HAL
open science

Capturing The Dynamic Association Between A Tris-Dipicolinate Lanthanide Complex And A Decapeptide: A Combined Paramagnetic NMR And Molecular Dynamics Exploration.

Sandrine Denis-Quanquin, Alessio Bartocci, Florence Szczepaniak, Francois Riobé, Olivier Maury, Elise Dumont, Nicolas Giraud

To cite this version:

Sandrine Denis-Quanquin, Alessio Bartocci, Florence Szczepaniak, Francois Riobé, Olivier Maury, et al.. Capturing The Dynamic Association Between A Tris-Dipicolinate Lanthanide Complex And A Decapeptide: A Combined Paramagnetic NMR And Molecular Dynamics Exploration.. Physical Chemistry Chemical Physics, 2021, 23 (19), pp.11224-11232. 10.1039/D0CP06570F . hal-03372877

HAL Id: hal-03372877

https://hal.science/hal-03372877v1

Submitted on 11 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Please do not adjust margins

ARTICLE

Capturing The Dynamic Association Between A Tris-Dipicolinate Lanthanide Complex And A Decapeptide: A Combined Paramagnetic NMR And Molecular Dynamics Exploration.†

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Sandrine Denis-Quanquin,[a] Alessio Bartocci,[a] Florence Szczepaniak,[a] Francois Riobé,[a] Olivier Maury,[a] Elise Dumont[a,c*] and Nicolas Giraud[b*]

In the realm of biomolecules, peptides can present a manifold of structures. Our study sheds new light on the highly dynamic structural interplay between a tris-dipicolinate lanthanide probe and a test decapeptide SASYKTLPRG. Whereas a rather monotous, electrostatically-driven association may have been expected, the combination of paramagnetic NMR and molecular dynamics simulations extensively captures interaction sites and their occupancy. This study reveals the importance of a large conformational sampling to reconcilate characteristic time in NMR with molecular dynamics simulations, where sampling at the microsecond range is needed. This study opens the door for a detailed mechanistic elucidation of the early steps of lanthanide complexe-peptide or lanthanide complexe-protein interaction or self-assembly processes.

Introduction

Interaction of lanthanide ions or complexes with biomolecules (proteins, peptides, but also lipid and DNA) is versatile involving binding tag,[1] supramolecular recognition,[2,3] and spans a large range of timely applications such as (chiral-) sensing[4] and even immuno assays.[5,6] Such applications often take advantage of the luminescence optical and chiroptical properties of lanthanide complexes.[7,8] Most probant results of lanthanide-biomolecules interactions have been obtained, yet structural insights are scarce. Indeed, only 72 structures of proteins in presence of lanthanide ions or complexes are currently reported in the Protein Data Bank (PDB). The paradigm may have changed recently as lanthanide ions have been proposed as agents for protein assisted co-crystallization, owing to their phasing[9,10,11] or even nucleating properties.[12] This calls even more for a rational view of interaction of lanthanide complexes with biomolecules. The recent concept of "molecular glues"[13,14] as chemical auxiliaries used to consolidate protein-protein interfaces, and notably may give rise to an increasingly large number of PDB structures co-crystallized lanthanide complexes.[15,16,17] Yet the crystallographic information will not suffice on its own and calls for an investigation the interaction of lanthanide complexes with biomolecules. A challenge is to restore a dynamic view of the non-covalent binding process. One of the best experimental method, relying on the most celebrated electronic properties of lanthanides, is paramagnetic nuclear magnetic resonance (NMR) which gave rise to the structural elucidation of several proteins based on the pseudo contacts shifts (PCS).[18,19,20] Whereas the complete, sole PCS-based structural elucidation is possible but remains challenging, such studies are usually coupled to docking[21] or molecular dynamics.[22,23] The structural elucidation of peptides has most often dealt with lanthanide ions which directly interact with side chains.

In this paper, we report an exhaustive structrual exploration of the dynamical association between a lanthanide probe and a disordered peptidic system. We have carried out a combined paramagnetic NMR and all-atom molecular dynamics investigation at the microsecond timerange to probe the interaction pattern. Both partners of this interaction process have been carefully selected on the basis of their interaction properties (Figure 1).

On the one hand, the tris-dipicolinate lanthanide complex $[Ln(DPA)_3]^{3-}$ is one of the first system of its kind whose abilty to develop supramolecular interactions with proteins has been reported.[11,24,25,21,26,27] Its interactions properties have paved the way for the development of versatile non-covalent tags that can be used (i) to induce protein crystallization, and/or (ii) to exploit the paramagnetic or anomalous scattering properties of the inserted lanthanide centre, for the purpose of probing protein structure. Furthermore, the structure of $[Ln(DPA)_3]^{3-}$ is simple, and exists as a racemic mixture of two enantiomers Λ and Δ that undergo a rapid interconversion in water at room temperature.[28]

[a.] Univ. Lyon, ENS de Lyon, CNRS UMR 5182, Université Claude Bernard Lyon 1, Laboratoire de Chimie, F69342, Lyon, France.
[b.] Université de Paris, Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, UMR CNRS 8601, Paris, France
[c.] Institut Universitaire de France, 5 rue Descartes, 75005 Paris
*Correspondence to elise.dumont@ens-lyon.fr; nicolas.giraud@u-paris.fr
†Electronic Supplementary Information (ESI) available: See DOI: 10.1039/x0xx00000x

Please do not adjust margins

Please do not adjust margins

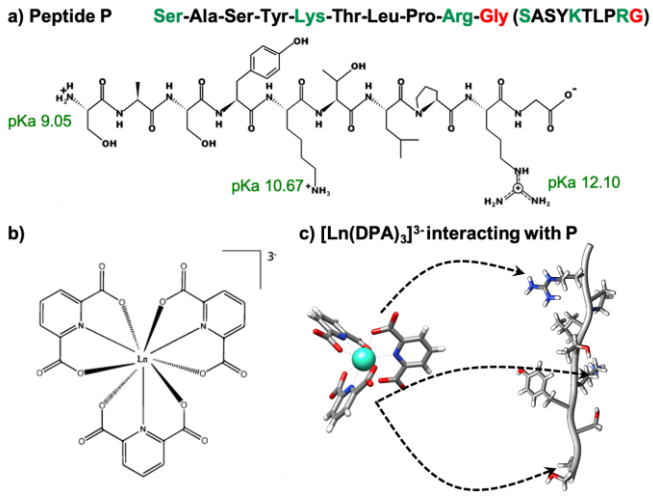

Figure 1. Structures of a) a decapeptide SASYKTLPRG denoted **P** hereafter, b) the tris-dipicolinate lanthanide complex and c) the complex **P**:$[Ln(DPA)_3]^{3-}$ (the Δ enantiomer is shown). The three positively-charged residues are denoted in green (serine S1 is uncapped) and correspond to the possible anchoring points with the lanthanide complex. The last residue is a C-terminal, uncapped, hence negatively-charged glycine (in red).

