8-2023

Development of Modeling Approaches for A Molecular Level Understanding of Ligand-Lanthanide-Water Systems

Sayani Biswas
sbiswas@clemson.edu

DEVELOPMENT OF MODELING APPROACHES FOR A MOLECULAR LEVEL UNDERSTANDING OF LIGAND-LANTHANIDE-WATER SYSTEMS

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Chemical Engineering

---

by
Sayani Biswas
August 2023

Accepted by:
Dr. Rachel B. Getman, Committee Chair
Dr. Eric Davis
Dr. Mark Roberts

ABSTRACT

Lanthanides (Ln) are a subset of rare earth elements (REEs) that are essential components in electric vehicles, smart phones, and wind turbines. Current REE recovery processes are time intensive and produce hazardous wastes. Developing sustainable recovery processes require new techniques that are economical and less wasteful, such as affinity-based extraction and separations. Enabling affinity-based recovery requires use of ligands designed to bind strongly and selectively to lanthanides. Developing design rules to make such ligands requires an understanding of the coordination environment of lanthanide-ligand complexes. In this work, we begin to develop a molecular level understanding of aqueous lanthanide-ligand systems with the ligand ethylenediaminetetraacetic acid (EDTA). We develop computational models of aqueous EDTA complexes of lanthanides La, Ce, Pr and Nd, informed by molecular dynamics (MD) and density functional theory (DFT). These models are developed so they can be used as the required reference structures for experimental techniques such as extended x-ray absorption fine structure (EXAFS) that provide structural information about metallic complexes. Our results suggest shortcomings in MD structures and indicate that DFT optimized structures prove to be reasonable models for aqueous Ln-EDTA complexes. These DFT structures are used to generate theoretical spectra that are fit against experimental EXAFS spectra to uncover information about bond distances and organization in aqueous Ln-EDTA complexes. Overall, this work contributes to developing a method to combine theoretical data with experimental data to understand the coordination

environment around aqueous phase ligand-lanthanide systems, which is a key step towards designing ligands for sustainable selective REE separations.

ACKNOWLEDGMENTS

I would first like to thank my advisor and committee chair Dr. Rachel Getman for being incredibly kind, supportive and encouraging throughout the period of this research and for providing guidance every step of the way. I would also like to thank members of my committee, Dr. Eric Davis and Dr. Mark Roberts, for their time, useful inputs and advice.

I would like to acknowledge the National Support Foundation (NSF) for their support (award 2133512) and our NSF ECO-CBET team collaborators at Case Western Reserve University and The Pennsylvania State University. I would also like to acknowledge the Research Computing and Data (RCD) group at Clemson Computing and Information Technology for maintaining resources on the Palmetto Supercomputing Cluster.

I would like to thank my past and current research group mates Hafeera, Sanchari, Rabbani, Stephen, Ricardo, Jiexin, Xiuting and Rohit, for everything I get to learn from them and the time and knowledge we get to share on the regular. Also a big thank you to Dr. Bruce, Casey, Joy and Alexandra for their help with administrative tasks. I am also grateful for friends in the Clemson ChBE department for their friendship and good times.

HUGE thank you to my mum, dad, sister and grandparents for their unconditional love and support and for always taking an interest in everything I do.

TABLE OF CONTENTS

Table of Contents (Continued) Page

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Lanthanides are the series of elements in the periodic table ranging from Lanthanum to Lutetium. The recovery of lanthanides is of interest because lanthanides are rare earth elements (REEs), which have unique magnetic and heat-resistant properties which make them essential components in electric vehicles, wind turbines and smart phones [1]. However, the recovery of REEs is challenging. This is because extraction and separation of REEs is difficult. Extraction of REEs is challenging because REEs occur in low concentrations in scattered deposits. Separation among REEs is challenging because they have identical ionic charges (+3) and similar sizes making it difficult to use traditional separation techniques that rely on charge and size. Currently, REEs are mined and purified via multiple stages of solvent extraction, which is time intensive, high in carbon dioxide emissions and produces hazardous wastes [2, 3].

Alternatively, phosphogypsum (PG) has been identified as a potential source of REEs, particularly Lanthanum (La), Cerium (Ce), Praseodymium (Pr) and Neodymium (Nd). Phosphogypsum is waste derived from phosphoric acid that is produced in large quantities at fertilizer manufacturing units. Currently, this waste is piped to nearby water bodies or stored in stacks in the soil, thereby contaminating these resources. In Florida alone, there is more than 1 billion tons of phosphogypsum stored in stacks. These storage stacks are estimated to contain more than 200 million tons of REEs [4] which is sufficient to supply the current annual consumption of REEs in the US at approximately 9000 tons. Treatment of PG for extraction of REEs using materials designed for REE separations could thus prove to be a sustainable source of REE supply. Moreover, the development of this

separation process could have the dual benefit of extracting valuable resources while treating hazardous wastes.

Currently, REEs are separated via multiple stages of solvent extraction where general metal coordinating ligands are used. This method is expensive and time intensive. Thus, to enable sustainable REE extraction and separation, new techniques that are economical and produce less waste, such as affinity based separations, must be developed. Enabling affinity-based separation techniques require ligands that bind strongly to lanthanides and exhibit selectivity towards individual lanthanides in aqueous environments. Developing design rules to make such ligands requires an understanding of the structure of aqueous ligand-lanthanide complexes and the influence of the coordination environment on the binding strengths and selectivity in these ligand-lanthanide complexes.

Ethylenediaminetetraacetic acid (EDTA) is a well known metal chelating ligand used in transition metal and REE separation processes [5, 6]. Therefore, there are works in literature that have studied the structure of EDTA complexes of lanthanides and other metals [7-12]. Since EDTA ligand systems are well studied, we choose EDTA as the ligand to benchmark our study of lanthanide-ligand-water systems. EDTA forms hexadentate complexes with metals in their +3 oxidation states, like Fe^{3+}, such that the metal is chelated by the six potential bonding sites of EDTA, four carboxyl and two amino groups [13]. A representation of this is shown in figure 1.

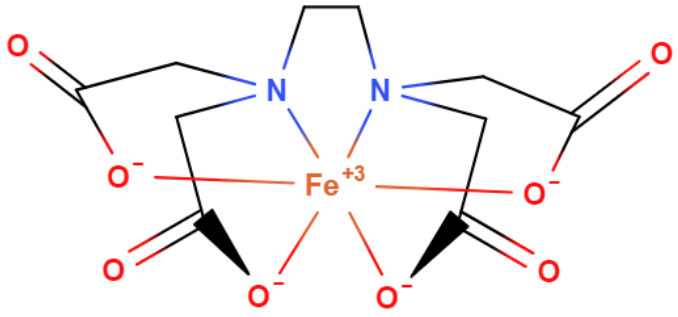

Figure 1: A schematic of an Fe^{3+} ion chelated by the six binding sites of EDTA where EDTA=black lines

The aqueous Ln-EDTA studies in literature reveal that the early lanthanides (including La, Ce, Pr and Nd) are all 9 coordinate in the first coordination shell, with 6 atoms coordinating to EDTA atoms and 3 to water atoms. However, the organization of these coordinating atoms and the interatomic bond distances in aqueous solution are unknown. Uncovering this information requires techniques that can study the local environment of these complexes in aqueous solution. Extended X-Ray Absorption Fine Structure (EXAFS) is an experimental technique that can characterize local surroundings of metallic complexes. However, determining the local structure using EXAFS requires data from a reference structure. Typically, reference structures for EXAFS are determined using X-ray Diffraction (XRD). However, XRD is performed on solid phase samples and does not represent structures in aqueous environments, making it an unsuitable technique to provide reference structures for aqueous phase Ln-EDTA complexes [24]. In this work, we address this challenge by using theoretically informed structures as references for fitting against experimental EXAFS data. Particularly, molecular dynamics (MD) and density functional theory (DFT) are used to develop tractable and accurate models of aqueous phase Ln-

EDTA structures. Theoretical spectra from these models are then generated and fit against experimental EXAFS spectra by our collaborators. The fits are then used to uncover information about the arrangement and bond distances in aqueous Ln-EDTA complexes. This uncovering is essential to understand the influence of structure of the ligand on the binding strengths and selectivity towards lanthanides. Thus, developing computational models that can be used as reference structures for EXAFS fitting is an important step towards designing selective ligands for sustainable REE separations.

2. RESEARCH DESIGN AND METHODS

This research is designed to learn about the complexation of lanthanide (Ln) cations La^{3+}, Ce^{3+}, $Pr^{3+}$and Nd^{3+} with the well-known metal chelator ethylenediaminetetraacetic acid (EDTA). The aim of this work is to develop models that are compatible with experimental EXAFS data to determine the aqueous phase structure of EDTA complexes of the lanthanides. Particularly, this research focuses on determining the coordination environment of EDTA complexed lanthanides. This chapter outlines the computational methods used to determine the structure of aqueous phase Ln-EDTA complexes. The computational methods presented here are combined with experimental work performed by our collaborators to determine the structures.

