Chemical Science

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: 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, DOI: 10.1039/D5SC02315G.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

View Article Online
DOI: 10.1039/D5SC02315G

Mining Peptides for Mining Solutions: Evaluation of Calcium-Binding Peptides for Rare Earth Element Separations

Farid F. Khoury[a], Bradley S. Heater[b], Daniel R. Marzolf[b], Sameera Abeyrathna[a], Jonathan W. Picking[b], Piyush Kumar[a], Steven A. Higgins[b], Randy Jones[b], Alan T. Lewis, Jr.[b], Katarzyna H. Kucharzyk[b,*], and Scott Banta[a,*]

[a] Department of Chemical Engineering, Columbia University, New York, NY, 10027

[b] Battelle Memorial Research Institute, Columbus, OH, 43201

Authors contributed equally

[*] Corresponding authors

E-mail: sbanta@columbia.edu

E-mail: kucharzyk@battelle.org

Abstract

Rare earth elements (REEs), which include the 15 lanthanides plus scandium and yttrium, are critical components commonly used in permanent magnets and play a significant role in electronics and green energy technologies. Due to the similarities of these ions, conventional separation processes are chemically- and energy-intensive and generate large quantities of waste. The lanthanides share physical and chemical similarities with calcium ions, which allow REEs to replace calcium in calcium-binding peptides and proteins. In this study, we conducted a bioinformatic search to identify calcium-binding peptides with high affinity to bind and separate

View Article Online
DOI: 10.1039/D5SC02315G

REEs. Seven unique domains representing different calcium-binding geometries were selected for evaluation. The results revealed a strong correlation between the charge of the binding loop and its affinity for REEs. We concluded that highly charged, aspartic acid-rich loops exhibit greater electrostatic repulsion, which creates higher affinity due to the increased stabilization effect of ion binding. Binding affinity across the lanthanide series was highest for ions with radii similar to that of calcium (~1 Å), consistent with the evolutionary optimization of calcium-binding proteins for selective ion recognition. While selectivity varied among proteins in solution, immobilized proteins demonstrated higher selectivity toward intermediate REEs. One notable candidate identified in the bioinformatic search was HEW5 from *Nocardioides zeae*. We leveraged the selectivity of HEW5 in a 7 mL column to demonstrate a single-stage, chelator-free separation of an equimolar lanthanum-neodymium mixture, achieving a high purity (>90%) and yield (90%) of REEs. Additionally, immobilized HEW5 was used to remove non-REE ions from a simulated leachate stream and separate lanthanum (>90% purity) from other REEs in a single separation stage.

Introduction

The use of rare earth elements (REEs) in energy, medicine, and defense applications makes them an indispensable resource.[1] REEs comprise the 15 lanthanides plus scandium and yttrium, which have similar physical and chemical properties, making separation processes considerably challenging.[1-3] Conventional separation methods rely on solvent extraction involving chelators, and due to the similarities of these ions, multiple separation stages are required to achieve high-purity individual REE products.[4-6] Solvent extraction of REEs is a chemical- and energy-intensive process that is only economical with high-grade feedstocks and generates large amounts of hazardous waste.[6, 7] The processing of REEs can have a substantial environmental impact,[8, 9] yet

View Article Online
DOI: 10.1039/D5SC02315G

the demand for REEs is expected to increase dramatically, with the International Energy Agency estimating the demand to triple by 2030.[3, 10] Therefore, new environmentally benign approaches are desired to satisfy future demand.

Biological approaches are attractive as they offer the potential to mitigate the environmental impact of current methods. Proteins can be readily engineered for customized functions, and to meet demand of large-scale applications.[11] It has been estimated that one-third of all proteins require a specific metal ion to perform their functions, and these binding interactions have evolved to be highly selective.[12, 13] Therefore, metalloproteins provide an exciting opportunity to develop new REE separation and recovery technologies.

Recently, a high-affinity lanthanide-binding protein,[14] lanmodulin (LanM), was identified from the methylotroph *Methylobacterium extorquens* following the discovery of a lanthanide-dependent methanol dehydrogenase in the same host.[15, 16] This *Mex*-LanM and, more recently, a *Hans*-LanM, are EF-hand domains with picomolar affinities to lanthanides that have demonstrated over a millionfold selectivity for lanthanides over non-REEs and some selectivity within the lanthanides series.[17] This has enabled single-stage separation of Nd and Dy ions.[18, 19] In multiple stages, *Mex*-LanM has showed the potential for high purity separation of Sc and grouped separation of other REEs. The extreme affinity of LanM for lanthanides is remarkable, but this can also lead to challenges for the elution of the REE ions. The release of bound ions requires sharp pH gradients or the use of strong chelators such as citrate or malonate.[20] This adds process costs and requires additional steps to separate and recover the ions from the chelators for recycling.

Trivalent lanthanide ions share physical and chemical similarities with divalent calcium ions, such as ionic radius, coordination geometry, and oxophilicity.[21, 22] These properties enable

View Article Online
DOI: 10.1039/D5SC02315G

lanthanides to replace calcium in many proteins, often with higher affinities.[22-24] Several studies have explored the loops of calmodulin (CaM) and other EF-hand domains and engineered them for lanthanide binding.[25-30] These peptides can be used to separate REEs from non-REEs and their moderate affinities compared to LanM enable simpler elution strategies. However, these peptides tend to lack selectivity within the lanthanide series. The most well-known example is the development of the lanthanide binding tag (LBT), which was created through a combinatorial screen of the calcium-binding loops of calmodulin.[29] These engineered LBTs exhibit nanomolar affinities for REEs and high selectivity within the lanthanide series, up to a 60-fold difference between La and Lu.[26, 31] LBTs have been successfully used for REE recovery via cell surface display.[32] However, in a chromatographic setting, their low binding capacity per peptide would necessitate large column volumes to achieve meaningful capture. Moreover, LBTs have been reported to lose approximately 50% of their binding capacity under moderately acidic conditions (pH 5).[32] While their selectivity is impressive, there are limited studies demonstrating effective separation of REE mixtures using LBTs in a chromatography format.

Recently, we reported that the Block V repeats-in-toxins (RTX) domain of adenylate cyclase from *Bordetella pertussis*, which has been engineered for calcium-induced hydrogel formation[33-35], calcium-controlled phase-separation peptides[36], and modulation of enzymatic activity[37]. RTX is also capable of binding REEs with high selectivity against non-REEs and can differentiate between the light lanthanides.[38-40] The ability of CaM and the RTX domain, two calcium-binding peptides with different binding modalities, to bind REEs and selectively separate them from non-REEs including calcium, motivates a broader search for calcium-binding peptides with alternative binding motifs. In this study, we explored a large library of calcium-binding domains to identify additional peptides capable of separating REEs. We performed a bioinformatic search to identify

View Article Online
DOI: 10.1039/D5SC02315G

potential high capacity, high affinity, and high selectivity REE binding domains. We characterized the most promising peptides and immobilized the best candidates for single-stage chelator-free separation of REEs. The most notable candidate was able to achieve high-purity and high-yield separation of La and Nd, two major components of bastnäsite ore leachate.

Results & Discussion

Bioinformatic search and candidate selection. The ability of calcium-binding peptides and proteins to coordinate lanthanides in their binding sites suggests that an enormous library of potential REE binders exists and can be explored. A bioinformatic search was performed, limited to the protein-coding genes of microorganisms from soils, since REEs may represent an underappreciated source of inorganic cofactors utilized by soil microbial communities.[41] The search results were refined to include protein-coding genes from genomes classified as thermophiles, acidophiles, or halophiles whose stability under extreme conditions may be advantageous in downstream industrial applications.[42-44] The protein sequences, retrieved from the proGenomes[45] database, were queried with profile hidden Markov models (HMMs)[46] constructed from seed sequences of known calcium/lanthanide binding domains (Table S1).

The search identified 3,035 calcium binding domains (CBDs), which consisted of 0.4% C2 domain, 3% epidermal growth factor (EGF), 3.3% Bγ-crystallin domain, 6.5% PPE-SVP, 42% Excalibur, and 45% thrombospondins (Figure S1). The search also identified over 19,000 pyrroloquinoline quinone (PQQ) calcium-binding domains, which were excluded due to the dependency of the metal coordination on the small molecule pyrroloquinoline quinone (PQQ). Three to five representative CBDs were chosen from each category, yielding 27 candidates. Of those, 14 candidates had high-confidence AlphaFold2[47] structures and predicted[48] to be

View Article Online
DOI: 10.1039/D5SC02315G

cytoplasmic. Seven representative domains were manually selected for characterization based on their similarity to other reported CBDs and their diverse binding geometries.

A0A7 (Uniprot ID: A0A7L4YJY0) was truncated from a 295 amino acid hypothetical protein from *Epidermidibacterium keratini*.[49] The parent protein comprises three repeating anti-parallel thrombospondin domains, and the selected sequence (S111–S196) forms an anti-parallel β-sheet domain with six calcium binding sites (Figure 1A). The calcium ions are modeled to bind in the loops connecting the β-strands with an octahedral binding geometry, where one of the coordinating oxygens extends from an adjacent loop.

HEW5 (Uniprot ID: A0A6P0HEW5) is a hypothetical protein presumed to regulate calcium ion concentrations in the cytosol of *Nocardioides zeae*[50]. The native protein is made of 137 amino acids, forming a β-sheet with eight octahedral calcium binding sites (Figure 1B). The truncated domain (P14–G125) excludes disordered regions not affiliated with calcium binding. The binding sites are highly conserved across the protein, with each loop bearing the amino acid sequence D(A/T)DGDGY(V/I/T/A)D. All of the coordinating oxygens in a binding site of HEW5 belong to the same binding loop, making the sites more independent than the binding sites in the A0A7 domain.

K3T(VV) and K3T(VN) (Uniprot ID: A0A1G7K3T7) were both truncated from a 400 amino acid hypothetical protein from *Ulvibacter litoralis*.[51] The protein is predicted to be a thrombospondin type 3 repeat and is made of a conserved sequence repeated in a chain 6.5 times with only slight variations. The N-terminal repeat (Unit A) was truncated to yield the K3T(VV) domain (V1–V63), and the C-terminal repeats (Units E-G) were trimmed to produce the K3T(VN) domain (V253–N402). The selected domains are aspartic acid-rich and bind four and eight calcium

View Article Online
DOI: 10.1039/D5SC02315G

ions, respectively. The domains lack secondary structure characteristics, and the calcium-binding sites are connected via disordered loop regions (Figure 1C and Figure 1D), unlike the β-strands in the A0A7 and HEW5 domains and α-helices in EF-hand domains. One key distinction between the two selected domains is a disulfide bridge linking the two units of the K3T(VN) domain.