On the other hand, we have chosen a decapeptide **P** whose sequence SASYKTLPRG makes it a candidate to bind $[Ln(DPA)_3]^{3-}$ without presenting a unique, somewhat trivial, association mode. Indeed, this peptide offers several key characteristics: (i) its sequence features two positively-charged side chain and an $–NH_3^+$ termini, which should promote competitive electrostatic interactions as **P** presents two extremal and one central anchoring points, (ii) a postulated structural flexibility, (iii) negatively-charged groups that can also act on the binding mode to minimize steric repulsion. A tyrosine lies as the fourth residue (Y4) since inspection proteins co-crystallized with lanthanide complexes has revealed possibility of π-stacking between dipicolinate moieties and aromatic residues.[16,29] Furthermore, **P** can be considered as an intrinsically disordered peptide (IDP). We note that its sequence offers three expected anchorage points upon interaction with $[Ln(DPA)_3]^{3-}$, whereas the C-terminal uncapped glycine G10 should give rise to a repulsive contribution.

Methodology

NMR Spectroscopy

The peptide **P** was purchased from the Protein Science Facility in Lyon (UMS3444). General procedures for the preparation of *tris*-dipicolinate lanthanide complexes ($[Na]_3[Ln(DPA)_3].xH_2O$ (Ln = Tb, Gd, Pr or Y)[28] are detailed in Supporting Information. The peptide was dissolved in a sodium phosphate buffer at a pH ranging from 6.5 to 7. D_2O was added to the sample and 1,4-dioxane was used as an internal reference. Samples were prepared in 5 mm tubes with a peptide concentration of 2 mM. Solutions of $[Na]_3[Ln(DPA)_3]$ in D_2O (Ln= Tb, Gd, Pr or Y) were prepared at a concentration of 100 mM for the titration experiments.

NMR experiments were carried out on a Bruker Avance III 400 MHz spectrometer equipped with a Prodigy cryoprobe with a z-axis gradient coil. The temperature was regulated at 298K. 1D ^1H, and 2D ^1H-^1H TOCSY spectra were recorded to fully characterize the peptide at every step of the titration. 2D ROESY spectra were acquired to assess any kind of secondary structure in the peptide. 2D ^1H-^{13}C HSQC spectra were acquired to characterize the peptide and for the observation of paramagnetic relaxation enhancements (PRE). All spectra were processed using Topspin 3.6 software.

2D ^1H-^1H TOCSY spectra were recorded using a MLEV-17 spin-lock sequence with a mixing time of 70 ms and a wet solvent suppression scheme. Typically, spectra were acquired with 256 t_1 increments, 1024 data points, a relaxation delay of 0.8 s and 8 to 32 scans, giving an overall experimental time between 40 min and 160 min. All spectra were zero-filled in the F1 spectral dimension to 1024 data points followed by forward linear prediction of 400 points. The baseline of the FID was corrected using a gaussian function to further suppress the residual water signal. Finally, a square sine bell window function (SSB = 3) was applied in both dimensions prior to Fourier transformation.

2D ^1H–^1H ROESY spectra were recorded using a spin lock of 200 ms and excitation sculpting for solvent suppression. Typically, spectra were acquired with 256 t_1 increments, 2048 data points, a relaxation delay of 0.8 s and 32 scans, giving an overall experimental time of 3h. All spectra were zero-filled in the F1 spectral dimension to 1024 data points followed by forward linear prediction of 512 points. Finally, a square cosine bell window function was applied in both dimensions prior to Fourier transformation.

2D ^1H-^{13}C HSQC spectra were recorded with a standard phase sensitive edited HSQC sequence using gradients and adiabatic pulses. Typically, spectra were acquired with 256 t_1 increments, 1024 data points, a relaxation delay of 1 s and 64 scans, giving an overall experimental time of 5h. All spectra were zero-filled in the F1 spectral dimension to 1024 data points followed by forward linear prediction of 400 points. The baseline of the FID was corrected using a gaussian function to further suppress the residual water signal. Finally, a square cosine bell window function was applied in both dimensions prior to Fourier transformation.

Molecular dynamics simulations

Explicit solvent, all-atom molecular dynamics simulations were performed by using the Amber18 software package[31], after parametrization of the *tris*-dipicolinate lanthanide complex enantiomers based on the metal centre parameters builder (MCPB) approach proposed for organometallic compounds by Li and Merz. The decapeptide **P** was built using the *tleap* available within *Ambertools*. Standard Amber force fields were applied: ff14SB[33] for the decapeptide and GAFF parameters for the ligand $[Y(DPA)_3]^{3-}$. Within the MCPB[32] method, the structure of the complex was optimized at the DFT-B3LYP/6-31G(d) level of theory by using the Gaussian[34] program, and the RESP charge assignment leads to a charge of 0.43e on the Y atom. The van der Waals parameters was set up to 1.602 Å, in agreement with previous values in the literature. This choice would be the same for an yttrium atom, and it is important to stress out here that

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Please do not adjust margins

we discard the role of the lanthanide centre as the coordination takes place in the second sphere.

The decapeptide **P,** isolated or interacting with the Δ or Λ enantiomers of the $[Y(DPA)_3]^{3-}$ complex, was simulated using the same computational protocol. The systems were explicitly solvated using TIP3P water in truncated octahedron boxes, with sodium (Na^+) and chloride (Cl^-) ions corresponding to a salt concentration of 0.2 M. For all the simulation boxes, minimization was performed including 5000 steps of steepest descent and 5000 steps of conjugate gradient. Then, the temperature was increased from 0 to 300 K in a 30 ps thermalization run (NVT), followed by a 1 ns equilibration performed in NPT conditions. During the rest of the simulation, the temperature was kept constant at 300 K using the Langevin thermostat with a collision frequency γ_{ln} equal to $1ps^{-1}$. Particle Mesh Ewald (PME) method was used to treat long-range interaction. During the equilibration and production run phases, a cut off of 10 Å was used. The bonds involving hydrogen were treated with the SHAKE constraints algorithm. For both enantiomers, four independent trajectories of 1 µs with different initial velocities were run (thus 4 µs in total), while for **P** alone, only one of 500 ns. *Cpptraj* module was used to perform cluster analysis, calculations of the distances and evaluation of surface accessible solvent area (SASA). The binding free energy ΔG_{bind} between **P** and both enantiomer molecules of the trisdipicolinate complex has been evaluated through the attach-pulling-release (APR[35]) method (see ESI).

Results and Discussion

NMR Study

P was fully characterized in solution using 2D ^1H-^1H TOCSY, NOESY, and ^{13}C-^1H HSQC experiments (the assignment of proton shifts as well as the NMR spectra recorded on **P** are shown in Figure S1 and Table S1) The ROESY spectra did not show any correlation suggesting a stable secondary structure, as expected for a peptide of this size (Figures S2 and S3). Four lanthanide ions (Terbium, Gadolinium, Yttrium and Praseodymium) were then chosen to probe the interaction process between **P** and $[Ln(DPA)_3]^{3-}$. A series of 2D ^1H-^1H TOCSY or ^{13}C-^1H HSQC spectra was acquired on a peptide sample to which we have added the lanthanide complex, for $[Ln(DPA)_3]^{3-}$:**P** ratios ranging from 0.2 to 5 (Ln = Tb, Gd, Pr or Y).