Initially, classical Molecular Dynamics (MD) simulations are undertaken to learn about the structural properties of Ln-EDTA complexes in aqueous solutions. Classical MD simulations are governed by equations of classical mechanics (i.e., not quantum) while Density Functional Theory (DFT) simulations are governed by quantum mechanics. MD simulations are undertaken initially because they have certain advantages over DFT simulations in that MD simulations are less expensive (require less computational resources), can assess multiple configurations and are less sensitive to initial structural models. However, given that the focus of this work is on the determination of local geometry around lanthanides, the use of quantum informed DFT simulations proves to be essential.

2.1 Structural Determination using a Combined Computational and Experimental Approach

Our experimental collaborators performed Extended X-ray Absorption Fine Structure (EXAFS) experiments to determine the first coordination shell of the lanthanides in aqueous Ln-EDTA complexes. The EXAFS experiments provide information about neighboring atoms of up to 5 Å from the Ln in the Ln-EDTA complexes. X-ray Absorption Spectroscopy (XAS) experiments are performed by our collaborators and used to generate experimental EXAFS spectra. To determine the aqueous Ln-EDTA structures, these EXAFS spectra need to be fitted against reference spectra. Conventionally, EXAFS spectra are fit using structures derived from X-ray Diffraction (XRD). However, XRD derived structures do not directly represent structures in solution since samples are crystallized prior to analysis [24]. To address this challenge, in this work, computationally derived structures are used to fit the EXAFS spectra. Both the MD and DFT derived structures are fit against the experimental EXAFS data individually. The flowcharts in figures 2 and 3 outline the process used to determine the aqueous Ln-EDTA complex structures using a combined MD/EXAFS and DFT/EXAFS approach, respectively.

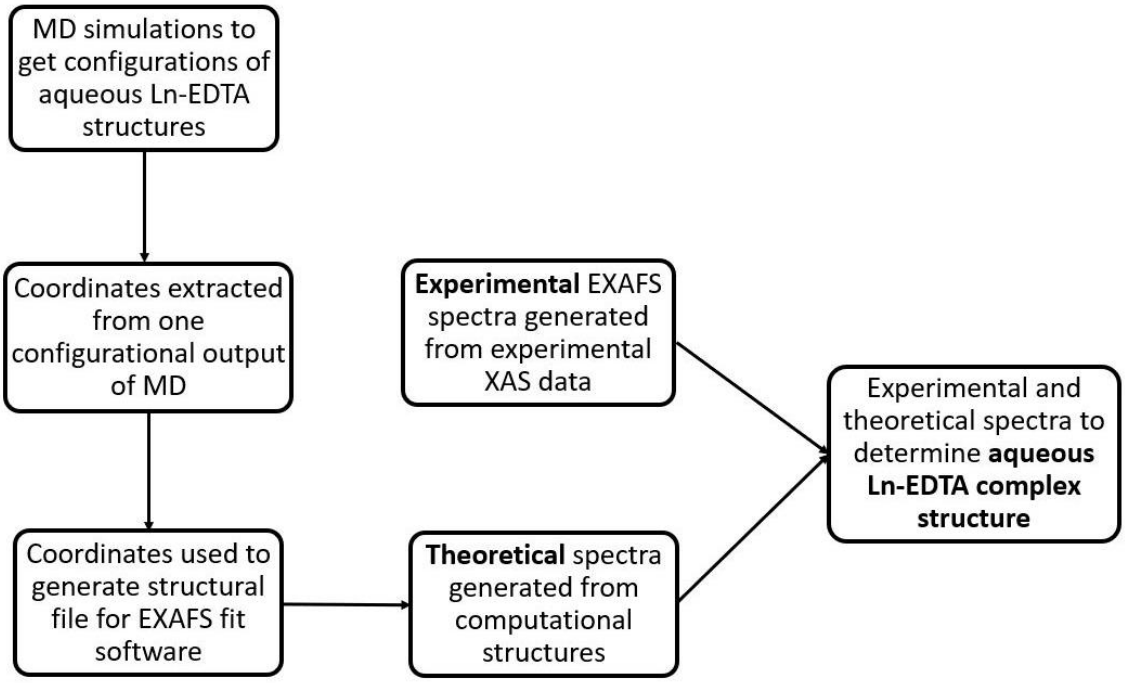

Figure 2: Flowchart describing process used to combine theoretical MD and experimental EXAFS to determine aqueous Ln-EDTA complex structures.

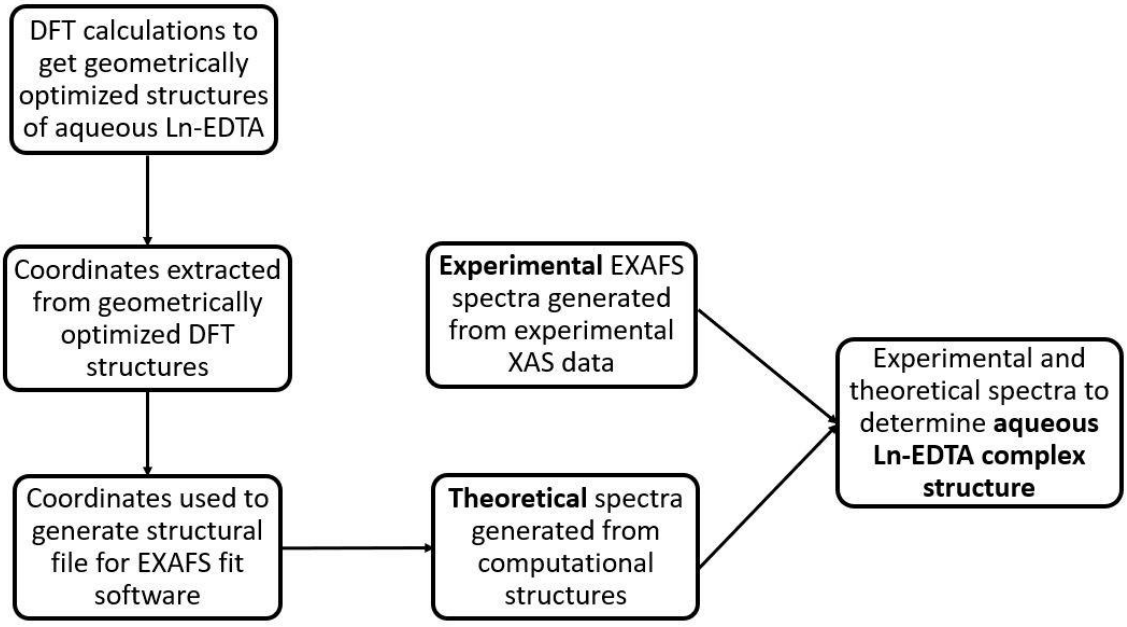

Figure 3: Flowchart describing process used to combine theoretical DFT and experimental EXAFS to determine aqueous Ln-EDTA complex structures.

2.2 Molecular Dynamics Simulations

Classical MD simulations are performed for an aqueous Ce-EDTA complex model with a 1:1 Ce:EDTA concentration, consistent with the ratio used in the experimental work. MD simulations are performed using the GROMACS package. EDTA is modelled in its completely deprotonated form (-4 charge) since it is known to form stable metal complexes in this form [14]. The Ce ion is modelled with a +3 charge. This results in an overall charge of -1 on the Ce-EDTA complex. Since MD uses periodic boundary conditions (PBC), the system used in simulations must be neutral. Thus, a Na ion with +1 charge is added to neutralize the system for simulations.

2.2.1 MD Parameters and Simulations

The structural file for EDTA is sourced from the chemical molecule database, PubChem [15]. Information about the topology of the EDTA structure such as partial atomic charges, atom types, dihedral and improper angles are sourced from a paper by Durand et al. [16]. Lennard-Jones parameters from CHARMM force field are employed for EDTA as well as the Ce^{3+} ion and the neutralizing Na^{+} ion [17, 18]. The simulation box is set to be 30.6 Å on each edge. Particle-mesh Ewald summation is used to model long range electrostatic interactions. The cutoff distance for Lennard-Jones and Coulombic interactions is set to 12 Å. The simulation box with the EDTA, lanthanide and neutralizing ion is solvated using the in built "solvate" function in GROMACS. This results in 950 water molecules and a water density of ~1028 kg/m^3 in the simulation box. The MD simulations are performed at 300 K. The initial model of the aqueous Ce-EDTA complex for MD simulations is shown in figure 4.