HJH0 (Uniprot ID: A0A6P0HJH0) is a 95 amino acid Excalibur calcium-binding domain-containing protein from *Nocardioides zeae*.[50] Excalibur domains are distant relatives of EF-hand domains with a conserved calcium-binding loop DxDxDGxxCE.[52] The cysteines form a disulfide bond, giving rigidity to the binding loop. The selected domain (G52–R95) omits the disordered region outside the binding pocket and can bind a single calcium ion (Figure 1E).

CaM(III, IV) and the RTX domains were not involved in the bioinformatics search and were manually selected for the study. CaM(III, IV) (PDB: 1CFF[53]) was selected as a representative EF-hand domain. It was truncated from the full-length CaM protein (M76–K148), yielding a domain with two calcium ion binding sites (Figure 1F). The RTX domain (PDB: 5CVW[54]) is the Block V domain of the adenylate cyclase protein from *Bordetella pertussis*. It was selected for its recently reported ability to bind REEs.[38, 40] The RTX domain forms a β-roll structure upon binding eight calcium ions comprised of a tandem repeat of a nine amino acid sequence GGxGxDx(L/F/I)x. The bound ions experience pentagonal-bipyramidal coordination geometry with one oxygen extending from an adjacent loop (Figure 1G). The Block V RTX domain was not truncated further to exclude the non-calcium binding portion of the C-terminus as this capping group influences the calcium binding by entropically stabilizing the motif.[55] The oligonucleotides and amino acid sequences for all constructs are tabulated in Table S2 and Table S3.

View Article Online
DOI: 10.1039/D5SC02315G

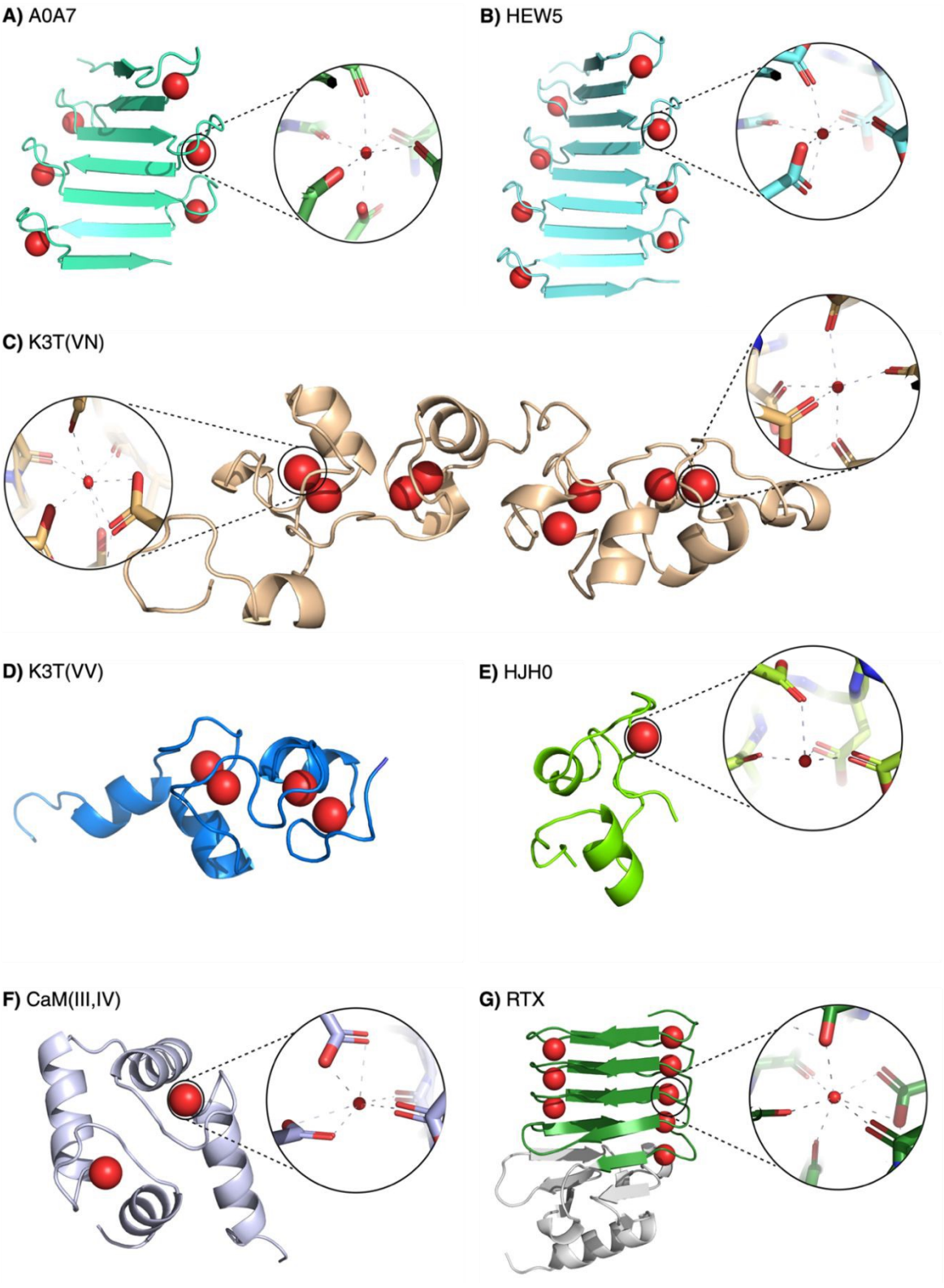

Figure 1: Predicted structures of the seven selected domains using AlphaFold3[56] with calcium ions (red spheres). A) A0A7 is from a repeating anti-parallel thrombospondin domain from

View Article Online
DOI: 10.1039/D5SC02315G

Epidermidibacterium keratini, B) HEW5 is a cytosolic calcium regulator from *Nocardioides zeae*, C) K3T(VN) and D) K3T(VV) are thrombospondin type 3 repeat domains from *Ulvibacter litoralis*, E) HJH0 is an excalibur calcium-binding domain-containing protein from *Nocardioides zeae*, CaM(III,IV) is an EF-hand domain truncated from full length calmodulin from *Xenopus laevis*, and G) RTX is the Block V domain of the adenylate cyclase protein from *Bordetella pertussis* (PDB ID 5CVW). PDB files for the predicted structures (A-F) can be found in the ESI.

Identification of high affinity domains. Peptide competition assays with a lanthanide responsive dye, xylenol orange (XO) (Kd = 1.3 ± 0.5 μM for Yb, Figure S2), revealed four sub-micromolar high-affinity sites for K3T(VN) and A0A7 domains, two for HEW5 and K3T(VV), and none for the CaM(III, IV), HJH0, and RTX domains at pH 6 (Figure 2, Table 1). The experiments were performed in a non-chelating (MES) buffer at pH 6 where REEs are soluble. The XO competition assays revealed two key observations. First, despite the similarity between the binding pockets of different proteins, their affinities for metal ions varied significantly. Second, even within a single protein with highly conserved binding loops, the predicted binding sites show varying affinities for the same ion. For instance, HEW5 is predicted to contain eight binding sites with the amino acid sequence D(A/T)DGDGY(V/I/T/A)D; however, the competition assays only indicated two sub-micromolar affinity sites. The affinity of a metalloprotein to an ion cannot be attributed to a single factor, and it is dependent on both the protein interactions and the metal ion. In the binding interaction, the metal ion and the sidechains in the binding loops are first dehydrated, releasing ordered water molecules, and then the protein and metal ion are associated.[57] Properties of the ions, such as the charge, hydration shell, and ionic radius play a role in the affinity of an ion and the specificity of a binding site.[58] Additionally, the charge of the binding loop, the composition of amino acids in the loop, and the overall structure of the proteins determine the affinity and specificity of the site.

View Article Online
DOI: 10.1039/D5SC02315G

The calcium-binding loops in many peptides are often disordered due to electrostatic repulsion from negatively charged aspartic and glutamic acids. The binding of the metal ion stabilizes the electrostatic repulsion and allows the protein to adopt a more compact structure.[57, 59-62] This stabilization correlates with increased binding affinities.[58, 63] For instance, A0A7 and HEW5, which have four aspartic acids per bind loops, showed higher affinities than the RTX domain, which has only one aspartic acid per loop. If the magnitude of the electrostatic stabilization is assigned to the isoelectric point of the protein, a clear pattern emerges: highly charged proteins with a low isoelectric point (A0A7, HEW5, K3T(VN), and K3T(VV)) have more high-affinity binding sites. In contrast, domains with higher isoelectric points (HJH0, CaM(III, IV), RTX) lacked high-affinity sites (Table 1). Additionally, entropic forces play a role in determining the affinities as upon metal ion binding, the ions and the sidechains in the binding pocket release their hydration shells, creating entropically favorable reactions. For instance, the binding of a lanthanide releases eight or nine water molecules, while calcium only releases six water molecules and this entropic driving force creates a higher affinity for REEs over calcium.[31] The composition of non-coordinating sidechains in the loop, which must be dehydrated in the binding reaction, and other structural elements of the protein also affect the affinities.

The coordination geometry of the ion is critical for determining both the selectivities and affinities of the binding sites.[57] The coordination geometries observed in the predicted structures for the selected domains varied between octahedral (CN=6, where CN stands for coordination number) and pentagonal-bipyramidal (CN=7), with all coordinating atoms being oxygens. Given potential errors in the structure prediction models (AlphaFold3), such minor variations in coordination geometries are unlikely to impact the affinities significantly. The charge of the coordinating oxygens, whether partial or full, and the presence of polydentate ligands likely play

View Article Online
DOI: 10.1039/D5SC02315G

a more prominent role in this case. The affinity gradients within the K3T(VN) and K3T(VV) can be explained by the charge of the binding loop. The domains have eight and four binding sites, respectively, while only half of the predicted sites are detected in the competition assay. A closer examination of the structure of the binding sites yields two distinctly different binding modes. The high-affinity binding sites are coordinated by seven oxygens in a pentagonal bipyramidal geometry with up to four fully charged oxygen ligands, while the other set of sites are coordinated by only five partially charged oxygen atoms (Figure 1C). For the other domains, sequential binding of the metal ions with cooperative binding could yield an affinity gradient within the protein.

Table 1: Overview of the characteristics and binding properties of selected proteins. The table is in descending order of maximum selectivity via FRET, determined as the maximum ratio of apparent k_d of any tested lanthanides. The error bars represent the standard deviations from three independent trials for xylenol orange and the propagated error (k_d ratio) for selectivity.