We have first used a Terbium complex to exploit the strong paramagnetic shifts induced by this metal even at a long distance.[36] Upon addition of $[Tb(DPA)_3]^{3-}$, we indeed observe a significant variation in most of the proton shifts in **P** (Figure 2). We find that the amplitude of this variation increases with the amount of $[Tb(DPA)_3]^{3-}$. Although at the highest ratio some proton signals were significantly broadened due to Paramagnetic Relaxation Enhacement (PRE) induced by the Terbium centre, their assignment was still possible without any ambiguity (see SI Table S2 for chemical shift and chemical shift variations of all proton resonances). Overall, these data show that **P** and $[Tb(DPA)_3]^{3-}$ are interacting.

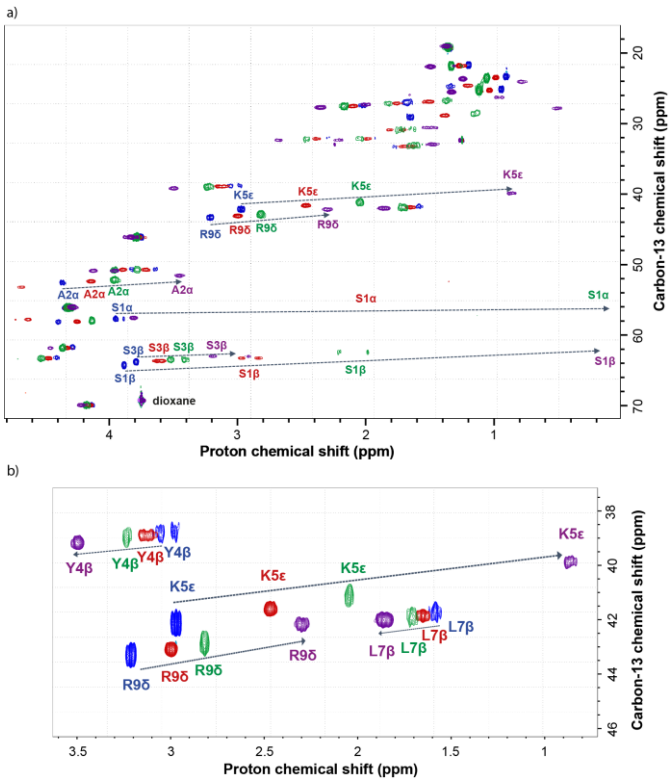

Figure 2. a) ^1H-^{13}C HSQC spectra recorded on the peptide sample (blue), and the peptide interacting with 0.5 (red), 1 (green) and 3 (purple) equivalents of $[Tb(DPA)_3]^{3-}$ b) zoom on the region between 44 and 38 ppm for ^{13}C. Assigned signals show some of the most significant chemical shift variations.

Moreover, the observation of a unique set of ^1H-^{13}C correlations with a characteristic linear shift reflects a fast exchange process between a free state and a bound state resulting from the interaction between **P** and $[Tb(DPA)_3]^{3-}$.

We remind here that the paramagnetic shift δ_{para} induced for any proton in **P** by the interaction with a paramagnetic complex is defined through the equation:

$$\delta_{para} = \delta^{bound}_{para} - \delta^{bound}_{dia} \quad [1]$$

where δ^{bound}_{para} is the chemical shift of the proton in the presence of the paramagnetic complex and δ^{bound}_{dia} is the chemical shift in the presence of the same complex incorporating a diamagnetic lanthanide metal. δ^{bound}_{dia} accounts for the conformational change induced in the molecule that is interacting with the paramagnetic compound. Furthermore, δ_{para} is the sum of two contributions called Fermi contact and pseudo-contact shifts. The Fermi contact comes from a through-bond interaction and is negligible in most situations. [36] Herein, the paramagnetic shift may be interpreted as the pseudo-contact shift, which depends both on the distance of the spin to the paramagnetic centre and on the orientation of the paramagnetic complex with respect to the interacting molecule.[37] For a fast exchange process, δ^{bound}_{para} and δ^{bound}_{dia} can be determined by fitting the model curve corresponding to the relevant interaction mechanism to the shifts measured upon titration.[38]

In the present case, we could model the oberved shifts evolutions by an equilibrium between a free and a "bound"

Please do not adjust margins

Please do not adjust margins

form, which we refer to as 1:1 equilibrium. An average affinity constant K_d= 8 ± 2 mM coud be determined from the analysis of the titration curves recorded for the different protons in **P**. To confirm this result, praseodymium was selected to monitor the evolution of the diffusion coefficient of the interacting species upon addition of $[Ln(DPA)_3]^{3-}$. This method has proven to be a complementary technique to chemical shift titration experiments.[25] The evolution of the observed diffusion coefficients for the peptide is also in good agreement with a 1:1 equilibrium, yielding a similar affinity constant value K_d=7 mM. We note however that it was not possible to exploit the diffusion data recorded for the complex, because the overall high fraction of the complex in the sample throughout the titration experiment leads to a weak variation in the observed diffusion coefficient for this molecule.

The analysis of the data recorded with the paramagnetic Terbium complex allowed for determining the paramagnetic shift in the bound form δ^{bound}_{Tb}. Finally, the same titration experiment was carried out using $[Y(DPA)_3]^{3-}$, which allowed for determining the diamagnetic contribution δ^{bound}_{Y} in order to throw off the contribution to the observed shift variation of the conformational changes induced in **P** upon supramolecular interactions (see SI Table S3 and Figures S4 and S5 for the analysis of titration and diffusion NMR experiments using a 1:1 equilibrium). Figure 3 shows the evolution of the apparent paramagnetic shift $\delta_{para} = \delta^{bound}_{Tb} - \delta^{bound}_{Y}$ that was determined for the different protons in **P** when it is interacting with $[Tb(DPA)_3]^{3-}$.

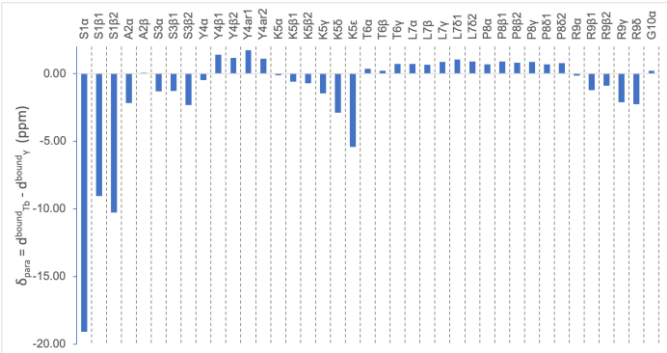

Figure 3. ^{1}H shifts variations $\delta^{bound}_{Tb} - \delta^{bound}_{Y}$ determined for all observed protons in **P** upon interaction with $[Na]_3[Ln(DPA)_3]$ (Ln = Tb, Y).