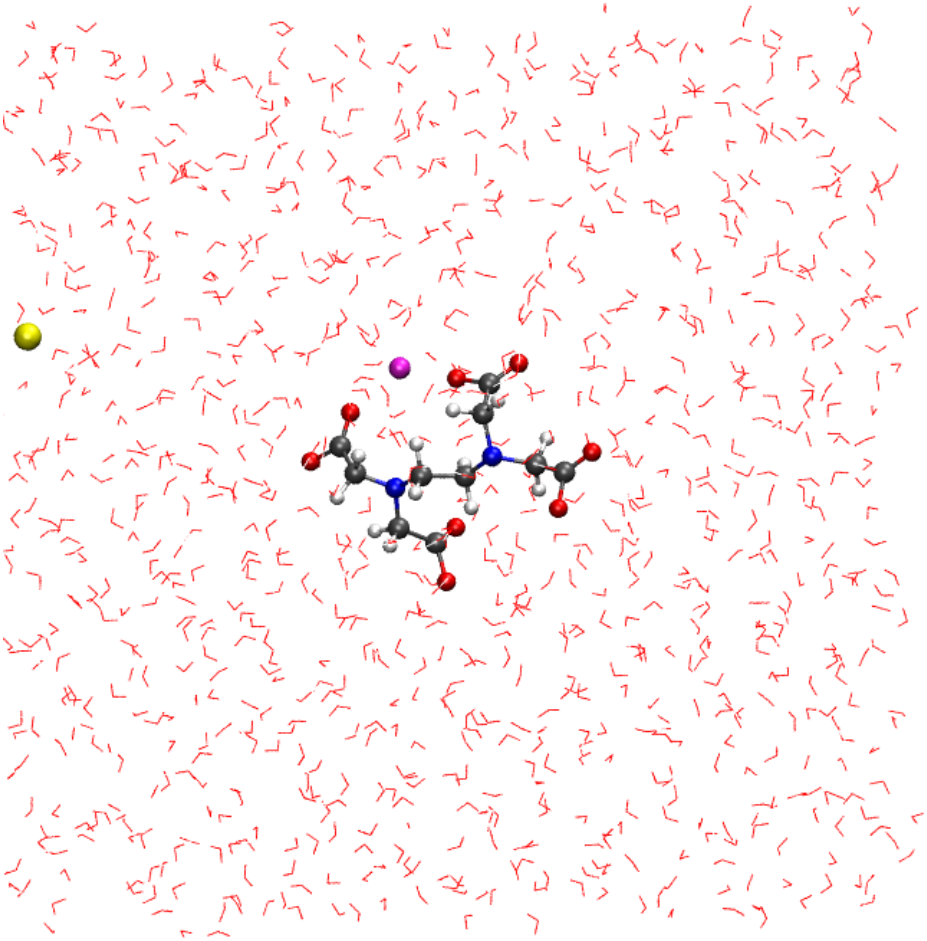

Figure 4: Initial input structure of aqueous Ce-EDTA complex for MD simulations where H2O molecules=red lines, sodium=yellow, nitrogen=blue, carbon=gray, hydrogen=white, oxygen=red and cerium=pink.

Initially, 50,000 steps of energy minimization are performed on the system. Following energy minimization, a constant volume equilibration is performed for 0.1 ns, followed by a constant pressure equilibration for 0.1 ns. Once the system is equilibrated, a 10 ns, constant pressure production run is carried out. The time step used in all the simulation runs is 2 fs. The model of the aqueous Ce-EDTA complex after MD simulations is shown in figure 5.

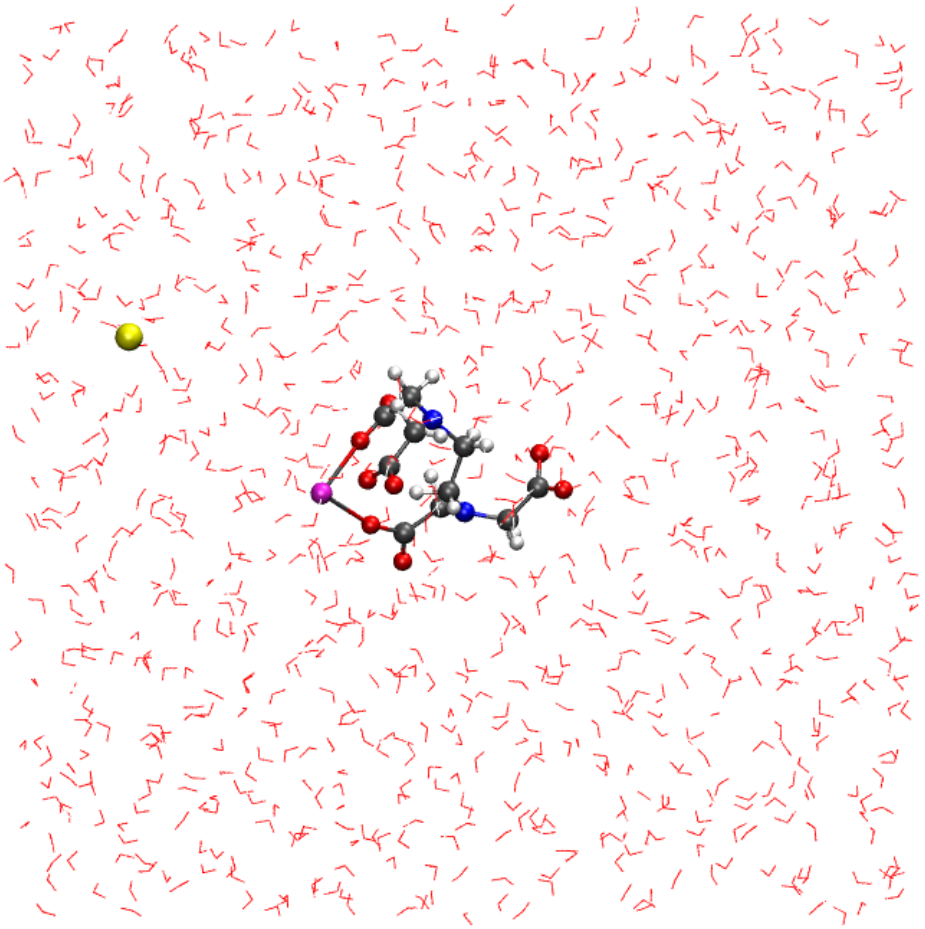

Figure 5: Structure of aqueous Ce-EDTA complex after MD simulations where H2O molecules=red lines, sodium=yellow, nitrogen=blue, carbon=gray, hydrogen=white, oxygen=red and cerium=pink.

2.3 Density Functional Theory Calculations

While MD simulations can provide information about the configuration of atoms in aqueous Ln-EDTA complexes, these do not provide information with quantum accuracy. Since the focus of this work is on the determination of the first coordination shell of the lanthanides, accurately capturing chemical bonding interactions is essential. Thus, DFT simulations are performed to accurately model the first coordination shells of the aqueous Ln-EDTA complexes.

EDTA is known to form stable metal complexes in its completely deprotonated form [14]. EDTA is hence modelled to be completely deprotonated with a –4 charge and each lanthanide cation is modelled with a +3 charge. This results in an overall charge of –1 on each Ln-EDTA complex. Since the DFT simulations undertaken in this work are non periodic, the system need not be neutral (i.e., the system can have an overall charge).

The initial structure of the Ln-EDTA complex is modelled such that the lanthanide cation is chelated by the six potential bonding sites of the EDTA (four carboxyl and two amino groups), similar to complexes that EDTA forms with metals in their +3 oxidation states, such as Fe^{3+} (figure 1) [13]. An example initial input structure used for the DFT simulations is pictured in figure 6.

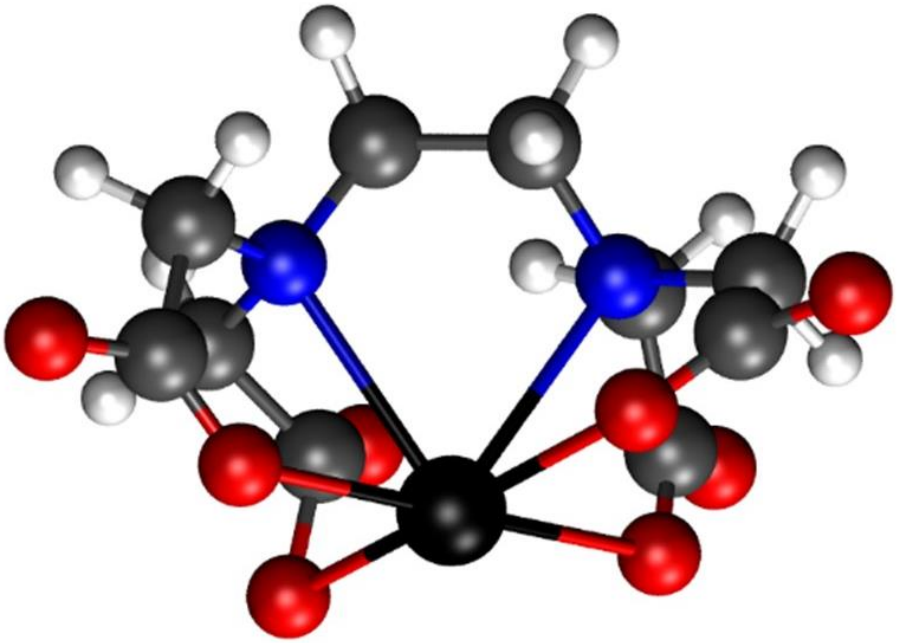

Figure 6: Initial input structure of La-EDTA complex for DFT simulations where nitrogen=blue, carbon=gray, hydrogen=white, oxygen=red, lanthanum=black.