Protein	Uniprot ID	# Residue (MW kDa)	pI	# Theoretical Sites	Capacity (# Theoretical Sites/100 AA)	# High-Affinity Sites via Xylenol Orange	# Sites via ITC	Max. Selectivity via FRET ↓
HEW5	A0A6P0HEW5(P14–G125)	117 (12.1)	2.98	8	6.8	2.3 ± 0.5	8.6 ± 0.6	3.4 ± 1.1
RTX	PDB: 5CVW	152 (15.9)	3.96	8	5.2	0.47 ± 0.12	5.1 ± 0.6	3.1 ± 0.5
A0A7	A0A7L4YJY0 (S111–S196)	87 (9.3)	2.69	6	6.9	3.7 ± 0.1	7.8 ± 0.7	1.7 ± 0.1
K3T(VN)	A0A1G7K3T7(V253–N402)	150 (15.6)	2.82	8	5.3	4.2 ± 0.3	ND	1.7 ± 0.2
CaM(III,IV)	PDB: 1CFF (M76–K148)	74 (8.5)	4.16	2	2.7	0.55 ± 0.10	ND	1.7 ± 0.3
K3T(VV)	A0A1G7K3T7(V1–V63)	63 (6.6)	2.82	4	6.3	1.9 ± 0.1	ND	1.3 ± 0.2
HJH0	A0A6P0HJH0 (G52–R95)	43 (4.7)	3.85	1	2.3	0.14 ± 0.10	ND	1.1 ± 0.1

View Article Online
DOI: 10.1039/D5SC02315G

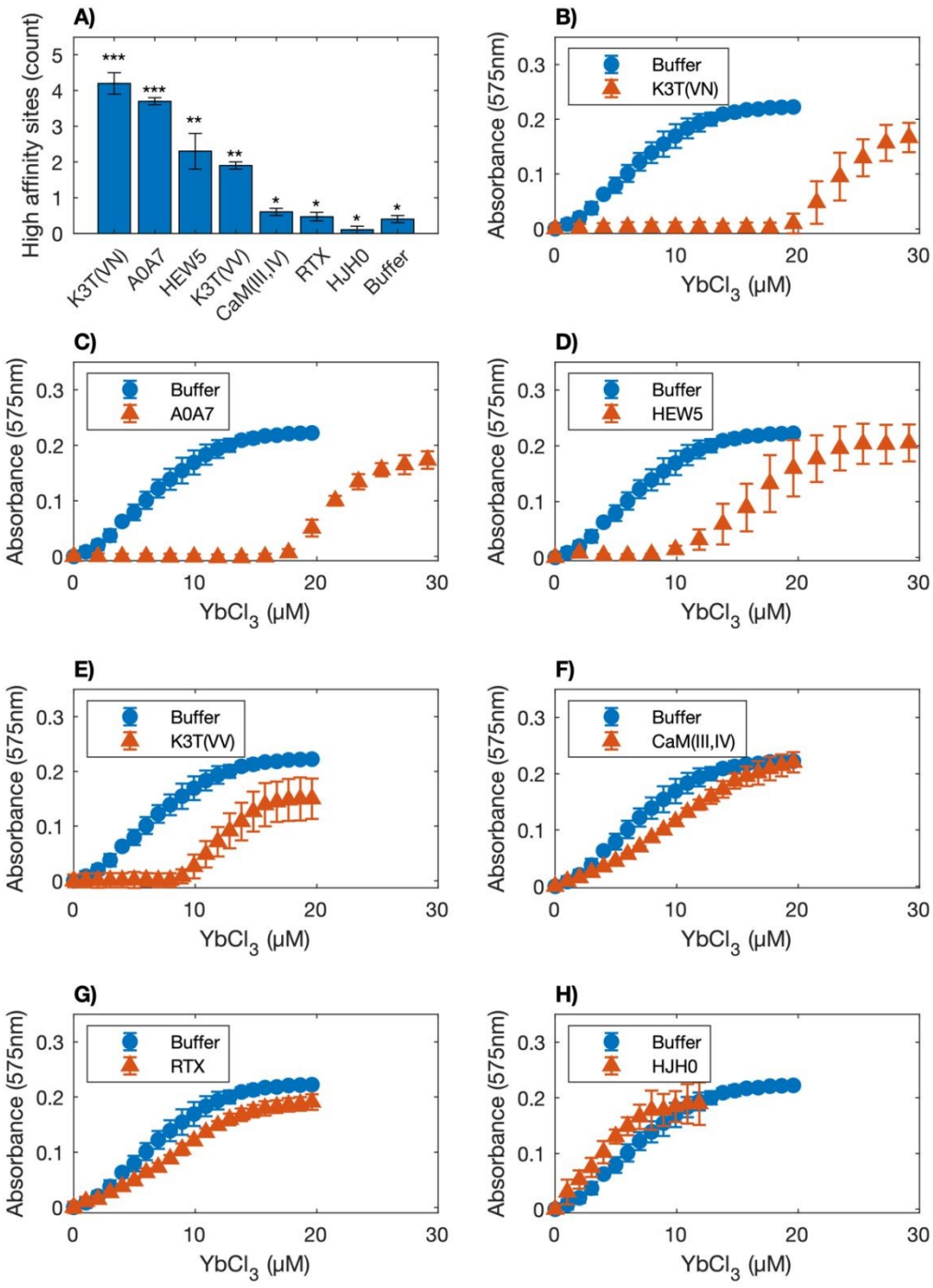

Figure 2: Competition experiments with xylenol orange (XO). A) Number of high affinity binding sites per domain. The (*), (**), and (***) represent three statistical groups where members of different groups have statistically significant differences between their means via a one-way ANOVA test (p-values are reported in Table S4). Titrations of XO with $YbCl_3$ for B) K3T(VN),

View Article Online
DOI: 10.1039/D5SC02315G

C) A0A7, D) HEW5, E) K3T(VV), F) CaM(III, IV), G) RTX, and H) HJH0. The error bars represent the standard deviations from three independent trials.

Identification of high selectivity domains. The selectivities of the proteins, defined as their ability to distinguish between ions within the lanthanide series, were initially tested using a high-throughput Förster resonance energy transfer (FRET)-based method utilizing the FRET pair cyan fluorescent protein (CFP) and enhanced yellow fluorescent protein (EYFP). The apparent dissociation constants ($k_{d,,apparent}$) were determined for $CeCl_3$ and $NdCl_3$ (light REEs), $DyCl_3$ (intermediate-heavy REE), and $YbCl_3$ (heavy REE) by fitting the normalized FRET efficiencies to the Hill equation (Table S5, Figure S3-S9). The maximum selectivities (apparent k_d ratio) between any two lanthanides for each domain are reported in Table 1. HEW5 and the RTX domains had the highest selectivities, with opposite preferences for light and heavy REEs (Figure 3). The FRET experiments were conducted under physiological pH conditions (Tris-HCl buffer pH 7.4) because the fluorescent proteins signals were significantly degraded at pH 6.

The maximum selectivity for the domains, measured by FRET, ranged from 1.1 for HJH0 to 3.4 for HEW5 (Table 1). The *Mex*-LanM had a maximum selectivity of about 3.0 when comparing the same ions in this study via the CFP/YFP FRET system,[18, 64] indicating HEW5 has a higher selectivity amongst the REEs screened. Therefore, the selectivity results are encouraging for protein-based separation of REEs.

While the CFP/YFP FRET method enables high-throughput screening of apparent affinities of REE-binding proteins at physiological pH, it is essential to recognize the limitations. The two large proteins (CFP and YFP) on the ends of the peptides may affect folding, which in turn impacts the affinities and selectivities of the binding sites, as previously reported for the entropic effect of the

View Article Online
DOI: 10.1039/D5SC02315G

C-terminus capping group of the RTX.[55] Additionally, although no precipitation was observed, heavy REEs are more prone to form insoluble hydroxides at physiological pH (Visual MINTEQ).

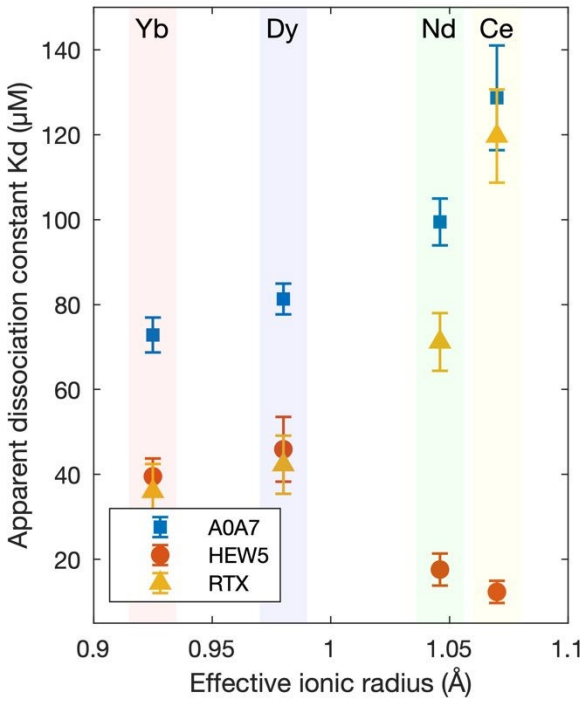

Figure 3: The apparent dissociation constants, determined via FRET, as a function of the effective ionic radius for the three most selective domains (A0A7, HEW5, and RTX). The error bars represent 95% confidence intervals.

View Article Online
DOI: 10.1039/D5SC02315G

Down-selection of candidates for chromatography application. Proteins can be covalently immobilized on chromatography resin for protein-based separation of REEs. However, the commercially available resins are expensive and often have low loading capacities, typically ranging from 5 to 30 mg of protein per mL of hydrated resin.[65] To compensate, the selected protein should be a high-capacity binding domain. Additionally, tight binding between the ligand and protein is essential to prevent premature elution, which can lead to product loss and reduce purity. Most importantly, the protein must have sufficient selectivity to separate REEs. Because the affinity and selectivity of the protein will vary as a function of pH and ionic strength, the optimal conditions for each protein need to be determined. This study focused on using a pH gradient to achieve the separation while avoiding costly chelators.

HJH0 and CaM(III, IV) were capable of binding the tested lanthanides as indicated by FRET, but neither domain exhibited high-affinity binding sites via XO competition assay. HJH0 could not differentiate between the tested ions, having the lowest measured selectivity (1.1 ± 0.1). In contrast, CaM(III, IV) demonstrated moderate selectivity (1.7 ± 0.3) with a preference for light REEs. The binding capacity of both proteins was the lowest among the domains, with only 2.5 binding sites per 100 amino acids. Consequently, HJH0 and CaM(III, IV) were excluded from further testing on chromatography columns.