The largest δ_{para} values are observed for Serine S1 protons, followed by Lysine K5(ε) and K5(δ) protons, at the extremity of the side chain, and Serine S3(HN) amide proton. We note however that for Arginine R9 lateral chain, which is expected to interact closely with the lanthanide complex,[24] we observe a rather weak δ_{para} value. Two interpretations can be considered: (i) Arginine R9 actually does not interact with the lanthanide complex, or (ii) it is interacting, but a dynamic process averages the paramagnetic shift that can be detected for this residue.

Furthermore, the Gadolinium complex was used to probe Paramagnetic Relaxation Enhancements (PREs) throughout the peptide structure. Although they do not induce any chemical shift changes, Gadolinium ions are known to lead to strong PREs that mainly depend on the distance between the nuclear spin and the paramagnetic centre. We have thus acquired ^{1}H-^{13}C HSQC experiments on a 1mM peptide solution with 0, 0.2 and 0.4 equivalents of $[Gd(DPA)_3]^{3-}$ to measure PREs over a broad range of distances.[39] The evolution of the peak ratio I_{para}/I_{dia} calculated for the different C-H correlations belonging to **P** is shown in Figure 4, for two different numbers of equivalents of $[Na]_3[Gd(DPA)_3]$.

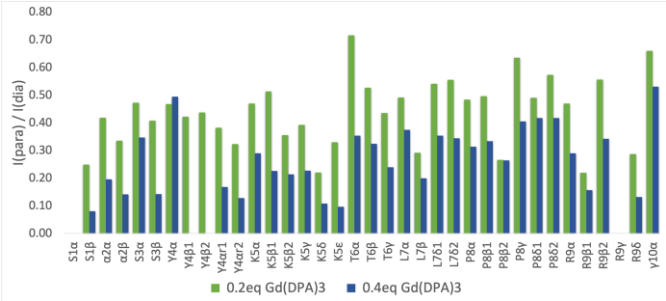

Figure 4. Relative intensity peak ratio (I_{para}/I_{dia}) in the ^{1}H-^{13}C HSQC spectrum of a solution of 1mM of P upon the addition of 0.2 (medium green) or 0.4 (dark blue) equivalents of $[Na]_3[Gd(DPA)_3]$. The intensity of the HSQC in the absence of any lanthanide complex is considered a good approximation of I_{dia}. L7γ is not shown because of a strong overlapping with K5δ.

Overall, we observe that the Guadolinium complex in solution induces a broadening of all the lines as reflected by the value of the I_{para}/I_{dia} ratios that are all below 80%. We also note that this broadening is not uniform. Some residues are indeed more impacted than others as illustrated on the I_{para}/I_{dia} ratios determined for 0.2 equivalent that show the disappearance of the S1(α) and R9(γ) protons signals. This increased PRE shows that some proton sites have been in closer contact with the paramagnetic centre. The analysis of the PRE distributions at 0.2 and 0.4 equivalents thus suggests that residues S1 and R9 were directly interacting with the lanthanide complex, as well as the region Y4-K5.

From the NMR study, we can conclude that we observe and quantify the interaction between **P** and the lanthanide complex $[Ln(DPA)_3]^{3-}$, which gives rise to a fast exchange process on the NMR timescale. The analysis of the PCS and PRE data highlight three regions that are in close contact with the complex: S1, Y4-K5, and R9. We remark however that the PCS values that are obtained for some residues suggest that the flexibility of **P** during the interaction process should be accounted for in order to fully explain them. To gain further insight into this system, in the following section we perform molecular dynamics simulations to unveil the dynamic features of the interaction process.

Molecular dynamics simulations of the decapeptide P and impact of binding with $[Ln(DPA)_3]^{3-}$

We first explore the structure of the explicitly-solvated decapeptide **P** to probe structural changes upon binding with the ligand. This decapeptide features two residues with a positively-charged side chain, Lysine K5 and Arginine R9 (boldfaced in green in Figure 1). The decapeptide **P** that was used in the NMR experiments is not capped at the N- and C-termini, leading to $–NH_3^+$ and $–COO^-$ terminal functional groups of Serine S1 and Glycine G10 that also play a role in defining the structure of **P**. Accordingly, **P** was left uncapped in our

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Please do not adjust margins

simulations. With three positive charged residues, two uncapped and polar ones prone to hydrogen bonding, **P** can be considered as a good testcase to delineate a competition as one, two (eventually the three) residues could interact with the carboxylate arms of the dipicolinate moiety, as suggested by the experimental NMR study.

As expected for an intrinsically-disordered peptide, **P** on its own spans a large numbers of structures. This is reflected in the distribution of the head-to-tail ($D_{h\text{-}t}$) interatomic distance measured between the carbon atoms of the N- and C-termini S1(α) and G10(α): it is centered around 17.2 ± 5.7 Å (Figure 5). The third cluster provides an exemple of a more compact fold of **P**, with a distance of 4.92 Å, owing to the formation of a salt bridge between S1 and G10. This structure is representative of the low-populated first peak at 5.5 Å (green line, Figure 5). Moreover, transient intermolecular hydrogen bonds between O:S3···H:L7 and O:Y4···H:T6 contribute to the folding of the decapeptide **P**.

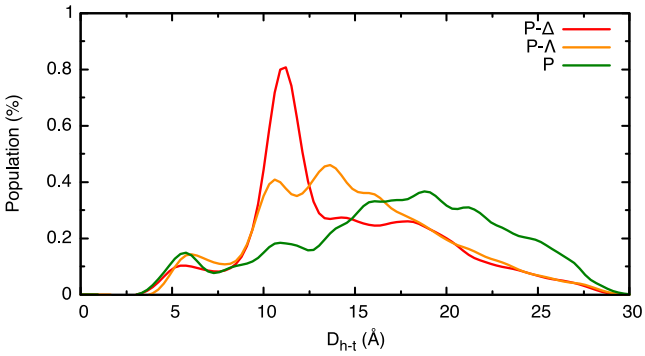

Figure 5. Overall simulation time head-tail distance $D_{h\text{-}t}$ in Å, between the carbon atoms of S1(α) and G10(α), reported as normalized histograms. The systems are represented as follows: red line for **P**-Δ, orange line for **P**-Λ and green line for the peptide **P** isolated.

The association between the tris-dipicolinate complex (charged -3) and the decapeptide **P** was then probed by molecular dynamics simulations, with two series of four trajectories of 1 μs, one series for the Δ enantiomer and one series for the Λ enantiomer, as reported in Table S4. The decapeptide **P** offers a priori at least three anchorage points upon interaction with $[Y(DPA)_3]^{3-}$ to trigger its "wrapping" upon binding with the tris-dipicolinate lanthanide. This fold can be seen by the decrease on the $D_{h\text{-}t}$ value (14.3 and 14.8 Å for the adducts **P**-Δ and **P**-Λ, respectively), as reported and Figure 5 and Table S4.