2.3.1 DFT Parameters

DFT calculations are performed using the Gaussian 16 software. The calculations are performed using the M06-L functional. This functional is chosen because it has shown good balance between computational tractability and accuracy for systems containing lanthanides in their +3 oxidation state [19]. The Def2-SVP basis set is used for H, C, O and N atoms of EDTA and water molecules [20] and Def2-TZVPP with associated effective core potentials (ECPs) is used for the lanthanide cations La, Ce, Pr and Nd [21, 22].

A different set of DFT parameters of SARC-DKH2 basis set with B3LYP functional are also known to perform well with lanthanide systems [23]. A major difference in the basis sets Def2-TZVPP and SARC-DKH2 is that Def2-TZVPP uses effective core potentials (ECPs) to model the core electrons while SARC-DKH2 treats all electrons explicitly. Test calculations are carried out using both set of DFT parameters to compare the bond distances and complexation energies of aqueous Ln-EDTA structures. These calculations show that the choice of basis set and functional results in a maximum difference of 0.06 Å in bond distances between the lanthanide and their coordinating atoms (figure 7), which is less than 2% of the total bond distance, and a maximum of ~222 kJ/mol on the calculated ion complexation energies, which is less than 4% of the total energies (table 1). The ion complexation energies are calculated for the reaction Eq. 1 using Eq. 2.

$$\mathrm{Ln}^{3+} + \mathrm{EDTA}^{4-} \rightarrow [\mathrm{Ln} - \mathrm{EDTA}]^{1-}$$

(Eq. 1.)

Ln ion complexation energy = Energy of Ln-EDTA complex – energy of EDTA – energy of Ln

(Eq. 2.)

Table 1. Complexation Energies of Ln-EDTA using different DFT modelling parameters.

Ln-EDTA	Complexation Energies (kJ/mol)			
	Using Def2-TZVPP and M06L	Using SARC-DKH2 and B3LYP	Difference of energies	Difference as % of energies
Ce-EDTA	-6204.52	-6311.82	-107.31	1.73
Nd-EDTA	-6302.23	-6523.90	-221.67	3.52

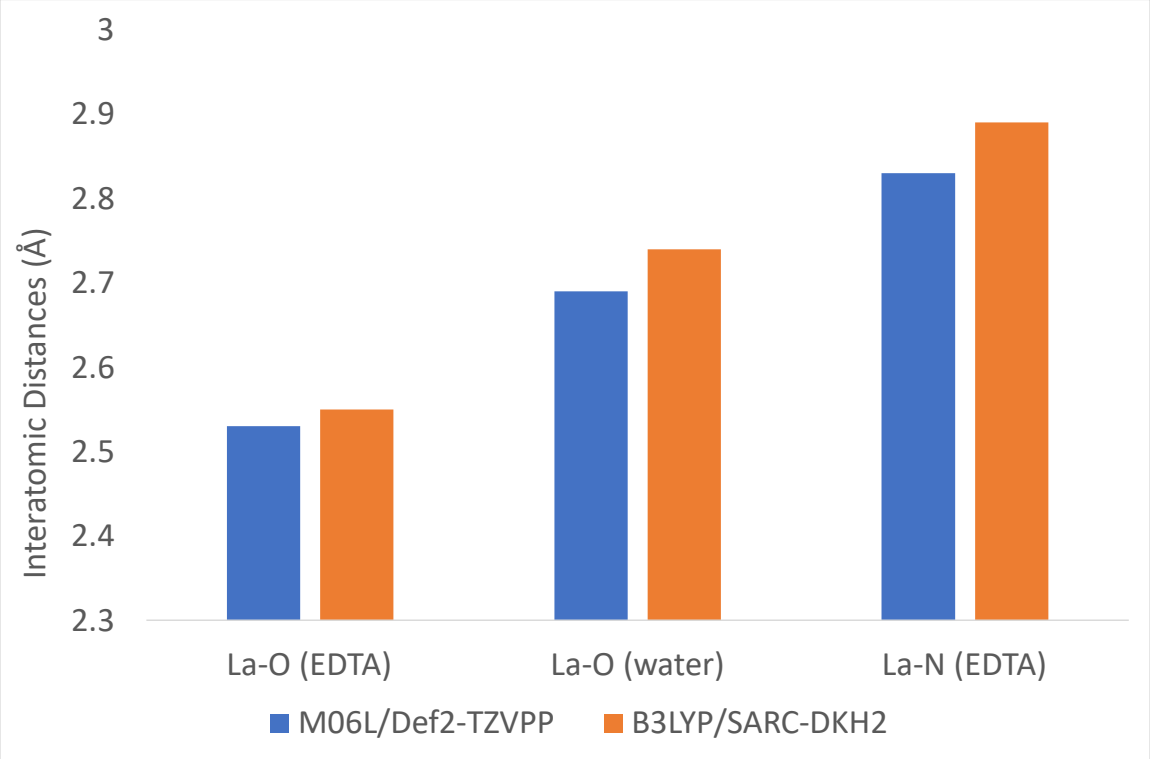

Figure 7: Comparison of bond distances between Lanthanum and coordinating atoms using different DFT parameters.

Several possible multiplicities (a property governed by the ion's electronic configuration resulting in different electronic energies) are considered for the lanthanide ions in the aqueous EDTA complex models. Specifically, multiplicities of 1, 3, 5 and 7 are

considered for La^{3+} and Pr^{3+} while multiplicities 2, 4, 6 and 8 are considered for Ce^{3+} and Nd^{3+}. One way to develop stable structures in DFT methods is to identify, for a given system, the properties which enable the lowest electronic energy in that system. Thus, the electronic energies are calculated at different multiplicities to compare and identify the lowest energy multiplicity for each isolated ion. The calculated energies of the different ions are presented in tables 1 and 2. The structures for each aqueous Ln-EDTA complex are calculated at the multiplicity that yields the lowest electronic energy for each isolated lanthanide cation.

Table 2: Energies of isolated La^{3+} and Pr^{3+} ions at multiplicities 1, 3, 5 and 7. The lowest energy in each case is highlighted in bold text.

Ln^{3+}	Electronic Configuration of Ln^{3+}	Energy (in Hartree) for possible multiplicities			
		1	3	5	7
La^{3+}	$[Xe]5s^2 5p^6$	**-30.18**	-29.41	-28.64	-27.94
Pr^{3+}	$[Xe]4f^2$	-515.57	**-515.62**	-515.04	-514.39

Table 3: Energies of isolated Ce^{3+} and Nd^{3+} ions at multiplicities 2, 4, 6 and 8.

Ln^{3+}	Electronic Configuration of Ln^{3+}	Energy (in Hartree) for possible multiplicities			
		2	4	6	8
Ce^{3+}	$[Xe]4f^1$	**-473.76**	-473.16	-472.50	-471.85
Nd^{3+}	$[Xe]4f^3$	-559.96	**-560.00**	-559.41	-558.78

2.3.2 DFT Calculations

First, geometric optimizations of the Ln-EDTA complex models (not in aqueous phase, i.e., where the influence of water is not accounted for) where all atoms in the model

are allowed to relax. Once an optimized geometry of the Ln-EDTA complex is obtained, the aqueous phase model is developed. To model the aqueous phase structure, explicit water molecules are added to the system one at a time and geometric optimization is repeated where all the atoms are allowed to relax again. These optimized Ln-EDTA complex structures with explicit water molecules are then modelled under implicit water solvent. Details of the explicit and implicit water modelling are outlined in section 2.3.3. Once the models of the aqueous Ln-EDTA complexes are developed, geometric optimizations are undertaken where all the atoms are allowed to relax. This is done for aqueous EDTA complexes of La^{3+}, Ce^{3+}, $Pr^{3+}$and Nd^{3+} to obtain the optimized DFT structures. The DFT optimized structure of aqueous La-EDTA is provided in figure 9. A sample input file for the DFT calculations in Gaussian 16 is included in appendix A.