K3T(VN) and K3T(VV) exhibited moderate selectivities, ranging from 1.3 to 1.7, with a preference for heavy REEs. These domains are high-capacity binders, with 5-6 binding sites per 100 amino acids, and contain multiple high-affinity binding sites. Despite meeting all the requirements, they present multiple binding modes that could complicate chromatography column performance. Specifically, half of the ions in these domains are coordinated in a pentagonal-bipyramidal (CN=7) geometry with multiple bidentate aspartic acids. The other half are

View Article Online
DOI: 10.1039/D5SC02315G

coordinated in a square-pyramidal (CN=5) geometry, with all coordinating oxygens being monodentate and partially charged. This variation in coordination could lead to differences in affinity, as indicated by XO results, and selectivity between sites, potentially resulting in broad and multipeak elution profiles in chromatography. For these reasons, the domains were not pursued for further characterization. However, their unique properties make them worthy of separate study.

A0A7, HEW5, and RTX domains were selected for further characterization and immobilization on chromatography resin. The three domains are high-capacity binders, with 5-7 binding sites per 100 amino acids, while EF-hand domains average less than three binding sites per 100 amino acids. The domains demonstrate high selectivity ranging from 1.7 to 3.4, and A0A7 and HEW5 display multiple high-affinity binding sites. Additionally, the binding loops of the domains are highly conserved and may be less likely to exhibit variation in selectivity across the protein. The repetitiveness and modularity of the sequences would enable the introduction of additional binding loops to increase the binding capacity per protein, as previously reported for the RTX domain.[66] The HEW5 and RTX are the most selective and have opposing preferences for light and heavy REEs, respectively, presenting an opportunity to use different scaffolds tailored to the composition of the source material.

Further characterization of A0A7, HEW5, and RTX via CD and ITC. The secondary structures of the peptides were examined via circular dichroism (CD) spectroscopy at pH 6. The domains appeared to be intrinsically disordered in their apo-state, characterized by the random coil peak at 200 nm (Figure 4). Upon the addition of calcium or neodymium, a broad peak developed between 210–220 nm, indicating the formation of β-sheets, which supports the AlphaFold3 predicted models (Figure 1). The peak at 200 nm largely remained after the addition of ions for both A0A7 and HEW5 despite the formation of β-sheets; the remaining peak is likely due to the

View Article Online
DOI: 10.1039/D5SC02315G

high loop content of the proteins or the presence of unordered regions. The spectra with titrated concentrations of calcium and neodymium are presented in SI (Figure S10). Significantly lower concentrations of Nd compared to Ca were needed to induce conformational changes in the domains, confirming that the domains have a higher affinity for lanthanides than for calcium.

View Article Online
DOI: 10.1039/D5SC02315G

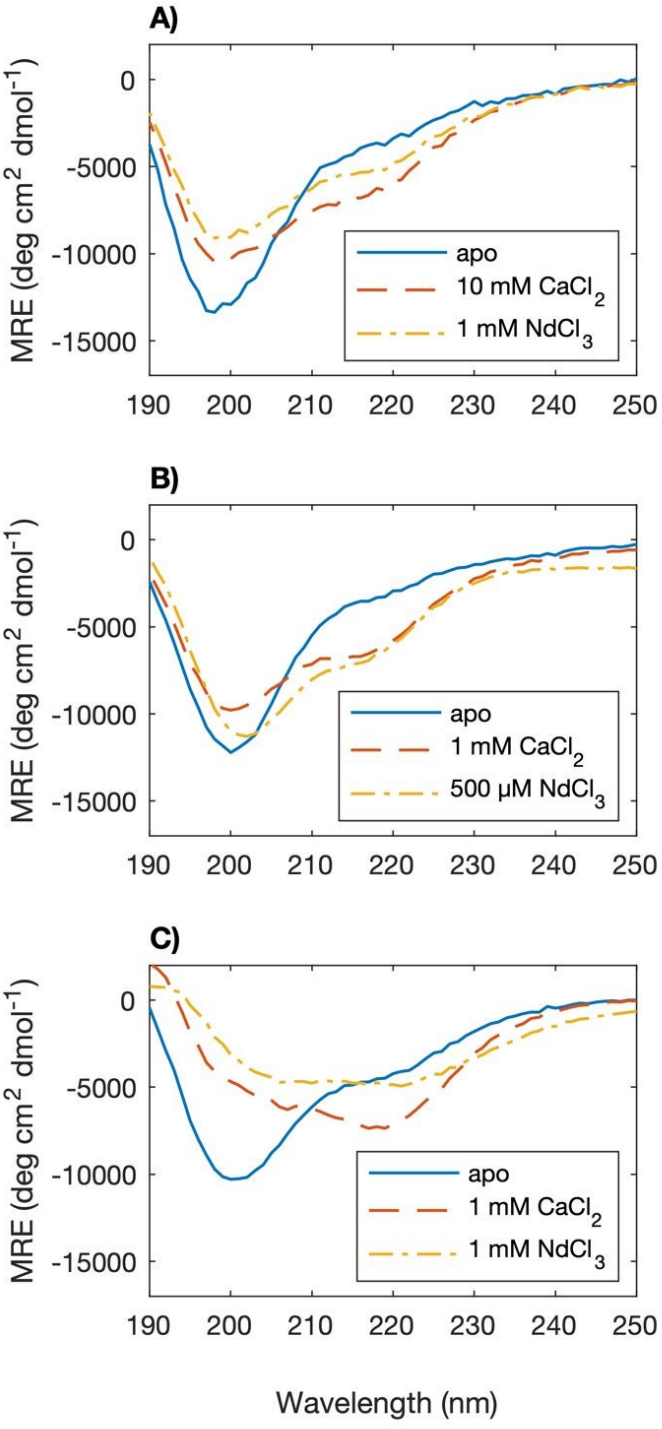

Figure 4: CD spectra of the proteins (60 µM) in their apo (—), Ca-bound (— —), and Nd-bound (— -) states for **A)** A0A7, **B)** HEW5, and **C)** RTX. The reported spectra are the means of three scans.

View Article Online
DOI: 10.1039/D5SC02315G

While the binding affinities were determined using a FRET-based method described above, the complications of having two larger proteins on either side and the requirement to run CFP/YFP FRET at near physiological pH conditions raised the need to pursue other direct routes with the unmodified peptides. Isothermal titration calorimetry (ITC) can be used to more directly measure the interaction between the ion and the binding site. The affinities of the down-selected proteins to REEs for nine lanthanide ions spanning light, intermediate, and heavy REEs in MES buffer (pH 6) were determined. HEW5 had the highest affinity with average dissociation constants ranging from 5–40 μM, with a preference for light REEs (Figure 5A). In contrast, the RTX domain preferred intermediate and heavy REEs, with the highest affinity for europium. The results for the RTX domain agree with previously determined apparent dissociation constants at pH 6 via FRET.[40] A0A7, similarly to HEW5, preferred light REEs and ranged from 10-70 μM, with the highest affinity for cerium and praseodymium. It is notable that the affinity of A0A7 to lanthanides had a different trend via ITC from what was measured via FRET, which could be attributed to change in pH conditions or that ITC measures the heat of the interaction between the amino acids and the ion while FRET measure the induced conformational change (apparent dissociation constants).

At low molar ratios, an initial increase in the enthalpy was observed. A similar response was reported for LanM peptides and other metalloproteins and is attributed to a large conformational change following the binding of the first ion.[67, 68] The data points were excluded from the isotherm fits; therefore, the estimated dissociation constants for A0A7 and RTX do not account for the first site. In case of negative cooperativity, as anticipated for the RTX, the affinity would be underestimated by ITC. The overall selectivity, the highest to lowest calculated lanthanide

View Article Online
DOI: 10.1039/D5SC02315G

dissociation constants ratio, was similar for the three domains and ranged from five to eight (Figure 5B). A summary of the calculated thermodynamic properties and individual isotherm fits are presented in SI (Table S6, Figure S11-S17).

Calcium interaction with the proteins could not be studied by ITC due to the weak affinity between the protein and calcium at pH 6. The resultant isotherms were incomplete sigmoidal curves (Figure S13, Figure S15, and Figure S17), and the dissociation constants were undetermined. The required amount of protein to increase the c-value ([protein]/k_d) to produce a suitable isotherm was prohibitively high. Therefore, the dissociation constants for the interactions between the proteins and calcium were estimated as the concentration of ions needed to induce conformational change in CD. The lanthanide to calcium selectivity, reported as the ratio of calcium ion required to induce a conformational change in CD to the average dissociation constant for lanthanide, was the highest for A0A7 and HEW5 at 60-fold selectivity for lanthanides over calcium (Figure 5B). In contrast, the selectivity of the RTX domain for lanthanides was only 15-fold over calcium. The results illustrate the ability of the domains to distinguish REEs from each other and to separate them from contaminants, including calcium.

View Article Online
DOI: 10.1039/D5SC02315G

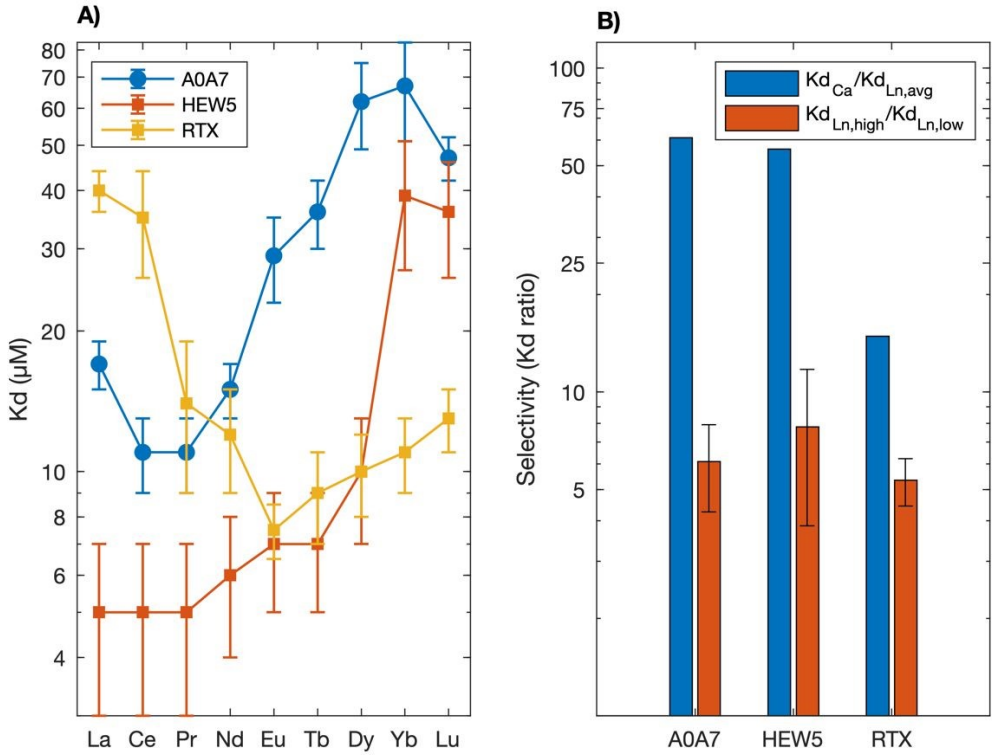

Figure 5: A) Average dissociation constants determined via ITC for A0A7, HEW5, and RTX. The error bars represent 95% confidence intervals. B) Selectivity of A0A7, HEW5, and RTX for lanthanides over calcium and maximum selectivity within the lanthanide series. The error bars represent the propagated error from the ratio of dissociation constants. Calcium to lanthanide ratios do not have error bars due to estimating the calcium dissociation constants as the concentration needed to induce a conformation change via CD.