The non-covalent interactions between **P** and the tris-dipicolinate lanthanide complex can also be quantified as associative by the decreasing values of SASA of **P** with respect to the intrinic value 1196 Å^2 (**P** alone).

The most populated structures iddue from the clusters analysis (46.0% and 37.4% for the Λ and Δ enantiomers respectively), shown leftside in Figure 6 (next page), correspond to an interaction of the trisdipicolinate with S1 and R9, whereas the central lysine K5 is solvent exposed. The other clusters, accounting more 40-50% of the total population correspond to an alternance of S1, K5 and R9 as two of them bind the tris-dipicolinate. Figure 7 provides the time evolution of three interatomic distances between the lanthanide ion and the centres of mass of the three residues S1, K5 and R9. It characterizes an highly dynamic association, as the positively-charged residues regularly swap to anchor the lanthanide.

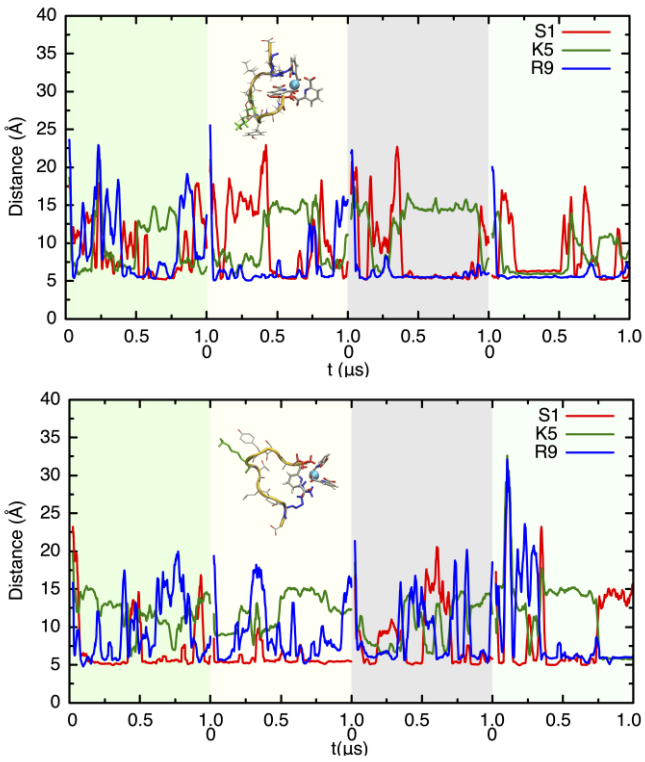

Figure 7. Time evolution of the distances (in Å) between S1, K5 and R9 centres of mass and the central lanthanide ion (green sphere) for P-Δ (upper panel) and P-Λ (lower panel) supramolecular adducts, along the four independent trajectories (collated). The averaged values are reported in Table S4 for all four different simulations.

Interaction with the N-terminal Serine S1 occurs through its uncapped -NH_3^+ group, which is prone to interact with an carboxylate oxygen of the tris-dipicolinate (the one not bounded to the lanthanide ion). The binding of the lanthanide complex with R9 corresponds to an electrostatic interaction, reinforced by a π-stacked mode (see the third clusters in Figures 6-b and 6-c). Interaction with the lysine K5 implies its NH_3^+ end group and the oxygens of the carboxylate arise. It is interesting to note that only one cluster, only for the Δ enantiomer, only weakly populated with an occupation of 13.5%, features the three interactions at the same time. This compact fold is barely observed for the Λ enantiomer. This indicates that the decapeptide **P** is too short and too dynamic to afford a stable association mode where the three-positively charge groups bind the trianionic complex.

Please do not adjust margins

Please do not adjust margins

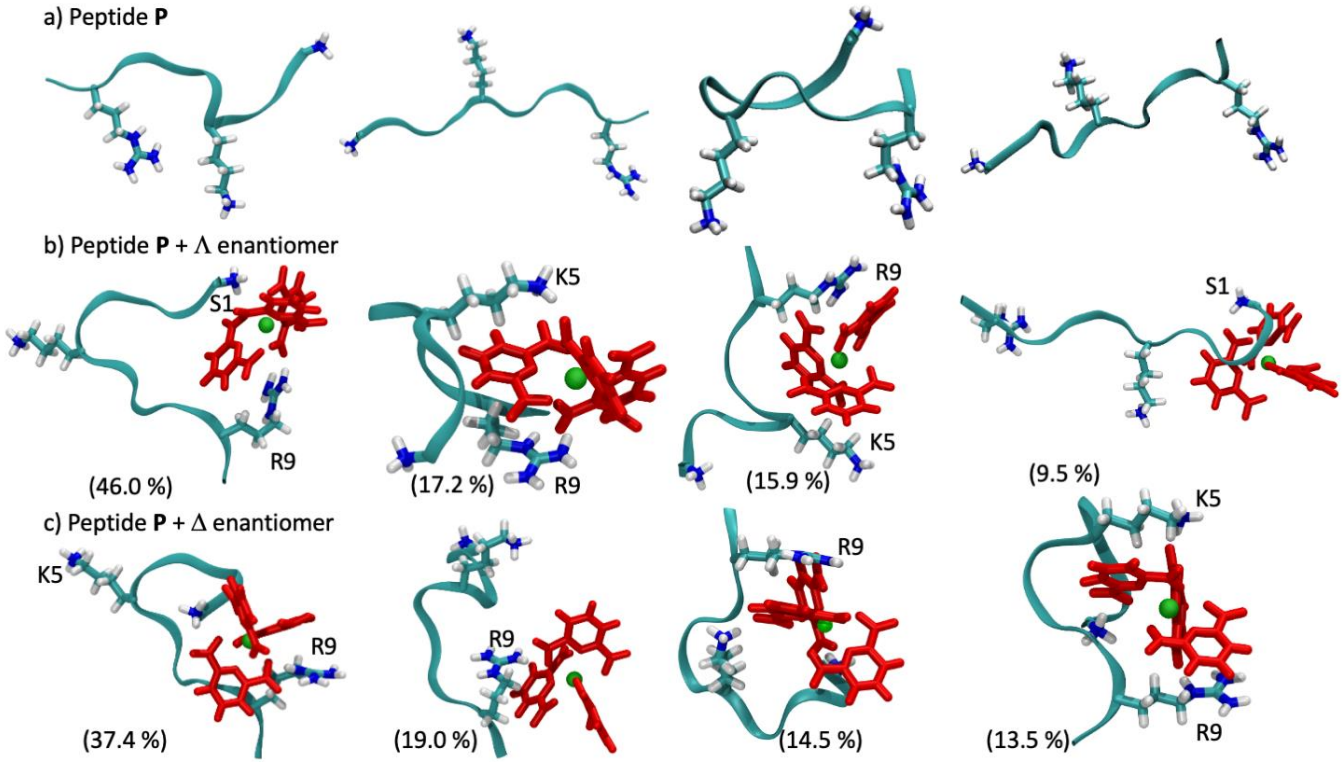

Figure 6. Four representative structures for **P**, isolated, interacting with the Δ or Λ enantiomer of $[Y(DPA)_3]^{3-}$ obtained from cluster analysis of all trajectories (population percentages are given in parenthesis). The water box and part of the hydrogen atoms are omitted for the sake of clarity. The lanthanide central ion is shown in green sphere, the dipicolinate ligands are shown in red.