2.3.3 Modelling Aqueous Ln-EDTA Complex

Since the Ln-EDTA complex is in aqueous phase, the solvent water molecules interact with the Ln-EDTA complex. One way to account for the solvent water molecules is to model each water molecule in solution explicitly. This is, however, computationally intractable. Thus, in this work, explicit water molecules are added to the model one at a time, and simultaneously the theoretical spectra from their optimized structures are fit against the experimental EXAFS spectra (the fitting procedure is undertaken by our collaborators). This process finds that the best fit models for the aqueous Ln-EDTA complexes are La-EDTA, Ce-EDTA and Pr-EDTA with 7 explicit water molecules and Nd-EDTA with 4 explicit water molecules.

These Ln-EDTA complex structures with explicit water molecules are then modelled under implicit water solvent using the Polarizable Continuum Model (PCM) method. Figure 8 shows the influence of including implicit solvation on the quality of the EXAFS fits. The quality is measured using a parameter called R-factor where a good quality fit typically has an R-factor of less than 0.02. Thus, DFT calculated structures having low R-factors are more similar to the experimentally observed structures.

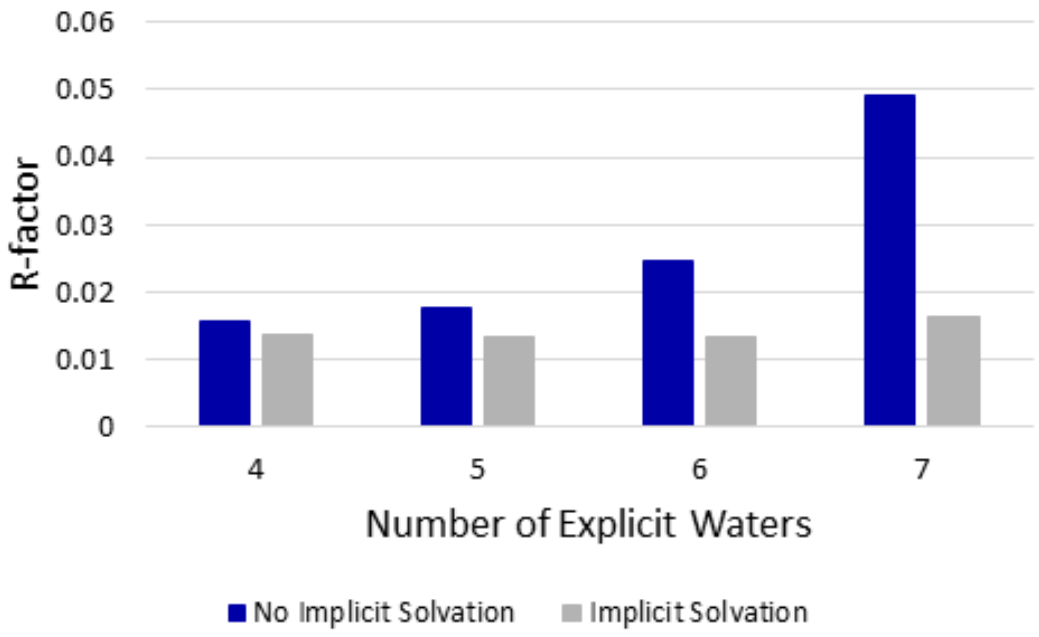

Figure 8: EXAFS statistical goodness of fit (R-factor) using DFT optimized structures as structural inputs. Simulations were run with 4-7 explicit waters and implicit solvation.

Figure 8 suggests that inclusion of implicit solvation is important to maintain a low R-factor, particularly when many explicit water molecules are included in the simulations. Therefore, to develop accurate models of aqueous Ln-EDTA complexes in this work, implicit solvation is included in addition to explicit water molecules.

Figure 9: The La-EDTA DFT optimized structure using a stick model to show the EDTA molecule and a ball and stick model to show coordinated water molecules where outer sphere water molecules are transparent and nitrogen=blue, carbon=gray, hydrogen=white, oxygen=red, lanthanum=black

3 RESULTS AND DISCUSSION

3.1 Aqueous Structures Determined by Molecular Dynamics and Experimental EXAFS

3.1.1 MD Derived Aqueous Ce-EDTA Complex Structure

The average aqueous structure of Ce-EDTA derived from MD (average of several configurations from the MD simulations) is found to have 9 coordinating atoms in the first shell of the Ce ion in the aqueous Ce-EDTA complex, consistent with literature. However, 6 of these coordinating atoms are O atoms of water molecules while the remaining 3 are O atoms of EDTA, inconsistent with previous published studies which suggest 3 of the 9 coordinating atoms are O atoms of water molecules while 4 are O atoms and 2 are N atoms of EDTA [8, 12, 16]. A snapshot of the configuration of derived from MD simulations is pictured in figure 10.

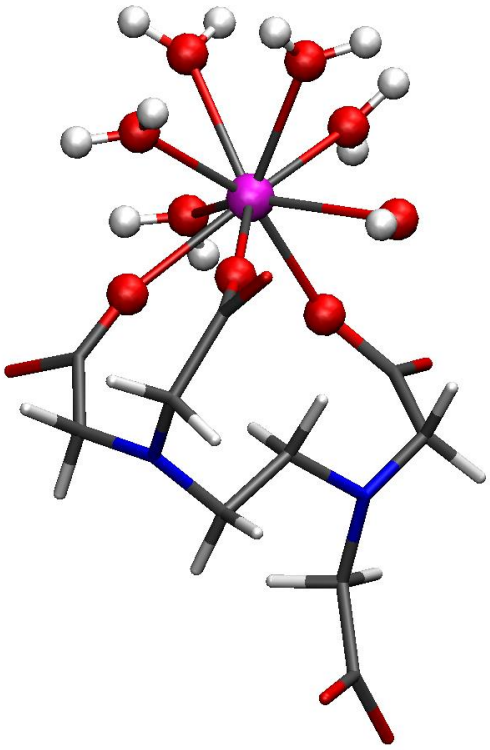

Figure 10: The Ce-EDTA MD structure using a stick model to show the EDTA molecule and a ball and stick model to show the coordinated water (element color key: red=O, blue=N, gray=C, white=H, pink=Ce).

3.1.2 Combined MD/EXAFS Approach

Once the MD derived structures are obtained, the coordinates are used to create a theoretical spectrum to fit against the experimental spectrum. This process is undertaken by our experimental collaborators. The preliminary combined MD and EXAFS fits indicate a poor fit which suggests that the MD derived structures do not align well with the experimentally derived structures. Particularly, the fits assess the bond distances between the Ln ion and atoms in its first coordination shell. Thus, a poor fit indicates that this coordination shell modelled in MD is not consistent with the experimentally observed coordination shell. This motivates the need to use DFT simulations instead.

3.2 Aqueous Structures Determined by Density Functional Theory and Experimental EXAFS

3.2.1 DFT Derived Aqueous La, Ce, Pr and Nd-EDTA Complex Structures

The preliminary MD/EXAFS fits suggest that quantum mechanic information is necessary to model the local environment of the aqueous Ln-EDTA complex accurately. The aqueous Ln-EDTA structures obtained using DFT are also found to have 9 coordinating atoms in the first shell of the Ln ions in their aqueous EDTA complexes (structures pictured in figure 11). For all the lanthanides, 4 of the coordinating atoms are oxygen atoms of EDTA, 2 are nitrogen atoms of EDTA and 3 are oxygen atoms of water molecules, which agrees with literature.

