCl 100 mM, ref. —. [h] $\log K_{app}$ for *Mex*-LanM in MOPS 30 mM pH 7.2, KCl 100 mM, measured by CD spectroscopy, ref. 81. [i] $\log K_{app}$ for *Mex*-LanM in acetate buffer 20 mM pH 5, KCl 100 mM, measured by CD spectroscopy, ref. 155. [j] $\log K_{app}$ for *Hans*-LanM acetate buffer 20 mM pH 5, KCl 100 mM, measured by CD spectroscopy, ref. 280. [k] $\log K_{app}$ for *Mex*-LanM KCl/$KCH_3COO^-$ buffer 100 mM pH 5.0, measured by UV-vis spectroscopy, ref. 155.

Table 16 Values from Fig. 19: selectivity of small chelators and some peptide/protein scaffolds for $Ln^{3+}$ over $Ca^{2+}$

| $Ln^{3+}$ | EDTA[a] | DTPA[a] | P3W[b] | EF loop[c] | CaM[d] | *Hans*-LanM[e] | *Mex*-LanM[f] | RTX[g] | PqqT[h] | PA[i] | 2SCC[j] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $La^{3+}$ | 15.5 | 19.5 | 5.20 | | | 10.2 | 11.3 | 4.1 | 6.2 | | 6.2 |
| $Ce^{3+}$ | 16.1 | 20.5 | | | | | | | | | |
| $Pr^{3+}$ | 16.3 | 21.1 | | | | | | 4.5 | | | |
| $Nd^{3+}$ | 16.6 | 21.6 | | | | 10.0 | 11.3 | 4.5 | | | |
| $Sm^{3+}$ | 17.0 | 22.3 | | | | | 11.2 | | | | |
| $Eu^{3+}$ | 17.2 | 22.4 | 5.2 | 6.8 | 8.00 | | | 4.6 | | 6.8 | |
| $Gd^{3+}$ | 17.3 | 22.5 | 5.2 | | | | 11.0 | | | | |
| $Tb^{3+}$ | 17.8 | 22.7 | | 6.8 | 8.8 | | 10.7 | 4.6 | | 6.3 | |
| $Dy^{3+}$ | 18.2 | 22.8 | | | | 8.6 | | 4.6 | | | |
| $Ho^{3+}$ | 18.4 | 22.8 | | | | | 10.6 | | | | |
| $Er^{3+}$ | 18.9 | 22.7 | | | | | | | | | |
| $Tm^{3+}$ | 19.6 | 22.7 | | | | | | | | | |
| $Yb^{3+}$ | 19.7 | 22.6 | | | | | | 4.6 | | | 6.4 |
| $Lu^{3+}$ | 20.0 | 22.4 | | | | | | 4.6 | | | |
| $Ca^{2+}$ | 11.0 | 10.7 | 4.5 | 4.6 | 6 | 4.2 | 3.2 | 3.3 | 3.8 | 3.7 | 1.7 |

[a] $\log \beta_{ML}$ from ref. 173. [b] $\log K_{app}$ in TRIS 5 mM pH 7.8, from ref. 76. [c] $\log K_{app}$ in HEPES 10 mM pH 7, NaCl 100 mM, from ref. 169. [d] $\log K_{app}$ in PIPES 25 mM pH 6.8, KCI 100 mM, from ref. 82. [e] $\log K_{app}$ for *Hans*-LanM acetate buffer 20 mM pH 5, KCl 100 mM, measured by CD spectroscopy, ref. 280. [f] $\log K_{app}$ for *Mex*-LanM in MOPS 30 mM pH 7.2, KCl 100 mM, measured by CD spectroscopy, ref. 81. [g] $\log K_{app}$ in bis-Tris 20 mM pH 6, NaCl 25 mM, ref. 129. [h] $\log K_{app}$ in HEPES 30 mM pH 7, NaCl 100 mM, from ref. 20. [i] $\log K_{app}$ in HEPES 10 mM pH 7, from ref. 41. [j] $\log K_{app}$ in Imidazole 30 mM pH 6.9, KCl 50 mM, from ref. 95.

## Acknowledgements

The authors acknowledge financial support from the French National Research Agency (ANR) through the ANR-23-CE07-0002 (LnZYME) and ANR-22-CE29-0026 (DREAMY) projects.

## References

1 S. V. Eliseeva and J.-C. G. Bünzli, *New J. Chem.*, 2011, **35**, 1165–1176.
2 Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (European Commission), M. Grohol and C. Veeh, *Study on the critical raw materials for the EU 2023: final report*, Publications Office of the European Union, 2023.
3 V. Balaram, *Geosci. Front.*, 2019, **10**, 1285–1303.
4 N. A. Fitriyanto, M. Fushimi, M. Matsunaga, A. Pertiwiningrum, T. Iwama and K. Kawai, *J. Biosci. Bioeng.*, 2011, **111**, 613–617.
5 Y. Hibi, K. Asai, H. Arafuka, M. Hamajima, T. Iwama and K. Kawai, *J. Biosci. Bioeng.*, 2011, **111**, 547–549.
6 T. Nakagawa, R. Mitsui, A. Tani, K. Sasa, S. Tashiro, T. Iwama, T. Hayakawa and K. Kawai, *PLoS One*, 2012, **7**, e50480.
7 A. Pol, T. R. M. Barends, A. Dietl, A. F. Khadem, J. Eygensteyn, M. S. M. Jetten and H. J. M. O. den Camp, *Environ. Microbiol.*, 2014, **16**, 255–264.
8 L. Ancel, A. Niedźwiecka, C. Lebrun, C. Gateau and P. Delangle, *C. R. Chim.*, 2013, **16**, 515–523.
9 E. Skovran, C. Raghuraman and N. C. Martinez-Gomez, *Curr. Issues Mol. Biol.*, 2019, 101–116.