Other residues transiently do not interact, or interact sparsely, with the tris-dipicolinate: Y4, the only aromatic residue of **P**, opportunistically π-stacks with dipicolinate moities (ca. 5% over time). Several, competitive non-covalent interactions can come into play, also, to trigger a partial folding with more compact conformations of **P**: a π-stacking between Y4 and K5 (ca. 50% of the trajectory, see Figure S8), and an hydrogen bond interaction between O:S1···H:L7 and O:S1···H:T6.

The structural inspection at the microsecond time range proves a highly dynamic and versatile interaction mode as $[Ln(DPA)_3]^{3-}$ interacts with **P**. As can be seen, the two enantiomers of the lanthanide complex behave structurally very similarly, with no chirality-enhanced for the association with the peptide **P**. This is confirmed by absolute free energy calculations of ΔG_{bind}, which is found similar for the two enantiomers, given the error bars, with values of -5.0±0.6 and -6.3±0.7 kcal.mol^{-1} (see Table S4).

Discussion

At this step it is interesting to draw up the overall landscape of the interaction process, as it can be addressed by the combination of NMR spectroscopy and MD simulations. On the one hand, it should be noted that both approaches have identified 3 major interaction sites along **P** that are residues S1, K5 and R9, but also a contribution from the Tyrosine Y4. In particular, there is a good agreement on the fact that S1 is interacting via the NH_3^+ group on its backbone, whereas for K5 and R9 their sidechain is coming into play. This atomic scale analysis is validated by both the distances measurements reported by MD, and the PCS and PRE perturbations probed by NMR. On the other hand, NMR and MS offer a complementary insight into the dynamic features of these supramolecular interactions. From the NMR point of view, the data extracted from the titration experiments are coherent with a 1:1 equilibrium in the fast exchange regime. Moreover, MD simulations suggest that this "bound" state should not be seen as a rigid, unique adduct, but results from a dynamic association between the lanthanide complex and **P**. The computed trajectories describe an interaction scenario with one highly populated structure (for each enantiomer) where $[Ln(DPA)_3]^{3-}$ interact through S1 and R9, and a series of less populated -transient- structures corresponding to weaker interaction modes, where S1, K5 and R9 alternatively interact with **P**. This

Please do not adjust margins

Please do not adjust margins

dynamic picture is useful to qualitatively explain the fluctuations in the PREs that were determined using the Gadolinium complex. In Figure 8, the PRE value is plotted as a function of the distance to the lanthanide centre for the different proton sites in the peptide. It is noticeable that the protons with the strongest attenuation are located on parts of the residues S1 and R9 that are directly interacting with the lanthanide complex in the most populated structure. Conversely, for the residues S3, T6, L7, P8 which are not key actors of the interaction process, we see higher PREs, with the general trend of the longer the distance, the higher the PRE. It is interesting to note that a third group of protons can be highlighted for K5 and Y4, with a longer mean distance and a rather small PRE value. These protons belong to the side chain of these residues that are interacting with the complex.

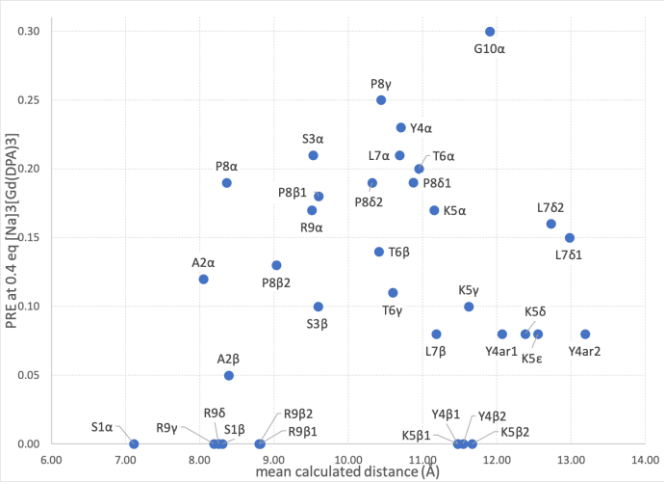

Figure 8. Correlation between the mean calculated distance from the different protons in **P** to the lanthanide complex calculated from the MD simulation, and the experimental PREs measured for the corresponding C-H correlations upon addition of 0.4 eq. of $[Na]_3[Gd(DPA)_3]$.

The rather small value of their PRE can be interpreted as arising from the fact that these protons spend time in close contact with the gadolinium metal, whereas the mean distance value that is rather long for such PRE value, can be explained by the fact that they belong to a highly flexible region of **P**, which explores a broad conformational space.

From these observations, it can be assumed that if on the time scale of the NMR experiment this system can be seen as 1:1 equilibrium undergoing a fast exchange process, MD simulations allow for addressing the fate of the "bound" state on a shorter time scale. The positioning of Tyrosine Y4 is interesting in that respect: our MD simulations characterize a p-cation interaction with the Lysine K5 side chain, with a low proximity to the lanthanide center (conformations for which the interatomic distance towards the center of mass of the Tyrosine Y4 is lower than 8 Å are populated to 10.0%). Even with a low population, such structures (shown in Figure S9) could explain the signals observed on the NMR experiments. The dynamic association between the lanthanide complex and the peptide **P** leads to an averaging of the NMR measurements that can be performed on this "bound" state, as it is illustrated by the affinity constant that can be extracted from titration experiments. We note that such averaging of PCS and PRE values upon interaction with a flexible system has already been reported[40]. The experimental Gibbs energy that can be evaluated from this affinity constant (ΔG = -2.9 ± 0.2 kcal.mol^{-1}), and the agreement with the computational estimation (-5.0 and -6.3 kcal.mol^{-1}) is encouraging: the overestimation based on all-atom MD simulations is expected here, due to the lack of polarizable force fields or the lack of sampling non-bounded states. A more integrated view on the free energy signature of peptides/lanthanide complexes will require more advanced exploration.

Conclusions

We report a combined NMR-molecular dynamics exploration of the highly dynamical association of a test decapeptide with the $[Ln(DPA)_3]^{3-}$ complex. Unlike what one could have anticipated, even this electrostatically driven system presents a non-trivial association mode, highly dynamic, which can be captured by the crosstalk between molecular dynamics and paramagnetic NMR. On the ns to µs timescale, MD simulations reveal that 3 residues in the peptide interact non simultaneously with the lanthanide complex, leading to one main structure alternating with less populated "bound" states with higher flexibility. These transient states are interconverting rapidly on the timescale that is relevant to NMR, hereby leading to the observation of an average "bound" state that is itself in fast exchange with a free state. Such complex picture of a supramolecular interaction process can only be addressed by the combination of NMR and MD, which both offer an atomic resolution insight of the interacting species, on complementary timescales.