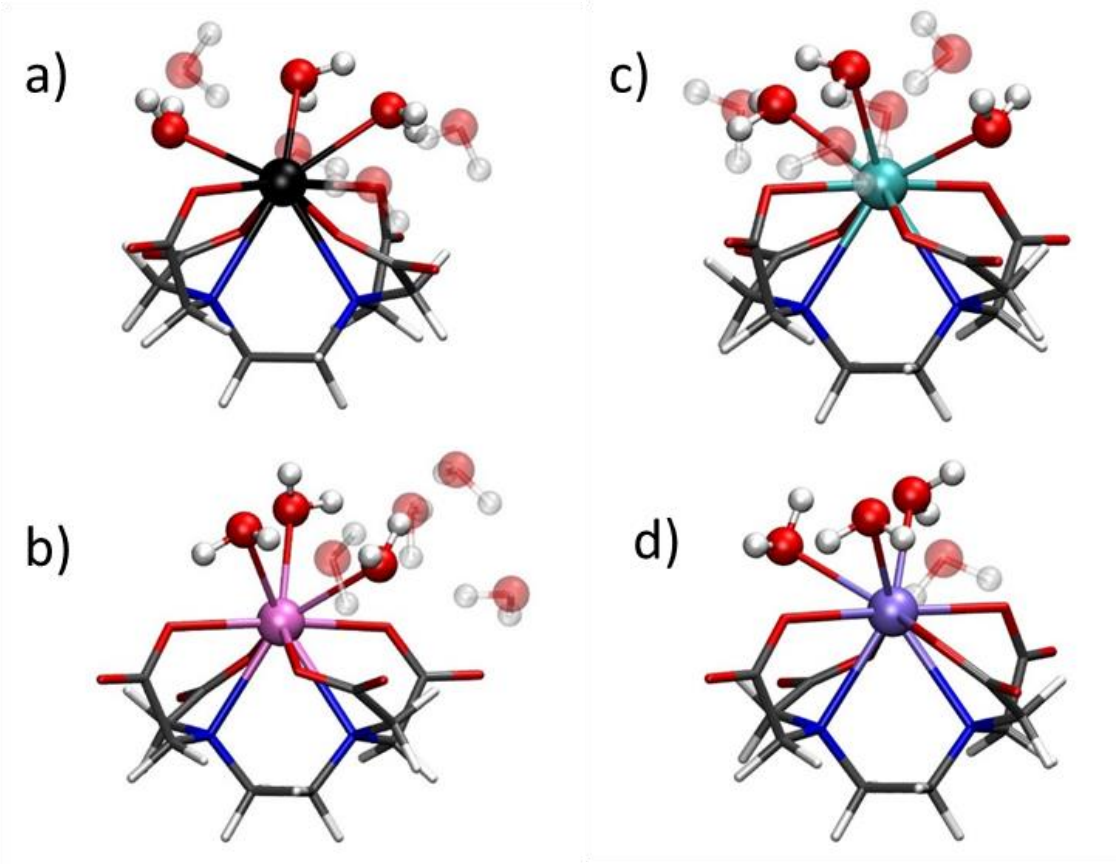

Figure 11: DFT optimized aqueous Ln-EDTA structures using a stick model to show the EDTA molecule and a ball and stick model to show the coordinated water molecules and outer sphere water molecules appear transparent (element color key: red=O, blue=N, gray=C, white=H, black=La, pink=Ce, teal=Pr, purple=Nd).
a) La-EDTA complex with 7 waters b) Ce-EDTA complex with 7 waters
c) Pr-EDTA complex with 7 waters d) Nd-EDTA complex with 4 waters

3.2.2 Combined DFT/EXAFS Approach

The coordinates of the DFT optimized structures are used to create theoretical spectra to fit against the experimental EXAFS spectra to uncover the structure of the first coordination shell of the lanthanides in their aqueous EDTA complexes. This process of fitting is conducted by our experimental collaborators. The quality of fits obtained using the DFT structures as theoretical structural references are indicated by an R-factor, which

is a statistical measure of the goodness of a fit. R-factors having a value of less than 0.02 are considered to be of high quality, which is found to be the case for all the aqueous Ln-EDTA complexes in this work. The interatomic distances between the lanthanides and atoms in their first coordination shell, obtained in this work using DFT and experimental EXAFS, is presented in figure 12.

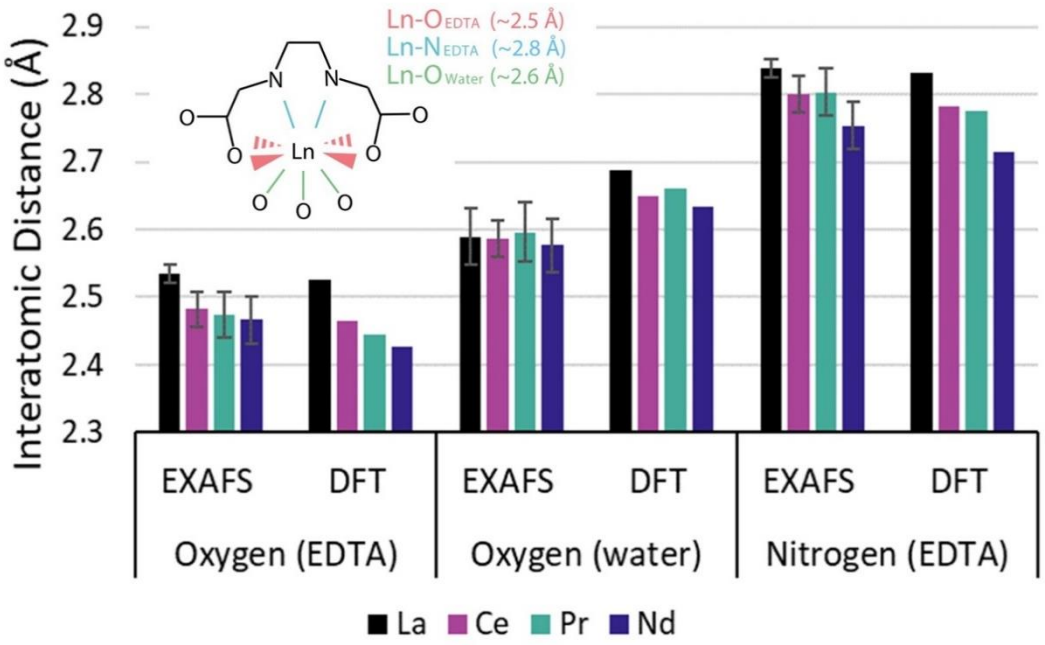

Figure 12: The average distance determined by EXAFS and DFT from the lanthanide atom to its inner sphere coordinating atoms: the four EDTA oxygen atoms, the three water oxygen atoms, and the two EDTA nitrogen atoms.

The trends of the bond distances are the same in both the DFT simulations and EXAFS experiments. This indicates that the DFT derived computational structures reasonably model the Ln-EDTA complexes as observed in aqueous solution. The trends show that the interatomic distances between the coordinating atoms of the ligands and the lanthanides decrease as we move across the lanthanide series. This trend has been observed in other

works involving lanthanides bound by ligands and is attributed to the decreasing ionic radius across the lanthanide series, also called as lanthanide contraction [25]. This agreement with experimental data and published work further strengthens the credibility of the DFT derived models as reference structures of aqueous phase Ln-EDTA complexes.

3.3 Comparison to other structural determination work in literature

3.3.1 Comparison with structural determination using X-ray diffraction

Prior works in literature studying complexation of lanthanides by EDTA have used solid phase X-Ray Diffraction (XRD) to determine structures [9-11, 26]. The best fit bond distances obtained in this work are compared to these XRD results from literature. This comparison is summarized in figure 13.

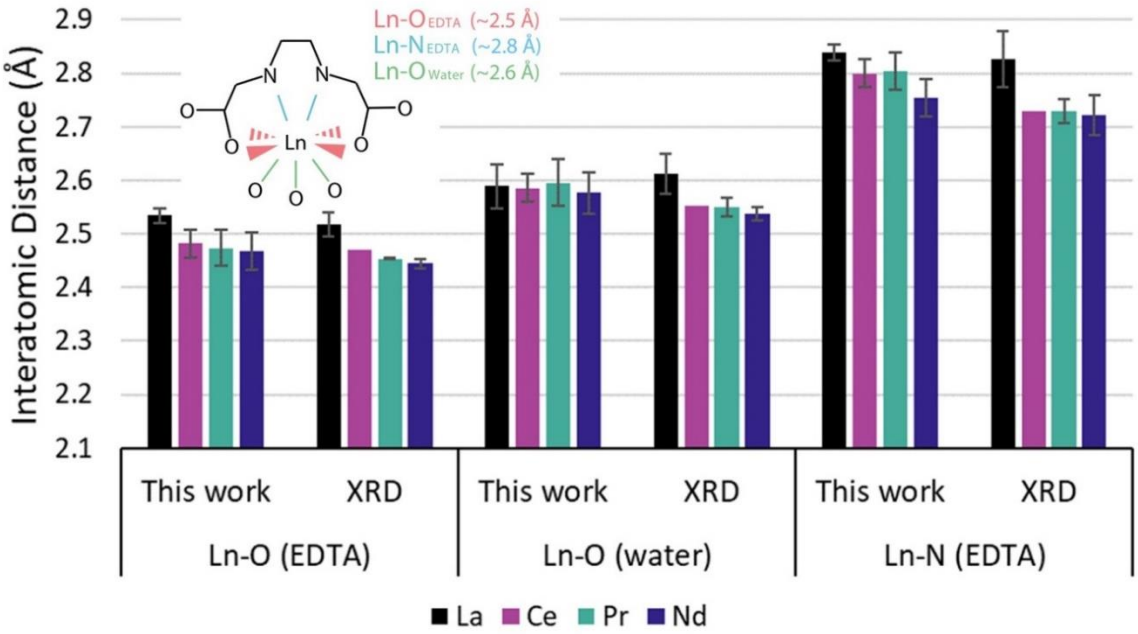

Figure 13: The average distances from the lanthanide atom to coordinating atoms as determined by the best fit EXAFS results of this study and by single crystal XRD from literature.