10 L. J. Daumann, *Angew. Chem., Int. Ed.*, 2019, **58**, 12795–12802.
11 E. R. Featherston and J. A. Cotruvo, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2021, **1868**, 118864.
12 L. J. Daumann, A. Pol, H. J. M. Op Den Camp and N. C. Martinez-Gomez, *Advances in Microbial Physiology*, Elsevier, 2022, vol. 81, pp. 1–24.
13 A. M. Webster and A. F. A. Peacock, *Chem. Commun.*, 2021, **57**, 6851–6862.
14 J. A. Mattocks and J. A. Cotruvo, *Chem. Soc. Rev.*, 2020, **49**, 8315–8334.
15 Q. Ye, D. Wang and N. Wei, *Trends Biotechnol.*, 2024, **42**, 575–590.
16 R. T. Kovacic, J. T. Welch and S. J. Franklin, *J. Am. Chem. Soc.*, 2003, **125**, 6656–6662.
17 S. Lim and S. J. Franklin, *Cell. Mol. Life Sci.*, 2004, **61**, 2184–2188.
18 S. J. Franklin and J. T. Welch, *Comments Inorg. Chem.*, 2005, **26**, 127–164.
19 A. S. Klein, F. Leiss-Maier, R. Mühlhofer, B. Boesen, G. Mustafa, H. Kugler and C. Zeymer, *J. Am. Chem. Soc.*, 2024, **146**(38), 25976–25985.
20 P. J. Thompson, D. G. Boggs, C. A. Wilson, A. T. Bruchs, U. Velidandla, J. Bridwell-Rabb and L. Olshansky, *Proc. Natl. Acad. Sci. U. S. A.*, 2024, **121**, e2405836121.
21 C. W. Breeze, Y. Nakano, E. C. Campbell, R. L. Frkic, D. W. Lupton and C. J. Jackson, *Acta Crystallogr., Sect. D:Struct. Biol.*, 2024, **80**, 289–298.
22 L. Götzke, G. Schaper, J. März, P. Kaden, N. Huittinen, T. Stumpf, K. K. K. Kammerlander, E. Brunner, P. Hahn, A. Mehnert, B. Kersting, T. Henle, L. F. Lindoy, G. Zanoni and J. J. Weigand, *Coord. Chem. Rev.*, 2019, **386**, 267–309.
23 R. Sánchez-Fernández, I. Obregon-Gomez, A. Sarmiento, M. E. Vázquez and E. Pazos, *Chem. Commun.*, 2024, **60**, 12650–12661.
24 M. A. Hay and C. Boskovic, *Chem. – Eur. J.*, 2021, **27**, 3608–3637.
25 R. G. Pearson, *J. Am. Chem. Soc.*, 1963, **85**, 3533–3539.
26 L. R. Morss, *Chem. Rev.*, 1976, **76**, 827–841.
27 P. D'Angelo, A. Zitolo, V. Migliorati, G. Chillemi, M. Duvail, P. Vitorge, S. Abadie and R. Spezia, *Inorg. Chem.*, 2011, **50**, 4572–4579.
28 R. D. Shannon, *Acta Crystallogr., Sect. A*, 1976, **32**, 751–767.
29 C. Ekberg and P. L. Brown, *Hydrolysis of Metal Ions*, John Wiley & Sons, Ltd, 2016.
30 H. C. Aspinall, *f-Block Chemistry*, Oxford University Press, 2020.
31 L. Ruckthong, M. L. Zastrow, J. A. Stuckey and V. L. Pecoraro, *J. Am. Chem. Soc.*, 2016, **138**, 11979–11988.
32 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.
33 S. Liu, E. R. Featherston, J. A. Cotruvo and C. R. Baiz, *Phys. Chem. Chem. Phys.*, 2021, **23**, 21690–21700.
34 S. A. Cotton, *C. R. Chim.*, 2005, **8**, 129–145.
35 A. Smerigan, S. Biswas, F. D. Vila, J. Hong, J. Perez-Aguilar, A. S. Hoffman, L. Greenlee, R. B. Getman and S. R. Bare, *Inorg. Chem.*, 2023, **62**, 14523–14532.
36 T. Summers, D. Zhang, J. Sobrinho, A. de Bettencourt-Dias, R. Rousseau, V.-A. Glezakou and D. Cantu, *ChemRxiv*, 2024, preprint, DOI: **10.26434/chemrxiv-2024-lgkjh**.
37 H. Lumpe, A. Menke, C. Haisch, P. Mayer, A. Kabelitz, K. V. Yusenko, A. Guilherme Buzanich, T. Block, R. Pöttgen, F. Emmerling and L. J. Daumann, *Chem. – Eur. J.*, 2020, **26**, 10133–10139.
38 K. Binnemans, *Coord. Chem. Rev.*, 2015, **295**, 1–45.
39 A. Beeby, I. M. Clarkson, R. S. Dickins, S. Faulkner, D. Parker, L. Royle, A. S. de Sousa, J. A. G. Williams and M. Woods, *J. Chem. Soc., Perkin Trans. 2*, 1999, 493–504.
40 R. M. Supkowski and W. D. W. Horrocks, *Inorg. Chim. Acta*, 2002, **340**, 44–48.
41 C. S. Bonnet, P. H. Fries, S. Crouzy, O. Sénèque, F. Cisnetti, D. Boturyn, P. Dumy and P. Delangle, *Chem. – Eur. J.*, 2009, **15**, 7083–7093.
42 G. Pintacuda, M. John, X.-C. Su and G. Otting, *Acc. Chem. Res.*, 2007, **40**, 206–212.
43 K. N. Allen and B. Imperiali, *Curr. Opin. Chem. Biol.*, 2010, **14**, 247–254.
44 W.-M. Liu, M. Overhand and M. Ubbink, *Coord. Chem. Rev.*, 2014, **273–274**, 2–12.
45 C. Nitsche and G. Otting, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2017, **98–99**, 20–49.
46 D. Joss and D. Häussinger, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2019, **114–115**, 284–312.
47 I. Baig, I. Bertini, C. Del Bianco, Y. K. Gupta, Y.-M. Lee, C. Luchinat and A. Quattrone, *Biochemistry*, 2004, **43**, 5562–5573.
48 I. Bertini, C. Del Bianco, I. Gelis, N. Katsaros, C. Luchinat, G. Parigi, M. Peana, A. Provenzani and M. A. Zoroddu, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 6841–6846.
49 N. C. J. Strynadka and M. N. G. James, *Annu. Rev. Biochem.*, 1989, **58**, 951–999.
50 A. Lewit-Bentley and S. Réty, *Curr. Opin. Struct. Biol.*, 2000, **10**, 637–643.
51 Z. Grabarek, *J. Mol. Biol.*, 2006, **359**, 509–525.
52 J. L. Gifford, M. P. Walsh and H. J. Vogel, *Biochem. J.*, 2007, **405**, 199–221.
53 D. B. Halling, B. J. Liebeskind, A. W. Hall and R. W. Aldrich, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**(9), 1216–1225.
54 H. G. Brittain, F. S. Richardson and R. B. Martin, *J. Am. Chem. Soc.*, 1976, **98**, 8255–8260.
55 B. Martin and F. S. Richardson, *Q. Rev. Biophys.*, 1979, **12**, 181–209.
56 W. D. Horrocks and D. R. Sudnick, *Acc. Chem. Res.*, 1981, **14**, 384–392.
57 M. A. Valentini and J. C. Wright, *Anal. Biochem.*, 1985, **150**, 47–57.
58 C.-L. A. Wang, P. C. Leavis, W. D. Horrocks and J. Gergely, *Biochemistry*, 1981, **20**, 2439–2444.
59 J. Gariepy, B. D. Sykes and R. S. Hodges, *Biochemistry*, 1983, **22**, 1765–1772.