The binding of the lanthanides to the proteins was an entropically driven interaction (Figure S11), and a similar behavior was observed for LBT.[31] The finding was expected since the dehydration of the ions and the sidechains, releasing water molecules into the solvent, and the burial of hydrophobic residues due to the folding of the disordered domains are entropically favorable processes. The REE ions and side chains are highly hydrated in solution, and the disruption of their tightly bound water shells during binding imposes a significant enthalpic

View Article Online
DOI: 10.1039/D5SC02315G

penalty, which must be entropically overcome to form the protein–ion complex. Interestingly, these domains evolved to bind calcium and have a preference for REEs. A0A7 and HEW5 prefer light REEs, while RTX showed a higher affinity for intermediate to heavy REEs. Replotting the dissociation constants (Figure 5A) as a function of the effective ionic radius,[69] where the radius of each ion is taken at appropriate coordination number (CN=6 for A0A7 and HEW5, and CN=7 for RTX, where CN stands for coordination number), the highest affinity for all three domains emerges at the same ionic radius (Figure 6), which closely matches the ionic size of calcium (~1 Å). These results suggest that the size of the ions is a major determining factor for the preference. It is important to note that coordination numbers were predicted via AlphaFold3 and should be confirmed by experimental validation.

View Article Online
DOI: 10.1039/D5SC02315G

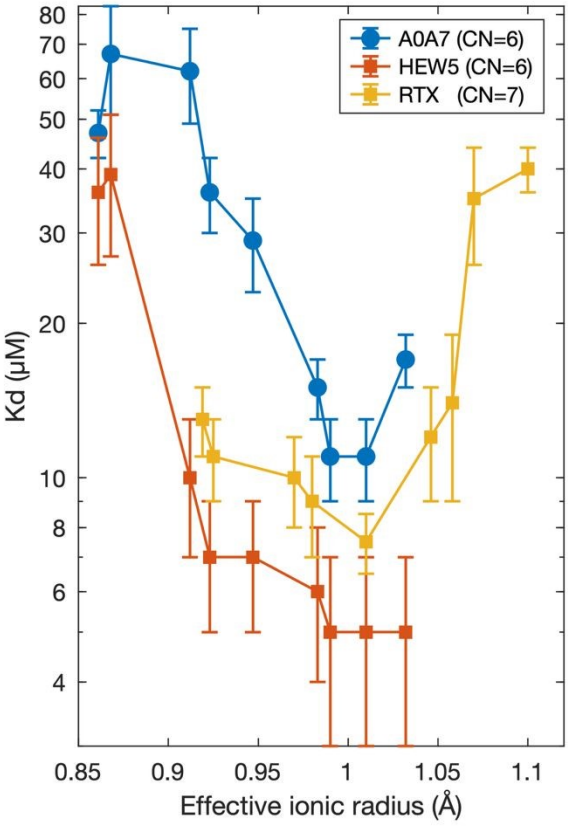

Figure 6: Dissociation constants as a function of effective ionic radius.[69] Average dissociation constants determined via ITC for A0A7, HEW5, and RTX. The coordination numbers (CN) were determined from AlphaFold3 predicted structures. The error bars represent 95% confidence intervals.

Characterization of immobilized proteins. The immobilization of proteins limits the conformational space accessible to unfolded proteins and reduces the entropic cost associated with folding.[70] Since the folding of the three domains with REEs is entropically driven, it is possible that the affinities and selectivities of the binding sites may differ from those measured in the bulk (non-immobilized) state. To study the effect of immobilization on the selectivities of the proteins,

View Article Online
DOI: 10.1039/D5SC02315G

the immobilized apparent dissociation constants for an equimolar mixture of five REEs (La^{+3}, Nd^{+3}, Sm^{+3}, Dy^{+3}, and Y^{+3}) were determined. The preference of A0A7 and HEW5 shifted from light REEs in the bulk state to intermediate REEs in the immobilized state (Figure 7). The immobilized apparent dissociation constants were calculated as the ratio of the total moles of binding sites (B_t, Figure S18) in the column (17 µmol) to the mean retention volumes (V_R),[71] which were determined by fitting the individual elution peaks to a normal distribution (Figure S19).

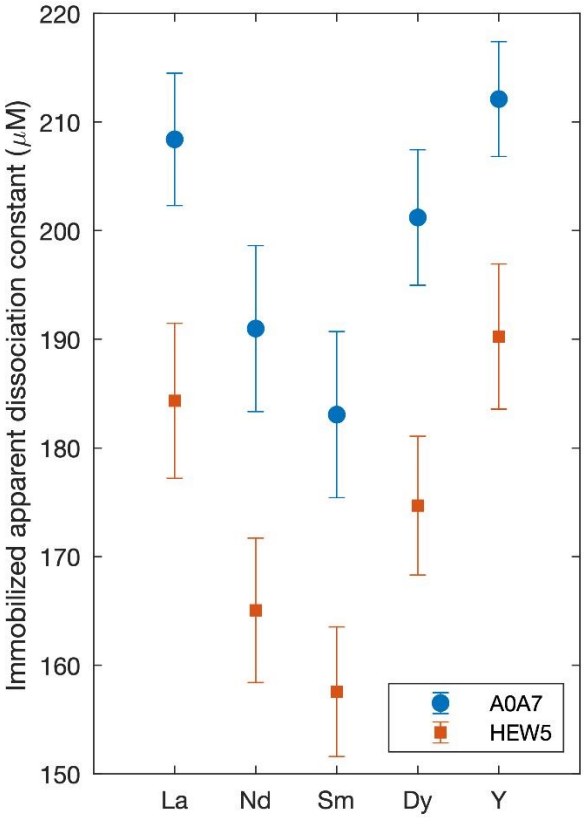

Figure 7: Immobilized apparent dissociation constants measured under isocratic elution conditions (pH 3) for A0A7 and HEW5. The error bars represent the propagated standard deviations from three independent trials. The uncertainties stem from the experimentally determined capacity ($\boldsymbol{B_t}$) and the standard deviations from the retention peaks ($\boldsymbol{V_R}$) of each

View Article Online
DOI: 10.1039/D5SC02315G

element. Triplicate measurements of apparent dissociation constants and elution profiles are presented in Supporting Information (Figure S19).

The elution profiles of the A0A7 and HEW5 domains were studied in MES buffer at pH 6, under the same conditions as the ITC experiment. However, as anticipated from the XO competition experiments, the tight binding of the two domains prevented facile REE elution. Optimal conditions for REE elution from these domains were found around pH 4, which is near the pKa of aspartic acid. In contrast, elution from the RTX domain began at just below pH 5. The RTX domain demonstrated poor separation efficacy, exhibiting broad and multipeak elution profiles (Figure S20). This behavior is likely due to a large affinity gradient of the binding sites within the RTX domain, as previously observed in studies using calcium.[38, 54] Additionally, ITC revealed a stoichiometry of five functional binding sites out of eight theoretical sites in the RTX domain (Table 1), suggesting that the remaining three weaker or non-functional sites may contribute to premature elution.

The selectivity of A0A7 and HEW5 shifted from a preference for light REEs in solution to intermediate REEs upon immobilization. This shift may be attributed to the entropic effect of anchoring the protein, the adoption of a new conformational state (which could cause an increase in coordination number favoring heavier REEs, as discussed in Figure 6), a shift in second-shell coordinating amino acids, and other factors such as column properties, pH, and environmental conditions. A recent study on the binding of lanthanides to the methanol dehydrogenase (*Xoxf*) using DFT showed a higher affinity to intermediate REEs, which is attributed to the higher spin state of these ions.[72] This change in selectivity creates an affinity gradient within the light REE series, which can be exploited for separation. Light REEs constitute the majority of REEs found

View Article Online
DOI: 10.1039/D5SC02315G

in bastnäsite, the most common rare earth source, according to the U.S. Geological Survey.[73] Since the RTX domain demonstrated poor separation efficacy under the conditions tested, and A0A7 and HEW5 exhibit similar selectivities, HEW5 was chosen for further separation tests due to its higher capacity and stronger affinity.

Separation of lanthanum and neodymium. Lanthanum and neodymium are the dominant components of bastnäsite leachate following cerium removal.[74] In a single stage, immobilized HEW5 separated an equimolar mixture of lanthanum and neodymium, achieving over 90% purity and yield for each element (Figure 8). The ability of HEW5 to separate the binary mixture with both high yield and purity in a single stage highlights its potential as an alternative to solvent extraction methods. The La/Nd mixture was loaded onto HEW5 column in MES buffer (pH 6) at nearly 25% of the column REE binding capacity, and the ions were eluted from the protein using a step gradient to pH 3 MES buffer. Although pH 3 falls outside the buffering range of MES, it was chosen to maintain consistency in the chemical composition of the buffers throughout the study.