In the XRD works, the first coordination shell of the lanthanide consists of 7 oxygen and 2 nitrogen atoms, suggesting a similar conformation to that found in this work. However, the comparisons suggest that the XRD distances are consistently shorter than the best fit bond distances calculated in this work. This is especially the case for water oxygen and EDTA nitrogen bond distances. The shorter distances between the coordinating atoms to the lanthanides in the XRD results could be attributed to a lack of solvent after crystallization in the XRD method [27].

3.3.2 Comparison with structural determination using molecular simulations

Some studies in literature have exclusively used molecular simulations to determine structures of aqueous Ln-EDTA complexes. A study using molecular dynamics (MD) determined the bond distances of the coordinating atoms to the lanthanide ions to be consistently ~0.1 Å shorter than the best fit distances calculated in this study [7]. This discrepancy could be a result of the use of classical (i.e., not quantum) force fields in the MD calculations or due to thermal disorder. Another study using ab-initio molecular dynamics (AIMD) determined distances between EDTA oxygen and nitrogen atoms to the lanthanides to be like those calculated in this work [8]. These comparisons suggest that the bond distances are influenced by the methods used to undertake molecular simulations.

4 CONCLUSION AND FUTURE WORK

In this work, we begin to develop a molecular level understanding of aqueous lanthanide-ligand systems. To do this, we determine the structures of aqueous Ln-EDTA complexes for the early lanthanides La, Ce, Pr and Nd by developing computational models of these systems using MD and DFT. These computational models are supplemented with data from EXAFS experiments performed by our collaborators.

Previous works in structural determination of aqueous Ln-EDTA complexes have used experimental techniques such as XRD or computational techniques like MD, DFT and AIMD individually. In this work, we have used DFT derived structures and combined this with EXAFS data to determine the aqueous solution phase structure of Ln-EDTA complexes eliminating the need to rely on solid phase experimental data. The results obtained in this work have been compared to results published in literature.

The EXAFS data is initially fit against both the MD and DFT data individually. The MD based structures result in poor fits against the experimental data. Based on these preliminary fitting results, the structures of aqueous Ln-EDTA complexes are determined using a combination of DFT and EXAFS. The EXAFS fits against the DFT structures are of high quality, indicated by an R-factor of less than 0.02. The experimental and computational results in this work show similar trends in bond distances. This indicates that the DFT derived computational structures reasonably model the Ln-EDTA complexes as observed in aqueous solution.

The EDTA complexes of the lanthanides La, Ce, Pr and Nd in this work are determined to be 9 coordinate in the first coordination shell. 6 of these sites are occupied by EDTA

atoms (4 carboxyl oxygen atoms and 2 nitrogen atoms) and 3 sites are occupied by oxygen atoms of water molecules. This is consistent with several works in literature.

The structural determination technique in this work, using a combination of DFT and EXAFS, is able to reasonably model the benchmark systems of aqueous phase Ln-EDTA complexes. Discrepancies in our results with published works have been discussed and can be primarily attributed to the difference in research design and methods. Thus, this combined DFT/EXAFS method can be extended to a variety of ligand-lanthanide-water systems for applications in selective REE recovery.

Eventually, calculations based on these derived structures can be undertaken to determine the thermodynamics of complexation of lanthanides with ligands in aqueous environments. Moreover, the determined structures can be used to provide insight into the thermodynamics of complexation and in the process develop a structure-property relationship, which is a key step towards designing ligands for sustainable selective REE separations.

REFERENCES

[1] Department of Energy. *Rare Earth Elements 101*. Available: https://www.energy.gov/sites/prod/files/2020/02/f71/Rare%20Earth%20Elements%20Infographic.pdf.

[2] A. Schreiber, J. Marx and P. Zapp, "Life Cycle Assessment studies of rare earths production - Findings from a systematic review," *Sci. Total Environ.*, vol. 791, pp. 148257, Oct 15, 2021. . DOI: 10.1016/j.scitotenv.2021.148257.

[3] A. Schreiber, J. Marx, P. Zapp, J. Hake, D. Voßenkaul and B. Friedrich, "Environmental Impacts of Rare Earth Mining and Separation Based on Eudialyte: A New European Way," *Resources*, vol. 5, *(4)*, 2016. . DOI: 10.3390/resources5040032.

[4] K. Kovler, "8 - radioactive materials," in *Toxicity of Building Materials*, F. Pacheco-Torgal, S. Jalali and A. Fucic, Eds. Woodhead Publishing, 2012, pp. 196-240Available: https://www.sciencedirect.com/science/article/pii/B9780857091222500085. DOI: 10.1533/9780857096357.196.

[5] K. Zhang, Z. Dai, W. Zhang, Q. Gao, Y. Dai, F. Xia and X. Zhang, "EDTA-based adsorbents for the removal of metal ions in wastewater," *Coord. Chem. Rev.*, vol. 434, pp. 213809, 2021. Available: https://www.sciencedirect.com/science/article/pii/S0010854521000436. DOI: 10.1016/j.ccr.2021.213809.

[6] B. Kronholm, C. G. Anderson and P. R. Taylor, "A Primer on Hydrometallurgical Rare Earth Separations," *Jom*, vol. 65, *(10)*, pp. 1321-1326, 2013. Available: https://doi.org/10.1007/s11837-013-0718-9. DOI: 10.1007/s11837-013-0718-9.

[7] S. Durand, J. Dognon, P. Guilbaud, C. Rabbe and G. Wipff, "Lanthanide and alkaline-earth complexes of EDTA in water: a molecular dynamics study of structures and binding selectivities," *J. Chem. Soc. , Perkin Trans. 2*, *(4)*, pp. 705-714, 2000. Available: http://dx.doi.org/10.1039/A908879B. DOI: 10.1039/A908879B.

[8] R. D. O'Brien, T. J. Summers, D. S. Kaliakin and D. C. Cantu, "The solution structures and relative stability constants of lanthanide–EDTA complexes predicted from computation," *Phys. Chem. Chem. Phys.*, vol. 24, *(17)*, pp. 10263-10271, 2022. Available: http://dx.doi.org/10.1039/D2CP01081J. DOI: 10.1039/D2CP01081J.

[9] X. Huang, X. Xu, W. Pan and R. Zeng, "Poly[aqua-[μ-N'-(carboxymethyl)ethylene-diamine--N,N,N'-triacetato]neodymium(III)," *Acta Crystallogr. Sect. E. Struct. Rep. Online*, vol. 64, *(Pt 9)*, pp. m1194, Aug 23, 2008. . DOI: 10.1107/S1600536808026445.

[10] R. Ragul and B. N. Sivasankar, "Synthesis, Structure, Antioxidant, and Antiviral Studies on N2H5[Ln(edta)(H2O)3].5H2O (Ln = Pr, Nd, and Sm)," *Synthesis and Reactivity in Inorganic, Metal-Organic, and Nano-Metal Chemistry,* vol. 43, *(4),* pp. 382-389, 2013. Available: https://doi.org/10.1080/15533174.2012.740736. DOI: 10.1080/15533174.2012.740736.

[11] D. Xiong, H. Chen, X. Yang and J. Zhao, "Hydrothermal synthesis and characterization of a new 1-D polymeric lanthanum ethylenediaminetetraacetate with less metal-aqua coordination: {[La(EDTA)(H2O)]2}n," *Inorg. Chim. Acta,* vol. 360, *(5),* pp. 1616-1620, 2007. Available: https://www.sciencedirect.com/science/article/pii/S0020169306005937. DOI: 10.1016/j.ica.2006.08.044.

[12] N. Sakagami, Y. Yamada, T. Konno and K. Okamoto, "Crystal structures and stereochemical properties of lanthanide(III) complexes with ethylenediamine-N,N,N',N'-tetraacetate," *Inorg. Chim. Acta,* vol. 288, *(1),* pp. 7-16, 1999.

[13] M. A. Zaitoun and C. T. Lin, "Chelating Behavior between Metal Ions and EDTA in Sol−Gel Matrix," *J Phys Chem B,* vol. 101, *(10),* pp. 1857-1860, 1997. Available: https://doi.org/10.1021/jp963102d. DOI: 10.1021/jp963102d.

[14] M. C. Yappert and D. B. DuPre, "Complexometric Titrations: Competition of Complexing Agents in the Determination of Water Hardness with EDTA," *J. Chem. Educ.,* vol. 74, *(12),* pp. 1422, 1997. Available: https://doi.org/10.1021/ed074p1422. DOI: 10.1021/ed074p1422.

[15] (). *PubChem Compound Summary for CID 6049, Edetic Acid*. Available: https://pubchem.ncbi.nlm.nih.gov/compound/6049.