60 K. D. Daughtry, L. J. Martin, A. Sarraju, B. Imperiali and K. N. Allen, *ChemBioChem*, 2012, **13**, 2567–2574.
61 S. Jain, J. T. Welch, W. D. Horrocks and S. J. Franklin, *Inorg. Chem.*, 2003, **42**, 8098–8104.
62 J. A. Mattocks, J. J. Jung, C.-Y. Lin, Z. Dong, N. H. Yennawar, E. R. Featherston, C. S. Kang-Yun, T. A. Hamilton, D. M. Park, A. K. Boal and J. A. Cotruvo, *Nature*, 2023, **618**, 87–93.
63 E. R. Featherston, E. J. Issertell and J. A. Cotruvo, *J. Am. Chem. Soc.*, 2021, **143**, 14287–14299.
64 M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen and B. Imperiali, *Angew. Chem., Int. Ed.*, 2004, **43**, 3682–3685.
65 K. Barthelmes, A. M. Reynolds, E. Peisach, H. R. A. Jonker, N. J. DeNunzio, K. N. Allen, B. Imperiali and H. Schwalbe, *J. Am. Chem. Soc.*, 2011, **133**, 808–819.
66 J. P. MacManus, C. W. Hogue, B. J. Marsden, M. Sikorska and A. G. Szabo, *J. Biol. Chem.*, 1990, **265**, 10358–10366.
67 J. Wöhnert, K. J. Franz, M. Nitz, B. Imperiali and H. Schwalbe, *J. Am. Chem. Soc.*, 2003, **125**, 13338–13339.
68 L. J. Martin, M. J. Hähnke, M. Nitz, J. Wöhnert, N. R. Silvaggi, K. N. Allen, H. Schwalbe and B. Imperiali, *J. Am. Chem. Soc.*, 2007, **129**, 7106–7113.
69 N. R. Silvaggi, L. J. Martin, H. Schwalbe, B. Imperiali and K. N. Allen, *J. Am. Chem. Soc.*, 2007, **129**, 7114–7120.
70 D. Barthelmes, M. Gränz, K. Barthelmes, K. N. Allen, B. Imperiali, T. Prisner and H. Schwalbe, *J. Biomol. NMR*, 2015, **63**, 275–282.
71 T. W. Victor, K. H. O'Toole, L. M. Easthon, M. Ge, R. J. Smith, X. Huang, H. Yan, Y. S. Chu, S. Chen, D. Gursoy, M. Ralle, B. Imperiali, K. N. Allen and L. M. Miller, *J. Am. Chem. Soc.*, 2020, **142**, 2145–2149.
72 M. Nitz, K. J. Franz, R. L. Maglathlin and B. Imperiali, *ChemBioChem*, 2003, **4**, 272–276.
73 L. J. Martin, B. R. Sculimbrene, M. Nitz and B. Imperiali, *QSAR Comb. Sci.*, 2005, **24**, 1149–1157.
74 L. J. Martin and B. Imperiali, in *Peptide Libraries: Methods and Protocols*, ed. R. Derda, Springer, New York, NY, 2015, pp. 201–220.
75 Y. Kim, J. T. Welch, K. M. Lindstrom and S. J. Franklin, *JBIC, J. Biol. Inorg. Chem.*, 2001, **6**, 173–181.
76 J. T. Welch, W. R. Kearney and S. J. Franklin, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 3725–3730.
77 P. Caravan, J. M. Greenwood, J. T. Welch and S. J. Franklin, *Chem. Commun.*, 2003, 2574–2575.
78 S. Lim and S. J. Franklin, *Protein Sci.*, 2006, **15**, 2159–2165.
79 S. W. Wong-Deyrup, Y. Kim and S. J. Franklin, *JBIC, J. Biol. Inorg. Chem.*, 2006, **11**, 17–25.
80 K. L. Harris, S. Lim and S. J. Franklin, *Inorg. Chem.*, 2006, **45**, 10002–10012.
81 J. A. Cotruvo, E. R. Featherston, J. A. Mattocks, J. V. Ho and T. N. Laremore, *J. Am. Chem. Soc.*, 2018, **140**, 15056–15061.
82 C.-L. A. Wang, R. R. Aquaron, P. C. Leavis and J. Gergely, *Eur. J. Biochem.*, 1982, **124**, 7–12.
83 E. C. Cook, E. R. Featherston, S. A. Showalter and J. A. Cotruvo, *Biochemistry*, 2019, **58**, 120–125.
84 F. Cisnetti, C. Lebrun and P. Delangle, *Dalton Trans.*, 2010, **39**, 3560–3562.
85 F. Cisnetti, C. Gateau, C. Lebrun and P. Delangle, *Chem. – Eur. J.*, 2009, **15**, 7456–7469.
86 A. Niedźwiecka, F. Cisnetti, C. Lebrun, C. Gateau and P. Delangle, *Dalton Trans.*, 2012, **41**, 3239–3247.
87 A. Niedźwiecka, F. Cisnetti, C. Lebrun and P. Delangle, *Inorg. Chem.*, 2012, **51**, 5458–5464.
88 J. Veliscek-Carolan, T. L. Hanley and K. A. Jolliffe, *RSC Adv.*, 2016, **6**, 75336–75346.
89 S. Safi, A. Jeanson, J. Roques, P. L. Solari, F. Charnay-Pouget, C. Den Auwer, G. Creff, D. J. Aitken and E. Simoni, *Inorg. Chem.*, 2016, **55**, 877–886.
90 I. V. Korendovych and W. F. DeGrado, *Q. Rev. Biophys.*, 2020, **53**, e3.
91 K. J. Koebke, T. B. J. Pinter, W. C. Pitts and V. L. Pecoraro, *Chem. Rev.*, 2022, **122**, 12046–12109.
92 P. Teare, C. F. Smith, S. J. Adams, S. Anbu, B. Ciani, L. J. C. Jeuken and A. F. A. Peacock, *Dalton Trans.*, 2018, **47**, 10784–10790.
93 D. N. Woolfson, *Advances in Protein Chemistry*, Academic Press, 2005, vol. 70, pp. 79–112.
94 W. D. Kohn, C. M. Kay and R. S. Hodges, *J. Pept. Res.*, 1998, **51**, 9–18.
95 W. D. Kohn, C. M. Kay, B. D. Sykes and R. S. Hodges, *J. Am. Chem. Soc.*, 1998, **120**, 1124–1132.
96 A. Kashiwada, K. Ishida and K. Matsuda, *Bull. Chem. Soc. Jpn.*, 2007, **80**, 2203–2207.
97 M. Samiappan, S. Alasibi, R. Cohen-Luria, A. Shanzer and G. Ashkenasy, *Chem. Commun.*, 2012, **48**, 9577–9579.
98 M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton and A. F. A. Peacock, *J. Am. Chem. Soc.*, 2014, **136**, 1166–1169.
99 M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton and A. F. A. Peacock, *Chem. Sci.*, 2016, **7**, 2207–2216.
100 L. N. Slope, M. G. Hill, C. F. Smith, P. Teare, F. J. de Cogan, M. M. Britton and A. F. A. Peacock, *Chem. Commun.*, 2020, **56**, 3729–3732.
101 L. N. Slope, O. J. Daubney, H. Campbell, S. A. White and A. F. A. Peacock, *Angew. Chem., Int. Ed.*, 2021, **60**, 24473–24477.
102 S. L. Newton, A. Franke, A. Zahl, G. Molinaro, A. Kenwright, D. J. Smith, I. Ivanovic-Burmazovic, M. M. Britton and A. F. A. Peacock, *Dalton Trans.*, 2023, **52**, 15665–15668.
103 C. W. A. Ende, H. Y. Meng, M. Ye, A. K. Pandey and N. J. Zondlo, *ChemBioChem*, 2010, **11**, 1738–1747.
104 A. F. A. Peacock, *J. Inorg. Biochem.*, 2025, **268**, 112903.
105 J. K. Myers, C. N. Pace and J. M. Scholtz, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 2833–2837.
106 J. S. Richardson and D. C. Richardson, *Science*, 1988, **240**, 1648–1652.
107 L. G. Presta and G. D. Rose, *Science*, 1988, **240**, 1632–1641.
108 L. Serrano and A. R. Fersht, *Nature*, 1989, **342**, 296–299.
109 Y. W. Deng, S. Y. Ro and A. C. Rosenzweig, *JBIC, J. Biol. Inorg. Chem.*, 2018, **23**, 1037–1047.