View Article Online
DOI: 10.1039/D5SC02315G

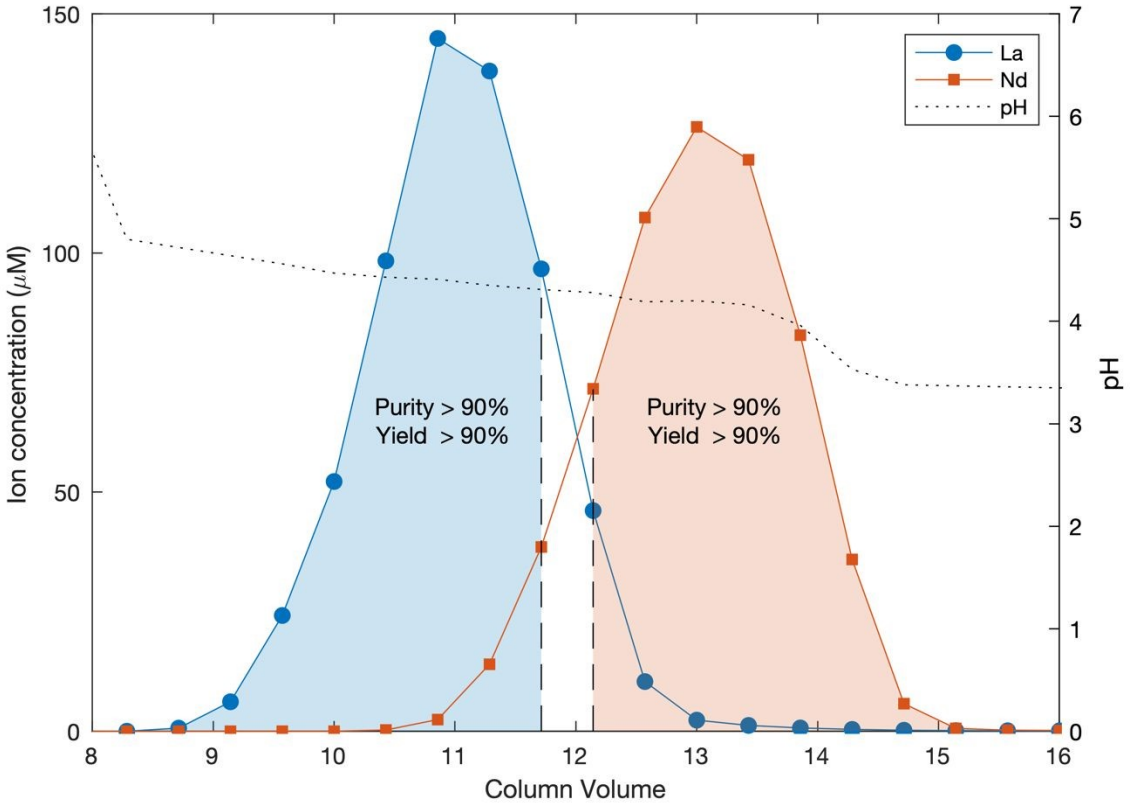

Figure 8: Single-stage separation of an equimolar mixture of lanthanum ($LaCl_3$) and neodymium ($NdCl_3$) using immobilized HEW5 (7-mL column volume). Shaded areas represent pooled fractions used for calculating purity (>90%) and yield (>90%) of each element. In this context, purity is defined as the molar percentage of the specified ion in the pooled fractions, while yield represents the molar ratio of the pure fraction to the total loaded onto the column. Results are representative of triplicate trials; additional details are provided in Supporting Information (Figure S21).

Bastnäsite ores are the primary source of REEs and are particularly rich in light lanthanides. The mine concentrate, or "refined ore," primarily contains cerium (Ce), lanthanum (La), neodymium (Nd), and praseodymium (Pr), along with minute amounts of heavier REEs such as samarium (Sm), europium (Eu), gadolinium (Gd), and yttrium (Y). It also includes significant

View Article Online
DOI: 10.1039/D5SC02315G

quantities of non-REEs, mainly calcium (Ca) and strontium (Sr).[75] Commonly, concentrated hydrochloric acid is used to leach the ore, releasing REEs into two streams: a cerium dioxide-rich stream and a REE chloride stream with lower cerium concentration.[74] These elements are separated from impurities and one another through multiple stages of solvent extraction—an energy-intensive and environmentally taxing process.[6-8]

To further demonstrate the capabilities of this platform, a synthetic solution simulating the chloride stream from the hydrochloric acid leaching of ores (Table S7) was separated using HEW5 with a loading of approximately 15% of the column capacity. Non-REE ions "impurities" were effectively separated from REE ions by initially operating the column at pH 5.5 where REEs remained bound to the column (Figure 9A, Zone A). A linear elution gradient from pH 4.75 to pH 3.5 was then applied followed by constant elution at pH 3.0. Lanthanum ($LaCl_3$) was separated from the remaining REEs (Zone B) with a purity ranging from 92–96% and a yield exceeding 80% (Figure 9B). The remaining REEs were group-eluted under pH 3.5 elution conditions (Zone C). In conventional processes, removing impurities alone would require multiple solvent extraction stages. Here, HEW5 demonstrates its ability to achieve impurity removal and separation of the most abundant REE in the mixture in a single stage, highlighting its efficiency and potential for simplifying REE separation workflow.

View Article Online
DOI: 10.1039/D5SC02315G

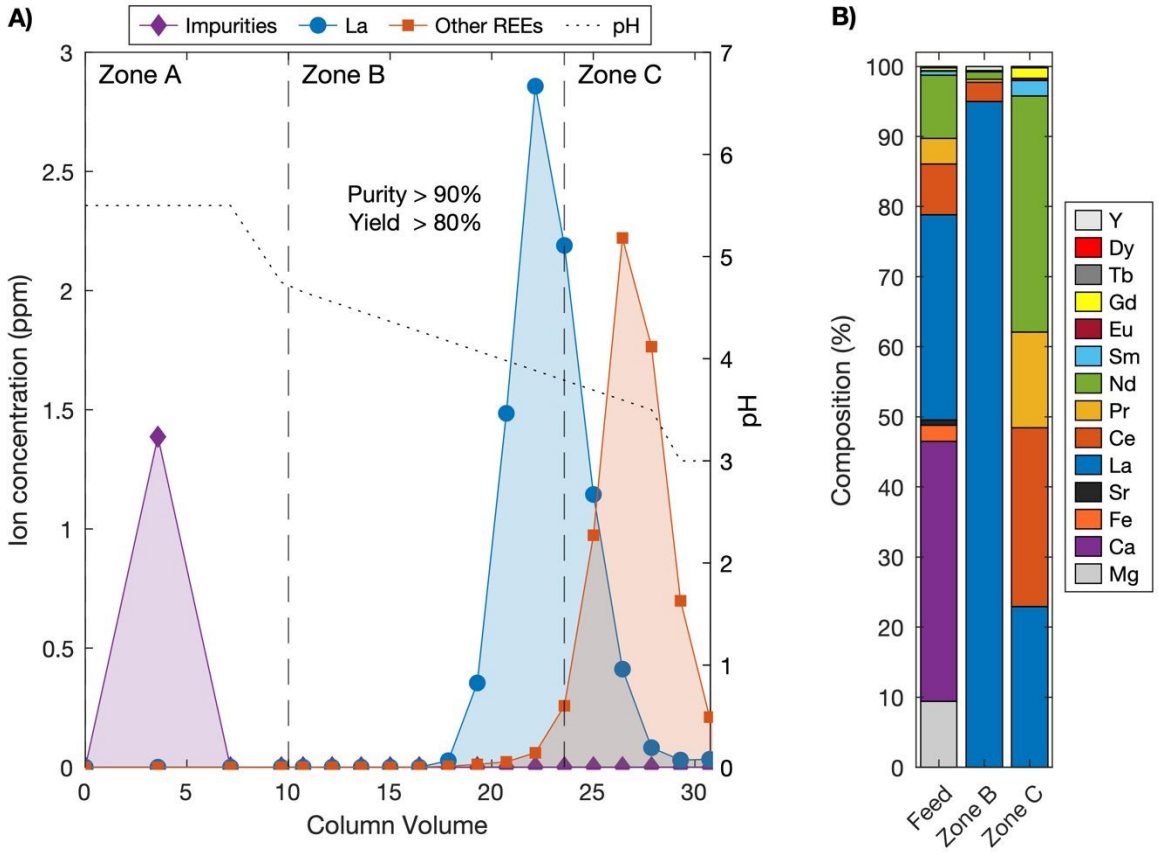

Figure 9: A) Single-stage separation of lanthanum from a simulated bastnäsite leachate solution using immobilized HEW5 achieves high yield and purity. **B)** Molar composition of Zone B, showing significant lanthanum enrichment (~95%) compared to the feed, which contains less than 30% lanthanum. Results are representative of triplicate trials; additional details are provided in Supporting Information (Figure S22).

View Article Online
DOI: 10.1039/D5SC02315G

Conclusions

REEs are critical for developing green energy technologies, such as magnets used in wind turbines and electric vehicles. However, due to their highly similar properties, separating REEs requires many stages of solvent extraction—processes that are both chemically and energy-intensive and generate large amounts of toxic waste.

The lanthanides share physical and chemical similarities with calcium ions, such as ionic size and coordination geometry, which allow REEs to substitute for calcium ions in calcium-binding proteins. This property opens the door to utilizing the extensive, well-studied library of calcium-binding proteins for protein-based separation of REEs.

In this study, we performed a bioinformatic search to identify promising candidates for REE binding and separation. The search results were narrowed to seven domains representing different coordination geometries of calcium-binding proteins. *In vitro* testing revealed a strong correlation between the charge of the binding loop (isoelectric point) and its affinity for REEs. We concluded that highly charged, aspartic acid-rich binding loops exhibit greater electrostatic repulsion, influencing the stabilization energy upon ion binding and correlating with affinity. Selectivity within the lanthanide series appeared to be a function of ionic size, with optimal binding near 1 Å, corresponding to the ionic size of calcium. While selectivity varied among proteins in solution, immobilized proteins demonstrated higher selectivity toward intermediate REEs. This behavior may be attributed to the higher valency states of intermediate REEs, entropic effects of immobilization, or the increased prevalence of carboxyl groups (aspartic acids) contributing to selectivity. Interestingly, a similar selectivity trend to the formation constants for acetate (Visual

View Article Online
DOI: 10.1039/D5SC02315G

MINTEQ) within the REE series was observed, which might suggest that REE preference is due to the strength of the REE-carboxylate interaction.

One notable candidate identified in the bioinformatic search was HEW5, a protein presumed to regulate calcium ion concentrations in the cytosol of *Nocardioides zeae*. HEW5 is a high-capacity binder, with nearly seven binding sites per 100 amino acids compared to fewer than three in EF-hand domains. Competition experiments with XO revealed sub-micromolar affinity of HEW5 for lanthanides and the highest selectivity within the lanthanide series among the tested proteins via FRET and ITC.

We leveraged the selectivity of HEW5 for a single-stage, chelator-free separation of an equimolar lanthanum-neodymium mixture, two major components of bastnäsite leachate. This process achieved high purity (>90%) and yield (90%) for both elements. Additionally, immobilized HEW5 was employed to remove non-REEs from a simulated leachate stream and separate lanthanum (>90% purity) from other REEs in a single stage under mild conditions.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its ESI. This includes the materials and methods section and data for FRET, CD, and ITC.