[16] S. Durand, J. Dognon, P. Guilbaud, C. Rabbe and G. Wipff, "Lanthanide and alkaline-earth complexes of EDTA in water: a molecular dynamics study of structures and binding selectivities," *J. Chem. Soc. , Perkin Trans. 2, (4),* pp. 705-714, 2000. Available: http://dx.doi.org/10.1039/A908879B. DOI: 10.1039/A908879B.

[17] W. Yu, X. He, K. Vanommeslaeghe and A. D. J. MacKerell, "Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations," *J. Comput. Chem.,* vol. 33, *(31),* pp. 2451-2468, Dec 5, 2012. . DOI: 10.1002/jcc.23067.

[18] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. J. Mackerell, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *J. Comput. Chem.,* vol. 31, *(4),* pp. 671-690, Mar, 2010. . DOI: 10.1002/jcc.21367.

[19] S. Grimmel, G. Schoendorff and A. K. Wilson, "Gauging the Performance of Density Functionals for Lanthanide-Containing Molecules," *J. Chem. Theory Comput.*, vol. 12, *(3)*, pp. 1259-1266, 2016. Available: https://doi.org/10.1021/acs.jctc.5b01193. DOI: 10.1021/acs.jctc.5b01193.

[20] F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy," *Phys. Chem. Chem. Phys.*, vol. 7, *(18)*, pp. 3297-3305, 2005. Available: http://dx.doi.org/10.1039/B508541A. DOI: 10.1039/B508541A.

[21] M. Dolg, H. Stoll and H. Preuss, "Energy-adjusted ab initio pseudopotentials for the rare earth elements," *J. Chem. Phys.*, vol. 90, *(3)*, pp. 1730-1734, 1989. Available: https://doi.org/10.1063/1.456066. DOI: 10.1063/1.456066.

[22] M. Dolg, H. Stoll and H. Preuss, "A combination of quasirelativistic pseudopotential and ligand field calculations for lanthanoid compounds," *Theoretica Chimica Acta*, vol. 85, *(6)*, pp. 441-450, 1993. Available: https://doi.org/10.1007/BF01112983. DOI: 10.1007/BF01112983.

[23] D. A. Pantazis and F. Neese, "All-Electron Scalar Relativistic Basis Sets for the Lanthanides," *J. Chem. Theory Comput.*, vol. 5, *(9)*, pp. 2229-2238, 2009. Available: https://doi.org/10.1021/ct900090f. DOI: 10.1021/ct900090f.

[24] Y. Chen, R. Rana, T. Sours, F. D. Vila, S. Cao, T. Blum, J. Hong, A. S. Hoffman, C. Fang, Z. Huang, C. Shang, C. Wang, J. Zeng, M. Chi, C. X. Kronawitter, S. R. Bare, B. C. Gates and A. R. Kulkarni, "A Theory-Guided X-ray Absorption Spectroscopy Approach for Identifying Active Sites in Atomically Dispersed Transition-Metal Catalysts," *J. Am. Chem. Soc.*, vol. 143, *(48)*, pp. 20144-20156, 2021. Available: https://doi.org/10.1021/jacs.1c07116. DOI: 10.1021/jacs.1c07116.

[25] R. B. Jordan, "Lanthanide Contraction: What is Normal?" *Inorg. Chem.*, vol. 62, *(9)*, pp. 3715-3721, Mar 6, 2023. . DOI: 10.1021/acs.inorgchem.2c03674.

[26] J. L. Hoard, B. Lee and M. D. Lind, "On the Structure-Dependent Behavior of Ethylenediaminetetraacetato Complexes of the Rare Earth Ln3+ Ions1," *J. Am. Chem. Soc.*, vol. 87, *(7)*, pp. 1612-1613, 1965. Available: https://doi.org/10.1021/ja01085a037. DOI: 10.1021/ja01085a037.

[27] A. Levina, R. S. Armstrong and P. A. Lay, "Three-dimensional structure determination using multiple-scattering analysis of XAFS: applications to metalloproteins and coordination chemistry," *Coord. Chem. Rev.*, vol. 249, *(1)*, pp. 141-160, 2005. Available: https://www.sciencedirect.com/science/article/pii/S0010854504002577. DOI: 10.1016/j.ccr.2004.10.008.

APPENDICES

Appendix A

Example Gaussian16 input file for geometry optimization of La-EDTA with 7 explicit waters and implicit water solvent

```
#p opt um06l/gen/auto geom=connectivity int=grid=ultrafine
SCRF(Solvent=Water)
pseudo=read scf=(xqc,intrep,maxconventionalcycle=350)
pop=(full,nbo,hirshfeld)

Title Card Required

-1 1
 O                 -4.49670000   -0.68946800   -0.02291100
 O                  0.67849200   -3.49249000    2.10380700
 O                  4.03177600    0.33933900    0.60729200
 O                 -0.87013800    4.33114500   -0.40678800
 O                 -2.40356900   -0.74597000   -0.84456700
 O                  0.97853500   -1.71109900    0.76263600
 O                  2.26921100    0.83053300   -0.68637600
 O                 -0.90001300    2.11708400   -0.79856100
 N                 -1.29968200   -0.55779400    1.61373500
 N                  0.72761400    1.62033300    1.31586600
 C                 -0.64972500    0.12328600    2.73615500
 C                 -0.23709100    1.54058800    2.40938400
 C                 -2.69477200   -0.16240100    1.44002600
 C                 -1.15688300   -2.00774400    1.75243600
 C                  2.07725100    1.18857700    1.66587300
 C                  0.71402700    2.95802200    0.72480300
 C                 -3.28818500   -0.58190600    0.08761600
 C                  0.29277700   -2.47540600    1.55790800
 C                  2.89080400    0.76284600    0.43491700
 C                 -0.46813900    3.19499600   -0.22267800
 H                 -1.30437500    0.12370900    3.63224800
 H                  0.23788700   -0.46376100    3.01974400
 H                 -1.12274400    2.12501700    2.10928400
 H                  0.14588900    2.03507700    3.32740000
 H                 -2.75979600    0.93779800    1.47601200
 H                 -3.34054700   -0.54341400    2.25600700
 H                 -1.76631300   -2.49351400    0.97175500
 H                 -1.54073400   -2.37424100    2.72437300
 H                  2.02735900    0.30172100    2.31487000
 H                  2.63648700    1.95787400    2.23676100
 H                  1.62447100    3.07624500    0.11835100
 H                  0.74082000    3.75100300    1.49740500
 La                 0.02924200   -0.10412200   -0.74617200
 O                 -1.40711000    0.60571400   -2.95369500
 H                 -1.42624300    1.49767100   -2.55462400
 H                 -2.16370500    0.18583700   -2.49984400
 O                  3.23187800   -2.05212300   -0.85072700
 H                  3.73872000   -1.29240900   -0.50298200
```

```
H                  2.56621100   -2.18594100   -0.14469500
O                  1.64907200   -0.91676100   -2.70356800
H                  2.26157500   -1.43971500   -2.10672100
H                  2.14996300   -0.09310900   -2.77909700
O                 -0.66926200   -2.31809500   -2.06951200
H                 -1.60968600   -2.07780500   -1.97989400
H                 -0.40456500   -2.05437700   -2.96048800
O                  1.34579920    1.15357485   -2.43760691
H                  2.30045715    1.06469008   -2.38936073
H                  1.10971057    2.08404481   -2.42823641
O                  0.83095936   -2.92431850   -2.53109834
H                  1.76108473   -3.02264051   -2.74743041
H                  0.29677015   -3.21186482   -3.27511380
O                 -0.43037371    2.31394149   -2.99726528
H                 -1.23790226    2.71905360   -3.32189096
H                  0.31675945    2.63501614   -3.50746436

1 17 2.0
2 18 2.0
3 19 2.0
4 20 2.0
5 17 1.5
6 18 1.5 33 1.0
7 19 2.0 33 1.0
8 20 1.5 33 1.0
9 11 1.0 13 1.0 14 1.0
10 12 1.0 15 1.0 16 1.0
11 12 1.0 21 1.0 22 1.0
12 23 1.0 24 1.0
13 17 1.0 25 1.0 26 1.0
14 18 1.0 27 1.0 28 1.0
15 19 1.0 29 1.0 30 1.0
16 20 1.0 31 1.0 32 1.0
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34 35 1.0 36 1.0
35
36
37 38 1.0 39 1.0
```

```
 38
 39
 40 41 1.0 42 1.0
 41
 42
 43 44 1.0 45 1.0
 44
 45
 46 47 1.0 48 1.0
 47
 48
 49 50 1.0 51 1.0
 50
 51
 52 53 1.0 54 1.0
 53
 54

@../H-basis-set.gbs
****
@../C-basis-set.gbs
****
@../N-basis-set.gbs
****
@../O-basis-set.gbs
****
@../La-basis-set.gbs
****
```