 

110 N. M. Good, M. Fellner, K. Demirer, J. Hu, R. P. Hausinger and N. C. Martinez-Gomez, *J. Biol. Chem.*, 2020, **295**, 8272–8284.
111 R. A. Schmitz, N. Picone, H. Singer, A. Dietl, K.-A. Seifert, A. Pol, M. S. M. Jetten, T. R. M. Barends, L. J. Daumann and H. J. M. Op den Camp, *mBio*, 2021, **12**, e01708–21.
112 B. Jahn, A. Pol, H. Lumpe, T. R. M. Barends, A. Dietl, C. Hogendoorn, H. J. M. Op den Camp and L. J. Daumann, *ChemBioChem*, 2018, **19**, 1147–1153.
113 N. M. Good, H. N. Vu, C. J. Suriano, G. A. Subuyuj, E. Skovran and N. C. Martinez-Gomez, *J. Bacteriol.*, 2016, **198**, 3109–3118.
114 M. Wehrmann, P. Billard, A. Martin-Meriadec, A. Zegeye and J. Klebensberger, *mBio*, 2017, **8**, e00570–17.
115 J. Huang, Z. Yu, J. Groom, J.-F. Cheng, A. Tarver, Y. Yoshikuni and L. Chistoserdova, *ISME J.*, 2019, **13**, 2005–2017.
116 M. Wehrmann, E. M. Elsayed, S. Köbbing, L. Bendz, A. Lepak, J. Schwabe, N. Wierckx, G. Bange and J. Klebensberger, *ACS Catal.*, 2020, **10**, 7836–7842.
117 N. M. Good, R. S. Moore, C. J. Suriano and N. C. Martinez-Gomez, *Sci. Rep.*, 2019, **9**, 1–12.
118 L. Wang, A. Hibino, S. Suganuma, A. Ebihara, S. Iwamoto, R. Mitsui, A. Tani, M. Shimada, T. Hayakawa and T. Nakagawa, *Enzyme Microb. Technol.*, 2020, **136**, 109518.
119 H. Lumpe, A. Pol, H. J. M. Op den Camp and L. J. Daumann, *Dalton Trans.*, 2018, **47**, 10463–10472.
120 C. Stines-Chaumeil, F. Mavré, B. Kauffmann, N. Mano and B. Limoges, *ACS Omega*, 2020, **5**, 2015–2026.
121 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, *Sci. Rep.*, 2020, **10**, 12663.
122 J. L. Hemmann, P. Keller, L. Hemmerle, T. Vonderach, A. M. Ochsner, M. Bortfeld-Miller, D. Günther and J. A. Vorholt, *J. Biol. Chem.*, 2023, **299**, 102940.
123 J. A. Mattocks, J. V. Ho and J. A. Cotruvo, *J. Am. Chem. Soc.*, 2019, **141**, 2857–2861.
124 A. M. Ochsner, L. Hemmerle, T. Vonderach, R. Nüssli, M. Bortfeld-Miller, B. Hattendorf and J. A. Vorholt, *Mol. Microbiol.*, 2019, **111**, 1152–1166.
125 W. B. Larrinaga, J. J. Jung, C.-Y. Lin, A. K. Boal and J. A. Cotruvo, *Proc. Natl. Acad. Sci. U. S. A.*, 2024, **121**, e2410926121.
126 S. J. Caldwell, I. C. Haydon, N. Piperidou, P.-S. Huang, M. J. Bick, H. S. Sjöström, D. Hilvert, D. Baker and C. Zeymer, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 30362–30369.
127 M. A. Blenner, O. Shur, G. R. Szilvay, D. M. Cropek and S. Banta, *J. Mol. Biol.*, 2010, **400**, 244–256.
128 W. Abdallah, K. Solanki and S. Banta, *ACS Catal.*, 2018, **8**, 1602–1613.
129 F. Khoury, Z. Su and S. Banta, *Inorg. Chem.*, 2024, **63**(29), 13223–13230.
130 J. A. Mattocks, J. L. Tirsch and J. A. Cotruvo, in *Methods in Enzymology*, ed. J. A. Cotruvo, Academic Press, 2021, vol. 651, pp. 23–61.
131 G. Deblonde, *Methods in Enzymology*, Academic Press, 2021, vol. 651, pp. 1–22.
132 W. R. Harris, C. J. Carrano and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 2722–2727.
133 E. Falcone and P. Faller, *Dalton Trans.*, 2023, **52**, 2197–2208.
134 Z. M. Anwar and H. A. Azab, *J. Chem. Eng. Data*, 2001, **46**, 613–618.
135 H. A. Azab, S. S. Al-Deyab, Z. M. Anwar, I. I. Abd El-Gawad and R. M. Kamel, *J. Chem. Eng. Data*, 2011, **56**, 2613–2625.
136 H. A. Azab, Z. M. Anwar and R. M. Kamel, *J. Solution Chem.*, 2016, **45**, 1516–1528.
137 P. Mandal, J. Kretzschmar and B. Drobot, *JBIC, J. Biol. Inorg. Chem.*, 2022, **27**, 249–260.
138 R. N. Goldberg, N. Kishore and R. M. Lennen, *J. Phys. Chem. Ref. Data*, 2002, **31**, 231–370.
139 K. J. Franz, M. Nitz and B. Imperiali, *ChemBioChem*, 2003, **4**, 265–271.
140 E. A. Alasadi, W. Choi, X. Chen, J. A. J. Cotruvo and C. R. Baiz, *ACS Chem. Biol.*, 2024, **19**, 1056–1065.
141 E. M. Jones, Y. Su, C. Sander, Q. A. Justman, M. Springer and P. A. Silver, *ACS Synth. Biol.*, 2024, **13**, 958–962.
142 L. L. Clainche, M. Figuet, V. Montjardet-Bas, S. Blanchard and C. Vita, *Biotechnol. Bioeng.*, 2006, **95**, 29–36.
143 S. M. Gutenthaler, S. Tsushima, R. Steudtner, M. Gailer, A. Hoffmann-Röder, B. Drobot and L. J. Daumann, *Inorg. Chem. Front.*, 2022, **9**, 4009–4021.
144 S. C. Zondlo, F. Gao and N. J. Zondlo, *J. Am. Chem. Soc.*, 2010, **132**, 5619–5621.
145 S. Balakrishnan and N. J. Zondlo, *J. Am. Chem. Soc.*, 2006, **128**, 5590–5591.
146 A. R. Urmey and N. J. Zondlo, *Biochemistry*, 2019, **58**, 2822–2833.
147 A. R. Urmey and N. J. Zondlo, *Free Radical Biol. Med.*, 2020, **152**, 166–174.
148 F. Gao, B. S. Thornley, C. M. Tressler, D. Naduthambi and N. J. Zondlo, *Org. Biomol. Chem.*, 2019, **17**, 3984–3995.
149 L. L. Liu and K. J. Franz, *J. Am. Chem. Soc.*, 2005, **127**, 9662–9663.
150 J. A. Mattocks, J. L. Tirsch and J. A. Cotruvo, *Methods in Enzymology*, Elsevier, 2021, vol. 651, pp. 23–61.
151 B. Mwewa, M. Tadie, S. Ndlovu, G. S. Simate and E. Matinde, *J. Environ. Chem. Eng.*, 2022, **10**, 107704.
152 S. Pramanik, S. Kaur, I. Popovs, A. S. Ivanov and S. Jansone-Popova, *Eur. J. Inorg. Chem.*, 2024, e202400064.
153 J. Haiech, C. B. Klee, J. G. Demaille and J. Haiech, *Biochemistry*, 1981, **20**, 3890–3897.
154 D. M. Park, D. W. Reed, M. C. Yung, A. Eslamimanesh, M. M. Lencka, A. Anderko, Y. Fujita, R. E. Riman, A. Navrotsky and Y. Jiao, *Environ. Sci. Technol.*, 2016, **50**, 2735–2742.
155 G. J.-P. Deblonde, J. A. Mattocks, D. M. Park, D. W. Reed, J. A. Cotruvo and Y. Jiao, *Inorg. Chem.*, 2020, **59**, 11855–11867.
156 T. L. Pham, M. R. Conde González, S. Fazliev, A. Kishore, P. Comba and F. Thomas, *ChemBioChem*, 2024, **25**, e202300715.
157 D. Ghosh, K.-H. Lee, B. Demeler and V. L. Pecoraro, *Biochemistry*, 2005, **44**, 10732–10740.