Author Information

Corresponding authors

View Article Online
DOI: 10.1039/D5SC02315G

Scott Banta – Department of Chemical Engineering, Columbia University, New York, NY, 10027. ORCID: 0000-0001-7885-0150

Katarzyna H. Kucharzyk – Battelle Memorial Research Institute, Columbus, OH, 43201. ORCID: 0000-0001-8524-2989

Authors

Farid F. Khoury – Department of Chemical Engineering, Columbia University, New York, NY, 10027. ORCID: 0000-0002-1116-6941

Bradley S. Heater – Battelle Memorial Research Institute, Columbus, OH, 43201. ORCID: 0000-0001-8851-4345

Daniel R. Marzolf – Battelle Memorial Research Institute, Columbus, OH, 43201.

Sameera Abeyrathna – Department of Chemical Engineering, Columbia University, New York, NY, 10027. ORCID: 0000-0002-1346-6796

Jonathan W. Picking – Battelle Memorial Research Institute, Columbus, OH, 43201.

Piyush Kumar – Department of Chemical Engineering, Columbia University, New York, NY, 10027. ORCID: 0009-0001-7277-7965

Steven A. Higgins – Battelle Memorial Research Institute, Columbus, OH, 43201. ORCID: 0000-0002-5209-5000

Randy Jones – Battelle Memorial Research Institute, Columbus, OH, 43201.

Alan T. Lewis, Jr. – Battelle Memorial Research Institute, Columbus, OH, 43201.

View Article Online
DOI: 10.1039/D5SC02315G

Author Contributions

S.B., K.H.K., B.S.H., and F.F.K. conceived this work. F.F.K. performed protein characterization experiments and data analysis. D.R.M. and S.A.H. performed bioinformatics search and in silico screening. B.S.H., S.A., and J.W.P. developed immobilization strategies and optimization of operation parameters. P.K. performed in vitro screening. B.S.H., J.W.P., R.J., and A.T.L., performed separation experiments and ICP analysis. S.B. and K.H.K. provided research supervision. All authors contributed to the writing and editing of the paper.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

This work was supported by DARPA's Environmental Microbes as a BioEngineering Resource (EMBER) program (FA8650-22-C-7213). The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. F.K. was supported by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP DGE-2036197) and the Columbia University Blavatnik Presidential Fellowship. We would like to thank Dr. Jerry Chang for providing access and training on ITC, and Wei Wang, Steven Martinez, and Christos Nikas for their assistance with FRET experiments. We also thank Mallori Herishko for her help proofreading and improving the readability of the manuscript. This research used the Chirascan V100 Circular Dichroism Spectrometer at The Precision Biomolecular Characterization Facility

View Article Online
DOI: 10.1039/D5SC02315G

(PBCF) in the Chemistry Department at Columbia University supported by NIH award (1S10OD025102-01).

Notes and References:

1. V. Balaram, Rare earth elements: A review of applications, occurrence, exploration, analysis, recycling, and environmental impact, *Geoscience Frontiers*, 2019, **10**, 1285-1303.
2. E. Elbashier, A. Mussa, M. Hafiz and A. Hawari, Recovery of rare earth elements from waste streams using membrane processes: An overview, *Hydrometallurgy*, 2021, **204**.
3. S.-L. Liu, H.-R. Fan, X. Liu, J. Meng, A. R. Butcher, L. Yann, K.-F. Yang and X.-C. Li, Global rare earth elements projects: New developments and supply chains, *Ore Geology Reviews*, 2023, **157**, 105428.
4. M. Gergoric, A. Barrier and T. Retegan, Recovery of Rare-Earth Elements from Neodymium Magnet Waste Using Glycolic, Maleic, and Ascorbic Acids Followed by Solvent Extraction, *Journal of Sustainable Metallurgy*, 2019, **5**, 85-96.
5. S. Pavón, A. Fortuny, M. T. Coll and A. M. Sastre, Neodymium recovery from NdFeB magnet wastes using Primene 81R·Cyanex 572 IL by solvent extraction, *Journal of Environmental Management*, 2018, **222**, 359-367.
6. Z. Chen, Z. Li, J. Chen, P. Kallem, F. Banat and H. Qiu, Recent advances in selective separation technologies of rare earth elements: a review, *Journal of Environmental Chemical Engineering*, 2022, **10**, 107104.
7. H. Pereira Neves, G. Max Dias Ferreira, G. Max Dias Ferreira, L. Rodrigues de Lemos, G. Dias Rodrigues, V. Albis Leão and A. Barbosa Mageste, Liquid-liquid extraction of rare earth elements using systems that are more environmentally friendly: Advances, challenges and perspectives, *Separation and Purification Technology*, 2022, **282**, 120064.
8. P. Zapp, A. Schreiber, J. Marx and W. Kuckshinrichs, Environmental impacts of rare earth production, *MRS Bull*, 2022, **47**, 267-275.
9. J. Bai, X. Xu, Y. Duan, G. Zhang, Z. Wang, L. Wang and C. Zheng, Evaluation of resource and environmental carrying capacity in rare earth mining areas in China, *Sci Rep*, 2022, **12**, 6105.
10. I. E. Agency, *Outlook for key energy transition minerals, rare earth elements*, International Energy Agency, 2024.
11. G. O. Ndochinwa, Q. Y. Wang, N. O. Okoro, O. C. Amadi, T. N. Nwagu, C. I. Nnamchi, A. N. Moneke and A. S. Odiba, New advances in protein engineering for industrial applications: Key takeaways, *Open Life Sci*, 2024, **19**, 20220856.
12. K. J. Waldron, J. C. Rutherford, D. Ford and N. J. Robinson, Metalloproteins and metal sensing, *Nature*, 2009, **460**, 823-830.
13. T. Dudev and C. Lim, Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins, *Chem Rev*, 2014, **114**, 538-556.
14. J. Cotruvo, E. Featherston, J. Mattocks, J. Ho and T. Laremore, Lanmodulin: A Highly Selective Lanthanide-Binding Protein from a Lanthanide-Utilizing Bacterium, *Journal of the American Chemical Society*, 2018, **140**, 15056-15061.

View Article Online
DOI: 10.1039/D5SC02315G

15. J. T. Keltjens, A. Pol, J. Reimann and H. J. Op den Camp, PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference, *Appl Microbiol Biotechnol*, 2014, **98**, 6163-6183.
16. F. Chu and M. E. Lidstrom, XoxF Acts as the Predominant Methanol Dehydrogenase in the Type I Methanotroph Methylomicrobium buryatense, *J Bacteriol*, 2016, **198**, 1317-1325.
17. G. J. Deblonde, J. A. Mattocks, D. M. Park, D. W. Reed, J. A. Cotruvo and Y. Jiao, Selective and Efficient Biomacromolecular Extraction of Rare-Earth Elements using Lanmodulin, *Inorg Chem*, 2020, **59**, 11855-11867.
18. Z. Dong, J. A. Mattocks, G. J. Deblonde, D. Hu, Y. Jiao, J. A. Cotruvo and D. M. Park, Bridging Hydrometallurgy and Biochemistry: A Protein-Based Process for Recovery and Separation of Rare Earth Elements, *ACS Cent Sci*, 2021, **7**, 1798-1808.
19. 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, Enhanced rare-earth separation with a metal-sensitive lanmodulin dimer, *Nature*, 2023, **618**, 87-93.
20. A. M. a. J. A. S. a. J. A. C. a. D. M. P. Ziye Dong and Joseph, Protein-based approach for high-purity Sc, Y, and grouped lanthanide separation, *Separation and Purification Technology*, 2024, **333**, 125919.
21. H. Brittain, F. Richardson and R. Martin, - Terbium(III) emission as a probe of calcium(II) binding sites in proteins, - *J. Am. Chem. Soc.*, 1976, DOI: - 10.1021/ja00441a060, - 8255.
22. M. Gross, G. Nelsestuen and R. Kumar, Observations on the binding of lanthanides and calcium to vitamin-d-dependent chick intestinal calcium-binding protein - implications regarding calcium-binding protein function, *Journal of Biological Chemistry*, 1987, **262**, 6539-6545.
23. Y. C. Sekharudu and M. Sundaralingam, A structure-function relationship for the calcium affinities of regulatory proteins containing 'EF-hand' pairs, *Protein Eng*, 1988, **2**, 139-146.
24. S. Edington, A. Gonzalez, T. Middendorf, D. Halling, R. Aldrich and C. Baiz, Coordination to lanthanide ions distorts binding site conformation in calmodulin, *Proceedings of the National Academy of Sciences of the United States of America*, 2018, **115**, E3126-E3134.
25. I. Bertini, I. Gelis, N. Katsaros, C. Luchinat and A. Provenzani, Tuning the affinity for lanthanides of calcium binding proteins, *Biochemistry*, 2003, **42**, 8011-8021.
26. M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen and B. Imperiali, Structural origin of the high affinity of a chemically evolved lanthanide-binding peptide, *Angew Chem Int Ed Engl*, 2004, **43**, 3682-3685.
27. Y. Ye, H. W. Lee, W. Yang, S. Shealy and J. J. Yang, Probing site-specific calmodulin calcium and lanthanide affinity by grafting, *J Am Chem Soc*, 2005, **127**, 3743-3750.
28. T. Nakatsukasa, Y. Shiraishi, S. Negi, M. Imanishi, S. Futaki and Y. Sugiura, Site-specific DNA cleavage by artificial zinc finger-type nuclease with cerium-binding peptide, *Biochem Biophys Res Commun*, 2005, **330**, 247-252.
29. M. Nitz, K. J. Franz, R. L. Maglathlin and B. Imperiali, A powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides, *Chembiochem*, 2003, **4**, 272-276.
30. Q. Ye, D. Wang and N. Wei, Engineering biomaterials for the recovery of rare earth elements, *Trends Biotechnol*, 2024, **42**, 575-590.
31. T. Hatanaka, N. Kikkawa, A. Matsugami, Y. Hosokawa, F. Hayashi and N. Ishida, The origins of binding specificity of a lanthanide ion binding peptide, *Sci Rep*, 2020, **10**, 19468.