We believe that this approach will be useful to address more complex supramolecular systems such as the nucleation steps that are preceding several aggregation or crystallization processes, because they involve transient species that can neither be captured in solution before the nucleation starts, These results open the way to the study of a broad range of molecular scaffolds, from proteins to amyloid peptides, where MD simulations could become an interesting way of pre-screening new systems before carrying out an NMR exploration. The development of methods to increase the analytical potential of this approach is in progress in our groups.

Conflicts of interest

"There are no conflicts to declare".

Acknowledgements

Classical MD simulations were performed using the resources of the Pole Scientifique de Modélisation Numériques (PSMN). A.B. is grateful for a post-doctoral grant awarded by the Fondation Maison de la Chimie (FMC). The authors thank the SYSPROD project and AXELERA Pôle de Compétitivité for financial support (PSMN Data Center).

Please do not adjust margins

Please do not adjust margins

References

(1) Sculimbrene, B. R.; Imperiali, B. Lanthanide-Binding Tags as Luminescent Probes for Studying Protein Interactions. *J. Am. Chem. Soc.* **2006**, *128* (22), 7346–7352. https://doi.org/10.1021/ja061188a.

(2) Meskers, S. C. J.; Dekkers, H. P. J. M. Enantioselective Quenching of Luminescence: Molecular Recognition of Chiral Lanthanide Complexes by Biomolecules in Solution. *J. Phys. Chem. A* **2001**, *105* (19), 4589–4599. https://doi.org/10.1021/jp004428w.

(3) Tsukube, H.; Shinoda, S. Lanthanide Complexes in Molecular Recognition and Chirality Sensing of Biological Substrates. *Chem. Rev.* **2002**, *102* (6), 2389–2404. https://doi.org/10.1021/cr010450p.

(4) Jastrząb, R.; Nowak, M.; Skrobańska, M.; Tolińska, A.; Zabiszak, M.; Gabryel, M.; Marciniak, Ł.; Kaczmarek, M. T. DNA as a Target for Lanthanide(III) Complexes Influence. *Coord. Chem. Rev.* **2019**, *382*, 145–159. https://doi.org/10.1016/j.ccr.2018.12.018.

(5) Rajendran, M.; Yapici, E.; Miller, L. W. Lanthanide-Based Imaging of Protein–Protein Interactions in Live Cells. *Inorg. Chem.* **2014**, *53* (4), 1839–1853. https://doi.org/10.1021/ic4018739.

(6) Zhang, K. Y.; Yu, Q.; Wei, H.; Liu, S.; Zhao, Q.; Huang, W. Long-Lived Emissive Probes for Time-Resolved Photoluminescence Bioimaging and Biosensing. *Chem. Rev.* **2018**, *118* (4), 1770–1839. https://doi.org/10.1021/acs.chemrev.7b00425.

(7) Bünzli, J.-C. G. Lanthanide Luminescence for Biomedical Analyses and Imaging. *Chem. Rev.* **2010**, *110* (5), 2729–2755. https://doi.org/10.1021/cr900362e.

(8) Carr, R.; Evans, N. H.; Parker, D. Lanthanide Complexes as Chiral Probes Exploiting Circularly Polarized Luminescence. *Chem Soc Rev* **2012**, *41* (23), 7673–7686. https://doi.org/10.1039/C2CS35242G.

(9) Girard, É.; Stelter, M.; Vicat, J.; Kahn, R. A New Class of Lanthanide Complexes to Obtain High-Phasing-Power Heavy-Atom Derivatives for Macromolecular Crystallography. *Acta Crystallogr. Sect. D* **2003**, *59* (11), 1914–1922. https://doi.org/10.1107/S0907444903020511.

(10) Silvaggi, N. R.; Martin, L. J.; Schwalbe, H.; Imperiali, B.; Allen, K. N. Double-Lanthanide-Binding Tags for Macromolecular Crystallographic Structure Determination. *J. Am. Chem. Soc.* **2007**, *129* (22), 7114–7120. https://doi.org/10.1021/ja070481n.

(11) Pompidor, G.; D'Aléo, A.; Vicat, J.; Toupet, L.; Giraud, N.; Kahn, R.; Maury, O. Protein Crystallography through Supramolecular Interactions between a Lanthanide Complex and Arginine. *Angew. Chem. Int. Ed.* **2008**, *47* (18), 3388–3391. https://doi.org/10.1002/anie.200704683.

(12) Engilberge, S.; Riobé, F.; Di Pietro, S.; Lassalle, L.; Coquelle, N.; Arnaud, C.-A.; Pitrat, D.; Mulatier, J.-C.; Madern, D.; Breyton, C.; Maury, O.; Girard, E. Crystallophore: A Versatile Lanthanide Complex for Protein Crystallography Combining Nucleating Effects, Phasing Properties, and Luminescence. *Chem Sci* **2017**, *8* (9), 5909–5917. https://doi.org/10.1039/C7SC00758B.

(13) McGovern, R. E.; Feifel, S. C.; Lisdat, F.; Crowley, P. B. Microscale Crystals of Cytochrome c and Calixarene on Electrodes: Interprotein Electron Transfer between Defined Sites. *Angew. Chem. Int. Ed.* **2015**, *54* (21), 6356–6359. https://doi.org/10.1002/anie.201500191.

(14) Rennie, M. L.; Fox, G. C.; Pérez, J.; Crowley, P. B. Auto-Regulated Protein Assembly on a Supramolecular Scaffold. *Angew. Chem. Int. Ed.* **2018**, *57* (42), 13764–13769. https://doi.org/10.1002/anie.201807490.

(15) Gruszczyk, J.; Fleurie, A.; Olivares-Illana, V.; Béchet, E.; Zanella-Cleon, I.; Moréra, S.; Meyer, P.; Pompidor, G.; Kahn, R.; Grangeasse, C.; Nessler, S. Structure Analysis of the Staphylococcus Aureus UDP-N-Acetyl-Mannosamine Dehydrogenase Cap5O Involved in Capsular Polysaccharide Biosynthesis. *J. Biol. Chem.* **2011**, *286* (19), 17112–17121. https://doi.org/10.1074/jbc.M110.216002.

(16) Engilberge, S.; Riobé, F.; Wagner, T.; Di Pietro, S.; Breyton, C.; Franzetti, B.; Shima, S.; Girard, E.; Dumont, E.; Maury, O. Unveiling the Binding Modes of the Crystallophore, a Terbium-Based Nucleating and Phasing Molecular Agent for Protein Crystallography. *Chem. – Eur. J.* **2018**, *24* (39), 9739–9746. https://doi.org/10.1002/chem.201802172.