158 J. A. Peters, K. Djanashvili, C. F. G. C. Geraldes and C. Platas-Iglesias, *Coord. Chem. Rev.*, 2020, **406**, 213146.
159 R. K. Wilharm, S.-Y. Huang, I. J. Gugger and V. C. Pierre, *Inorg. Chem.*, 2021, **60**, 15808–15817.
160 A. Roca-Sabio, M. Mato-Iglesias, D. Esteban-Gómez, É. Tóth, A. de Blas, C. Platas-Iglesias and T. Rodríguez-Blas, *J. Am. Chem. Soc.*, 2009, **131**, 3331–3341.
161 A. Hu, S. N. MacMillan and J. J. Wilson, *J. Am. Chem. Soc.*, 2020, **142**, 13500–13506.
162 S. L. Wu and W. D. W. Horrocks, *J. Chem. Soc., Dalton Trans.*, 1997, 1497–1502.
163 A. Hu, E. Aluicio-Sarduy, V. Brown, S. N. MacMillan, K. V. Becker, T. E. Barnhart, V. Radchenko, C. F. Ramogida, J. W. Engle and J. J. Wilson, *J. Am. Chem. Soc.*, 2021, **143**, 10429–10440.
164 T. Hatanaka, N. Kikkawa, A. Matsugami, Y. Hosokawa, F. Hayashi and N. Ishida, *Sci. Rep.*, 2020, **10**, 19468.
165 H. Singer, B. Drobot, C. Zeymer, R. Steudtner and L. J. Daumann, *Chem. Sci.*, 2021, **12**, 15581–15587.
166 J. A. Mattocks, J. L. Tirsch and J. A. Cotruvo, *Methods in Enzymology*, Elsevier, 2021, vol. 651, pp. 23–61.
167 Z. Dong, J. A. Mattocks, J. A. Seidel, J. A. Cotruvo and D. M. Park, *Sep. Purif. Technol.*, 2024, **333**, 125919.
168 J. Seidel, P. Diep, Z. Dong, J. A. Cotruvo and D. M. Park, *JACS Au*, 2024, **4**(11), 4273–4284.
169 C. S. Bonnet, M. Devocelle and T. Gunnlaugsson, *Org. Biomol. Chem.*, 2012, **10**, 126–133.
170 G. J.-P. Deblonde, J. A. Mattocks, H. Wang, E. M. Gale, A. B. Kersting, M. Zavarin and J. A. Cotruvo, *J. Am. Chem. Soc.*, 2021, **143**, 15769–15783.
171 G. J.-P. Deblonde, J. A. Mattocks, Z. Dong, P. T. Wooddy, J. A. Cotruvo and M. Zavarin, *Sci. Adv.*, 2021, **7**, eabk0273.
172 G. J.-P. Deblonde, K. Morrison, J. A. Mattocks, J. A. Cotruvo, M. Zavarin and A. B. Kersting, *Environ. Sci. Technol.*, 2023, **57**, 20830–20843.
173 R. M. Smith and A. E. Martell, *Sci. Total Environ*, 1987, **64**, 125–147.
174 J. A. Mattocks, J. A. Cotruvo and G. J.-P. Deblonde, *Chem. Sci.*, 2022, **13**, 6054–6066.
175 S. Özçubukçu, K. Mandal, S. Wegner, M. P. Jensen and C. He, *Inorg. Chem.*, 2011, **50**, 7937–7939.
176 H. Singer, R. Steudtner, A. S. Klein, C. Rulofs, C. Zeymer, B. Drobot, A. Pol, N. Cecilia Martinez-Gomez, H. J. M. Op den Camp and L. J. Daumann, *Angew. Chem., Int. Ed.*, 2023, **62**, e202303669.
177 A. Shah, M. J. Taylor, G. Molinaro, S. Anbu, M. Verdu, L. Jennings, I. Mikulska, S. Diaz-Moreno, H. E. L. Mkami, G. M. Smith, M. M. Britton, J. E. Lovett and A. F. A. Peacock, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**, e2219036120.
178 C. S. Bonnet, M. Devocelle and T. Gunnlaugsson, *Chem. Commun.*, 2008, 4552–4554.
179 S. K. Drake and J. J. Falke, *Biochemistry*, 1996, **35**, 1753–1760.
180 M. C. Heffern, L. M. Matosziuk and T. J. Meade, *Chem. Rev.*, 2014, **114**, 4496–4539.
181 A. J. Amoroso and S. J. A. Pope, *Chem. Soc. Rev.*, 2015, **44**, 4723–4742.
182 E. Mathieu, A. Sipos, E. Demeyere, D. Phipps, D. Sakaveli and K. E. Borbas, *Chem. Commun.*, 2018, **54**, 10021–10035.
183 U. Cho and J. K. Chen, *Cell Chem. Biol.*, 2020, **27**, 921–936.
184 J. H. S. K. Monteiro, *Molecules*, 2020, **25**, 2089.
185 R. Lengacher, K. E. Martin, D. Śmiłowicz, H. Esseln, P. Lotlikar, A. Grichine, O. Maury and E. Boros, *J. Am. Chem. Soc.*, 2023, **145**, 24358–24366.
186 J.-C. G. Bünzli, *Chem. Rev.*, 2010, **110**, 2729–2755.
187 S. V. Eliseeva, *Angew. Chem., Int. Ed.*, 2015, **54**, 8598.
188 A. D'Aléo, F. Pointillart, L. Ouahab, C. Andraud and O. Maury, *Coord. Chem. Rev.*, 2012, **256**, 1604–1620.
189 J.-C. G. Bünzli, *Coord. Chem. Rev.*, 2015, **293–294**, 19–47.
190 D. Kovacs and K. E. Borbas, *Coord. Chem. Rev.*, 2018, **364**, 1–9.
191 J. De Jersey, P. J. Morley and R. B. Martin, *Biophys. Chem.*, 1981, **13**, 233–243.
192 M. Latva, H. Takalo, V.-M. Mukkala, C. Matachescu, J. C. Rodríguez-Ubis and J. Kankare, *J. Lumin.*, 1997, **75**, 149–169.
193 W. D. Horrocks, J. P. Bolender, W. D. Smith and R. M. Supkowski, *J. Am. Chem. Soc.*, 1997, **119**, 5972–5973.
194 E. Falcone, P. Gonzalez, L. Lorusso, O. Sénèque, P. Faller and L. Raibaut, *Chem. Commun.*, 2020, **56**, 4797–4800.
195 A. M. Reynolds, B. R. Sculimbrene and B. Imperiali, *Bioconjugate Chem.*, 2008, **19**, 588–591.
196 J. Wang, X. Liu, K. Li, T. Shi, Q. Xu, T. Peng, Q. Huang, Z. Gao, H. Zhou, W. Lu and J. Wang, *J. Am. Chem. Soc.*, 2025, **147**, 15205–15215.
197 M. Starck, J. D. Fradgley, S. Di Vita, J. A. Mosely, R. Pal and D. Parker, *Bioconjugate Chem.*, 2020, **31**, 229–240.
198 H.-F. Chau, Y. Wu, W.-Y. Fok, W. Thor, W. C.-S. Cho, P. Ma, J. Lin, N.-K. Mak, J.-C. G. Bünzli, L. Jiang, N. J. Long, H. L. Lung and K.-L. Wong, *JACS Au*, 2021, **1**, 1034–1043.
199 J.-H. Choi, G. Fremy, T. Charnay, N. Fayad, J. Pécaut, S. Erbek, N. Hildebrandt, V. Martel-Frachet, A. Grichine and O. Sénèque, *Inorg. Chem.*, 2022, **61**, 20674–20689.
200 C. Cepeda, S. A. Denisov, D. Boturyn, N. D. McClenaghan and O. Sénèque, *Inorg. Chem.*, 2021, **60**, 17426–17434.
201 M. Isaac, L. Raibaut, C. Cepeda, A. Roux, D. Boturyn, S. V. Eliseeva, S. Petoud and O. Sénèque, *Chem. – Eur. J.*, 2017, **23**, 10992–10996.
202 Y.-H. Yeung, H.-F. Chau, H.-Y. Kai, W. Zhou, K. H.-Y. Chan, W. Thor, L. J. Charbonnière, F. Zhang, Y. Fan, Y. Wu and K.-L. Wong, *Adv. Opt. Mater.*, 2024, **12**, 2302070.
203 C. Cepeda, L. Raibaut, G. Fremy, S. V. Eliseeva, A. Romieu, J. Pécaut, D. Boturyn, S. Petoud and O. Sénèque, *Chem. – Eur. J.*, 2020, **26**, 13476–13483.
204 C. Penas, J. L. Mascareñas and M. E. Vázquez, *Chem. Sci.*, 2016, **7**, 2674–2678.
205 B. R. Sculimbrene and B. Imperiali, *J. Am. Chem. Soc.*, 2006, **128**, 7346–7352.
206 K. Y. Huang, L. Cardenas, A. D. Ellington and D. J. F. Walker, *Nat. Commun.*, 2024, **15**, 9200.

 

207 L. Raibaut, W. Vasseur, G. D. Shimberg, C. Saint-Pierre, J.-L. Ravanat, S. L. J. Michel and O. Sénèque, *Chem. Sci.*, 2017, **8**, 1658–1664.
208 E. Pazos, M. Goličnik, J. L. Mascareñas and M. E. Vázquez, *Chem. Commun.*, 2012, **48**, 9534–9536.
209 R. Sánchez-Fernández, A. Sánchez-Temprano, D. Esteban-Gómez and E. Pazos, *ChemBioChem*, 2023, **24**, e202300072.
210 E. Falcone, B. Vileno, M. Hoang, L. Raibaut and P. Faller, *J. Inorg. Biochem.*, 2021, **221**, 111478.
211 M. J. Scheuermann, C. R. Forbes and N. J. Zondlo, *Biochemistry*, 2018, **57**, 6956–6963.
212 J. Wahsner, E. M. Gale, A. Rodríguez-Rodríguez and P. Caravan, *Chem. Rev.*, 2019, **119**, 957–1057.
213 M. E. Ladd, P. Bachert, M. Meyerspeer, E. Moser, A. M. Nagel, D. G. Norris, S. Schmitter, O. Speck, S. Straub and M. Zaiss, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2018, **109**, 1–50.
214 P. Caravan, *Chem. Soc. Rev.*, 2006, **35**, 512–523.
215 P. Caravan, J. J. Ellison, T. J. McMurry and R. B. Lauffer, *Chem. Rev.*, 1999, **99**, 2293–2352.
216 D. H. Powell, O. M. N. Dhubhghaill, D. Pubanz, L. Helm, Y. S. Lebedev, W. Schlaepfer and A. E. Merbach, *J. Am. Chem. Soc.*, 1996, **118**, 9333–9346.
217 J. J. Yang, J. Yang, L. Wei, O. Zurkiya, W. Yang, S. Li, J. Zou, Y. Zhou, A. L. W. Maniccia, H. Mao, F. Zhao, R. Malchow, S. Zhao, J. Johnson, X. Hu, E. Krogstad and Z.-R. Liu, *J. Am. Chem. Soc.*, 2008, **130**, 9260–9267.
218 S. Xue, H. Yang, J. Qiao, F. Pu, J. Jiang, K. Hubbard, K. Hekmatyar, J. Langley, M. Salarian, R. C. Long, R. G. Bryant, X. P. Hu, H. E. Grossniklaus, Z.-R. Liu and J. J. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 6607–6612.
219 C. S. Bonnet, P. H. Fries, S. Crouzy and P. Delangle, *J. Phys. Chem. B*, 2010, **114**, 8770–8781.
220 Y. Liu, D. Gao, Y. He, J. Ma, S. Y. Chong, X. Qi, H. J. Ting, Z. Luo, Z. Yi, J. Tang, C. Chang, J. Wang, Z. Sheng, H. Zheng and X. Liu, *Nat. Commun.*, 2024, **15**, 9834.
221 D. Li, M. Kirberger, J. Qiao, Z. Gui, S. Xue, F. Pu, J. Jiang, Y. Xu, S. Tan, M. Salarian, O. Ibhagui, K. Hekmatyar and J. J. Yang, *Invest. Radiol.*, 2024, **59**, 170–186.
222 F. Pu, M. Salarian, S. Xue, J. Qiao, J. Feng, S. Tan, A. Patel, X. Li, K. Mamouni, K. Hekmatyar, J. Zou, D. Wu and J. J. Yang, *Nanoscale*, 2016, **8**, 12668–12682.
223 M. Salarian, H. Yang, R. C. Turaga, S. Tan, J. Qiao, S. Xue, Z. Gui, G. Peng, H. Han, P. Mittal, H. E. Grossniklaus and J. J. Yang, *Biomaterials*, 2019, **224**, 119478.
224 S. Tan, H. Yang, S. Xue, J. Qiao, M. Salarian, K. Hekmatyar, Y. Meng, R. Mukkavilli, F. Pu, O. Y. Odubade, W. Harris, Y. Hai, M. L. Yushak, V. M. Morales-Tirado, P. Mittal, P. Z. Sun, D. Lawson, H. E. Grossniklaus and J. J. Yang, *Sci. Adv.*, 2020, **6**, eaav7504.
225 H. Yang, S. Tan, J. Qiao, Y. Xu, Z. Gui, Y. Meng, B. Dong, G. Peng, O. Y. Ibhagui, W. Qian, J. Lu, Z. Li, G. Wang, J. Lai, L. Yang, H. E. Grossniklaus and J. J. Yang, *Cancer Gene Ther.*, 2022, **29**, 1827–1839.
226 O. Y. Ibhagui, D. Li, H. Han, G. Peng, M. L. Meister, Z. Gui, J. Qiao, M. Salarian, B. Dong, Y. Yuan, Y. Xu, H. Yang, S. Tan, G. Satyanarayana, S. Xue, R. C. Turaga, M. Sharma, Y. Hai, Y. Meng, K. Hekmatyar, P. Sun, G. Sica, X. Ji, Z.-R. Liu and J. J. Yang, *Chem. Biomed. Imaging*, 2023, **1**, 268–285.
227 Q. N. Do, J. S. Ratnakar, Z. Kovács and A. D. Sherry, *ChemMedChem*, 2014, **9**, 1116–1129.
228 G. Angelovski and É. Tóth, *Chem. Soc. Rev.*, 2017, **46**, 324–336.
229 J. Lux and A. D. Sherry, *Curr. Opin. Chem. Biol.*, 2018, **45**, 121–130.
230 C. S. Bonnet and É. Tóth, *Curr. Opin. Chem. Biol.*, 2021, **61**, 154–169.
231 H. Lu, A. Chen, X. Zhang, Z. Wei, R. Cao, Y. Zhu, J. Lu, Z. Wang and L. Tian, *Nat. Commun.*, 2022, **13**, 7948.
232 M. Isaac, A. Pallier, F. Szeremeta, P.-A. Bayle, L. Barantin, C. S. Bonnet and O. Sénèque, *Chem. Commun.*, 2018, **54**, 7350–7353.
233 K. P. Malikidogo, A. Pallier, F. Szeremeta, C. S. Bonnet and O. Sénèque, *Dalton Trans.*, 2023, **52**, 6260–6266.
234 M. Sethurajan, E. D. van Hullebusch, D. Fontana, A. Akcil, H. Deveci, B. Batinic, J. P. Leal, T. A. Gasche, M. Ali Kucuker, K. Kuchta, I. F. F. Neto, H. M. V. M. Soares and A. Chmielarz, *Crit. Rev. Environ. Sci. Technol.*, 2019, **49**, 212–275.
235 K. Binnemans, P. T. Jones, T. Müller and L. Yurramendi, *J. Sustainable Metall.*, 2018, **4**, 126–146.
236 K. Binnemans, P. T. Jones, B. Blanpain, T. Van Gerven and Y. Pontikes, *J. Cleaner Prod.*, 2015, **99**, 17–38.
237 S. Pramanik, S. Kaur, I. Popovs, A. S. Ivanov and S. Jansone-Popova, *Eur. J. Inorg. Chem.*, 2024, e202400064.
238 N. C. Martinez-Gomez, H. N. Vu and E. Skovran, *Inorg. Chem.*, 2016, **55**, 10083–10089.
239 S. Dev, A. Sachan, F. Dehghani, T. Ghosh, B. R. Briggs and S. Aggarwal, *Chem. Eng. J.*, 2020, **397**, 124596.
240 R. M. Brown, A. Mirkouei, D. Reed and V. Thompson, *Renewable Sustainable Energy Rev.*, 2023, **173**, 113099.
241 M. Chakankar, F. Lederer, R. Jain, S. Matys, S. Kutschke and K. Pollmann, *Management of Electronic Waste*, John Wiley & Sons, Ltd, 2024, pp. 375–405.
242 X. Qian, C. Ma, H. Zhang and K. Liu, *Bioorg. Chem.*, 2024, **143**, 107040.
243 M. Vítová and D. Mezricky, *World J. Microbiol. Biotechnol.*, 2024, **40**, 189.
244 H. Singer, R. Steudtner, I. Sottorff, B. Drobot, A. Pol, H. J. M. O. den Camp and L. J. Daumann, *Chem. Commun.*, 2023, **59**, 9066–9069.
245 I. Chae, A. Shivkumar, F. M. Doyle and S.-W. Lee, *Nano Lett.*, 2024, **24**, 9946–9952.
246 I. Chae, F. M. Doyle and S.-W. Lee, *TMS 2025 154th Annual Meeting & Exhibition Supplemental Proceedings*, Springer, Cham, 2025, pp. 1139–1148.
247 D. M. Park, A. Brewer, D. W. Reed, L. N. Lammers and Y. Jiao, *Environ. Sci. Technol.*, 2017, **51**, 13471–13480.
248 A. Brewer, E. Chang, D. M. Park, T. Kou, Y. Li, L. N. Lammers and Y. Jiao, *Environ. Sci. Technol.*, 2019, **53**, 7714–7723.
249 A. Brewer, A. Dohnalkova, V. Shutthanandan, L. Kovarik, E. Chang, A. M. Sawvel, H. E. Mason, D. Reed, C. Ye, W. F.

Hynes, L. N. Lammers, D. M. Park and Y. Jiao, *Environ. Sci. Technol.*, 2019, **53**, 13888–13897.
250 D. Park, A. Middleton, R. Smith, G. Deblonde, D. Laudal, N. Theaker, H. Hsu-Kim and Y. Jiao, *Sep. Purif. Technol.*, 2020, **241**, 116726.
251 E. Chang, A. W. Brewer, D. M. Park, Y. Jiao and L. N. Lammers, *Appl. Geochem.*, 2020, **112**, 104478.
252 X. Xie, K. Yang, Y. Lu, Y. Li, J. Yan, J. Huang, L. Xu, M. Yang and Y. Yan, *J. Hazard. Mater.*, 2022, **438**, 129561.
253 M. Gut, T. Wilhelm, O. Beniston, S. Ogundipe, C.-C. Kuo, K. Nguyen and A. Furst, *Adv. Mater.*, 2025, **37**, 2412607.
254 F. F. Khoury, B. S. Heater, D. R. Marzolf, S. S. Abeyrathna, J. W. Picking, P. Kumar, S. A. Higgins, R. Jones, A. T. Lewis, K. H. Kucharzyk and S. Banta, *Chem. Sci.*, 2025, **16**, 15333–15346.
255 P. K. R. Tay, A. Manjula-Basavanna and N. S. Joshi, *Green Chem.*, 2018, **20**, 3512–3520.
256 R. A. Scheel, J. K. Sahoo, L. D. Morton, Z. Xia, J. R. Blum, Z. Martin-Moldes and D. L. Kaplan, *ACS ES&T Eng.*, 2024, **4**(9), 2135–2144.
257 Z. Dong, J. A. Mattocks, G. J.-P. Deblonde, D. Hu, Y. Jiao, J. A. Cotruvo and D. M. Park, *ACS Cent. Sci.*, 2021, **7**, 1798–1808.
258 Z. Hussain, S. Kim, J. Cho, G. Sim, Y. Park and I. Kwon, *Adv. Funct. Mater.*, 2022, **32**, 2109158.
259 Z. Hussain, D. Dwivedi and I. Kwon, *Front. Bioeng. Biotechnol.*, 2024, **12**, DOI: **10.3389/fbioe.2024.1385845**.
260 Q. Ye, X. Jin, B. Zhu, H. Gao and N. Wei, *Environ. Sci. Technol.*, 2023, **57**, 4276–4285.
261 L. Johnson and C. E. Duval, *Langmuir*, 2025, **41**, 9581–9589.
262 Q. Ye, X. Jin, B. Zhu, H. Gao and N. Wei, *Environ. Sci. Technol.*, 2023, **57**, 4276–4285.
263 H. Jin, D. M. Park, M. Gupta, A. W. Brewer, L. Ho, S. L. Singer, W. L. Bourcier, S. Woods, D. W. Reed, L. N. Lammers, J. W. Sutherland and Y. Jiao, *ACS Sustainable Chem. Eng.*, 2017, **5**, 10148–10155.
264 J. T. Welch, M. Sirish, K. M. Lindstrom and S. J. Franklin, *Inorg. Chem.*, 2001, **40**, 1982–1984.
265 M. Sirish and S. J. Franklin, *J. Inorg. Biochem.*, 2002, **91**, 253–258.
266 Y. Kim and S. J. Franklin, *Inorg. Chim. Acta*, 2002, **341**, 107–112.
267 L. Wang, K. Liu, Z. Song, H. Do, L. Yang, J. Wu, L. Jiang and H. Yu, *Colloids Surf., B*, 2025, **251**, 114596.
268 M. J. Allen, *Synlett*, 2016, 1310–1317.
269 H. Pellissier, *Coord. Chem. Rev.*, 2017, **336**, 96–151.
270 Y. Qiao and E. J. Schelter, *Acc. Chem. Res.*, 2018, **51**, 2926–2936.
271 A. Prieto and F. Jaroschik, *Curr. Org. Chem.*, 2022, **26**, 6–41.
272 N. Mahieu, J. Piątkowski, T. Simler and G. Nocton, *Chem. Sci.*, 2023, **14**, 443–457.
273 R. Bhimpuria, R. Charaf, K. Ye, A. Thapper, H. Sathyan, M. Ahlquist, L. Hammarström and K. E. Borbas, *Chem*, 2025, 102547.
274 R. Bhimpuria, M. Tomar, A. Thapper, M. Ahlquist and K. E. Borbas, *Chem*, 2025, 102450.
275 F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis and T. R. Ward, *Chem. Rev.*, 2018, **118**, 142–231.
276 P. Kouba, P. Kohout, F. Haddadi, A. Bushuiev, R. Samusevich, J. Sedlar, J. Damborsky, T. Pluskal, J. Sivic and S. Mazurenko, *ACS Catal.*, 2023, **13**, 13863–13895.
277 P. Notin, N. Rollins, Y. Gal, C. Sander and D. Marks, *Nat. Biotechnol.*, 2024, **42**, 216–228.
278 A. Winnifrith, C. Outeiral and B. L. Hie, *Curr. Opin. Struct. Biol.*, 2024, **86**, 102794.
279 K. K. Yang, Z. Wu and F. H. Arnold, *Nat. Methods*, 2019, **16**, 687–694.
280 J. A. Mattocks, J. J. Jung, C.-Y. Lin, Z. Dong, N. H. Yennawar, E. R. Featherston, C. S. Kang-Yun, T. A. Hamilton, D. M. Park, A. K. Boal and J. A. Cotruvo, *Nature*, 2023, **618**, 87–93.

 