View Article Online
DOI: 10.1039/D5SC02315G

32. D. Park, D. Reed, M. Yung, A. Eslamimanesh, M. Lencka, A. Anderko, Y. Fujita, R. Riman, A. Navrotsky and Y. Jiao, Bioadsorption of Rare Earth Elements through Cell Surface Display of Lanthanide Binding Tags, *Environmental Science & Technology*, 2016, **50**, 2735-2742.
33. K. Dooley, Y. Kim, H. Lu, R. Tu and S. Banta, Engineering of an Environmentally Responsive Beta Roll Peptide for Use As a Calcium-Dependent Cross-Linking Domain for Peptide Hydrogel Formation, *Biomacromolecules*, 2012, **13**, 1758-1764.
34. K. Dooley, B. Bulutoglu and S. Banta, Doubling the Cross-Linking Interface of a Rationally Designed Beta Roll Peptide for Calcium-Dependent Proteinaceous Hydrogel Formation, *Biomacromolecules*, 2014, **15**, 3617-3624.
35. B. Bulutoglu, S. J. Yang and S. Banta, Conditional Network Assembly and Targeted Protein Retention via Environmentally Responsive, Engineered β-Roll Peptides, *Biomacromolecules*, 2017, **18**, 2139-2145.
36. O. Shur, K. Dooley, M. Blenner, M. Baltimore and S. Banta, A designed, phase changing RTX-based peptide for efficient bioseparations, *Biotechniques*, 2013, **54**, 197-+.
37. W. Abdallah, K. Solanki and S. Banta, Insertion of a Calcium-Responsive β-Roll Domain into a Thermostable Alcohol Dehydrogenase Enables Tunable Control over Cofactor Selectivity, *Acs Catalysis*, 2018, **8**, 1602-1613.
38. G. R. Szilvay, M. A. Blenner, O. Shur, D. M. Cropek and S. Banta, A FRET-based method for probing the conformational behavior of an intrinsically disordered repeat domain from Bordetella pertussis adenylate cyclase, *Biochemistry*, 2009, **48**, 11273-11282.
39. B. Bulutoglu, K. Dooley, G. Szilvay, M. Blenner and S. Banta, Catch and Release: Engineered Allosterically Regulated β-Roll Peptides Enable On/Off Biomolecular Recognition, *ACS Synth Biol*, 2017, **6**, 1732-1741.
40. F. Khoury, Z. Su and S. Banta, Rare Earth Element Binding and Recovery by a Beta Roll-Forming RTX Domain, *Inorg Chem*, 2024, **63**, 13223-13230.
41. Z. Yu and L. Chistoserdova, Communal metabolism of methane and the rare Earth element switch, *J Bacteriol*, 2017, **199**, e00328-00317.
42. P. A. Fields, Review: Protein function at thermal extremes: balancing stability and flexibility, *Comp Biochem Physiol A Mol Integr Physiol*, 2001, **129**, 417-431.
43. M. de Champdoré, M. Staiano, M. Rossi and S. D'Auria, Proteins from extremophiles as stable tools for advanced biotechnological applications of high social interest, *J R Soc Interface*, 2007, **4**, 183-191.
44. T. J. Ashaolu, T. Malik, R. Soni, M. A. Prieto and S. M. Jafari, Extremophilic Microorganisms as a Source of Emerging Enzymes for the Food Industry: A Review, *Food Sci Nutr*, 2025, **13**, e4540.
45. D. R. Mende, I. Letunic, J. Huerta-Cepas, S. S. Li, K. Forslund, S. Sunagawa and P. Bork, proGenomes: a resource for consistent functional and taxonomic annotations of prokaryotic genomes, *Nucleic Acids Res*, 2017, **45**, D529-D534.
46. S. R. Eddy, Accelerated Profile HMM Searches, *PLoS Comput Biol*, 2011, **7**, e1002195.
47. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P.

View Article Online
DOI: 10.1039/D5SC02315G

Kohli and D. Hassabis, Highly accurate protein structure prediction with AlphaFold, *Nature*, 2021, **596**, 583-589.

48. V. Thumuluri, J. J. Almagro Armenteros, A. R. Johansen, H. Nielsen and O. Winther, DeepLoc 2.0: multi-label subcellular localization prediction using protein language models, *Nucleic Acids Res*, 2022, **50**, W228-W234.
49. D. G. Lee, M. E. Trujillo, S. Kang, J. J. Nam and Y. J. Kim, Epidermidibacterium keratini gen. nov., sp. nov., a member of the family Sporichthyaceae, isolated from keratin epidermis, *Int J Syst Evol Microbiol*, 2018, **68**, 745-750.
50. S. P. Glaeser, J. A. McInroy, H. J. Busse and P. Kämpfer, Nocardioides zeae sp. nov., isolated from the stem of Zea mays, *Int J Syst Evol Microbiol*, 2014, **64**, 2491-2496.
51. O. I. Nedashkovskaya, S. B. Kim, S. K. Han, M. S. Rhee, A. M. Lysenko, E. Falsen, G. M. Frolova, V. V. Mikhailov and K. S. Bae, Ulvibacter litoralis gen. nov., sp. nov., a novel member of the family Flavobacteriaceae isolated from the green alga Ulva fenestrata, *Int J Syst Evol Microbiol*, 2004, **54**, 119-123.
52. D. J. Rigden, M. J. Jedrzejas and M. Y. Galperin, An extracellular calcium-binding domain in bacteria with a distant relationship to EF-hands, *FEMS Microbiol Lett*, 2003, **221**, 103-110.
53. B. Elshorst, M. Hennig, H. Försterling, A. Diener, M. Maurer, P. Schulte, H. Schwalbe, C. Griesinger, J. Krebs, H. Schmid, T. Vorherr and E. Carafoli, NMR solution structure of a complex of calmodulin with a binding peptide of the Ca2+ pump, *Biochemistry*, 1999, **38**, 12320-12332.
54. L. Bumba, J. Masin, P. Macek, T. Wald, L. Motlova, I. Bibova, N. Klimova, L. Bednarova, V. Veverka, M. Kachala, D. I. Svergun, C. Barinka and P. Sebo, Calcium-Driven Folding of RTX Domain β-Rolls Ratchets Translocation of RTX Proteins through Type I Secretion Ducts, *Mol Cell*, 2016, **62**, 47-62.
55. M. Blenner, O. Shur, G. Szilvay, D. Cropek and S. Banta, Calcium-Induced Folding of a Beta Roll Motif Requires C-Terminal Entropic Stabilization, *Journal of Molecular Biology*, 2010, **400**, 244-256.
56. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C. C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, 2024, **630**, 493-500.
57. E. E. Snyder, B. W. Buoscio and J. J. Falke, Calcium(II) site specificity: effect of size and charge on metal ion binding to an EF-hand-like site, *Biochemistry*, 1990, **29**, 3937-3943.
58. T. Dudev and C. Lim, Metal binding affinity and selectivity in metalloproteins: insights from computational studies, *Annu Rev Biophys*, 2008, **37**, 97-116.
59. Q. Tan, Y. Ding, Z. Qiu and J. Huang, Binding Energy and Free Energy of Calcium Ion to Calmodulin EF-Hands with the Drude Polarizable Force Field, *ACS Phys Chem Au*, 2022, **2**, 143-155.
60. E. A. Grzybowska, Calcium-Binding Proteins with Disordered Structure and Their Role in Secretion, Storage, and Cellular Signaling, *Biomolecules*, 2018, **8**.

View Article Online
DOI: 10.1039/D5SC02315G

61. L. Kalmar, D. Homola, G. Varga and P. Tompa, Structural disorder in proteins brings order to crystal growth in biomineralization, *Bone*, 2012, **51**, 528-534.
62. V. N. Uversky, J. R. Gillespie and A. L. Fink, Why are "natively unfolded" proteins unstructured under physiologic conditions?, *Proteins*, 2000, **41**, 415-427.
63. A. W. Maniccia, W. Yang, J. A. Johnson, S. Li, H. Tjong, H. X. Zhou, L. A. Shaket and J. J. Yang, Inverse tuning of metal binding affinity and protein stability by altering charged coordination residues in designed calcium binding proteins, *PMC Biophys*, 2009, **2**, 11.
64. J. A. Mattocks, J. V. Ho and J. A. Cotruvo, A Selective, Protein-Based Fluorescent Sensor with Picomolar Affinity for Rare Earth Elements, *J Am Chem Soc*, 2019, **141**, 2857-2861.
65. A. A. Homaei, R. Sariri, F. Vianello and R. Stevanato, Enzyme immobilization: an update, *Journal of Chemical Biology*, 2013, **6**, 185-205.
66. O. Shur and S. Banta, Rearranging and concatenating a native RTX domain to understand sequence modularity, *Protein Engineering Design & Selection*, 2013, **26**, 171-180.
67. S. M. Gutenthaler, S. Tsushima, R. Steudtner, M. Gailer, A. Hoffmann-Röder, B. Drobot and L. J. Daumann, Lanmodulin peptides - unravelling the binding of the EF-Hand loop sequences stripped from the structural corset, *Inorg Chem Front*, 2022, **9**, 4009-4021.
68. F. Bou-Abdallah and T. R. Giffune, The thermodynamics of protein interactions with essential first row transition metals, *Biochim Biophys Acta*, 2016, **1860**, 879-891.
69. R. D. Shannon, Revised effective ionic-radii and systematic studies of interatomic distances in halides and chalcogenides, *Acta Crystallographica Section a*, 1976, **32**, 751-767.
70. T. A. Knotts, N. Rathore and J. J. de Pablo, An entropic perspective of protein stability on surfaces, *Biophys J*, 2008, **94**, 4473-4483.
71. S. a. D.-T. M.-D. Ohlson, *Weak Affinity Chromatography (WAC)*, John Wiley & Sons Ltd, 2017.
72. R. Friedman, Preferential Binding of Lanthanides to Methanol Dehydrogenase Evaluated with Density Functional Theory, *J Phys Chem B*, 2021, **125**, 2251-2257.
73. B. S. Van Gosen, P. L. Verplanck, R. R. Seal Ii, K. R. Long and J. Gambogi, *Rare-earth elements*, Report 1802O, Reston, VA, 2017.
74. J. He, Y. Li, X. Xue, H. Ru, X. Huang and H. Yang, Leaching of fluorine and rare earths from bastnaesite calcined with aluminum hydroxide and the recovery of fluorine as cryolite, *RSC Advances*, 2017, **7**, 14053-14059.
75. X. Sun, Z. Li, D. Meng, X. Huang, Z. Feng, M. Wang and Z. Yanyan, Thermal decomposition behavior of Mountain Pass rare earth concentrate in air/CO2 atmosphere, *Journal of Rare Earths*, 2024, DOI: https://doi.org/10.1016/j.jre.2024.04.002.

View Article Online
DOI: 10.1039/D5SC02315G

The authors confirm that the data supporting the findings of this study are available within the article and its ESI.