(17) Engilberge, S.; Wagner, T.; Santoni, G.; Breyton, C.; Shima, S.; Franzetti, B.; Riobé, F.; Maury, O.; Girard, E. Protein Crystal Structure Determination with the Crystallophore, a Nucleating and Phasing Agent. *J. Appl. Crystallogr.* **2019**, *52* (4), 722–731. https://doi.org/10.1107/S1600576719006381.

(18) Allegrozzi, M.; Bertini, I.; Janik, M. B. L.; Lee, Y.-M.; Liu, G.; Luchinat, C. Lanthanide-Induced Pseudocontact Shifts for Solution Structure Refinements of Macromolecules in Shells up to 40 Å from the Metal Ion. *J. Am. Chem. Soc.* **2000**, *122* (17), 4154–4161. https://doi.org/10.1021/ja993691b.

(19) Jia, X.; Yagi, H.; Su, X.-C.; Stanton-Cook, M.; Huber, T.; Otting, G. Engineering [Ln(DPA)3]3− Binding Sites in Proteins: A Widely Applicable Method for Tagging Proteins with Lanthanide Ions. *J. Biomol. NMR* **2011**, *50* (4), 411. https://doi.org/10.1007/s10858-011-9529-x.

(20) Bertini, I.; Luchinat, C.; Parigi, G.; Pierattelli, R. NMR Spectroscopy of Paramagnetic Metalloproteins. *ChemBioChem* **2005**, *6* (9), 1536–1549. https://doi.org/10.1002/cbic.200500124.

(21) Saio, T.; Yokochi, M.; Kumeta, H.; Inagaki, F. PCS-Based Structure Determination of Protein–Protein Complexes. *J. Biomol. NMR* **2010**, *46* (4), 271–280. https://doi.org/10.1007/s10858-010-9401-4.

(22) Schmitz, C.; Vernon, R.; Otting, G.; Baker, D.; Huber, T. Protein Structure Determination from Pseudocontact Shifts Using ROSETTA. *J. Mol. Biol.* **2012**, *416* (5), 668–677. https://doi.org/10.1016/j.jmb.2011.12.056.

(23) Tu, K.; Gochin, M. Structure Determination by Restrained Molecular Dynamics Using NMR Pseudocontact Shifts as Experimentally Determined Constraints. *J. Am. Chem. Soc.* **1999**, *121* (40), 9276–9285. https://doi.org/10.1021/ja9904540.

(24) Dumont, E.; Pompidor, G.; D'Aléo, A.; Vicat, J.; Toupet, L.; Kahn, R.; Girard, E.; Maury, O.; Giraud, N. Exploration of the Supramolecular Interactions Involving Tris-Dipicolinate Lanthanide Complexes in Protein Crystals by a Combined Biostructural, Computational and NMR Study. *Phys Chem Chem Phys* **2013**, *15* (41), 18235–18242. https://doi.org/10.1039/C3CP53671H.

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Please do not adjust margins

(25) Denis-Quanquin, S.; Riobé, F.; Delsuc, M.-A.; Maury, O.; Giraud, N. Paramagnetic DOSY: An Accurate Tool for the Analysis of the Supramolecular Interactions between Lanthanide Complexes and Proteins. *Chem. – Eur. J.* **2016**, *22* (50), 18123–18131. https://doi.org/10.1002/chem.201603460.

(26) Wei, Z.; Yang, Y.; Li, Q.-F.; Huang, F.; Zuo, H.-H.; Su, X.-C. Noncovalent Tagging Proteins with Paramagnetic Lanthanide Complexes for Protein Study. *Chem. – Eur. J.* **2013**, *19* (18), 5758–5764. https://doi.org/10.1002/chem.201204152.

(27) Su, X.-C.; Liang, H.; Loscha, K. V.; Otting, G. [Ln(DPA)3]3− Is a Convenient Paramagnetic Shift Reagent for Protein NMR Studies. *J. Am. Chem. Soc.* **2009**, *131* (30), 10352–10353. https://doi.org/10.1021/ja9034957.

(28) Ouali, N.; Bocquet, B.; Rigault, S.; Morgantini, P.-Y.; Weber, J.; Piguet, C. Analysis of Paramagnetic NMR Spectra of Triple-Helical Lanthanide Complexes with 2,6-Dipicolinic Acid Revisited: A New Assignment of Structural Changes and Crystal-Field Effects 25 Years Later. *Inorg. Chem.* **2002**, *41* (6), 1436–1445. https://doi.org/10.1021/ic010801i.

(29) Talon, R.; Nauton, L.; Canet, J.-L.; Kahn, R.; Girard, E.; Gautier, A. Clicked Europium Dipicolinate Complexes for Protein X-Ray Structure Determination. *Chem Commun* **2012**, *48* (97), 11886–11888. https://doi.org/10.1039/C2CC36982F.

(30) Tancrez, N.; Feuvrie, C.; Ledoux, I.; Zyss, J.; Toupet, L.; Le Bozec, H.; Maury, O. Lanthanide Complexes for Second Order Nonlinear Optics: Evidence for the Direct Contribution of f Electrons to the Quadratic Hyperpolarizability. *J. Am. Chem. Soc.* **2005**, *127* (39), 13474–13475. https://doi.org/10.1021/ja054065j.

(31) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688. https://doi.org/10.1002/jcc.20290.

(32) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599–604. https://doi.org/10.1021/acs.jcim.5b00674.

(33) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian˜16 Revision B.01*; 2016.

(35) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Gilson, M. K. The SAMPL5 Host–Guest Challenge: Computing Binding Free Energies and Enthalpies from Explicit Solvent Simulations by the Attach-Pull-Release (APR) Method. *J. Comput. Aided Mol. Des.* **2017**, *31* (1), 133–145. https://doi.org/10.1007/s10822-016-9970-8.

(36) Pintacuda, G.; John, M.; Su, X.-C.; Otting, G. NMR Structure Determination of Protein−Ligand Complexes by Lanthanide Labeling. *Acc. Chem. Res.* **2007**, *40* (3), 206–212. https://doi.org/10.1021/ar050087z.

(37) Nitsche, C.; Otting, G. Pseudocontact Shifts in Biomolecular NMR Using Paramagnetic Metal Tags. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *98–99*, 20–49. https://doi.org/10.1016/j.pnmrs.2016.11.001.

(38) Fielding, L. NMR Methods for the Determination of Protein–Ligand Dissociation Constants. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 219–242. https://doi.org/10.1016/j.pnmrs.2007.04.001.

(39) Yagi, H.; Loscha, K. V.; Su, X.-C.; Stanton-Cook, M.; Huber, T.; Otting, G. Tunable Paramagnetic Relaxation Enhancements by [Gd(DPA)3]3− for Protein Structure Analysis. *J. Biomol. NMR* **2010**, *47* (2), 143–153. https://doi.org/10.1007/s10858-010-9416-x.

(40) Pell, A. J.; Pintacuda, G.; Grey, C. P. Paramagnetic NMR in Solution and the Solid State. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *111*, 1–271. https://doi.org/10.1016/j.pnmrs.2018.05.001.

Please do not adjust margins

Please do not adjust margins

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins