A Biomimetic Approach for Light Harvesting using Lanthanide Binding Coiled Coils

by

Oliver James Daubney

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Chemistry
University of Birmingham
College of Engineering and Physical Sciences
January 2019

Abstract

Natural inspiration has led to the design and synthesis of a series of novel coiled coil peptides with lanthanide binding sites and chromophores for energy capture and transfer. Their structural and photophysical properties were studied and the complex with the most intense lanthanide emission was explored further as a model system. The use of different chromophores was desired to further expand the lanthanide sensitisation and redox behaviour already exploited by tryptophan. Therefore, a computational methodology for predicting emission spectra from molecular structure was developed. This was validated by comparing the predictions with experimental results for a series of polyaromatic hydrocarbons.

Using the model peptide system a small library of coiled coils was prepared with different chromophores in place of the tryptophan residue. This involved the synthesis of a novel non-natural amino acid and the inclusion of this species in the previously designed sequence. The resulting family of peptides were studied for both their structural and photophysical properties. Throughout this work various novel chromophore to lanthanide interactions have been discovered, expanding the field of energy transfer. This work has led the way for the preparation of a small light harvesting module based upon a coiled coil framework.

Acknowledgements

I would like to thank Dr. Anna Peacock for the opportunity to carry out the work herein reported. To Liam Cox and Graham Worth, thank you for the all the advice, help, but above all your honesty about work and the academic life. My thanks to Kevin Bown and Robert Reeve for your support and guidance throughout this project, I hope that the final results may prove useful to you.

To Scott White and Louise Male for your help with solving crystals of small and big molecules. To Rian Griffiths, thank you for organizing the informative workshops on mass spectrometry and having the patience to teach the fine art of electrospray ionization. Thank you to Hannah Campbell for the preliminary work on the crystallisation of the HC02 peptide, it greatly accelerated the work and I know setting up trays was not fun, and to Cecile Soto in the initial work on predicting the emission of anthracene. Thank you also Marcus Taylor for the friendship, great collaborative work on the models for emission prediction and exciting conversations on the potential of computational chemistry. I am very grateful to Chi, Peter and Cecile not only for their great help and advice but also for the friendship, unconditional support and great times spent together. Without you, mass spectrometry and NMR would not have been as welcoming and helpful. To Allen Bowden, thank you for looking after the chromatography lab and making sure that everything ran as smoothly as possible.

To the DSTL, my thanks for the funding which enabled the opportunity to work on this project, the experience provided and your patience. I would also like to thank Diamond Light Source for beam time and the staff of beamlines i03 and i04 for their support.

Louise Slope and Sarah Newton, thank you so much for putting up with me when I started, and all the help with training but mostly for your friendship, especially to Louise for being a friend and support in the good and more challenging times.

To my family for support over all these years because without you I would not be this insane. Thank you for listening to crazy things even when you did not understand them and nodding encouragingly when I sounded excited. Finally, my greatest thanks to my fiancée, Ana Almeida, without whom I would definitely not have stayed the course and whose support and kindness has been a light in very dark times. Thank you so much for sharing a lab, an office, many dinners, and soon our lives together.

Table of Contents

Chapter 1

INTRODUCTION

1.1 INTRODUCTION

By 2012 the estimated world energy demand had reached 17 TW and this requirement will continue to increase throughout the century as more of the world have access to developed lifestyles.[1] This energy requirement needs to be met in order to continue our economic and technological growth. There are typically two strategies to achieve this goal the first being to directly increase energy supply and the second is to decrease energy demand by improving device efficiency. While the former of these strategies has received a great deal of attention, the question of enhancing efficiency tends to be left to the commercial developers. Nevertheless, the pressing need for a reliable source of energy should itself generate research and innovation in both sourcing and using energy efficiently. While fossil fuels have provided a convenient and easy source of energy over the last 150 years, certain limitations of this approach are now understood, as well as, the folly of wasting such a valuable resource by only using it for combustion (c.f. synthetic precursors). Fossil fuels are a round-about route to access solar energy, however, direct solar flux provides the earth with 175 PW, greater than 10,000 times humanities current annual requirements (Fig. 1). Indeed, solar energy is the most abundant and readily available source of energy on earth. This has generated a great deal of scientific interest to determine if it is possible to efficiently harvest solar energy directly, without relying on the mechanisms used to generate fossil fuels or indeed renewable sources such as wood and wind.

Considering photovoltaics or solar panels, it is often thought that electricity generation is the primary goal of solar capture. To think this, is to stray from the original objective, which is to provide energy whether it be chemical or electrical. This drives an interest in the

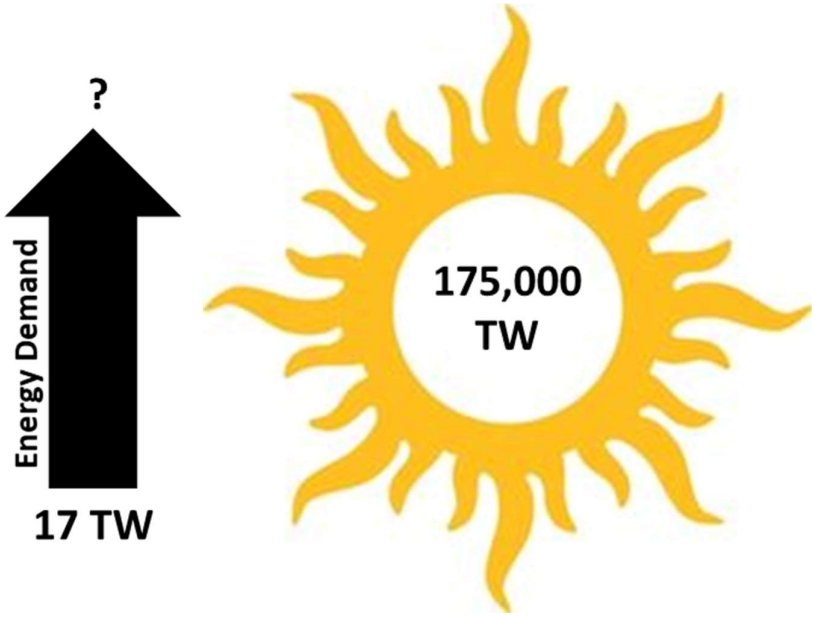

Figure 1: The present energy demand of humanity is around 17 TW and this is set to increase dramatically. The energy from solar flux to earth is estimated to be 175,000 TW which is significantly greater than present demand.[1]

'photosynthesis' of fuels and fine chemicals using light as the energetic source. Of course, the ideal system would be able to carry out both functions, to generate electricity or chemical potential dependent on requirements. It is interesting to note that a review by Oelgemoller *et al.* has shown the breadth of research into the direct use of sunlight for chemical synthesis via solar reactors.[2] While this is not a particularly versatile approach, it has tremendous potential to reduce industrial energy demands from conventional sources for fine chemical production. Before looking at the approach nature takes to harvest light, it is worth briefly considering current approaches to solar harvesting.

1.2 CURRENT WORK AND LIMITATIONS

Current work in the area of light harvesting, for the most part, focuses on electricity generation. This is achieved either by a form of photovoltaic cell which is based upon semiconductor

technology, or a dye-sensitised cell which is a derivative of a photovoltaic that makes use of a pigment to facilitate electron transfer to a semiconductor (Fig. 2).[3] Both these systems lead to solid state devices with relatively long operational lifetimes ~25 years. Though recent studies indicated that as long as the photovoltaic cell is cooled it could have both an extended working lifetime and increased efficiencies.[4]

At present, photovoltaics are the most efficient and effective solution to the problem of generating electricity from light.[5] Indeed, most of the efficiency problems can be hugely improved by careful engineering of the overall system. For example a panel of approximately 20% efficiency, if incorporated with a water cooling system will also harvest thermal energy raising efficiency whilst extending the life of the system.

Dye-sensitised solar cells are driving a new and possibly inexpensive interest in light harvesting.[6] The tailoring of pigments to harvest light and generate a potential difference is particularly exciting and mimics, in some part, what is achieved by nature. This is very useful when considering applications that do not just require a potential difference for electricity generation but where the charge separated state or photo-excited states could be used for catalysis. A large amount of research in this area has been dedicated to the use of ruthenium functionalised materials due to their high efficiencies, despite the obvious cost and availability limitations which result in impractical systems. While both these systems have great potential for electricity generation, the principles of a dye-sensitised solar cell are better suited for chemical production. In order to design a light harvesting system, tailorable absorption profiles for specific wavelengths of light will be required and water solubility will allow aqueous catalysis for use of the harvested energy to synthesise fuels or fine-chemicals.

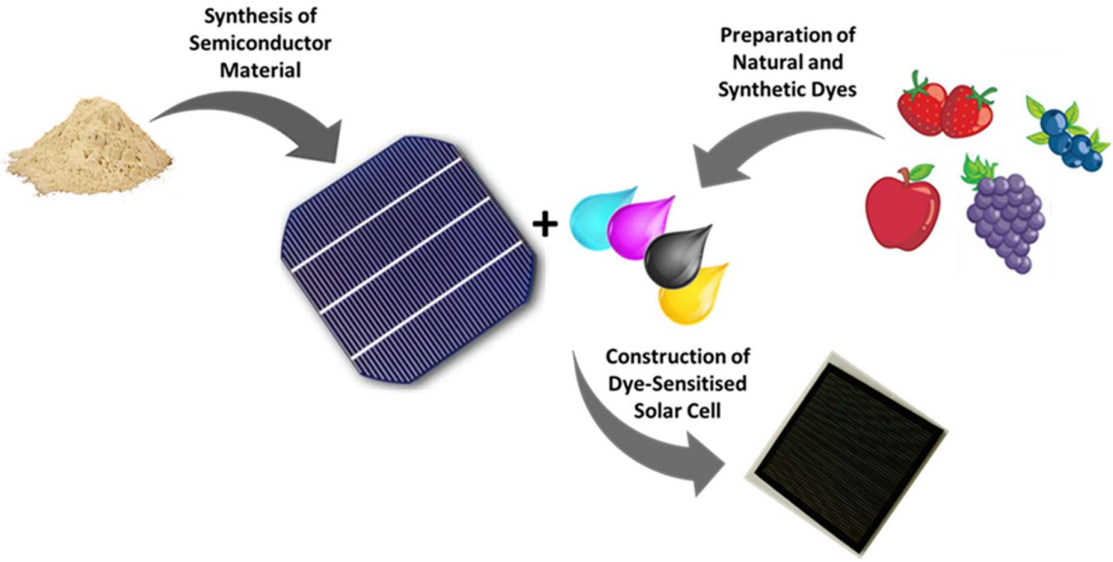

Figure 2: A representation of the preparation of a photovoltaic and a dye-sensitised solar cell. The photovoltaic makes use of a semiconductor to achieve charge separation, whereas a dye-sensitised solar cell utilises pigments coated on the semiconducting surface to enhance the efficiency of charge separation. Fruit pulps have been found to be relatively successful as the pigment layer.

1.3 NATURAL LIGHT HARVESTING COMPLEXES

Nature utilises complex protein structures to carry out specific functions with an elegant degree of control and selectivity. It also claims the most experienced and knowledgeable pedigree when considering light harvesting processes and photosynthesis. Interestingly, natural systems only tend to absorb between 3-6% of available solar radiation, however, the subsequent transfer and synthesis processes occur with efficiencies nearing unity.[7] These efficiencies are not just intrinsically necessary for the host organism but for almost every other form of life on the planet. To this end, it is worth focusing on a well characterised light harvesting apparatus, such as that of *Rhodoblastus acidophila* (Fig. 3).[8] This will allow the elucidation of key features to consider when designing a biologically inspired structure capable of harvesting light.

The light harvesting complex is a protein architecture that supports various coenzyme species which, due to its ingenious design, allows simultaneous multiphoton absorption and transfer to a reaction centre. While some chemical processes can be induced by the absorption of a single photon (usually of high energy), it is much more efficient to absorb multiple lower energy photons and then make use of the stored energy. This is for two reasons, firstly because the complex can make use of lower energy light (which is more abundant in solar radiation at sea level) and secondly, because it avoids damaging over-excitation pathways which result in photobleaching of the system. An example of the use of multiple photons is that every photosynthetic generation of an oxygen molecule (O_2) requires 8 photons (though there are some mechanisms that suggest 6).[9] Therefore, a light harvesting complex must be capable of both absorbing light efficiently and also holding onto the energy it has absorbed.

Considering the structure of the complex, it is immediately apparent that the system is highly symmetric being made up of nine repeat units that interact to form the familiar cylindrical shape of the scaffold (Fig. 4). This implies that the system can undergo self-assembly from supramolecular interactions between relatively simple subunits. The symmetry also implies that precise distances can be designed and replicated into a system by careful structuring of a subunit. This precise control of distance and environment is particularly important when placing chromophores into a structure, as their properties can be intrinsically manipulated to form an absorbing energy surface quite removed from their native state.

Secondly, the subunit is made up of two α-helices interacting to form a coiled coil dimer (Fig. 5). This structure has two specific environments; an external hydrophobic face with high membrane affinity for positioning in a chromatophore and a core with residues positioned to

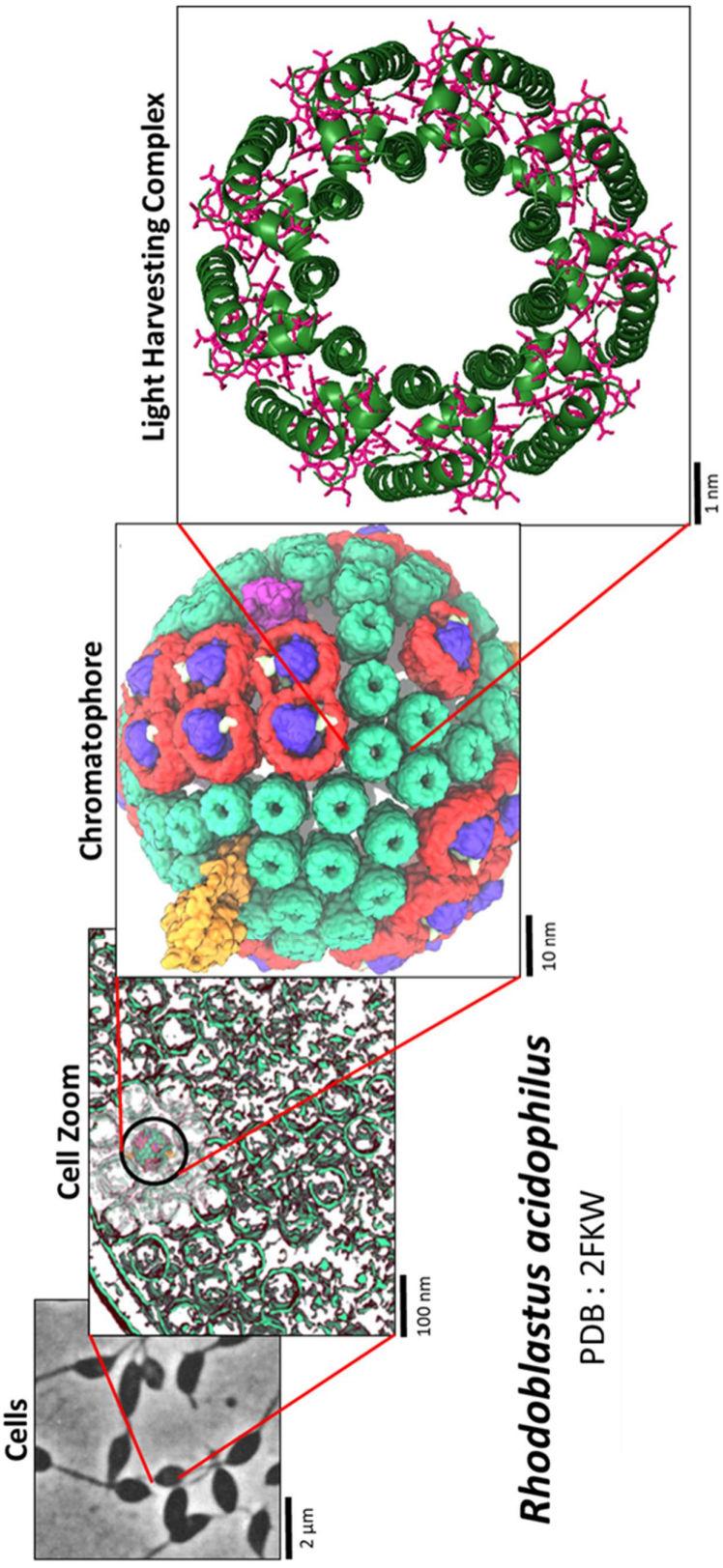

Figure 3: A photograph of the cells of Rhodoblastus acidophilus upon their discovery in 1969. A computational microscope overlay illustrating the cellular positions of the chromatophore and a zoom of the computational chromatophore structure with the photosystems (green), reaction centres (red and blue) and an ion channel (yellow). The light harvesting complex is the crystallographic structure which was determined from a sample prepared in a membrane (PDB: 2FKW). The protein structure (green) is furnished with chlorophyll derivatives (pink), the other cofactors are omitted for clarity.

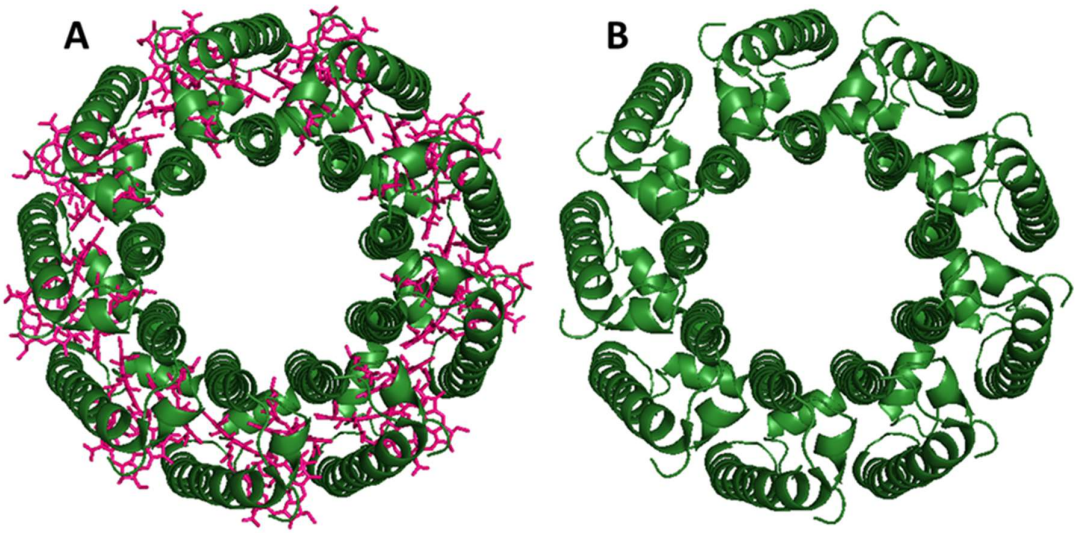

Figure 4: A cartoon representation for the crystal structure of the light harvesting complex from Rhodoblastus acidophilus (PDB: 2FKW) both A) with and B) without the chlorophyll derivatives for clarity. The nine-fold symmetry can be seen around the complex with a dimer repeat unit.

host various natural pigments. This internal environment has been well studied, with the selective hosting of specific molecules being probed both experimentally and computationally. Modelling by Noy *et al.* of the coordination of chlorins and haems onto these surfaces,[10] coupled with biological exclusion studies, have demonstrated that histidine residues are particularly useful to stabilise chlorophylls within the structure.[11] This has been further confirmed by x-ray crystal structures of the photosystem in a membrane, illustrating chlorophylls and related species (bacteriochlorophylls) within binding proximity to histidine residues.[12] The central Mg(II) ion is coordinated towards the nitrogen lone pair of the histidine imidazole ring. As each dimer binds three chlorophyll species, the symmetry of the whole light harvesting complex indicates that there are 27 chlorophylls present. Thus the subunit both positions and stabilises chromophores internally while also self-assembling into a complex machine within a membrane.

Thirdly, to study in detail the chromophores within the complex would take a great deal of time and has been the focus of many exceptional books and papers, yet it can be useful to note a few

elementary features.[13] While the focus has been on chlorophylls within the structure, there are other chromophores present such as carotenoids which we will consider for their stabilising effect. Carotenoids broaden the spectrum of light absorbed by the light harvesting complex and their short excited-state lifetimes are stabilised by energy transfer to chlorophylls, reducing losses due to solvent vibrations and thus increasing useful energy capture.[14] This proximity is also important for back transfer from chlorophylls as they can generate a reactive triplet which

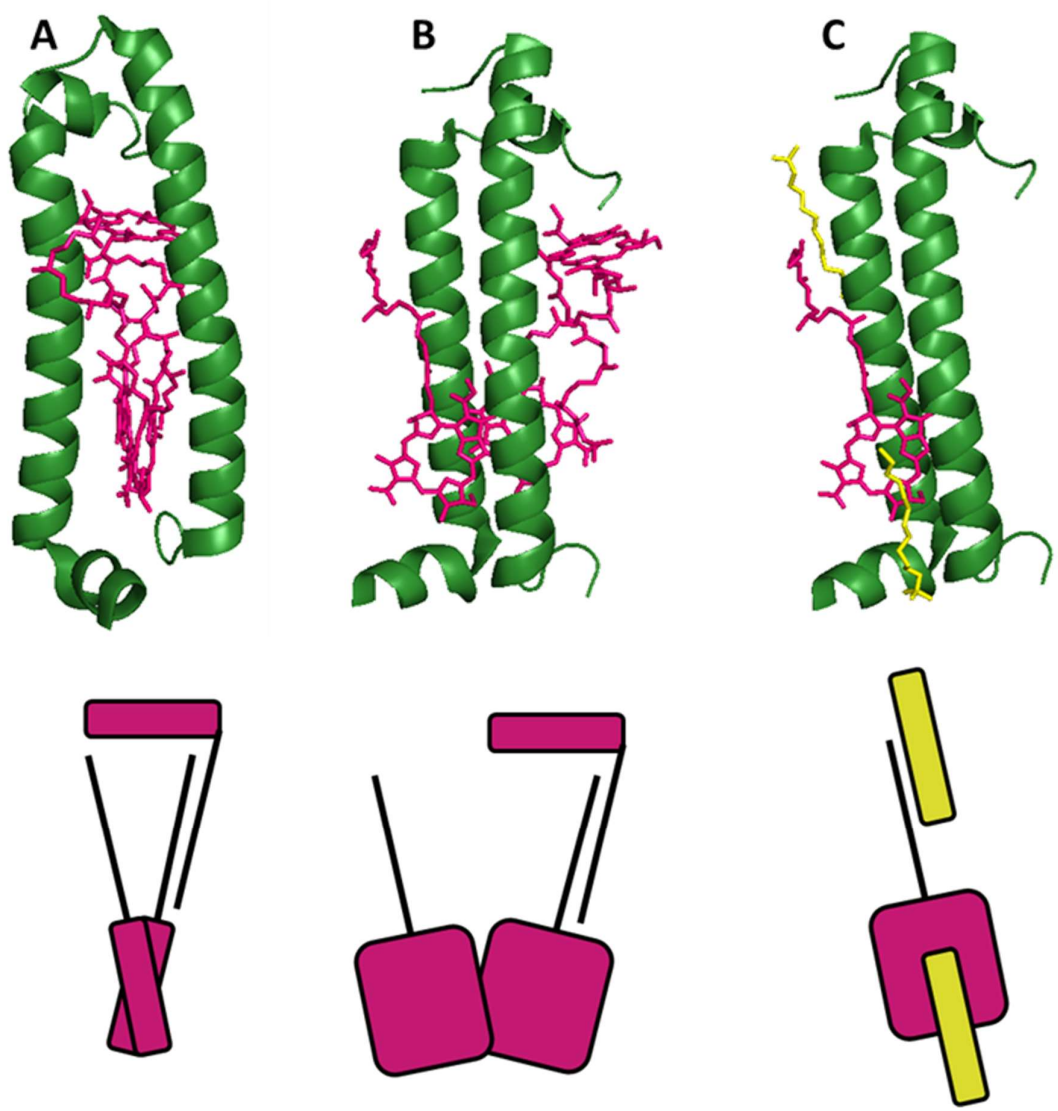

Figure 5: The dimer unit (green) illustrating the position of the chlorophyll derivatives (pink) and a cartoon depiction with both the head group and tail group interactions; A) side view, B) front view. C) The carotene cofactors (yellow) and their interaction with the chlorophyll derivative (pink) both as cartoon from the crystal structure (PDB: 2FKW) and cartoon schematic.

can interact with molecular oxygen to generate singlet oxygen.[14,15] The rapid quenching of this triplet by locally situated carotenoids helps prevent the build-up of reactive singlet oxygen in the photosystem. Thus, chlorophylls and carotenoids have a mutual interaction that enables high turnover of desirable excited states due to enhanced spectral sensitivity, while retaining protection against damaging over-excitation. The mutual behaviour of these chromophores and the behaviour of the whole light harvesting complex is quite unique. When chromophores are positioned correctly, close enough for orbital overlap, they cease to behave as a single entity and can be considered as one vast species undergoing orbital mixing.[16] Indeed, the closest idea to this in contemporary work is that of band-gap theory. For simplicity, one can consider the array of chromophores like a wire in which electronic excitation leads to simultaneous excitation at the reaction centre. Designing chromophores to mimic this is both challenging and interesting, as it moves away from the more traditional ideas of a single pigment dominating colour and spectral properties.[17]

Despite the elegance of the light harvesting complex architecture there are some fundamental limitations for light harvesting processes. This is mostly due to the electron transfer mechanism at the reaction centre with the chromophores arranged as a triad of species each carrying out a specific operation (Fig. 6). The donor absorbs the photon and becomes excited, the acceptor receives an electron via an electron transfer process resulting in charge separation, the stabiliser undergoes electron transfer to the donor thus stabilising the charge separated state. It has been found that the highest efficiency transfer occurs when the donor is situated as close to the acceptor as possible.[18] Spatial control in proteins allows this separation distance to be modulated, consequently tuning the energy transfer rates.

As previously discussed the light harvesting complex has a tailored design which optimises the utility of ambient light at sea level (and below).[19] One of the key problems for the development of light harvesting materials is the inefficient use of the full spectrum present at sea level. This can be overcome by synthetically incorporating metal centres and non-natural cofactors (chromophores and amino acids) into a protein structure to develop light harvesting proteins with enhanced energy capture and storage properties. To this end a versatile scaffold is required. Chromophores have been attached to a wide range of structures such as polymers,[20] gels,[21] and large dendrimeric antenna complexes.[22] However, in the case of polymers and gels it is almost impossible to achieve a monodisperse and well defined morphological system. Whereas with proteins it is common place, even essential, to have a purely monodisperse species of defined length and composition.[23] Perhaps one of the most appealing features of a protein derived system is the availability of the required elements such as carbon, nitrogen, oxygen and hydrogen for the main structure and metals such as magnesium, zinc, and iron for the reaction centres, all of which are inexpensive and abundant.[24]

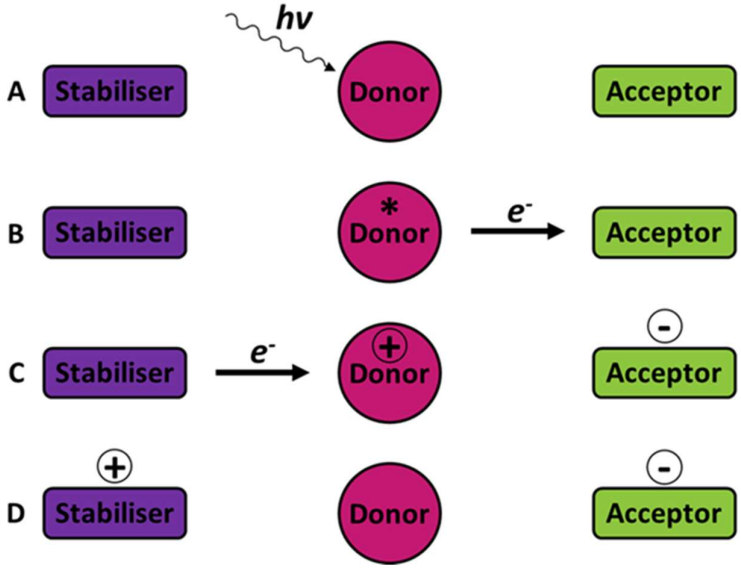

Figure 6: The process of stabilising charge separation in the reaction centre. A) The donor absorbs a photon and becomes excited, B) the donor undergoes electron transfer to the acceptor resulting in a charge separated state. C) The stabiliser undergoes electron transfer to the positively charged donor resulting in greater separation of the charges and thus D) stabilisation.[18]

The metal ions of interest in this thesis are the lanthanides. This is due to a number of key features regarding their photophysics which will be discussed in detail.

1.4 PHOTOPHYSICS AND THE LANTHANIDES

Before considering the lanthanides and their appealing photophysical properties it is necessary to discuss general photophysical principles (Fig. 7). The interaction of light with matter results in a number of processes, which will be considered in turn. The process of absorption involves the excitation of an electron in the molecule of interest from a low energy state to a higher energy state. The energy difference between these two states can only be equal to, or less than, the energy provided by the exciting photon. A couple of quantum mechanically derived selection rules restricts the absorption event from a lower state to a higher energy state. The first rule is that it is forbidden for an electron to alter spin state during a transition ($\Delta S = 0$), this is known as the spin rule. The second rule is that it is forbidden for an electronic transition to result in redistribution of electrons within a given subshell such as p, d or f orbitals if the species has a centre of symmetry ($\Delta l = \pm 1$), this is known as the Laporte rule.

A direct product of the forbidden nature of spin changes is that the most populated excited states have the same spin state as the original ground state electron, these excited states are known as singlets ($2S + 1 = 1$). It is possible to achieve states with inverted spin via intersystem crossing from the singlet state, these excited states are known as triplets ($2S + 1 = 3$). These excited states, both singlet and triplet, are unstable and will decay overtime by a number of potential mechanisms. The mechanisms of decay can be grouped into radiative and non-

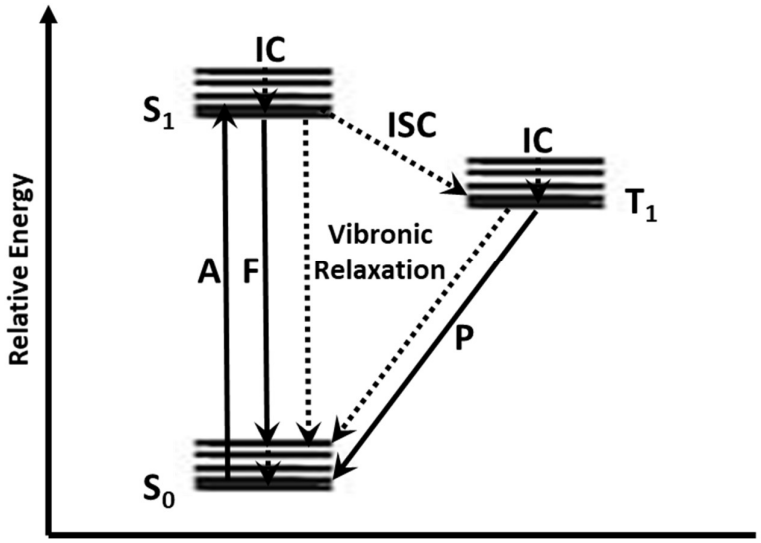

Figure 7: A Jablonski diagram illustrating key photophysical phenomena. The ground state (S_0), first excited singlet state (S_1) and first excited triplet state (T_1) are shown with vibrational levels. The radiative absorption (A), fluorescence (F) and phosphorescence (P) processes are shown as solid lines. The non-radiative processes of internal conversion (IC), intersystem crossing (ISC) and vibronic relaxation are shown as dashed lines.

radiative processes, with associated rates (k_r and k_{nr} respectively). The radiative deexcitation process is fluorescence for a singlet and phosphorescence for a triplet. The state transitions for emission as fluorescence or phosphorescence are restricted by the same selection rules as previously discussed. The non-radiative deexcitation processes are vibrational relaxation, intersystem crossing and intermolecular quenching. The emission deexcitation can be quantified by the lifetime of the state (τ) or time it takes to decay by 1/e. This can be expressed for experimentally observed fluorescence (τ_f^{obs}) by considering the rates of vibronic relaxation (k_{vib}), fluorescence (k_f), intersystem crossing (k_{isc}) and quenching (k_q). It is valuable to note the difference between the fundamental and fluorescence lifetime (τ_f) and experimentally observed lifetime (*see* Equations 1 and 2).

$$\tau_f^{obs} = \frac{1}{k_r + k_{nr}} = \frac{1}{k_f + k_q + k_{isc} + k_{vib}} \qquad \text{Equation 1}$$

$$\tau_f = \frac{1}{k_f} \qquad \text{Equation 2}$$

Another metric to consider with photophysical processes is the quantum yield which gives a measure to the efficiency of each of these processes. The fluorescence quantum yield (Φ^f) can be expressed by considering the number of photons absorbed and the number of photons emitted as fluorescence (*see* Equation 3).

$$\Phi^f = \frac{\text{Number of Photons Emitted}}{\text{Number of Photons Absorbed}} \qquad \text{Equation 3}$$

This is a useful equation to consider the efficiency of the fluorescence process, however, it can be more useful to consider the mechanisms that reduce the efficiency of fluorescence. This can be expressed as a function of the rates of fluorescence (k_f), quenching (k_q), intersystem crossing (k_{isc}) and vibrational relaxation (k_{vib}) to give the fluorescence quantum yield (*see* Equation 4). It should be noted that this equation does not consider dipole-dipole energy transfer mechanisms for simplicity, however, a term can be added to take this into account. These rates (k_n) are fundamentally linked to the previously discussed lifetimes (τ_n) by equation 5.

$$\Phi^f = \frac{k_f}{k_f + k_q + k_{isc} + k_{vib}} \qquad \text{Equation 4}$$

$$k_n = \frac{1}{\tau_n} \qquad \text{Equation 5}$$

The lifetime of fluorescence emission is typically very short (1 ps – 100 ns) indicating that the rate of fluorescence is very fast. Interestingly, the lifetime of phosphorescence emission is

typically much longer (10 μs – 100 s) due to radiative transition being spin forbidden. Having considered the basic fundamental principles of absorption and emission it is useful to apply this to the lanthanides and their intrinsic features with regard to photophysical phenomena.

The absorption process for the lanthanides is due to 4f-4f transitions, however, these are forbidden by the Laporte selection rule, that electrons cannot be redistributed within their subshell of origin. The forbidden nature of these transitions result in very low molar absorption coefficients ($\varepsilon < 10\ M^{-1}\ cm^{-1}$) implying that it is hard to directly excite a lanthanide ion.[25] Therefore lanthanides are often indirectly excited via energy transfer from locally situated chromophores.[26] This process is known as sensitisation and the chromophore is a sensitiser. A Jablonski diagram can illustrate the general sensitisation of the visible emitting lanthanides;

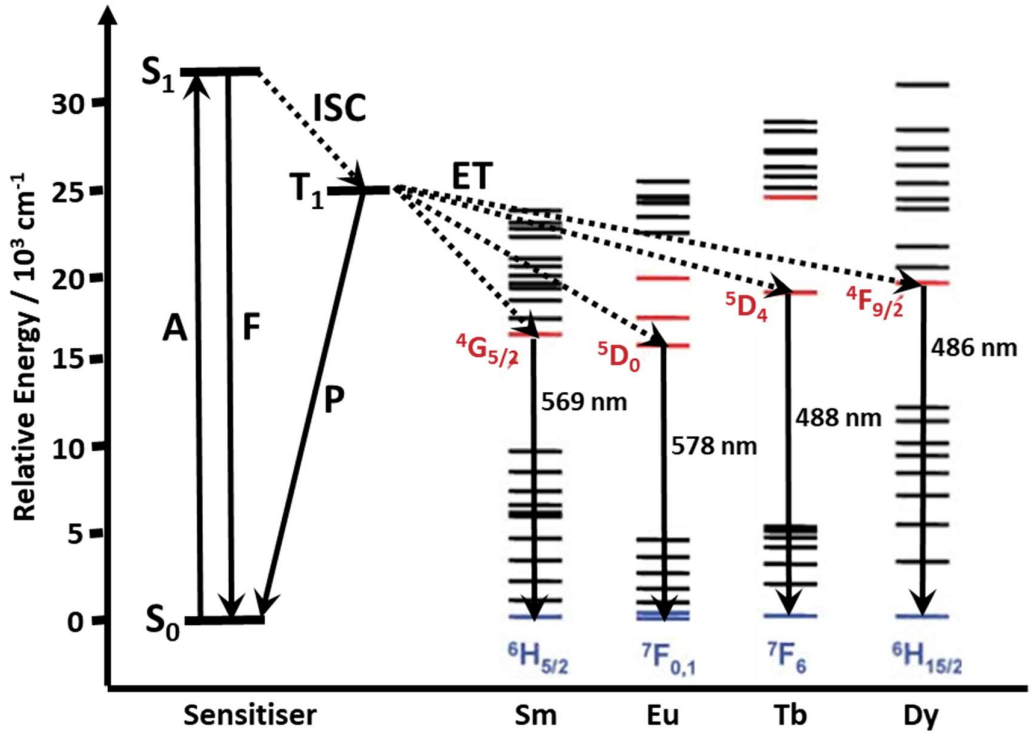

Figure 8: A Jablonski diagram with the ground state (S_0), first excited singlet state (S_1) and first excited triplet state (T_1) of a sensitiser are shown without vibrational levels for clarity. The visible emitting energy states for samarium, europium, terbium and dysprosium are shown with a single transition indicated for each lanthanide and the respective emission wavelength. The radiative absorption (A), fluorescence (F) and phosphorescence (P) processes are shown as solid lines. The non-radiative processes of intersystem crossing (ISC) and energy transfer (ET) are shown as dashed lines.

samarium, europium, terbium and dysprosium (Fig. 8). The chromophore is excited resulting in population of a singlet state. This state can decay via the processes previously discussed; however, the mechanism of interest is intersystem crossing to generate an excited triplet state.[27] As with the singlet, this can decay via emission (phosphorescence), non-radiative relaxation or energy transfer to the lanthanide of interest. If the decay is via energy transfer the result is a successfully sensitised lanthanide in an excited state. The efficiency of sensitisation (E_{sens}) can be described by the product of the intersystem crossing quantum yield (Φ_{isc}) and energy transfer quantum yield (Φ_{ET}) in equation 6.

$$E_{sens} = \Phi_{isc} \times \Phi_{ET} \qquad \text{Equation 6}$$

The efficiency of energy transfer is dependent on a number of key principles. The energy of the triplet state must be greater than the desired lanthanide excited state. It has been found that the triplet should be ideally 2500 – 3500 cm^{-1} greater than the lanthanide state.[28] If the energy difference is less than 1850 cm^{-1} then back transfer from the lanthanide excited state to the sensitiser triplet can occur reducing the efficiency of the sensitising process.[29]

The emission from lanthanide 4f-4f excited states is also forbidden, for the reasons previously discussed, which leads to the long lifetimes of these states as long as quenching processes such as solvent vibrations are minimised. The lanthanides have three classifications of emission profile. Firstly, the non-emissive lanthanum, lutetium and gadolinium do not have available 4f-4f transitions due to their electronic structure or the transition energy is high resulting in UV emission. Secondly, the visible emitting samarium, europium, terbium and dysprosium with distinctive spectral lines in the visible spectrum. The energy states for the visible emitting

lanthanides are shown in figure 8 with a main transition indicated. Finally, the remaining lanthanides have emission in the near-infrared region.

The ability to sensitise lanthanide emission using another chromophore was first reported by S. Weissman utilising a β-diketonate ligand to sensitise europium emission.[26] This led to further studies elucidating mechanisms of energy transfer,[30] which has been followed by the development of a wide range of small molecule lanthanide sensitising complexes. These

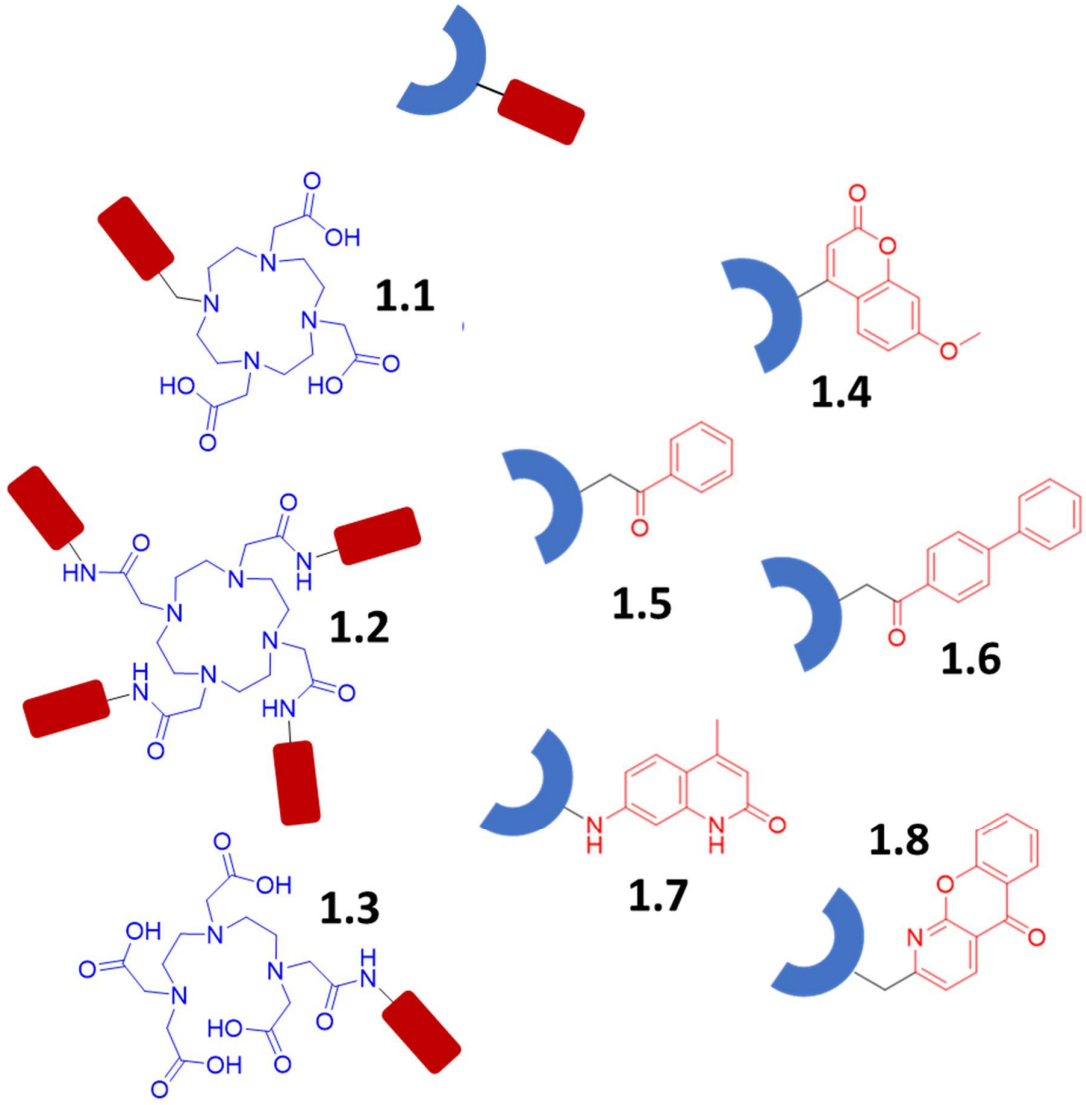

*Figure 9: A selection of polycarboxylate/polycarbamide coordination domains (**1.1** – **1.3**) and chromophores (**1.4** – **1.8**) typically utilised for preparing lanthanide sensitising complexes. The sensitising chromophores are derivatives of coumarin (**1.4**), benzaldehyde (**1.5**), biphenylaldehyde (**1.6**), carbostyril (**1.7**) and xanthone (**1.8**).*

complexes typically consisted of a chromophore connected to a coordination moiety, or as part of a coordination sphere to ensure close proximity to the bound lanthanide ion. These complexes, and the principles of sensitisation, have been extensively reviewed by Bünzli *et al.* in multiple publications.[31] Solution state visible emitting lanthanide complexes are of interest due to their high energy states and the variety of chromophores utilised as sensitisers, therefore a number will be considered here.

*Table 1: The triplet energy for the chromophores **1.4** – **1.8**.*

Chromophore	Triplet State Energy / cm^{-1}
1.4	26040 [33]
1.5	25300 [35]
1.6	20800 [35]
1.7	23500 [32]
1.8	24800 [34]

The first type of complex involves separate coordination and sensitising domains (Fig. 9). The lanthanides tend to form complexes with high coordination numbers (7 - 9) to hard ligands such as oxygen and nitrogen. This has led to common coordination domains involving the use of polycarboxylate/polycarbamide macrocycles (**1.1 – 1.3**), and related nitrogen chelators, with appended oxygen donor ligands such carboxylic acid and the carboxylate of amides. These ligands have been combined with various chromophores such as carbostyril,[32] coumarin,[33] xanthone,[34] benzaldehyde,[35] and biphenylaldehyde groups (**1.4 – 1.8**).[35] These chromophores were all used to sensitise emission from both terbium and europium trications. A consideration of their triplet state energy (*see* Table 1) demonstrates that they all have sufficient energy to sensitise both the $^5D_4$ state of terbium (>20500 cm^{-1}) and the $^5D_0$ state of europium (>17300 cm^{-1}). This assumes reasonable intersystem crossing yields to generate the triplet state. It is interesting to note that the triplet state energy for biphenylaldehyde at 20800 cm^{-1} was only

slightly greater than the $^5D_4$ state of terbium at 20500 cm^{-1}. This led to successful sensitisation; however, the quantum yield for terbium emission was reported to be low due to back transfer from the lanthanide to the sensitising chromophore.[35] This small collection of chromophores clearly demonstrates the importance of understanding the energetic profile, such as the triplet energy, of desired sensitisers, in order to predict efficient energy transfer to lanthanides.

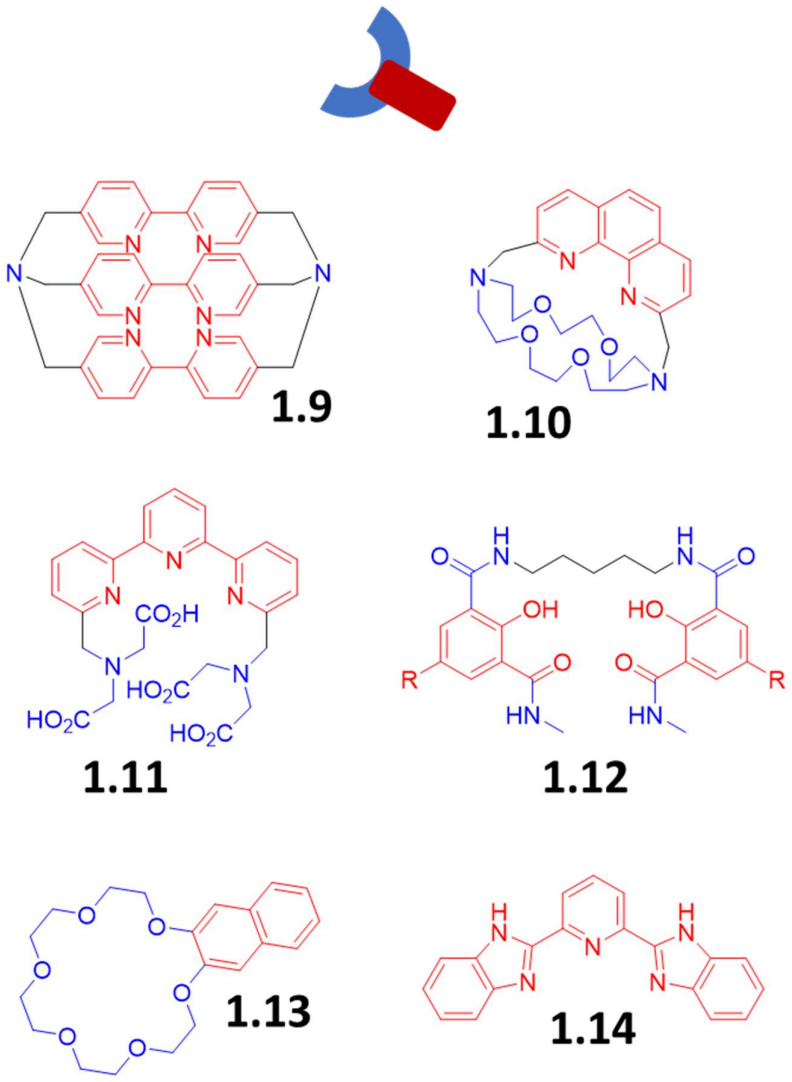

*Figure 10: A selection of lanthanide sensitising complexes with a cryptate (**1.9**), a phenanthroline derivative (**1.10**), tripyridyl complex (**1.11**), a substituted hydroxyisophthalamide ligand (**1.12**), a crown ether functionalised with naphthalene (**1.13**) and a nitrogen containing aromatic ligand (**1.14**). The chromophoric components are coloured red and the binding components are coloured blue. The structure of these complexes results in the chromophore also engaging in binding.*

The second type of complex is where the chromophore is directly involved in coordination to the lanthanide, usually via heteroatoms such as oxygen and nitrogen in an aromatic system. Preparation of cryptate complexes with pyridine ligands to encapsulate a europium trication, led to a large number of similar species such as a phenanthroline derivative (Fig. 10).[36] The cryptate **1.9** was utilised to probe the photophysics of bound terbium and europium to elucidate the mechanism of sensitisation.[30] The use of aromatic donor ligands has led to the preparation of multiple complexes such as **1.11**.[37] This motif has been used to append functional groups to alter the photophysics of the complex or provide sensing capabilities such as the addition of anthracene for singlet oxygen detection.[38] A good example of substituent effects on the photophysics of a lanthanide can be seen in the work of Samuel *et al.* with *p*-substituted 2-hydroxyisophthalamide ligands (**1.12**).[39] As with previously studied systems the energy of the triplet state was altered by varying substituents. The quantum yield of sensitised terbium emission was noted to increase as the energy of the chromophore triplet state increased. This was thought to be due to a decrease in energy back transfer as the energy gap between the terbium $^5D_4$ state and sensitiser triplet state increased.

The use of crown ethers has also been explored with an interesting study exploring the effect of orientation of a naphthalene sensitiser to the bound terbium or europium ion.[40] This demonstrated that the orientation between the chromophore and lanthanide ion is not trivial, however, for simplicity in calculations this is often approximated by considering the degrees of freedom for a tethered sensitiser.[41] The structures considered here involved a single lanthanide ion in a one-to-one ratio. However, there are a wide range of chromophoric molecules that form complexes with multiple ligands. Indeed, early complexes made use of *β*-diketone ligands which involved four monomers coordinating to a single lanthanide ion.[42] A

nitrogen donating aromatic ligand, **1.14**, was used to demonstrate the detrimental effects of back transfer from a bound europium ion to the sensitiser resulting in a decrease in energy transfer efficiency the more monomer was present.[43] Having considered the photophysics of lanthanides it remains to discuss the use of protein as a building block in design.

1.5 PROTEIN DESIGN PRINCIPLES

Natural proteins have been used directly to construct systems, indeed, an interesting light harvesting system was prepared that made use of a mutated or bioengineered light harvesting complex for electricity generation (Fig. 11). Zhang *et al.* were successful in incorporating a light harvesting system into a titanium dioxide nanotube and zinc oxide nanowire surface.[44] This was achieved by mutating the photosystem to include an amino acid sequence that had high affinity for zinc oxide,[45] allowing direct adhesion to the surface. Nanostructuring has led to vastly improved efficiencies due to the increased surface area provided by a honeycomb effect when using nanotubes. This approach has resulted in a system which markedly improved features such as photocurrent density of up to 362 μA cm^{-2} or 3.62 A m^{-2} when exposed to simulated sunlight. Previous studies using non-realistic conditions, such as intense

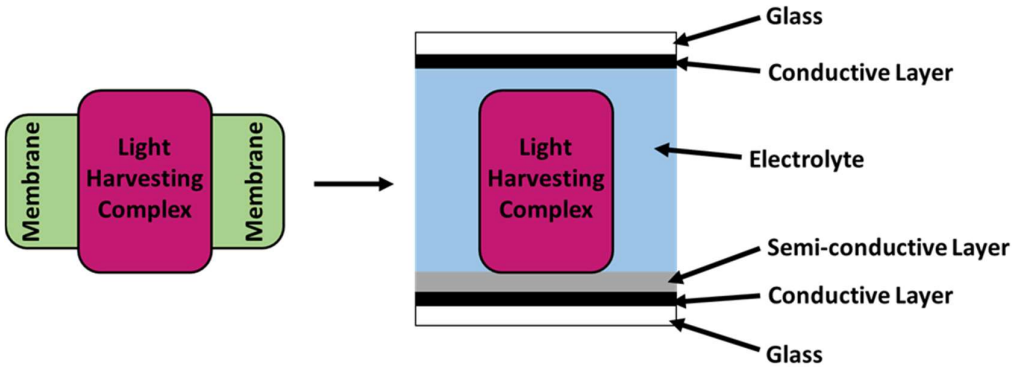

Figure 11: A schematic demonstrating the use of a light harvesting complex from a biological system in a biophotovoltaic system by being mounted on a semiconductor.

monochromatic light tuned to the absorption maximum of the light harvesting complex, only managed to achieve photocurrent density values of 0.3 μA cm^{-2} and 0.1 μA cm^{-2}.[46] This research has been further advanced by using graphene as a very thin conducting but transparent surface which can potentially allow layers to be stacked,[47] creating a membrane type surface with polyaniline stabilising the protein structure,[48] and the use of visible-NIR functional photosystems from purple bacteria to create low light active cells.[49] An indication of the interest in this field is the publication of reviews into both the use of photosystems[50] and natural pigments[51] for biophotovoltaic type systems. This has been successful for electricity generation, however, re-engineering complexes for advanced chemical behaviour such as catalysis would be challenging. It should also be noted that the field of bio-photovoltaics has been hindered by two key problems, the difficulty in isolating biological systems without denaturing proteins and the poor efficiencies of the resultant surfaces for light to electricity conversion.[52] Thus it may be more favourable to design a protein complex from the ground up.

Designing proteins from first principles allows the system to be purpose built, indeed, the elegance of protein design has attracted a great deal of interest for multiple reasons. The first of which must be the shear breadth of potential in this field. Nature itself is a testament to the versatility of proteins and what can be achieved using a relatively small library of amino acids. Researchers have used proteins for structure,[53] cofactor binding,[54] catalytic activity,[55] functional materials and much more.[56] All these applications rely upon a relatively small number of motifs in the secondary structure made up predominantly of α-helices and β-sheets. There are also regions of random coiling and loops, however, these are usually flexible and disordered. Both the two main types of secondary structure, α-helices and β-sheets can be designed by following certain established sets of rules.

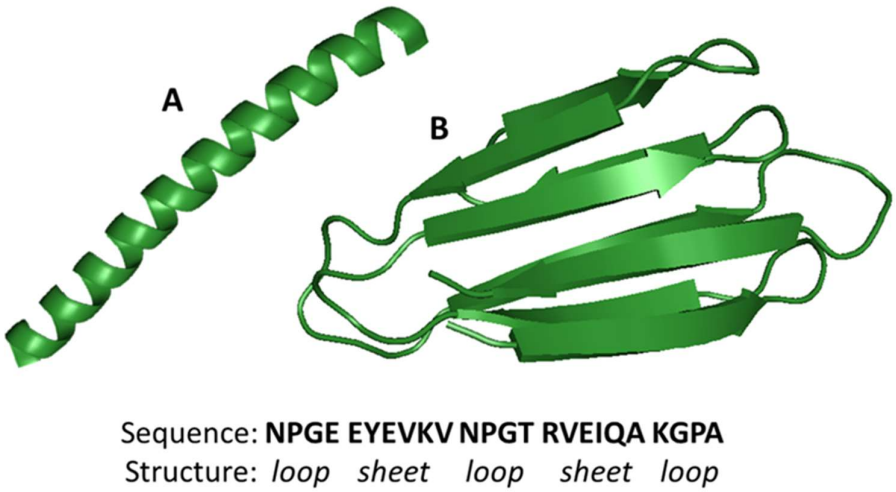

Figure 12: Cartoon representation of α-helix and B) a solution NMR structure of a de novo designed β-sheet protein (PDB: 6E5C). The sequence is for part of the β-sheet protein indicating the clear use of proline and glycine residues for the loops and alternative hydrophobic and hydrophilic residues for the sheet domains.

α-Helices are characterised by hydrogen bonding along the backbone between an N-H and a C=O group of amino acids positioned between 3 or 4 residues apart.[57] The helical coiling is right-handed due to geometric restraint of the left-handed amino acids.[58] The main design feature for a helix is a repeat of positive and negative polar residues around the helix to stabilise the turn via salt bridging while also using helix inducing residues such as alanine (Fig. 12).[59] One of the first synthetic examples of this was the work by S. Marqusee and R. L. Baldwin to synthesise a monomeric helix with glutamate and lysine residues stabilising alanine turns.[60] There are certain residues such as glycine and proline, which are heavily disfavoured for inducing helix formation and should be avoided in the design process.[61]

β-Sheets are characterised by lateral strands arranged adjacent to each other and again furnished with stabilising hydrogen bonds.[62] Indeed, this stability is often the cause of aggregation and particulate formation. There is very little literature on the specific design of β-sheets, possibly due to their tendency to aggregate and the greater variation in structural possibilities of even the simplest unit.[63] Nevertheless, work by Y. Yan and B. W. Erickson was

successful in designing and synthesising a β-sheet dimer using solid phase peptide synthesis.[64] The strategy adopted was to have a repeating unit of non-polar followed by polar residues for the length of a strand, followed by a few residues to execute a turn (Fig. 12), these could be combined to give the β-sheet component. Interestingly, large aromatic residues such as phenylalanine, tyrosine, and tryptophan are very prevalent in the core of natural β-sheets.

Both α-helices and β-sheets can be used and combined with loops and turns to create a huge number of exquisitely controlled three dimensional structures with precise spatial positioning such as β-hairpins, Greek keys, ψ loops, and helix-turn-helix domains to name a few (Fig. 13). However, this report will focus on the extensively studied and well understood topology of coiled coil structures which require only helix components, avoiding difficulties with β-sheet domains. The use of coiled coils is also appealing because there are excellent rules to guide their design and many examples to help inform new synthesis.

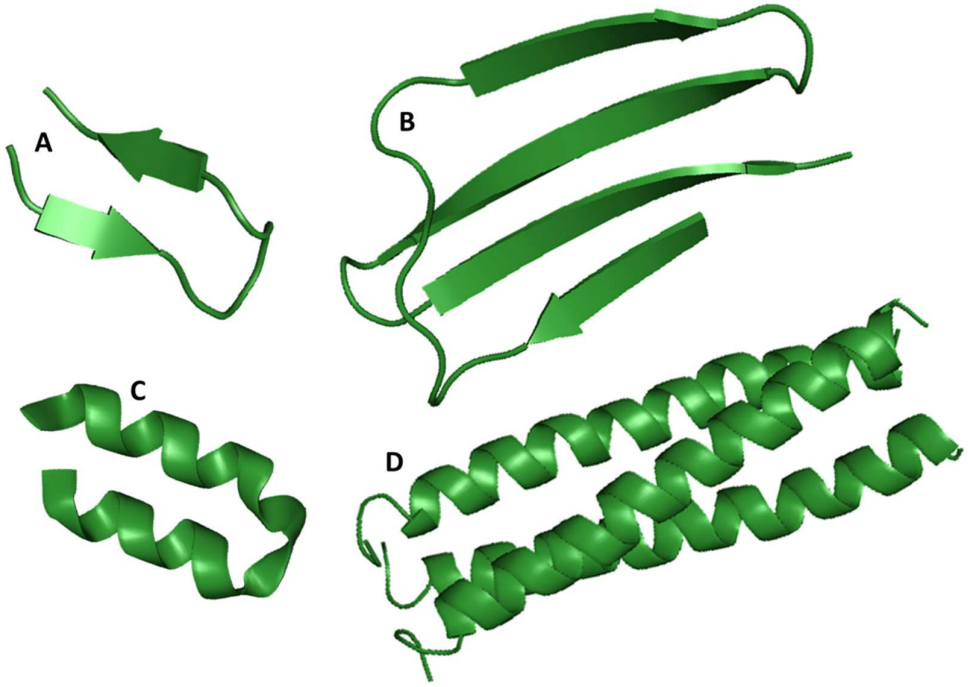

Figure 13: Cartoon representations for a number of different protein structural motifs; A) β-hairpin (PDB: 2EVQ), B) Greek key of p-hydroxybenzoate hydroxylase (PDB: 1PHH), C) helix-turn-helix or a de novo peptide (PDB: 1VRZ) and D) a coiled coil domain from a chicken cartilage matrix protein (PDB: 1AQ5).

1.6 COILED COILS AS STRUCTURAL BUILDING BLOCKS

Coiled coils are superhelical constructs of α-helices that weave together to optimise enthalpic (e.g. salt bridging) and entropic (hydrophobic packing) effects. Indeed, coiled coils are the spontaneous products of hydrophobic burying due to the hydrophobic residue, polar residue repeated unit required for a helical turn. Therefore it can be quite challenging to design a helix that will not supercoil in an aqueous environment. These constructs result in spatially well-defined environments that have both a tailorable hydrophobic core and hydrophilic surface (Fig. 14). The number of α-helices in a coiled coil can typically vary between two and up to seven, this is due to the loss of hydrophobicity in greater than heptameric coiled coils as water inclusion in the core destabilises the system.[65] The two distinctive chemical environments of the interior and exterior have the potential to make these structures an exciting multifunctional

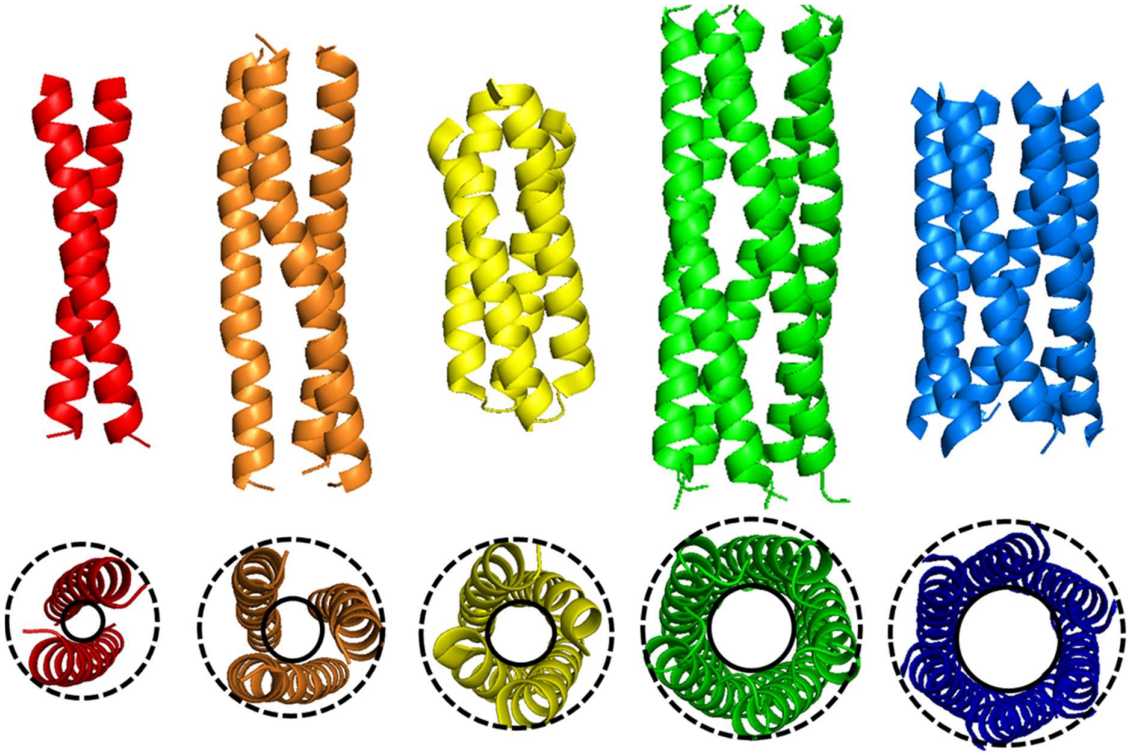

Figure 14: The structures of a series of coiled coils from dimer to hexamer illustrating the increase in the core hydrophobic cavity as more helices are incorporated. Hydrophobic (solid line) and hydrophilic (dashed line) surfaces are shown.

supramolecular building block.[66] This has resulted in extensive study of the design principles that can be used to tame this natural motif. Excellent reference material for the scope of these systems has been provided by the work of Woolfson *et al.*,[67] demonstrating the versatility of coiled coils for the protein engineer. Conveniently, biology has provided a plethora of natural systems that make use of the precise structure of coiled coils, which aided in both the discovery of the supercoiling phenomena and acted as a schoolmaster for uncovering many of the key design principles. Thus, before discussing the rules for designing coiled coils from a linear sequence, it is advantageous to understand the history of their discovery.

To uncover this story it is necessary to venture into the work of a titan in the field of structural chemistry, L. Pauling. As a champion of the modelling approach to solve chemical problems he led the discovery of the structure of a polypeptide to be α-helical.[57] While the helix was accepted rapidly, in part due to its elegance, it raised questions about the potential for interaction between helices to form superstructures. The next step involves one of the famous discoverers of the structure of DNA, F. Crick, who discussed ideas about the potential for supercoiling to occur, indeed he coined the phrase coiled coil. Both L. Pauling and F. Crick published independent mathematical proofs for the existence of a coiled coil structure in 1952.[68] However, it was not until 1983 that the first experimentally determined sequence of a coiled coil was reported by Hanukoglu *et al.*, a domain of keratin.[69] Following this many coiled coil domains were observed in natural systems but a single study was particularly important. In 1993, Harbury *et al.*, published a landmark paper illustrating that the oligomeric state of a coiled coil could be designed by selecting certain core residues (Table 2).[70] The bulkier the side chain the higher the oligomeric state, for example valine residues lead to dimer formation but isoleucine residues direct for trimer formation and the bulkiest leucine residues

Table 2: The GCN4 peptide mutants utilised to observe the effect of hydrophobic core residues on the directing on coiled coil oligomeric states.

Peptide	Sequence	Oligomeric State
GCN4-p1	Ac–R MKQ**L**EDK **V**EE**L**LSK **V**YH**L**ENE **V**AR**L**KKL **V**GER–NH_2	*Dimer*
GCN4-pII	Ac–R MKQ**I**EDK **I**EE**I**LSK **I**YH**I**ENE **I**AR**I**KKL **I**GER–NH_2	*Trimer*
GCN4-pLI	Ac–R MKQ**I**EDK **L**EE**I**LSK **L**YH**I**ENE **L**AR**I**KKL **L**GER–NH_2	*Tetramer*

direct for tetramers. This really opened up the concept of purpose built coiled coils, leading to many novel and bespoke systems being designed.

Considering the design principles the synthesis of coiled coil proteins has been made particularly accessible by following a linear design strategy which utilises a heptameric residue repeat unit (Fig. 15). The linear repeat unit would be $-(I_a\text{-}I_b\text{-}I_c\text{-}I_d\text{-}I_e\text{-}I_f\text{-}I_g)_n-$ where ***n*** is double the number of desired turns in the coil and ***I*** is the specific amino acid residues.[71] Hydrophobic residues are placed at the core positions and hydrophilic residues are placed in the external positions. For example, if one were to be constructing a trimer, hydrophobic residues such as

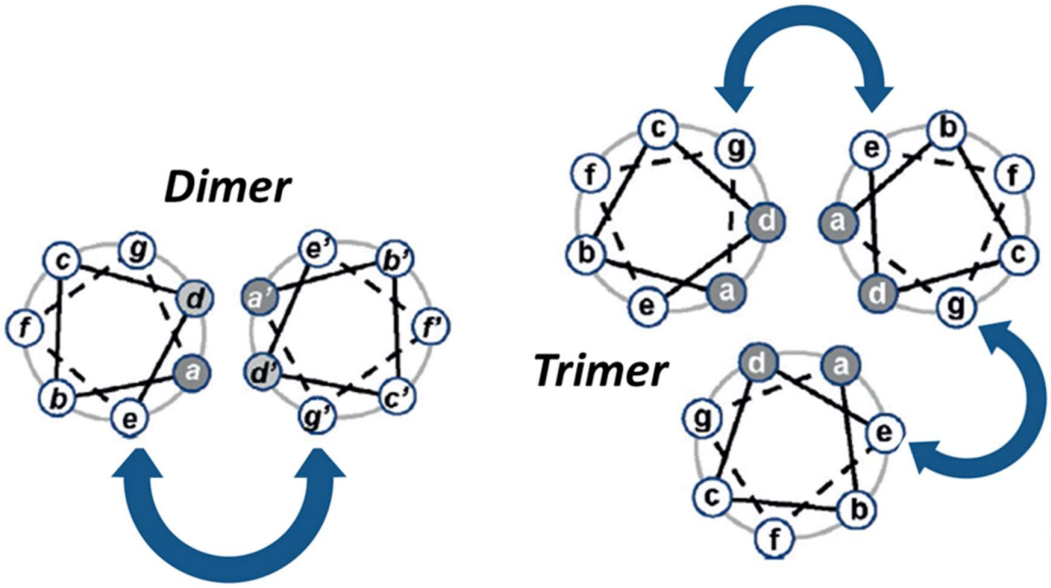

Figure 15: Top-down view of a schematic representation of the positions of linearly placed peptide residues in a coiled coil sequence, both a dimer and trimer super-helix. The blue arrows indicate salt bridging interactions and the shaded positions are involved in hydrophobic interactions within the core.

isoleucine would be placed at positions ***a*** and ***d***, as discussed previously these positions can be used to take advantage of the hydrophobic effect to direct the oligomeric state. Hydrophilic or helix inducing residues would be placed at positions ***b***, ***c*** and ***f***. Salt bridging is often employed at positions to encourage specific super helix formation, once again for a trimer this would be at positions ***e*** and ***g***. Structural design and synthesis of these systems have been extensively studied by the Woolfson Group among others, and they have collated a very useful database of coiled coils, their recent work has been expanding the potential oligomeric states available for designed coiled coils to include both new hexamers and heptamers.[72]

Synthetic coiled coil complexes have been designed, created and studied, from dimers to hexamers. Indeed, much of the work in coiled coil research now focuses on the use of the design principles to build functional systems, such as metal binding complexes.[73] Databases of these systems provide an invaluable resource when attempting to design a superhelix from natural amino acids. The principles of metal coordination have been applied to allow templating of coiled coil systems, taking a poorly folded system which upon inclusion of the metal folds efficiently.[74] Often the ligands required to bind a metal are not hydrophobic, therefore their placement has to be considered carefully in order to avoid excessive destabilisation. Interestingly, the binding of metals into a coiled coil is an extensive field with ligands being designed and placed for the binding of Pb(II), Hg(II), As(III), Bi(III), Zn(II), Cd(II), Ni(II), Co(II), and the lanthanides to name a few.[75] This has been advanced by the design of systems that use the metal to both template and carry out dynamic functions in the system, just like a protein.[76] This discussion will focus on coiled coil systems utilised for energy transfer or light harvesting.

Research conducted by Waters *et al.* has studied the energy transfer efficiencies between Ru(II) and Os(II) complexes mounted on a coiled coil scaffold.[77] The coiled coil dimers supramolecular structure was used to precisely position the donor/acceptor chromophores in space to facilitate energy transfer (Fig. 16). The peptide α-helices were decorated with bipyridyl ligands via click chemistry which in turn coordinated the desired metal ions. Excitingly the donor Ru(II) complex with light of wavelength 444 nm resulted in energy being transferred to the acceptor Os(II) complex. It was found that the closer the complexes were in space, the more efficient the energy transfer. The use of ruthenium in coiled coils is well established with a bimetallic coiled coil trimer that was capped at both ends with different metals by careful selection of ligands for the binding sites and steric design of the alpha-helix.[78] At one end a non-natural bipyridyl ligand which binds a Ru(II) ion as a trimer, while at the other end histidine residues bind a Cu(II) ion. While this is exciting because it utilises a combination of natural and synthetic motifs, the inclusion of a Ru(II) ion has gained particular interest because the metal complexes are known to have broad absorption profiles and hence are useful for light harvesting.[79]

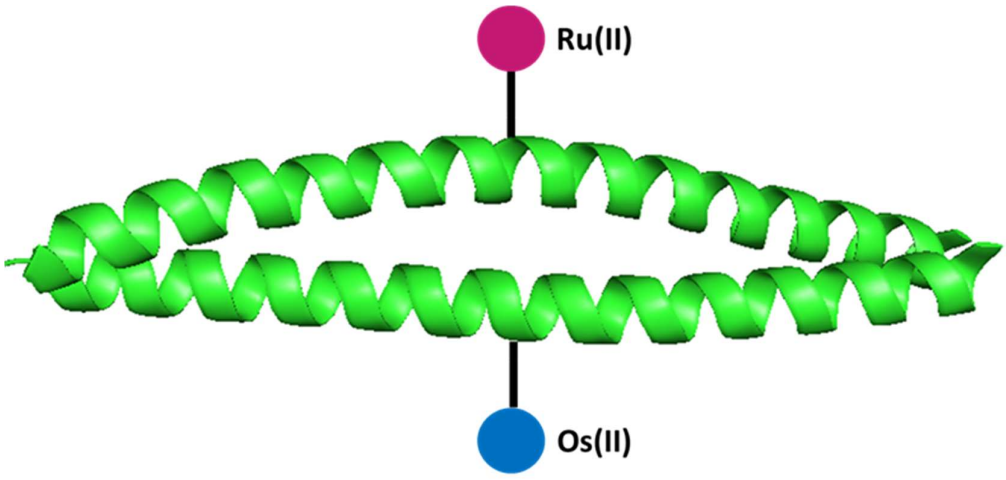

Figure 16: The peptide dimer systems used to position Ru(II) and Os(II) precisely in space and measure the energy transfer between the two species.

Various different chromophores have been coordinated to coiled coils and rather than using compounds analogous to chlorophyll a great deal of work has been done using zinc substituted chlorin and porphyrins. They have been shown to coordinate to a coiled coil tetramer through histidine ligation.[80] It was determined that the equivalence of histidines to bound zinc chlorins was 1:1 implying that only a single histidine residue was required to tether the chromophore. This was further confirmed by studies using pyridine as an axial ligand for another zinc chlorophyll derivative which stabilised as a 5-coordinate species.[81] Zinc chlorins have been particularly favoured because they retain the high absorption and energy transfer properties associated with chlorophylls, while being easier to handle and easier to coordinate.[82]

Characteristic work by the Dutton group involves the construction of coiled coil maquettes, made up of four α-helices, that play host to a multitude of cofactors (Fig. 17).[83] This species is of particular interest due to its apparent versatility for functional design from electron transfer between iron centred haems and zinc porphyrins, to zinc chlorophyll derivative pairs with enhanced absorption and high energy transfer rates. Previous work in this area made use of a maquette to bind four parallel haems which were shown to behave as pairs of interacting species exhibiting redox behaviour typical of native haems.[84] This work has been advanced by the specific inclusion of a number of structurally diverse haems, indicating the broad flexibility of the binding sites.[85] Zinc substituted chlorophylls have also been studied by Noy *et al.* and bound in a tetrameric coiled coil that stabilises an interacting pair.[86] This complex has been shown to behave in a similar way to a protein stabilised chlorophyll pair, with dramatically reduced molecular fluorescence and an increase in the relaxation kinetics in keeping with rapid energy transfer between the pair.

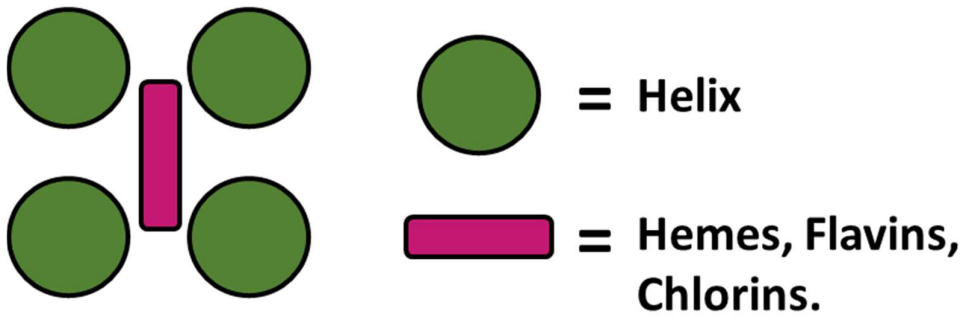

Figure 17: A schematic of the maquette with a tetramer coiled coil domain (green) and an internal cavity for sequestering various chromophores (pink).

The use of porphyrins has been explored because they have an extended π-system analogous to chlorophyll molecules and similarly broad/intense absorption profiles. Different metals such as cobalt and zinc have been incorporated into the porphyrin binding site and then coordinated to ligands in a coiled coil structure. The work of Ogawa *et al.* was to design a coiled coil dimer to attach external cobalt centred porphyrins, however, this design instead formed a trimer but also retained the capability of coordinating the metalloporphyrins.[87] The external coordination of the cobalt centred porphyrins was facilitated by pyridyl interactions. Two trimers were able to coordinate to the opposite axial faces of the metal resulting in superstructure formation. This was especially satisfying as it was possible to see the nanoscale effects by the resultant macroscale materials, fibre formation of around 1mm in length. Alongside this work a coiled coil dimer was used to construct macroscale materials when exposed to this same cobalt complex.[88]

While absorption of light and the transfer of energy are essential requirements of all light harvesting systems, it is equally important that charge separation can occur. This is both to make use of the harvested light and to prevent the system from photo-bleaching. Recent research has been carried out to place a naphthoquinone moiety in close proximity to chromophores in a coiled coil tetramer, creating a charge separation unit.[89] The

naphthoquinone species was placed precisely in a coiled coil structure by including a non-natural amino acid in the peptide sequence. Photoexcitation of a bound tetrapyrrole derivative resulted in a long-lived triplet that was capable of reducing naphthoquinone to the anionic semiquinone form. The charge separation was observed by a decrease in the lifetime of the tetrapyrrole triplet state as it was quenched by the naphthoquinone species. Interestingly, this effect was not observed when the naphthoquinone species had been reduced by ammonia prior to excitation.

All the coiled coil examples so far discussed lack application and are, at best, proof of concept. This last example is representative of a large number of catalytic systems attempting to mimic hydrogenases. It should be noted that the peptide is not photosensitive in these systems as it typically serves as a scaffold for the ferredoxin catalyst and is in solution with a sacrificial photosensitive agent. A popular method to chemically store harvested solar energy is to split water into hydrogen and oxygen. This has been achieved using hydrogenase active site mimics that split water via a ferredoxin catalyst. A recent example has employed a single helix which was functionalised with a non-natural dithiol amino acid residue that stabilised the diiron-hexacarbonyl reaction centre of the hydrogenase mimic.[90] When this was in aqueous solution with both a photosensitizer such as tris(bipyridine)ruthenium(II) and a sacrificial reducing agent such as ascorbate it catalysed the splitting of the surrounding water molecules, when excited with activating light. However, this system has a low turnover number typical of the hydrogenase mimics at present, which would suggest that a commercial device is not on the horizon. A review by Ghirlanda *et al.* has recorded the recent advances in hydrogenase mimics and clearly shows the value of a coiled coil scaffold to position and replicate the protein environment.[91]

1.7 SUMMARY

By observing natural systems, one concludes that the subunit design and selection of chromophores are essential for the efficient capture and use of energy. This involves careful structuring of both the steric and energetic environments to allow efficient packing of a designed peptide while retaining photophysical properties. Exciting developments in protein synthesis, especially the ease of solid phase synthesis, has driven research interest in attempting to mimic natures precise structural and energetic control. This project proposes that systems could be prepared that can harvest light to be used for both energy generation purposes and potentially for catalytic applications. The ease of designing coiled coils with reliable results implies that they are a sensible starting point in the design of a peptidic light harvesting module. Indeed, it is also important to note that the solid phase synthesis regime does not limit our designs to natural amino acids, but allows the use of interesting unnatural species with exotic properties and chromophoric handles.

1.8 AIMS AND OBJECTIVES

The aim of this project is to explore the feasibility of using coiled coils to place chromophores precisely allowing their photophysical characteristics to be studied. The considerations regarding the exquisite precision that can be achieved with the use of amino acids as building blocks informs the first objective of this project. This was to make use of the design principles for coiled coil proteins to prepare a small library of peptide structures to position chromophores and lanthanide ions precisely in space. The initial test structures will make use of tryptophan

as a sensitising species for emission from terbium ions. These structures will need to be characterised thoroughly to help understand their light harvesting potential. Secondly, the potential of these new peptides to sensitise other visible emitting lanthanides such as samarium, europium and dysprosium will be studied to evaluate their use as emissive tags and the capability of the biologically relevant amino acid, tryptophan, to sensitise their emission.

Following the design and construction of a model system with natural amino acids, the use and synthesis of non-natural amino acids functionalised with chromophores will be explored. This will result in the preparation of a small library of peptides functionalised with different sensitisers. These new peptides will be used to sensitise the visible emitting lanthanides. This will develop a system for studying the interaction between localised chromophores and bound lanthanide ions. Indeed, the behaviour of each of these systems will inform the future development of rudimentary coiled coil light harvesting complexes that perform other functions rather than simply reemitting the trapped energy.

1.9 REFERENCES

[1] A. Kleidon, L. Miller and F. Gans, *Solar Energy for Fuels, Topics in Current Chemistry*, 2016, **371**, 1-22.

[2] M. Oelgemoller, *Chem. Rev.*, 2016, **116**, 9664-9682.

[3] T. V. Arjunan and T. S. Senthil, *Materials Technology*, 2013, **28**, 9-14.

[4] i) S. Odeh and M. Behnia, *Heat Transfer Eng.*, 2009, **30**, 499-505, ii) A. Royne, C. J. Dey and D. R. Mills, *Sol. Energy Mater. Sol. Cells*, 2005, **86**, 451-483, iii) S. Krauter, *Sol. Energy Mater. Sol. Cells,* 2004, **82**, 131-137.

[5] i) D. Yang and H. Yin, *IEEE Trans. Energy Convers.*, 2011, **26**(2), 662-670; ii) L. E. Chaar, L. A. Lamont and N. E. Zein, *Renew. Sustainable Energy Rev.*, 2011, **15**, 2165-2175.

[6] A. Hagfeldt, G. Boschloo, L. Sun, L. Kloo and H. Pettersson, *Chem. Rev.*, 2010, **110**, 6595-6663.

[7] i) X. -G. Zhu, S. P. Long and D. R. Ort, *Annu. Rev. Plant Biol.*, 2010, **61**, 235-261, ii) V. Tiwari, W. K. Peters and D. M. Jonas, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 1203-1208.

[8] R. Croce and H. van Amerongen, *Nat. Chem. Bio.*, 2014, **10**, 492-501.

[9] B. A. Osborne and R. J. Geider, *New Phytol.*, 1987, **106**, 631-644.

[10] P. Braun, E. Goldberg, C. Negron, M. von Jan, F. Xu, V. Nanda, R. L. Koder and D. Noy, *Proteins*, 2011, **79**, 463-476.

[11] i) A. Dudkowiak, T. Kusumi, C. Nakamura and J. Miyake, *J. Photochem. Photobiol. A*, 1999, **129**, 51-55; ii) A. Dudkowiak, C. Nakamura, T. Arai, J. Miyake, *J. Photochem. Photobiol. B*, 1998, **45**, 43-50.

[12] M. Z. Papiz, V. Cherezov, J. Clogston and M. Caffrey, *J. Mol. Biol.*, 2006, **357**, 1605-1618.

[13] C. Curutchet and B. Mennucci, *Chem. Rev.*, 2017, **117**, 294-343.

[14] A. Krieger-Liszkay, C. Fufezan and A. Trebst, *Photosynthesis Research*, 2008, **98**, 551-564.

[15] A. Telfer, *Phil. Trans. R. Soc. Lond. B*, 2002, **357**, 1431-1440.

[16] T. Mirkovic, E. E. Ostroumov, J. M. Anna, R. van Grondelle and G. D. Scholes, *Chem. Rev.*, 2017, **117**, 249-293.

[17] P. Trouillas, J. C. Sancho-Garcia, V. D. Freitas, J. Gierschner, M. Otyepka and O. Dangles, *Chem. Rev.*, 2016, **116**, 4937-4982.

[18] A. Punnoose, L. A. McConnell, W. Liu, A. C. Mutter and R. L. Koder, *PLoS ONE*, 2012, **7**, E36065.

[19] G. D. Scholes, G. R. Fleming, A. Olaya-Castro and R. van Grondelle, *Nat. Chem.*, 2011, **3**, 763-774.

[20] i) C. N. Fleming, K. A. Maxwell, J. M. DeSimone, T. J. Meyer and J. M. Papanikolas, *J. Am. Chem. Soc.*, 2001, **123**, 10336-10347; ii) D. A. Friesen, T. Kajita, E. Danielson and T. Meyer, *J. Inorg. Chem.*, 1998, **37**, 2756-2762.

[21] i) S. S. Babu, K. K. Kartha and A. Ajayaghosh, *J. Phys. Chem. Lett.*, 2010, **1**, 3413-3424; ii) S. Bhattacharya and S. K. Samanta, *Langmuir*, 2009, **25**, 8378-8381.

[22] i) D. L. Andrews, *J. Mater. Res.*, 2012, **27**, 627-638; ii) G. D. D'Ambruoso and D. V. McGrath, *Adv. Polym. Sci.*, 2008, **214**, 87-147.

[23] J. S. Lin, J. C. Albrecht, R. J. Meagher, X. Wang and A. E. Barron, *Biomacromolecules*, 2011, **12**, 2275-2284.

[24] A. J. Hurd, R. L. Kelly, R. G. Eggert and M. H. Lee, *Mater. Res. Soc. Bull.*, 2012, **37**, 505-510.

[25] K. Binnemans and C. Görller-Waldrand, *Chem. Phys. Lett.*, 1995, **235**, 163-174.

[26] S. I. Weissman, *Chem. Phys.*, 1942, **10**, 214-217.

[27] i) G. A. Crosby, R. E. Whan and R. M. Alire, *J. Chem. Phys.*, 1961, **34**, 743-748; ii) W. R. Dawson, J. L. Kropp and M. W. Windsor, *J. Chem. Phys.*, 1966, **45**, 2410-2418.

[28] S. V. Eliseeva and J. C. G. Bünzli, *Chem. Soc. Rev.*, 2010, **39**, 189-227.

[29] M. Latva, H. Takalo, V. M. Mukkala, C. Matachescu, J. C. Rodriguez-Ubis and J. Kankare, *J. Lumin.*, 1997, **75**, 149-169.

[30] B. Alpha, R. Ballardini, V. Balzani, J. –M. Lehn, S. Perathoner and N. Sabbatini, *Photochem. Photobiol.*, 1990, **52**, 299-306.

[31] i) J. -C. G. Bünzli and C. Piguet, *Chem. Soc. Rev.*, 2005, **34**, 1048-1077; ii) J. -C. G. Bünzli, *Coord. Chem. Rev.*, 2015, **293**, 19-47; iii) J. -C. G. Bünzli, *Eur. J. Inorg. Chem.*, 2017, 5058-5063.

[32] i) J. G. Reifenberger, G. E. Snyder, G. Baym and P. R. Selvin, *J. Phys. Chem. B*, 2003, **107**, 12862-12873; ii) D. Kovacs, X. Lu, L. S. Meszaros, M. Ott, J. Andres and K. E. Borbas, *J. Am. Chem. Soc.*, 2017, **139**, 5756-5767.

[33] A. K. R. Junker, L. R. Hill, A. L. Thompson, S. Faulkner and T. J. Sørensen, *Dalton Trans.*, 2018, **47**, 4794-4803.

[34] P. Atkinson, K. S. Findlay, F. Kielar, R. Pal, D. Parker, R. A. Poole, H. Puschmann, S. L. Richardson, R. A. Stenson, A. L. Thompson and J. Yu, *Org. Biomol. Chem.*, 2006, **4**, 1707-1722.

[35] G. Zucchi, A. –C. Ferrand, R. Scopelliti and J. –C. G. Bünzli, *Inorg. Chem.*, 2002, **41**, 2459-2465.

[36] B. Alpha, J. –M. Lehn and G. Mattis, *Angew. Chem. Int. Ed. Engl.*, 1987, **26**, 266-267.

[37] A. Bettencourt, P. S. Barber, S. Viswanathan, *Coord. Chem. Rev.*, 2014, **273**, 165-200.

[38] B. Song, G. Wang, M. Tan and J. Yuan, *J. Am. Chem. Soc.*, 2006, **128**, 13442-13450.

[39] A. P. S. Samuel, J. Xu and K. N. Raymond, *Inorg. Chem.*, 2009, **48**, 687-698.

[40] i) S. Bhattacharyya, M. B. Roy and S. Ghosh, *Chem. Phys.*, 2004, **300**, 295-304; ii) P. S. Sardar, S. Samanta, M. B. Roy and S. Ghosh, *Mol. Phys.*, 2008, **106**, 827-840.

[41] i) R. E. Dale and J. Eisinger, *Biopolymers*, 1974, **13**, 1573-1605; ii) D. B. VanBeek, M. C. Zwier, J. M. Shorb and B. P. Krueger, *Biophys. J.*, 2007, **92**, 4168-4178.

[42] S. Sato and M. Wada, *Bull. Chem. Soc. Jpn.*, 1970, **43**, 1955-1962.

[43] F. R. Goncalves e Silva, R. Longo, O. L. Malta, C. Piguet and J. -C. G. Bünzli, *Phys. Chem. Chem. Phys.*, 2000, **2**, 5400-5403.

[44] A. Mershin, K. Matsumoto, L. Kaiser, D. Yu, M. Vaughn, M. K. Nazeeruddin, B. D. Bruce, M. Graetzel and S. Zhang, *Sci. Rep.*, 2012, **2**, 1-6.

[45] K. Kjaergaard, J. K. Sorensen, M. A. Schembri and P. Klemm, *Appl. Env. Microbiol.*, 2000, **66**, 10-14.

[46] i) C. J. Faulkner, S. Lees, P. N. Ciesielski, D. E. Cliffel and G. K. Jennings, *Langmuir*, 2008, **24**, 8409-8412. ii) P. N. Ciesielski, A. M. Scott, C. J. Faulkner, B. J. Berron, D. E. Cliffel and G. K. Jennings, *ACS Nano*, 2008, **2**, 2465-2472.

[47] D. Gunther, G. LeBlanc, D. Prasai, J. R. Zhang, D. E. Cliffel, K. I. Bolotin and G. K. Jennings, *Langmuir*, 2013, **29**, 4177-4180.

[48] E. A. Gizzie, J. S. Niezgoda, M. T. Robinson, A. G. Harris, G. K. Jennings, S. J. Rosenthal and D. E. Cliffel, *Energy Environ. Sci.*, 2015, **8**, 3572-3576.

[49] Q. Fu, C. Zhao, S. Yang and J. Wu, *Mater. Lett.*, 2014, **129**, 195-197.

[50] K. Nguyen and B. D. Bruce, *Biochem. Biophys. Acta*, 2014, **1837**, 1553-1566.

[51] i) G. Calogero, A. Bartolotta, G. D. Marco, A. D. Carlo, F. Bonaccorso, *Chem. Soc. Rev.*, 2015, **44**, 3244-3294. ii) H. Hug, M. Bader, P. Mair and T. Glatzel, *Applied Energy*, 2014, **115**, 216-225. iii) A. Hagfeldt, G. Boschloo, L. Sun, L. Kloo and H. Pettersson, *Chem. Rev.*, 2010, **110**, 6596-6663.

[52] A. J. McCormick, P. Bombelli, R. W. Bradley, R. Thorne, T. Wenzel and C. J. Howe, *Energy Environ. Sci.*, 2015, **8**, 1092-1109.

[53] Q. Luo, C. Hou, Y. Bai, R. Wang and J. Liu, *Chem. Rev.*, 2016, **116**, 13571-13632.

[54] S. Shoda, H. Uyama, J. Kadokawa, S. Kimura and S. Kobayashi, *Chem. Rev.*, 2016, **116**, 2307-2413.

[55] J. Stubbe and W. A. van der Donk, *Chem. Rev.*, 1998, **98**, 705-762.

[56] i) R. S. Hodges, *Biochem. Cell Biol.*, 1996, **74**, 133-154; ii) R. de la Rica and H. Matsui, *Chem. Soc. Rev.*, 2010, **39**, 3499-3509.

[57] L. Pauling, R. B. Corey and H. R. Branson, *Proc. Natl. Acad. Sci. USA*, 1951, **37**, 205-211.

[58] i) J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips and V. C. Shore, *Nature*, 1960, **185**, 422-427; ii) J. D. Dunitz, *Angew. Chem. Int. Ed.*, 2001, **40**, 4167-4173.

[59] C. J. Russell, D. S. King, T. E. Thorgeirsson and Y. K. Shin, *Protein Eng.*, 1998, **11**, 539-547.

[60] S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 8898-8902.

[61] C. N. Pace and J. M. Scholtz, *Biophys. J.*, 1998, **75**, 422-427.

[62] F. R. Salemme, *Prog. Biophys. Molec. Biol.*, 1983, **42**, 95-133.

[63] M. H. Hecht, *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 8729-8730.

[64] Y. Yan and B. W. Erickson, *Protein Sci.*, 1994, **3**, 1069-1073.

[65] i) A. Lupas, *TIBS*, 1996, **21**, 375-382; ii) A. Lupas and M. Gruber, *Adv. Protein Chem.*, 2005, **70**, 37-78.

[66] i) S. Zhang, *Nat. Biotech.*, 2003, **21**(10), 1171-1178. ii) P. Burkhard, J. Stetefeld and S. V. Strelkov, *Trends in Cell Biology*, 2001, **11**(2), 82-88. iii) K. Beck and B. Brodsky, *J. Struct. Biol.*, 1998, **122**, 17-19.

[67] D. Woolfson, *Adv. Protein Chem.*, 2005, **70**, 79-112.

[68] i) F. H. C. Crick, *Nature*, 1952, **170**, 882-883, ii) L. Pauling and R. B. Corey, *Nature*, 1953, **171**, 59-61.

[69] I. Hanukoglu and E. Fuchs, *Cell*, 1983, **33**, 915-924.

[70] i) P. B. Harbury, T. Zhang, P. S. Kim and T. Alber, *Science*, 1993, **262**, 1401-1407, ii) P. B. Harbury, P. S. Kim and T. Alber, *Nature*, 1994, **371**, 80-83.

[71] i) C. Cohen and D. A. D. Parry, *Proteins*, 1990, **7**, 1-15. ii) A. J. Gamble and A. F. A. Peacock, in *Protein Design*, ed V. Kohler, Springer Science, New York, 2nd edn., 2014, vol. 1216, pp. 211-231.

[72] N. R. Zaccai, B. Chi, A. R. Thomson, A. L. Boyle, G. J. Bartlett, M. Bruning, N. Linden, R. B. Sessions, P. J. Booth, R. L. Brady and D. N. Woolfson, *Nat. Chem. Biol.*, 2011, **7**, 935-941.

[73] F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum and V. L. Pecoraro, *Chem. Rev.*, 2014, **114**, 3495-3578.

[74] A. J. Doerr and G. L. McLendon, *Inorg. Chem.*, 2004, **43**, 7916-7925.

[75] i) D. Ghosh and V. L. Pecoraro, *Inorg. Chem.*, 2004, **43**, 7902-7915. ii) S. Chakraborty, D. S. Touw, A. F. A. Peacock, J. Stuckey and V. L. Pecoraro, *J. Am. Chem. Soc.*, 2010, **132**, 13240-13250. iii) M. Matzapetakis, D. Ghosh, T. –C. Weng, J. E. Penner-Hahn and V. L. Pecoraro, *J. Biol. Inorg. Chem.*, 2006, **11**, 876-890. vi) M. R. Ghadiri, C. Soares and C. Choi, *J. Am. Chem. Soc.*, 1992, **114**, 825-831. v) M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton and A. F. A. Peacock, *Chem. Sci.*, 2016, **7**, 2207-2216.

[76] i) A. F. A. Peacock, *Curr. Opin. Chem. Biol.*, 2013, **17**, 934-939. ii) M. L. Zastrow and V. L. Pecoraro, *Coord. Chem. Rev.*, 2013, **257**, 2565-2588.

[77] D. J. Wilger, S. E. Bettis, C. K. Materese, M. Minakova, G. A. Papoian, J. M. Papanikolas and M. L. Waters, *Inorg. Chem.*, 2012, **51**, 11324-11338.

[78] M. R. Ghadiri and M. A. Case, *Angew. Chem. Int. Ed. Engl.*, 1993, **32**(11), 1594-1597.

[79] i) Z. Ji, G. Natu and Y. Wu, *Appl. Mater. Interfaces*, 2013, **5**, 8641-8648. ii) F. Odobel and H. Zabri, *Inorg. Chem.*, 2005, **44**, 5600-5611.

[80] M. R. Razeghiford and T. Wydrzynski, *Biochemistry*, 2003, **42**, 1024-1030.

[81] S. Sasaki, T. Mizoguchi, and H. Tamiaki, *Tetrahedron*, 2005, **61**, 8041-8048.

[82] M. Taniguchi and J. S. Lindsey, *Chem. Rev.*, 2016, **117**, 344-535.

[83] T. A. Farid, G. Kodali, L. A. Solomon, B. R. Lichtenstein, M. M. Sheehan, B. A. Fry, C. Bialas, N. M. Ennist, J. A. Siedlecki, Z. Zhao, M. A. Stetz, K. G. Valentine, J. L. R. Anderson, A. J. Wand, B. M. Discher, C. C. Moser and P. L. Dutton, *Nat. Chem. Bio.*, 2013, **9**, 826-836.

[84] J. M. Shifman, B. R. Gibney, R. E. Sharp and P. L. Dutton, *Biochemistry*, 2000, **39**, 14813-14821.

[85] L. A. Solomon, G. Kodali, C. C. Moser and P. L. Dutton, *J. Am. Chem. Soc.*, 2014, **136**, 3192-3199.

[86] I. Cohen-Ofri, M. van Gastel, J. Grzyb, A. Brandis, I. Pinkas, W. Lubitz and D. Noy, *J. Am. Chem. Soc*., 2011, **133**, 9526-9535.

[87] D. V. Zaytsev, F. Xie, M. Mukherjee, A. Bludin, B. Demeler, R. M. Breece, D. L. Tierney and M. Y. Ogawa, *Biomacromolecules*, 2010, **11**(10), 2602-2609.

[88] I. M. M. Carvalho and M. Y. Ogawa, *J. Braz. Chem. Soc.*, 2010, **21**(7), 1390-1394.

[89] B. R. Lichtenstein, C. Bialas, J. F. Cerda, B. A. Fry, P. L. Dutton and C. C. Moser, *Angew. Chem. Int. Ed.*, 2015, **54**, 13626-13629.

[90] A. Roy, C. Madden and G. Ghirlanda, *Chem. Commun.*, 2012, **48**, 9816-9818.

[91] M. Faiella, A. Roy, D. Sommer and G. Ghirlanda, *Peptide Science*, 2013, **100**(6), 558-571.

Chapter 2

STRUCTURAL DESIGN AND CHARACTERISATION OF NOVEL COILED COIL SCAFFOLDS

2.1 INTRODUCTION

The structural potential of peptides and more specifically coiled coils has been explored in Chapter 1. The use of coiled coils as ligands has found a particular advantage over conventional small molecule species, as they can offer a greater steric and environmental control, in much the same way as enzymes. This has led to a great deal of research into metal ion binding using coiled coil structures to create functional systems. Nevertheless, while at present there are no commercial uses for coiled coil technologies, they have found extensive use as model systems to study biological enzyme sites and binding domains. Several systems have been studied, such as anhydrases, hydrogenases and copper enzymes (type 1 and 2) as enzymatic mimics.[1,2] A notable example has been the construction of a coiled coil, zinc binding, carbonic anhydrase mimic which was utilised to study the turn-over mechanism.[1] Indeed, coiled coils have proven to be excellent model systems for biological study in both metallated and non-metallated states.[3] The use of metal ions within coiled coil systems has not just been limited to natural mimics but has made use of biologically unavailable metals such as rhodium,[4] ruthenium,[5] osmium,[5] uranium,[6] mercury and lead to name a few.[7] The field of metallated coiled coils has been extensively reviewed by F. Yu *et al.* and M. Zastrow *et al.* and therefore will not be considered in detail here.[8]

Natural systems make use of bound cofactors such as chromophores and metal ions to add varied functionality and photophysical properties, allowing manipulation of photons within a constrained system. This can also be achieved by using metal ions not traditionally available to nature, such as the lanthanides. The appeal of these metal ions is primarily two-fold. Firstly, the varied and well characterised photophysics of the lanthanides offers a range of potential

functionality. This can be both for useful systems employing redox chemistry and energy transfer as well as luminescence and lifetime analysis, which are frequently used in assays and diagnostics.[9] Secondly, the similar coordination chemistry across the lanthanides allows the same binding site design to be used to coordinate much of the series. This will reduce the design requirement while enabling the construction of a wide and varied range of systems by exploiting a 'plug-and-play' type of approach. There are a few coiled coil systems that have been used to bind lanthanides and they merit further consideration.

The first reported instance of lanthanides bound to coiled coils was an initial study on the interaction of lanthanide trications with an acid rich coiled coil domain.[10] The results reinforced the idea that lanthanides preferentially bound to acidic sidechains such as glutamates. This was further advanced by the introduction of acidic non-natural amino acids to achieve excellent binding, however, this approach is expensive, complex and tends to limit commercial application.[11] Work within the Peacock group has been to develop a coiled coil trimer that exhibited a novel lanthanide binding site constructed from natural amino acids, MB1-2 (Fig. 1).[12] This site made use of three aspartate residues to provide a 3- charge for electrostatic ionic binding to the tricationic lanthanide ions and three asparagine residues to provide a complete 9-coordinate hard oxygen environment. It is of particular interest for a new peptide design as it is relatively simple to include in a trimeric coiled coil and can be utilised to bind across the lanthanide series.[13] Unfortunately, the transfer efficiencies and photophysical behaviour of the MB series is particularly weak, with energy transfer between tryptophan and Tb(III) ions being recorded as $0.08 \pm 0.07\%$,[13] therefore this system alone would not be a useful light harvest module. Attempts to crystallise the MB series have proven to be unsuccessful, therefore it would be advantageous if a new design could readily crystallise in order to obtain the solid

state structure of the intact binding site. This leads to the consideration of another literature system of interest, CoilSer (CS),[14] which was successfully crystallised by B. Lovejoy *et al.* and although the solid state structure was found to be antiparallel, both solution state studies and further work have illustrated that it is parallel. Thus this sequence has been used prolifically as a crystallisation motif, enabling the study of zinc, lead, mercury and arsenic binding within CS derivative systems.[15] Due to the proven robustness of the CS system and it's well characterised structure it would be logical to base a new design on the structure of this peptide. Indeed, recent work by L. Ruckthong *et al.*, demonstrated that adding a heptad to the original CS peptide resulting in a long derivative, CS+, was both crystallisable and also parallel (*see* Table 1).[15]

Table 1: The sequences and associated charges on a single strand of the literature CS and MB1-2 species, the designed HC series and the model CS+ sequence. The charges are calculated for the peptides in aqueous solution at pH 7.0. The tryptophan residues are highlighted in red and the binding site residues are shown as bold and underlined.

Peptide	Sequence	Charge
CS	Ac-E **W**EALEKK LAALESK LQALEKK LEALEHG-NH_2	***-2***
MB1-2	Ac-G IAAIEQK IAA**N**E**W**K **D**AAIEQK IAAIEQK IAAIEQK G-NH_2	***-1***
CS+	Ac-E **W**EALEKK LAALESK LQALEKK LQALEKK LEALEHG-NH_2	***-1***
HC0C	Ac-E **W**EAIEKK IAAIESK IQAIEKK IQAIEKK IEAIEHG-NH_2	***-1***
HC01	Ac-E **W**EA**N**EKK **D**AAIESK IQAIEKK IQAIEKK IEAIEHG-NH_2	***-2***
HC02	Ac-E **W**EAIEKK IAA**N**ESK **D**QAIEKK IQAIEKK IEAIEHG-NH_2	***-2***
HC03	Ac-E **W**EAIEKK IAAIESK IQA**N**EKK **D**QAIEKK IEAIEHG-NH_2	***-2***
HC04	Ac-E **W**EAIEKK IAAIESK IQAIEKK IQA**N**EKK **D**EAIEHG-NH_2	***-2***

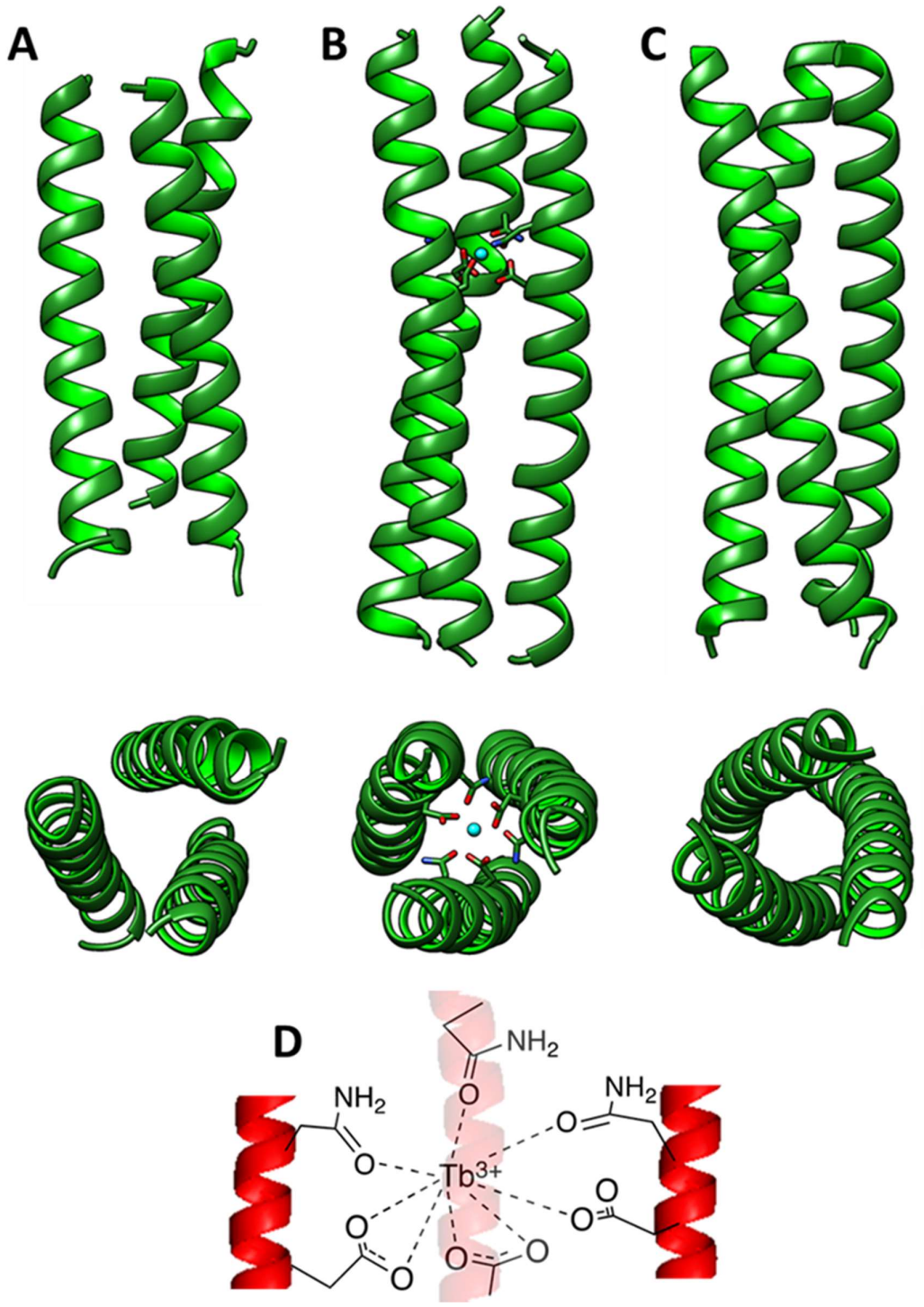

Figure 1: The side on and top down view of the solid state structures of literature peptides A) CoilSer generated from the pdb 1COS,6 B) MB1-2 cartoon generated from computational model,4 and C) CS+ cartoon generated from the long CoilSer derivative pdb 5KB2. The lanthanide binding site of MB1-2 is shown with both sticks and spheres (terbium is cyan, oxygen is red, nitrogen is blue and carbon green). D) A schematic of the binding site design with three anions from the carboxylic acid residues interacting with the cationic terbium ion.

2.2 PEPTIDE DESIGN

Having determined that the MB series and CS peptide make a natural pairing for a new design it remained to combine the desirable features from each system. As previously discussed it was beneficial to retain the designed lanthanide binding site of the MB series, thus requiring the design to be three α-helices forming a parallel coiled coil in a trimeric fashion. The binding site was retained with aspartate and asparagine residues to retain the same coordination environment as previously reported, namely three negative charges for ionic bonding to the tricationic lanthanide species and a 9-coordinate oxygen binding environment. The trimer structural restriction is to maintain the coordination sphere presented by the two residues on each of the three α-helices. The core binding site should also exclude water molecules to limit energy losses from vibrational modes.

In addition to the metal binding and trimeric features it would be ideal that the peptide should be very water soluble from pH 6-8, this is to balance the stability of the MB binding site that would be compromised by protonation and to avoid lanthanide hydroxide formation that occurs at higher pH values.[16] The system should also retain a chromophoric handle to enable rapid and simple concentration determination. As previously mentioned it was particularly useful to analyse the structures of the scaffolds by crystallographic techniques. A method to improve the chances of crystallisation is to functionalise the surface of the designed coiled coil to induce interactions or metal ion coordination between the superstructures. This can be achieved with sidechains such as hard oxygen donors from glutamates or nitrogen donors from histidine residues, which have been used to induce crystal matrix formation/nucleation and are found as key external features on the CS system.

Having discussed the required parameters for a simple structure, it remains to design the linear sequence to give the corresponding super-structure. The principles behind coiled coil heptad repeats has already been discussed at length (*see* Chapter 1.5), this heptad repeat regime will be used to build and understand the new sequences. By combining the designs of the MB series and the CS peptide, it resulted in the replacement of the core leucine residues from CS with isoleucine residues at the ***a*** and ***d*** positions, as in the MB series. This provides the hydrophobic core of the coiled coil system which would still be trimer directing due to the steric bulk of the isoleucine sidechains.[17] The CS sequence was further utilised to provide the residues of the ***b*** and ***c*** position which are typically utilised for helix inducing and water solubilising functions, therefore alanine and glutamine residues are the obvious choices.[18] The ***e*** and ***g*** positions are commonly used for salt bridging residues to stabilise interhelical interactions between the α-helices in a coiled coil and which typically consists of a lysine and glutamate pairing. Indeed this motif was present in both the reported peptides and was retained in the design. This left only the external ***f*** positions to assign and a few deviations from the normal heptad repeat. The CS peptide had a highly functionalised surface with the ***f*** positions containing polar residues such as histidine and lysine, these were thought to be important for metal coordination to the surface in order to crystallise the peptide, so were also retained. A histidine residue was placed near the C-terminus to coordinate zinc metal ions to facilitate crystal formation which was thought to be important for the original crystallographic success of CS.[14] A tryptophan residue was retained close to the N-terminus, providing the chromophoric tag thus completing the design of the control HC0C peptide. The lanthanide binding site of MB1-2 was incorporated by placing an asparagine and aspartate in the ***d*** position of one heptad and the ***a*** position of the subsequent heptad respectively, this site could be translated up and down the sequence.[19] Contrary to the MB series, the tryptophan was not moved with the binding site. It should be

noted that all the peptides discussed herein have the N-terminus acetylated and the C-terminus amidated. This limits repulsion between peptide helices at either end when they form a parallel coiled coil due to the lack of like charges. These features resulted in the hybrid designs collectively known as the HC series (*see* Table 1).

The overall charge on the peptides at pH 7 was determined from the number of glutamate, aspartate and lysine residues in the sequences (*see* Table 1). The resultant overall charge of -2 for the HC series indicates that the native peptides would be very soluble in water and have a high affinity for positively charged species such as lanthanide trications. Upon binding a trication, the designed trimer unit would still have an overall charge of -3 (-1 from each α-helix) therefore would be expected to be very soluble. The control HC0C peptide had an overall charge of -1 which implies a trimer species charge of -3, this suggests that the control will also be soluble in water. The distribution of lysine and glutamate residues is relatively even through-out the sequence, indicating that the overall charge on the peptide is not particularly localised, therefore the whole system should be soluble without highly polar and apolar domains.

2.3 PREDICTIVE ANALYSIS OF THE PEPTIDE DESIGNS

In order to predict the behaviour of the newly designed sequences, the coiling propensity of each sequence was analysed using a Markov model-based programme called MARCOIL.[20] This programme predicts the helical coiled coil structure from protein primary sequences. Initially, the control peptide was compared with CS, MB1-2 and the extended CS derivative

(CS+) to ascertain the relative coiled propensity of the new design before including a binding site. The results clearly indicated that the three peptides; MB1-2, CS and HCOC would be well folded with over 70% of the peptide structure expected to be well formed helices (Fig. 2A). X-ray diffraction and circular dichroism data have determined that CS and MB1-2 respectively are α-helical, therefore it seemed a valid assumption that HC0C will also be helical.

Having determined that the control system had a high coiling propensity, a rough prediction of the disruptive influence of the hydrophilic residues in the core at the various binding site

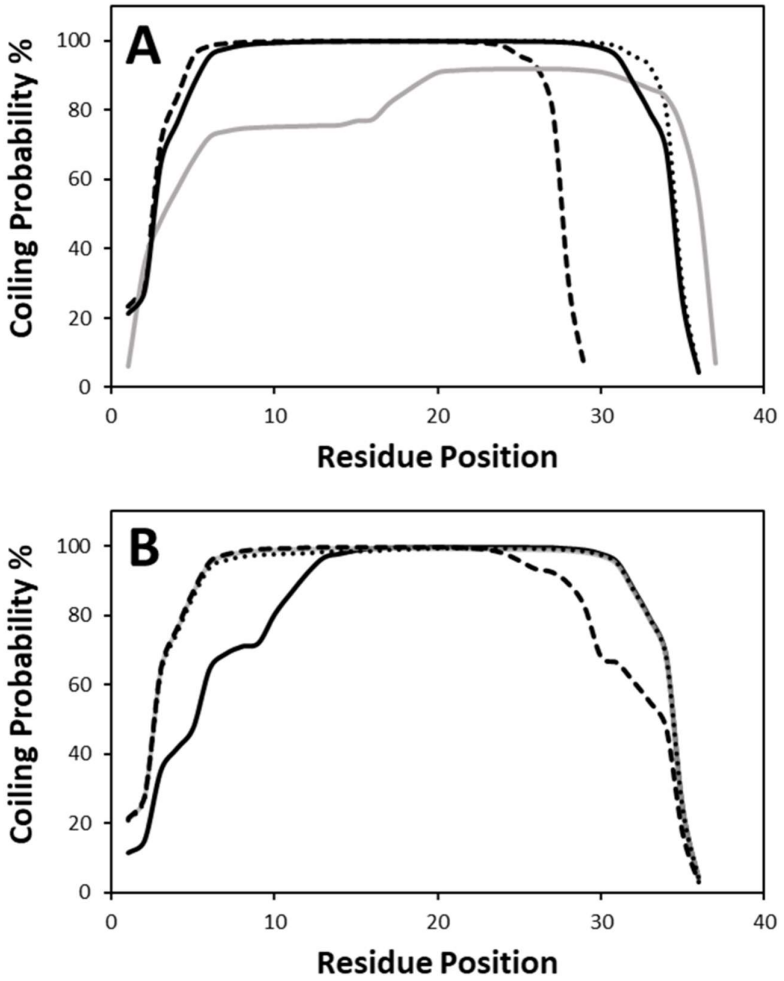

Figure 2: The MARCOIL profiles of A) CS (dashed black), CS+ (dotted black), MB1-2 (solid grey), HC0C (solid black) and B) HC01 (solid black), HC02 (dotted black), HC03 (solid grey) and HC04 (dashed black) indicating the helical coiling probability by residue position.

positions was studied (Fig 2B). Both HC01 and HC04 are predicted to have less well formed coiled coil structures for the respective termini at which the binding site is placed due to termini being generally more flexible. However, supercoiling of the other heptads is thought to stabilise this disruption. The HC02 and HC03 systems are predicted to retain helical structure despite the central placement of the binding sites. This seems unlikely as it is well characterised for the MB series that the centrally placed sites result in poorly folded native peptides due to the hydrophilic residues disrupting core interactions.[19] Therefore, we would expect to see a similar behaviour for the HC series. This would imply that HC0C, HC01 and HC04 would be well structured helical systems with some flexibility at the respective binding site termini, whilst the HC02 and HC03 are unlikely to be well folded due to the centrally placed binding sites.

Further to the helical propensity it was necessary to determine the probability of these helices forming trimeric coiled coils, therefore the programme LOGICOIL was used to predict the oligomeric states of the peptides in Table 2.[21] Each of the studied peptides favoured a trimeric structure with exceptionally high propensities, as demonstrated by the low values for both dimeric and tetrameric oligomeric states (*see* Table 2). The LOGICOIL output represents a probability of oligomeric state formation, values that are within 0.3 of each other indicate that multiple species are likely to be present. For example, MB1-2 has the trimer as the most likely species but a parallel dimer is also relatively high probability suggesting that this might form a mixed system. Pleasingly, the values for all the HC peptides indicated a very clear favouring of trimeric structures. This was not entirely unexpected as the ***a*** and ***d*** positioned isoleucine residues are known to direct the supramolecular structure towards trimers.[17] The valid prediction for the literature systems gave further confidence that the new designs would also form coiled coil trimers.

Table 2: The LOGICOIL output for the oligomeric domain propensity of the literature peptides CS and MB1-2, the designed HC series and the model CS+ system. Higher values indicate higher probability of formation.

Peptide	Antiparallel Dimer	Parallel Dimer	Trimer	Tetramer
CS	0.93	1.03	1.58	0.85
MB1-2	0.99	1.00	1.23	0.79
CS+	0.93	1.04	1.52	0.86
HC0C	1.00	0.77	1.88	0.74
HC01	1.00	0.81	1.74	0.76
HC02	1.00	0.81	1.74	0.76
HC03	1.00	0.81	1.74	0.76
HC04	1.00	0.81	1.74	0.76

2.4 PEPTIDE SYNTHESIS AND CIRCULAR DICHROISM ANALYSIS

The HC peptide sequences were prepared via a solid-phase peptide synthesis methodology and purified by reverse-phase high pressure liquid chromatography. The mass and purity of each of the peptides was determined by matrix-assisted laser desorption/ionisation mass spectrometry and analytical high pressure liquid chromatography respectively (*see* Chapter 8.4 and S2-S6).

The secondary structure, and sometimes more advanced structure, can be determined by monitoring the circular dichroism profile of a sample. This can be illustrated by observing the resultant spectrum from the control HC0C peptide system (Fig. 3). The molar ellipticity minimum at 222 nm (θ_{222}) and 208 nm (θ_{208}) are clear features pertaining to α-helical domains.[22] Furthermore, it has begun to be accepted that if the ratio $\theta_{222}/\theta_{208}$ is greater than one

it indicates the formation of a super-structure such as a coiled coil.[23] This is because the interaction of the helices resulting in n-π* transitions, implying that a ratio greater than one would indicate supramolecular coiled coil formation leading to orbital overlap between strands. The percentage of the peptide sequence that is helical can be determined from the concentration of the peptide solution and the molar ellipticity at 222 nm (θ_{222}).[22] The non-metallated HC0C was found to be 93% ± 1% helical suggesting that the majority of the system is well structured and that ~7% is probably disordered due to flexibility at the termini of the coil. This result gives an upper limit to the helical content that would be expected to be present in the other metal binding coiled coils of the HC series.

Considering the non-metallated profiles for the peptides HC01-HC04 the effect of the binding site becomes apparent (Fig. 4). For both HC02 and HC03 the circular dichroism profiles exhibit little secondary structure with only 15% ± 1% and 16% ± 1% in α-helical domains respectively, though it should be noted that at these poorly folded states the calculations do not represent any actual helical structure. This indicates that these peptides are not well structured or preformed for binding a metal ion. This is due to the centrally placed binding site consisting of

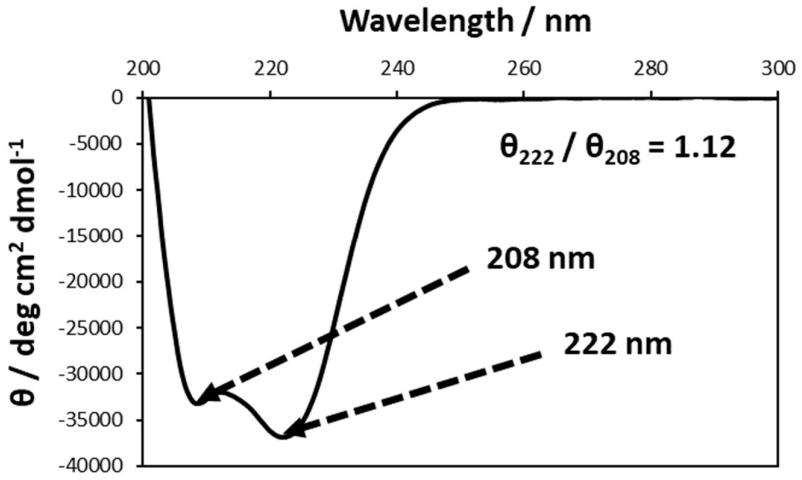

Figure 3: The circular dichroism spectrum of HC0C (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0) with the minima at 208 and 222 nm. The $\theta_{222}/\theta_{208}$ ratio being greater than one indicates a coiled coil system.

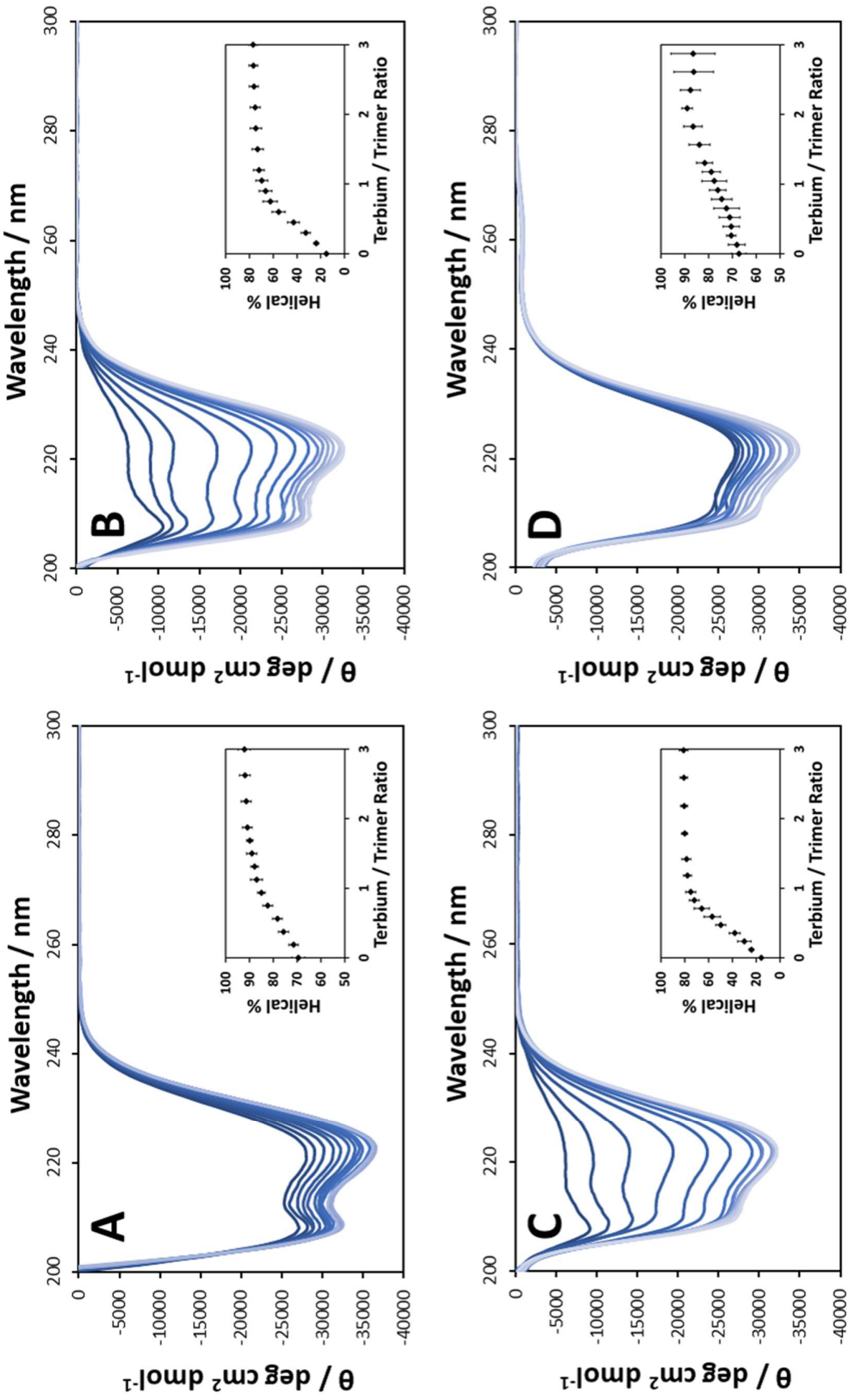

Figure 4: The circular dichroism profiles of A) HC01, B) HC02, C) HC03, D) HC04 peptide systems upon the addition of Tb(III) ions. The inserts show the change in helical content upon the addition of Tb(III) ions as a ratio of metal ion to peptide trimer. The samples were peptide (90 μM) in HEPES buffer (10 mM, pH 7.0) and errors reported are the standard deviation of three repeats.

the hydrophilic asparagine and aspartate residues being buried in the hydrophobic core, resulting in disrupted core heptads which are more stabilising than terminal heptads.[24] Interestingly, both HC01 and HC04 exhibited similar profiles to the control peptide with remarkably well folded systems at 70% ± 3% and 67% ± 3% respectively. This indicated that they were well structured α-helices in solution and the ratio of $\theta_{222}/\theta_{208}$ being greater than one also suggested that both HC01 and HC04 formed coiled coils, perhaps with preformed binding sites.

After characterising the non-metallated peptides, the binding of terbium was studied by titrating terbium trichloride solutions into samples of HC01-HC04 and recording the change in the circular dichroism spectra. In all four cases the α-helical content increased upon the addition of Tb(III) ions, as indicated by a decrease in the minima at 208 and 222 nm, suggesting that metal binding was inducing a structural change in the system. This change was most clearly manifested for both HC02 and HC03 due to the previously destabilised non-metallated systems (Fig. 4B, C). The initial helical content which was particularly low at 15% ± 1% and 16% ± 1% for HC02 and HC03 respectively, upon the addition of one equivalence of Tb(III) ions increased dramatically to 70% ± 5% and 75% ± 4% respectively, indicating that a coiled coil trimer was binding a single metal ion with the metal ion acting to template the peptide structure. Using the molar ellipticity at the maximum folded state and the initial molar ellipticity, the fraction of trimer bound (f) was calculated (*see* Chapter 8.9). This allowed for the determination of the log of the binding constants ($\log K_a$) for both HC02 and HC03 which were found to be 5.13 ± 0.11 and 5.17 ± 0.10 respectively, when fitted to a monomer-to-trimer model (Fig. 5). These binding constants are very similar and within the error of the standard deviation of three repeats indicating that the sites of HC02 and HC03 behave in the same manner, as

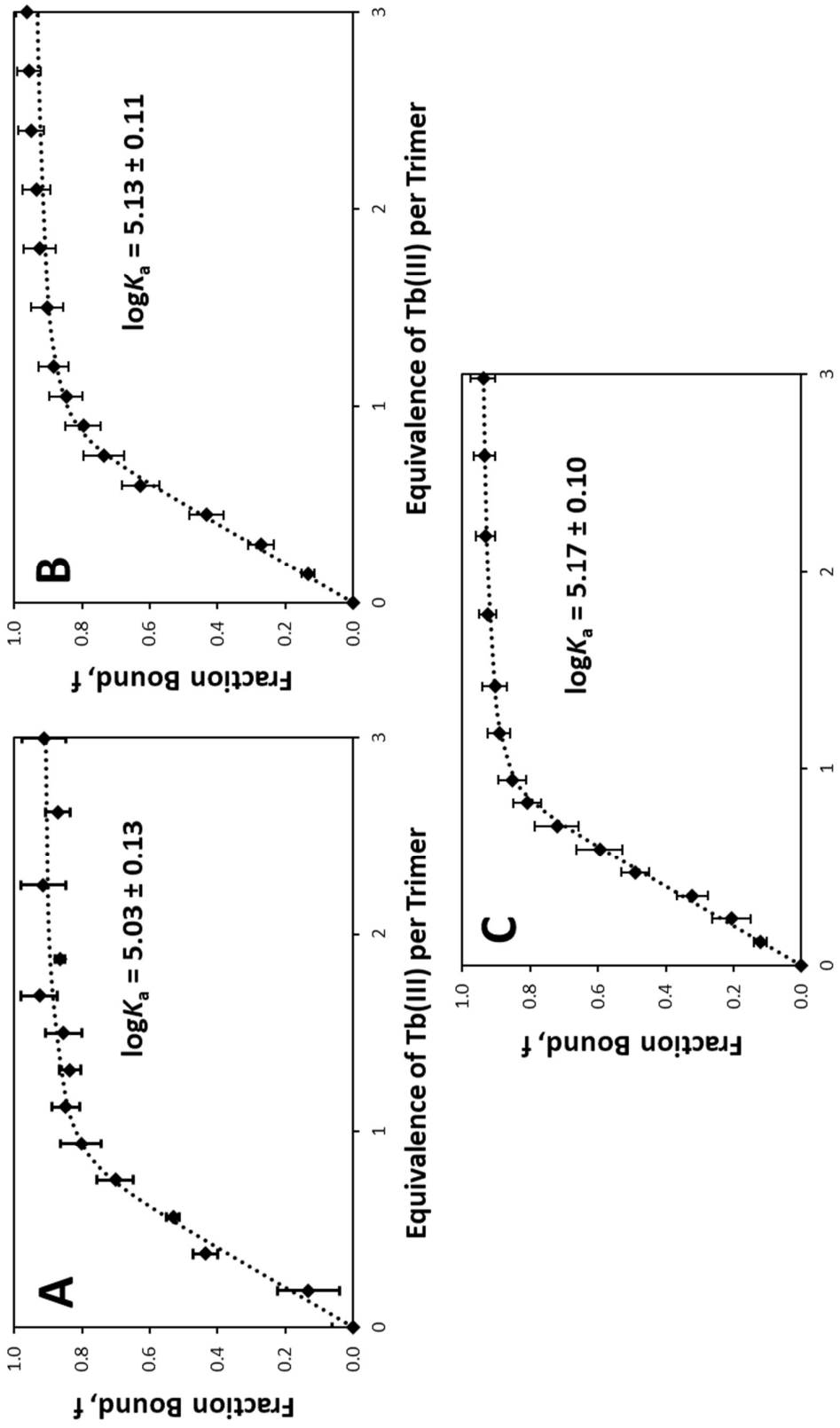

Figure 5: The binding profiles of A) HC01, B) HC02 and C) HC03 peptide systems upon the addition of Tb(III) ions. The binding constant fits (dotted lines) have been calculated from a three-to-one binding equation. The samples were peptide (90 μM) in HEPES buffer (10 mM, pH 7.0) and errors reported are the standard deviation of three repeats.

would be expected. These binding constants result in binding energies ($\Delta G^{\ominus}_{bind}$) of -28.8 ± 0.6 kJ mol^{-1} and -29.0 ± 0.6 kJ mol^{-1} respectively (*see* Table 3) which is consistent with weak binding, however, the values are the same magnitude as that previously observed for MB1-2.[19] Interestingly, it was also possible to observe the enhancement in coiled coil assembly by monitoring the ratio $\theta_{222}/\theta_{208}$ upon the addition of Tb(III). The overall coiled coil structure is enhanced up to one equivalence of Tb(III) to trimer and then plateaus (*see* S13-S14), indicating that only one Tb(III) ion per trimer induces coiled coil templating. This is particularly exciting as it demonstrates that the designed peptides are behaving as intended.

The HC01 peptide system exhibits extensive α-helical structure before the addition of metal ions, which, as previously discussed indicates that the species is a templated system. Due to the well-structured non-metallated state the resulting change in folding upon titration with Tb(III) is quite small, with only ~20% change, but noticeably demonstrating that 1 equivalence of metal ions bind to a trimer unit (Fig. 4A). Indeed, monitoring the $\theta_{222}/\theta_{208}$ ratio it becomes apparent that the first equivalence of metal ions also results in an enhanced coiled coil structure, suggesting that the already templated structure improves upon the addition of metal. This is in agreement with the hypothesis that upon effective metal binding at the N-terminus of the coiled coil, folding is enhanced resulting in the metal stabilising further terminal coiling. The enhancement in coiling was very reproducible with little error compared to the change in α-helical content (*see* S13). Therefore, the $\theta_{222}/\theta_{208}$ ratio was used to monitor binding for HC01 and the resultant profile was fitted, as before (*see* Chapter 8.9), to a three-to-one thermodynamic binding event (Fig. 5A). The determined log of the binding constant was found to be 5.03 ± 0.13 and a binding energy of -28.2 ± 0.7 kJ mol^{-1}. This is remarkably similar to the values ascertained for both HC02 and HC03 (*see* Table 3) suggesting that while the site

modulation can alter the behaviour of the non-metallated peptides, the overall binding affinity is relatively unchanged. Considering that these systems have dramatically different non-metallated states, yet they result in similar binding energies, this would suggest that the predominant driving force is enthalpic. This is particularly surprising as the binding site is composed primarily of hard oxygen donors, not too dissimilar from the aqueous solvent environment. However, the ionic interaction between the three aspartate -1 charges and the terbium trication would result in a similar binding energy for all three systems.

The binding constant values for HC01, HC02 and HC03 are all remarkably similar in the range of 10^5. This has certain implications for the speciation of samples, for example when the peptide concentration is 90 μM and the Tb(III) concentration is 30 μM, it would result in 25 μM concentration of trimer complex or 83% of the Tb(III) is involved in complexation and 17% free in solution.

Table 3: The log of the binding constants and the binding energies for HC01, HC02 and HC03 with Tb(III). The errors represent the standard deviation of three repeats.

Peptide	Binding Constant ($\log K_a$)	Binding Energy / kJ mol^{-1}
HC01	5.03 ± 0.13	-28.2 ± 0.7
HC02	5.13 ± 0.11	-28.8 ± 0.6
HC03	5.17 ± 0.10	-29.0 ± 0.6

The HC04 system starts as well folded α-helices in a coiled coil environment as evidenced by the minima at 222 nm and 208 nm. This implied that it would behave in a similar manner to that observed for HC01. However, there were no obvious changes upon the addition of metal ions, beyond a steady increase in α-helical content. Despite this, it could be considered that the

system reached a plateau at two equivalents of Tb(III) to a trimer unit, but the large error associated with the final points makes this unclear. The extent of coiling did not increase, such that the $\theta_{222}/\theta_{208}$ ratio remained relatively unchanged through-out the titration (*see* S14). This could be due to one of two reasons, the first being that the peptide system is well templated and that the binding site does not disrupt the coiled coil, therefore upon binding little change is observed. The second possibility is that the designed core binding site is not successfully formed perhaps due to a more stable arrangement of each helix without requiring the hydrophilic sidechains to be in close proximity. The circular dichroism profiles of the HC series both before and after the addition of Tb(III) indicates that all except HC04 were behaving as designed. The two discussed possibilities to explain the HC04 results can be differentiated by observing the complexed species by mass spectrometry. If the HC04 species is observed with Tb(III) bound, this suggests that the binding site is intact and thus implies the first hypothesis.

2.5 MASS SPECTROMETRY OF COILED COIL LANTHANIDE COMPLEXES

The use of mass spectrometry to observe supramolecular complexes has been studied extensively, however, complexes that are stabilised by only van der Waals and salt bridging interactions have proved challenging to observe.[25] This is due to the energy requirement for ionisation typically exceeding the stabilisation of the complex, resulting in fragmentation, creating an interplay between size and ionisation energy. While larger systems are generally harder to ionise they have more non-covalent interactions, therefore, they can be ionised with greater energy without fragmentation but small non-covalently interacting coiled coil complexes proved challenging to observe due to the ease with which they fragmented upon

ionisation. There has been work carried out to observe leucine zipper domains and oligomeric states by electrospray mass spectrometry.[26] This was particularly encouraging as it implied that the HC series could also be observed by mass spectrometry. Indeed, MB1-2 was observed with Tb(III) and Gd(III) ions bound by mass spectrometry, however, the non-metallated system has not be observed.[13] Considering the circular dichroism data which suggested complex formation in the presence of Tb(III), it should be possible to observe these coordinated supramolecular species, but perhaps more crucially the non-metallated complexes of HC0C, HC01 and HC04. It was necessary to remove non-volatile salts from the samples to aid ionisation, therefore the HC peptides were prepared in an ammonium acetate buffer. The metallated samples were generated by the addition of 1.3 equivalents of Tb(III) before being analysed. However, due to concentration limitations of the technique the sample would ideally have a concentration of complex below 10 μM, in order to prevent saturation of the detector. For this reason the peptide concentration was 9 μM implying that the maximum concentration of trimer complex could be 3 μM. Unfortunately the binding constants previously discussed suggest that even with 1.3 equivalents of Tb(III) only 50% of the peptide will be binding the metal ions. However, this would be sufficient to determine the nature of the complex by mass spectrometry. A crude attempt at ionisation using matrix-assisted laser desorption/ionisation with a sinapinic acid matrix resulted in observing only monomer species. This was probably because the trimer was too delicate and upon ionisation broke up into monomer units. However, the strategy utilising electrospray ionisation with the ammonium acetate buffer was then attempted with success.

Utilising the control peptide, HC0C, as an example a 5+ charged trimer species was clearly observable with an overall mass of 12465.5 amu (Fig. 6A). This indicated a complex structure of three monomer units of 4155.2 amu, the mass of HC0C. The odd charged species was

particularly easy to identify as there were no overlapping clusters present. This is not so trivial for a 3+ trimer where a monomer of 1+ charge would disguise the trimer species. The HC0C spectrum also showed a 3+ dimeric complex with a monoisotopic species mass of 8310.6 amu. However, rather than being a structure in solution this may have been the result of the trimeric species losing an α-helix under the ionising conditions. To determine if this was the case a tandem mass spectrometry experiment was carried out, selecting for the 5+ trimer species and maintaining the same ionisation energy as the previous experiment (Fig. 6B). This demonstrated that the trimer degraded to give a monomer and dimer species. The ion count ratio for the monomer and dimer species was found to be one, confirming their source as the trimer parent species. Nevertheless, the most important comparison was that of the ratio of trimer/dimer ion in both the previous and tandem mass spectrometry experiment. These values were 0.60 and 0.53 respectively, both very comparable, suggesting that fragmentation of the trimer was the primary source of the observed dimer.

The mass spectra for both the non-metallated and metallated HC01 systems were very similar to that of the HC0C species. There were prominent ions consistent with trimeric species for both samples (Fig. 7) with the respective increase in mass for the metallated sample indicating a single Tb(III) ion bound. The ion count of the HC01 trimer species for both the non-metallated and metallated samples was found to be very similar to that observed for HC0C, when using the same ionisation energy. This suggested that the HC01 peptide is a well-formed trimer, even before Tb(III) binding. A tandem mass spectrometry experiment was also run for the metallated HC01 that demonstrated the same observations as for HC0C (*see* S15-S17); interestingly the metal ion was seen to remain with the dimer species with fragmentation yielding the metallated dimer and the non-metallated monomer ions.

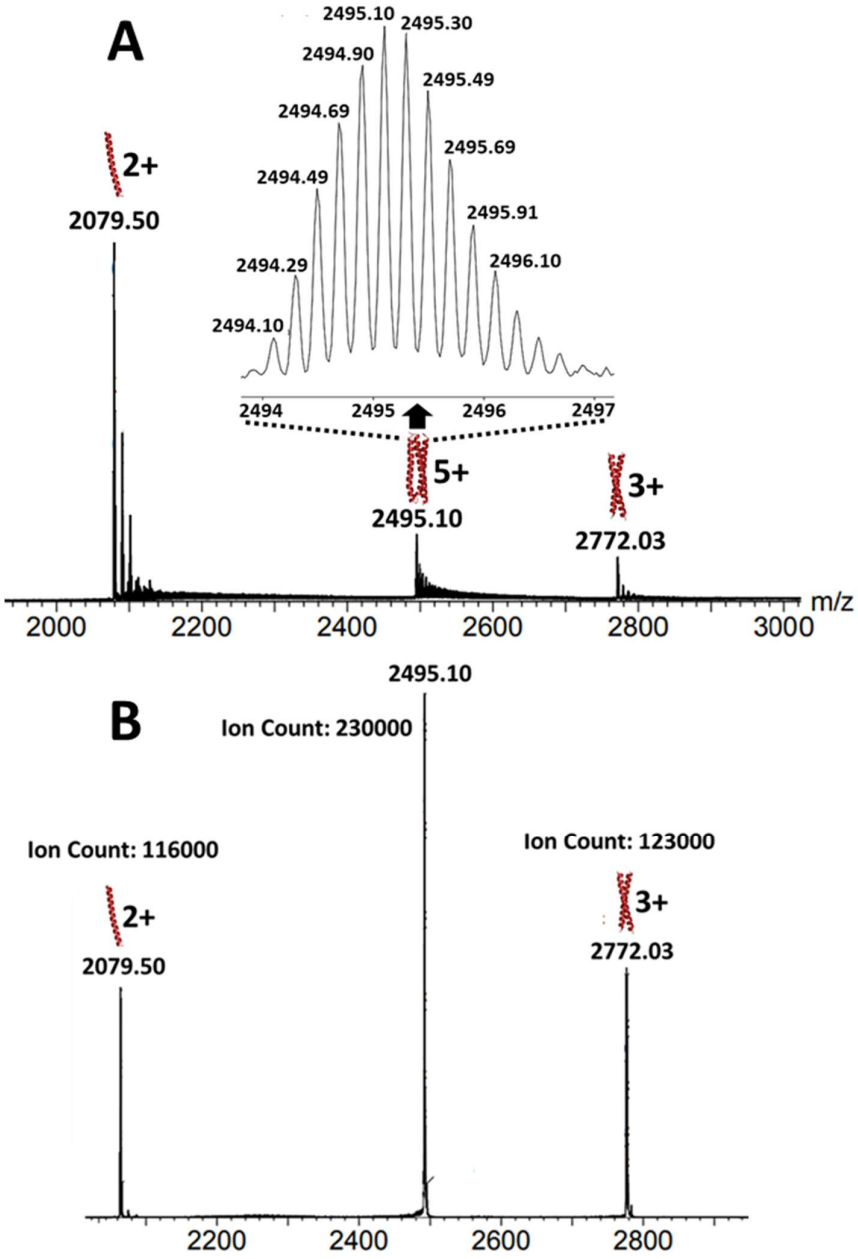

Figure 6: A) The electrospray ionisation mass spectrum of HC0C with the insert showing greater detail for the 5+ trimer isotopic peaks. B) The tandem mass spectrum for the locked ion at 2494 m/z indicating the resultant dimer and monomer result from the fragmentation of the trimer. The sample was HC0C (9 μM) in ammonium acetate buffer (5 mM, pH 6.8).

When considering the spectra for HC02 and HC03, as with the circular dichroism data they behave in a very similar way. While the trimer is visible in the native state for both peptides, the ion counts were very low indicating that this species is either weakly interacting (ionisation energy is close to dissociation energy) or the species is not present at a particularly high concen-

Table 4: The monoisotopic m/z values for the recorded spectra of the HC series with the resultant trimer and monomer masses.

Peptide	Monoisotopic (m/z)	Charge (z)	Monoisotopic Trimer Mass[a]	Monoisotopic Monomer Mass
HC0C	2494.10	5+	12465.5	4155.2
HC01	2527.22	5+	12631.1	4157.4
HC02	2527.20	5+	12631.0	4157.4
HC03	2527.29	5+	12631.5	4157.5
HC04	2496.08	5+	12475.4	4158.5

[a]*The systems that have Tb(III) bound have an increase of 159 amu relative to the trimer mass.*

-tration. Thus, upon the addition of Tb(III) the intensity of the trimer increases dramatically, in agreement with the formation of either a more stable species or present at higher concentrations. This confirms that the relatively unstructured species observed by circular dichroism becomes a well-folded coiled coil upon the addition of Tb(III). The tandem mass spectrometry experiment for the metallated species confirmed a metallated dimer and non-metallated monomer both present in equal measure resulting from the parent trimer species.

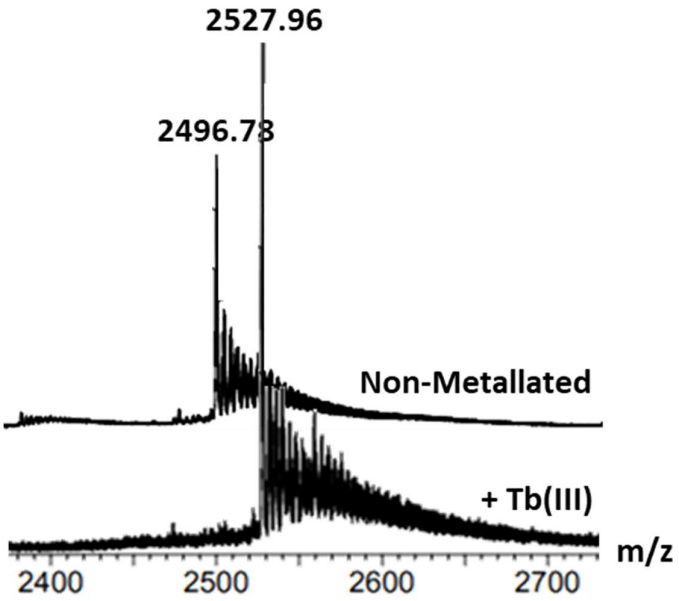

Figure 7: The electrospray mass spectra for the HC01 trimer both with and without 1.3 equivalences of Tb(III), indicating the increase in trimer mass upon the metal binding event. The sample was HC01 (9 μM) in ammonium acetate buffer (5 mM, pH 6.8).

Interestingly, the HC04 peptide was observed as a non-metallated complex, just as had been the case for HC01. Surprisingly, it was not observed with metal present suggesting that this peptide is very stable as a trimer without metal indicating a potential lack of a binding site. Together with the circular dichroism profiles observed for HC04 it supports a lack of metal binding, thus it is likely that the designed parallel coiled coil with a binding site is not forming and that another trimeric assembly is forming instead. One option is that an antiparallel arrangement could be present as this would result in high α-helical content but the core coordination sphere would be disrupted and therefore not functioning. The mass spectra of the complexes clearly indicated that HC0C, HC01-HC03 all behaved as intended and indeed the observation of the trimers by mass gave clear evidence for the adopted oligomeric states of the designed sequences, in support of the LOGICOIL predictions and the original designs. However, the aberrant behaviour of HC04 merits further investigation.

2.6 CRYSTAL GROWTH

Primary screening for the crystallisation conditions for HC02 was carried out by H. Campbell in unpublished work utilising the JCSG+ crystal screen kits and it was determined that a specific set of conditions seemed to generate reasonable hits for HC02.[27] The conditions involved PEG3350, ammonium chloride and zinc acetate, of notable importance was the zinc which was thought to be particularly important in the initial crystal formation steps to the external histidine residues in the designed sequences.[28] Indeed, zinc coordination is a well-established motif for both function and structure in natural proteins.[29] The initial screening conditions were further optimised to allow repeated growth of crystals for HC02 in the presence

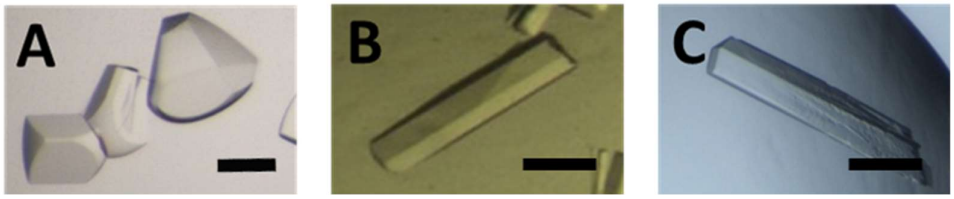

Figure 8: Photographs of crystals of A) HC02 as rounded triangular prisms, B) HC03 as long triangular prisms and C) HC0C as long triangular prisms. These crystals were grown from a combined solution of peptide (16 mg mL^{-1} peptide, 6 mM terbium chloride and 5 mM zinc acetate) and precipitant (0.21 M ammonium chloride and 25% v/v poly(ethylene) glycol 3350) in a 1:1 ratio.. The black bar indicates 0.2 mm.

of Tb(III) to ensure the binding site was populated (Fig. 8). The resultant crystals were large (0.6 × 0.2 mm) and the quality appeared sufficient for structural analysis. Pleasingly, both HC03 and HC0C crystallised exceptionally well from the same conditions with crystals of similar dimensions in multiple crops. However, attempts to crystallise either HC01 or HC04 proved to be more challenging.

Concerning HC01 and HC04 it was considerably trickier to instigate crystal growth, with high concentrations of peptide resulting in precipitate but lower concentrations yielding nothing but empty wells. The resultant solution involved PEG400 doped samples of the previous hit conditions to yield HC04 crystals that were sizeable and appeared to be useable (Fig. 9B). Attempts to crystallise HC01 have resulted in crystalline sprays (Fig. 9A) which are very hard to handle. Even with the preparation of individual plates by excising from the sprays, the plates proved to have such high mosaicity as to be unusable. Considering the poor crystal growth,

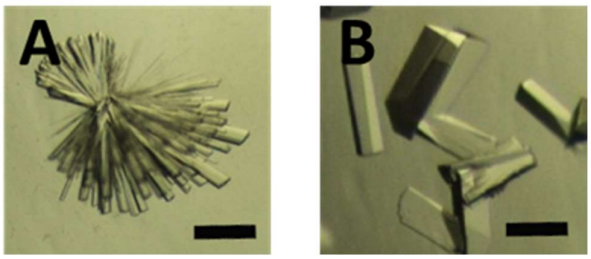

Figure 9: Photographs of crystals of A) HC01 as a spray of thin plates and B) HC04 as long triangular prisms. These crystals were grown from a combined solution of peptide (16 mg mL^{-1} peptide, 6 mM terbium chloride and 5 mM zinc acetate) and precipitant (0.21 M ammonium chloride, 25% v/v poly(ethylene) glycol 3350 and 8% v/v poly(ethylene) glycol 400) in a 1:1 ratio. The black bar indicates 0.2 mm.

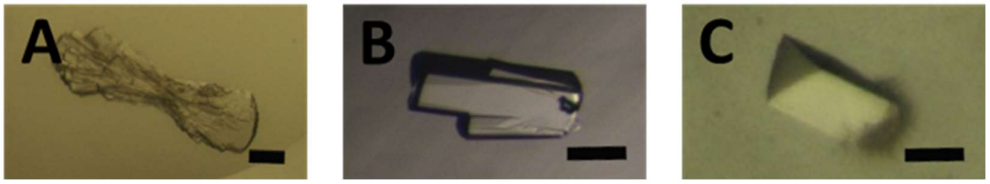

Figure 10: Photographs of A) the amorphous structure from HC03 grown in the presence of Yb(III), B) the triangular prisms of HC03 grown with Eu(III) and C) the triangular prisms of HC03 grown with Gd(III). These crystals were grown from a combined solution of peptide (16 mg mL^{-1} peptide, 5 mM zinc acetate and either 6 mM ytterbium chloride, gadolinium chloride or europium chloride) and precipitant (0.21 M ammonium chloride and 25% v/v poly(ethylene) glycol 3350) in a 1:1 ratio. The black bar indicates 0.1 mm.

seeding with the sprays was attempted into stocks of varying concentration of HC01, however, this produced either nothing or the same result as previously observed. The difficulties experienced with HC01 and HC04 may be due to disruption at the terminal sites resulting in a lack of packing at these positions, preventing good crystal nucleation and growth.

Considering the similarity of coordination behaviour across the lanthanides it was investigated whether crystals could be grown in the presence of lanthanides other than Tb(III). Crystals of HC02 and HC03 were successfully grown in the presence of Eu(III) and Gd(III) ions, however, in the presence of other lanthanides such as Yb(III) and La(III) this proved unsuccessful, resulting in poorly structured amorphous species (Fig. 10A). This may have been a problem associated with concentrations of the metal ion and peptide samples, however, the same result was observed for a number of different iterations. It may also be that the more charge dense Yb(III) or the large La(III) disrupt the structure too much to allow for good crystal growth. Attempts were made to crystallise the peptides in the presence of only lanthanide ions, however, none of these conditions were successful. This indicated that the zinc ions were essential to crystal growth and this further implied that the histidine residues were important for interhelical bridging and nucleating crystal growth. It should be noted that while histidine and zinc have a relatively high affinity, there are also many other coordinating residues such as glutamates available.

The crystals had to be cryo-protected before being frozen in liquid nitrogen. This prevented water in the peptide matrix forming ice-crystals, which could both damage the structure of the peptide crystals and lead to diffraction from amorphous ice layers resulting in poor data collection.[30] The cryo-protectant was a mixture of ethylene glycol and low molecular weight poly(ethylene glycol). These molecules disrupt hydrogen bonding in water preventing rapid crystalline solidification upon cooling, thus reducing damage to the peptide-water matrix. Of all the crystals that were protected, only one showed any signs of shocking (catastrophic loss of structure) and therefore was not mounted. Crystals of HC0C, HC02, HC03 and HC04 were mounted in loops, frozen and placed in a puck to be transported to the synchrotron.

2.7 X-RAY FLUORESCENCE AND EDGE ANALYSIS

X-ray fluorescence spectroscopy is a technique that makes use of the unique core electron structure of each element to provide an elegant tool for elemental analysis. This is carried out by first exciting a core electron with a primary high energy excitation source. The instability generated by this event results in higher orbital electrons 'falling' to fill the hole. The transitions associated with the electron movement results in a measureable emission which provides a unique signature for any element present. In order to assess the presence of Tb(III) within the crystal matrix a crystal of HC03, which had been grown in the presence of Tb(III) ions, it was excited with high energy electromagnetic radiation of 12658 eV (Fig. 11). The resultant X-ray fluorescence spectrum had distinctive peaks at 5626 eV, 6292 eV, 7002 eV, 7388 eV and 8098 eV indicating the presence of terbium, and 8633 eV and 9632 eV indicating the presence of zinc. These results were found to be in agreement with the literature.[31]

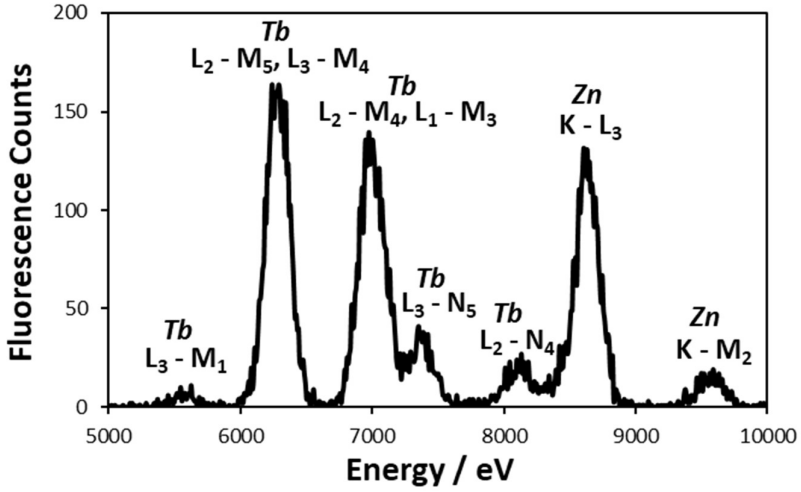

Figure 11: Fluorescence emission spectrum from a HC03 crystal grown in the presence of both Tb(III) and Zn(II) after irradiation with energy of 12658 eV. The peaks are labelled with the metal and transition responsible for the emission.

Following the successful X-ray fluorescence spectroscopy indicating the presence of Tb(III) and Zn(II), an absorption edge scan was carried out. This involved observing the absorption of the crystal when it was excited with radiation from 7200 – 7800 eV. The clear absorption edge present at 7519 eV was from the L_3 edge of terbium due to excitation of the electrons populating the 2p (J=3/2) orbitals.[32] This provided further supporting evidence of the presence of Tb(III) in the samples. Interestingly, the absorption profile had clearly defined x-ray absorption near edge structure and extended X-ray absorption fine structure, however, this was not explored further due to the redundancy afforded by the crystal structure data.

This edge scan also provided the data required to set the wavelength for crystal structure collection in order to maximise the terbium scattering compared to the zinc scattering. This would allow easy identification and positioning of the different ions by their scattering intensities. Predictive profiles for the edges of terbium and zinc (*see* S25) clearly demonstrated that at any wavelength within the beam range the scattering for Tb(III) would be considerably greater than for Zn(II).[33]

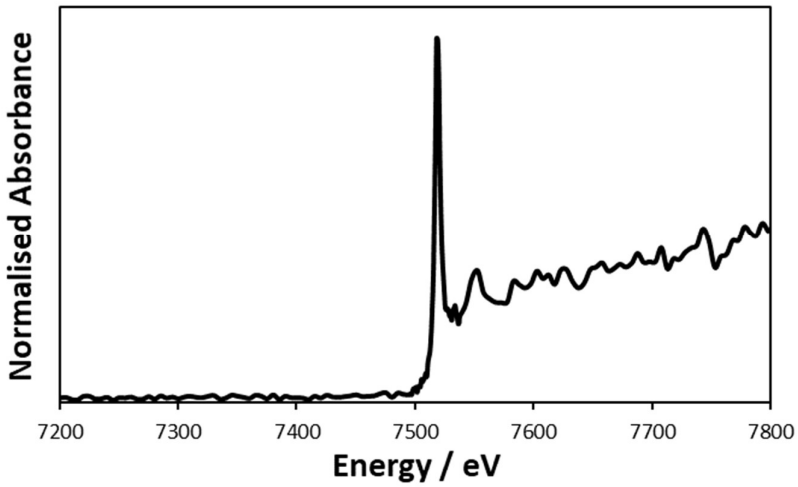

Figure 12: The absorption edge profile for a HC03 crystal with Tb(III) bound, scanned from 7200-7800 eV.

2.8 STRUCTURE DETERMINATION

The diffraction pattern images for the HC0C, HC02, HC03 and HC04 peptides were collected and the resultant datasets of the images were analysed and merged (Fig. 13). The raw data was found to be of good resolution with spots that were uniform in shape indicating the patterns should be easily identifiable. The overall statistics for the data collection are shown in Table 5.

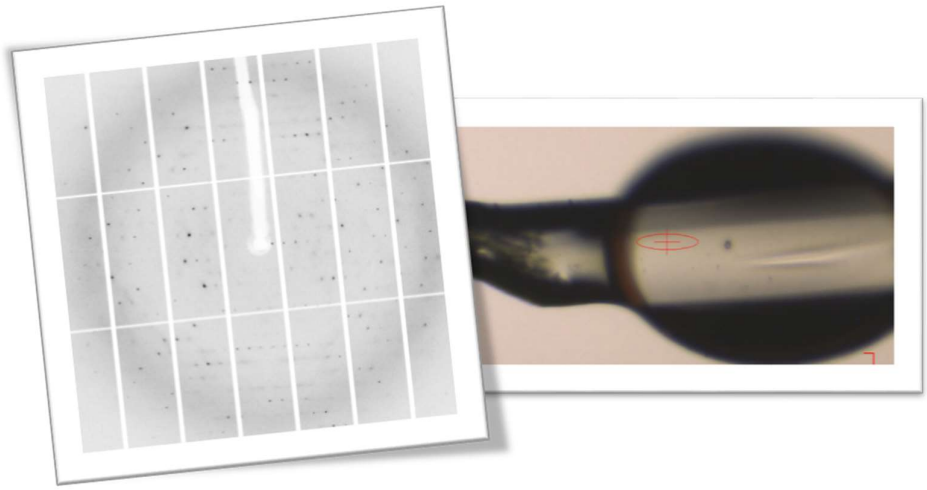

Figure 13: An example of the diffraction pattern from a single image and a photograph of a mounted HC0C crystal in a loop on the beamline.

Table 5: The crystal properties and data collection parameters associated with the HC series. The values in parenthesis are for the highest resolution set.

	HC02	**HC03**	**HC04**	**HC0C**
Wavelength (Å)	0.9282	0.9282	0.9282	0.9795
Resolution Range (Å)	41.5-2.1 (2.2-2.1)	41.4-2.1 (2.2-2.1)	21.6-3.0 (3.1-3.0)	30.0-2.1 (2.2-2.1)
Space Group	*H3*	*H3*	*H3*	*H3*
Unit Cell				
a=b, c	103.4, 46.9	103.2, 46.7	101.4, 47.2	103.8, 46.2
α=β, γ	90, 120	90, 120	90, 120	90, 120
No. Total Reflections	49925 (5130)	100183 (10555)	34800 (3295)	92072 (8228)
No. Unique Reflections	10820 (1073)	10821 (1097)	3598 (357)	10834 (1086)
Multiplicity	4.6 (4.8)	9.3 (9.6)	9.7 (9.2)	8.5 (7.6)
Completeness (%)	100 (100)	99.8 (99.6)	99.5 (100)	99.9 (99.8)
I / σ(I)	13.2 (2.8)	16.3 (2.59)	13.2 (5.8)	8.6 (1.0)
Wilson B ($Å^2$)	54.0	52.9	71.4	50.2
R_{merge}	0.08 (0.45)	0.08 (0.90)	0.16 (0.43)	0.14 (2.46)
$CC_{1/2}$	0.99 (0.93)	1.00 (0.91)	1.00 (0.87)	1.00 (0.35)

The resolution across the peptides was overall relatively good (2.1 Å) except for HC04 at 3.0 Å. Nevertheless, all these datasets are high enough resolution to determine a number of key structural features. The $CC_{1/2}$ and R_{merge} values were acceptable, suggesting that the diffraction data had good agreement, however, at higher resolution the quality is poorer due to the high mosaicity of the crystals that were used. This high mosaicity may have been due to the cryo-protectant conditions utilised, hence it is worth investigating alternative conditions to improve the crystal quality. Overall the completeness of the datasets was very good with all peptides having greater than 99%. The space groups for all the peptides were determined to be *H3*,

which is particularly useful as the high symmetry associated with this space group implies that a shorter rotation is required to collect a full set of data. The data was merged and reduced in iMosflm to give the complete reflection files.[34] Attempts to collect data from a single crystal plate drawn from the previously discussed HC01 crystalline sprays was unsuccessful in generating useable data. The data had to be truncated at very high resolutions ~6-9 Å due to the incredibly high R_{merge} values and resulted in very poor quality datasets. Nothing useful was found from any subsequent HC01 sample, most probably due to the very poor crystal formation causing high mosaicity.

The collected datasets were phased within the Phenix interface utilising the high anomalous scattering due to the large terbium ions present in the crystal structure.[35] This yielded both anomalous scattering and density maps which allowed for rapid analysis (Fig. 14). It was immediately apparent from the density that there were three α-helices in close proximity

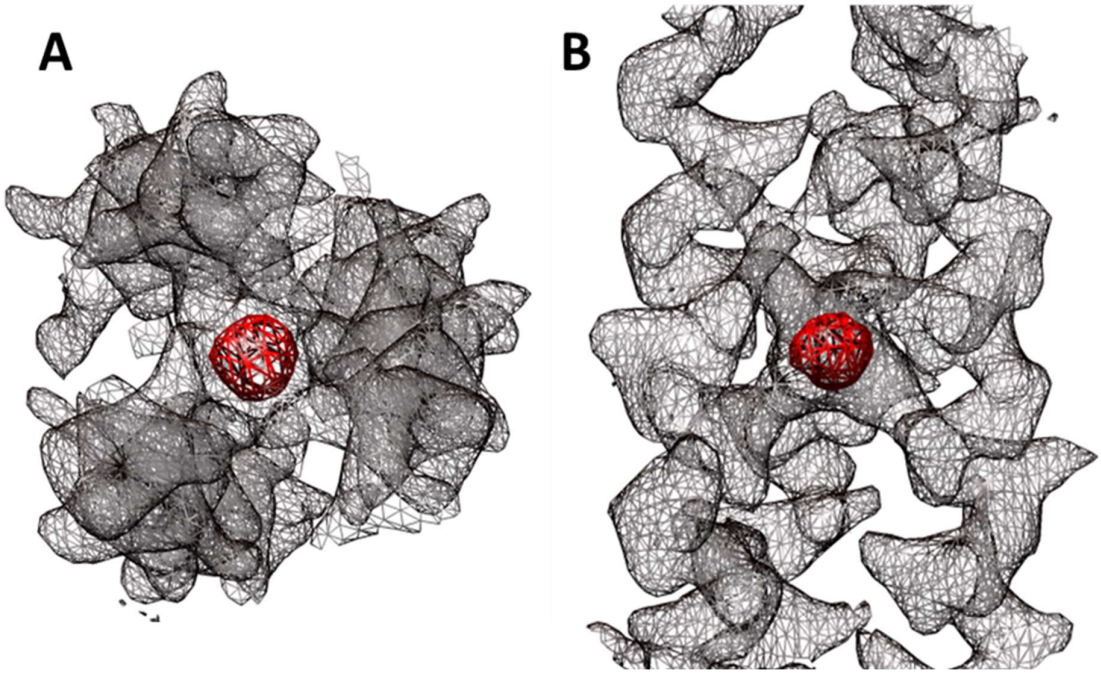

Figure 14: The electron density maps (black mesh, r.m.s.d. 1.4) for HC02 with the anomalous scattering map (red mesh, r.m.s.d. 3.5) indicating the position of a bound Tb(III) ion, A) top down view and B) side view.

forming a trimeric coiled coil assembly. Furthermore, the anomalous scattering map can be used to place a Tb(III) ion at the core of the coiled coils for HC02, HC03 and HC04. The position of the metal ion, as indicated by the anomalous scattering, was in agreement with the designed modulation of the binding site down the coiled coil. In addition, α-helices of the coiled coils were clearly arranged parallel to each other as the side-chains were all angled in the same direction, resulting in a 'Christmas tree' effect as illustrated by the electron density map of HC0C (Fig. 15). The system with the greatest chance to have ambiguous alignment, either parallel or antiparallel, was the control HC0C as this design did not have a metal binding site to template the parallel structure. Therefore to ensure this system was indeed parallel a further two crystals from different droplets were analysed and the density plots confirmed that all the crystals resulted in parallel solid state structures.

The modelling process was facilitated by fitting the expected sequence to the available density, this resulted in three α-helices being placed. The metal ions were easily positioned by overlaying the anomalous scattering map with the electron density map. This rough model was then further refined by manual rotation and translation of residues using the coot package that yielded a series of refined coordinate files which, while this work was carried out by O.

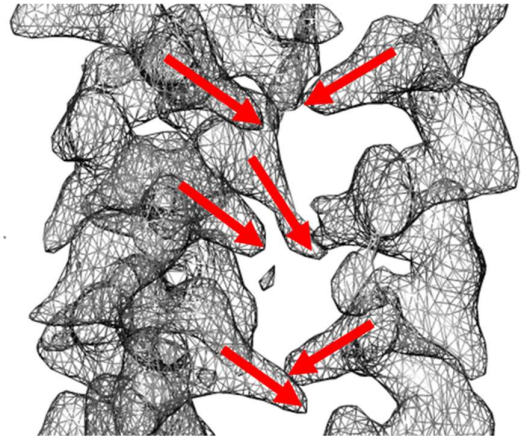

Figure 15: The electron density map (black mesh, r.m.s.d. 1.5) of HC0C indicating the same directionality of the side chains suggesting a parallel alignment of helices forming the coiled coil.

Daubney, it was subsequently checked by Dr. White. These files provided the solid state crystal structures of HC0C, HC02, HC03 and HC04 all with Tb(III) and Zn(II) ions coordinated (Fig.16). The refinement statistics and fit parameter can be found in Table 6.

The models made some of the previous observations easier to characterise and there are certain general features that applied to all the structures. They all adopted a parallel trimeric arrangement of α-helical monomers to give the desired coiled coils, in agreement with the LOGICOIL predictions and original designs. Interhelical salt bridging is also clearly visible between each α-helix (Fig. 17A) which is characterised by a distance of 4 Å or less between the two bridging residues, in this case lysine and glutamate.[36] The core is made up of predominantly hydrophobic isoleucine residues, as expected from their ***a*** and ***d*** positions (Fig. 17B). The termini of all the studied peptides are more dynamic as exhibited by the high B-factors associated with the residues at these positions. This is not unexpected as previous studies of coiled coils have demonstrated that the ends can be frayed.[37] The Ramachandran plots (*see* S26-S29) demonstrated that the histidine residues were distorted from the conformation of well folded α-helices, most likely due to the influence of neighbouring residues with interhelical bridging metal ion coordination environments. The presence of metal ions in external sites is most likely the cause of successful crystal growth.

The external surface of the peptide is particularly functionalised with residues at both the top and bottom of the helix engaging in metal coordination, indeed this was intended as it would encourage initial crystal formation. While the Zn(II) was added to coordinate to the histidine residues, it was apparent that the histidine residues are not involved in the metal ion coordination at the termini. The external sites are complex and somewhat difficult to assign

Table 6: The crystallographic refinement statistics for the HC series.

	HC02	HC03	HC04	HC0C
No. of Non-H Atoms	996	926	909	1042
No. of Waters	111	41	25	144
R_{work} / R_{free}	0.20 / 0.22	0.22 / 0.24	0.21 / 0.25	0.22 / 0.24
CC_{work} / CC_{free}	0.97 / 0.96	0.95 / 0.98	0.96 / 0.97	0.95 / 0.97
R.m.s.d.				
Bonds (Å)	0.005	0.004	0.006	0.003
Angles (°)	0.55	0.81	0.89	0.51
Ramachandran[a]	97 / 3 / 0 / 0	99 / 1 / 0 / 0	98 / 2 / 0 / 0	98 / 2 / 0 / 0
Average *B* (Å^2)	77.1	74.3	73.4	63.9

[a]*The values indicate the percentage of residues in Ramachandran favoured/allowed/outliers/rotamer outliers*.

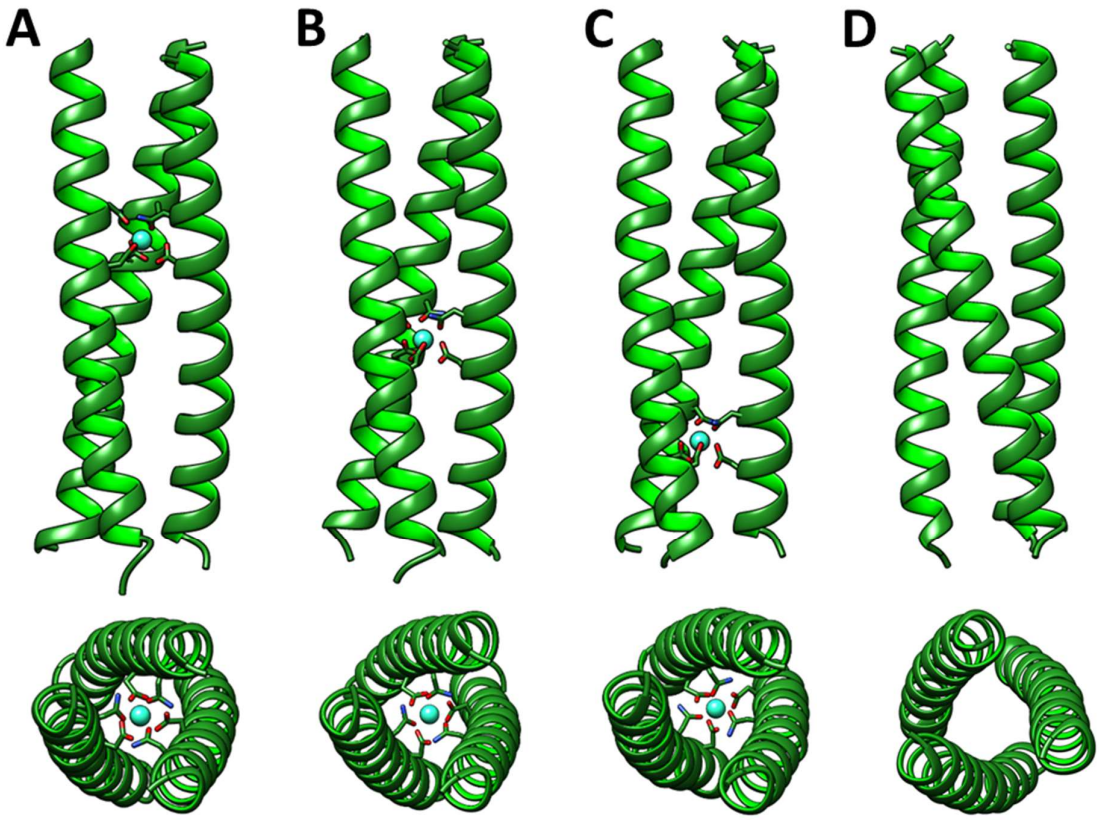

Figure 16: The crystallographic structures shown in cartoon ribbon with the core binding sites illustrated (terbium is cyan, oxygen is red, nitrogen is blue and carbon is green) for A) HC02, B) HC03, C) HC04 and D) HC0C with both a side and top down view.

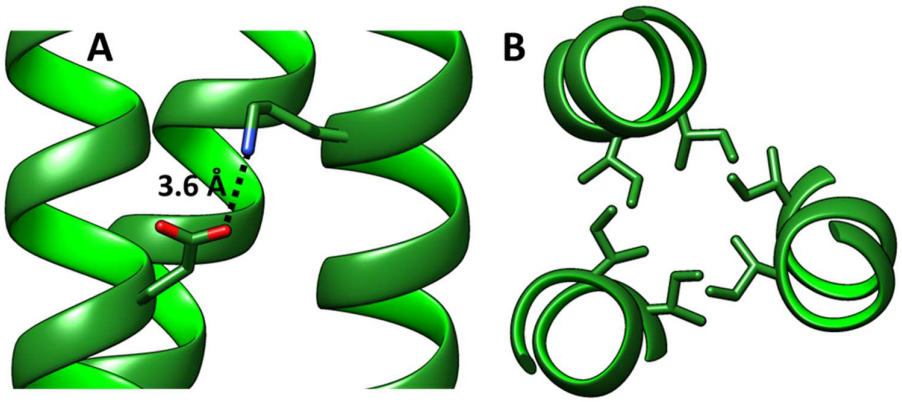

Figure 17: The crystallographic structure of HC0C indicating an A) interhelical salt-bridge between a lysine and glutamate residue and B) the hydrophobic core of the trimer consisting of packed isoleucine residues.

due to their dynamic nature. However, they appear to consist of four glutamate residues, two from the top of an α-helix of one coiled coil and two from the bottom of an α-helix of another coiled coil. The coordination mode of the acidic sidechains varies and water molecules are also involved in the external site coordination sphere. The identity of the metal ions was easily determined from their anomalous scattering profile with the large intensities clearly indicating the presence of a majority of Tb(III) and a lower intensity indicating the presence of some Zn(II) sites (Fig. 18). This is particularly interesting as attempts to grow crystals without Zn(II) proved unsuccessful, therefore these external sites must be essential to crystal nucleation and growth. The importance of these sites may also indicate why there were difficulties in growing both HC01 and HC04 systems. The terminally situated core binding sites would restrict the flexibility of the required glutamate residues thus restricting the interhelical metal binding required for crystal formation.

The core lanthanide binding sites are well resolved for both HC02 and HC03, and less so for the lower resolution HC04 data (Fig. 19). Each α-helix has both a coordinating acid and amide functional group from the aspartate and asparagine residues, respectively. The identity of the metal in the designed site was determined from the anomalous scattering to be a Tb(III) ion.

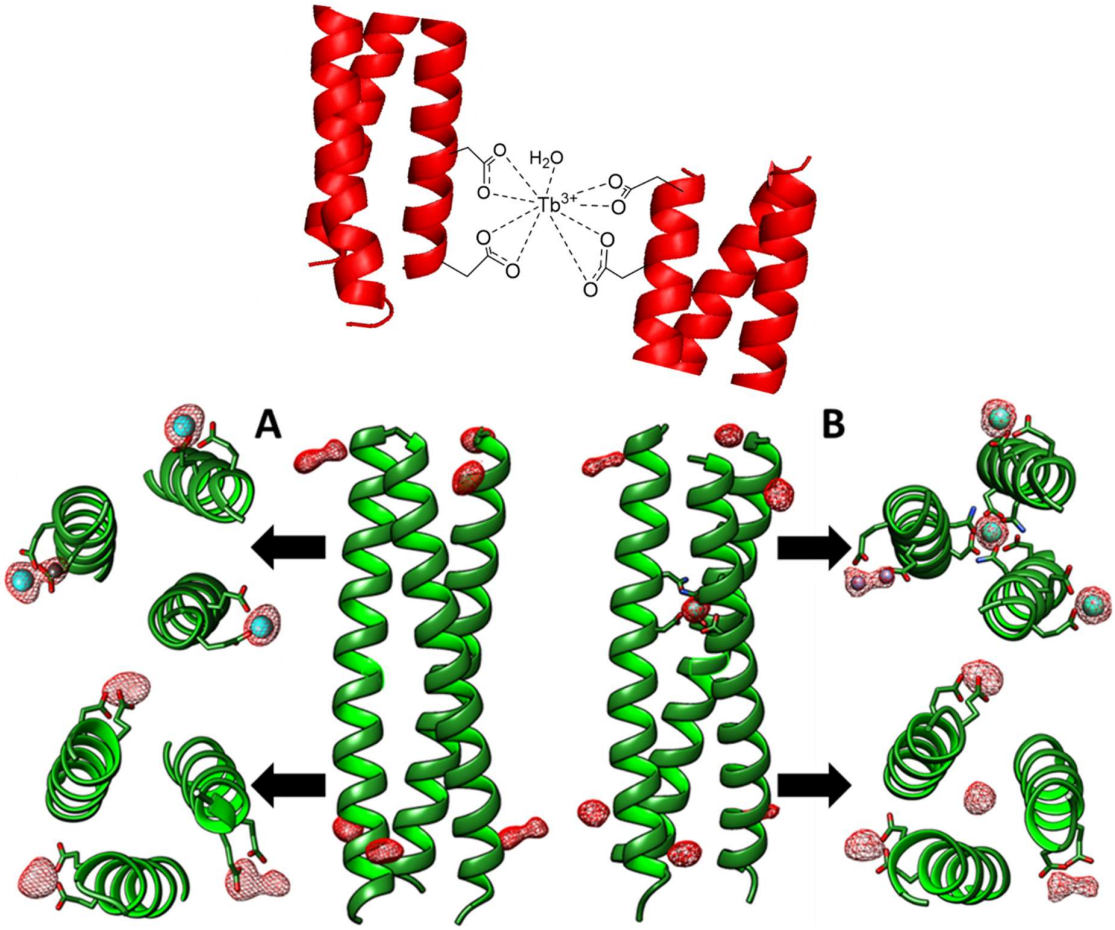

Figure 18: A schematic illustrating the interhelical bridging external binding sites and the crystallographic structure of A) HC0C and B) HC02 with the anomalous scattering map (red mesh, r.m.s.d. 4.0) indicating the positions of the bound metal ions. The size of the sphere is directly related to the anomalous scattering intensity and therefore indicates either the presence of Tb(III) (cyan spheres) or Zn(II) ions (grey spheres).

Following the placement of the metal ion the distance from each coordinating atom to the Tb(III) ion was measured (*see* Table 7). Typically all the distances were found to be between 2.0 and 2.7 Å, with an average of 2.3 Å in agreement with previous peptide and carboxylate ligands coordinated to Tb(III).[38] Concerning the asparagine residues there is no discernible indication in the electron density of whether the nitrogen or oxygen from the amide are the coordinating atom. However, the Tb-X distance for asparagine are in agreement with the shorter Tb-O literature distances and not with the longer Tb-N distances.[39] This is further supported by the knowledge that lanthanide trications are regarded as very hard metal ions due to their ionic bonding character and would thus be expected to favour the harder oxygen atom

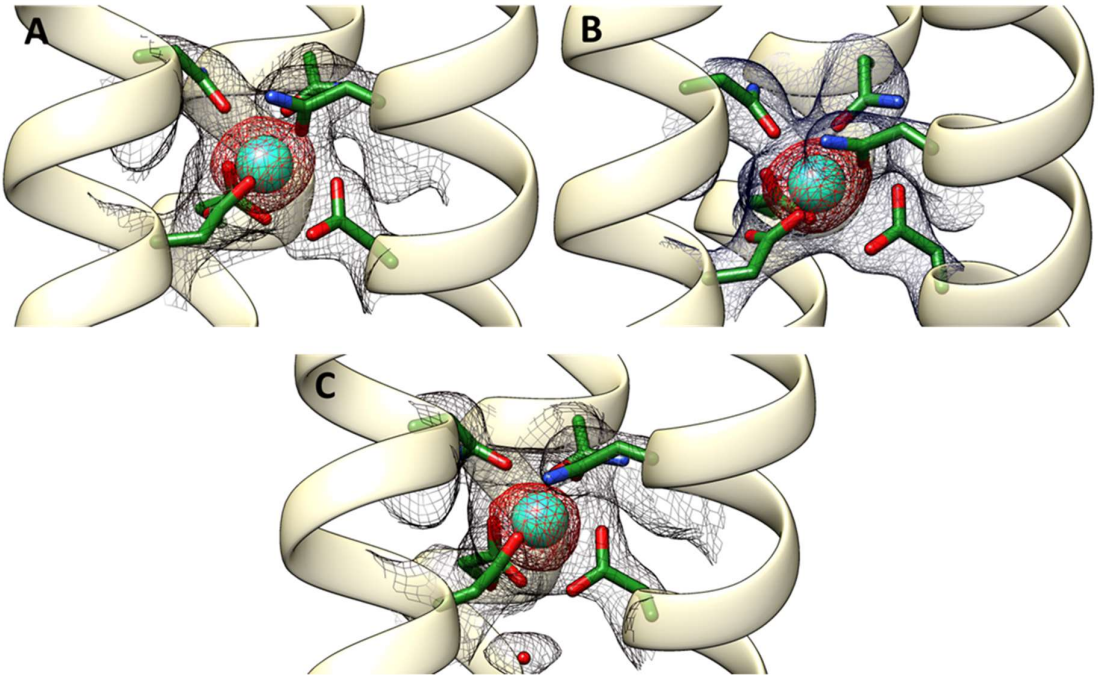

Figure 19: The crystallographic structure of the core binding sites for A) HC02, B) HC03 and C) HC04 with the electron density map (black mesh, r.m.s.d. 1.1) and anomalous scattering map (red mesh, r.m.s.d. 3.5) overlaid to indicate the side chain and metal ion placement.

over the relatively softer nitrogen.[40] Therefore, the models have been generated in which all the designed binding residues have an oxygen interacting with the Tb(III) ion. Interestingly, the varied distances to the metal ion from coordinating residues indicates a lack of site symmetry.

There is a particularly long Tb-O distance in each system, at 2.8, 2.8 and 3.0 Å for HC02, HC03 and HC04 respectively. These long distances indicate that the atom is perhaps not just involved in coordination with the Tb(III) ion but also the local environment. This suggests that the Tb(III) ion is in a 8-9 coordinate geometry with some ambiguity as to the 9 coordinate state. A localised water was observed close to the binding site in each case, indeed, the environment of the water molecule was very similar for all the systems (Fig. 20). For both HC02 and HC04 there appeared to be an interaction between an aspartate residue and a water molecule as indicated by the short distance between them. This would weaken the bonding to the Tb(III)

Table 7: The measured terbium to oxygen distances for the designed core binding sites of HC02, HC03 and HC04. The corresponding chain identity is in parenthesis. Error of measured distances is ± 0.1 Å.

Tb(III)-Oxygen Distances (Å)[a]	HC02	HC03	HC04
Asn(A)	2.4	2.5	2.0
Asn(B)	2.3	2.3	2.2
Asn(C)	2.1	2.3	2.5
Asp(A)	2.5, 2.3	2.7, 2.1	3.0, 2.3
Asp(B)	2.7, 1.8	2.7, 2.3	2.7, 2.1
Asp(C)	2.8, 2.5	2.7, 2.8	2.5, 2.1

[a] *The error in reported measurements can be found in Table 6 (0.005, 0.004 and 0.006 Å for HC02, HC03 and HC04 respectively), due to high B-factors associated with the model these were reported to one decimal place.*

ion and thus result in a longer Tb-O distance. In all cases the localised water molecule was situated too far away from the Tb(III) ion to be interacting directly. Nevertheless, dynamic water molecules would not be observed within the crystal structure, therefore water molecules in close proximity to the bound Tb(III) ion may still be possible. This will be important for understanding photophysics and quenching of the bound lanthanides in later chapters.

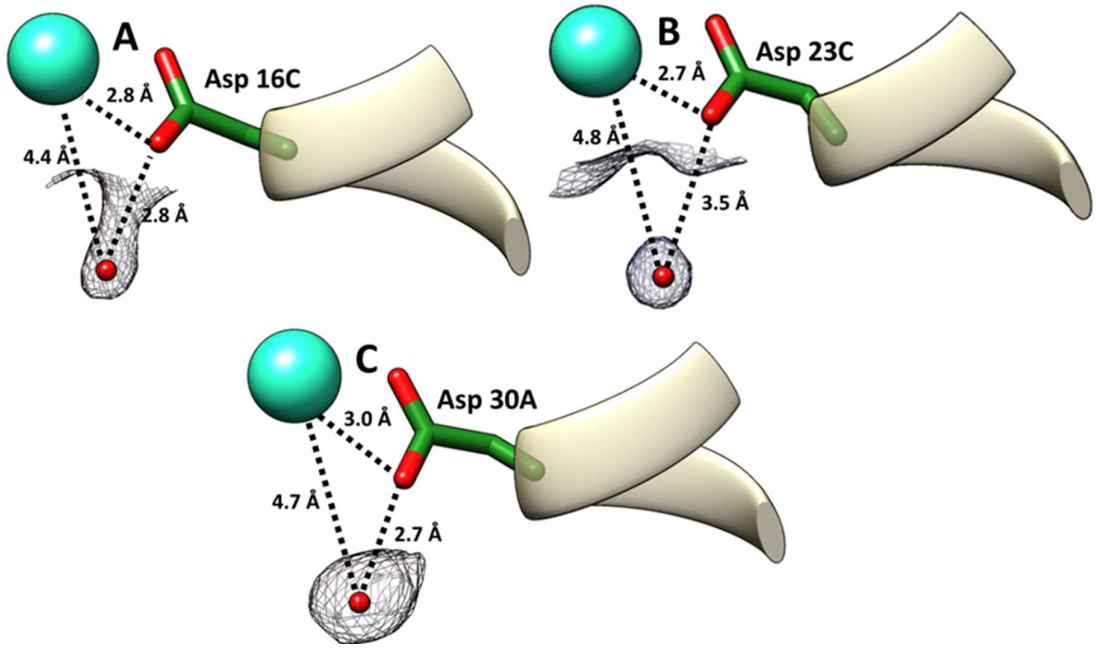

Figure 20: The crystallographic structure of a local water close to the bound Tb(III) for A) HC02, B) HC03 and C) HC04. The electron density map (black mesh, r.m.s.d. 0.9) and measured distances (black dotted line) are overlaid for clarity. Terbium is cyan, oxygen is red and carbon is green.

2.9 CONCLUSIONS

A new series of coiled coil trimers containing a lanthanide binding site was designed and successfully synthesised. Combining the lanthanide binding capabilities of the MB series together with the stability and crystallisation propensity of the CS series, the HC series aspired to be an ideal system to structurally understand an aspartate and asparagine functionalised lanthanide binding site. Solution state analysis utilising circular dichroism proved the presence of well-formed coiled coils for HC0C, HC01 and HC04 before the addition of Tb(III) ions. The structure of both HC02 and HC03 was disordered before the addition of one equivalence of Tb(III) ions after which they were well folded coiled coil trimers. This validated that the central sites of HC02 and HC03 disrupt the core and that the Tb(III) ions are able to template the trimer formation, as designed. To further characterise the peptides, electrospray mass spectrometry was used to observe the intact trimeric complexes. While, HC01-HC03 were successfully observed with Tb(III) ions bound, once again HC04 was only observed as a non-metallated trimeric species even after the addition of excess Tb(III) ions. This suggested that the integrity of the terminal binding site for HC04 may be compromised in the solution state.

Crystals of HC01-HC04 and HC0C were successfully grown, unfortunately the HC01 crystals were small and poor quality, therefore data was collected for all but HC01. All the studied peptides crystallised as parallel trimers of α-helices to form the desired coiled coils. HC02-HC04 all exhibited a core binding site with Tb(III) present, coordinated to oxygen donors from each of the binding site residues. The asparagine residues were monodentate, which was modelled to be the oxygen of the amide, and the aspartate residues were bidentate with two oxygen donors from the acid side chain. There was some ambiguity regarding the coordination

number, as one of the Tb-O distances was particularly long for each system suggesting that it could be considered 8/9 coordinate. External binding sites at the both the *N*- and *C*- termini, composed of glutamate residues coordinated to Tb(III) and Zn(II) ions, were responsible for the interactions between coiled coil species and are most likely the source of crystal nucleation and growth. Having studied the structures of the HC peptides, it remains to explore the photophysical properties of these systems.

2.10 REFERENCES

[1] M. L. Zastrow, A. F. A. Peacock, J. A. Stuckey and V. L. Pecoraro, *Nat. Chem.*, 2012, **4**, 118-123.

[2] i) D. Shiga, D. Nakane, T. Inomata, Y. Funahashi, H. Masuda, A. Kikuchi, M. Oda, M. Noda, S. Uchiyama, K. Fukui, K. Kanaori, K. Tajima, Y. Tanako, H. Nakamura and T. Tanaka, *J. Am. Chem. Soc.*, 2010, **132**, 18191-18198; ii) M. Faiella, A. Roy, D. Sommer and G. Ghirlanda, *J. Pept. Sci.*, 2013, **100**, 558-571.

[3] R. B. Hill, D. P. Raleigh, A. Lombardi and W. F. DeGrado, *Acc. Chem. Res.*, 2000, **33**, 745-754.

[4] Z. T. Ball, *Curr. Opin. Chem. Biol.*, 2015, **25**, 98-102.

[5] D. J. Wilger, S. E. Bettis, C. K. Materese, M. Minakova, G. A. Papoian, J. M. Papanikolas and M. L. Waters, *Inorg. Chem.*, 2012, **51**, 11324-11338.

[6] L. Zhou, M. Bosscher, C. Zhang, S. Ozcubukcu, L. Zhang, W. Zhang, C. J. Li, J. Liu, M. P. Jensen, L. Lai and C. He, *Nat. Chem.*, 2014, **6**, 236-241.

[7] i) G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado and V. L. Pecoraro, *J. Am. Chem. Soc.*, 1997,

119, 6195-6196; ii) V. Cangelosi, L. Ruckthong and V. L. Pecoraro, *Met. Ions. Life Sci.*, 2017, **17**, 1-60.

[8] i) M. L. Zastrow and V. L. Pecoraro, *Coord. Chem. Rev.*, 2013, **257**, 2565-2588; ii) F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum, V. L. Pecoraro, *Chem. Rev.*, 2014, **114**, 3495-3578.

[9] i) J. -C. G. Bunzli, *Chem. Rev.*, 2010, **110**, 2729-2755; ii) R. D. Teo, J. Termini and H. B. Gray, *J. Med. Chem.*, 2016, **59**, 6012-6024.

[10] W. D. Kohn, C. M. Kay and R. S. Hodges, *J. Peptide Res.*, 1998, **51**, 9-18.

[11] i) A. Kashiwada, K. Ishida and K. Matsuda, *Bull. Chem. Soc. Jpn.*, 2007, **80**, 2203-2207; ii) A. Niedzwiecka, F. Cisnetti, C. Lebrun and P. Delangle, *Inorg. Chem.*, 2012, **51**, 5458-5464.

[12] M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton and A. F. A. Peacock, *J. Am. Chem. Soc.*, 2014, **136**, 1166-1169.

[13] i) M. R. Berwick, PhD Thesis, University of Birmingham, 2016; ii) L. N. Slope, PhD Thesis, University of Birmingham, 2018.

[14] i) K. T. O'Neil and W. F. DeGrado, *Science*, 1990, **250**, 646-651; ii) B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado and D. Eisenberg, *Science*, 1993, **259**, 1288-1293.

[15] L. Ruckthong, M. L. Zastrow, J. A. Stuckey and V. L. Pecoraro, *J. Am. Chem. Soc.*, 2016, **138**, 11979-11988.

[16] D. E. Chirkst, O. L. Lobacheva and I. V. Berlinskii, *Russ. J. Phys. Chem. A*, 2010, **84**, 2047.

[17] J. M. Fletcher, A. L. Boyle, M. Bruning, G. J. Bartlett, T. L. Vincent, N. R. Zaccai, C. T. Armstrong, E. H. C. Bromley, A. R. Thomson and D. N. Woolfson, *ACS Synth. Biol.*, 2012, **1**, 240-250.

[18] M. Meier, J. Stetefeld and P. Burkhard, *J. Struct. Biol.*, 2010, **170**, 192-201.

[19] M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton and A. F. A. Peacock, *Chem. Sci.*, 2016, **7**, 2207-2216.

[20] i) M. Delorenzi and T. Speed, *Bioinformatics*, 2002, **18**, 617-625; ii) L. Zimmermann, A. Stephens, S.-Z. Nam, D. Rau, J. Kubler, M. Lozajic, F. Gabler, J. Soding, A. N. Lupas and V. Alva, *J. Mol. Biol.*, 2018, **430**, 2237-2243; iii) MARCOIL programme, https://bcf.isb-sib.ch/webmarcoil/webmarcoilC1.html, (accessed July 2018).

[21] i) T. L. Vincent, P. J. Green and D. N. Woolfson, *Bioinformatics*, 2013, **29**, 69-76; ii) LOGICOIL programme, http://coiledcoils.chm.bris.ac.uk/LOGICOIL, (accessed July 2018).

[22] Y. H. Chen, J. T. Yang and K. H. Chau, *Biochemistry*, 1974, **13**, 3350-3359.

[23] i) N. E. Zhou, C. M. Kay and R. S. Hodges, *J. Biol. Chem.*, 1992, **267**, 2664-2670; ii) N. E. Zhou, C. M. Kay and R. S. Hodges, *Biochemistry*, 1992, **31**, 5739-5746; iii) T. T. Zheng, A. Boyle, H. R. Marsden, D. Valdink, G. Martelli, J. Raap and A. Kros, *Org. Biomol. Chem.*, 2015, **13**, 1159-1168.

[24] G. D. Crescenzo, J. R. Litowski, R. S. Hodges and M. D. O'Connor-McCourt, *Biochemistry*, 2003, **42**, 1754-1763.

[25] i) M. Mann, *Org. Mass Spectrom.*, 1990, **25**, 575-587; ii) A. J. R. Heck and R. H. H. van den Heuvel, *Mass Spectrom. Rev.*, 2004, **23**, 368-389.

[26] H. Wendt, E. Durr, R. M. Thomas, M. Przybylski and H. R. Bosshard, *Protein Sci.*, 1995, **4**, 1563-1570.

[27] H. Campbell, MSci Thesis, University of Birmingham, 2016.

[28] A. Laganowsky, M. Zhao, A. B. Soriaga, M. R. Sawaya, D. Cascio and T. O. Yeates, *Protein Sci.*, 2011, **20**, 1876-1890.

[29] M. Laitaojo, J. Valjakka and J. Janis, *Inorg. Chem.*, 2013, **52**, 10983-10991.

[30] D. W. Rodgers, *Structure*, 1994, **2**, 1135-1140.

[31] J. A. Ibers and W. C. Hamilton, *International Tables For X-ray Crystallography*, The Kynoch Press, Birmingham, 1974, vol. 4.

[32] R. C. Miake-Lye, S. Doniach and K. O. Hodgson, *Biophys. J.*, 1983, **41**, 287-292.

[33] Edgeplots Web Tool, http://skuld.bmsc.washington.edu/scatter/AS_form.html, (accessed December 2018).

[34] i) T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. W. Leslie, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2011, **67**, 271-281; ii) M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2011, **67**, 235-242.

[35] i) A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674; ii) P. D. Adams, R. W. Grosse-Kunstleve, L. – W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2002, **58**, 1948-1954.

[36] S. Kumar and R. Nussinov, *ChemBioChem*, 2002, **3**, 604-617.

[37] i) M. E. Holtzer, E. G. Lovett, D. A. d'Avignon and A. Holtzer, *Biophys. J.*, 1997, **73**, 1031-1041; ii) B. Triplet and R. S. Hodges, *J. Struct. Biol.*, 2002, **137**, 220-235.

[38] i) D. Parker, R. S. Dickins, H. Puschmann, C. Crossland and J. A. K. Howard, *Chem. Rev.*, 2002, **102**, 1977-2010; ii) B. Barja, R. Baggio, M. T. Garland, P. F. Aramendia, O. Pena and M. Perec, *Inorganica Chim. Acta*, 2003, **346**, 187-196; iii) M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen and B. Imperiali, *Chem. Int. Ed.*, 2004, **43**, 3682-3685.

[39] i) P. A. Stabnikov, G. I. Zharkova, A. I. Smolentsev, N. V. Pervukhina and V. V. Krisyuk, *J. Struct. Chem.*, 2011, **52**, 560-567; ii) S. Shintoyo, T. Fujinami, N. Matsumoto, M. Tsuchimoto, M. Weselski, A. Bienko and J. Mrozinski, *Polyhedron*, 2015, **91**, 28-34.

[40] J.-C. G. Bunzli, *J. Coord. Chem.*, 2014, **67**, 3706-3733.

Chapter 3

EXPLORING TRYPTOPHAN TO LANTHANIDE ENERGY TRANSFER

3.1 INTRODUCTION

The processes of energy capture and transfer are essential for the machinery of life. Natural systems make use of chromophores, proteins and membranes to absorb and transport light energy to enable the synthesis of carbohydrates, hormones and other fine chemicals.[1] After studying the structure of the designed HC series it was necessary to understand how energy would behave, and could be manipulated, within these systems. In nature's toolbox of amino acids the most efficient chromophore is the indole of tryptophan. This species has an absorption profile that allows relatively efficient absorption at 280 nm (ε_{280} = 5690 M^{-1} cm^{-1}).[2] While tryptophan is not a practical chromophore for harvesting solar light due to the low intensity of high energy UV-light at sea level, it does permit absorption and energy transfer phenomena to be studied within the coiled coil motif. Therefore, taking advantage of the tryptophan that was already incorporated in the sequence of the HC series, it was anticipated that this would enable the study of rudimentary light harvesting.

The binding of lanthanides within the HC structures provided a useful template for the assembly of the coiled coil trimer but this was not the only property of interest. As discussed previously (*see* Chapter 1), the lanthanides possess some unique and attractive photophysical features such as narrow emission bands, which allows easy identification of emission without excessive signal overlap, a common problem associated with the broad emission of organic species.[3] These emission profiles are also relatively environment independent, a useful characteristic when attempting to predict energy transfer and behaviour within a system. An interesting feature of the emissive states is that they exhibit very long lifetimes making them particularly appealing as energy traps for efficient handling of absorbed energy.[4] One of the

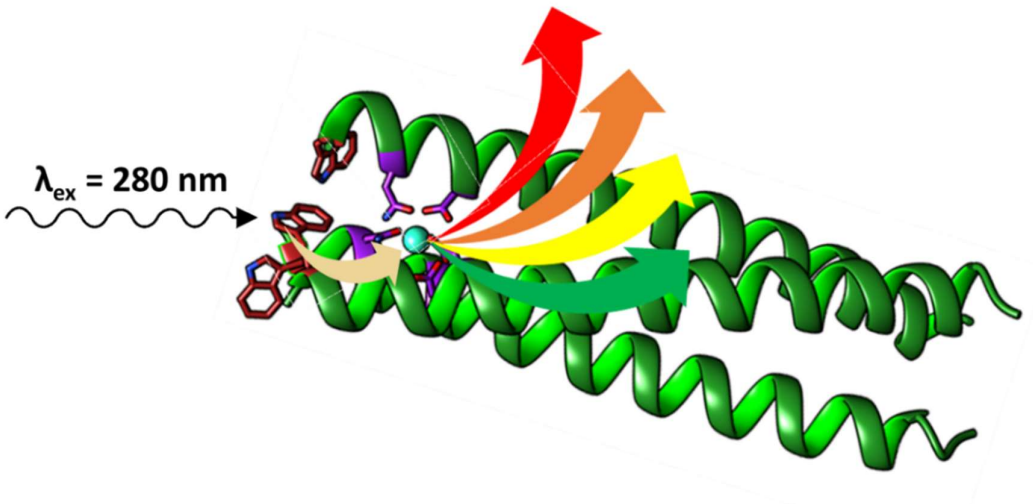

Figure 1: A schematic of the proposed simple light harvesting system with the tryptophan residues sensitising emission from different lanthanide ions.

most important and unique properties of the lanthanides is that of long range energy transfer, as demonstrated by the lanthanide resonance energy transfer distances for terbium. This behaviour where a donor species is capable of transferring energy over exceptionally large distances (>10 nm) with surprisingly high efficiency is particularly exciting when designing something akin to an antenna domain that can direct energy to a reaction centre.[5] Due to these unique photophysical properties certain lanthanides such as terbium and europium have found extensive use as probes in a diverse array of systems from solution state samples in assays to amorphous glass materials.[6] Indeed, they have been routinely used as luminescent probes for biological interactions.[7] However, a significant limitation of lanthanide species such as Tb(III) is that they have very poor absorption coefficients due to the Laporte forbidden 4f-4f transitions.[8] To circumvent the poor absorption of the lanthanides, sensitising chromophores have been used to indirectly excite them.[9] The key features of a successful sensitising species relates to criteria of the excited state chromophore. The quantum yield of emission from a Ln(III) sensitised by a chromophore (Φ_{Ln}^{Sens}) can be described as the product of the efficiency of sensitisation from the chromophore (E_{sens}) and the luminescence quantum yield of the

Ln(III) species (Φ_{Ln}^{Ln}).[10] To increase the efficiency of sensitisation there are a number of desirable features to incorporate into the chromophore. Firstly, the chromophore should exhibit efficient intersystem crossing to generate the useful triplet state species. Secondly, the energy of the triplet state should be greater than the acceptor lanthanide state by 2500 - 3500 cm^{-1} to facilitate energy transfer.[11] However, care should be taken to ensure that the energy of the sensitising triplet state is at least 1850 cm^{-1} greater than the lanthanide acceptor state. This is to prevent unwanted back transfer from the sensitised lanthanide to the sensitising chromophore.[12] The use of tryptophan as a sensitiser for both Tb(III) and Eu(III) has been well established for a number of different systems.[13]

With regards to taking advantage of lanthanide photophysics in biological systems, two approaches have been used, such as utilising pre-existing metal binding sites or attaching a tag with the metal already bound. The use of natural calcium binding sites was expected because of the similar size of lanthanide trications and Ca(II) ions, as well as some similarity in their coordination chemistry which suggested that interchangeable binding could occur. Studies in which lanthanides were bound in natural proteins, with calcium binding pockets, were utilised in the discovery of the ability of tryptophan or tyrosine to sensitise Tb(III) and Eu(III) emission.[13]

The study of lanthanides within protein systems has resulted in a number of interesting designs and discoveries, however, the photophysical aspects were explored most thoroughly by the Imperiali group. The lanthanide binding tag (LBT), a synthetic peptide system inspired by previous natural proteins, showcased excellent energy transfer and very strong emission intensity for Tb(III) ions.[14]

Subsequent work within the Peacock group involved the development of the MB series, a series of lanthanide binding coiled coil peptides. These peptides demonstrated different photophysical behaviour dependent upon the location of the site within the coil.[15, 16] Contrary to the HC design, the tryptophan residues of the MB series were translated down the coiled coil, being kept adjacent to the binding site. Due to the tryptophan residues being externally situated, ~11 Å from the bound metal, energy transfer efficiencies were very low and thus emission intensity from the coordinated Tb(III) ions was also low. The centrally located binding sites in the second and third heptads of the coiled coil resulted in low hydration states, but despite the protected metal ion resulting in low solvent induced quenching effects, the emission intensity was still very low due to the distance factors previously discussed. Interestingly, both terminal sites behaved differently, this is due to the *N*-terminal site being situated closer to the end of the coiled coil than the corresponding *C*-terminal binding site. This implies that the *C*-terminal site will behave more like a core site with relatively little exposure to the solvent environment. Whereas the *N*-terminal site will be more solvent exposed and dynamic. This same variability in behaviour is anticipated to be carried through to the HC series due to the same location translation of the binding site throughout the structure. Yet the retention of the tryptophan residues near the N-termini of the α-helices of the HC series are anticipated to alter the photophysical behaviour considerably. Furthermore, the tryptophan residues may act as a cap to the N-terminus preventing the water ingress to the bound metal in the HC01 system.

Before exploring the use of tryptophan to sensitise lanthanides it is advantageous to study the behaviour of the chromophore for both absorption and emission characteristics.

3.2 TRYPTOPHAN ABSORPTION

In order to understand the absorption characteristics of the HC series, it is necessary to consider the primary absorber, tryptophan. The chromophoric motif of the tryptophan residue is an aromatic heterocycle known as indole (Fig. 2). The aromatic nature of this species leads to low lying energetic transitions such as the π→π* transition. These available transitions result in

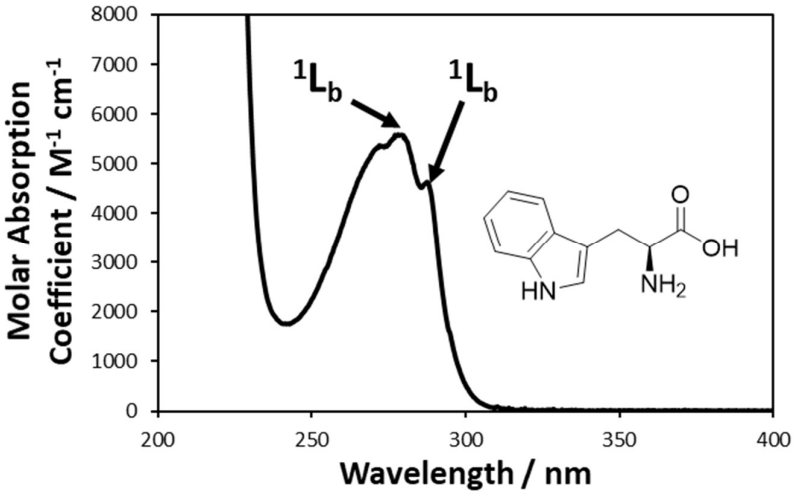

Figure 2: The molecular structure and absorption profile of tryptophan. The observable fine-structure from the ^1L_b state is indicated with arrows. The sample consisted of tryptophan (90 μM) in aqueous urea (7 M).

observable features in the absorption spectrum from 240 - 300 nm due to a π - π* transition.[17] The inherent lack of symmetry across the aromatic system causes the absorption profile to be anisotropic, resulting in two distinct states which can be populated. These are known as the ^1L_a and ^1L_b states with energies dependent upon the solvent environment.[18] The ^1L_b state does not absorb at lower energy than c.a. 290 nm, but the ^1L_a state can absorb up to 300 nm, this difference can be used to selectively excite the ^1L_a state. Interestingly, the ^1L_b state is not particularly sensitive to the solvent, whereas the ^1L_a state is very sensitive due to the solvent interactions with the amine moiety. The fine structure visible in the absorption profile, such as the λ_{max} at 290 and 280 nm, is primarily due to the ^1L_b state.

While the focus has been on single photon absorption it is known that tryptophan residues can be excited by 560 nm light, thus indicating a two photon absorption mechanism.[19] Although possible, the two photon absorption process is relatively inefficient due to the very low absorption cross section of the tryptophan aromatic system. This implies that unless high intensity 560 nm light is used, this phenomenon will not be observed and therefore is unlikely to be important when considering a tryptophan functionalised light harvesting complex.

3.3 TRYPTOPHAN EMISSION

The emission profiles of proteins are typically dominated by the influence of tryptophan, with smaller contributions coming from tyrosine and phenylalanine residues. Conveniently, these emissions which further increase the complexity of a system can be minimised in designed peptides by utilising only a single type of chromophore within the complex.

Considering that the absorption of the $\pi - \pi^*$ transition for tryptophan was demonstrated to occur via two states, ^1L_a and ^1L_b, it is to be expected that they will influence the emission spectrum. The 90º difference in the polarisation axis between these states is the fundamental reason for unequal emission of light along two different polarisation vectors. This phenomenon is known as fluorescence anisotropy and is an essential factor required to understand the photophysics of tryptophan due to the influence of solvent or local environment on the energy of the ^1L_a and ^1L_b states (Fig. 3). In polar solvents, the energy of the ^1L_a state lies below the ^1L_b due to hydrogen bonding with the amine of the indole. This implies two things, firstly energy absorbed to excite to the ^1L_b state will be transferred through a cascade to the ^1L_a state and

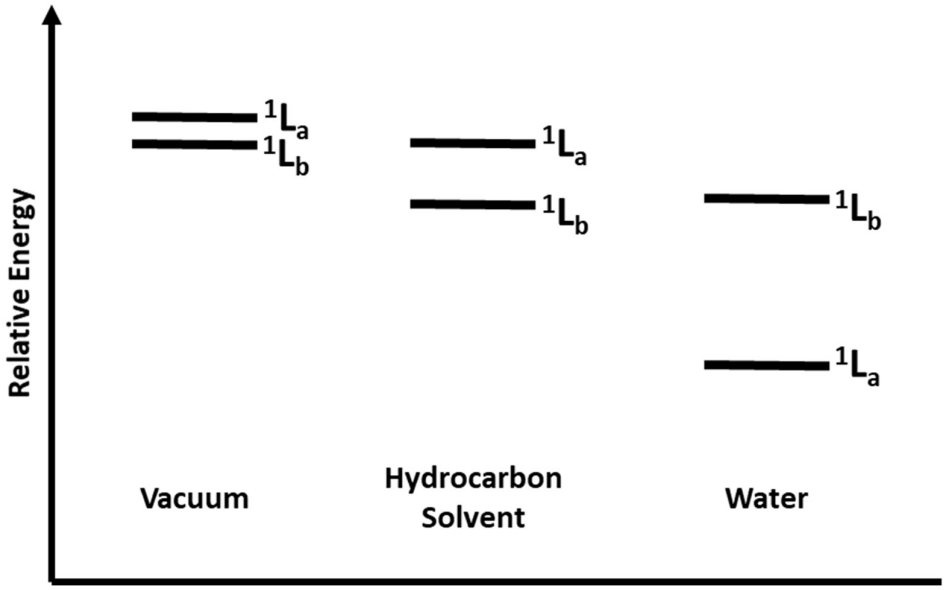

Figure 3: A cartoon representation of the energies associated with the ^1L_a and ^1L_b states of tryptophan in vacuum, hydrocarbon solvent and water. If both states are populated there is a favourable transfer from $^1L_a \rightarrow {}^1L_b$ in hydrocarbon solvents but in water this is reversed.

consequently that emission in polar solvents is dominated by the ^1L_a state. Interestingly, the ^1L_a state is unstructured so we observe broad emission with few features in polar solvents. If the environment becomes apolar, the structured ^1L_b state is lower in energy resulting in the reverse cascade and well resolved fine structure observed in the blue-shifted emission. However, because the ^1L_a state involves interaction with the solvent, it is intrinsically sensitive to the solvent and furthermore the local environment. This interplay between the ^1L_a and ^1L_b states results in tryptophan residues being sensitive to solvent, pH, temperature, protein/solvent induced electric fields, and quenchers where the quenching species could be different chromophores, other residues (e.g. amide backbone), or other tryptophan residues.

Most protein work is carried out in water which is a polar protic solvent and thus interacts strongly with the ^1L_a excited state (Fig. 3). This results in efficient quenching of the tryptophan emission in aqueous environments which was dramatically illustrated by the use of deuterated water to demonstrate the increase in the fluorescence quantum yield of solvent exposed

tryptophan residues.[20] If solvent, temperature and pH are kept constant, the sensitivity of tryptophan can be exploited to elucidate various features about its local environment and solvent exposure.[21] The greater the intensity of the hypsochromic shift of the emission, the more protected the tryptophan residue is from the solvent. This tends to suggest the burying of the tryptophan residue within the hydrophobic core of a protein.

Analysing the emission profiles of both non-metallated and metallated HC01 systems it became apparent that upon metal binding the tryptophan residues went from solvent exposed to partially protected by the peptide bulk (Fig. 4). This was to be expected since upon binding

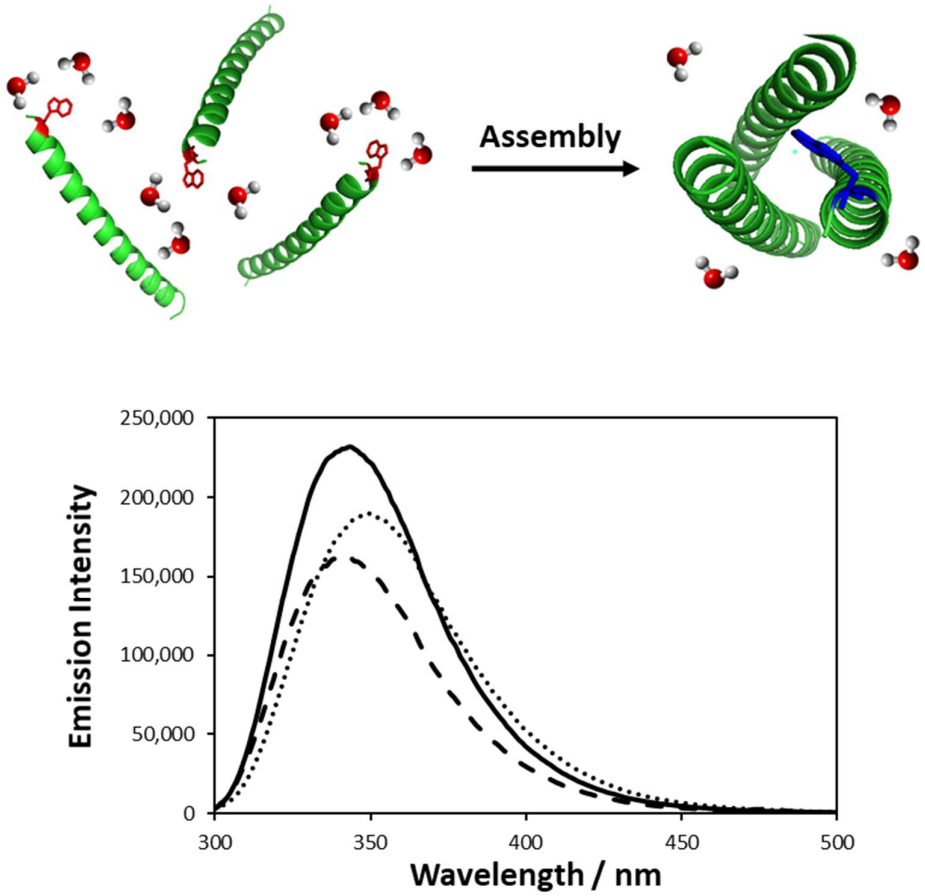

Figure 4: A schematic indicating the solvent exposed nature of the tryptophan residues before assembly leading to the partially burying upon assembly, the disassembled structure has been exaggerated to clarify the concept. The emission profile of non-metallated HC01 (dotted line), metallated HC01 (dashed line) and HC0C (solid line) indicating the hypsochromic shift from non-metallated to metallated HC01. The samples were peptide (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0), the metallated samples included 1.3 equivalences of Tb(III).

of Tb(III) it would induce further coiling at the *N*-terminus thus forcing the ***a*** position tryptophan residues to become more buried. Pleasingly, the tryptophan residues of the control peptide HC0C were found to have the same emission profile as that observed for the metallated HC01. This clearly indicated that the residues were in a similar environment to the metallated coiled coil with the tryptophan residues partially protected from the solvent. These results further supported the structural information previously ascertained in Chapter 2 and confirm the templating effect afforded by the metal ion.

Due to the close proximity of the three tryptophan residues in the HC series there were concerns that this could have resulted in the formation of an exciplex involving either two or all of the indole aromatic systems. The typical feature of an exciplex formation is a large bathochromic shift in the emission spectra resulting in a λ_{max} of ~400 nm.[22] This shift was not observed for either HC0C or metallated HC01 indicating that this interaction was not occurring.

3.4 SOLUTION STATE ENERGY TRANSFER MECHANISMS

After considering the absorption and emission properties of the primary absorber, tryptophan, the mechanisms by which energy can be transported within this system need to be understood. Energy transfer in solution is dominated by three main mechanisms; Dexter Electron Transfer (DET), Förster Resonance Energy Transfer (FRET), and the classic emission and absorption regime. The final of these mechanisms is the result of the excited species emitting a photon followed by subsequent absorption via a second species in solution. The emission orientation is random resulting in very inefficient transfer, except where the species have very high

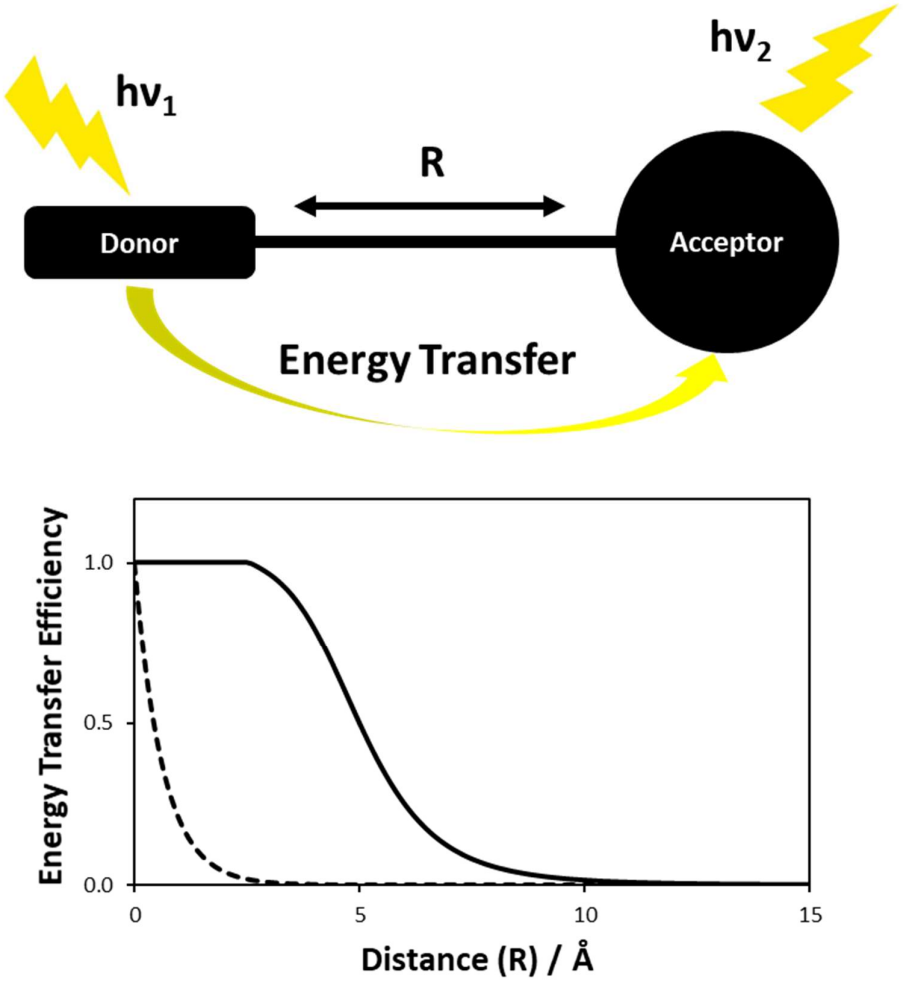

Figure 5: A schematic illustrating the distance (R) between the donor and the acceptor. The plot shows the distance dependence of both DET and FRET efficiencies (solid line) and just the DET process (dashed line). This demonstrates the dominance of DET processes at very close proximity and the importance of FRET at greater separation distances. A FRET constant of 5 Å was applied for this plot.

absorption coefficients or the solutions are highly concentrated. This mechanism can become very important for solid state samples, but will not be discussed here. One must consider both DET and FRET mechanisms as a function of distance (Fig. 5), where the greater the distance between the two species the less efficient the transfer.

DET is a non-radiative quenching mechanism that involves the excited state electron from a donor molecule being transferred to the acceptor molecule while a simultaneous transfer of a

ground state electron from the acceptor molecule to the donor molecule occurs.[23] This process is exceptionally distance sensitive because orbital overlap is required, therefore it only occurs at separation distances of less than ~6.0 Å. This can be understood by considering the rate of electron transfer (k_{DET}) as the separation (R) between the donor and acceptor species increases (Equation 1).

$$k_{DET} \alpha e^{-4R} \qquad \text{(Equation 1)}$$

Whilst DET leads to very high transfer at small separation distances and is utilised extensively in natural light harvesting complexes, it is particularly difficult to build a complex with components in such close proximity. In the previously described HC peptides all transfer distances will be greater than 6.0 Å, therefore this mechanism is unlikely to account for any observed transfer.

FRET is a non-radiative energy transfer process between a donor which absorbs efficiently and an acceptor located within close proximity.[24] The donor excited state relaxes as the acceptor is excited via dipolar coupling. This interaction is coulombic and as a result is sensitive to distance, though not as drastically as DET. One can consider the rate of energy transfer (k_{FRET}) as a function of the distance (R) between the donor and acceptor species (Equation 2).

$$k_{FRET} \alpha R^{-6} \qquad \text{(Equation 2)}$$

A more useful measure of FRET is the efficiency of energy transfer (E_{FRET}) as described by the separation distance (R) and the distance at which 50% of the energy is transferred (R_0)

which can be defined by equation 3. These distances typically range from 2.5 – 50.0 Å and it should be noted that at distances of less than 2.5 Å it is unlikely that FRET is particularly important over DET like processes and related orbital mixing.

$$E_{\mathrm{FRET}} = \frac{1}{1+(\frac{R}{R_0})^6} \qquad \text{(Equation 3)}$$

The interaction between a donor and acceptor can be seen in the graphical depiction of a representative E_{FRET} equation (R_0 is set to 5 Å). There is an inverse 6th power relationship between the donor and acceptor distance and the overall transfer efficiency. The R_0 value is dependent upon the donor and acceptor species and a high R_0 value is desirable as this leads to a more efficient transfer. If the usual limits of zero and infinity are applied to equation 3 it becomes apparent that as R_0 tends to zero the efficiency tends to zero but as R_0 tends to infinity the efficiency tends to 1. Therefore, the ability to predict the value of R_0 is essential for the understanding of FRET mechanics and allows the selection of useful FRET couples.

3.5 PREDICTING FRET EFFICIENCY

The literature distance at which 50% energy transfer from tryptophan to Tb(III) occurs is typically ~3.4 Å, which is too small to be of any use for energy harvesting applications due to less than 10% energy transfer over distances of 1 nm or greater.[13] In order to increase the transfer and thus the utility of this system it is necessary to be able to predict FRET phenomena and its efficiency. Further to this, the use of lanthanides in biological systems has typically been limited to Tb(III) and Eu(III) while there are other visible emitting lanthanides such as

Dy(III) and Sm(III) that may be able to be sensitised by excited tryptophan residues. The energy of the triplet state of tryptophan is ideally situated to sensitise these additional ions (Fig. 6). The required energy to excite Sm(III) is similar to that of Eu(III) and that required for Dy(III) is similar to that of Tb(III), therefore it would be expected that both Dy(III) and Sm(III) could be sensitised. The energy transfer mechanism between tryptophan and terbium has been extensively studied, clearly demonstrating that the transfer event occurs primarily from the triplet state of tryptophan.[25] However, the use of the fluorescence spectrum of tryptophan was shown to be successful for determining the R_0 values.[13] For this reason the rudimentary FRET predictions were performed as a guide to potential energy transfer couples. Interestingly, the success of these calculations indicates that while triplet states are primarily responsible for transfer, the singlet state is also interacting with the lanthanide. Indeed, it has been

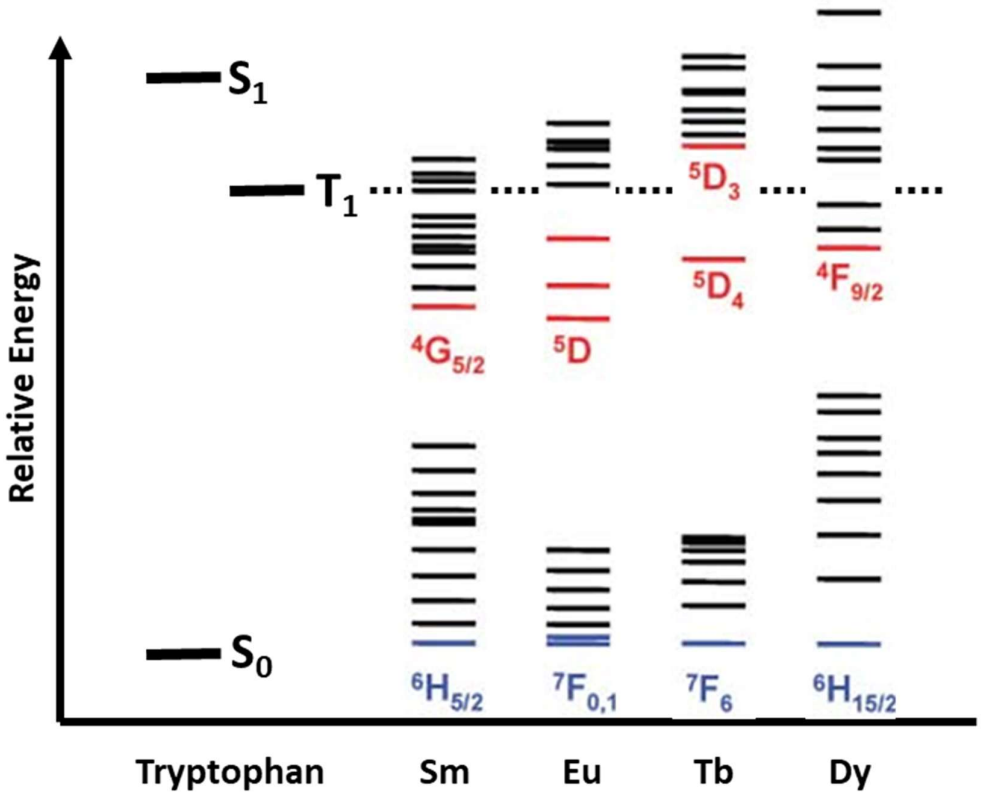

Figure 6: A Jablonski diagram illustrating the relative energy of the excited singlet and triplet states for tryptophan compared with the energy of the visible emitting lanthanides. The dashed line clearly shows the triplet energy is sufficient for excitation of the selected lanthanides.

demonstrated that lanthanides induce intersystem crossing in tryptophan residues.[26] Therefore, it may be that the induced triplet is rapidly quenched by the lanthanide resulting in effective transfer. Using a simplistic approximation of energy transfer, the distance at which 50% of the energy is transferred can be calculated according to equation 4. Care should be taken with this calculation as theoretical determination of FRET efficiency can often appear greater than experimentally observed values. This is due to the spectral overlap integral calculation failing to take into account secondary quenching mechanisms and alterations to chromophore behaviour upon conformational changes. The variables to consider are the fluorescence quantum yield of the donor species (Q_D), the dipole orientation between acceptor and donor (κ), the refractive index of the material (n) and the spectral overlap integral between the emission of the donor and absorption of the acceptor (J).

$$R_0 = (0.2108 \times \frac{Q_D \kappa^2 J}{n^4})^{1/6} \qquad \text{(Equation 4)}$$

This equation can be simplified considerably when some careful assumptions are made. Firstly, the fluorescence quantum yield of tryptophan is known to be between 0.16 and 0.39 dependent upon its environment, this helps set the upper and lower limits of the potential R_0 values.[21, 27] Further to this the refractive index of protein material typically falls between 1.4 and 1.5.[28] As a result both the fluorescence quantum yield and refractive index can be replaced by equations 5a and 5b with upper and lower bounds for these values.

$$\text{Upper: } R_0 = (0.02140 \times \kappa^2 J)^{1/6} \qquad \text{(Equation 5a)}$$

$$\text{Lower: } R_0 = (0.00666 \times \kappa^2 J)^{1/6} \qquad \text{(Equation 5b)}$$

Though the dipole orientation is somewhat more complex to handle (adopting values between 0 – 4), a rough approximation of $^2/_3$ is often used and should be appropriate in these conditions.[29] It should be noted that this is possibly the largest source of error within the calculation. This facilitates a further simplification allowing the window of FRET efficiency to be calculated with relative ease. All the assumptions relate to the environment of the FRET pair which should remain relatively similar with different lanthanides bound. This implies that any deviation in the assumptions will generate a systematic error, which would be both obvious to detect and fail to prevent trend analysis between the experimental and theoretical results.

$$\text{Upper: } R_0 = (0.00951 \times J)^{1/6} \qquad \text{(Equation 6a)}$$

$$\text{Lower: } R_0 = (0.00296 \times J)^{1/6} \qquad \text{(Equation 6b)}$$

Using both equations 6a and 6b, it is possible to calculate the upper and lower bounds of FRET efficiency dependent only upon the spectral overlap integral. The spectral overlap value is the main factor that determines the overall FRET efficiency as it can take values from 0 – 10^{15}, hence any error in the previous assumptions is usually inconsequential. The spectral overlap integral (J) can be described by the integral product of the normalised donor emission spectrum ($f_D(\lambda)$), the molar extinction coefficient of the acceptor at each wavelength ($\varepsilon_A(\lambda)$) and the individual wavelengths raised to a 4th index (λ^4).

$$J = \int f_D(\lambda)\varepsilon_A(\lambda)\lambda^4 . \delta\lambda \qquad \text{(Equation 7)}$$

The FRET efficiency between a donor and acceptor can be manipulated by three variables (Equation 7). Firstly, the distance between the donor and acceptor species (R) is very important

due to the intrinsic distance limitations. However, this would stay relatively constant as different lanthanides are added to the same designed system, apart from a slight alteration due to the difference in ionic radii. This is a useful feature to acknowledge as it could be modified during any peptide redesign. Secondly, the molar absorptivity (ε_A) of the acceptor limits the possible transfer efficiency, but as the consideration here is transfer to a metal ion there is little that can be done to improve the low lanthanide absorption coefficients. Therefore, this factor also remains relatively constant. The final variable is the overlap between the normalised donor emission spectrum and the acceptor absorption spectrum which offers tremendous potential for variability and thus will be considered in further detail.

In order to determine the spectral overlap values a basis set of the desired lanthanides absorption profiles were collected (*see* S32-S33). The trichlorides of Sm(III), Eu(III), Tb(III) and Dy(III) were dissolved in water and the absorption spectra recorded. The aqueous solvent environment implied that the lanthanides would have a fully hydrated first coordination sphere therefore interacting with oxygen donors. While this may not be the best model, the differences in the absorption profiles between the oxygen rich aqueous lanthanides and the oxygen rich designed binding site are likely to be minimal, especially due to the spatially buried nature of the f-orbitals resulting in very small ligand effects.[3] Lanthanide absorption coefficients are very low and this necessitated the use of concentrated solutions (0.1 M) for the absorption measurements. The high concentration resulted in Rayleigh scattering, though this was manually removed by carrying out a power function reduction on the raw data to remove scattering artefacts. Interestingly, the power functions had indices between 2 and 3, suggesting that the lanthanides aggregated into particles, responsible for the scattering, of a size that was close to the range of measured absorption wavelengths ~200 - 600 nm across.[30]

The emission spectra of tryptophan had to be selected carefully due to the solvent sensitive nature of the chromophore. Considering that the energy transfer efficiency measurements were to be taken using a HC derivative it made sense to use the emission profile of a control species that exhibited the same emission profile as the HC peptides but without the quenching effects of a metal present. The HC0C peptide system represented a clear solution, as the emission spectrum has already been demonstrated to be remarkably similar to that of metallated HC01 (Fig. 4). This removed any potential quenching interactions to bound metal ions in HC01 that may have influenced the emission profile and thus the spectral overlap calculations.

After overlaying the emission spectrum of HC0C with the absorption spectra of Sm(III), Eu(III), Tb(III) and Dy(III) and calculating the spectral overlap integral, it remained to determine the R_0 values using equations 6a and 6b (*see* Table 1). Pleasingly the predicted values for both Tb(III) and Eu(III) are in good agreement with the literature suggesting that the calculations were valid.[13] The resultant R_0 values clearly suggested that both Dy(III) and Sm(III) should be successfully sensitised by a locally situated tryptophan residue.

To support these predictions further, reports of Kirk *et al.* utilising a small molecule lanthanide complex showed energy transfer from an indole to Dy(III) and Sm(III).[31] Previous work with tryptophan observed that chemical oxidation using peroxides could sensitise Tb(III), Eu(III), Dy(III) and Sm(III) and that photoacoustic excitation of tryptophan co-crystallised with lanthanides resulted in measurable emission.[32, 33] Despite all this, there appears to be no reported use of Dy(III) and Sm(III) sensitised emission from tryptophan in biological systems. This may be due to amide backbone quenching as well as the distance dependence which would have made this phenomenon very hard to observe. It should be noted that while the R_0 values

indicate that the most efficient transfer of energy should be to Sm(III) and Dy(III), this does not provide information about the resultant emission intensity as various deexcitation pathways may reduce the yield of photons emitted from the differing lanthanide species.

Table 1: The predicted R_0 values for energy transfer between tryptophan and the respective lanthanides compared with the literature values for terbium and europium.[13] Errors in R_0^{theory} are from upper and lower limits of the calculation.

Lanthanide	R_0^{theory} / Å	$R_0^{literature}$ / Å
Sm(III)	4.77 ± 0.46	-
Eu(III)	4.27 ± 0.42	3.75 ± 0.05
Tb(III)	3.78 ± 0.37	3.41 ± 0.17
Dy(III)	4.71 ± 0.46	-

3.6 PHOTOPHYSICAL STUDY OF THE HC SERIES

In order to observe the predicted sensitisation from tryptophan to the various lanthanides of interest it was necessary to determine the system with the greatest emission intensity from Tb(III). This was carried out by irradiating each of four peptides, HC01 - HC04, with 280 nm light to selectively excite the tryptophan residues. The excited tryptophan residues would then sensitise any locally situated Tb(III) ions. The excitation was not of the Tb(III) ions directly, therefore the observed emission would be only from sensitised species. Titration with aliquots of terbium chloride was carried out followed by recording the emission spectra from 450 - 700 nm (Fig. 7). The assumption was that as the binding site was moved further from the tryptophan residues, the Tb(III) emission would decrease resulting in HC01 being most efficient and HC04

the least efficient. However, the experimental observations (Fig. 7) were more complex to interpret. It should be noted that previous discussion regarding speciation (*see* Chapter 2) would suggest that with 90 μM peptide at the 1 equivalence of Tb(III) ions to trimer ~85% would be bound, therefore it would be expected that 15% would be free in solution.

HC01 behaved as expected with the addition of Tb(III) resulting in a plateau at one equivalence of metal ions to three equivalence of peptide monomer, further supporting the mass spectrometry and circular dichroism evidence for the formation of a trimer species in solution. The binding isotherm was fitted to a three-to-one model (*see* Chapter 8.9) resulting in a determined log of the binding constant ($\log K_a$) of 5.25 ± 0.22. This was in agreement with previously reported data for this binding site in the MB1-1 system with a $\log K_a$ of 5.30 ± 0.15.[16] When comparing the emission from bound Tb(III) in both the HC01 and MB1-1 systems the residual tryptophan emission from HC01 is much less apparent than for MB1-1. This suggests that the HC01 bound Tb(III) emission is greater than that of the MB1-1 system or that the tryptophan emission has been quenched more efficiently in the HC01 system.

Interestingly both the HC02 and HC03 systems demonstrated significantly reduced emission intensity up to one equivalence of metal ion to peptide trimer, this was followed by a dramatic increase in terbium emission attributed to non-specific binding close to the tryptophan residues. The initial lack of emission was reasoned to be the distance from the designed binding site to the tryptophan resulting in no energy transfer. For both HC02 and HC03 this distance is considerably greater than the literature R_0 values at ~18 Å and ~28 Å respectively, measured from the crystal structures (*see* Chapter 2). The lack of sensitised Tb(III) emission up to one equivalent implies that the strongest binding site in the system is too far away to be efficiently

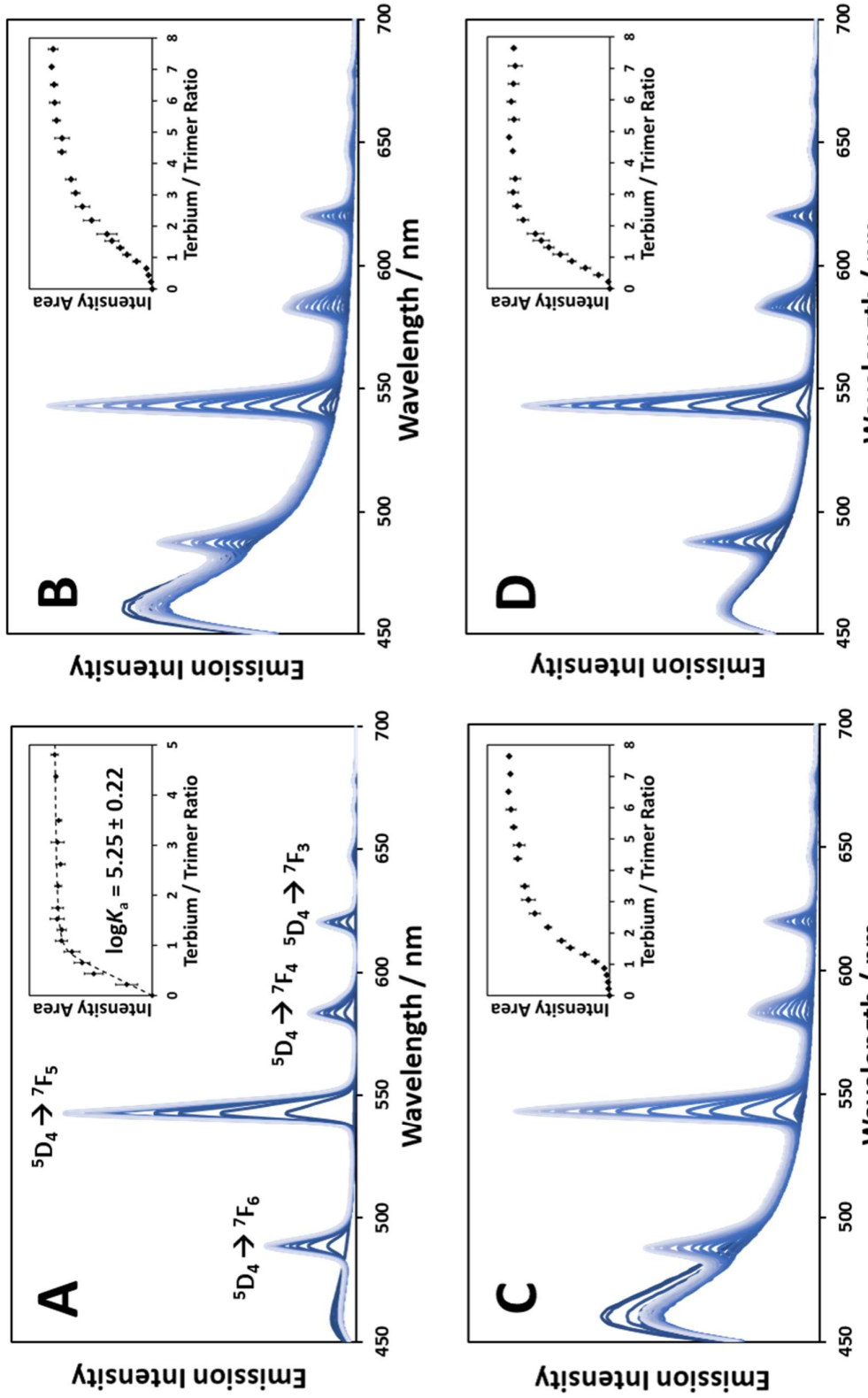

Figure 7: The emission profiles of Tb(III) as it is titrated into A) HC01, B) HC02, C) HC03, and D) HC04 systems, excitation wavelength was 280 nm. The inserts show the increase in intensity area as Tb(III) is titrated into the respective systems, the reported error is the standard deviation of three repeats. The intensity area is the area under the peak at 545 nm. The samples were peptide (90 μM) in HEPES buffer (10 mM, pH 7.0).

sensitised, such as the designed site in the second or third heptads respectively. The subsequent dramatic increase in emission is consistent with Tb(III) binding to the external sites at the top and bottom of the HC series, made up of glutamate residues, which can coordinate metals as evident from the crystallographic structures (*see* Chapter 2). However, this was unexpected as it was not anticipated that the weak interactions between coiled coils in the crystal lattice would be retained in solution. This coordination was thought to be the reason that Tb(III) ions could be situated close enough to the tryptophan residues to result in observable emission. Elucidating specific distance information from this data would be particularly challenging as these external sites are all likely to be exceptionally dynamic. Nevertheless, the initial low emission indicates that the external sites must have weaker binding than the designed core sites as exhibited by the binding preference for the core binding site. This has been previously shown by circular dichroism titrations, with only a single equivalent of metal ion required to induce folding, having $\log K_a$ of 5.11 ± 0.06 and 5.06 ± 0.10 for HC02 and HC03 respectively. It should be noted that the mass spectrometry of the complexes of HC02 and HC03 (*see* Chapter 2) also only observed a trimer with a single metal ion bound. Therefore, it would be anticipated that the binding constants for the external sites would have $\log K_a < 5$.

The behaviour of HC04 is notably different to that observed for HC01 - HC03 as demonstrated by the rapid increase in Tb(III) emission, suggesting that no core binding was occurring. While this was not in agreement with the crystallographically determined structure of HC04 (*see* Chapter 2), it is important to note that the solid state structure does not always represent what is present in solution. Indeed, the lack of a core binding site corresponds very favourably with the previously reported mass spectrometry and circular dichroism data. Therefore, it was proposed that the designed binding site of HC04 was not intact, resulting in the emission being

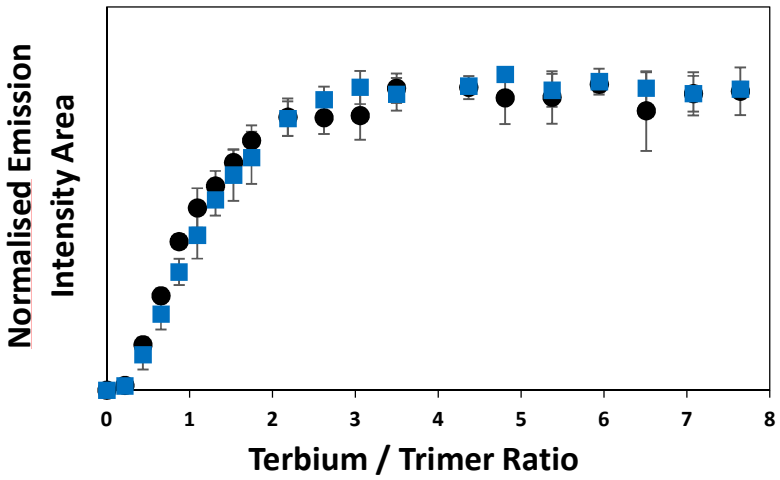

Figure 8: The emission from bound Tb(III) ions in both HC04 (blue squares) and HC0C (black circles). The samples were peptide (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0) and reported errors are the standard deviation of three repeats of the isotherm.

due entirely to external binding. This would imply that the control peptide HC0C would exhibit the same behaviour by binding Tb(III) to non-specific external sites. Upon titrating Tb(III) with the HC0C peptide the emission profile was found to behave almost identically to the HC04 peptide (Fig. 8). This clearly demonstrated the lack of a core binding site in the HC04 system, indicating that the previously observed parallel crystal structure, with a well-defined internal binding site, may be disrupted in solution. Therefore it was postulated that HC04 may exist in an antiparallel conformation in solution. This was further confirmed by considering the overall emission intensity of the peptides (*see* Table 2). If the external sites had all been the same it would be expected that the final plateau for the intensity of the emission for all, except HC01, would be the same. This is because external sites are closer to the tryptophan than the designed sites, therefore resulting in the dominant Tb(III) emission. However, both HC0C and HC04 exhibit dramatically increased intensity, which suggests that the external site environments are different. An antiparallel arrangement would result in more sites close to tryptophan residues due to their presence at both the top and bottom of the coiled coil. This also suggested that the control peptide, HC0C, may be antiparallel with similar binding behaviour to HC04.

Table 2: Final emission intensity area of the 545 nm peak for HC02, HC03, HC04 and HC0C. The ratio of the peak area was determined against HC04 to demonstrate the difference between the peptide systems. The samples were peptide (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0).

	HC02	**HC03**	**HC04**	**HC0C**
Final Emission Area:	354316	396514	524166	512497
Ratio:	0.68	0.76	1.00	0.98

It is clear from this study that the HC01 system is most suited to probe the interaction of other lanthanides with tryptophan residues. This is because firstly the binding site is intact and secondly because the high efficiency of energy transfer allows easy observation of weak signals. Unfortunately, it has also became apparent that HC0C is not a valid control for the HC01 system due to its apparent antiparallel nature resulting in the tryptophan residues being in subtly different environments.

3.7 BINDING SITE HYDRATION STATE

An essential part of increasing the efficiency of a lanthanide system is to remove sources of quenching. Water is well known to be a quencher for lanthanide excited states via O-H bond vibrations.[34] Utilising the secondary sphere ligand control afforded by larger structures such as peptides it is possible to limit the number of water molecules capable of accessing the bound metal ion.[16] The hydration state or number of water molecules around the metal can be determined by measuring certain properties of the ion itself.

One method to determine the shielded nature of the different binding environments across the HC series would be to monitor the luminescence lifetime of the terbium emission in each site

compared with the naked ion in aqueous solution. If the lifetime is longer than the aqueous species then the bound ion can be regarded as more shielded than the totally solvated species. However, this technique can give a false sense of hydration if other quenchers are present, for example other chromophores in the system. This is because the excited lanthanide can be rapidly quenched by certain chromophores resulting in a very short emission lifetimes, the expected result of a Tb(III) in a highly hydrated state. Fortunately, the HC peptides do not have a chromophore that would result in this problem therefore a direct analysis can be undertaken.

The emission lifetimes (τ^H) of the Tb(III) ions were recorded for each peptide at a third equivalence of metal ions to trimer in order to limit artefacts from non-specific binding (*see* Table 3). This is particularly important as the metal ion concentration is low enough that the favoured binding site will be occupied with little residency in other sites. Indeed, the speciation at such a low concentration of Tb(III), even with a binding constant of 10^5, leads to over 99% of the metal bound. Therefore, when considering the previous luminescence titrations with Tb(III), the ions for HC02 and HC03 are not very emissive due to very poor sensitisation. The excitation wavelength was 280 nm to selectively excite the tryptophan residues in the designed complexes. However, it is important to note that this does not favour the Tb(III) bound in the designed binding site for every system as both HC02 and HC03 are at a distance too great for efficient sensitisation to occur. It becomes immediately apparent that HC01-HC03 all have much longer lifetimes of ~2.0 ms compared to the aqueous Tb(III) at 0.4 ms. This suggests that the binding sites of HC01-HC03 are very solvent protected resulting in the observed long lived emission. In contrast, both HC04 and HC0C emission lifetimes can be fitted biexponentially, indicating two components responsible for the emission. The shortest of these is the aqueous Tb(III) and should be considered as unbound ions. The second components are significantly

Table 3: The luminescence lifetime of the Tb(III) emission at 545 nm for each of the HC peptides in both water (τ^H) and deuterium oxide (τ^D). The number of water molecules in the first coordination sphere of the metal ion as determined from the lifetimes. The samples were peptide (90 μM), HEPES buffer (10 mM, pH 7.0) and $TbCl_3$ (10 μM) in either water or deuterium oxide. Reported errors are the standard deviation of three repeats.

Peptide	τ^H_1 / ms	τ^H_2 / ms	τ^D_1 / ms	No. of Water Molecules
HC01	1.95 ± 0.01	-	2.27 ± 0.13	**0**
HC02	2.00 ± 0.02	-	2.46 ± 0.03	**0**
HC03	1.85 ± 0.05	-	2.25 ± 0.05	**0**
HC04	1.39 ± 0.04	0.62 ± 0.01	2.18 ± 0.04	**1, 5**
HC0C	1.23 ± 0.03	0.44 ± 0.01	2.87 ± 0.03	**2, 9**
Aqueous	0.43 ± 0.01	-	2.26 ± 0.07	**9**

longer than aqueous Tb(III) ions but not as long as that observed for HC01 - HC03. This behaviour would be expected of external sites that would have somewhat protected Tb(III) ions because of the coordinated ligands, but would still have directly coordinated water molecules that can quench the hydrated, but not fully solvated ion.

The lifetime of Tb(III) emission gives an idea of the hydration state, though it cannot give an exact value for the number of water molecules coordinated. Work by Beeby and Horrocks *et al*. elucidated an equation that related the decreased efficiency of O-D vibrations to quench terbium emission in deuterated water to the number of water molecules in the first coordination sphere of Tb(III).[34, 35] Using equation 8 and the experimentally determined lifetimes of the Tb(III) emission in both H_2O (τ^H) and D_2O (τ^D) it was possible to elucidate approximately how many water molecules were around the metal ion in each binding site (*see* Table 3).

$$\text{No. of Water Molecules} = \frac{5}{\tau^H} - \frac{5}{\tau^D} - 0.3 \qquad \text{(Equation 8)}$$

As expected, HC01 - HC03 all exhibit zero water molecules in the first coordination sphere indicating that the site is solvent protected and allows the potential for efficient use of excited state lanthanides. Indeed, considering the speciation of the complex and that the preferential site would be poorly sensitised for both HC02 and HC03, this gives strong evidence that the initial poor emission observed in the luminescence titration was due to the presence of the metal in the designed core binding site. Regarding HC04 and HC0C, both have water molecules around the Tb(III) ion indicating that the sites are external and almost ligand saturated, but that some water molecules can still access the ions. Further to this, the presence of more than just totally solvated ions in the HC0C control system clearly proves the presence of other binding sites within the designed sequence. Interestingly, HC01 has zero waters whereas the previous top site peptide MB1-1 had three water molecules coordinated.[16] A possible explanation for this difference is that the tryptophan residues at the top of the HC01 peptide could effectively close the top and reduce solvent access to the metal ion.

3.8 DYSPROSIUM AND SAMARIUM SENSITISED LUMINESCENCE

The peptide system that exhibited the highest Tb(III) emission and retained a solvent protected binding site was HC01. Therefore this system offered the best potential to observe even weak transfer and was chosen to validate the previous tryptophan to Dy(III) and Sm(III) transfer predictions. Characterisation with Eu(III) was carried out for completeness. Upon the addition of Dy(III), Sm(III) or Eu(III), followed by excitation of the tryptophan (λ = 280 nm) it was clear that all three systems resulted in successful energy transfer (Fig. 8). Dy(III) was particularly exciting with the three bands at 478 nm, 571 nm and 660 nm clearly visible. This

is the first reported instance of Dy(III) being sensitised in a peptide or protein system. While the emission was observed for both bound Sm(III) and Eu(III) ions, when exciting the sample at 280 nm, it was of particularly weak intensity. The Eu(III) spectrum has characteristic emission with peaks at 592 nm, 616 nm and 695 nm (Fig. 8B). The hypersensitive $^5D_0 \rightarrow {}^7F_2$ transition at 616 nm is particularly intense, resulting in an overall red emission, indicating that the Eu(III) is bound in a low symmetry environment.[36] This is in agreement with the discussed structures of the binding sites that exhibited low symmetry. The Sm(III) spectrum had weak but characteristic peaks at 562 nm, 600 nm, and 642 nm (Fig. 8D). As further confirmation of these emission results, the luminescence spectra of a literature species, the lanthanide binding tag (LBT), with Dy(III) and Sm(III) were recorded (*see* S34).[14] This peptide system demonstrated similar results indicating that this phenomenon was not just limited to HC01, but was intrinsically related to the tryptophan residue placement. In addition to this, the previously reported MB1-2 lanthanide binding peptide was also used to further validate this energy transfer phenomenon.[15] While the Dy(III) emission was observed clearly, the Sm(III) emission was too weak to be observed (*see* S35). This was most likely because the distance between the tryptophan residues and lanthanide ion in MB1-2 at ~11 Å is considerably greater than for the HC01 or LBT systems at ~8 Å and 6 Å respectively. The raw emission from HC01, especially when compared with the residual tryptophan emission, demonstrated a qualitative improvement in the sensitisation of Tb(III) by HC01 over that previously observed for the MB1-2 system.[15]

Considering the luminescence curves for HC01 it is apparent that all four lanthanides plateau at one equivalence of lanthanide to three equivalence of peptide (Fig. 9). As with Tb(III), this indicates the successful formation of the coiled coil trimer upon binding of the metal ions. The

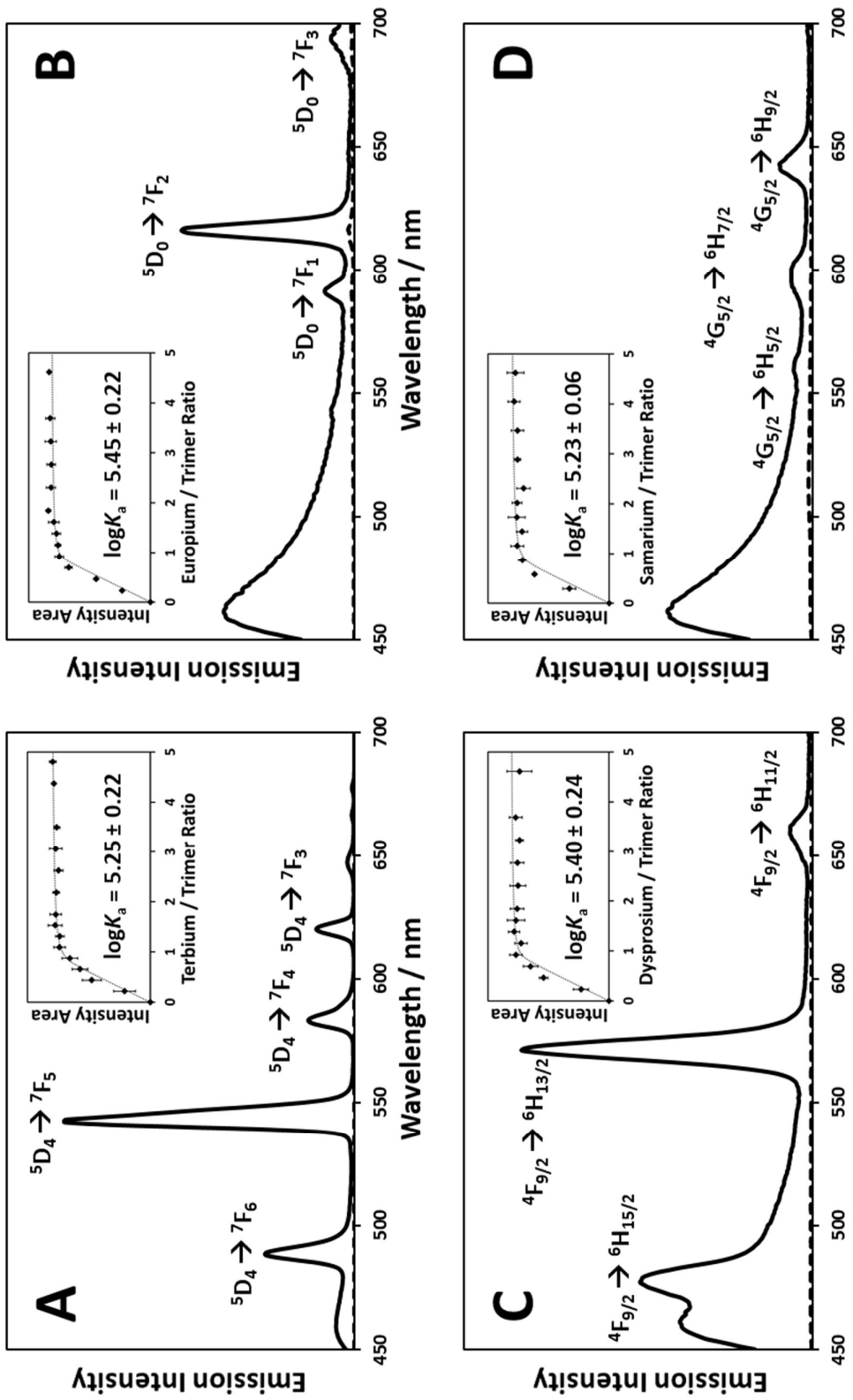

Figure 9: The emission profiles of HC01 (solid line) and aqueous controls (dashed line) with A) Tb(III), B) Eu(III), C) Dy(III) and D) Sm(III) after excitation with 280 nm light. The inserts correspond to the change in emission intensity area at 545 nm for Tb(III), 616 nm for Eu(III), 571 nm for Dy(III) and 643 nm for Sm(III) upon the addition of the respective lanthanides. The log of the association constant was determined by fitting the data to a three-to-one binding equation. Samples were peptide (90 μM), HEPES buffer (10 mM, pH 7.0) and the respective metal ions, errors are the standard deviation of three repeats.

binding constants for all four lanthanides were similar indicating no observable selectivity. This is expected as lanthanide binding selectivity is generally achieved via size differences but Tb(III), Eu(III), Dy(III) and Sm(III) are all similarly sized species.[37] The ability to interchangeably bind lanthanides is particularly useful for this model system, allowing simple comparisons to be made between different lanthanide ions.

3.9 LUMINESCENCE LIFETIMES OF SAMARIUM, DYSPROSIUM AND EUROPIUM

Following the successful emission from all four lanthanides, the lifetimes of the emissions, when the samples were excited with 280 nm light, were analyzed and fitted to exponential components (*see* Table 4). Each lifetime was considerably longer than the free aqueous ion which, as before, can be explained by the Ln(III) excited state being protected from solvent quenching resulting in long-lived emission.[38] This is further evidence of the Ln(III) ion being bound inside the peptide core, therefore the designed binding site is intact.

Table 4: The emission lifetimes for the dominant peak of the HC01 bound lanthanides, 545 nm for Tb(III), 616 nm for Eu(III), 571 nm for Dy(III) and 643 nm for Sm(III). The aqueous lanthanide emission lifetimes are provided for comparison.[38] Samples were peptide (90 µM) in aqueous HEPES buffer (10 mM, pH 7.0) and the respective lanthanide trications (10 µM). The reported errors are the standard deviation of three repeats.

Lanthanide	HC01 Bound Lifetime / µs	Aqueous Lifetime / µs
Samarium	25.46 ± 0.02	2.69 ± 0.01
Europium	1100 ± 4	112.4 ± 0.7
Terbium	1950 ± 10	442 ± 2
Dysprosium	27.06 ± 0.02	2.49 ± 0.02

The most exciting aspect of the lanthanide emission lifetimes is the clear potential for time-gated experiments.[39] The exceptionally long lifetimes of the Tb(III) and Eu(III) ions at 1950 ± 10 μs and 1100 ± 4 μs respectively allows the collection of spectral data after 200-500 μs to selectively observe just their signals. This has proven exceptionally useful especially due to their relatively weak emissions. With the intermediate lifetimes of the Dy(III) and Sm(III) ions at 27.06 μs and 25.46 μs respectively it is proposed that a further channel could be observed by time-gating at 10 μs and 200-500 μs, to observe first the Dy(III) or Sm(III) species followed by the Tb(III) and Eu(III) species. This is of particular interest for biological assays as it would add an extra feature or redundant dataset that can be collected and monitored within a single experiment, further increasing either the validity or scope of the results. Indeed, considering the relatively high intensity of the Dy(III) emission it could easily be paired with a Tb(III) system for simple multichannel experiments.[40] Further to this, the narrow emission bands of these lanthanides reduces the problem of spectral overlap, reducing erroneous results.

3.10 ELECTRON TRANSFER

During this chapter, the primary consideration has been the process of energy transfer via resonance based mechanisms. However, these mechanisms are effected by a number of limitations, especially when considering energy loss via multistep cascades and the steady red-shifting to lower energy with each transfer. On occasion this relentless loss of captured energy can be countered with additive processes such as multiphoton absorption and up-conversion. Though it should be noted that these are pseudo-solutions as energetically they still result in loss. Therefore, more stable energy storage and transport solutions are required.

Nature makes use of substrates that are reduced by electron transfer as a means to store and transport higher energy species, such as in the use of quinones.[41] Electron transfer from donor species to lanthanides has been recorded resulting in reduction from Ln(III) to Ln(II) species.[42] This potential to separate charges between the bound lanthanide ion and the tryptophan indole was of exceptional interest as this opened up the potential for electron transfer and high energy transport mechanisms. This section will explore the potential of the natural amino acid tryptophan to engage in photoinduced electron transfer with various lanthanides.

Early observation of a proposed electron transfer process between tryptophan and Yb(III) was made by Horrocks *et al*.[43] Still, this was a deduction due to the lack of any potential spectral overlap. Thus, there was no direct observation to support the charge transfer mechanism. In order to predict the potential for charge transfer from tryptophan to any of the lanthanides it was necessary to carry out some thermodynamic calculations. By considering the energetic potential of the excited state tryptophan (E_{Trp*}), the overall energy available to the system can be considered 3.9 eV.[43] The Gibbs free energy of reduction for each Ln(III) to Ln(II),[3] and of oxidation of tryptophan ($E_{(Trp^{+\cdot}/Trp)}$) at 1.13 eV,[44] can be used to determine the feasibility of the electron transfer process (ΔG_{Ln}) from equation 11.[45]

$$\Delta G_{Ln} = E_{(Trp^{+\cdot}/Trp)} - E_{Trp*} - E_{Ln(III)/Ln(II)} \qquad \text{(Equation 11)}$$

The results of this calculation gave the feasibility of the charge transfer process for each of the lanthanide trications (Fig. 10). This clearly indicates that only Eu(III), Yb(III), Sm(III) and Tm(III) are expected to engage in charge transfer with tryptophan with feasibilities of -2.25

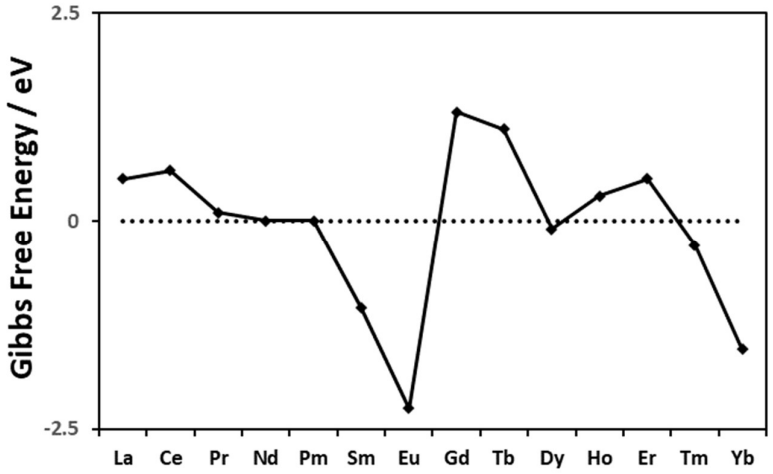

Figure 10: The predicted Gibbs free energy of electron transfer from an excited state tryptophan residue to a lanthanide trication.

eV, -1.54 eV, -1.04 eV and -0.29 eV respectively. It is interesting to note that both Eu(III) and Yb(III) are known to undergo this process, however, for Sm(III) and Tm(III) this interaction has not been previously reported. Therefore, to probe this mechanism the quenching of tryptophan emission, excited at 280 nm, in HC01 was observed upon the addition of these lanthanides to demonstrate that the reduction in emission was much greater than expected from the previously discussed resonance transfer. The effect of Dy(III) was also analysed as this value was just about feasible (ΔG_{Dy} = -0.09 eV) from the prediction, thus, it was likely that the rudimentary nature of the prediction implied that this would not undergo charge transfer.

All four of the studied lanthanides, Eu(III), Yb(III), Sm(III) and Tm(III), resulted in appreciable quenching of the tryptophan fluorescence (Fig. 11), far greater than could be explained by the previous resonance mechanisms. Considering the predictions from equation 11, it would suggest that electron transfer is indeed occurring between these lanthanides and the localised tryptophan residues, resulting in observable quenching. The change was also consistent with a single metal ion binding to a trimeric peptide unit, as expected for HC01. Pleasingly, the impact of Dy(III) upon the fluorescence of the tryptophan was unnoticeable

(Fig. 12), with no measureable change occurring across the studied range. This is the expected result of a system with low efficiency resonance energy transfer rather than electron transfer. The lack of quenching for tryptophan in the presence of Dy(III) suggests that no charge transfer is occurring. This is particularly useful as it suggests that the predictions were set at the right energy as none of the other systems had a feasibility greater than Dy(III). This implies that Eu(III), Yb(III), Sm(III) and Tm(III) are the only lanthanide trications capable of accepting an electron from the tryptophan excited state. It should be noted, however, that the observed quenching does not prove that the mechanism is one of charge transfer but implies it as was previously reported by Horrocks *et al.*[43]

The most dramatic quenching is observed for the Eu(III) titration resulting with 78% of the tryptophan emission being quenched. It should be noted that this consideration of the three tryptophan residues within the trimer as identical may be too simple as the 78% quenching may represent two proximal residues being quenched 100%. Both Yb(III) and Sm(III) were able to quench 59% and 44% of the tryptophan emission respectively. As before, this could indicate a percentage of quenching across the three localised tryptophan residues or the complete quenching of the closest one or two residues respectively. The quenching afforded by Tm(III) was particularly weak at 13%, but this was anticipated as the feasibility was predicted to be possible but the rate of transfer would be expected to be low. The tryptophan quenching was also used to monitor the binding event, creating a set of isotherms for each of the four lanthanides (Fig. 13). Each of the binding isotherms were fitted to a three-to-one model (*see* Chapter 8.9) resulting in the determination of their binding constants. The $\log K_a$ values for Eu(III), Yb(III), Sm(III) and Tm(III) were found to be 5.16 ± 0.16, 5.18 ± 0.13, 5.24 ± 0.15, and 5.02 ± 0.08 respectively. In agreement with the previous measurements both from circular

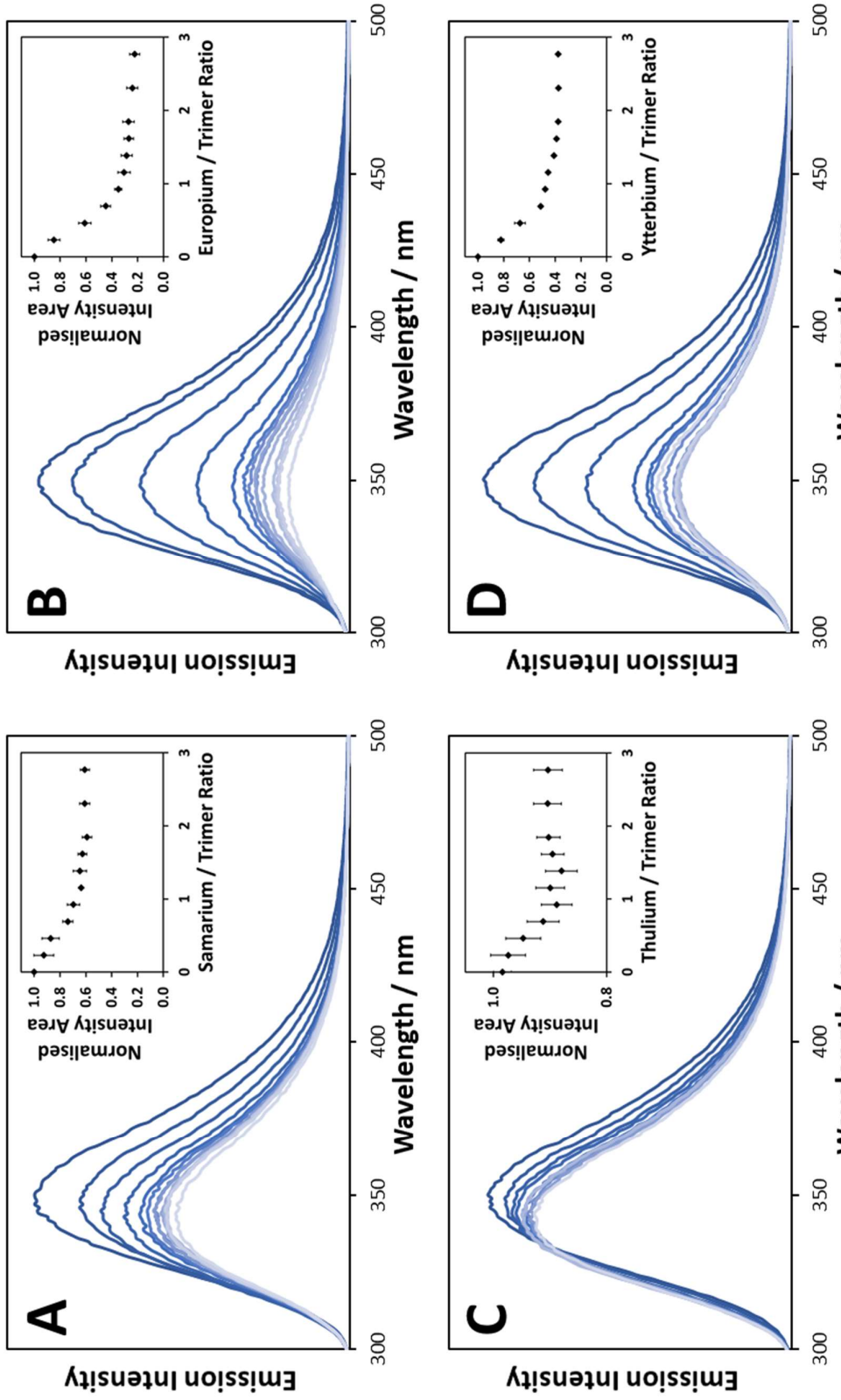

Figure 11: The resultant fluorescence quenching of the tryptophan residues in HC01 by A) Sm(III), B) Eu(III), C) Tm(III) and D) Yb(III). Excitation wavelength was 280 nm. The inserts show the decrease in normalised intensity area for the emission peak upon the addition of aliquots of the respective lanthanides. The samples consisted of peptide (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0), the reported errors are from the standard deviation of three repeats.

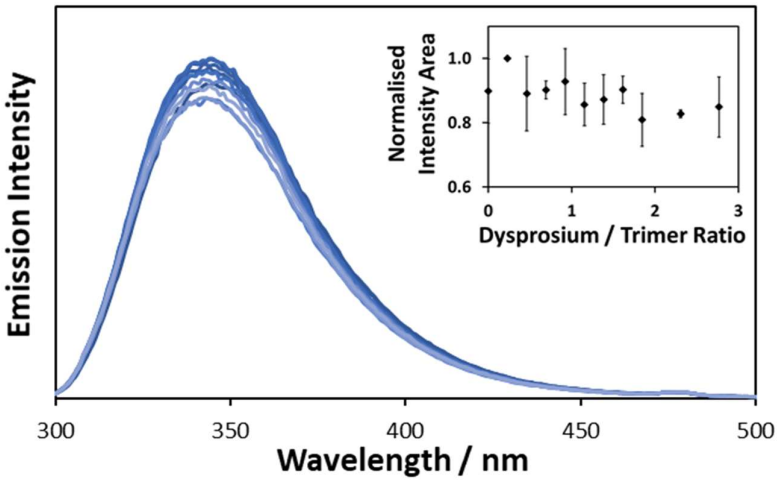

Figure 12: The resultant fluorescence quenching of the tryptophan residues in HC01 by Dy(III). The insert shows the decrease in normalised intensity area for the emission peak upon the addition of aliquots of the Dy(III) ions. The samples consisted of peptide (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0), the reported errors are from the standard deviation of three repeats.

dichroism (*see* Chapter 2) and luminescence. This was particularly interesting as the considerably smaller Yb(III) ion seemed to bind with a similar affinity to the other studied lanthanides. Indeed, this lack of selectivity suggested that the binding site had a significant degree of flexibility to accommodate a range of lanthanide ions with similar binding energies.

3.11 INFRARED EMISSION

Following the tryptophan quenching titrations it was clearly apparent that resonance energy transfer could not explain the efficiency of the quenching mechanism. The work of Horrocks *et al.* had observed the weak infrared emission from Yb(III) when located in close proximity to an excited tryptophan residue.[43] This observation had been attributed to an electron transfer mechanism resulting in a Yb(II) species that undergoes back transfer to the tryptophan regenerating the Yb(III) species, but in an excited state. The relaxation from this excited state Yb(III) was thought to be responsible for the observed broad emission at 980 nm. Therefore if

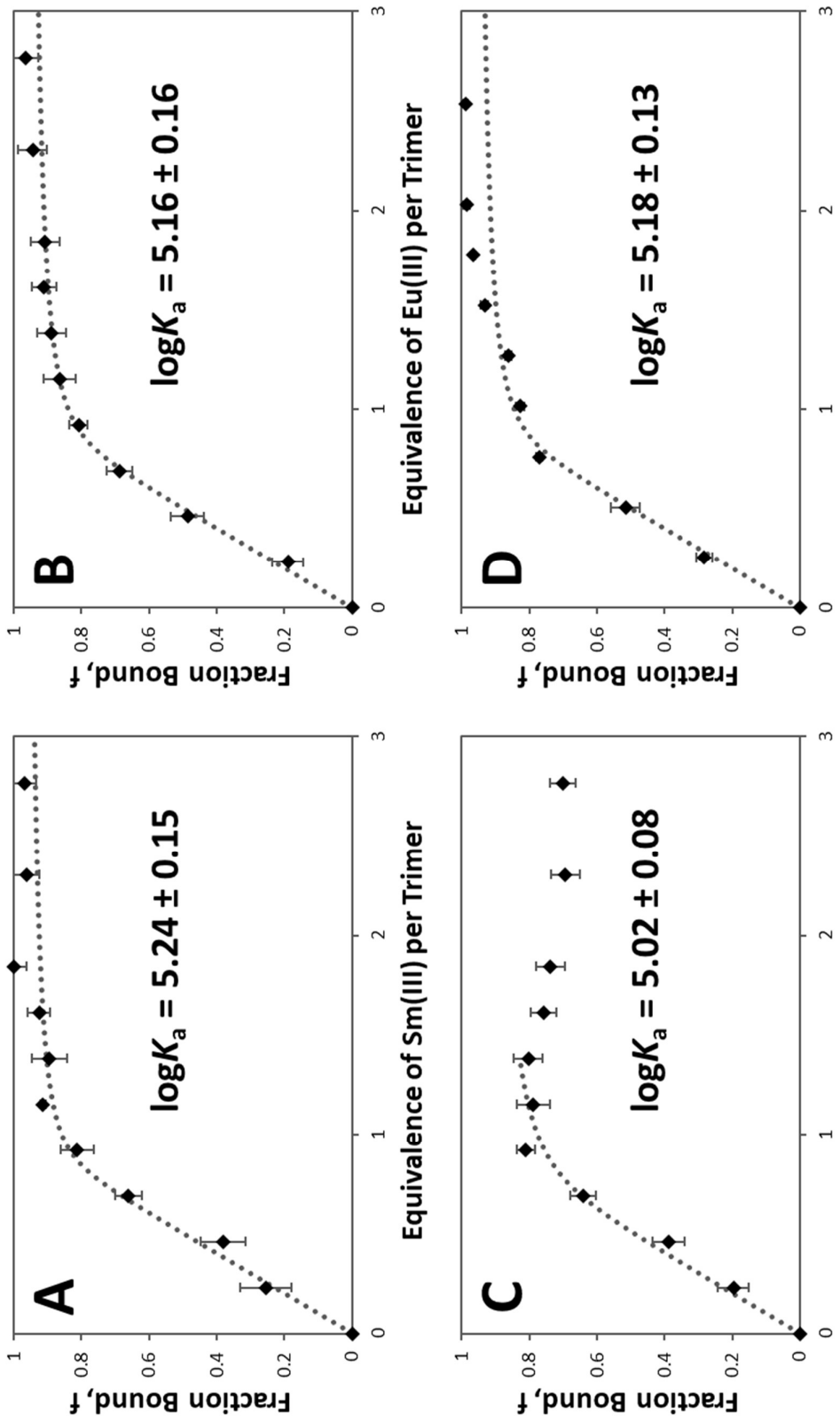

Figure 13: The binding isotherms for A) Sm(III), B) Eu(III), C) Tm(III) and D) Yb(III) in HC01 as determined from the tryptopham quenching profiles. The isotherms were fitted to a three-to-one model and the binding constants are given with an error due to the fitting of the upper and lower limits of the isotherm data. For the Tm(III) data the first points were given a greater weight value due to the aberrant behaviour of the plateau.

a similar process were occurring in the HC01 system it would be expected that Yb(III) emission could be observed. Pleasingly, upon exciting the tryptophan residues at 280 nm the near-infrared emission was observed at 980 nm with exceptional clarity compared to that previously reported (Fig. 14A). This is most likely due to the presence of three localised tryptophan residues compared to the single chromophore in the previously studied system. The overall emission is relatively featureless and the excitation profile (Fig. 14B) clearly illustrates that the tryptophan is absorbing the energy and transferring it to the Yb(III) for emission in the near-infrared region. While electron transfer could lead to an excited state Yb(III) ion which then emits, this is not the case for the other lanthanides. This short study is in agreement with the previous work carried out on proteins with lanthanides bound, however, it also does not directly observe the electron transfer event. This could be probed by utilising fast techniques such as photoelectron spectroscopy to monitor the change in binding energy of the tryptophan after excitation.

3.12 CONCLUSIONS

Throughout this chapter various photophysical aspects of the HC series have been studied. Simple analysis of the tryptophan emission profile for HC0C, metallated and non-metallated HC01 indicates that upon metal binding the tryptophan residues adopt a partially solvent protected conformation. The titration of Tb(III) into each of the HC peptides confirmed that HC01 - HC03 behaved as intended, albeit with some additional external binding. The external binding sites observed in the luminescence titration have been reasoned to be due to the glutamate rich *N*- and *C*-termini of the HC peptides. Indeed, these glutamate rich regions were

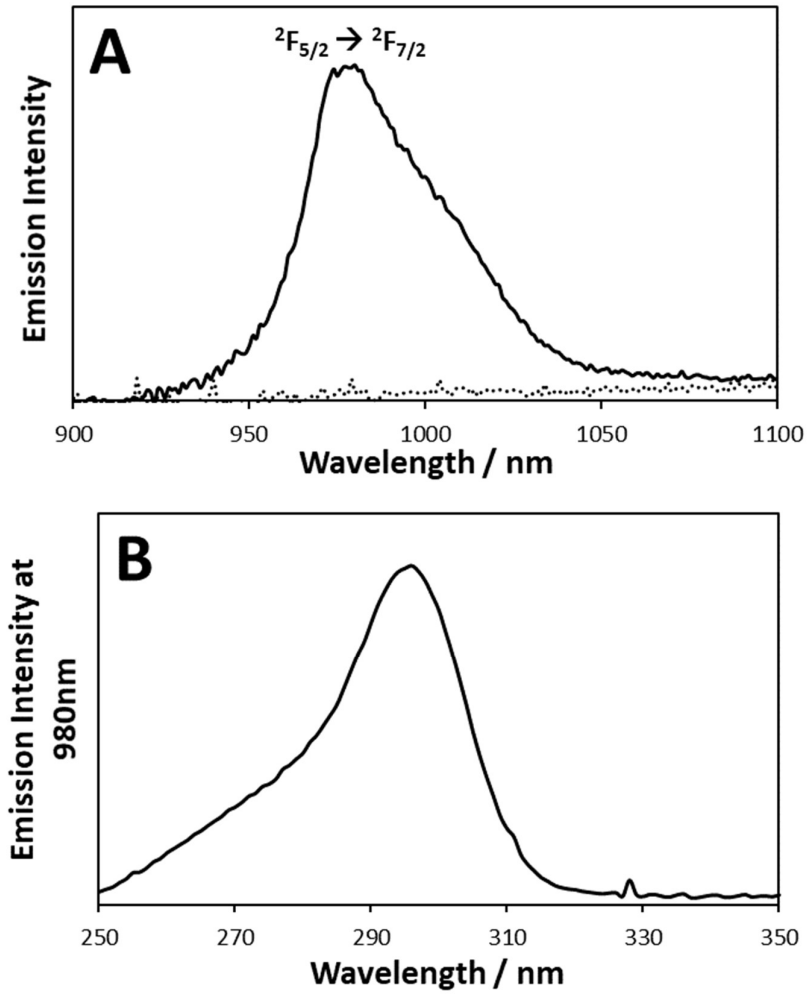

Figure 14: A) The emission profile of HC01 with bound Yb(III) after excitation at 280 nm, the dotted line is a Yb(III) baseline in water. B) The excitation spectra of HC01 with bound Yb(III) monitoring the emission at 980 nm. The sample consisted of peptide (500 µM) in aqueous HEPES (10 mM, pH 7.0) and YbCl3 (150 µM).

observed in the solid state crystal structures as binding Tb(III) in between coiled coil helices. The luminescence behaviour of both HC0C and HC04 suggested that they both formed antiparallel trimers with the core binding site of HC04 disrupted and not intact. The system with the greatest intensity of Tb(III) emission was HC01 and so this was used to validate the predicted Dy(III) and Sm(III) sensitisation by tryptophan. Sensitisation of both Dy(III) and Sm(III) by tryptophan was found to be successful with observable emission from both lanthanide species. This is exciting as it adds the potential for two more lanthanide ions to be incorporated into biological systems as probes sensitised by tryptophan.

A brief consideration was given to other mechanisms resulting in the dramatic quenching of the tryptophan emission. Of particular note was both Eu(III) and Yb(III) which resulted in 78% and 59% quenching respectively. The behaviour of Yb(III) was especially important as no resonance mechanisms could explain this transfer (due to a lack of spectral overlap), therefore it was reasoned to be due to electron transfer. The study was supported by a rudimentary thermodynamic analysis which illustrated that only Eu(III), Yb(III), Sm(III) and Tm(III) could engage in this transfer with tryptophan. Pleasingly, all four demonstrated quenching of the tryptophan emission, however, the electron transfer event was not directly monitored. The observation of Yb(III) emission at 980 nm was encouraging as it suggested that a Yb(II) species had been formed which then decayed to give an excited state Yb(III) species. This redox behaviour will be further studied for the development of charge transfer cascades.

These studies have demonstrated that HC01 is a good model system to study lanthanide-tryptophan interactions in a biomimetic framework. Nonetheless, it is apparent that for use as a light harvesting system further improvement can be made. Inferring from the predictive calculations it is clear that the emission spectra of the donor species is the primary factor that can be easily modified. This feature and its potential for improvement will be considered in greater detail elsewhere.

3.13 REFERENCES

[1] K. Jensen, P. E. Jensen and B. L. Moller, *Trends Plant Sci.*, 2012, **17**, 60-63.

[2] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley and T. Gray, *Protein Sci.*, 1995, **4**, 2411-2423.

[3] S. Cotton, *Lanthanide and Actinide Chemistry*, John Wiley and Sons Ltd., Chichester, 2006.

[4] i) G. E. Peterson and P. M. Bridenbaugh, *J. Opt. Soc. Am.*, 1963, **53**, 301-302; ii) W. DeW. Horrocks, G. F. Schmidt, D. R. Sudnick, C. Kittrell and R. A. Bernheim, *J. Am. Chem. Soc.*, 1997, **99**, 2378-2380.

[5] i) P. R. Selvin, T. M. Rana and J. E. Hearst, *J. Am. Chem. Soc.*, 1994, **116**, 6029-6030; ii) P. R. Selvin and J. E. Hearst, *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 10024-10028; iii) P. R. Selvin, *Annu. Rev. Biophys. Biomol. Struct.*, 2002, **31**, 275-302.

[6] J. -L. Adam, *Chem. Rev.*, 2002, **102**, 2461-2476.

[7] F. S. Richardson, *Chem. Rev.*, 1982, **82**, 541-552.

[8] S. I. Weissman, *J. Chem. Phys.*, 1942, **10**, 214.

[9] J. –C. G. Bunzli and C. Piguet, *Chem. Soc. Rev.*, 2005, **34**, 1048-1077.

[10] J. -C. G. Bunzli and S. V. Eliseeva, *Basics of Lanthanide Photophysics* in *Lanthanide Luminescence*, ed. P. Hanninen and H. Harma, Springer, Berlin, 2010, **7**, 1-45.

[11] S. V. Eliseeva and J. C. G. Bunzli, *Chem. Soc. Rev.*, 2010, **39**, 189-227.

[12] M. Latva, H. Takalo, V. M. Mukkala, C. Matachescu, J. C. Rodriguez-Ubis and J. Kankare, *J. Lumin.*, 1997, **75**, 149-169.

[13] i) W. DeW. Horrocks and A. P. Snyder, *Biochem. Biophys. Res. Commun.*, 1981, **100**, 111-117; ii) W. DeW. Horrocks and W. E. Collier, *J. Am. Chem. Soc.*, 1981, **103**, 2856-2862; iii) J. Bruno, W. DeW. Horrocks and R. J. Zauhar, *Biochemistry*, 1992, **31**, 7016-7026.

[14] i) K. J. Franz, M. Nitz and B. Imperiali, *ChemBioChem*, 2003, **4**, 265-271; ii) M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen and B. Imperiali, *Angew. Chem. Int. Ed.*, 2004, **43**, 3682-3685.

[15] M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton and A. F. A. Peacock, *J. Am. Chem. Soc.*, 2014, **136**, 1166-1169.

[16] M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton and A. F. A. Peacock, *Chem. Sci.*, 2016, **7**, 2207-2216.

[17] E. H. Strickland, J. Horwitz, E. Kay, L. M. Shannon, M. Wilchek and C. Billups, *Biochemistry*, 1971, **10**, 2631-2638.

[18] P. R. Callis, *Methods Enzymol.*, 1997, **278**, 113-150.

[19] Y. P. Meshalkin, *Quantum Electron.*, 1996, **26**, 536-537.

[20] i) L. Stryer, *J. Am. Chem. Soc.*, 1966, **88**, 5708-5712; ii) L. P. McMahon, W. J. Colucci, M. L. McLaughlin and M. D. Barkley, *J. Am. Chem. Soc.*, 1992, **114**, 8442-8448.

[21] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer Science, New York, 2006.

[22] T. Keleti, *FEBS Lett.*, 1970, **7**, 280-282.

[23] i) D. L. Dexter, *J. Chem. Phys*., 1953, **21**, 836; ii) C. B. Murphy, Y. Zhang, T. Troxler, V. Ferry, J. J. Martin and W. E. Jones Jr., *J. Phys. Chem. B*, 2004, **108**, 1537-1543.

[24] i) T. Forster, *Ann. Phys. (Berlin)*, 1948, **437**, 55-75; ii) E. A. Jares-Erijman and T. M. Jovin, *Nature Biotech.*, 2003, **21**, 1387-1395; iii) K. E. Sapsford, L. Berti and I. L. Medintz, *Angew. Chem. Int. Ed.*, 2006, **45**, 4562-4588.

[25] i) J.-L Martini, C. Tetreau, F. Pochon, H. Tourbez, J.-M. Lentz and D. Lavalette, *Eur. J. Biochem.*, 1993, **211**, 467-473; ii) B. D. Schlyer, D. G. Steel and A. Gafni, *J. Biol. Chem.*, 1995, **270**, 22890-22894; iii) S. Ghosh, A. Misra, A. Ozarowski and A. H. Maki, *J. Phys. Chem. B*, 2003, **107**, 11520-11526.

[26] i) V. Anantharaman and J. Chrysochoos, *J. Less Common Metals*, 1983, **93**, 59-66; ii) V. Anantharaman and J. Chrysochoos, *J. Less Common Metals*, 1983, **93**, 67-72; iii) M. D. Gaye-Seye and J. J. Aaron, *Biomed. Chromatogr.*, 1999, **13**, 171-172.

[27] P. R. Callis and J. T. Vivian, *Chem. Phys. Lett.*, 2003, **369**, 409-414.

[28] R. Barer and S. Joseph, *J. Cell. Sci.,* 1954, **s3-95**, 399-423.

[29] M. Khrenova, I. Topol, J. Collins and A. Nemukhin, *J. Biophys.*, 2015, **108**, 126-132.

[30] C. J. Rieger and F. G. Carpenter, *J. Res. Natl. Bur. Stand.*, 1959, **63A**, 205-211.

[31] W. R. Kirk, W. S. Wessels and F. G. Prendergast, *J. Phys. Chem.*, 1993, **97**, 10326-10340.

[32] M. Kaczmarek and S. Lis, *Int. J. Photoenergy*, 2007, **2007**, 42582.

[33] Y. Yuetao, S. Qingde, Z. Huazhang and Z. Guiwen, *Spectrochim. Acta A*, 1998, **54**, 645-649.

[34] A. Beeby, I. M. Clarkson, R. S. Dickins, S. Faulkner, D. Parker, L. Royale, A. S. Sousa, J. A. G. Williams and M. Woods, *J. Chem. Soc. Perkin Trans. 2*, 1999, **3**, 493-504.

[35] R. M. Supkowski and W. DeW. Horrocks, *Inorg. Chim. Acta*, 2002, **340**, 44-48.

[36] i) K. W. Lei, W. S. Liu and M. Y. Tan, *Spectrochim. Acta A*, 2007, **66**, 118-125; ii) K. Binnemans, *Coord. Chem. Rev.*, 2015, **295**, 1-45.

[37] i) M. Seitz, A. G. Oliver and K. N. Raymond, *J. Am. Chem. Soc.*, 2007, **129**, 11153-11160; ii) L. Tei, Z. Baranyai, E. Brucher, C. Cassino, F. Demicheli, N. Masciocchi, G. B. Giovenzana and M. Botta, *Inorg. Chem.*, 2010, **49**, 616-625.

[38] S. Lis, T. Kimura and Z. Yoshida, *J. Alloys Compd.*, 2001, **323**, 125-127.

[39] J. G. Bunzli, *Chem. Rev.*, 2010, **110**, 2729-2755.

[40] S. B. Meshkova, A. V. Kiriyak, M. P. Tsvirko and V. P. Gorodnyuk, *J. Anal. Chem.*, 2008, **63**, 840-843.

[41] i) H. Kurreck and M. Huber, *Angew. Chem. Int. Ed. Engl.*, 1995, **34**, 849-866; ii) Y. Ding, Y. Li and G. Yu, *Chem.*, 2016, **1**, 790-801.

[42] i) T. Inada, Y. Funasaka, K. Kikuchi, Y. Takahashi and H. Ikeda, *J. Phys. Chem. A*, 2006, **110**, 2595-2600; ii) D. Kovacs and K. E. Borbas, *Coord. Chem. Rev.*, 2018, **364**, 1-9.

[43] W. DeW. Horrocks, J. P. Bolender, W. D. Smith and R. M. Supkowski, *J. Am. Chem. Soc.*, 1997, **119**, 5972-5973.

[44] P. Meers, *Biochemistry*, 1990, **29**, 3325-3330.

[45] D. Rehm, A. Weller, *Isr. J. Chem.*, 1970, **8**, 259-271.

Chapter 4

EXPLORING THE UTILITY OF NON-NATURAL AMINO ACIDS IN THE CONSTRUCTION OF LIGHT HARVESTING MODULES

4.1 INTRODUCTION

One of the key appeals for the use of protein in the design of a complex systems, is the exquisite way in which residues can be precisely placed in a complex spatial arrangement using only a simple linear sequence. However, a distinct limitation is that the range of sidechain functionalities is typically restricted to the 20 canonical amino acids. There are a limited number of commercially available non-natural amino acids such as naphthalene functionalised species (Fig. 1) for use in research, but as such these tend to be expensive and often result in difficulties with the solid phase methodologies. Therefore it is worth considering synthetic strategies to access interesting chromophoric motifs. It should be noted that the exploration of non-natural amino acids and their distinct advantages for protein and supramolecular design has been discussed at length in the literature.[1]

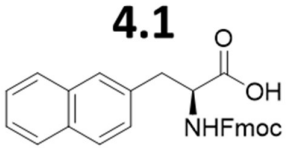

Figure 1: The molecular structure of an available naphthalene functionalised amino acid with a fluorenylmethyloxycarbonyl protecting group that is compatible with solid phase peptide synthesis.

In this chapter a synthetic approach to tyrosine derived non-natural amino acids will be developed and discussed, due to the interesting synthetic potential of this species. This strategy will then be used to prepare a novel coumarin functionalised amino acid for use in peptide synthesis. The effect of a hydrophobic sidechain from a non-natural naphthalene functionalised amino acid will be demonstrated and the feasibility of including non-natural amino acids into a series of HC01 derivatives is studied. The use of both naphthalene and coumarin species was due to their similar size to the indole of tryptophan and the literature supporting their use for sensitising Tb(III) and Eu(III) emission.[2] A small library of HC01 derivatives is prepared and their respective structures assessed utilising both solution and solid state techniques.

4.2 OBSERVING THE HYDROPHOBIC EFFECT

The commercially available naphthalene amino acid, **4.1**, was incorporated as the sensitising chromophore into an MB1-2 derivative system as this was the most well studied system within the group. This involved the same trimer design as previously reported,[3] with a core binding site between the second and third heptads (*see* Table 1). Unlike the HC series, this novel peptide had the chromophore located adjacent to the binding site. While the synthesis and purification of the MB1-2,14Nap were successful, its behaviour in solution was particularly distinct. Before the addition of Tb(III) ions the peptide was observed by circular dichroism to be lacking in any significant helical structure (Fig. 2). Disturbingly, upon the addition of Tb(III) no enhancement in helical folding was observed, even following excess addition (10 equivalents). This suggested that the binding site was no longer intact and therefore was not able to template the formation of a trimeric coiled coil. The only difference between the previously characterised MB1-2 system and this peptide was the naphthalene functionalised amino acid in place of the previous tryptophan residue. Thus it is probable that the more hydrophobic nature of the

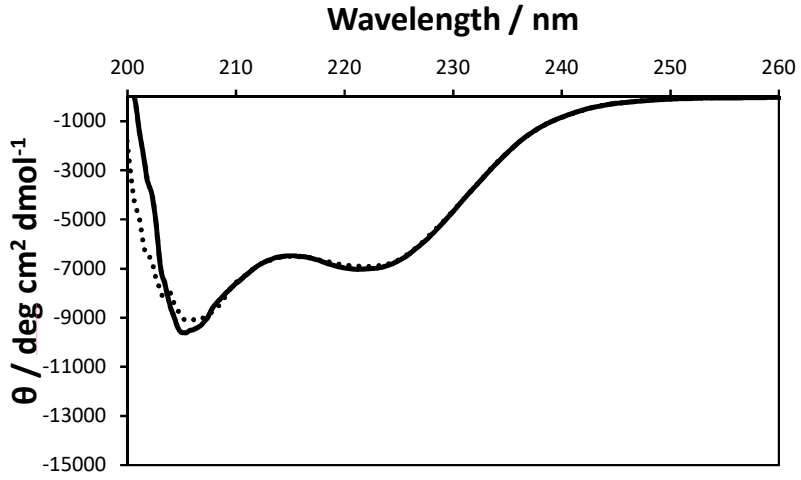

Figure 2: The circular dichroism profile for MB1-2,14Nap both before (solid line) and after (dotted line) the addition of 10 equivalents of Tb(III). Samples were peptide monomer (90 μM) in aqueous HEPES buffer (10 mM, pH 7.0).

naphthalene coupled with the hydrophilic binding site residues, resulted in the thermodynamic spontaneity for the binding site formation to be lost in favour of burying the naphthalene sidechains. Interestingly, the synthesis of this peptide demonstrated the ease of including this non-natural amino acid using solid phase peptide synthesis. Therefore it was still desirable to consider non-natural residues as tools to introduce chromophores into the HC series.

4.3 TYROSINE DERIVED NON-NATURAL AMINO ACIDS

In order to position chromophores within a peptide system it was desirable to prepare unnatural amino acids that could supplement the limited range available from commercial sources. The synthetic protocol was desired to be both versatile for large substrate scope, while also being relatively straight-forward. Typically, the most time consuming and synthetically challenging steps are to impart enantiopurity to a product, therefore it seemed sensible to use a reagent which already possessed the necessary chirality. Of particular interest was the phenol side chain of tyrosine, which has been reported to undergo a Pechmann condensation reaction.[4] This is an interesting approach to generate coumarin functionalised amino acids, dramatically improving previous synthetic procedures as this involved but a single step to generate a useful product. It should be noted that the phenol of tyrosine is not particularly activated, due to the donating effect being from the *para*-positioned functionality, which would lead to relatively low yields for the Pechmann reaction.[5]

The preparation of coumarin functionalised amino acids is not new, having been explored previously by F. Bennett *et al.* and P. Kele *et al.*[6] However, the strategy applied through all

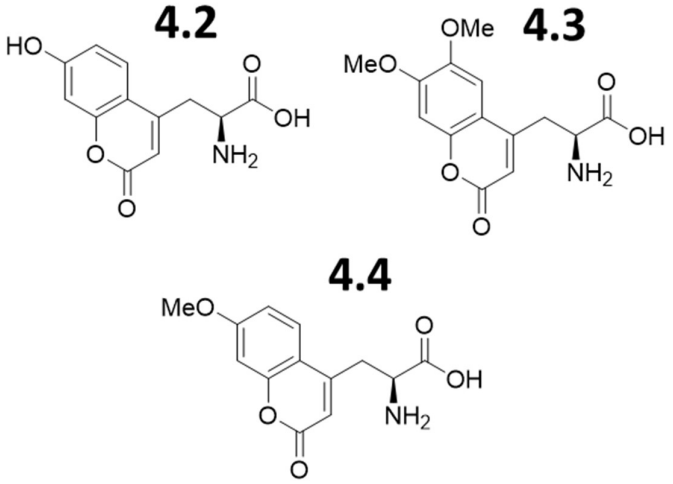

Figure 3: Literature examples of coumarin functionalised amino acids. The biological functionality is built onto the coumarin at the 4-position, resulting in similar structures from multiple strategies.

these reactions, involved building the amino acid functionality onto a desired coumarin via the 4-position to yield compounds such as **4.2** and **4.3** (Fig. 3). This resulted in relatively long synthetic strategies of 5-8 steps with correspondingly low yields. An improvement on this strategy was the use of the natural amino acid glutamate to provide the initial enantiopurity,[7] nevertheless, this also involved a number of steps to prepare the ester functionality which could then react with a phenol to form the coumarin product **4.4**. As such this method was also time consuming with relatively low overall yields. Thus the strategy employed by A. Fonseca *et al.* seemed to be the most elegant as it utilised tyrosine as the phenol and via one step could

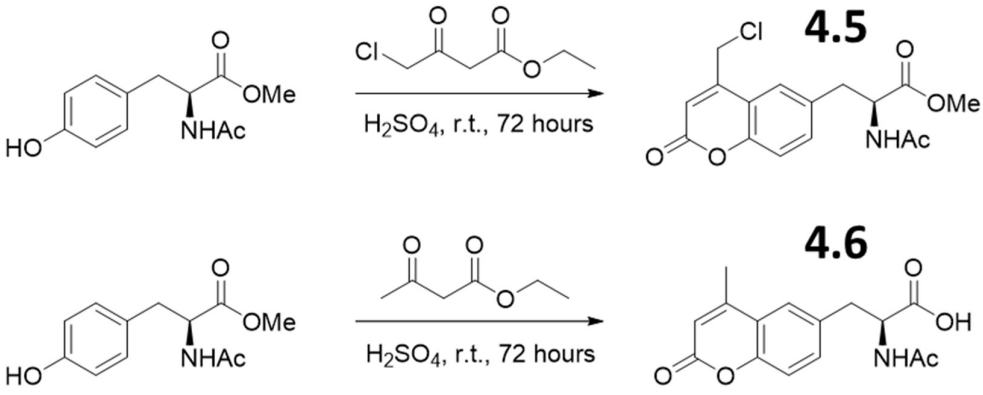

*Figure 4: The literature reaction to generate the N-acetyl-L-(4-chloromethyl)coumarin-6-yl-alanine methyl ester **5.5** product followed by the planned reaction to generate the novel N-acetyl-L-(4-methyl)coumarin-6-yl-alanine **5.6** product.*

generate a usable product, **4.5**.[4] It should be noted that the methyl ester was not deprotected, thus the reported product was not directly useful for routine solid phase peptide synthesis. There seemed to be no reason that the deprotection could not be performed so this strategy was carried forward.

The intention was to repeat the literature methodology and synthesize **4.5**, followed by deprotection of the methyl ester to generate the useful carboxylic acid, and to ensure there were no complications with this hydrolysis. However, this species would not be particularly useful for peptide synthesis due to the reactivity of the halogen functionality under both cleavage and deprotection conditions. Therefore, it is desirable to synthesise a novel non-natural amino acid from tyrosine with a relatively inert coumarin sidechain, such as a 4-methylcoumarin derivative to give **4.6**. The 4-methylcoumarin was particularly appealing as the ethyl acetoacetate reagent is a common industrial feedstock. Fortunately the literature repeat was successful, though it should be noted that the high reported yields of 55% did not seem reproducible as multiple attempts resulted in a low yield of typically 23%. Interestingly, the deprotection of the methyl ester occurred during the workup with the use of sodium bicarbonate to neutralise the sulphuric acid present. Therefore attempts to synthesize the more inert 4-methylcoumarin were carried out by reacting ethyl acetoacetate with acetyl and methyl protected tyrosine in concentrated sulphuric acid to generate **4.6** via a Pechmann condensation (Fig. 4). Pleasingly, the product crystallised out of ethanol and water as small square plates and the structure was confirmed by x-ray diffraction (Fig. 5). It was particularly exciting as this novel species could be directly used as an *N*-terminal cap in solid phase peptide synthesis. Unfortunately, the yield for this reaction was low at 3.9% therefore there is significant room for optimisation in future work including a more versatile protection for the nitrogen of the amide.

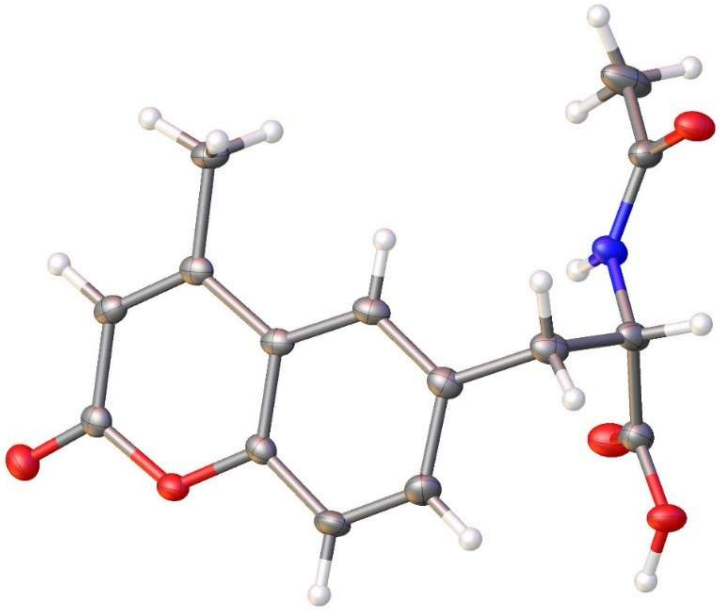

Figure 5: The crystallographic structure of N-acetyl-L-(4-methyl)coumarin-6-yl-alanine. Ellipsoids drawn at 50% probability.

4.4 INCORPORATION OF NON-NATURAL AMINO ACIDS INTO PEPTIDES

The use of a commercially available naphthalene amino acid to prepare the MB1-2,14Nap peptide raised the possibility of synthesizing a HC01 derivative, HC01-2Nap (*see* Table 1). While the hydrophobicity of the naphthalene sidechain was thought to be a problem when positioned adjacent to a lanthanide binding site spanning the second and third heptads, the HC01 system has the chromophore instead located at a top core site. Indeed, this change in location should lead to a stabilisation of the coiled coil as burying of the naphthalene sidechains would be eminently favourable, perhaps even helping to template the designed binding site. In addition to the naphthalene species, having successfully synthesized the novel coumarin functionalised amino acid, it was with great excitement that a HC01 derivative with the coumarin chromophore at the *N*-terminus was prepared, HC01-2Cou. The small library of

peptides was completed with the addition of the natural amino acid tyrosine, to create a system with a phenol chromophoric tag at the top site, HC01-2Y. Thus, three novel HC01 derivatives (*see* Table 1), two of which contained non-natural amino acids, were studied to determine both their structural and photophysical properties.

Table 1: The sequences for the designed peptides with the binding site residues in bold and underlined and the chromophoric species in red and italics.

Peptide	Sequence
MB1-2,14Nap	Ac-G IAAIEQK IAA**N**E***Nap***K **D**AAIEQK IAAIEQK IAAIEQK **G**-NH_2
HC01-2Y	Ac-E ***Y***EA**N**EKK **D**AAIESK IQAIEKK IQAIEKK IEAIEHG-NH_2
HC01-2Nap	Ac-E ***Nap***EA**N**EKK **D**AAIESK IQAIEKK IQAIEKK IEAIEHG-NH_2
HC01-2Cou	Ac-***Cou***EA**N**EKK **D**AAIESK IQAIEKK IQAIEKK IEAIEHG-NH_2

The three HC01 derivatives; HC01-2Nap, HC01-2Cou and HC01-2Y were successfully prepared via a solid-phase peptide synthesis methodology and purified by reverse-phase high pressure liquid chromatography. The mass and purity of each of the peptides was determined by matrix-assisted laser desorption/ionisation mass spectrometry and analytical high pressure liquid chromatography respectively (*see* S7 - S11).

4.5 CIRCULAR DICHROISM STUDY

To establish the structural behaviour in solution of each of the HC01 derivatives, their circular dichroism profiles were measured both before and after the addition of Tb(III) ions (Fig. 6). Interestingly, all the systems exhibited an increasing in folding of between 10-20% upon the addition of one equivalence of Tb(III) ions to the trimer unit before reaching a plateau. This suggests that the lanthanide binding behaviour for all three systems is as expected, with a single

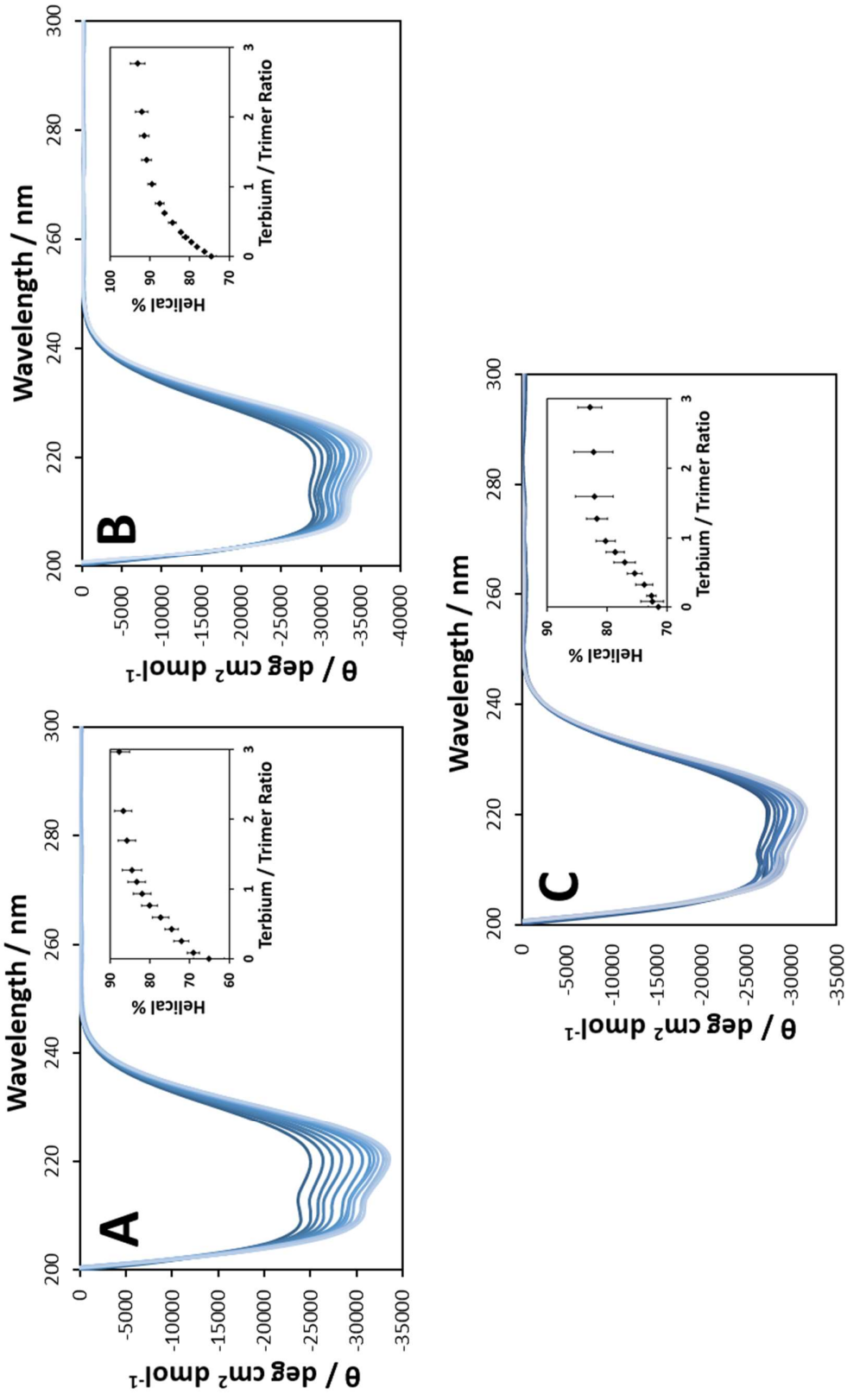

Figure 6: The circular dichroism profiles for A) HC01-2Y, B) HC01-2Nap and C) HC01-2Cou upon the addition of Tb(III) ions. Samples were peptide monomer (90 μM) in aqueous HEPES (10 mM, pH 7.0), metallated samples involved the addition of Tb(III) up to three equivalents. The insets represent change in helical content upon the increase in Tb(III) to peptide trimer ratio. The error bars are the standard deviation of three isotherm repeats.

ion being bound in the designed site. Considering the native state of both the peptide systems HC01-2Nap and HC01-2Cou, they exhibit very high helical content of 75% ± 1% and 72% ± 2% respectively before the addition of metal ion. This is in agreement with the well folded native HC01 peptide (70% ± 3%) indicating that both HC01-2Nap and HC01-2Cou would behave in a similar manner to the original HC01 peptide scaffold. The HC01-2Y system was slightly less well folded at 65% ± 3%, suggesting that the size and solubility of the chromophore sidechain could have an effect upon the helical structure of these complexes. All the assemblies achieved greater than 80% helical content upon the addition of Tb(III) ions suggesting that the final metallated systems are of a similar structure throughout this small library. This was not a dramatic change in folding, clearly suggesting that these coiled coils have a degree of preformed structure before the addition of metal ions.

The increase in helical content was utilised to observe the binding event (*see* Chapter 8.9), resulting in a plot of the fraction of trimer bound as a function of the addition of Tb(III) ions (Fig. 7). This allowed the determination of the binding constants for HC01-2Y and HC01-2Nap with Tb(III). The isotherm for HC01-2Cou was somewhat more complicated with more than one event occurring, therefore fitting to a three-to-one binding model was not sufficient. The 4-methylcoumarin sidechain is particularly bulky and likely involves a rearrangement upon binding the metal ion. Pleasingly the determined log of the binding constants for HC01-2Y and HC01-2Nap of 4.93 ± 0.10 and 4.96 ± 0.06 respectively, were in agreement for the binding constants determined for HC01 with Tb(III) of 5.03 ± 0.13 from the circular dichroism titration and 5.25 ± 0.22 from the luminescence titration. This suggests that the small library of peptides behaves in a similar way to the HC01 system. For HC01-2Cou, it should be noted that the assumption that folding equates to binding is only true in certain situations. Therefore when

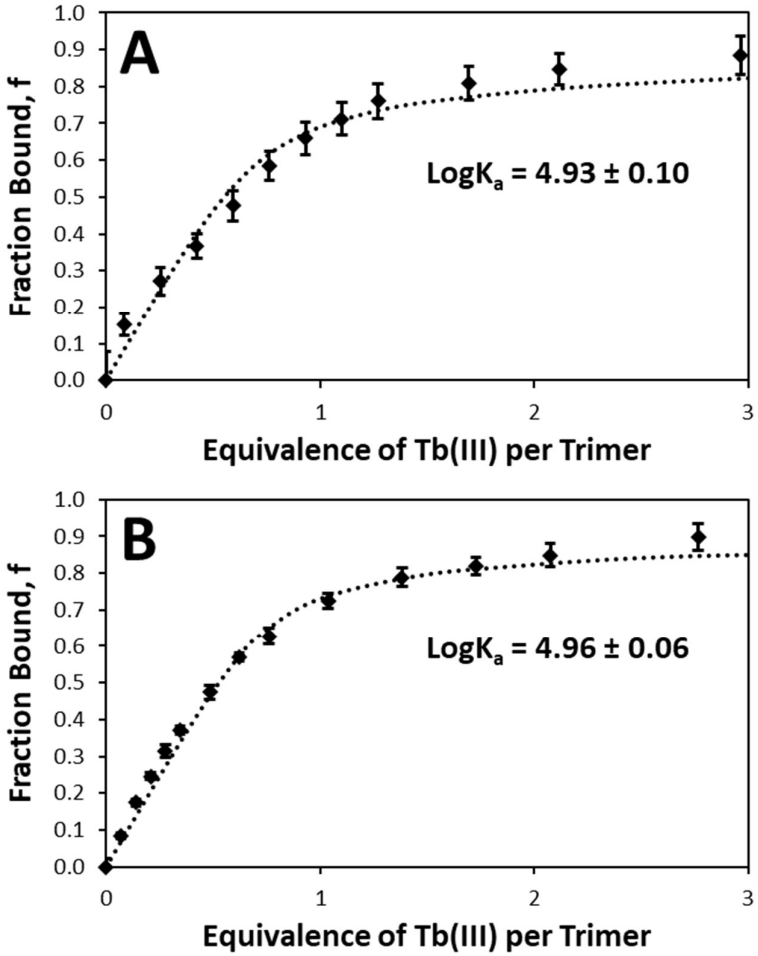

Figure 7: The binding profiles for both A) HC01-2Y and B) HC01-2Nap with Tb(III) ions as monitored by the change in circular dichroism spectra. The dotted line illustrates a thermodynamic fit for the binding constant values shown. The error of the log of the binding constant is due to fitting of the upper and lower limits of the binding profiles.

the folding arrangement required large sidechains to migrate, the steric factors may mask the binding event by apparent unfolding, folding or lack of net change.

4.6 MASS SPECTROMETRY STUDY

Further insights into the structural behaviour of the HC01 derivatives was achieved by observing the native complexes by mass spectrometry (*see* Table 2). Each of the systems were

observed by utilising electrospray ionisation to generate a 5+ species (*see* S23-S24). Upon analysis these species were found to be comprised of three monomer units, in agreement with the preformed coiled coil trimer structure. Thus despite the circular dichroism results suggesting that HC01-2Y was slight less well folded, all three systems appear to have a trimeric structure present, further supporting the hypothesis that the coiled coils are preformed. This is to be expected as, while the *N*-terminal top sites can have a degree of freedom due to the hydrophilic residues that comprise the binding motif, the hydrophobic nature of the chromophores result in burying to reduce solvent interaction. Therefore, it is likely that the chromophores act to 'cap' the *N*-terminus reducing flexibility.

Table 2: The observed monoisotopic +5 charged species for each of the peptides and the resultant trimer and monomer masses.

Peptide	Monoisotopic (m/z)	Charge (z)	Monoisotopic Trimer Mass	Monoisotopic Monomer Mass
HC01-2Y	2482.6	5+	12408.0	4136.0
HC01-2Nap	2503.0	5+	12510.0	4170.0
HC01-2Cou	2444.8	5+	12219.0	4073.0

4.7 CRYSTALLOGRAPHIC STUDY

Having synthesised a small library of peptides it was desirable to ascertain the solid state structures. Using the strategies employed in Chapter 2, crystals of both HC01-2Nap and HC01-2Y with Tb(III) ions were successfully grown (Fig. 8). The crystal analysis and data collection were carried out at Diamond Light Source on the i04 beamline. Interestingly the symmetry and unit cell of the two crystals were both *P41,21,2* but differed from the previous *H3* symmetry of the other HC peptides. This is most likely due to the *N*-terminal binding site disrupting the

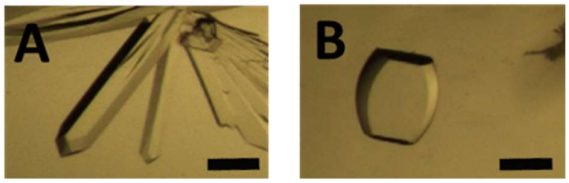

Figure 8: Photographs of the crystals of both A) HC01-2Y as long triangular prisms and B) HC01-2Nap as curved plates. These crystals were grown from a combined solution of peptide (16 mg mL^{-1} peptide, 6 mM terbium chloride and 5 mM zinc acetate) and precipitant (0.21 M ammonium chloride and 25% v/v poly(ethylene) glycol 3350) in a 1:1 ratio. The black bar indicates 0.2 mm.

previous packing in the crystal matrix. Indeed, this may explain why it has been so challenging to crystallise HC01, as there is likely to be competition between the different packing orders. The collected datasets were processed within the Phenix interface,[8] utilising Phaser and Coot packages to both phase and build the models for the proposed electron density maps.

Table 3: The crystallographic data and collection statistics for both the HC01-2Y and HC01-2Nap peptide crystals. The values in parenthesis are for the highest resolution set.

	HC01-2Y	**HC01-2Nap**
Wavelength (Å)	0.9795	0.9795
Resolution Range (Å)	40.3-2.2 (2.3-2.2)	41.3-2.2 (2.3-2.2)
Space Group	P 41 21 2	P 41 21 2
Unit Cell		
a = b, c	90.2, 45.0	88.2, 46.8
α = β = γ	90	90
No. of Total Reflections	218376 (21293)	212172 (20848)
No. of Unique Reflections	9859 (944)	9823 (948)
Multiplicity	22.1 (22.6)	21.6 (22.0)
Completeness (%)	99.0 (97.7)	99.8 (99.7)
I / σ(I)	19.5 (3.8)	20.8 (4.1)
Wilson B (Å^2)	57.2	55.1
R_{merge}	0.09 (0.84)	0.09 (0.86)
$CC_{1/2}$	1.00 (0.99)	1.00 (0.96)

As before the metal ions were easily placed due to their strong anomalous scattering signals (*see* Chapter 2) and this greatly simplified the model building process. The statistics for the final models of the HC01-2Y and HC01-2Nap with Tb(III) bound can be found in Table 4.

Table 4: The refinement statistics for the models of both HC01-2Y and HC01-2Nap with bound Tb(III) fitted to the density maps.

	HC01-2Y	**HC01-2Nap**
No. of Non-H Atoms	881	953
No. of Waters	1	68
R_{work} / R_{free}	0.22 / 0.26	0.23 / 0.26
CC_{work} / CC_{free}	0.96 / 0.97	0.96 / 0.98
R.m.s.d.		
Bonds (Å)	0.007	0.003
Angles (º)	0.99	0.78
Ramachandran[a]	100 / 0 / 0 / 0	99 / 1 / 0 / 0
Average *B* (Å^2)	82.1	77.0

[a]*Indicates the percentage of residues in favoured/allowed/outliers/rotamer outlier positions.*

In agreement with the previous HC crystal structures (*see* Chapter 2), the structures of HC01-2Y and HC01-2Nap with Tb(III) were found to be parallel coiled coils comprised of three α-helices (Fig. 9). The core of the coiled coil domains was found to be interlocking isoleucine sidechains fitting with the 'knobs into holes' logic of hydrophobic burying driving protein folding. In addition to the structurally directing isoleucine residues, salt bridges were observed between the e-position glutamate residues and g-position lysine residues on adjacent α-helices.

The binding site was comprised of six residues for both systems, an aspartate residue and asparagine residue from each of the three helices (Fig. 10). The aspartate residues were bidentate via the two oxygen atoms of the carboxylic acid sidechain, with an average Tb-O

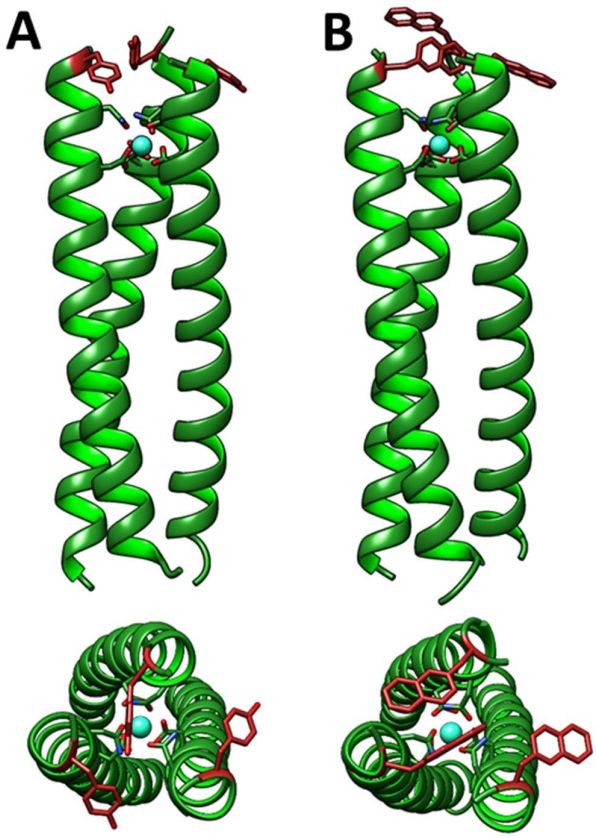

Figure 9: The crystallographically determined models for both A) HC01-2Y and B) HC01-2Nap with Tb(III) bound in the core binding site from a side view and top view. The chromophoric residue is in red and the Tb(III) ion is cyan, the binding site aspartate and asparagine side chains are shown (green for carbon, blue for nitrogen and red for oxygen).

distance of 2.3 Å (*see* Table 5). The asparagine residues appeared to only coordinate via a single donor atom which was modelled to be the oxygen of the amide side chain for reasons previously discussed (*see* Chapter 2). The average Tb-O distance for the asparagine residues was found to be 2.4 Å which was also in agreement with reported literature values.[9] Therefore the binding site was found to be remarkably similar to the HC peptides previously characterised with a Tb(III) ion, however, both systems did not exhibit a water molecule close (~3 Å) to the residues of the coordination sphere which was observed for HC02, HC03 and HC04 (*see* Chapter 2). Nevertheless the slightly long Tb-O distance for an asparagine residue in both HC01-2Y and HC01-2Nap resulted in there being the same ambiguity in coordination number leading to an 8/9 coordinate site. Overall, there appeared to be little difference in the binding

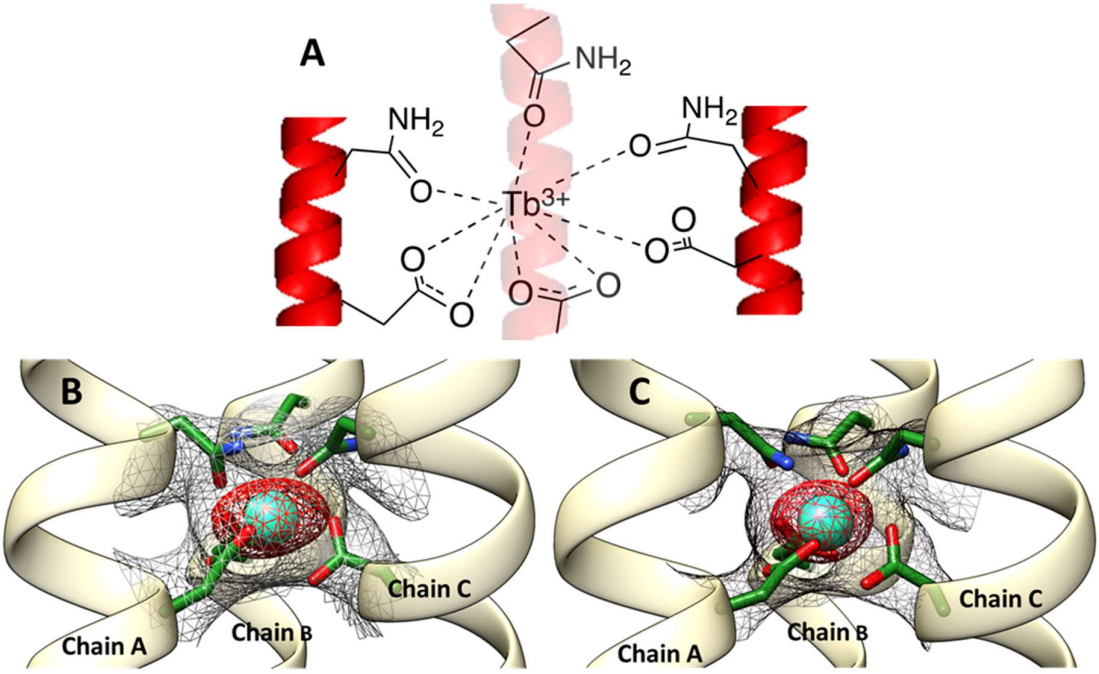

Figure 10: A) The idealised schematic diagram of the binding site with a terbium trication illustrating the bidentate aspartate and monodentate asparagine residues. The crystallographically determined model of the binding site for both B) HC01-2Y and C) HC01-2Nap. The electron density is shown as a black mesh (r.m.s.d 1.4) and the anomalous scattering signal is shown as a red mesh (r.m.s.d. 4.0). The Tb(III) ion is shown as a cyan sphere, the oxygen atoms are red, nitrogen is blue and carbon is green.

site environment between these modified HC01 derivatives and the previously studied HC series. This was reassuring as it implied that the HC01 design was particularly robust, allowing non-natural amino acids to be included with relatively little impact on the final structure.

Table 5: The measured Tb-O distances from the crystallographically determined models of both HC01-2Y and HC01-2Nap. The corresponding chain identity is in parenthesis. Error of measured distances is ± 0.1 Å.

Tb(III)-Oxygen Distances (Å)[a]	HC01-2Y	HC01-2Nap
Asn(A)	2.7	2.7
Asn(B)	2.1	2.3
Asn(C)	2.4	2.2
Asp(A)	2.4, 2.0	2.5, 2.2
Asp(B)	2.6, 2.4	2.3, 2.0
Asp(C)	2.5, 2.2	2.5, 2.2

[a]The error in reported measurements can be found in Table 4 (0.007 and 0.004 Å for HC01-2Y and HC01-2Nap respectively), due to high B-factors associated with the model these were reported to one decimal place.

The distances from the bound Tb(III) ion to the nearest chromophore was measured for each system in order to allow calculation of the energy transfer constants in future work (*see* Chapter 6). The distance from the Tb(III) ion to the closest naphthalene of HC01-2Nap was 7.0 Å and the distance from the Tb(III) ion to the closest phenol of HC01-2Y was 7.5 Å. Therefore in both cases the distance is short enough that even weak sensitisation should be visible.

4.8 CONCLUSIONS

In this chapter a novel coumarin functionalised amino acid, suitable for *N*-terminal capping, has been prepared using a simple synthetic strategy. This species was then utilised to prepare a HC01 derivative which had a 4-methylcoumarin sidechain in place of the tryptophan residue. The use of a commercially available naphthalene functionalised amino acid demonstrated that the hydrophobicity of the sidechain can dramatically alter the behaviour of a designed system. Therefore, it is particularly important to consider the sequence position of this amino acid, as placing it into the solvent would result in structural distortion as the naphthalene moiety is buried. Fortunately the use of this amino acid was successful in the synthesis of a HC01 derivative, as the naphthalene scaffolds were partially protected from the solvent, reducing the hydrophobic effect. The final HC01 derivative contained tyrosine to give a phenol chromophore at the *N*-terminus. This small library of derivatives was characterised by circular dichroism and mass spectrometry to understand their dynamic structures. This indicated that all the systems formed coiled coil trimers which bound a single Tb(III) ion requiring three peptide monomers. Both HC01-2Y and HC01-2Nap were successfully crystallised and the structures were found to be parallel coiled coils with a core lanthanide binding site. The

distance from the closest phenol or naphthalene functionality to the Tb(III) ion was measured and found to be 7.5 Å and 7.0 Å, respectively. The design and preparation of a small library of HC01 derived systems has illustrated the versatility of this peptide to the introduction of non-natural amino acids at the *N*-terminal chromophore site. Thus, having completed the structural analysis it remained to study the photophysics of these novel peptide complexes.

4.9 REFERENCES

[1] i) T. Hohsaka and M. Sisido, *Curr. Opin. Chem. Biol.*, 2002, **6**, 809-815; ii) K. E. Schwieter and J. N. Johnston, *J. Am. Chem. Soc.*, 2016, **138**, 14160-14169; iii) N. A. Tavenor, M. J. Murnin and W. S. Horne, *J. Am. Chem. Soc.*, 2017, **139**, 2212-2215.

[2] i) W. –N. Wu, W. –B. Yuan, N. Tang, R. –D. Yang, L. Yan and Z. –H. Xu, *Spectrochim. Acta A*, 2006, **65**, 912-918; ii) J. Andres and A. –S. Chauvin, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15981-15994; iii) N. K. Al-Rasbi, H. Adams and F. O. Suliman, *Dyes Pigm.*, 2014, **104**, 83-88.

[3] M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton and A. F. A. Peacock, *J. Am. Chem. Soc.*, 2014, **136**, 1166-1169.

[4] A. S. C. Fonseca, M. S. T. Goncalves and S. P. G. Costa, *Amino Acids*, 2012, **43**, 2329-2338.

[5] S. Sethna and R. Phadke, *The Pechmann Reaction* in *Organic Reactions*, John Wiley & Sons, Inc., 2004.

[6] i) F. A. Bennett, D. J. Barlow, A. N. O. Dodoo, R. C. Hider, A. B. Lansley, M. J. Lawrence, C. Marriott and S. S. Bansai, *Tetrahedron Lett.*, 1997, **38**, 7449-7452; ii) P. Kele, G. Sui, Q. Huo and R. M. Leblanc, *Tetrahedron Asymmetry*, 2000, **11**, 4959-4963.

[7] M. -P. Brun, L. Bischoff and C. Garbay, *Angew. Chem. Int. Ed.*, 2004, **43**, 3432-3436.

[8] i) A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674; ii) P. D. Adams, R. W. Grosse-Kunstleve, L. – W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2002, **58**, 1948-1954.

[9] i) D. Parker, R. S. Dickins, H. Puschmann, C. Crossland and J. A. K. Howard, *Chem. Rev.*, 2002, **102**, 1977-2010; ii) B. Barja, R. Baggio, M. T. Garland, P. F. Aramendia, O. Pena and M. Perec, *Inorganica Chim. Acta*, 2003, **346**, 187-196; iii) M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen and B. Imperiali, *Angew. Chem. Int. Ed.*, 2004, **43**, 3682-3685.

Chapter 4 – Exploring the Utility of Non-Natural Amino Acids in the Construction of Light Harvesting Modules

Chapter 5

DIVERSE CHROMOPHORES FOR LANTHANIDE SENSITISATION

5.1 INTRODUCTION

Within biological applications the ability to utilise different chromophores provides greater flexibility to spectroscopic experiments allowing the use of multiple channels to simultaneously observe many effects within a single sample. The lanthanides have been used extensively to study biological systems, however, as discussed (*see* Chapter 3) they generally require sensitising chromophores positioned in close proximity to have useful emission. The purpose of this chapter is to explore lanthanide photophysics with a variety of chromophores as the donor species within the designed coiled coil system.

Previously it has been demonstrated that the indole of tryptophan **5.1** provided a useful and interesting chromophore to sensitise localised lanthanides.[1] Indeed, this area has been explored and developed within this thesis (*see* Chapter 3). After the successful synthesis of a novel coumarin functionalised amino acid and the availability of a naphthalene functionalised amino acid **5.3** it was possible to explore the novel photophysics available within the optimised coiled coil design. The versatility of peptide synthesis with the amino acids (Fig. 1) was explored by incorporating them into HC01 derivatives, building a small library of coiled coils with varying

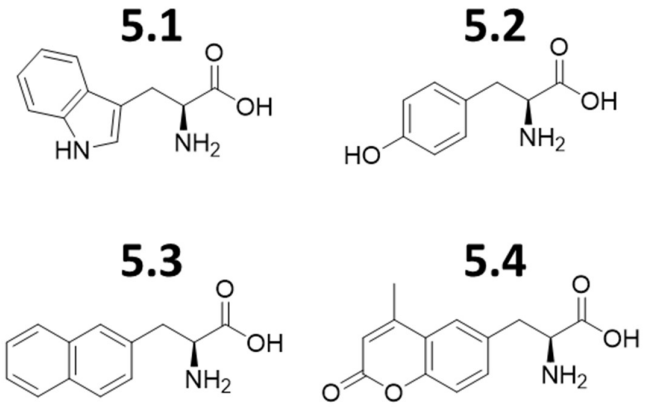

Figure 1: Chemical structures of 5.1) tryptophan, 5.2) tyrosine, 5.3) 3-(2-naphthyl)-L-alanine and 5.4) 4-methylcoumarin functionalised amino acid 6-(4-methylcoumarin)-L-alanine.

chromophores (*see* Chapter 4). Having demonstrated that the structure of the coiled coils were intact, it remained to study the photophysical properties of the new systems.

The knowledge that tyrosine **5.2** can be used to sensitise both Eu(III) and Tb(III) is well established, having been studied previously by Horrocks *et al.*[2] However, the use of tyrosine has not been explored as a sensitiser for Dy(III) or Sm(III). Therefore, it was of particular interest to see if, as with tryptophan, it were possible to expand the scope of this natural amino acid for the sensitisation of other visible emitting lanthanides.

Naphthalene and coumarin derivatives have both been studied with respect to their ability to transfer energy to lanthanides such as Tb(III) and Eu(III). Indeed, a comparison of the triplet state energies for both naphthalene and coumarin indicate that they should be capable of sensitising the visible emitting lanthanides (Fig. 2).[3] However, it should be noted that the energy levels for the singlet and triplet states of naphthalene and coumarin will not be identical to the amino acids **5.3** and **5.4,** but slightly different due to the additional functional groups.

The use of naphthalene as a sensitiser for the lanthanides was first observed by utilising naphthalene solubilised in micelles which were then deposited in an aqueous solution of Tb(III) ions to sensitise emission.[4] This was further advanced by the use of a lanthanide binding cyclodextrin moiety **5.5** which made use of sequestered hydrocarbons, such as naphthalene, from solution to sensitise the emission of bound Tb(III) or Eu(III) ions.[5] Improvements in the yield of emission were made by tethering multiple cyclodextrin moieties to a lanthanide binding core.[6] All these examples make use of naphthalene as an untethered species to enable energy transfer. Nonetheless, the use of naphthalene derivatives covalently attached to

lanthanide chelating crown ligands, such as **5.6**, quickly enabled the study of a large number of properties including the importance of the orientation of the naphthalene donor to both Tb(III) and Eu(III) acceptor species for energy transfer.[7] Indeed, naphthalene functionalised systems have been used to directly chelate the donor and acceptor species, as is the case with

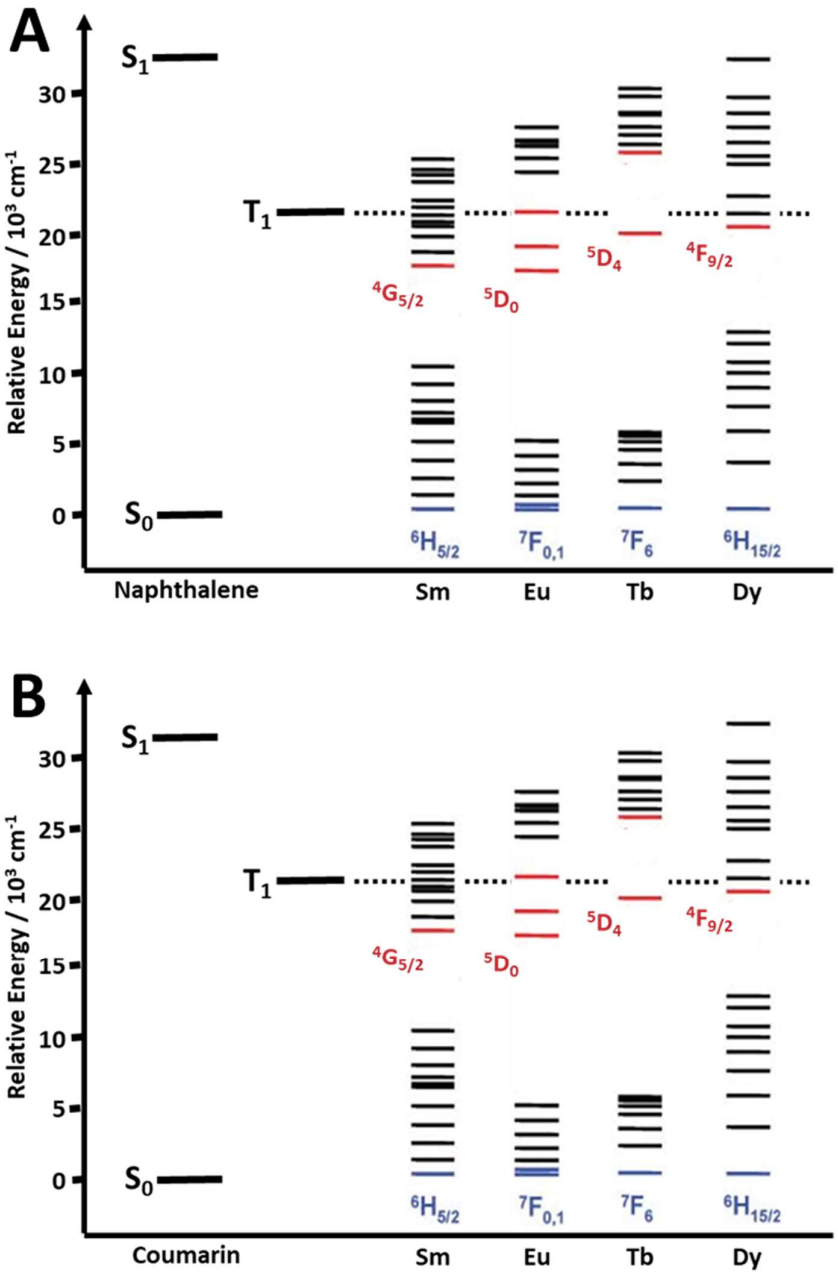

Figure 2: The energy of first singlet and triplet excited states for A) naphthalene and B) coumarin with respect to the energy of the excited states for the visible emitting lanthanides; samarium, europium, terbium and dysprosium.

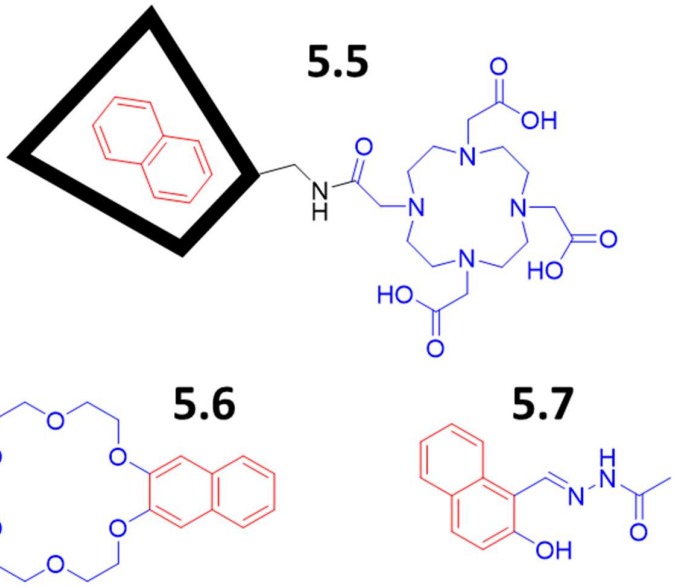

Figure 3: The molecular structures of 5.5) the polycarboxylate lanthanide binding species functionalised with cyclodextrin (trapezoid) to hold the sensitising naphthalene, 5.6) crown ether functionalised with naphthalene and 5.7) a Schiff base monomer that requires three units to bind a lanthanide ion. The lanthanide binding domain is in blue and the sensitising species is in red.

5.7, resulting in highly emissive materials.[8] While all the systems that have previously been discussed utilised Tb(III) or Eu(III) ions it has recently been discovered that naphthalene can be used to sensitise both Sm(III) and Dy(III) emission.[9] This is useful as it enables a single probe to be used to observe multiple lanthanide ions. Despite these results, lanthanide sensitisation by naphthalene in a protein environment has not been described.

The first observation of a coumarin species successfully sensitising a lanthanide was made by P. Tarassoff and N. Filipescu from the triplet state transfer of coumarinium cations to Eu(III) ions in a solution of alkylsulphonic acids.[10] Later work led to the design of multiple small molecule assemblies usually based around ether or carboxylate coordinating species,[11] such as crowns **5.8** or polycarboxylates **5.9** respectively,[12] linked to a coumarin derived sensitiser. These have found use as biological tags or for energy transfer mechanistic studies.[13] Previously coumarin derived systems were only used to sensitise Eu(III) or Tb(III) ions, but as with naphthalene it has recently been demonstrated that energy transfer between coumarin

derivatives and Sm(III) and Dy(III) is possible.[14] Though it should be noted that these systems made use of direct coordination between the coumarin donor and the lanthanide acceptor resulting in highly favourable transfer mechanisms. A particularly thorough distance dependence relationship was determined by J. Andres and A. Chauvin,[15] involving coumarin derivatives which were attached to a lanthanide binding moiety using polymer linkers of varying lengths **5.10-5.12**. Interestingly this study demonstrated that the length of the linker was important for the solid state samples but that transfer efficiencies were relatively similar across the study in the solution state. This was reasoned to be due to the flexibility of the chain in solution, therefore, it would be expected that distance measurements could still be carried out in the relatively rigid HC01-2Cou peptide system functionalised with **5.4**.

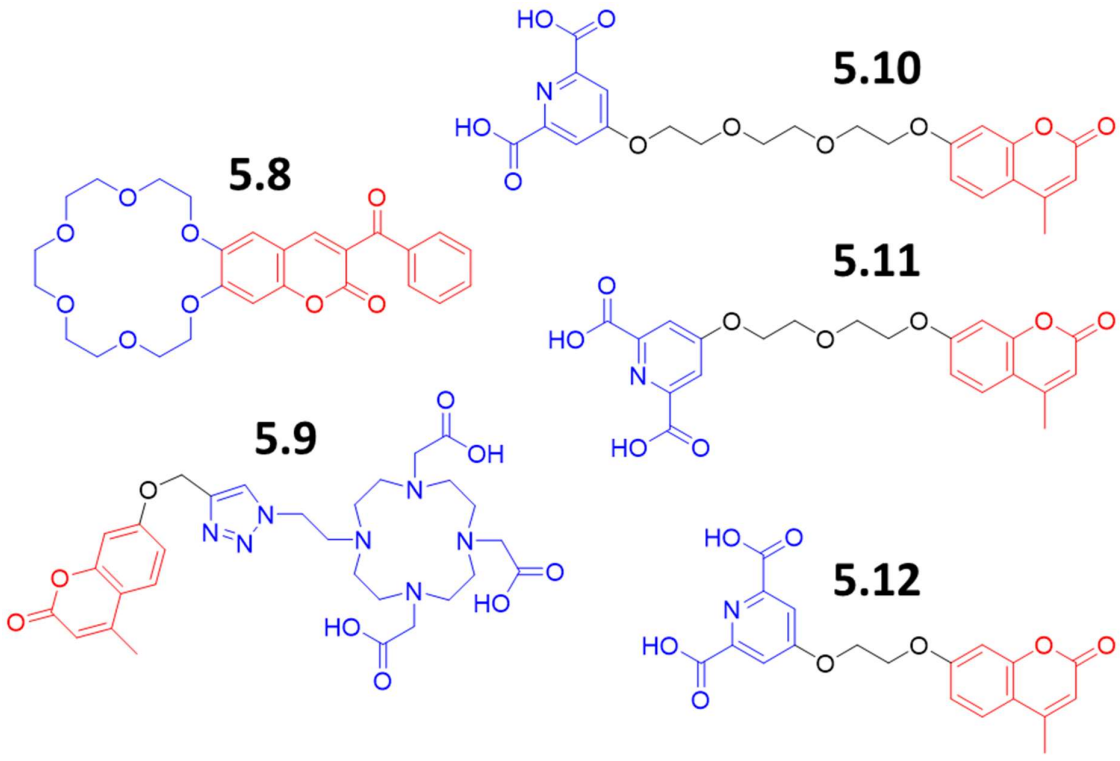

Figure 4: The molecular structures of 5.8) a crown ether functionalised with a coumarin derivative, 5.9) a polycarboxylate lanthanide binding species with an appended coumarin derivative and 5.10-5.12) a series of lanthanide binding monomers with varied polymer lengths to the sensitising 4-methylcoumarin. It should be noted that it requires three monomer units to bind a lanthanide ion. The lanthanide binding domain is in blue, the sensitising species is in red and the linker is in black.

It is particularly exciting that no work has been carried out to explore the use of naphthalene or coumarin functionalised amino acids for sensitising locally bound lanthanides. This is because the amino acid functionality allows simple and direct incorporation into protein and peptide systems. Therefore, this work offers the potential to provide new tools for sensitising lanthanide probes in biologically relevant systems by expanding the scope of available FRET pairs, increasing diversity of species and therefore the versatility of energy transfer dependent experiments.

5.2 ABSORPTION

The molecular structures of phenol, coumarin and naphthalene can be used to infer some details regarding their absorption behaviour. Firstly, the hydroxyl functional group of tyrosine and the lactone of the coumarin suggest that both these systems can interact directly via hydrogen bonding to protic polar solvents such as water. This implies that vibrational modes will not be well resolved resulting in broad, relatively featureless absorption profiles. The hydrophobic naphthalene is not expected to interact extensively with the solvent and therefore would retain fine-structure from various vibrational modes. Pleasingly, it is immediately apparent from the absorption spectra (Fig. 5) that both HC01-2Y and HC01-2Cou have generally broad profiles whereas HC01-2Nap retains some vibrational fine-structure, albeit broadened.

The absorption profile for HC01-2Y exhibits a single feature, a relatively narrow band with a maximum at 277 nm. This corresponds to the $S_0 \rightarrow S_1$ transition, primarily $\pi \rightarrow \pi^*$ in character,for the tyrosine residues in the system.[16] The absorption profile for HC01-2Cou while

being broad has two distinct maxima at 319 and 276 nm which corresponds to $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions respectively. Both these transitions are due to $\pi \rightarrow \pi^*$ transitions of the 4-methylcoumarin moieties at the N-termini of the α-helices.[17] The behaviour of the HC01-2Nap system is somewhat more complicated due to the forbidden nature of the $S_0 \rightarrow S_1$ transition for naphthalene. The symmetrical nature of the naphthalene results in no change in the electronic dipole upon exciting the lowest lying $\pi \rightarrow \pi^*$ transition, thus the transition is symmetry forbidden.[18] However, vibrational modes can allow this transition to be observed between 300-325 nm, though very weakly. Therefore, the primary transfer that is observed is the $S_0 \rightarrow S_2$ transition at ~270 nm with retained vibrational fine-structure.

While all the species had a maxima at ~280 nm, the coumarin was interesting as it also exhibited a longer wavelength, lower energy, band up to 350 nm. This implies that the HC01-2Cou system can be excited by lower energy radiation at up to 350 nm. This would allow the 4-methylcoumarin to be selectively excited in the presence of tyrosine or tryptophan residues, making it a versatile addition to these natural building blocks.

5.3 EMISSION

The emission profiles of the three species; HC01-2Y, HC01-2Nap and HC01-2Cou, can be seen in figure 6. The profile of tyrosine is consistent with the mirror image of the relatively featureless absorption with a maximum at 302 nm. However, there appears to be a broad long-wavelength emission at 408 nm, which is possibly due to the presence of tyrosine phosphorescence or the formation of a tyrosine excimer which could be favoured by the closely

packed chromophores at the N-termini.[19] Indeed, this phenomenon is most closely modelled by polytyrosine,[20] indicating that in the HC01-2Y system there are multiple tyrosine residues in close proximity. The coumarin emission when excited at 320 nm is also broad and featureless, with a maximum at 392 nm. As with the absorption profile the naphthalene emission exhibits very clear fine structure leading to two emission bands with maxima at 334 and 326 nm. This is due to retention of vibrational fine structure for naphthalene as hydrophilic solvent broadening is not very efficient for the hydrophobic hydrocarbon. The excitation spectra for these species (*see* S49, S59 and S63) clearly indicates the chromophoric source and reinforced the lower energy capability to excite emission for HC01-2Cou.

The addition of Lu(III) ions to any of these systems resulted in only a slight decrease in the overall emission and very little change in the profile, except for HC01-2Nap. The naphthalene functionalised system appears to have the features of an excimer with broadened emission up to 470 nm. Upon the addition of metal the emission classically from the monomer at 334 and 326 nm decreases and the emission from an excimer type species at 390 nm increases.[21] The increase in an excimer is to be expected as the binding of the metal ion will result in a tighter

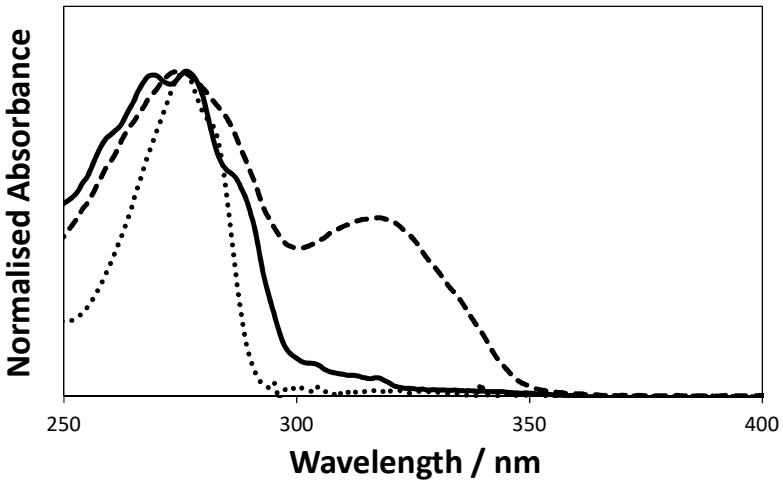

Figure 5: The normalised absorption spectra of HC01-2Y (dotted line), HC01-2Nap (solid line) and HC01-2Cou (dashed line). Samples were 0.5 mM peptide monomer, in 7M aqueous urea.

assembly at the N-termini. Interestingly, all the native emission profiles cover a similar wavelength range suggesting that the same energy transfer mechanisms that operated between tryptophan and the lanthanides may work effectively. Indeed, this is expected considering the literature for naphthalene, coumarin and tyrosine sensitising Tb(III) successfully.

Table 1: Summary of absorption and emission features for HC01-2Y, HC01-2Nap and HC01-2Cou both native and metallated with Lu(III) ions.

	Absorbance λ_{max} / nm	**Emission λ_{max} / nm**
HC01-2Y		
Native	277	302, 408
Metallated	-	304, 408
HC01-2Nap		
Native	288, 277, 270, 259	326, 334, 351, 390
Metallated	-	326, 334, 351, 390
HC01-2Cou		
Native	319, 276	392
Metallated	-	392

5.4 TERBIUM LUMINESCENCE TITRATIONS

The use of all three systems to sensitise Tb(III) emission was quickly ascertained and utilised to carry out titrations (Fig. 7). This allowed the binding event to be monitored as the Tb(III) emission could be used as a fluorescent informer on the localised proximity to the sensitising chromophore. Indeed, due to the preformed nature of the HC01 derivatives it is difficult to measure the binding event by circular dichroism (*see* Chapter 4) because of the small change in folding whereas the fluorescence signal is much more distinct.

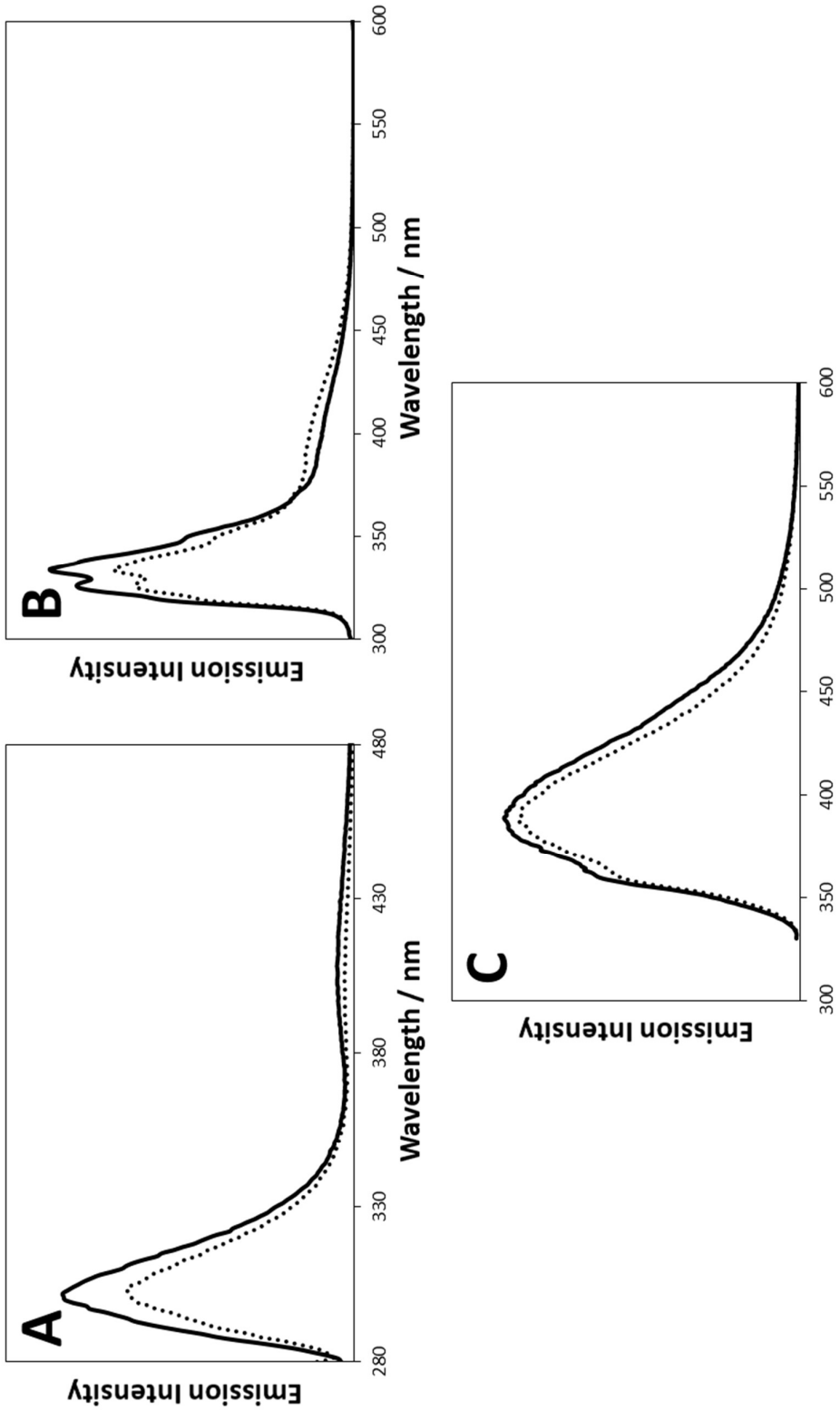

Figure 6: The emission spectra for A) HC01-2Y, B) HC01-2Nap and C) HC01-2Cou with both native (solid line) and metallated (dotted line) samples. Samples were peptide monomer (90 μM) in aqueous HEPES (10 mM, pH 7.0), metallated samples involved the addition of Lu(III) (35 μM). The excitation wavelength was 280 nm for all samples.

The binding isotherm for HC01-2Y with Tb(III) clearly demonstrated the presence of a single distinct binding site requiring three peptide monomer units to form a trimer. This was illustrated by the increase in Tb(III) emission up to one equivalence of the lanthanide per trimer unit. There is a noticeable continuation in the increase of terbium emission upon the further addition of Tb(III) ions. This suggests that the previously discussed external sites are active in solution (*see* Chapter 3). Indeed, it is likely that the only reason these sites were not observed from the emission of the tryptophan functionalised HC01 is due to the very high emission intensity sensitised from the designed site masking the external site emission. This also suggests that the tyrosine of HC01-2Y is a much weaker donor than the tryptophan of HC01, as previously observed.[2]

Interestingly, the binding isotherms for both HC01-2Nap and HC01-2Cou with Tb(III) had much clearer profiles due to the strong sensitisation afforded by the naphthalene and coumarin moieties when compared with the weakly sensitising phenol of tyrosine. In both cases it was apparent that one equivalence of Tb(III) ions per trimer unit was required to satisfy the binding site. A clear plateau was reached after one equivalents was added, unlike the tyrosine system. However, it is expected that the external binding are almost certainly still in place but, as with HC01, the sensitised effect is almost unnoticeable against the strong emission from the designed core binding site. All three of the binding profiles indicated that the peptide systems were behaving as designed, like the previously studied HC01.

The binding profiles for both HC01-2Nap and HC01-2Cou were fitted to a three-to-one equation (*see* Chapter 8.9) resulting in the determination of the binding constants (Fig. 7). The log of the binding constants for both HC01-2Nap and HC01-2Cou were both noticeably greater

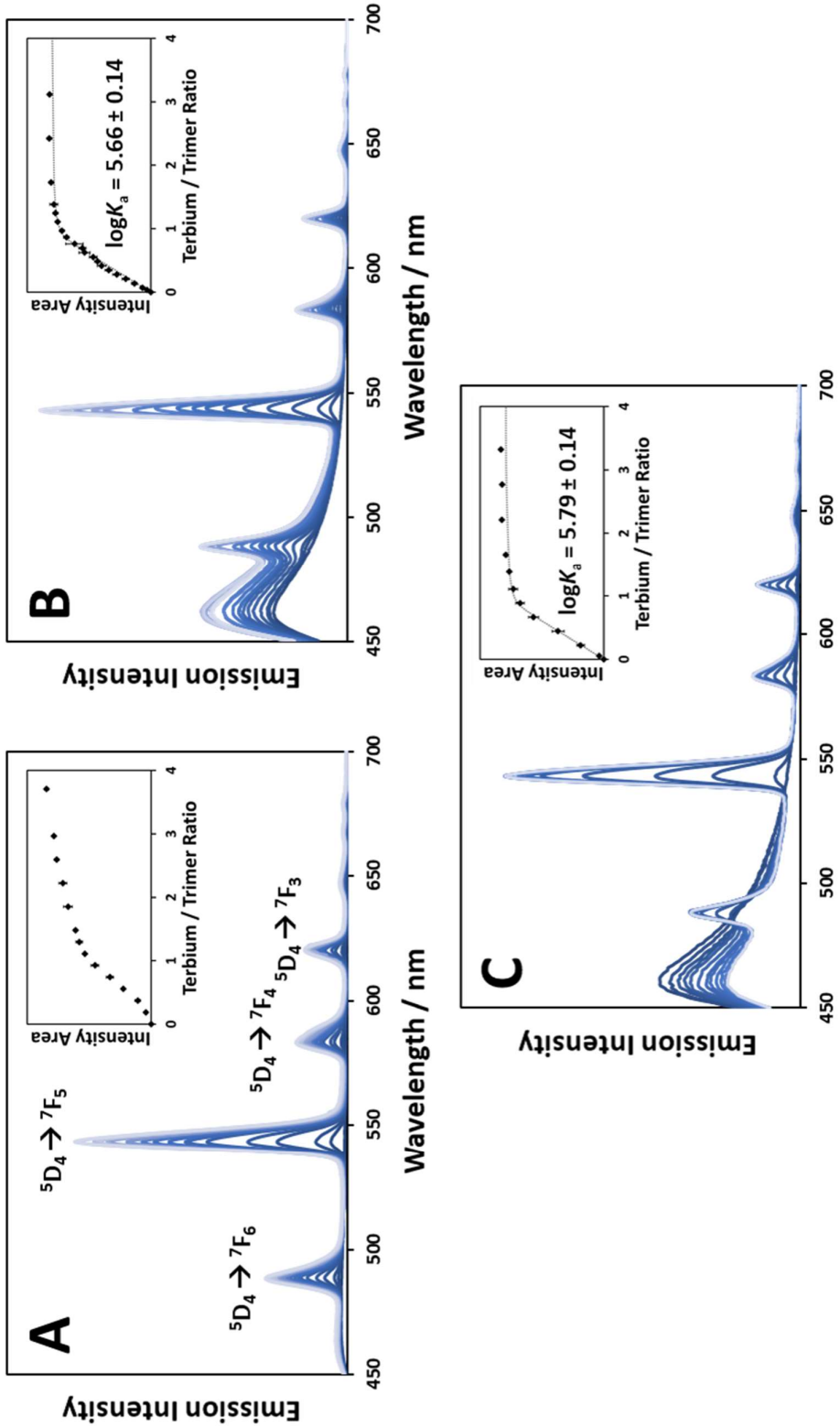

Figure 7: The emission of Tb(III) as it is titrated into A) HC01-2Y, B) HC01-2Nap and C) HC01-2Cou. The inserts indicate the emission intensity area of the peak at 545 nm against the ratio of Tb(III) to peptide trimer. The binding constant was determined from fitting the isotherms to a three-to-one equation. Samples were peptide monomer (90 μM) in aqueous HEPES (10 mM, pH 7.0). Error bars are the standard deviation of three repeats. Both HC01-2Y and HC01-2Nap were excited with 280 nm light while HC01-2Cou was excited with 320 nm light.

than that of HC01 with Tb(III) at 5.68 ± 0.14 and 5.80 ± 0.14 respectively when compared to 5.03 ± 0.13 for HC01. This implies that the site may be more preformed as there is little difference between the binding sites across the HC01 derivatives except for the N-terminal positioned chromophore residue. The larger and more hydrophobic naphthalene and coumarin sidechains would result in closer packing at the N-terminus, triggering binding site preformation and thus leading to the more feasible binding energies associated with the HC01-2Nap and HC01-2Cou peptides. The HC01-2Y data was not fitted to this model due to the clear presence of emission artefacts, most probably due to competitive binding to external sites throughout the titration.

5.5 LANTHANIDE SENSITISATION

After the successful sensitisation of Tb(III) from the various chromophores it remained to assess their ability to sensitise the other visible emitting lanthanides such as Dy(III), Sm(III) and Eu(III). Considering HC01-2Y it was apparent that, in addition to Tb(III), the phenol donor was able to sensitise emission from both Eu(III) and Dy(III) ions resulting in their distinct emission profiles (Fig. 8A). This is the first reported instance of tyrosine sensitising emission from Dy(III) and is an exciting addition to the photophysical potential of this natural amino acid. Attempts to sensitise Sm(III) emission were unsuccessful resulting in no greater enhancement in signal than aqueous Sm(III) ions. It should be noted that the intensity of the emission from each lanthanide did not indicate the efficiency of energy transfer as the non-radiative deexcitation pathways can differ between species.

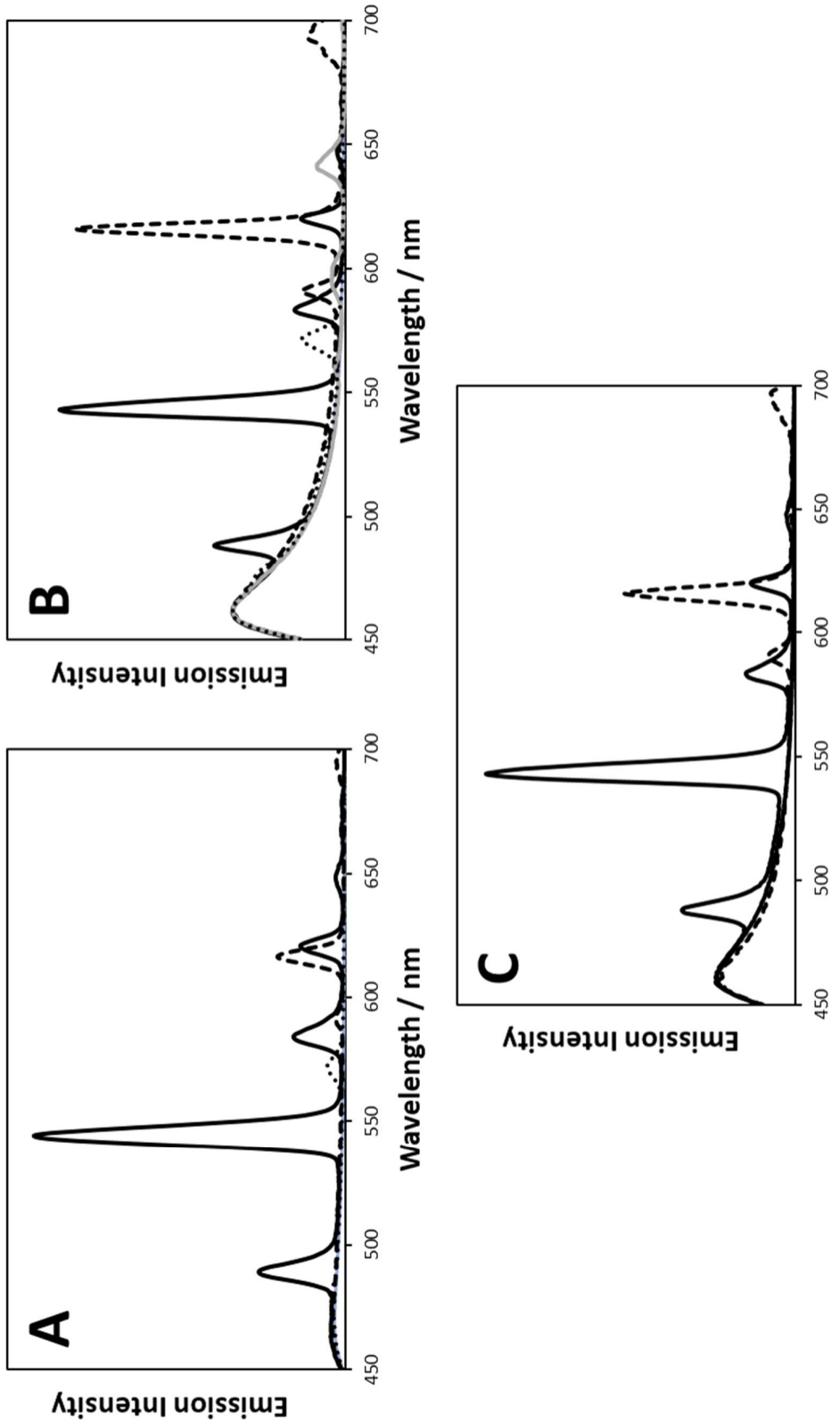

Figure 8: The emission of visible emitting lanthanides Tb(III) (black solid line), Eu(III) (blacked dashed line), Dy(III) (black dotted line) and Sm(III) (grey solid line) in A) HC01-2Y, B) HC01-2Nap and C) HC01-2Cou. Samples were peptide monomer (90 μM) in aqueous HEPES (10 mM, pH 7.0) with respective Ln(III) (35 μM). Both HC01-2Y and HC01-2Nap were excited with 280 nm light while HC01-2Cou was excited with 320 nm light.

The naphthalene moiety was clearly able to sensitise all four of the visible emitting lanthanides as exhibited by the strong emission profiles of these species. The emission of Eu(III) is particularly strong considering it is known to be hampered by dominant non-radiative quenching mechanisms. This observation therefore implies that the transfer efficiency to Eu(III) may be particularly high.

When considering HC01-2Cou, the coumarin is able to sensitise both Tb(III) and Eu(III). This is very exciting as the lanthanides were excited via energy transfer from the 4-methylcoumarin pumped at 320 nm, which though still considered very high energy radiation for biological systems, it is considerably lower than the 280 nm radiation used for the other systems in this thesis. The excitation spectra of both Tb(III) and Eu(III) bound in HC01-2Cou (*see* S60), indicates that sensitisation can be carried out up to 350 nm excitation wavelength. However, attempts to sensitise Dy(III) and Sm(III) proved unsuccessful. This may seem counter-intuitive due to the literature example demonstrating successful energy transfer.[14] Nevertheless, it must be noted that the successful complex had direct orbital overlap between the coumarin and the lanthanide species enabling different energy transfer processes such as Dexter electron transfer to dominate. The HC01 system and its derivatives make use of a through space energy transfer mechanism due to the relatively long distances involved. All similar mechanistic systems have not reported energy transfer to Dy(III) or Sm(III).[11, 12, 13, 15]

On a side note, the Eu(III) emission of both HC01-2Cou and HC01-2Nap have characteristic emission at 616 nm from the hypersensitive $^5D_0 \rightarrow {}^7F_2$ transition. This indicates that the Eu(III) is bound in a low symmetry environment as previously observed (*see* Chapter 3).[22] This is

further supported by the solid state structure for HC01-2Nap that exhibits Tb(III) bound in a low symmetry site.

5.6 LANTHANIDE LIFETIMES

The lifetime of the lanthanide emission was recorded for all the sensitised ions in each system to infer details about the local environments of the metal ions. A literature set of aqueous lifetimes are included to compare totally solvent exposed lifetimes with the given results.[23]

All the studied systems exhibited considerably longer lifetimes for each bound ion when compared with the aqueous samples indicating that the metal ion is located in a more solvent protected site. The exception to this is the particularly short lifetime of 507.9 ± 1.5 μs for Tb(III) bound by HC01-2Nap. This suggests the presence of a localised quenching species, however, it is unlikely to be water due to the lack of noticeable quenching for the other bound lanthanides. Therefore the only alternative is the process of quenching by the naphthalene donor via a back transfer mechanism. Indeed, the phenomenon of back transfer to a donor

Table 2: The lifetimes for bound lanthanide ions in each of the peptides. The emission wavelength monitored for Tb(III) was 545 nm, Eu(III) was 616 nm, Dy(III) was 573 nm and Sm(III) was 643 nm. The literature aqueous lanthanide lifetimes are included for comparison.[23] Samples were peptide (90 μM) aqueous HEPES buffer (10 mM, pH 7.0) and lanthanide trications (10 μM). Reported errors are the standard deviation of three repeats. Both HC01-2Y and HC01-2Nap were excited with 280 nm light while HC01-2Cou was excited with 320 nm.

Lanthanide	HC01-2Y Lifetime / μs	HC01-2Nap Lifetime / μs	HC01-2Cou Lifetime / μs	Aqueous Lifetime / μs
Samarium	-	26.97 ± 0.12	-	2.69 ± 0.01
Europium	738.8 ± 9.0	886.3 ± 1.3	803.3 ± 5.7	112.4 ± 0.7
Terbium	1327 ± 17	507.9 ± 1.5	1209 ± 31	430 ± 20
Dysprosium	25.34 ± 0.28	21.87 ± 0.42	-	2.49 ± 0.02

ligand has been observed for multiple lanthanide luminescent systems.[24] As previously discussed, the longer lifetimes observed for the lanthanides bound in these systems would be suitable for time-gated experiments. Even the shorter lifetime of the HC01-2Nap bound Tb(III) would still be appropriate adding further redundancy to labelled time-gated experiments.

5.7 HYDRATION STATE

The lanthanide lifetime results demonstrated that dramatic quenching was occurring for Tb(III) bound in the HC01-2Nap system. In order to assess whether the quenching was due to water molecules a lifetime experiment was recorded for each peptide with Tb(III) bound in the presence of both water and deuterium oxide. The resultant changes in lifetime between the deuterated and hydrated water allow for the determination of the number of water molecules in the first coordination sphere (*see* Chapter 3).[25] The results are summarised in table 3, with an aqueous control to determine the number of water molecules around solvated Tb(III) ions.

For HC01-2Y there was a single water in the first coordination sphere explaining the slight quenching observing in the hydrated lifetimes. The phenol sidechain of tyrosine is smaller than the indole of the previously used tryptophan, therefore it is likely to be less effective at blocking solvent access to the N-terminal lanthanide binding site. This may set a lower limit for the size of the chromophore to adequately hinder water ingress to the binding site. It should also be noted that the hydroxyl group of phenol is likely to engage in hydrogen bonding to water, thus encouraging water in close proximity to the chromophore. Indeed, the hydroxyl- group of the phenol may be responsible for the quenching due to exchange during the experiment.

Table 3: The lifetimes of Tb(III) bound in the respective peptide systems with the determined number of water molecules in the first coordination sphere. The samples were peptide (90 μM), HEPES buffer (10 mM, pH 7.0), $TbCl_3$ (10 μM) in water or deuterium oxide. The reported errors are the standard deviation of three repeats. Both HC01-2Y and HC01-2Nap were excited with 280 nm light while HC01-2Cou was excited with 320 nm light.

Peptide	τ^{H}_{1} / ms	τ^{D}_{1} / ms	No. of Water Molecules
HC01-2Y	1327 ± 17	2120 ± 76	**1**
HC01-2Nap	507.9 ± 1.5	549.0 ± 2.0	**0**
HC01-2Cou	1209 ± 31	1805 ± 18	**1**
Aqueous	430 ± 20	2260 ± 70	**9**

Interestingly, the hydrated state of HC01-2Nap bound Tb(III) indicates that there are no water molecules within the first coordination sphere. This strongly supports the previously discussed back transfer from the excited state Tb(III) to the locally situated naphthalene species as there are no locally situated solvent molecules to account for the quenching. Indeed, upon a decrease in temperature others have reported that the Tb(III) lifetime increases, suggesting that the back transfer is via a thermally activated transfer from the excited metal to the naphthyl triplet state.[26, 6] This could be further studied by degassing the solvent to displace dissolved oxygen, which is well understood to quench triplet states.[27] The removal of dissolved oxygen would increase the lifetime of the triplet states, which in turn, would be expected to increase the lifetime of the Tb(III) emission. The use of flash photolysis to populate the naphthyl triplet state would allow the lifetime of this state to be measured both in the presence and absence of oxygen. This state would also act as a donor to bound Tb(III) ions and thus the presence of any long lived emission, with a rate of ~1000 s^{-1}, from the naphthyl moiety would suggest back transfer from excited Tb(III).[28]

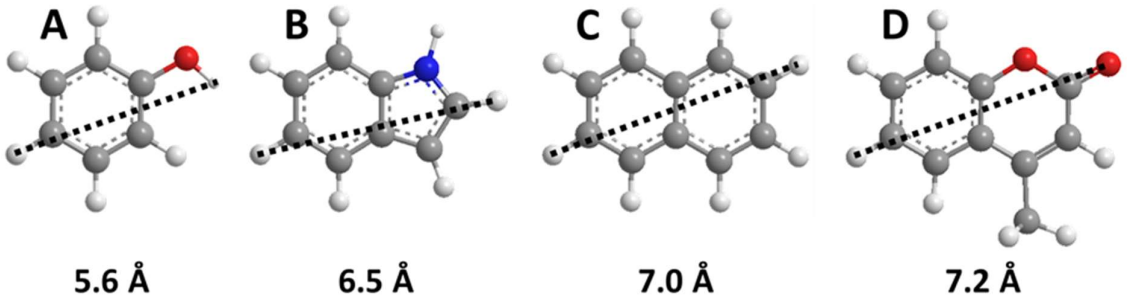

Figure 9: The molecular structures of A) phenol, B) indole, C) naphthalene and D) 4-methylcoumarin with the longest distance illustrated by a dashed black line, measured values given below.

The HC01-2Cou system exhibits a single bound water molecule in the Tb(III) first coordination sphere. Considering the larger size of the 4-methylcoumarin chromophore it is possible that the three sidechains do not pack well within the three stranded coiled coil. Therefore, leaving space for water ingress to the lanthanide binding site. Pairing this result with that observed for tyrosine and comparing the sidechains (Fig. 9) it is likely that there is an ideal size for blocking water access to the binding site. This would be expected to be larger than phenol but smaller than 4-methylcoumarin.

5.8 CONCLUSIONS

The use of diverse chromophoric amino acids to sensitise lanthanides in the optimised HC01 peptide system has been explored. None of the incorporated chromophores appear to hinder the lanthanide binding activity of the designed site. Indeed, luminescence titrations further supported the circular dichroism data (*see* Chapter 4) in showcasing that the incorporation of these new scaffolds retained, and in some cases improved, the binding activity. This demonstrates the versatility of the HC01 type backbone as an excellent protein model system to achieve an understanding of lanthanide photophysics. A ‘goldilocks’ zone for the size of the

chromophore to successfully prevent water molecules coordinating to a lanthanide in the core binding site was determined. While it is probable that the phenol was too small to block the N-terminus sufficiently, it is likely that the 4-methylcoumarin sidechain was too large to pack efficiently thus allowing water ingress.

The use of tyrosine to sensitise emission from Dy(III) ions was demonstrated for the first time, expanding the photophysical versatility of this natural amino acid. Further to this, a study of lanthanide sensitisation by both HC01-2Nap and HC01-2Cou was achieved, demonstrating that the naphthalene functionalised amino acid has the capability to sensitise the visible emitting lanthanides; Tb(III), Eu(III), Dy(III) and Sm(III). The 4-methylcoumarin amino acid was able to sensitise emission from both Eu(III) and Tb(III) requiring excitation at 320 nm (up to 350 nm) rather than the more damaging 280 nm radiation used for the other systems. Indeed, this is a significant improvement, though still damaging for biological samples. As with tryptophan (*see* Chapter 3), the toolbox of FRET pairs has been expanded allowing a greater variety of amino acids to be utilised with lanthanide probes.

5.9 REFERENCES

[1] i) W. DeW. Horrocks and A. P. Snyder, *Biochem. Biophys. Res. Commun.*, 1981, **100**, 111-117; ii) W. DeW. Horrocks and W. E. Collier, *J. Am. Chem. Soc.*, 1981, **103**, 2856-2862.

[2] J. Bruno, W. DeW. Horrocks, R. J. Zauhar, *Biochemistry*, 1992, **31**, 7016-7026.

[3] i) T. Hashimoto, N. Nakano and K. Hirao, *J. Chem. Phys.*, 1996, **104**, 6244-6258; ii) B. H. Jhun, K. Ohkubo, S. Fukuzumi and Y. You, *J. Mater. Chem. C*, 2016, **4**, 4556-4567; iii) Y. Y.

Pan, J. Huang, Z. M. Wang, D. W. Yu, B. Yang and Y. G. Ma, *RSC Adv.*, 2017, **7**, 26697-26703.

[4] J. R. Escabi-Perez, F. Nome and J. H. Fendler, *J. Am. Chem. Soc.*, 1977, **99**, 7749-7754.

[5] M. A. Mortellaro and D. G. Nocera, *J. Am. Chem. Soc.*, 1996, **118**, 7414-7415.

[6] P. J. Skinner, A. Beeby, R. S. Dickins, D. Parker, S. Aime and M. Botta, *J. Chem. Soc., Perkin Trans. 2*, 2000, **0**, 1329-1338.

[7] i) S. Bhattacharyya, M. B. Roy and S. Ghosh, *Chem. Phys.*, 2004, **300**, 295-304; ii) P. S. Sardar, S. Samanta, M. B. Roy, and S. Ghosh, *Mol. Phys.*, 2008, **106**, 827-840.

[8] i) W. –N. Wu, W. –B. Yuan, N. Tang, R. –D. Yang, L. Yan and Z. –H. Xu, *Spectrochim. Acta A*, 2006, **65**, 912-918; ii) N. K. Al-Rasbi, H. Adams and F. O. Suliman, *Dyes Pigm.*, 2014, **104**, 83-88.

[9] i) S. Comby, F. Stomeo, C. P. McCoy and T. Gunnlaugsson, *Helv. Chim. Acta*, 2009, **92**, 2461-2473; ii) E. Debroye, S. Laurent, L. V. Elst, R. N. Muller and T. N. Parac-Vogt, *Chem. Eur. J.*, 2013, **19**, 16019-16028.

[10] P. G. Tarassoff and N. Filipescu, *J. C. S. Chem. Comm.*, 1975, **0**, 208-209.

[11] M. T. Alonso, E. Brunet, C. Hernandez and J. C. Rodriguez-Ubis, *Tetrahedron Lett.*, 1993, **34**, 7465-7468.

[12] D. Kovacs, X. Lu, L. S. Meszaros, M. Ott, J. Andres and K. E. Borbas, *J. Am. Chem. Soc.*, 2017, **139**, 5756-5767.

[13] i) E. Heyduk and T. Heyduk, *Anal. Biochem.*, 1997, **248**, 216-227; ii) J. C. Rodriguez-Ubis, M. -T. Alonso, O. Juanes and E. Brunet, *Luminescence*, 2000, **15**, 331-340; iii) M. -T. Alonso, E. Brunet, O. Juanes and J. C. Rodriguez-Ubis, *J. Photochem. Photobiol. A*, 2002, **147**, 113-125; iv) A. K. R. Junker, L. R. Hill, A. L. Thompson, S. Faulkner and T. J. Sorensen, *Dalton Trans.*, 2018, **47**, 4794-4803.

[14] O. Guzman-Mendez, F. Gonzalez, S. Bernes, M. Flores-Alamo, J. Ordonez-Hernandez, H. Garcia-Ortega, J. Guerrero, W. Qian, N. Aliaga-Alcalde and L. Gasque, *Inorg. Chem.*, 2018, **57**, 908-911.

[15] J. Andres and A. –S. Chauvin, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15981-15994.

[16] J. M. Antosiewicz and D. Shugar, *Biophys. Rev.*, 2016, **8**, 151-161.

[17] i) R. H. Goodwin and B. M. Pollock, *Arch. Biochem. Biophys.*, 1954, **49**, 1-6; ii) R. H. Abu-Eittah and B. A. H. El-Tawil, *Can. J. Chem.*, 1985, **63**, 1173-1179.

[18] H. Spooner and G. P. Nordheim, *Discuss. Faraday Soc.*, 1950, **9**, 19-26.

[19] i) T. Keleti, *FEBS Lett.*, 1970, **7**, 280-282; ii) A. C. Samuels, J. O. Jensen and H. F. Hameka, *J. Mol. Struct. Theochem*, 1998, **454**, 25-30.

[20] E. Yeargers, F. R. Bishai, L. Augenstein, *Biochem. Biophys. Res. Commun.*, 1966, **23**, 570-575.

[21] i) K. Uchida, M. Tanaka and M. Tomura, *J. Lumin.*, 1979, **20**, 409-414; ii) M. Pabst, B. Lunkenheimer and A. Kohn, *J. Phys. Chem. C*, 2011, **115**, 8335-8344.

[22] i) K. W. Lei, W. S. Liu and M. Y. Tan, *Spectrochim. Acta A*, 2007, **66**, 118-125; ii) K. Binnemans, *Coord. Chem. Rev.*, 2015, **295**, 1-45.

[23] S. Lis, T. Kimura and Z. Yoshida, *J. Alloys Compd.*, 2001, **323**, 125-127.

[24] M. Latva, H. Takalo, V. –M. Mukkala, C. Matachescu, J. C. Rodriguez-Ubis and J. Kankare, *J. Lumin.*, 1997, **75**, 149-169.

[25] i) A. Beeby, I. M. Clarkson, R. S. Dickins, S. Faulkner, D. Parker, L. Royale, A. S. Sousa, J. A. G. Williams and M. Woods, *J. Chem. Soc. Perkin Trans. 2*, 1999, **3**, 493-504; ii) R. M. Supkowski and W. DeW. Horrocks, *Inorg. Chim. Acta*, 2002, **340**, 44-48.

[26] A. Beeby, D. Parker and J. A. G. Williams, *J. Chem. Soc., Perkin Trans. 2*, 1996, **0**, 1565-1579.

[27] C. Grewer and H. -D. Brauer, *J. Phys. Chem.*, 1994, **98**, 4230-4235.

[28] J. A. G. Williams, PhD Thesis, University of Durham, 1995.

Chapter 6

EMISSION PREDICTION OF SMALL POLYAROMATIC HYDROCARBONS

6.1 INTRODUCTION

The work presented in this chapter is collaborative and was started alongside C. Soto an Erasmus student and M. Taylor a PhD student in the Worth group. The computational model was extensively improved by M. Taylor and further details can be found in his thesis.[1]

The photophysical properties of a molecule have been used to characterise and prepare functional materials such as fluorescent probes,[2] optical amplifiers and photocatalysts.[3] Absorption and emission phenomena are crucial for understanding many fundamental processes and can play an important role in reaction mechanism elucidation.[4] Interestingly, the absorption of light by molecules has been modelled extensively; indeed, there are many excellent programmes and packages of software to predict the absorption spectra of organic species with varying levels of complexity.[5] However, the modelling of emission spectra appears to be limited to small studies, generally involved in determining specific photophysical effects. Indeed, there are many examples of detailed and explicit emission prediction or fitting of experimental spectra to analyse specific states and photophysical phenomena.[6] Nevertheless, the methodologies presented in these publications are time consuming, demanding on processing capabilities and require a great deal of input from the user to give viable results. This is clearly not appropriate for rapid analysis of large libraries of compounds. While density functional theory based predictions are routinely used these tend to result in featureless Gaussian profiles as vibrational modes are not considered. Thus the absorption and especially the emission properties of molecules are often neglected as a means to characterise a species, despite the fact that it can provide useful insights into molecular structure, stability, dynamics and reactivity. But while the diagnostic features of emission and absorption spectra for

molecular identification are often neglected, it is usual to have a spectroscopic requirement for a new material as in the Forster resonance energy transfer materials discussed previously (*see* Chapter 3). Fortunately it is much less challenging to model the photophysical properties of selected structures than to attempt to select properties and design back to a structure. Therefore it should be possible to design and use a methodology that would allow rudimentary screening of a large number of molecules for their expected emission profiles.

Considering the importance of emission prediction for the development of energy transfer pairs, emissive materials or even molecular recognition it seemed necessary to prepare a methodology to provide fast, yet reliable emission prediction results. To this end, a simple workflow for these predictions has been prepared and the viability of the methodology tested using a series of simple aromatic species. The intention was two-fold, firstly that the method would not be demanding with respect to processing power and secondly to retain vibrational fine structure, as this is particularly important for spectral overlap calculations.

6.2 MODELLING ABSORPTION AND EMISSION PROCESSES

The theory of absorption and emission can be simply described by considering the behaviour of an electron residing in the highest occupied molecular orbital of a chromophore. Upon excitation with light, the molecule will absorb a photon resulting in the electron being promoted from the ground state into a higher energy excited state (Fig. 1). This process is directly responsible for the reduction in transmission monitored to gain an absorption spectrum. The excited state is unstable and therefore the electron will return to the ground state by a

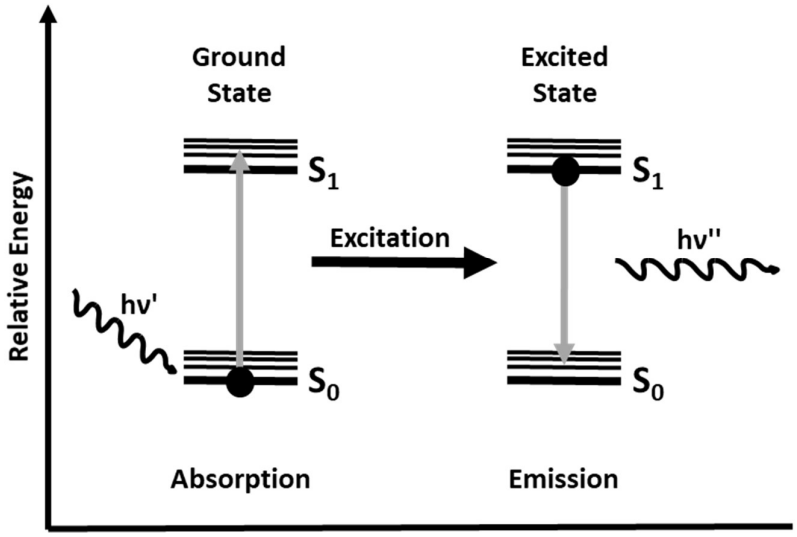

Figure 1: A Jablonski diagram illustrating the absorption and emission processes of a molecule with the absorption of a high energy photon (hv') promoting an electron (black dot) from the ground state (S_0) into an excited state (S_1). The relaxation of the excited electron back to the ground state (S_0) results in the emission of a lower energy photon (hv'').

combination of radiative (the process of emitting a photon) and non-radiative pathways. Thus, the emission process can be considered as almost the reverse of that of the absorption.

The approach to predicting an emission spectrum is not dissimilar to that of an absorption spectrum as both require the use of quantum dynamic methods. The prediction of an absorption spectrum can be summarized by a series of smaller problems (Fig. 2). The following is a short description of the process utilised by M. Taylor in the modelling of small molecule spectra.[1] Starting with a geometry optimised species of interest, this model made use of a simplified vibronic coupling Hamiltonian requiring only the Hessian matrices of the atomic Cartesian coordinates at the ground state (S_0) equilibrium geometry and Franck-Condon point, resulting in dramatically reduced processing times. However, it is recognized that this sacrifices the usual rigor of the explicit calculation involving the fitting of a series of points along desired vibrational normal modes, which captures any anharmonicity. In the interests of processing requirements, the move away from explicit fitting was acceptable but this may also result in

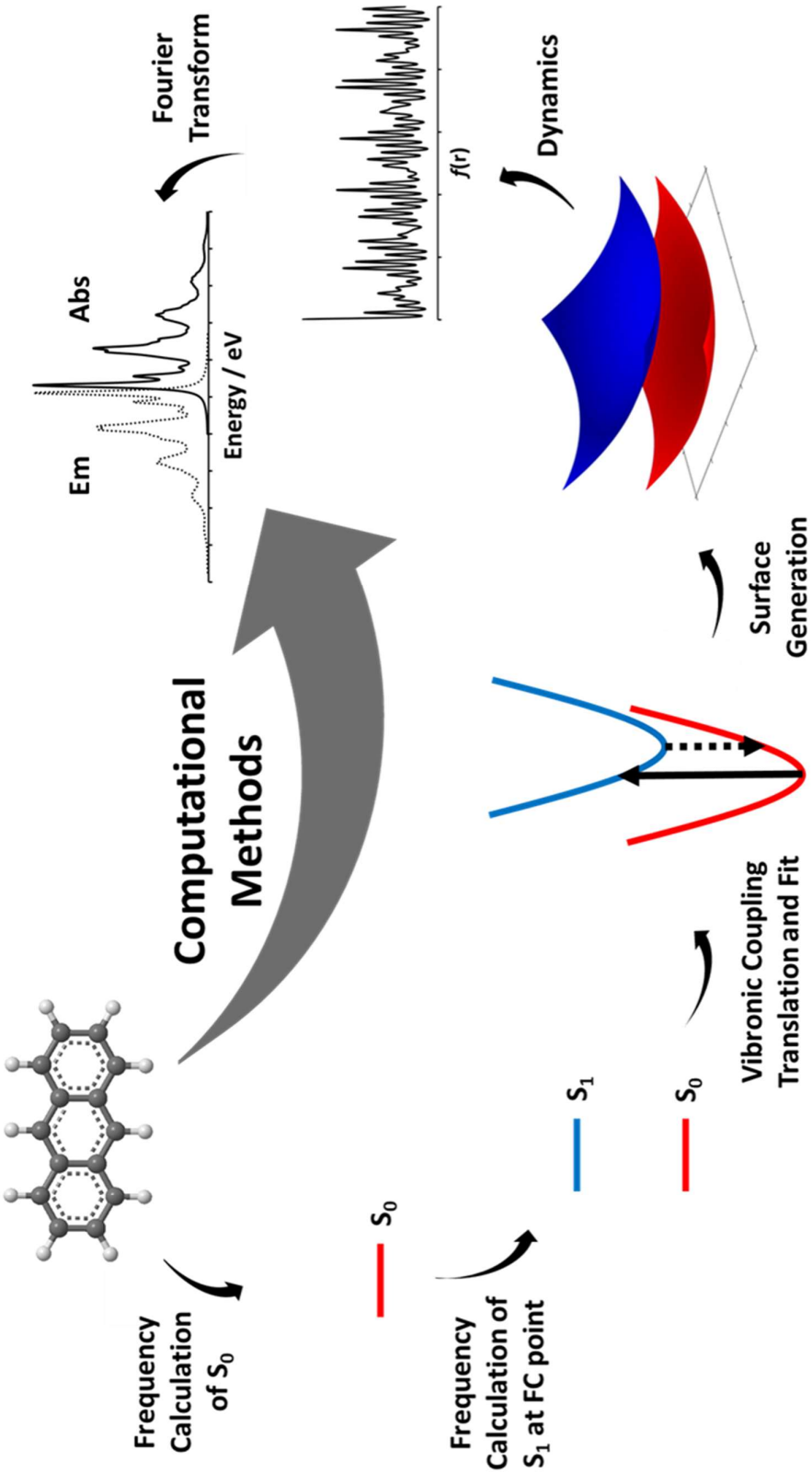

Figure 2: A schematic representation of the processes required to predict either the absorption or emission properties of a molecules. This involves starting with a geometry optimised molecule to generate ground state and excited state orbitals. Including vibrational modes allows the generation of surfaces. The process of modelling absorption or emission then depends upon the selected dynamics run on the surfaces. The resultant autocorrelation function f(r) can undergo a Fourier transform to give the respective spectra.

some vibrational modes being ignored alongside vibronic coupling to other excited states. The use of symmetry arguments, in agreement with the principles of Fermi resonance, also enabled further simplification with only the linear coupling constants along totally-symmetric modes having non-zero matrix elements.[7] This dramatically reduced the number of vibrational modes that had to be considered further simplifying the process. Simulated spectra are then obtained through wave-packet propagation using the powerful multiconfigurational time-dependent Hartree (MCTDH) algorithm used routinely for wave-packet propagation calculation.[8] These Gaussian wave-packets initiate from excitation to the Frank-Condon point on the S_1 state and are then allowed to propagate for 300 femtoseconds providing information about the wave-packet dynamics. The resultant autocorrelation function, which is the overlap between the initial and propagated wave-function, can undergo a Fourier transform to yield the absorption spectrum. The required quantum dynamics method and algorithm have been discussed at length elsewhere.[8]

To predict the emission spectrum the wave-packet is first relaxed on the excited state surface from the Frank-Condon point to the excited state minima. This process provides the ground state vibrational wave-function in the excited state position, requiring relaxation to the ground state in order to generate the emission spectrum. This is achieved by de-excitation of the wave-packet vertically downwards onto the ground state surface. As before a Fourier transform of the autocorrelation function provides the more familiar emission spectrum. In order to correct the energy of the emission, the final energy when relaxing on the excited state surface was subtracted from the result. This energy is further scaled by calculating the difference between the S_1 minima calculated in the gas phase and in solution, using an implicit dielectric constant to match the experimental solvent.[9] In all cases, the computational spectra were broadened to

allow accurate comparison with the experimental data by multiplying with a damping function to simulate the dissipation of oscillation energy. This work made use of the highly accessible QUANTICS suite of programmes opening up the potential for free and accessible software preparation.[8] Though it should be noted that all electronic structure calculations employed the use of Gaussian09.[10] All the calculations made use of TD-DFT and B3LYP levels of theory using a 6-31G** basis set.

6.3 VALIDATION WITH AROMATIC PROGRESSION

In order to determine the utility of this methodology and to ensure that the simplifications did not adversely affect the results, it was necessary to validate the computational spectra against experimental data. The initial study focused upon an aromatic progression of linear complexity from benzene to pentacene. This was due to well-understood photophysical behaviour of these species as well as the extensive computational work to determine the symmetry and vibrational operators, which facilitated the process of testing the method postulated here. Indeed, the computational work carried out on these species meant that problems could be relatively easily analysed through the methodology. As the aromatic systems increases in size, the emission energy for the $S_0 \leftarrow S_1$ transition decreases due to the narrowing of the energy gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital. Indeed, this relationship has been studied extensively by E. Clar alongside the ideas of aromatic dilution.[11]

Overall the progression was modelled successfully with a steady decrease in emission energy of the $S_0 \leftarrow S_1$ transition from benzene to pentacene. There was a clear trend in the experimental

position of the $S_0 \leftarrow S_1$ transition for anthracene, tetracene and pentacene that was mimicked by the computational prediction (Fig. 3). These systems agreed remarkably well with the experimental data having only a small deviation in the energetic position. This deviation to lower energy is to be expected from computational methods, which consider molecules in an optimised gas-phase model, but at low temperatures. The experimental use of solvent and room temperature conditions can explain this deviation. The vibrational fine-structure is also broadened in the experimental results due to thermal vibration of bonds and solvent interaction which is a noticeable feature in all the comparisons.[12] The fine-structure present in the computational spectra can be reproduced by including only a small number of vibrational modes. These totally symmetric modes exhibit large displacements in the excited state minima compared to the ground state. The excellent reproducibility of the spectra with the absence of modes with different symmetry is indicative of little vibronic coupling between the states influencing the electronic environment.

For both benzene and naphthalene the long wavelength transition ($S_1 \leftarrow S_0$) is forbidden by electric dipole selection rules due to the symmetrical nature of the naphthalene. This results in no change in the electronic dipole upon exciting the lowest lying $\pi \rightarrow \pi^*$ transition, implying that it is symmetry forbidden.[13] However, the transition can be weakly observed due to Herzberg-Teller and vibronic coupling resulting in the nuclear positions being considered in the potential electronic overlap.[14] Despite no consideration of such coupling computationally, the model is still able to successfully reproduce the main vibrational progression from a ring breathing mode. The lack of vibronic coupling is evident in the simulated naphthalene spectrum and the electronic structure calculations fail to accurately reproduce the energies of the S_1 and S_2 states.

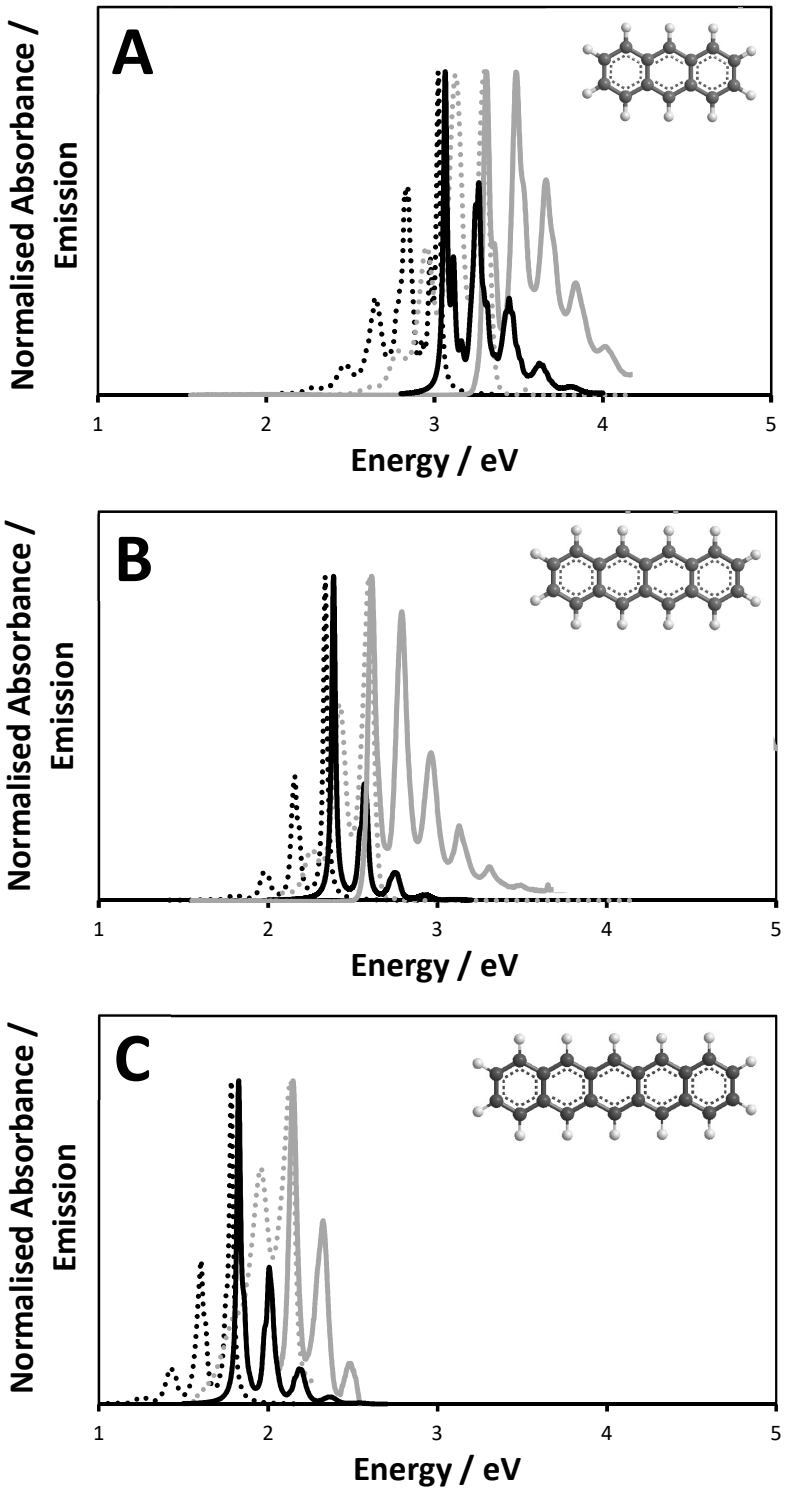

Figure 3: The computationally predicted emission (black dotted line) and absorption (black solid line) spectra for A) anthracene, B) tetracene and C) pentacene. This can be compared with the experimentally measured emission (grey dotted line) and absorption (grey solid line) spectra for samples of the respective chromophores (5 μM) in hexane.

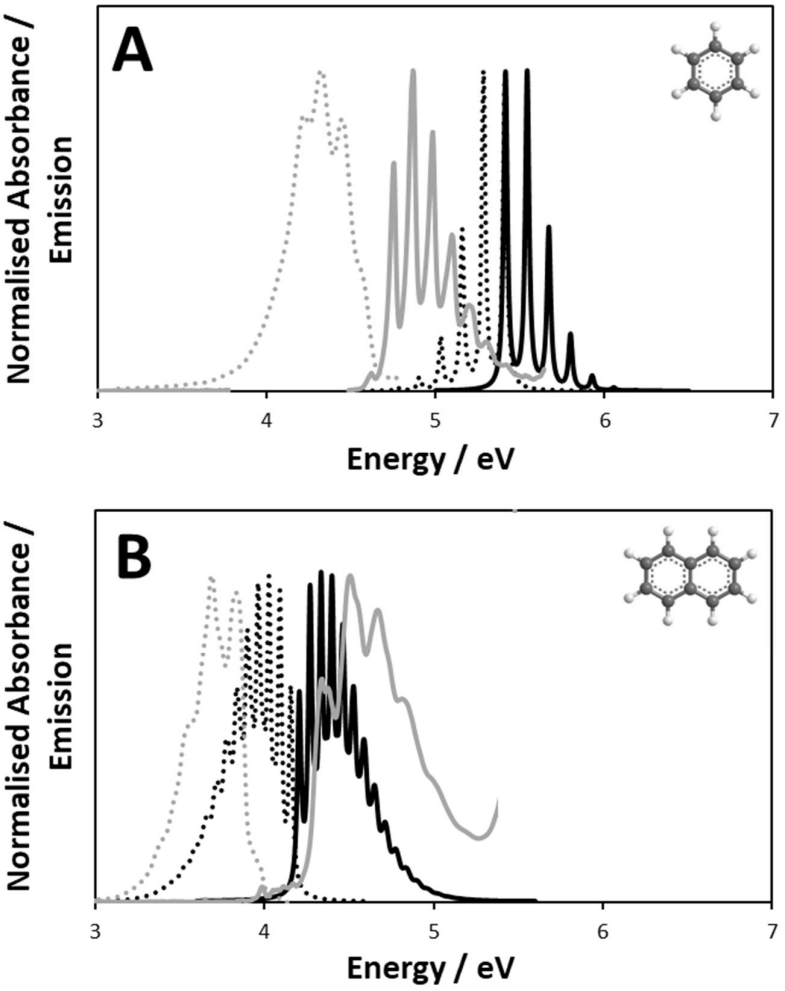

Figure 4: The computationally predicted emission (black dotted line) and absorption (black solid line) spectra for A) benzene and B) naphthalene. This can be compared with the experimentally measured emission (grey dotted line) and absorption (grey solid line) spectra for samples of the respective chromophores (5 μM) in hexane.

6.4 VALIDATION WITH SIMPLE ANTHRACENE DERIVATIVES

Following the study of sequential aromatics the impact of simple functional groups was explored utilising anthracene as a parent structure. Initially this involved two modifications to the anthracene structure, the addition of a methyl or hydroxymethyl group at the 9-position (Fig. 5). Considering the electron donating nature of these functional groups it would be

expected that the 9-methylanthracene would be more electron rich than the 9-hydroxymethylanthracene which, in turn, would be more electron rich than anthracene. This is because the hydroxyl group is electron withdrawing for the methyl group resulting in a reduction in its donating capabilities. Therefore, the 9-methylanthracene would require less energy to excite the electron rich ground state to the excited state, hence this transition would occur at a longer wavelength than that of 9-hydroxymethylanthrancene.

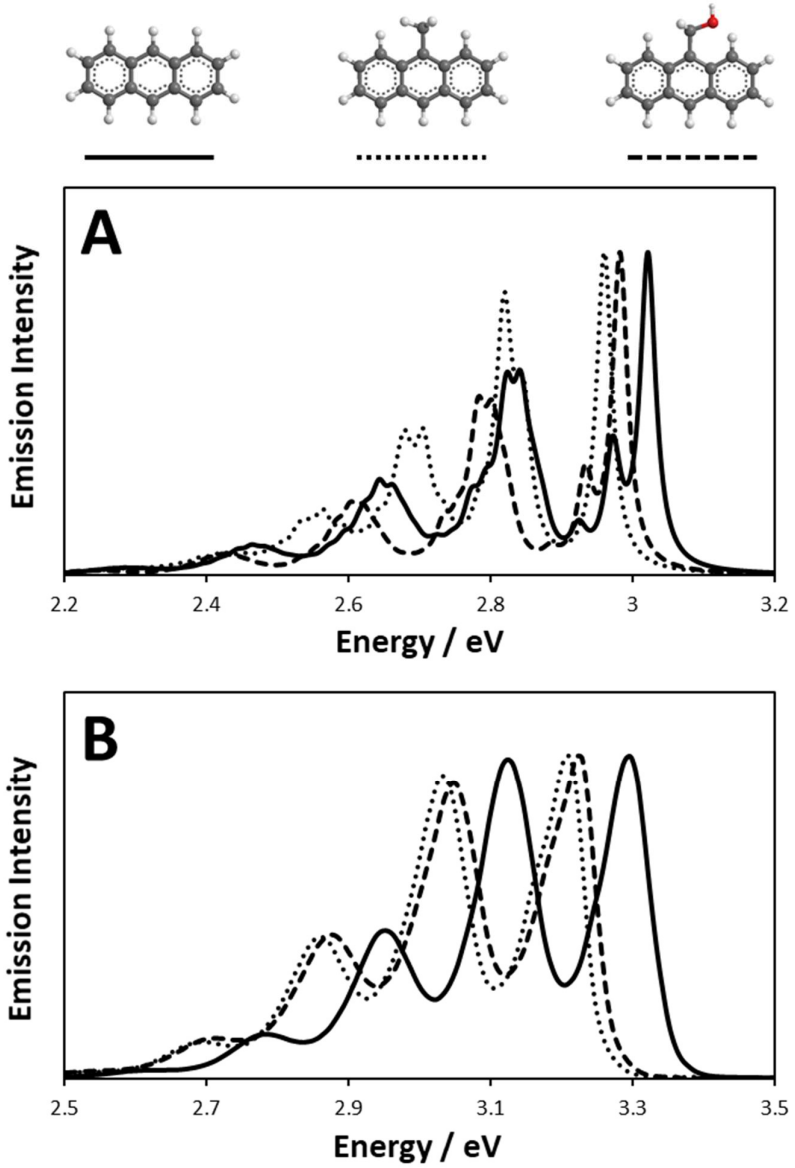

Figure 5: The emission spectra of anthracene (black solid line), 9-methylanthracene (black dotted line) and 9-hydroxymethylanthracene (black dashed line) for both the A) computational predicted and B) experimental determined results.

The experimental spectra indicated a shift to lower energy emission for both species, in agreement with the donating capability of both the functional groups. Pleasingly, the 9-methylanthracene $S_0 \leftarrow S_1$ transition was predicted at slightly lower energy than the 9-hydroxylmethylanthracene as expected. However, the differences between the two emission profiles were incredibly small. Therefore, it was with great excitement that the model successfully determined that the 9-methylanthracene would have a lower energy emission than the 9-hydroxymethylanthracene and that both would be lower than the unfunctionalised anthracene. As before, there was a slight energy difference between the computational and experimental results which has already been discussed.

6.5 – VALIDATION WITH SOLVENT SENSITIVE ANTHRACENE DERIVATIVES

The introduction of aldehyde and carboxylic acid functionalities at the 9-position resulted in extended π-systems, this was evidenced by a distinct red-shift to lower energy in the experimental emission profile due to the resultant increase in HOMO π-orbital energy and decrease in LUMO π*-orbital energy. This decrease in emission energy was also clearly evident in the computational predictions. Interestingly, the prediction suggested that the carboxylate functionality would result in the lowest energy emission, which would be expected when only considering resonance stabilisation but upon including solvent interaction with the functional groups this becomes more complicated. Both the two functionalities demonstrated key limitations of the methodology involving solvent effects.

Firstly, both the acid and aldehyde functional groups can interact with solvent directly resulting in the energy of the molecular orbitals to shift dependent upon the solvent environment. The

effect of solvent upon the emission profile due to more intrinsic interactions such as orbital overlap results in solvatochromic behaviour. However, to consider this would be far too demanding for a simple model and as such has not be carried out here. Secondly, understanding the intermolecular behaviour in different solvents is important when interpreting the results. For example, the carboxylic acid formed a dimer species in ethanol and water, but in basic water the dimer was broken due to generation of the anthracene carboxylate anion, resulting in

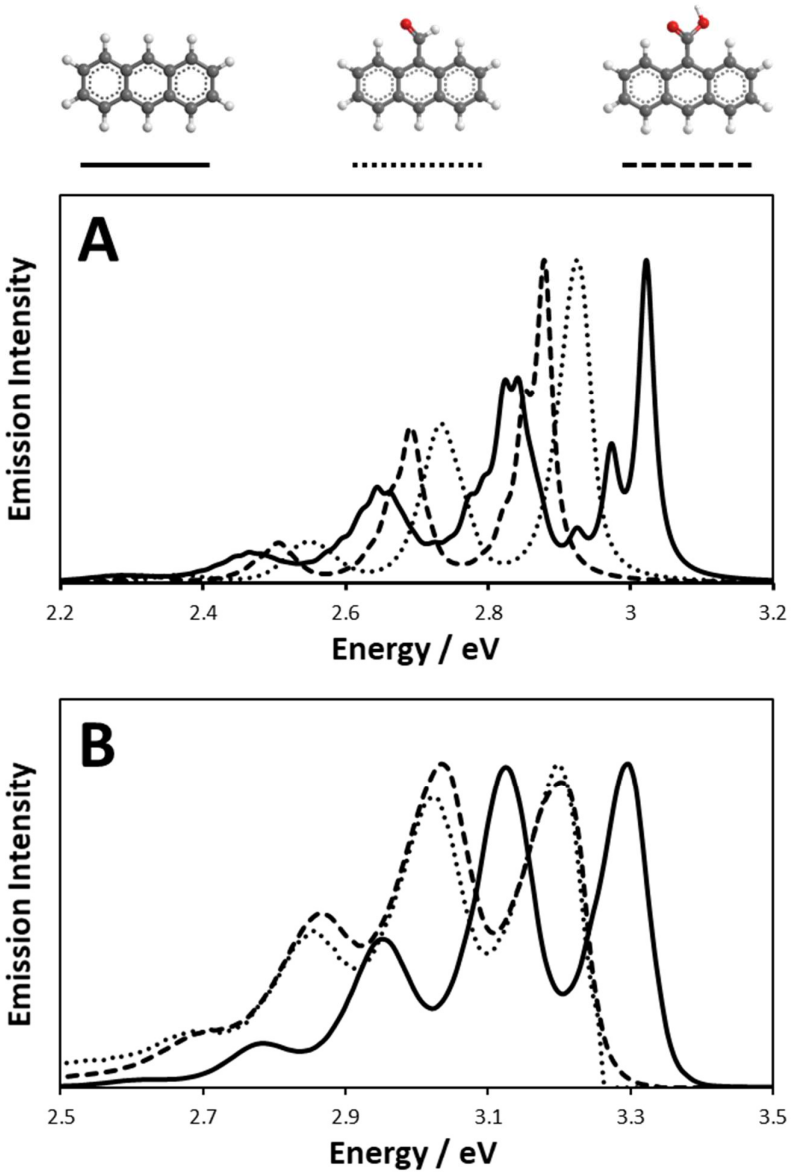

Figure 6: The emission spectra of anthracene (black solid line), 9-methylanthracene (black dotted line) and 9-hydroxymethylanthracene (black dashed line) for both the A) computational predicted and B) experimental determined results.

the loss of stabilising hydrogen bonds.[15] While the methodology employed herein failed to accurately predict the emission of the dimer species due to the model handling each molecule discreetly, the monomer was successfully determined. Considering the simplicity of this model it is unlikely to be able to include a feature to understand molecular assembly especially as this is an extensive domain of computational chemistry itself.

This demonstrates a key limitation of the methodology, with multiple aspect of solvent influence being ignored. However, it should be noted that many of these areas have their own domains within computational chemistry, due to their complexity, hence a simple model would not do them justice. The possibility of non-adiabatic effects and the energetic shifting of states due to solvent interactions are neglected due to the complexity of their consideration. Nevertheless, further calculations could be assigned to various functional groups in different solvents such as carbonyls or acids to give an approximation of shift and broadening effects. It should be noted that this is a key area requiring knowledge from the software user, thus while simplifications have been made this methodology still requires careful thought about the inputs, especially regarding potential solvent effects.

6.6 VALIDATION WITH HALOGENATED ANTHRACENE DERIVATIVES

The final validations were carried out using a series of anthracene species with halogens at the 9-position (Fig. 7). Theoretically all the halogens have a negative inductive effect, which results in electron density being drawn away from the aromatic system. However, the halogens also have electrons in the p-orbitals which can be donated into the aromatic system but this is

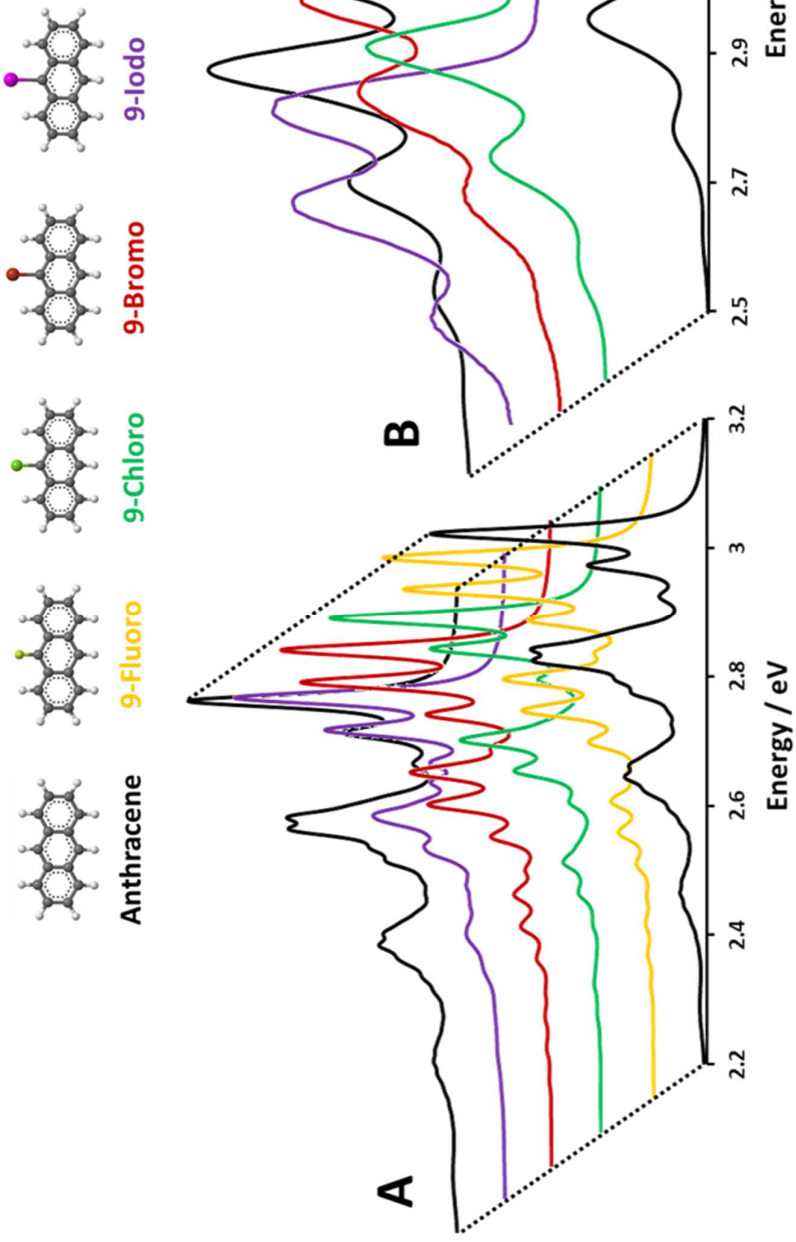

Figure 7: The A) computational predictions and B) experimentally measured emission spectra for the halogenated anthracene species. The spectra of anthracene (black line), 9-fluoroanthracene (yellow line), 9-chloroanthracene (green line), 9-bromoanthracene (red line) and 9-iodoanthracene (purple line) are illustrated. The samples for the experimental emission were the respective chromophores (5 μM) in hexanes. The spectra were normalised to allow direct comparison of the profile.

typically less effective. Therefore, the effect on the gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital would be expected to be an interplay between the electronegative and p-orbital overlap effects. With 9-fluoroanthracene it is expected that there would be poor p-orbital overlap but a very strong electronegative inductive effect removing electron density from the ring system. This would result in it being harder to excite from the highest occupied molecular orbital, hence more energy would be required. Pleasingly, the computational prediction demonstrate this, with greater energy being observed for the $S_0 \leftarrow S_1$ emission band than for the unfunctionalised anthracene. Unfortunately, experimental work with fluorinating agents was not possible, however, literature emission of 9-fluoroanthracene clearly indicated a shift to higher energy compared with anthracene.[16]

Considering the other derivatives, experimentally, the 9-chloroanthracene, 9-bromoanthracene and 9-iodoanthracene all have lower energy emission than anthracene with a sequential decrease down the series. This implies that the interplay between electronegative inductive effects and the p-orbital donation favours overall electron density donation to the aromatic system. Indeed, from chlorine to iodine the overall polarizability increases while the electronegativity decreases. This suggests that simultaneously both the electronegative inductive effect decreases while the electron donating potential increases. This was modelled perfectly with the 9-iodoanthracene exhibiting the lowest energy emission as expected.[17] The improved vibrational fine structure of the computational results suggests that the broader peaks in the experimental spectra are composed from several vibrational progressions rather than a single mode. It should be noted that the decrease in emission energy is more pronounced in the experimental spectra than the predictions. This could be due to the prediction merely considering the change in electronegativity and not the donating capabilities.

6.7 COMPUTATIONAL SPECTRA FOR FRET PREDICTIONS

In order to ascertain the utility of the computational emission predictions for use in determining R_0 values, the spectra of both anthracene and naphthalene were used to calculate the R_0 values for transfer to Tb(III) ions. The selection of these two spectra was due to their overlap with the Tb(III) absorption profile, a crucial feature to ensure valid spectral overlap calculations. The other linear polyaromatic hydrocarbons had emission profiles that were too low energy (long wavelength) to have sufficient overlap, thus being unsuitable for this calculation. This was then compared with the R_0 values determined from the experimental spectra recorded for these species. The R_0 values were calculated as previously discussed (*see* Chapter 3), with the only alteration to the assumptions being to change the fluorescence quantum yields of the donors to the literature values. Thus for naphthalene and anthracene the values were set as 0.23 and 0.36 respectively.[18] The computational model struggled with the predictions for naphthalene, however, this is useful as it provides a control to see the impact of poor agreement in the emission spectra and how that affects the calculated R_0 values. Throughout all the emission predictions there was a small energy discontinuity which appeared to be relatively constant and is easiest to see from anthracene-pentacene (Fig. 3). Therefore, as this appeared to be a systematic deviation, in both cases a 0.27 eV correction was applied to the computational emission profiles.

The determination of the Förster resonance energy transfer R_0 values was carried out as previously described (*see* Chapter 3) and made use of spectral overlap calculations. Pleasingly the R_0 values for anthracene to Tb(III) calculated from both computational and experimental spectra resulted in exceptionally agreeable values (Table 1). This was to be expected due to the

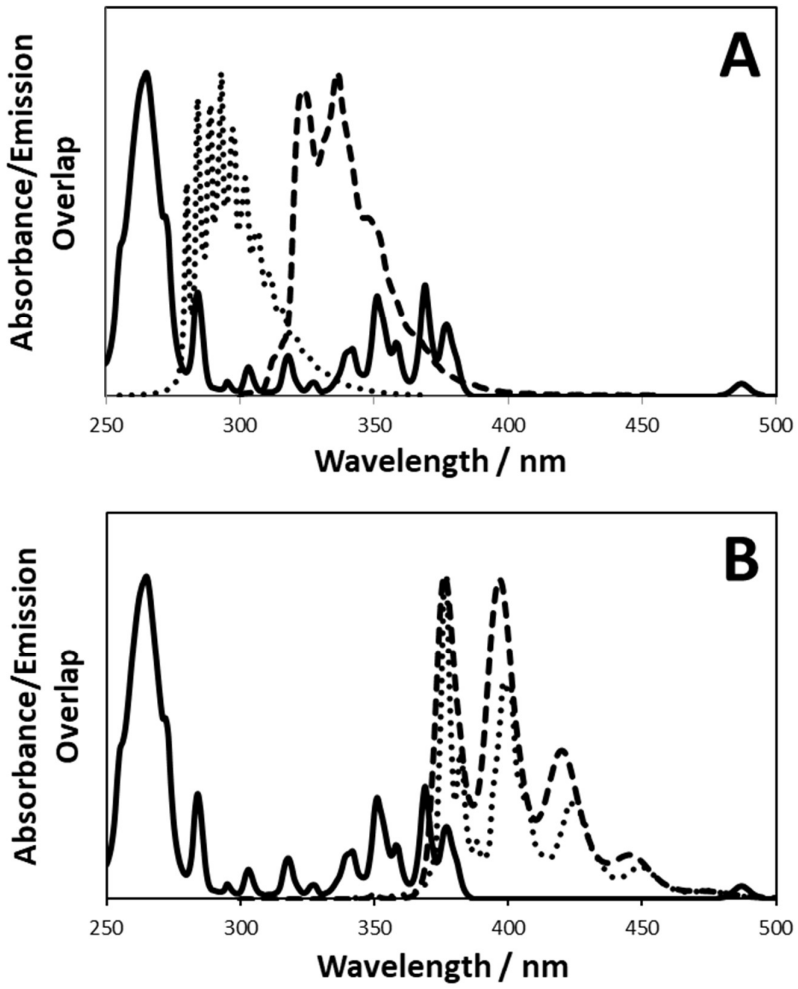

Figure 8: The spectral overlap of A) naphthalene emission and B) anthracene emission with the absorption profile of Tb(III) in aqueous solution (solid line). Both the computational predictions (dotted line) and experimentally determined (dashed line) spectra are illustrated. The experimental spectra were from samples of naphthalene (5 μM) and anthracene (5 μM) in hexane.

excellent agreement between the respective emission spectra as can be seen from the overlap plots (Fig. 8). The experimentally determined R_0 value for naphthalene to Tb(III) was 3.8 Å, as expected considering that the emission profile covers a similar wavelength range to tryptophan. Thus the higher energy, shorter wavelength emission of the prediction resulted in displacement of the overlap avoiding the ideal Tb(III) absorption bands which lead to the short R_0 prediction. While there is a clear discrepancy between the computational and experimental result it is interesting to note that even a poor emission prediction will result in, what appears to be, an acceptable R_0 value, therefore care must be taken when interpreting these results that

the initial data is valid as it is particularly hard to detect errors in R_0 values. These calculations demonstrate the potential of this predictive methodology, but also highlight the limitations of the present logic for simple emission prediction. Indeed, the error associated with a Forster resonance energy transfer distance measurement is usually between 1.0 - 0.5 Å for particularly accurate experiments. Therefore, the observed error in prediction is disturbingly small raising some concerns about the viability of the whole Forster resonance energy transfer experiment.

Table 1: The Forster resonance energy transfer constants R_0 for transfer from naphthalene and anthracene to Tb(III) as calculated from the computationally or experimentally determined emission spectra.

	Computational R_0 / Å	**Experimental R_0 / Å**
Naphthalene	2.8 ± 0.1	3.8 ± 0.1
Anthracene	3.9 ± 0.1	3.9 ± 0.1

6.8 CONCLUSIONS

In summary, the ability to simply predict the emission of various polyaromatic hydrocarbons has been demonstrated. This is particularly exciting as emission prediction does not appear to be a readily available computational technique and it offers the potential to inform experimental design. This is in contrast to the typical computational role, that is, to perform the function of mechanistic elucidation on the experimentally determined systems of interest. This model will also provide utility for applications such as assay design, helping to expand potential luminescent probes available. Indeed, it seems that the model is particularly useful for computational analysis, allowing rapid modelling of a system which then informs the decision on the methods required for a more rigorous approach.

The validity of the methodology was explored utilising functionalised anthracene derivatives. This demonstrated both the potential scope and limitations of the present mode. Overall the preliminary results have proved to be encouraging, especially with regard to the retention of vibrational fine structure within this methodology. Future work on this methodology could include the introduction of heteroatoms into the modelled structures and expanding the predictions to include triplet state emission. It should be noted that while the computational predictions are very useful the methodology is still not accessible to an experimentalist drawing a structure and rapidly screening emission profiles. Therefore a key aspect of future work will involve designing a simple interface for the preparation of a molecular structure and the addition of various simple parameters, such as the presence of a solvent, before uploading to a server to achieve an accurate prediction.

6.9 REFERENCES

[1] M. Taylor, PhD thesis, University of Birmingham, 2019.

[2] i) T. Ha and P. Tinnefeld, *Annu. Rev. Phys. Chem.*, 2012, **63**, 595-617; ii) E. M. S. Stennett, M. A. Ciuba and M. Levitus, *Chem. Soc. Rev.*, 2014, **43**, 1057-1075.

[3] i) M. Artiglia, P. D. Vita and M. Potenza, *Optical and Quantum Electronics*, 1994, **26**, 585-608; ii) J. C. G. Bunzli and C. Piguet, *Chem. Soc. Rev.*, 2005, **34**, 1048-1077; iii) V. Singh, I. J. C. Beltran, J. C. Ribot and P. Nagpal, *Nano Lett.*, 2014, **14**, 597-603; iv) D. J. Wang, J. Zhang, L. Guo, H. D. Shen, F. Fu, G. L. Xue and Y. F. Fang, *J. Inorg. Mater.*, 2015, **30**, 683-693.

[4] i) F. Siebert, *Mikrochim. Acta*, 1997, **S14**, 43-50; ii) M. Hunger and J. Weitkamp, *Angew. Chem. Int. Ed.*, 2001, **40**, 2954-2971.

[5] J. Y. Zhou, J. J. Krich, I. Kassal, A. S. Johnson and A. A. Guzik, *Ultrafast Spectroscopy – Quantum Information and Wavepackets*, IOP Publishing Ltd, Bristol, 2014.

[6] i) P. Jensen and P. R. Bunker, *Computational Molecular Spectroscopy*, John Wiley and Sons Ltd., London, 2000; ii) D. C. Lathey, MSci thesis, Marshall University, 2005.

[7] E. Fermi, *Z. Phys.*, 1931, **71**, 250-259.

[8] i) G. A. Worth, K. Giri, G. W. Richings, I. Burghardt, M. H. Beck, A. Jackle and H. -D. Meyer, *The Quantics Package*, 2015, **Version 1.1**, University of Birmingham; ii) M. H. Beck, A. Jackle, G. A. Worth and H. -D. Meyer, *Phys. Rep.*, 2000, **324**, 1-105.

[9] A. DeFusco, N. Minezawa, L. V. Slipchenko, F. Zahariev and M. S. Gordon, *J. Phys. Chem. Lett.*, 2011, **2**, 2184-2192.

[10] Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian Inc., Wallingford CT, 2016.

[11] i) E. Clar, *Polycyclic Hydrocarbons,* Academic Press, London, 1964; ii) E. Clar, *The Aromatic Sextet*, John Wiley and Sons Ltd., London, 1972.

[12] i) M. A. Omary and H. H. Patterson, *Luminescence Theory* in *Encyclopedia of Spectroscopy and Spectrometry*, ed. J. C. Lindon, G. E. Tranter and D. Koppenaal, Academic Press, London, 1st edn, 2000, pp. 1186-1207; ii) D. Pentlehner, C. Greil, B. Dick and A. Slenczka, *J. Chem. Phys.*, 2010, **133**, 114505.

[13] R. Lubart and M. B. Lindner, *Spectrochim. Acta*, 1982, **38A**, 953-959.

[14] i) G. Herzberg and E. Teller, *Z. Phys. Chem. Abt. B*, 1933, **21**, 410; ii) R. Renner, *Z. Phys.*, 1934, **92**, 172; iii) G. J. Small, *J. Chem. Phys.*, 1971, **54**, 3300-3306, iv) V. A. Kuz'mitskii, *Optics and Spectroscopy*, 2006, **101**, 666-672.

[15] M. S. A. Abdel-Mottaleb, H. R. Galal, A. F. M. Dessouky, M. El-Nagger, D. Mekkawi, S. S. Ali and G. M. Attya, *Int. J. Photoenergy*, 2000, **2**, 47-53.

[16] R. P. Steiner and J. Michl, *J. Am. Chem. Soc.*, 1978, **100**, 6861-6867.

[17] i) C. G. Morgante and W. S. Struve, *Chem. Phys. Lett.*, 1979, **68**, 272-275; ii) R. L. Pineault, C. G. Morgante and W. S. Struve, *J. Photochem.*, 1981, **17**, 435-449.

[18] I. B. Berlman, *Handbook of Fluorescence Spectra of Aromatic Molecules*, Academic Press, London, 1965.

Chapter 7

CONCLUSIONS AND FUTURE WORK

7.1 INTRODUCTION

The research presented in this thesis involved the construction of novel coiled coil peptide systems for their elegant spatial positioning properties. These structures were subsequently studied with a particular interest in the interaction between tethered chromophores and bound lanthanide ions. This allowed an understanding of the principles of energy transfer in biological systems and the potential for these to be used in the construction of simple light harvesting complexes. Throughout the course of this research interesting discoveries have been made regarding structural features such as the lanthanide binding site, photophysical phenomenon involving new lanthanide sensitisation, rapid computational modelling of emission spectra, synthesis of a novel non-natural amino acid with a coumarin motif and finally incorporation of non-natural amino acids into exciting peptide scaffolds. The present work and potential of each of these areas will be considered in turn.

7.2 PEPTIDE STRUCTURE

The design of a series of novel coiled coil peptides, inspired by both the MB series studied previously in this group,[1] and well characterised coil-Ser derivatives,[2] was successfully achieved. This resulted in a series of coiled coils with core lanthanide binding sites, which crystallised remarkably well resulting in all but one peptide having a structure successfully determined. The structure of the lanthanide binding site was observed for the first time allowing useful insights into the coordination environment of bound Tb(III) ions. It was apparent that every residue in the designed binding site motif was necessary providing at least one donor

atom. The parallel nature of the helices resulted in the successful formation of the binding site and from both luminescence and circular dichroism titrations the log of the binding constants (LogK_a) were found to be typically ~5 for all the lanthanides measured. This is particularly weak binding and indeed the issue of weak binding is one of the key limitations of this type of system for certain applications as exemplified by the MB series with Gd(III).[1] This leads to otherwise promising MRI activity not being feasible due to the toxicity of free Gd(III).[3] Therefore to improve binding it may be beneficial to make use of the chelate effect and design a protein comprised of a linear bundle with the binding site incorporated. However, a key difficulty is the required parallel nature of the helices for the binding site to function. The most efficient design for an α-helical bundle would be antiparallel as this would not require two long loop sections. Therefore, using the crystal structure data from the parallel site, it may be possible to redesign and model an antiparallel site with the same preformed coordination environment. This approach could also require the use of expression for synthesis as a long sequence is not yet compatible with resin based solid-phase peptide synthesis.[4] This is a useful and important step as commercial peptides/proteins are generally prepared via expression and not the more expensive solid-phase techniques.

7.3 TRYPTOPHAN AND TYROSINE TO LANTHANIDE PHOTOPHYSICS

Utilising the naturally occuring amino acid, tryptophan, it was possible to study the photophysical behaviour of both Tb(III) and Eu(III) ions bound within the coiled coil trimer. Indeed, the use of tryptophan as a sensitiser for Tb(III) and Eu(III) has been well documented,[5] however, in addition to Tb(III) and Eu(III), the use of tryptophan as a viable sensitiser for both

Dy(III) and Sm(III) in peptide/protein systems was predicted and experimentally verified. Tryptophan has been found to be useful as a broad sensitiser for the visible emitting lanthanides. Exciting results indicating that tryptophan could sensitise near-infrared emission from Yb(III) are in agreement with observations made by Horrocks *et al*.[6] Thus it can be concluded that tryptophan can be a viable sensitiser for emission from Sm(III), Dy(III) and Yb(III) as well as the more traditionally utilised Tb(III) and Eu(III). Interestingly, it was also observed that tyrosine could successfully sensitise emission from Dy(III), expanding the utility of this natural amino acid as well.

Further to the sensitising potential of tryptophan it has also been made apparent that interesting redox behaviour is not just limited to Eu(III) and Yb(III),[6] which have been discussed in the literature, but also includes Sm(III) and weakly Tm(III). All four of these lanthanides can be reduced from Ln(III) to Ln(II) by excited state tryptophan via electron transfer processes. Nevertheless, it should be noted that the change in oxidation state has not been observed directly but is hypothesised from a proposed mechanism. Therefore a key aspect of future work is to test this mechanism utilising fast techniques such as photoelectron spectroscopy.

Overall, this aspect of the research has increased the potential utility of the natural amino acids, tryptophan and tyrosine, for sensitising lanthanides. This, in turn, expands the potential luminescence probes available and thus a large number of biological systems can be studied. One can imagine using the HC peptides as tags to biological systems which could be a viable route to construct assays. However, it should be noted that this would be very complex and likely to alter the behaviour of the system being studied. Therefore it would be advantageous to test the assay viability of the discovered photophysical effects by using well understood

systems such as the lanthanide binding tag (LBT).[7] Indeed, the sensitisation of Dy(III) and Sm(III) bound in a LBT has already been demonstrated in this thesis (*see* S34). This would allow rapid adoption of the potential luminescence probes, as they would only require the addition of the different lanthanide ion and alteration of spectroscopic filter settings. Another possibility for future work would be to explore the potential for incorporating the lanthanide binding site reported here and elsewhere, into a biological system of interest. However, the lack of linearity in the binding site design makes this also particularly complicated. This reiterates the previous discussion of the structure of the HC series, namely a redesign of the binding site for an antiparallel motif, which would enable incorporation into proteins by adding a linear sequence capable of forming a lanthanide binding helical domain.

7.4 COMPUTATIONAL EMISSION PREDICTION

The computational work carried out in collaboration with the Worth group has proved to be a versatile tool for accurately predicting the emission spectra of a number of small molecules. The reproducibility of spectra for simple polyaromatic hydrocarbons has been particularly encouraging. Unfortunately, there are a number of severe limitations, as previously discussed, regarding solvent effects and that the only consideration is for singlet state energy transfer. As this was, for the most part, a proof of concept methodology it has performed remarkably well, especially with the retention of vibrational fine-structure in the predicted emission spectra.

Future work on this methodology should include a simple work-around for various functional group to solvent interactions, especially more polar functionalities. A simple addition would

be the consideration of triplet states and the resultant phosphorescence spectra prediction as the generation of these surfaces is not too challenging and the dynamics would remain unchanged. However, the primary limitation at present is that the methodology is not, at present, easily accessible for the experimentalist. Therefore a key aspect of the work required for the model, is to design an interface to allow relatively straightforward input of the molecular structure for rapid emission prediction.

7.5 SYNTHESIS OF A NON-NATURAL AMINO ACID

The synthetic strategy applied to prepare the 4-methylcoumarin functionalised amino acid has potential as there are a great many ethyl acetoacetate derivatives which could be reacted with the phenol functionalised side chain of tyrosine, though three key limitations should be noted. Firstly, the overall yields for this reaction were quite low, typically 3.9% for the *N*-acetyl-*L*-(4-methyl)coumarin-6-yl-alanine derivative and 23% for the *N*-acetyl-*L*-(4-chloromethyl) coumarin-6-yl-alanine derivative, but this could be improved by optimisation. Secondly, from a peptide synthesis perspective it is important to consider hydrophobicity and steric factors associated with non-natural amino acids, as this will influence folding and protein stability. This is important because this strategy results in relatively large amino acid side chains, therefore it will be useful to consider the effect of functional groups appended to the ethyl acetoacetate feedstocks. Thirdly, at present the *N*-acetyl protecting group limits the use of this species for *C*-terminal capping. However, optimisation should result in a strategy that would allow the Pechmann reaction to be applied to *N*-fluorenylmethoxycarbonyl- protected feedstocks.

7.6 LIGHT HARVESTING MODULE DESIGN AND SYNTHESIS

In the pursuit of a light-harvesting module, a number of key features are particularly attractive regarding the HC01 derivatives. The HC01 systems have the capacity to absorb light, albeit of high energy, and transport this captured energy to a bound lanthanide species. The buried nature of the lanthanide binding site, reduces solvent exposure, thus generating long lived excited states for the lanthanide emission, from ~20 μs to 2 ms. This potentially allows a type of storage mechanism for captured light which is usually lost in femtosecond to nanosecond timescales for organic species. The variety of chromophores used to construct the HC01 derivatives has given some interesting properties to each system. Perhaps the most important is the shift to lower excitation energy required for HC01-2Cou allowing the use of a broader range of light wavelengths, or to be selectively excited in the presence of other chromophores, such as those considered here. Further work is required to establish the efficiency of the energy transfer from the donor chromophores to the acceptor lanthanide ions. This can be achieved by utilising a standard to measure the quantum yield of emission from the lanthanide ions.[8]

Interestingly, the potential electron transfer from tryptophan to various lanthanides provides a method of achieving charge separation. Advantage could be taken of this phenomenon to use a redox cascade to reduce energy losses via back transfer and exploit the redox properties of these excited states to carry out chemistry. However, electron transfer efficiencies are usually incredibly distance sensitive therefore it is important to model the potential transfer as the distance between even two heptads of ~1 nm may be too great for this to occur.[9]

7.7 CONCLUSION

Throughout this thesis the study of light interaction with several designed synthetic peptide scaffolds was achieved. A structural and photophysical understanding of new lanthanide binding coiled coils was attained, improving the knowledge of this field beyond previous endeavours. The toolbox of lanthanide sensitising chromophores was expanded and an optimised model system was constructed, HC01. The computational tools for emission prediction have proven to be both versatile and accurate, showcasing a promising methodology for generating spectra from molecular structures. This has the potential to revolutionise the study of large libraries of molecules, allowing rapid analysis before time-consuming experimental work is undertaken. Non-natural amino acid chromophores combined with the versatility of the model system HC01 has resulted in the construction of a number of small, compact light harvesting modules that exhibit high energy transfer from the primary chromophore to a bound lanthanide ion. This has paved the way for the preparation of an efficient light harvesting module to act as a power source for nanomachines or catalysts.

7.8 REFERENCES

[1] i) M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton and A. F. A. Peacock, *J. Am. Chem. Soc.*, 2014, **136**, 1166-1169, ii) M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton and A. F. A. Peacock, *Chem. Sci.*, 2016, **7**, 2207-2216.

[2] i) K. T. O'Neil and W. F. DeGrado, *Science*, 1990, **250**, 646-651, ii) B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado and D. Eisenberg, *Science*, 1993, **259**, 1288-1293, iii) L. Ruckthong, M. L. Zastrow, J. A. Stuckey and V. L. Pecoraro, *J. Am. Chem. Soc.*, 2016, **138**, 11979-11988.

[3] i) G. E. Jackson, S. Wynchank and M. Woudenberg, *Magn. Reson. Med.*, 1990, **16**, 57-66, ii) M. Rogosnitzky and S. Branch, *Biometals*, 2016, **29**, 365-376.

[4] W. Chan and P. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, 2000.

[5] i) W. DeW. Horrocks and A. P. Snyder, *Biochem. Biophys. Res. Commun.*, 1981, **100**, 111-117, ii) W. DeW. Horrocks and W. E. Collier, *J. Am. Chem. Soc.*, 1981, **103**, 2856-2862, iii) J. Bruno, W. DeW. Horrocks and R. J. Zauhar, *Biochemistry*, 1992, **31**, 7016-7026.

[6] W. D. Horrocks, J. P. Bolender, W. D. Smith and R. M. Supkowski, *J. Am. Chem. Soc.*, 1997, **119**, 5972-5973.

[7] i) K. J. Franz, M. Nitz and B. Imperiali, *ChemBioChem*, 2003, **4**, 265-271, ii) M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen and B. Imperiali, *Angew. Chem. Int. Ed.*, 2004, **43**, 3682-3685.

[8] M. Xiao and P. R. Selvin, *J. Am. Chem. Soc.*, 2001, **123**, 7067-7073.

[9] M. Sisido, S. Hoshino, H. Kusano, M. Kuragaki, M. Makino, H. Sasaki, T. A. Smith and K. P. Ghiggino, *J. Phys. Chem. B*, 2001, **105**, 10407-10415.

Chapter 8

MATERIALS AND EXPERIMENTAL METHODS

8.1 MATERIALS

Chemicals were used as received and purchased from; Sigma Aldrich (sinapinic acid, diethyl ether, ethyl acetoacetate, ethyl 4-chloroacetoacetate, *L*-tyrosine methyl ester, *N*-fluorenylmethoxycarbonyl-3-(2-naphthyl)-*L*-alanine, terbium chloride hexahydrate, thulium chloride hexahydrate, ytterbium chloride hexahydrate, neodymium chloride hexahydrate, gadolinium chloride hexahydrate, zinc acetate, poly(ethylene) glycol 3350, naphthalene, anthracene, tetracene, pentacene, 9-methylanthracene, 9-hydroxymethylanthracene, 9-chloroanthracene, 9-bromoanthracene, 9-iodoanthracene, 9-anthracenecarboxaldehyde, 9-anthracenecarboxylic acid, dimethyl sulfoxide-d_6, deuterium oxide, sulphuric acid, ammonium acetate, toluene, ethanol, ammonium chloride and ethylene glycol), Pepceuticals (*N*-fluorenylmethoxycarbonyl protected amino acids, *O*-(benzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate, synthesis grade *N,N*-dimethylformamide, and 20% piperidine in dimethylformamide premix), AGTC Bioproducts Ltd. (rink amide 4-methylhydrylamine resin, *N*-methyl-2-pyrrolidone, hydroxybenzotriazole, *N*,*N*-diisopropylethylamine and dichloromethane), Acros Organics (xylenol orange sodium salt, europium chloride hexahydrate, acetic anhydride, triisopropylsilane and trifluoroacetic acid), Rathburn Chemicals Ltd. (>99.9% *N,N*-dimethylformamide), Fisher Scientific Ltd. (high pressure liquid chromatography grade water and acetonitrile, 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid, urea, ethylenediaminetetraacetic acid, and glacial acetic acid), Alfa Aesar (dysprosium chloride hexahydrate), VWR International (hexane, acetone, lutetium chloride hexahydrate), Strem Chemicals UK Ltd. (samarium chloride hexahydrate). All water, except that used for high pressure liquid chromatography, was ultrapure grade obtained from a Millipore-Elix-Gradient A10 system.

8.2 PEPTIDE SYNTHESIS

8.2.1 AUTOMATED

The methodology for peptide synthesis have been discussed extensively by previous members of the Peacock group.[1] Peptide synthesis was facilitated by standard solid phase peptide synthesis protocols utilising a CEM Liberty Blue Automated Peptide Synthesiser.[2]

The protected amino acids were *N*-(9-fluorenylmethyloxycarbonyl)-*L*-alanine, N_α-(9-fluorenylmethoxycarbonyl)-N_γ-trityl-*L*-asparagine, *N*-(9-fluorenylmethyloxycarbonyl)-*L*-aspartic acid 4-^{t}butyl ester, N_α-(9-fluorenylmethoxycarbonyl)-N_δ-trityl-*L*-glutamine, *N*-(9-fluorenylmethyloxycarbonyl)-*L*-glutamic acid 5-^{t}butyl ester, *N*-(9-fluorenylmethyloxycarbonyl)-glycine, N_α-(9-fluorenylmethoxycarbonyl)-*N*-trityl-*L*-histidine, *N*-(9-fluorenylmethyloxycarbonyl)-*L*-isoleucine, N_α-(9-fluorenylmethoxycarbonyl)-N_ε-^{t}butyloxycarbonyl-*L*-lysine, *N*-(9-fluorenylmethyloxycarbonyl)-*O*-^{t}butyl-*L*-serine, N_α-(9-fluorenylmethoxycarbonyl)-*N*-^{t}butyloxycarbonyl-*L*-tryptophan, *N*-(9-fluorenylmethyloxycarbonyl)-*O*-^{t}butyl-*L*-tyrosine. These species were dissolved in *N,N*-dimethylformamide to give 0.2 M solutions. Rink amide 4-methylhydrylamine resin (0.368 g, 0.68 meq g^{-1}) was agitated in *N,N*-dimethylformamide (10 mL) for 20 minutes to swell the resin. The swollen resin was then loaded into the automated synthesiser reaction vessel. All methods on the automated synthesiser were carried out under an inert nitrogen atmosphere. The *N*-fluorenylmethyloxycarbonyl protecting groups were removed using 20% piperidine

and hydroxybenzotriazole (0.1 M) in *N,N*-dimethylformamide, facilitated by microwave methods (60 W, 75 °C, 180 s or 30 s). Following deprotection, the next amino acid in the sequence was coupled in the presence of an activator such as *O*-(benzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (0.5 M) in *N,N*-dimethylformamide and an activator base such as *N,N*-diisopropylethylamine (2 M) in *N*-methyl-2-pyrrolidone. The couplings were facilitated by microwave methods, typically 10 minutes coupling time (30 W, 75 °C, 600 s), though for histidine the coupling temperature was reduced (30 W, 50 °C, 600 s) but the time remained 10 minutes. The first residue to be attached to the resin was double coupled and the final three heptads were also double coupled. Following any reaction, either deprotection or coupling, the vessel was washed with *N,N*-dimethylformamide (3 x 5 mL, 5 s drain time). After couplings the manifold was also washed with *N,N*-dimethylformamide (2 x 4 mL, 5 s drain time).

Upon completion of the sequence synthesis, the final residue was left deprotected ready for capping. The resin was removed from the automated synthesiser and swollen in *N,N*-dimethylformamide (10 mL) for 20 minutes before the addition of a capping mixture consisting of *N,N*-diisopropylethylamine (5 mL) and acetic anhydride (5 mL). The mixture was agitated for 30 minutes, then filtered and the resin washed with dichloromethane (3 x 7 mL) and diethyl ether (3 x 7 mL) before air drying for 2 hours.

The dry capped peptide on resin was placed in a falcon tube prior to the addition of a cleavage cocktail consisting of trifluoroacetic acid (18 mL), water (1 mL) and triisopropylsilane (1 mL). The tube was agitated for 3 hours before filtering and

washing the solids with trifluoroacetic acid (2-3 mL). The filtrate was triturated in ice-cold diethyl ether (4 x 40 mL) and placed in the freezer for 30 minutes. The solid crude peptide was isolated by centrifugation (Eppendorf centrifuge 5702 with A-4-38 and rotational speed of 4400 rpm for 10 minutes), the diethyl ether was decanted and a further addition of ice-cold diethyl ether was added (4 x 20 mL). The tubes were agitated for 20 minutes and placed in the freezer for 30 minutes prior to centrifugation and decanting of the diethyl ether. This yielded the crude peptide as a white/cream solid. The crude peptide was dissolved in aqueous acetonitrile (1:1), typically requiring 25-30 mL before filtering through a Millex syringe-driven filter unit (0.45 μm pore size) in preparation for high pressure liquid chromatography purification.

8.2.2 MANUAL

The only peptides that required manual synthesis were HC01-2Nap, HC01-2Cou and MB1-2,14Nap. All manual couplings only took place following the bulk of the peptide having been synthesised in an automated system (*see* **8.2.1**)

The rink amide 4-methylhydrylamine resin with the partially constructed peptide was removed from the reaction vessel, after the deprotection stage of the residue immediately preceding the position required for a non-natural amino acid. The resin was placed in a falcon tube and *N,N*-dimethylformamide (10 mL) was added. To the falcon tube was added a solution of *O*-(benzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (2.5 mL, 0.5 M) in *N,N*-dimethylformamide,

N,*N*-diisopropylethylamine (0.75 mL) and *N*-methyl-2-pyrrolidone (0.5 mL). The tube was then shaken for 10 minutes before the addition of a solution of the non-natural amino acid (6 mL, 0.2 M) in *N,N*-dimethylformamide followed by further shaking for 2 hours at room temperature. The resin was then filtered and washed with *N,N*-dimethylformamide (3 x 7 mL). For HC01-2Cou the resin was then washed with dichloromethane (3 x 7 mL) and diethyl ether (3 x 7 mL) before air-drying for 1 hour and storing in preparation for cleavage (*see* **8.2.1**). For both MB1-2,14Nap and HC01-2Nap the resin was placed in a falcon tube with piperidine (18 mL, 20% v/v) in *N,N*-dimethylformamide and shaken for 20 minutes. The resin was then filtered, washed with *N,N*-dimethylformamide (3 x 7 mL) and placed in a falcon tube with *N,N*-dimethylformamide (10 mL), *O*-(benzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (2.5 mL, 0.5 M) in *N,N*-dimethylformamide, *N*,*N*-diisopropylethylamine (0.75 mL) and *N*-methyl-2-pyrrolidone (0.5 mL) before shaking for 10 minutes. This was followed by the addition of a solution of the next amino acid (6 mL, 0.2 M) in *N,N*-dimethylformamide and shaken for 2 hours at room temperature. The resin was then filtered and washed with *N,N*-dimethylformamide (3 x 7 mL). For MB1-2,14Nap the resin was then placed back in the reaction vessel of the CEM Liberty Blue Automated Peptide Synthesiser to proceed with the synthesis of the sequence (*see* **8.2.1**). For HC01-2Nap the resin was placed in a falcon tube with piperidine (18 mL, 20% v/v) in *N,N*-dimethylformamide and shaken for 20 minutes. The resin was then filtered and washed with *N,N*-dimethylformamide (3 x 7 mL) before following the capping and cleaving procedures discussed previously (*see* **8.2.1**).

8.3 HIGH PRESSURE LIQUID CHROMATOGRAPHY

8.3.1 PREPARATIVE

The crude peptide was purified by reverse phase high pressure liquid chromatography on a preparative Waters XBridge C_{18} column. The eluent A consisted of water and trifluoroacetic acid (0.05% v/v) and the eluent B consisted of acetonitrile and trifluoroacetic acid (0.05% v/v). The method involved going from 0 – 100% eluent B over 40 minutes followed by a wash cycle. Flow rate was maintained at 10 mL min^{-1} throughout the method and the observed detector wavelengths were set at 210 and 280 nm. The resultant pure fractions were rotary evaporated to remove the acetonitrile and lyophilised on a freeze dryer overnight to yield the pure peptide as fluffy white powder. Typical yields of the final purified peptides ranged between 10% and 40%.

8.3.2 ANALYTICAL

A sample of the purified peptide in aqueous acetonitrile (50% v/v) was placed into the auto sampler for analysis by reverse phase high pressure liquid chromatography on an analytical Waters XBridge C_{18} column. The eluent A consisted of water and trifluoroacetic acid (0.05% v/v) and the eluent B consisted of acetonitrile and trifluoroacetic acid (0.05% v/v). The method involved going from 0 – 100% eluent B over 40 minutes followed by a wash cycle. Flow rate was maintained at 1 mL min^{-1} throughout the method and the observed detector wavelengths were set at 210 and 280

nm. Note that the signals present at ~4 minutes are instrumental artefacts and were present in water blanks as well as the sample. The typical retention times for all peptides in this thesis were between 18 and 23 minutes and the purities were found to be >95% (*see* Table 1). The analytical chromatograms for each peptide can be found in the supporting information (*see* S2 – S12).

8.4 MASS SPECTROMETRY

8.4.1 MATRIX ASSISTED LASER DESORPTION IONISATION

A solution of the peptide (9 μM) in aqueous acetonitrile (1:1) was added to a solution of sinapinic acid (10 mg mL^{-1}) in a 1:1 ratio. The resultant solution (2 μL) was placed on a matrix assisted laser desorption ionisation plate and the solvent was allowed to evaporate. Data collection was carried out on a Waters Micromass Micro Spectrometer. The monoisotopic masses are reported in table 1 and the spectra can be found in the supporting information (*see* S2 – S12). All the data was found to be consistent with successfully synthesised peptide sequences.

8.4.2 ELECTROSPRAY IONISATION

Mass spectrometry of the trimer complexes was carried out on a Synapt G2S mass spectrometer fitted with a TirVersa NanoMate chip-based spray controller and

Nanoflow+. Samples of peptide (9 μM) in ammonium acetate buffer (10 mM, pH 6.4) were added in 5 - 10 μL sprays, with a gas pressure of 0.15 psi and a voltage of 1.75 - 1.80 kV. It was essential to maintain the backing vacuum at 6.5 x 10^{-6} torr and the valve temperature at 100 °C. Detected ionisation parameters were carried out utilising a local method, file identity – OJD_2017. Stepwave Offsets were varied between 11 - 25 and the TrapCE and TransferCD were generally set to 20, though these were not always applied. Gas controls were particularly important for larger oligomeric states with trimers generally visible between 3.0 - 5.0 gas flow but tetramers requiring >5.0. The capture range was set to 2000 - 3000 m/z; measured spectra can be found in the supporting information (*see* S15 – S24).

Table 1: The retention time in minutes for the analytical high pressure liquid chromatography of the named peptides with their analytical high pressure liquid chromatography determined purity (%). The monoisotopic masses from the matrix assisted laser desorption ionisation mass spectrometry.

Peptide	Retention Time / min	Purity %	Monoisotopic Mass
HC0C	20.3	>99	4155.8
HC01	17.9	>99	4158.8
HC02	17.2	>99	4158.8
HC03	17.6	97	4158.7
HC04	18.6	>99	4159.1
HC01-2Y	18.2	>99	4135.5
HC01-2Nap	18.6	98	4169.6
HC01-2Cou	18.5	98	4072.3
MB1-2	19.0	97	4001.2
MB1-2,14Nap	18.6	98	4012.8
LBT	12.4	97	1986.7

8.5 GENERAL SAMPLE PREPARATION

Lanthanide chloride solutions (1 mM) were prepared in ultrapure water (10 mL), and their accurate concentrations determined by titration in triplicate using xylenol orange indicator and ethylenediaminetetraacetic acid titration as reported by Barge *et al.* to yield the lanthanide stock solution.[3] The lanthanide can be terbium, europium, samarium, dysprosium, ytterbium, thulium, and lutetium.

Peptide stock solutions were prepared in deionised water, typically 1 mg mL^{-1}. The peptide concentration was determined by taking an aliquot of the stock (10 μL) and placing it in a 1.5 mL eppendorf with an aqueous solution of urea (690 μL, 7 M) to give a total volume of 700 μL. The vial was centrifuged for 3 minutes using an Eppendorf MiniSpin with a 45° fixed angle rotor F-45-12-11 and a rotational speed of 13,400. The solution was left to stand for 10 minutes before 600 μL were placed in a quartz cuvette prior to UV-vis characterisation to determine the peptide concentration via the Beer-Lambert law (*see* **8.6**).

Experimental samples required to support the computational work presented in Chapter 6 were prepared by accurately weighing samples and diluting with the necessary solvent (toluene, hexane or basic ethanol) into volumetric flasks, to yield the final concentrations reported in Table 2.

8.6 UV-VIS ABSORPTION SPECTROSCOPY

Ultraviolet and visual spectrum absorption profiles were recorded on a Shimadzu 1800 UV Spectrophotometer, using single beam mode and a medium scan speed. The slit width was set at 1.0 nm and the data interval was set at 1.0 nm. All samples were placed in a quartz UV-vis cuvette (pathlength of 1 cm).

Peptide samples were scanned from 500 - 200 nm with a baseline of urea (690 µL, 7 M) and water (10 µL). The baseline was removed from the sample absorption profile and the peptide

Table 2: The concentration of prepared polyaromatic hydrocarbon samples with the solvent utilised to dissolved them and the excitation wavelength used to obtain an emission spectrum.

Compound	Concentration / µM	Solvent	Excitation Wavelength / nm
Naphthalene	45	Hexane	275
Anthracene	34	Hexane	350
Tetracene	15	Toluene	420
Pentacene	5	Toluene	533
9-Chloroanthracene	30	Hexane	350
9-Bromoanthracene	25	Hexane	350
9-Iodoanthracene	30	Hexane	-
9-Methylanthracene	72	Hexane	330
9-Hydroxymethyl anthracene	22	Hexane	330
9-Anthracene carboxylic acid	45	Basic Water (pH 10)	365
9-Anthracene carboxaldehyde	19	Ethanol	372

concentrations were determined based on tryptophan absorbance at 280 nm (5690 M^{-1} cm^{-1}),[4] tyrosine absorbance at 274 nm (1405 M^{-1} cm^{-1}),[4] 2-naphthyl absorbance at 289 nm (3860 M^{-1} cm^{-1}) and 4-methylcoumaryl absorbance at 320 nm (4120 M^{-1} cm^{-1}).[5] The molar absorptivity for both 2-naphthyl and 4-methylcoumaryl can be found in the supporting information (*see* Chapter 9). The concentration of the sample was calculated utilising the molar absorptivity of the chromophore present on the peptide, this in turn was used to calculate the original stock concentration.

Experimental samples required to support the computational work presented in Chapter 6 were used directly and scanned from 800 - 200 nm. The baseline samples were the solvent required to dissolve the compound of interest (*see* Table 2).

8.7 CIRCULAR DICHROISM SPECTROSCOPY

The circular dichroism spectra were recorded using a JASCO J-715 Spectropolarimeter which required the optical chamber to be purged with nitrogen for ~20 minutes prior to use. The optics were kept under a nitrogen atmosphere throughout the experiment. For each sample measurement there were three scans recorded from 350-190 nm using a 200 nm min^{-1} scan rate, 1 nm bandwidth and 1 s response time. The sample chamber temperature was maintained at 20 °C throughout the experiment.

The samples consisted of the aqueous peptide (90 μM) and 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid buffer (10 mM, pH 7.0) with the baseline sample being just the 2-[4-(2-

hydroxyethyl)piperazin-1-yl]ethanesulfonic acid buffer (10 mM, pH 7.0). The samples were placed in a 1 mm pathlength quartz cuvettes and if any addition was made to the sample (e.g. terbium during a titration) it was left to equilibrate for 10 minutes.

The observed ellipticity, θ_{obs} (mdeg) was converted into molar ellipticity, (θ), and is reported in units of deg $dmol^{-1}$ cm^2. The helical content was determined following the method reported by Y. Chen *et al.* and using the minimum molar ellipticity determined by J. Myers *et al.*[6] The change in molar ellipticity was analysed as an isotherm (*see* **8.9**).

8.8 LUMINESCENCE SPECTROSCOPY

8.8.1 UV-VIS LUMINESCENCE FOR PEPTIDE SAMPLES

Emission spectra were recorded in a 1 cm path length quartz cuvette using an Edinburgh Instruments Fluorescence FL920 system with a 450 W xenon arc lamp and a Hamamatsu R928 photomultiplier tube. Two interchangeable gratings blazed at 500 nm and 1200 nm were fitted to the emission monochromator. When monitoring the emission between 300 - 500 nm a 305 nm long pass filter was utilised with an excitation slit width of 3 nm, an emission slit width of 1 nm and a 0.5 s dwell time. When monitoring the emission from 460 - 800 nm a 455 nm long pass filter was utilised with an excitation slit width of 3 nm, an emission slit width of 5 nm and 0.5 s dwell time. Samples of peptide (90 μM) in 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid buffer (10 mM, pH 7.0) were prepared and placed in 1 cm pathlength quartz

cuvettes. If any addition was made to a sample (e.g. terbium during a titration) it was stirred with a micro stirrer bar and stirrer plate for 10 minutes prior to recording an emission spectrum. The solutions of peptide were excited with energy dependent upon the chromophore present; tryptophan was excited at 280 nm, tyrosine was excited at 275 nm, naphthyl was excited at 280 nm and coumaryl at 320 nm. Excitation profiles were obtained by observing the emission maxima for tryptophan at 350 nm, tyrosine at 303 nm (thus no filter was used for HC01-2Y emission), naphthyl at 333 nm, coumaryl at 389 nm, terbium at 545 nm, europium at 616 nm, dysprosium at 571 nm and samarium at 643 nm. The excitation wavelength for tryptophan, tyrosine and naphthyl was scanned from 200 - 300 nm with a 305 nm long pass filter used for both tryptophan and naphthyl between the sample and the emission slit. The excitation wavelength for the coumaryl was scanned from 200 - 360 nm. The excitation wavelengths scanned for the peptide bound lanthanide species was dependent upon the chromophore attached to the peptide. For all lanthanide excitation measurements a 455 nm long pass filter was placed between the sample and the emission slit. The data was collected and processed using F900 spectrometer analysis software.

8.8.2 INFRARED LUMINESCENCE FOR PEPTIDE SAMPLES

Emission spectra were recorded in a 1 cm path length quartz cuvette using Edinburgh Instruments Fluorescence FL920 system with a 450 W xenon arc lamp and a Hamamatsu R928 photomultiplier tube. The detector was cooled to -70 °C prior to running any experiments and this temperature was maintained throughout the

experiments. Two interchangeable gratings blazed at 500 nm and 1200 nm were fitted to the emission monochromator. When monitoring the emission between 900 - 1100 nm a 780 nm long pass filter was utilised with an excitation slit width of 5 nm, an emission slit width of 15 nm and a 5.0 s dwell time. Samples of peptide (500 μM) in 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid buffer (10 mM, pH 7.0) were prepared and placed in 1 cm pathlength quartz cuvettes. If any addition was made to a sample (e.g. ytterbium) it was stirred with a micro stirrer bar and stirrer plate for 10 minutes prior to recording an emission spectrum. The solutions of tryptophan functionalised peptide were excited with 280 nm light. The excitation profile of HC01 with ytterbium bound was obtained by observing the emission maxima for ytterbium at 980 nm while scanning the excitation wavelength for tryptophan 250 - 350 nm with a 780 nm long pass filter between the sample and the emission slit. The data was collected and processed using F900 spectrometer analysis software.

8.8.3 UV-VIS LUMINESCENCE FOR SMALL MOLECULES

Emission spectra were recorded in a 1 cm path length quartz cuvette using Edinburgh Instruments Fluorescence FL920 system with a 450 W xenon arc lamp and a Hamamatsu R928 photomultiplier tube. Two interchangeable gratings blazed at 500 nm and 1200 nm were fitted to the emission monochromator. When monitoring the emission between 300 - 600 nm, a 305 nm long pass filter was utilised with an excitation slit width of 3 nm, an emission slit width of 1 nm and a 0.5 s dwell time. Solutions of the experimental samples required for the computational work were

prepared as previously discussed (*see* **8.5**). These solutions (1 mL) were placed in a 1cm pathlength quartz cuvettes and excited with varying wavelengths to obtain the emission spectra (*see* Table 3). The data was collected and processed using F900 spectrometer analysis software.

8.9 ISOTHERM FITTING

Both the circular dichroism and luminescence titration data can be manipulated to give useful isotherms monitoring the fraction of trimers bound (*f*) against the concentration of the terbium ions (*see* Equation 1B). For the circular dichroism data; θ indicates the molar ellipticity observed for each sample at 222 nm, θ_0 indicates the initial molar ellipticity at 222 nm and θ_{max} indicates the maximum molar ellipticity at 222 nm. For the luminescence data; *I* indicates the emission intensity or intensity area observed for each sample, I_0 indicates the initial emission intensity or intensity area and I_{max} indicates the maximum emission intensity or intensity area. This relationship assumes that the end of the binding event is directly related to the end of a change in the observed phenomena.

$$f = \frac{Concentration\ of\ Bound\ Trimer}{Concentration\ of\ Bound\ and\ Free\ Trimer} \qquad \text{(Equation 1A)}$$

$$f = \frac{\theta - \theta_0}{\theta_{max} - \theta_0} = \frac{I - I_0}{I_{max} - I_0} \qquad \text{(Equation 1B)}$$

For the most part the isotherm clearly indicated that the binding event was complete upon the ratio addition of one lanthanide ion to three peptide chains. Therefore the data was fitted to a three-to-one binding model (Scheme 1).

$$\text{Peptide} \quad + \quad 1/3\ \text{Tb(III)} \underset{K_d}{\overset{K_a}{\rightleftharpoons}} 1/3\ \text{Peptide}_3\text{-Tb(III)}$$

Scheme 1: The stoichiometric relationship involved in the binding of Tb(III) ions to any of the designed peptides. The binding constant (K_a) is associated with the forward reaction and the dissociation constant (K_d) is associated with the reverse reaction.

This scheme can be used to derive an equation for determining the dissociation constant which can be further rearranged to require only three variables (Equation 2); the concentration of the terbium trications, [Tb(III)], the total concentration of peptide present, [peptide], and the fraction bound (*f*).

$$K_d = \frac{([peptide]-f[peptide])([Tb(III)]-\frac{f[peptide]}{3})^{\frac{1}{3}}}{(\frac{f[peptide]}{3})^{\frac{1}{3}}} \qquad \text{(Equation 2)}$$

This dissociation constant can be rearranged to calculate the concentration of terbium trications required to induce the binding event (*see* Equation 3D). This was used to construct the fits to the experimentally collected data.

$$K_d{}^3 = \frac{([peptide]-f[peptide])^3([Tb(III)]-\frac{f[peptide]}{3})}{(\frac{f[peptide]}{3})} \qquad \text{(Equation 3A)}$$

$$\frac{fK_d{}^3[peptide]}{3} = ([peptide]-f[peptide])^3([Tb(III)]-\frac{f[peptide]}{3}) \qquad \text{(Equation 3B)}$$

$$\frac{fK_d{}^3}{3(1-f)^3[peptide]^2} = [Tb(III)]-\frac{f[peptide]}{3} \qquad \text{(Equation 3C)}$$

$$[Tb(III)] = f(\frac{K_d{}^3}{3(1-f)^3[peptide]^2}+\frac{[peptide]}{3}) \qquad \text{(Equation 3D)}$$

Once the dissociation constant (K_d) had been successfully fitted, the binding constant (K_a) was elucidated from the inverse relationship (*see* Equation 4).

$$K_a = \frac{1}{K_d} \qquad \text{(Equation 4)}$$

Throughout the thesis the log of the binding constant ($\log K_a$) is utilised as this value is a more accessible representation of the binding. The errors associated with the binding constant values are from the fits to both the upper and lower limits of the isotherm data. The isotherm data error bars come from the standard deviation of three repeats for each of the titrations.

8.10 LANTHANIDE LIFETIMES

The lanthanide lifetimes were determined for the peptide complexed lanthanides utilising an Edinburgh Instruments Fluorescence FL920 system equipped with a μF flashlamp light source (50 Hz). The bound lanthanide species were Tb(III), Eu(III), Dy(III) and Sm(III) in samples of aqueous peptide (90 μM) in 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid buffer (10 mM, pH 7.0) with the lanthanide trications (10 μM). The lanthanide concentration was deliberately lower than the expected 1:3 ratio to ensure that all emission signals came from bound lanthanide species. The Tb(III) emission was monitored at 545 nm, the Eu(III) at 616 nm, the Dy(III) at 571 nm and the Sm(III) at 643 nm. A 455 nm long pass filter was placed between the sample and the emission slit. Data collection for Tb(III) and Eu(III) was over a 10 ms time range with a lamp trigger delay of 0.01 ms. Data collection for Dy(III) and Sm(III) was over 100 μs with a lamp trigger delay of 1 μs. The excitation wavelength for the peptide

chromophore was dependent upon the species; tryptophan was excited at 280 nm, tyrosine at 275 nm, naphthyl at 280 nm and coumaryl at 320 nm. The data was collected and processed using F900 spectrometer analysis software with the tailfit model (*see* Equation 5). Data was fitted to preferentially a monoexponential single component model with a second component added if required as indicated by the χ^2 value. The decay model is a function of the time, *t*, and kinetic parameters. The pre-exponential factor (*B*) is made up of instrumental and sample parameters. The measured lifetime, τ is the time taken for the fluorescence to decay by 1 / e. Representative fits of the lifetimes for each system and tables of data can be found in the supporting information (*see* S38 - S46, S50 - S52, S55 – S58 and S61 – S62).

$$R(t) = B_1 e\left(\frac{-t}{\tau_1}\right) + B_2 e\left(\frac{-t}{\tau_2}\right) \qquad \text{(Equation 5)}$$

The lifetime experiment was repeated with Tb(III) in each of the peptide systems but utilising deuterium oxide for all solutions including the buffer, Tb(III) ions and peptide. All experimental parameters were retained except the time range was extended to 20 ms. The resultant data was fitted as described above and utilised to determine the number of water molecules in the first coordination sphere according to the work of A. Beeby *et al.* (*see* Equation 6).[7]

$$\text{Hydration Number} = 5.0\, ms \times \left(\frac{1}{\tau_H} - \frac{1}{\tau_D} - 0.06\, ms^{-1}\right) \qquad \text{(Equation 6)}$$

8.11 PEPTIDE CRYSTALLOGRAPHY

8.11.1 CRYSTAL GROWTH

Crystals of HC0C, HC02, HC03, HC01-2Y and HC01-2Nap with lanthanide ions were prepared using the sitting-drop vapour-diffusion technique at 18 °C with a drop consisting of 3 μL peptide stock solution (16 mg mL^{-1} peptide, 6 mM lanthanide chloride and 5 mM zinc acetate) mixed with 3 μL precipitant solution (0.21 M ammonium chloride, 25% v/v poly(ethylene) glycol 3350). The lanthanide was predominantly terbium, however, for HC02 and HC03 there was some success with europium and gadolinium. For HC0C, HC02 and HC03 triangular prism-like crystals belonging to the space group *H3* grew within a few weeks to a maximum dimension of approximately 0.6 × 0.2 mm. For HC01-2Y crystals of rectangular plate like shapes grew within a few days, whereas the crystals of HC01-2Nap were much smoother with dramatically curved edges. Both HC01-2Y and HC01-2Nap belonged to the space group *P41 21 2*. Crystals were transferred to a stabilization solution based on the mother liquor and were then cryo-protected by increasing the concentration of ethylene glycol (up to 25%) before flash-cooling and storage in a liquid nitrogen cooled dewar.

Crystals of HC04 with terbium were prepared using the sitting-drop vapour-diffusion technique at 18 °C with a drop consisting of 3 μL peptide stock solution (20 mg mL^{-1} peptide, 6 mM lanthanide chloride and 5 mM zinc acetate) mixed with 3 μL precipitant solution (0.21 M ammonium chloride, 20% v/v poly(ethylene) glycol 3350, 8% v/v poly(ethylene) glycol 400). Crystals of triangular prism like shapes grew within a few

weeks belonging to the space group *H3*. The crystals were transferred and stored as previously discussed.

8.11.2 X-RAY FLUORESCENCE MEASUREMENTS, ABSORPTION EDGE SCAN AND DIFFRACTION MEASUREMENTS

The datasets were all collected at Diamond Light Source, on the i04 or i03 beamlines. Crystals were place in numbered unipucks compatible with a BART robotic sample changer. The samples were kept under liquid nitrogen and automatically mounted in the beamline. The mounted crystals were maintained at cryo-temperatures by a cryo stream of N_2. The beamlines were equipped with Pilatus 6M-F (i04) and Pilatus3 6M (i03) detectors.

A crystal of HC03 with Tb(III) bound was analysed by x-ray fluorescence on the i04 beamline with a beam energy of 12658.2 eV (transmission 1.6%) and data collection occurred for 1.0 seconds. Sample counts recorded at 14051 compared to a background of 1184. The data was analysed manually, comparing the energy of noticeable peaks to those expected for ions present such as Tb(III) and Zn(II). This was followed by the HC03 bound Tb(III) absorption edge being scanned from 6800 - 7800 eV, focusing on the L_3 absorption band. The sudden increase in absorption was used to monitor the distinct core electronic signature for Tb(III) L_3 excitation. The peptide diffraction datasets were collected with scans around the phi axis and beam transmission at 100%. The overall collection time was between 180 and 300 seconds with the number of

images collected typically between 1500 and 3000. Each image was collected from between 0.08 and 0.04 seconds of beam exposure. The wavelength of the beam was typically between 0.9282 and 0.9795 Å. The technique utilised was single-wavelength anomalous diffraction.

8.11.3 X-RAY DIFFRACTION DATA PROCESSING

The datasets were indexed and integrated with iMosflm,[8] reduced with POINTLESS,[9] scaled using AIMLESS and structure amplitudes were calculated using CTRUNCATE.[10, 11] The Tb(III) ions provided heavy metal anomalous signals, from which their positions were located and phases determined using Phaser within the Phenix pipeline.[12, 13] This resulted in clearly interpretable electron density maps with particularly clear definition around the Tb(III) binding sites. Coot was used for manual refinement while phenix.refine with MLHL targets was used for intervals of automated refinement.[14, 15] Both manual and automated strategies were combined to complete the peptide models. Data-collection, refinement and validation statistics are presented in relevant chapters (*see* Chapter 2 and Chapter 4). Figures were generated in UCSF Chimera.[16]

8.12 SMALL MOLECULE SYNTHESIS

8.12.1 *N*-ACETYL-L-TYROSINE METHYL ESTER (1)

This synthesis is based on a literature procedure.[17] To a solution of *L*-tyrosine methyl ester (10.0 g, 0.05 mol) in acetic acid (100 mL) was added acetic anhydride (9 ml, 0.2 mol), and the mixture was refluxed for 10 min and allowed to stand at room temperature for 2 h. Acetic acid was evaporated *in vacuo*, and the residue was crystallized from methanol and water to yield *N*-acetyl-*L*-tyrosine methyl ester as small white needles. Yield: 10.8 g, 91.1%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 1.79 (3H, s, NAc-CH_3), 2.74 (1H, dd, J = 13.8, 9.1 Hz, β-CH_2), 2.86 (1H, dd, J = 13.8, 5.7 Hz, β-CH_2), 3.57 (3H, s, OCH_3), 4.34 (1H, m, α-CH), 6.65 (2H, d, J = 8.4 Hz, ArH), 6.98 (2H, d, J = 8.4 Hz, ArH), 8.28 (1H, d, J = 7.6 Hz, NH), 9.24 (1H, s, ArOH) ppm; ^{13}C NMR (400 MHz, DMSO-d_6): δ = 22.26 (NAc-CH_3), 36.05 (β-CH_2), 51.74 (OCH_3), 54.00 (α-CH), 115.05 (Ar-CH), 127.23 (Ar-CC), 130.02 (Ar-CH), 155.99 (Ar-COH), 169.33 (CONH), 172.38 (COOMe) ppm; IR (neat): ν = 3546, 3320, 3149, 1729, 1658, 1516, 1232, 820, 690 cm^{-1}; MS (ESI): *m/z* 237.10 (M^+), 178.06 (M^+ – COOMe), 147.04 (M^+ – C_6H_5OH), 107.05 (M^+ – CH(NAc)COOMe).

8.12.2 *N*-ACETYL-L-4-METHYLCOUMARYLALANINE (2)

To a solution of *N*-acetyl-*L*-tyrosine methyl ester **1** (5.0 g, 21 mmol) in aqueous 70% sulphuric acid (25 mL) was added ethyl acetoacetate (3.5 mL, 27 mmol). The reaction

mixture was stirred at room temperature for 72 hours, poured into ice, and the precipitate was filtered and washed with ice-cold water (2 x 10 mL), saturated sodium bicarbonate solution (2 x 10 mL) and finally water (2 x 10 mL). The collected precipitate was recrystallized from an ethanol and water mixture to yield *N*-acetyl-*L*-4-methylcoumarylalanine as square plates of pale cream crystals. Yield: 0.24 g, 3.9%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 1.78 (3H, s, NAc-CH_3), 2.42 (3H, d, *J* = 1.2 Hz, Ar-CH_3), 2.91 (1H, dd, *J* = 13.8, 9.7 Hz, *β*-CH_2), 3.13 (1H, dd, *J* = 13.8, 4.8 Hz, *β*-CH_2), 4.45 (1H, m, *α*-CH), 6.39 (1H, s, Ar-H), 7.31 (1H, d, *J* = 8.5 Hz, Ar-H), 7.48 (1H, dd, *J* = 8.5, 2.0 Hz, Ar-H), 7.63 (1H, d, *J* = 2.0 Hz, Ar-H) 8.24 (1H, d, *J* = 8.2 Hz, NH), 12.73 (1H, s, COOH) ppm; ^{13}C NMR (400 MHz, DMSO-d_6): δ = 18.07 (Ar-CH_3), 22.33 (NAc-CH_3), 36.04 (*β*-CH_2), 53.34 (*α*-CH), 114.40 (Ar-CH), 116.20 (Ar-CH), 119.26 (Ar-CC), 125.74 (Ar-CCH_2), 132.94 (Ar-CH), 133.89 (Ar-CCH_3), 151.71 (Ar-CH), 153.14 (Ar-CO), 159.82 (Ar-COO), 169.27 (CONH), 173.03 (COOH) ppm; IR (neat): ν = 3350, 2935, 2511, 1722, 1611, 1541, 1174, 826, 625 cm^{-1}; HRMS (ES): *m/z* calculated for $C_{15}H_{16}NO_5$, $[M+H]^+$: 290.1028, found 290.1027. Single crystal x-ray data can be found in the supporting information (*see* S72 – S77).

8.12.3 *N*-ACETYL-L-4-CHLOROMETHYLCOUMARYLALANINE (3)

To a solution of *N*-acetyl-*L*-tyrosine methyl ester **1** (5.0 g, 21 mmol) in aqueous 70% sulphuric acid (25 mL) was added ethyl 4-chloroacetoacetate (4 mL, 30 mmol). The reaction mixture was stirred at room temperature for 72 hours, poured into ice, and the precipitate was filtered and washed with ice-cold water (2 x 10 mL), saturated sodium

bicarbonate solution (2 x 10 mL) and finally water (2 x 10 mL). The collected precipitate was purified by silica column chromatography with an eluent of dichloromethane and methanol (9:1) to yield *N*-acetyl-*L*-4-chloromethyl-coumarylalanine as a pale cream powder. Yield: 1.6 g, 23%; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 1.78 (3H, s, CH_3), 2.92 (1H, dd, *J* = 13.8, 9.8 Hz, *β*-CH_2), 3.14 (1H, dd, *J* = 13.8, 4.6 Hz, *β*-CH_2), 4.43 (1H, m, *α*-CH), 5.00 (2H, s, CH_2Cl), 6.68 (1H, s, Ar-CH), 7.37 (1H, d, *J* = 8.5 Hz, Ar-CH), 7.53 (1H, dd, *J* = 8.5, 2.0 Hz, Ar-CH), 7.74 (1H, d, *J* = 2.0 Hz, Ar-CH), 8.23 (1H, d, *J* = 8.2 Hz, NH) ppm; ^{13}C NMR (400 MHz, DMSO-d_6): δ = 22.37 (NAc-CH_3), 36.01 (*β*-CH_2), 41.16 (Ar-CH_2Cl), 53.39 (*α*-CH), 115.50 (Ar-CH), 116.56 (Ar-CH), 116.72 (Ar-CC), 125.60 (Ar-CCH_2), 133.36 (Ar-CH), 134.12 (Ar-CCH_2Cl), 150.52 (Ar-CH), 152.10 (Ar-CO), 159.67 (Ar-COO), 169.32 (CONH), 172.97 (COOH) ppm; IR (neat): ν = 3415, 3351, 3271, 3089, 2562, 1716, 1624, 1545, 1230, 1174, 958, 904, 826 cm^{-1}; HRMS (ES): *m/z* calculated for $C_{15}H_{13}NO_5Cl$, $[M]^+$: 322.0482, found 322.0481.

8.13 REFERENCES

[1] i) M. R. Berwick, PhD Thesis, University of Birmingham, 2016; ii) L. N. Slope, PhD Thesis, University of Birmingham, 2018 and iii) S. L. Newton, PhD Thesis, University of Birmingham, 2018.

[2] W. C. Chan and P. D. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, New York, 2002.

[3] A. Barge, G. Cravotto, E. Gianolio and F. Fedeli, *Contrast Medio Mol. Imaging*, 2006, **1**, 184.

[4] H. Edelhoch, *Biochemistry*, 1967, **6**, 1948-1954.

[5] i) S. Hamai, *Bull. Chem. Soc. Jpn.*, 1982, **55**, 2721-2729; ii) R. X. F. Ren, N. C. Chaudhuri, P. L. Paris and E. T. Kool, *J. Am. Chem. Soc.*, 1996, **118**, 7671-7678.

[6] i) Y. H. Chen, J. T. Yang and K. H. Chau, *Biochemistry*, 1974, **13**, 3350-3359; ii) J. K. Myers, C. N. Pace and J. M. Scholtz, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 2833-2837.

[7] A. Beeby, I. M. Clarkson, R. S. Dickins, S. Faulkner, D. Parker, L. Royle, A. S. DeSousa, J. A. G. Williams and M. Woods, *J. Chem. Soc. Perkin Trans. 2*, 1999, **2**, 493-504.

[8] i) T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. W. Leslie, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2011, **67**, 271-281; ii) M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2011, **67**, 235-242.

[9] i) P. R. Evans, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2006, **62**, 72-82; ii) P. R. Evans, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2011, **67**, 282-292.

[10] P. R. Evans and G. N. Murshudov, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2013, **69**, 1204-1214.

[11] i) S. French and K. Wilson, *Acta Crystallogr.*, *Sect A: Found. Crystallogr.*, 1978, **34**, 517-525; ii) J. E. Padilla and T. O. Yeates, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2003, **59**, 1124-1130; iii) A. N. Popov and G. P. Bourenkov, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2003, **59**, 1145-1153.

[12] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.

[13] P. D. Adams, R. W. Grosse-Kunstleve, L. –W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2002, **58**, 1948-1954.

[14] P. Emsley and K. Cowtan, *Acta Crystallogr.*, *Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126-2132.

[15] P. V. Afonine, R. W. Grosse-Kunstleve and P. D. Adams, *CCP4 Newsl. Protein Crystallogr.*, 2005, **42**, 8.

[16] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.

[17] X. W. Liao, W. Liu, W. F. Dong, B. H. Guan, S. Z. Chen and Z. Z. Liu, *Tetrahedron*, 2009, **65**, 5709-5715.

Chapter 9

SUPPORTING INFORMATION

9.1 SYNTHESISED PEPTIDE CHARACTERISATION

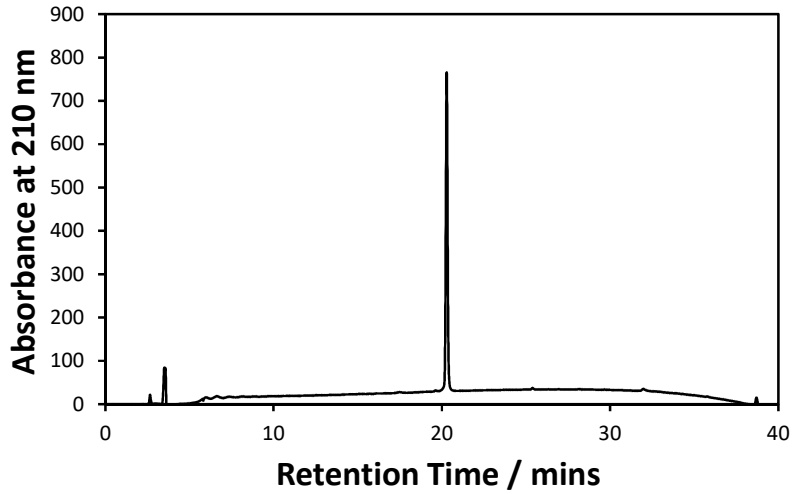

Figure 1: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC0C. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

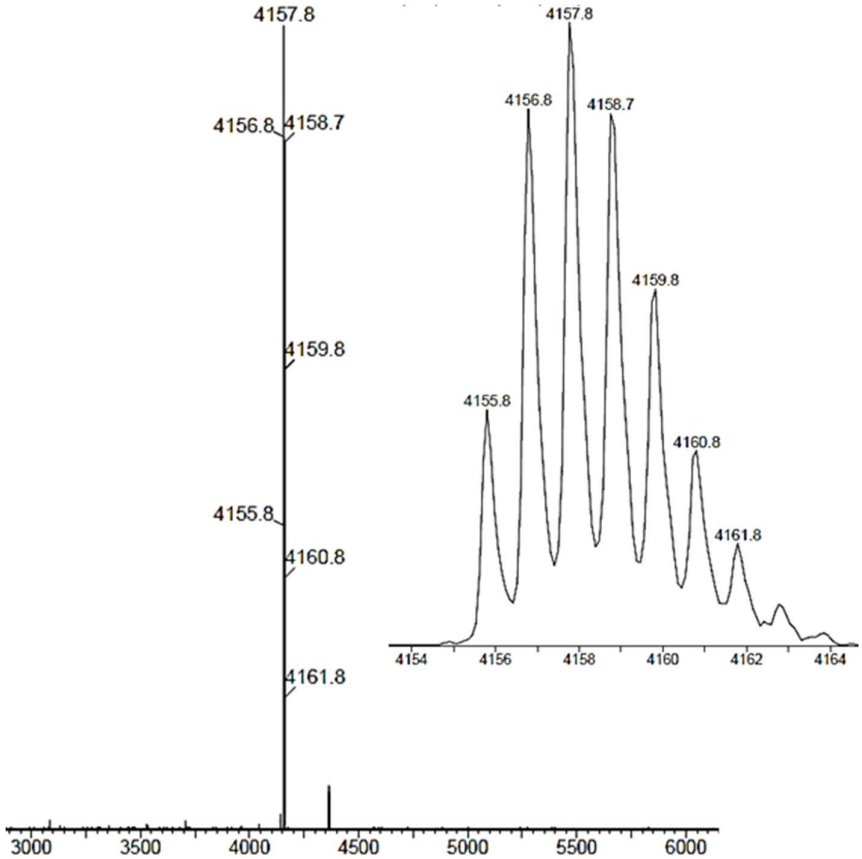

Figure 2: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC0C in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

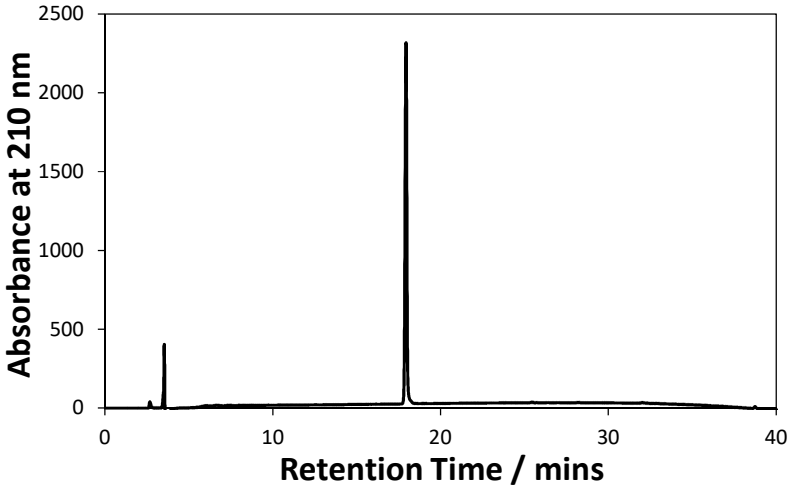

Figure 3: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC01. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

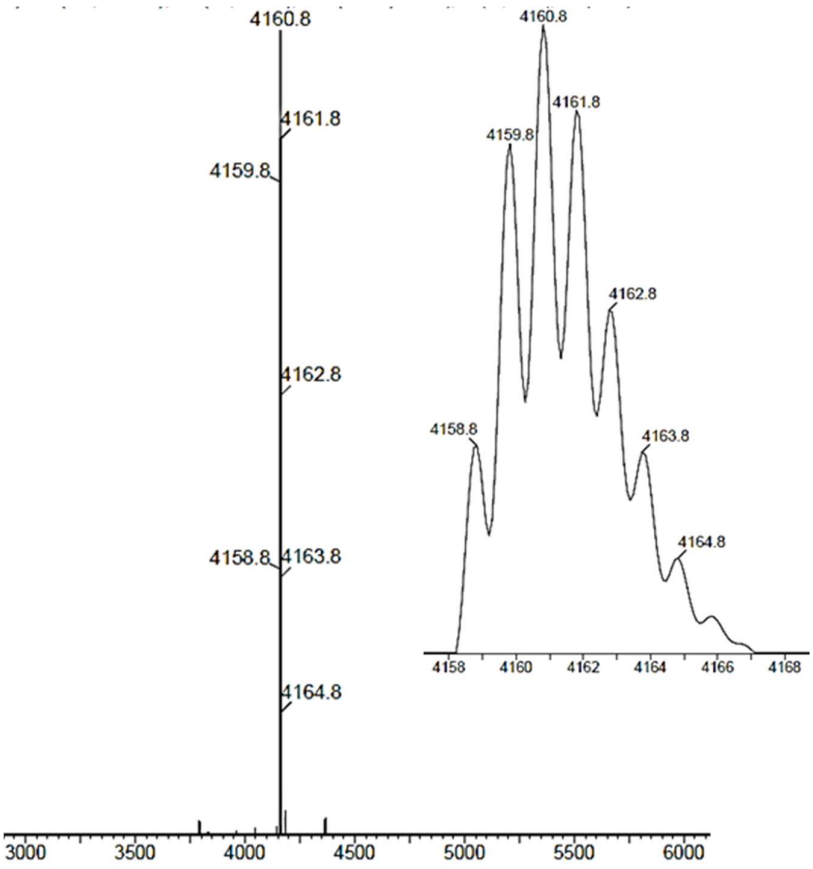

Figure 4: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC01 in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

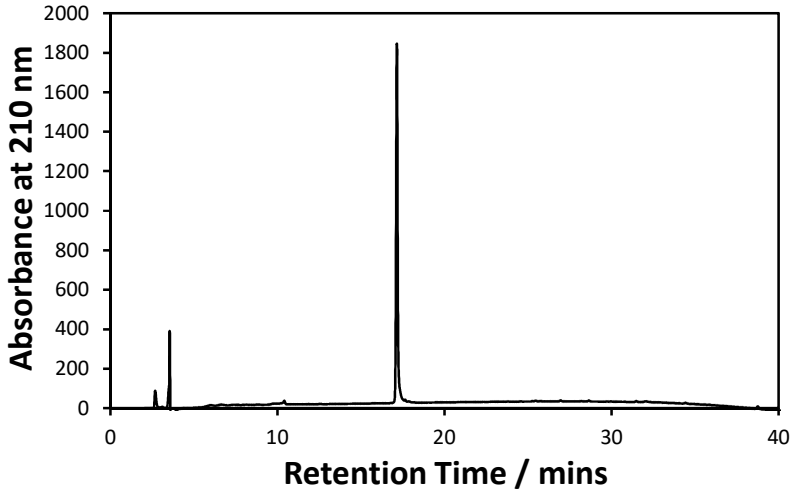

Figure 5: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC02. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

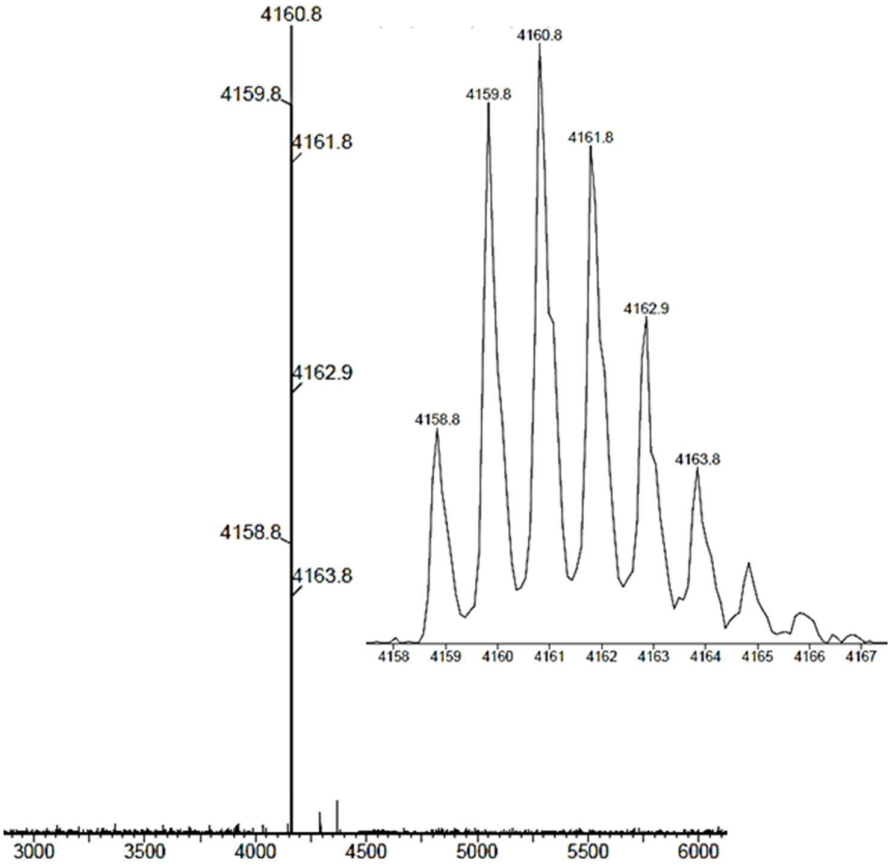

Figure 6: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC02 in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

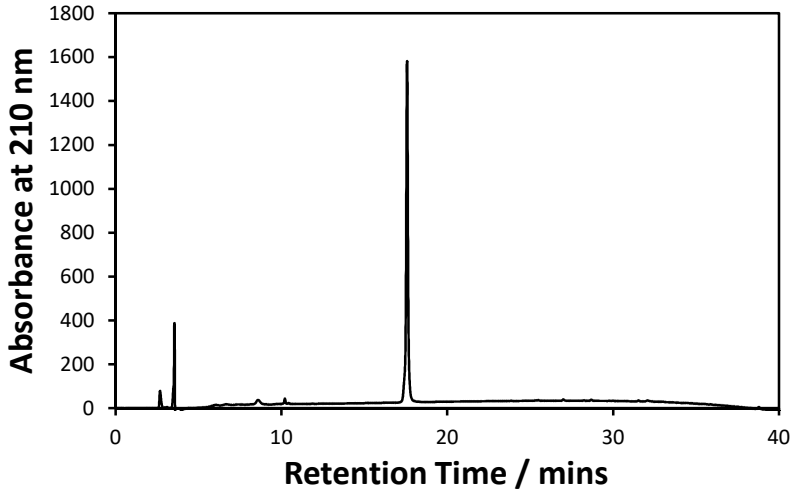

Figure 7: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC03. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

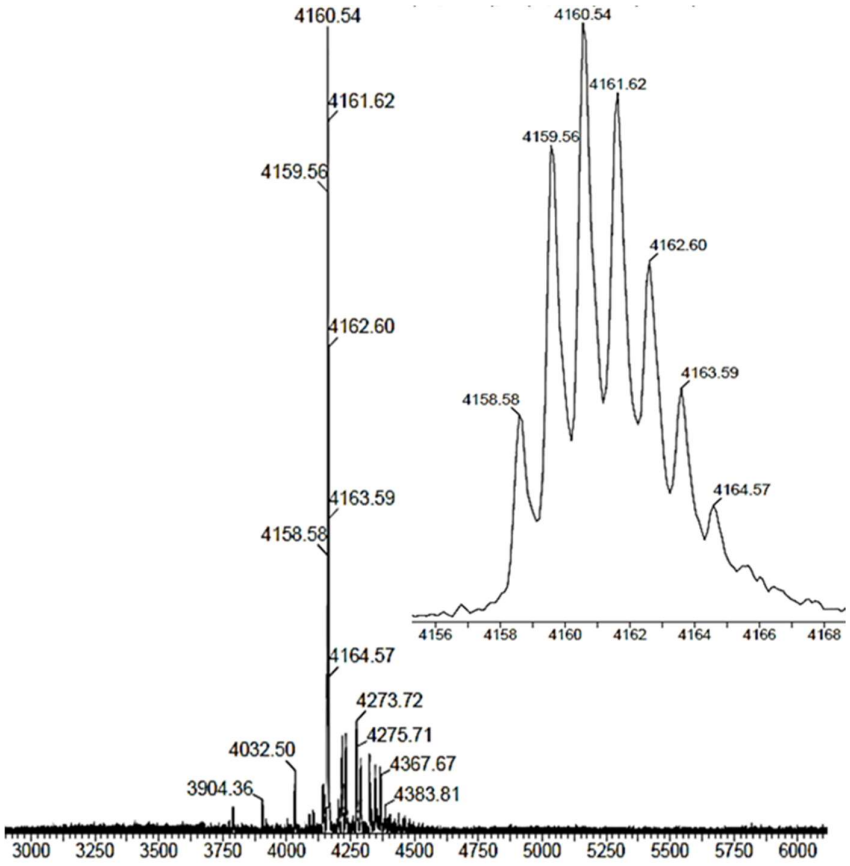

Figure 8: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC03 in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

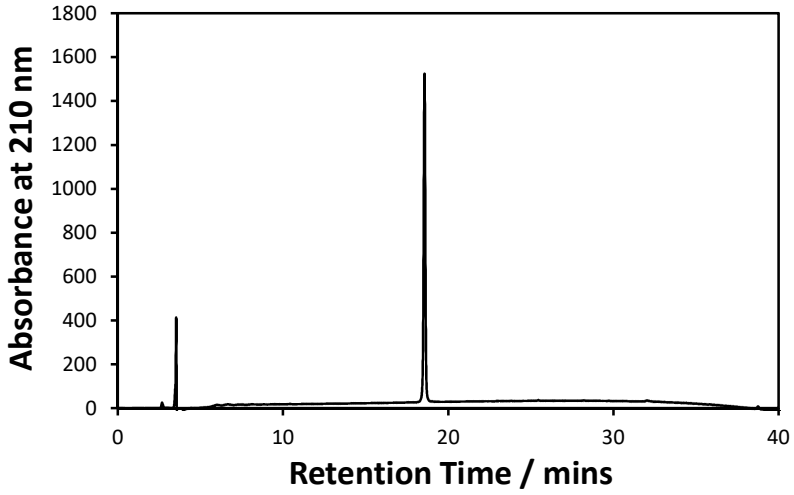

Figure 9: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC04. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

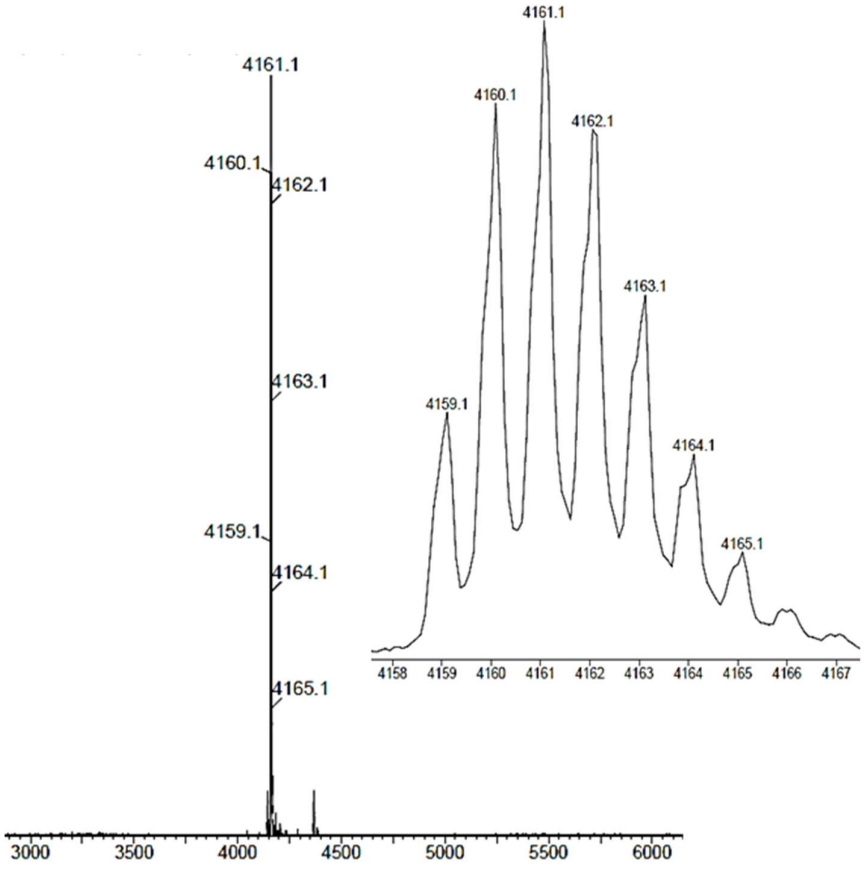

Figure 10: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC04 in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

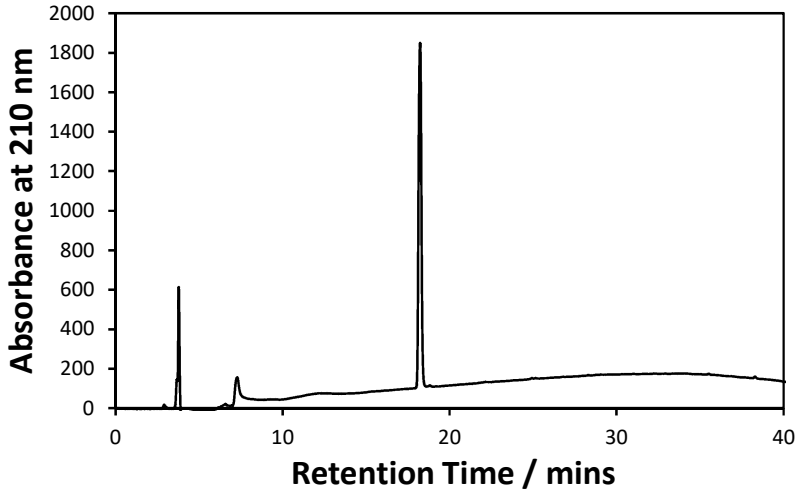

Figure 11: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC01-2Y. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

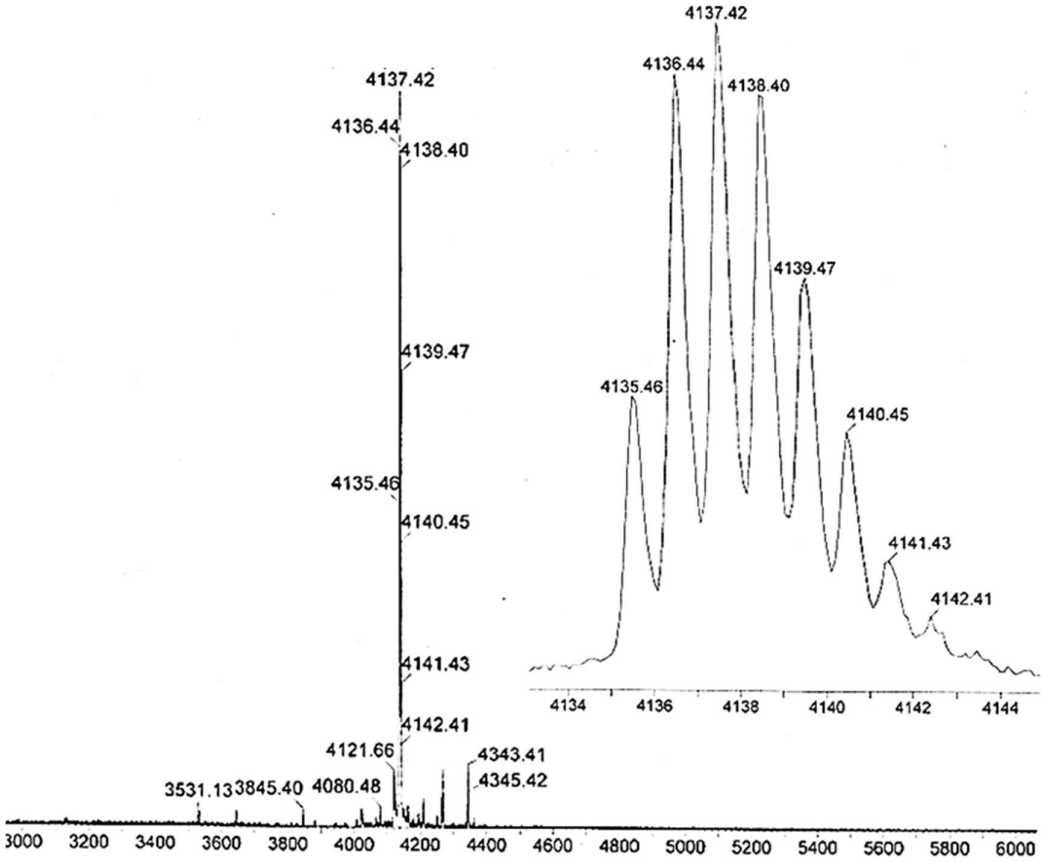

Figure 12: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC01-2Y in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

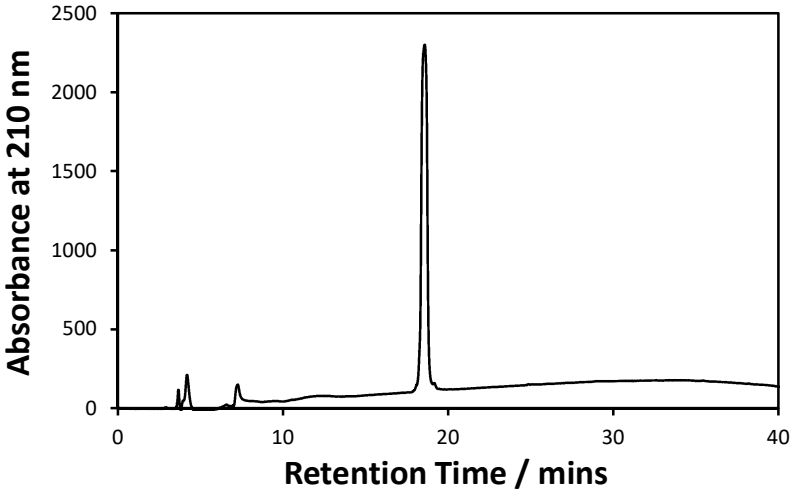

Figure 13: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC01-2Nap. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

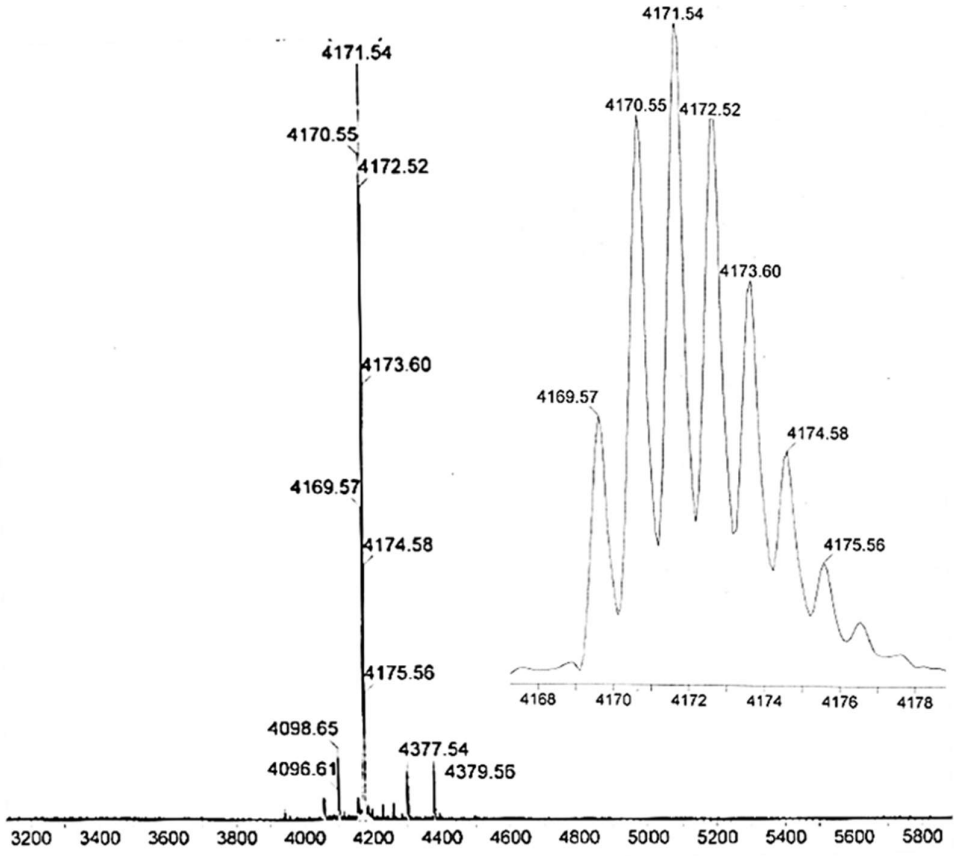

Figure 14: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC01-2Nap in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

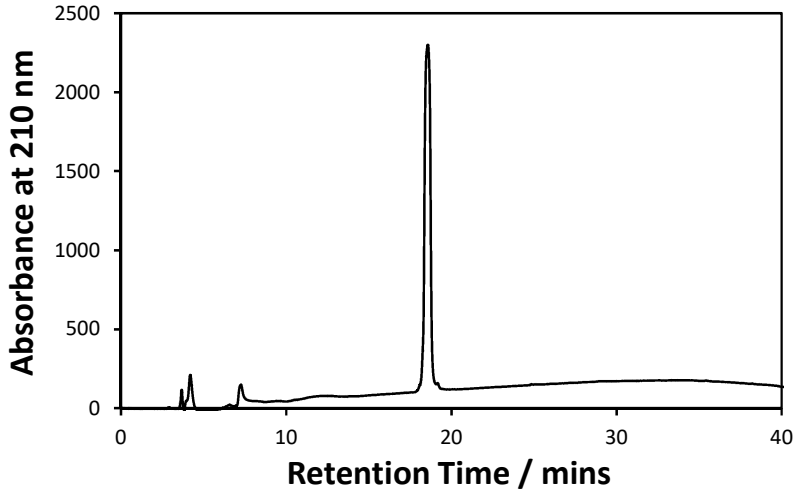

Figure 15: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure HC01-2Cou. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

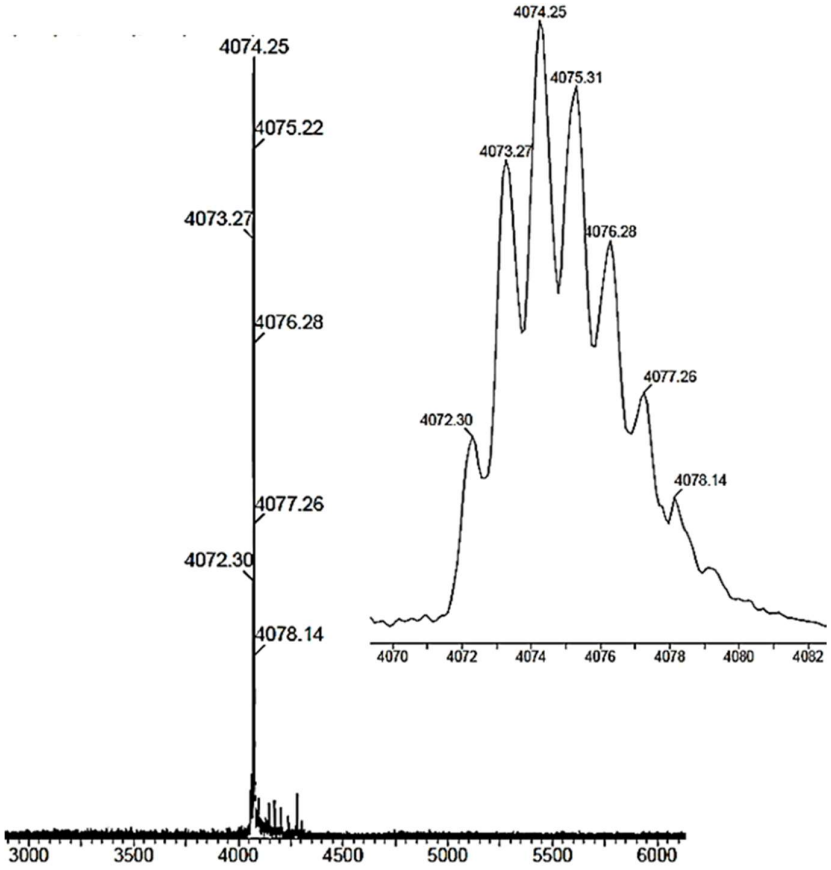

Figure 16: The matrix-assisted laser desorption/ionisation mass spectrum for pure HC01-2Cou in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

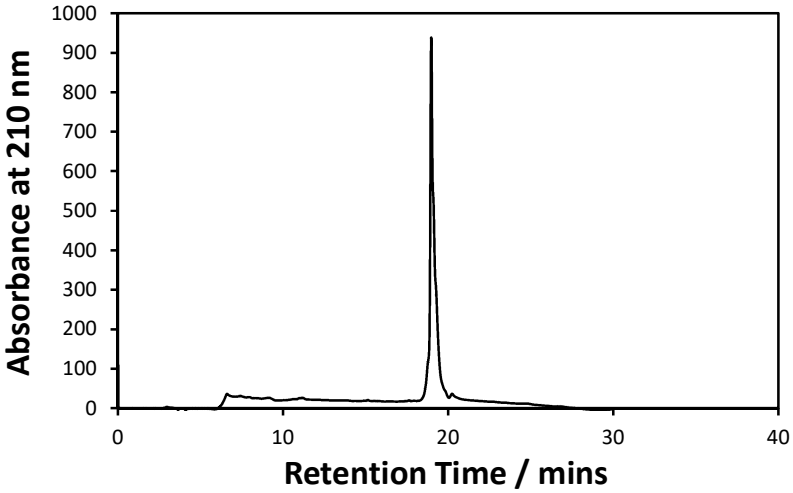

Figure 17: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure MB1-2. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

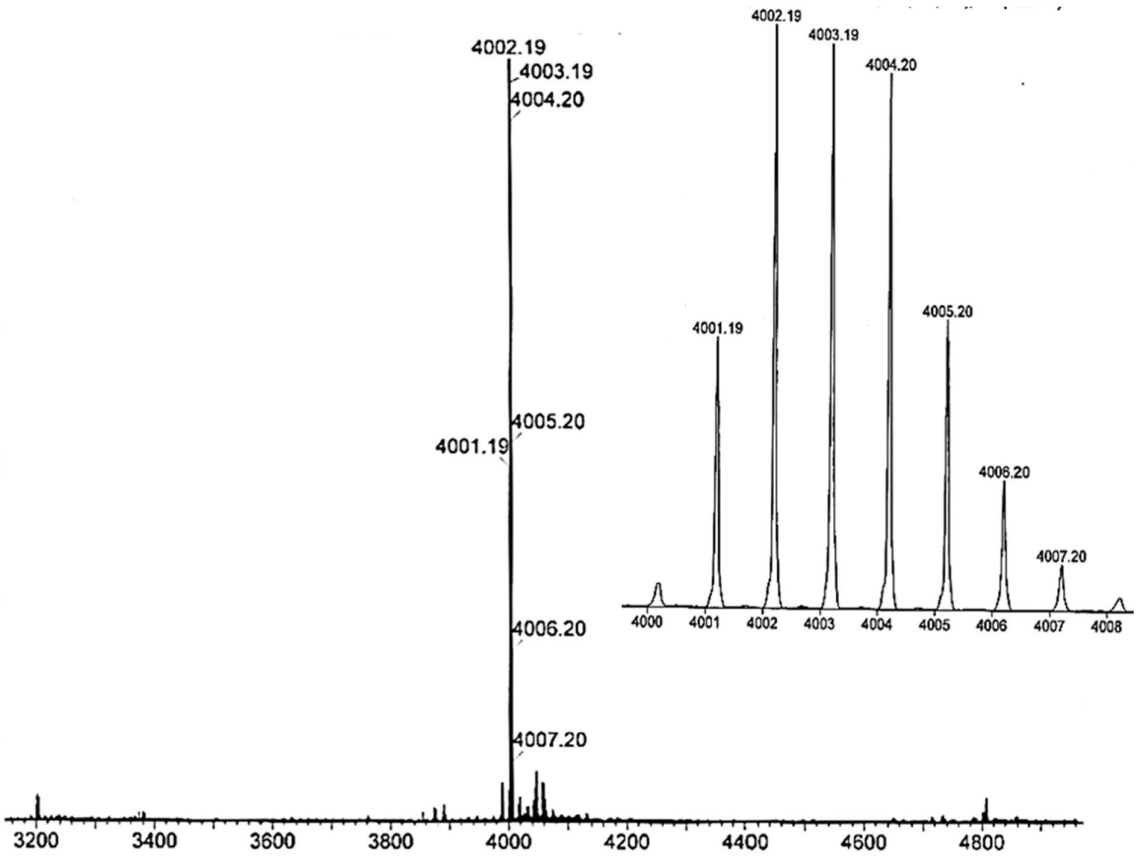

Figure 18: The matrix-assisted laser desorption/ionisation mass spectrum for pure MB1-2 in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

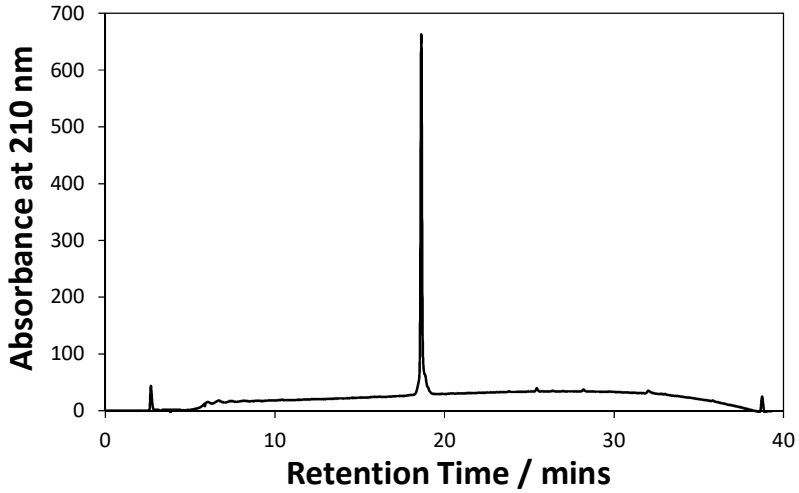

Figure 19: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure MB1-2,14Nap. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

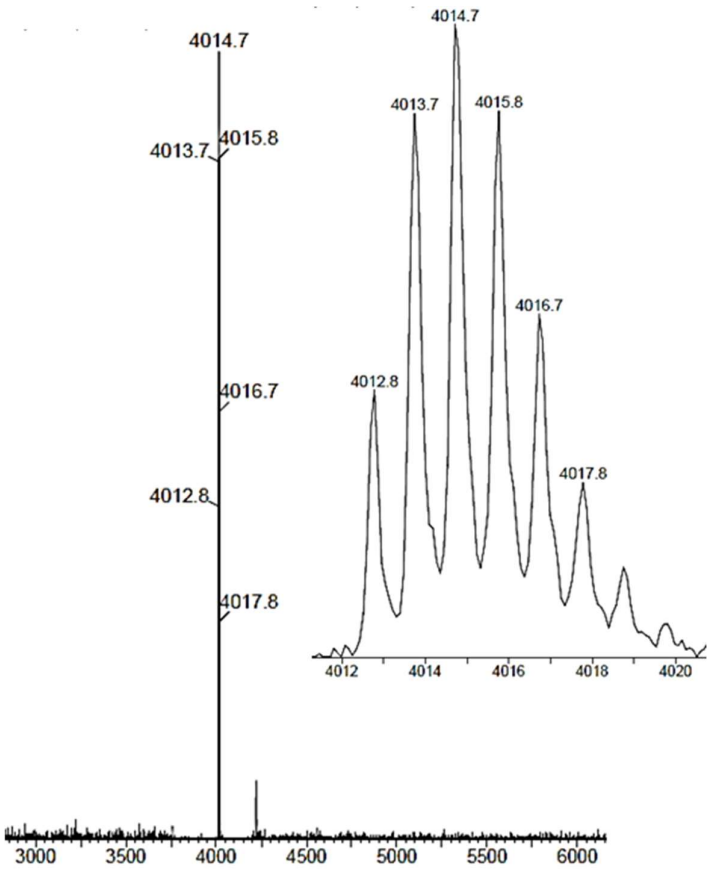

Figure 20: The matrix-assisted laser desorption/ionisation mass spectrum for pure MB1-2,14Nap in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

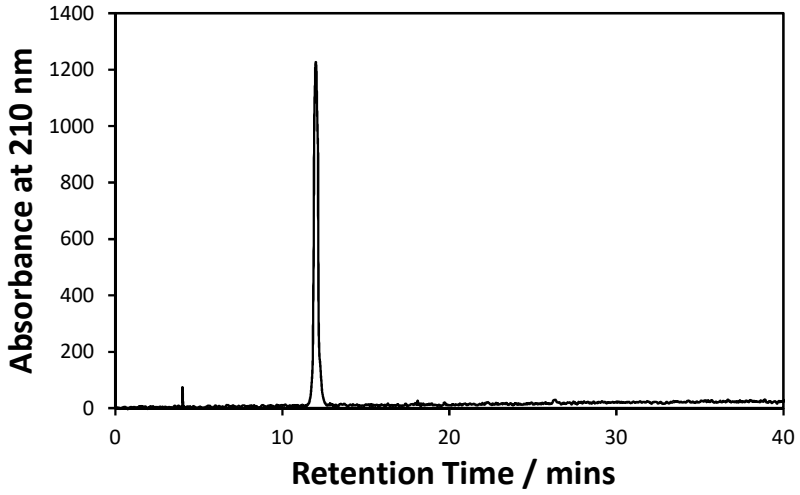

Figure 21: The C_{18} reverse phase analytical high pressure liquid chromatogram for pure LBT. The absorbance was monitored at 210 nm and the solvent system went from 100% water with 0.05% trifluoroacetic acid to 100% acetonitrile with 0.05% trifluoroacetic acid over 40 minutes.

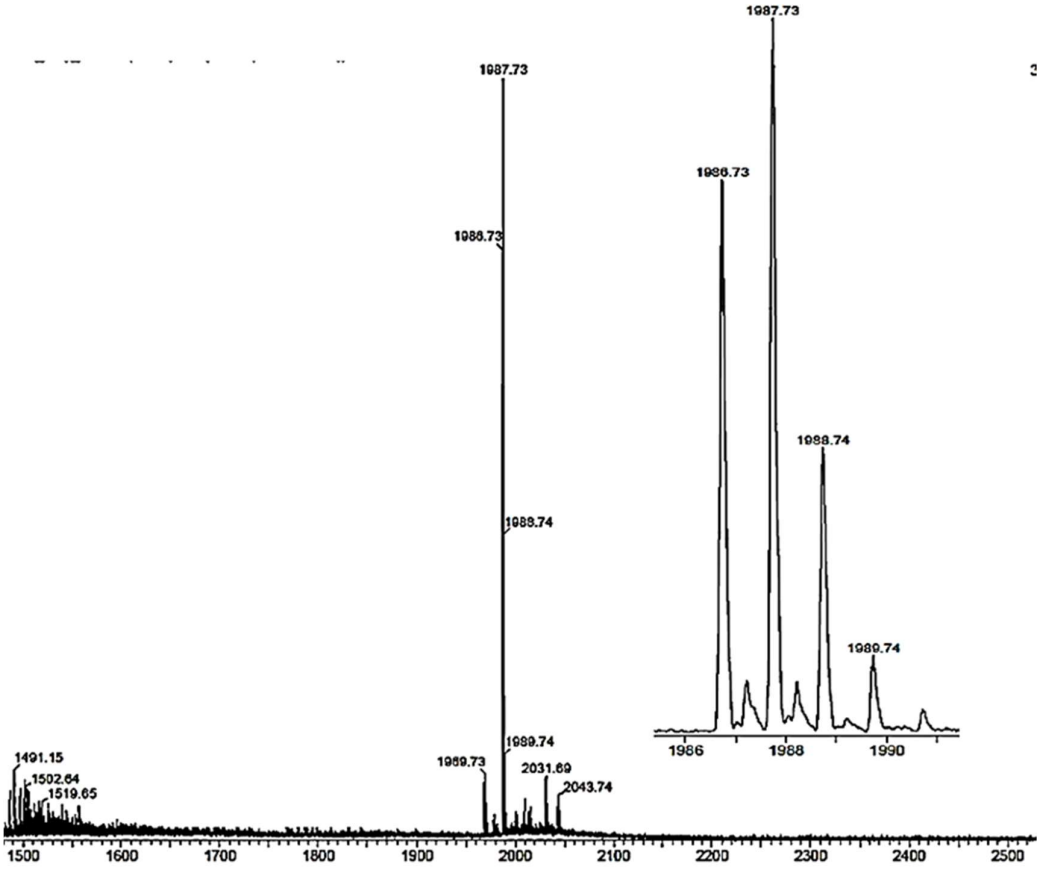

Figure 22: The matrix-assisted laser desorption/ionisation mass spectrum for pure LBT in a sinapinic acid matrix. The insert shows the isotope distribution of the +1 charge peak cluster.

9.2 CIRCULAR DICHROISM COILED COIL RATIO (θ_{222} / θ_{208})

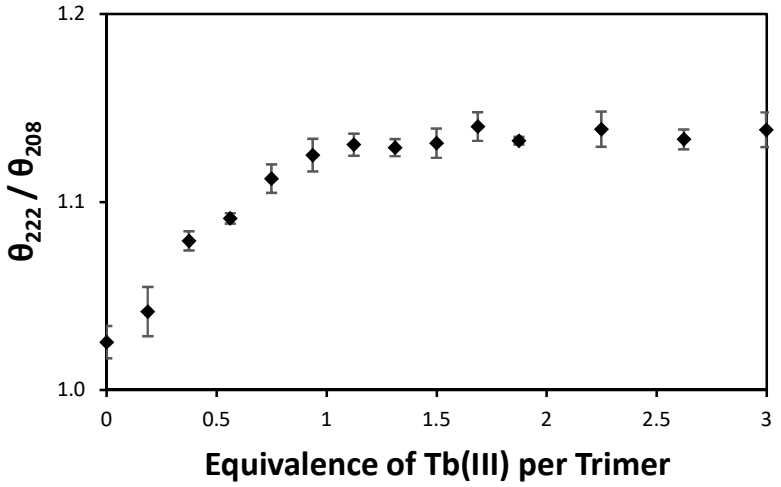

Figure 23: The increase in the ratio of molar ellipticity at 222 nm and 208 nm for HC01 upon the addition of Tb(III), which indicates the formation of coiled coil superstructures. Sample was of HC01 (90 μM), HEPES buffer (10 mM, pH 7.0) and the addition of Tb(III) up to 3 equivalences in water.

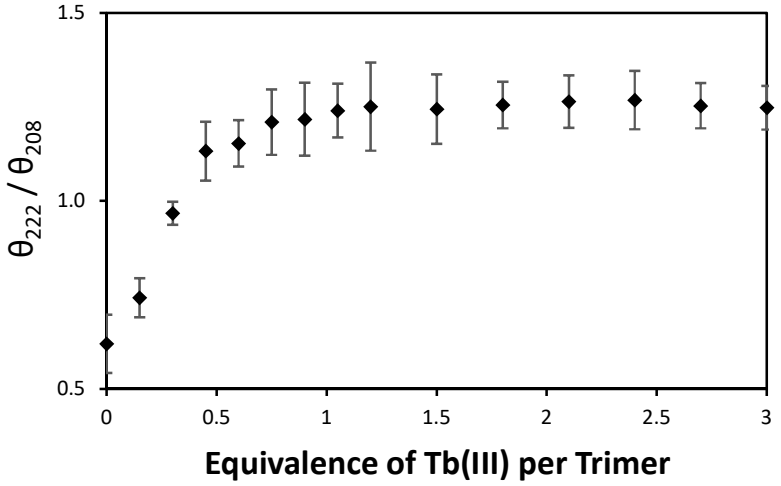

Figure 24: The increase in the ratio of molar ellipticity at 222 nm and 208 nm for HC02 upon the addition of Tb(III), which indicates the formation of coiled coil superstructures. Sample consisted of HC02 (90 μM), HEPES buffer (10 mM, pH 7.0) and the addition of Tb(III) up to 3 equivalences in water.

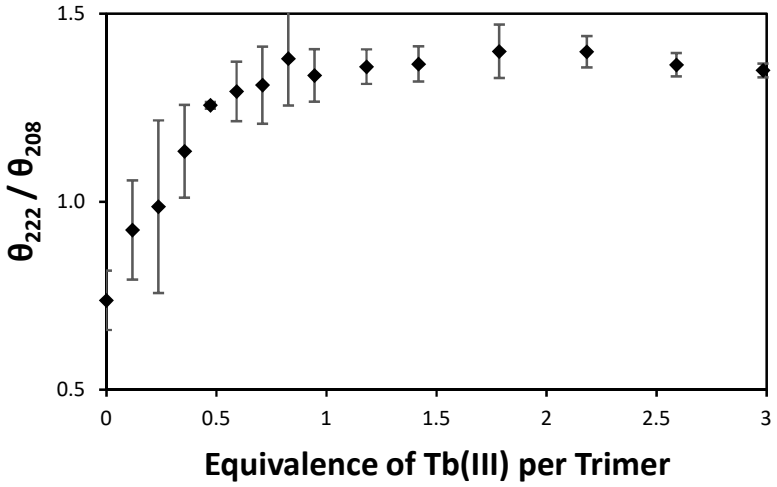

Figure 25: The increase in the ratio of molar ellipticity at 222 nm and 208 nm for HC03 upon the addition of Tb(III), which indicates the formation of coiled coil superstructures. Sample consisted of HC03 (90 μM), HEPES buffer (10 mM, pH 7.0) and the addition of Tb(III) up to 3 equivalences in water.

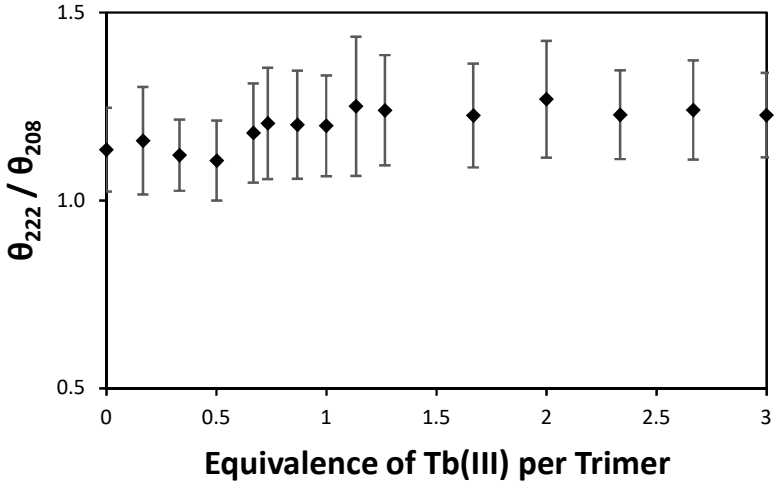

Figure 26: The ratio of molar ellipticity at 222 nm and 208 nm for HC04 upon the addition of Tb(III), which indicates no change in the formation of coiled coil superstructures. Sample consisted of HC04 (90 μM), HEPES buffer (10 mM, pH 7.0) and the addition of Tb(III) up to 3 equivalences in water.

9.3 ELECTROSPRAY IONSIATION MASS SPECTROMETRY

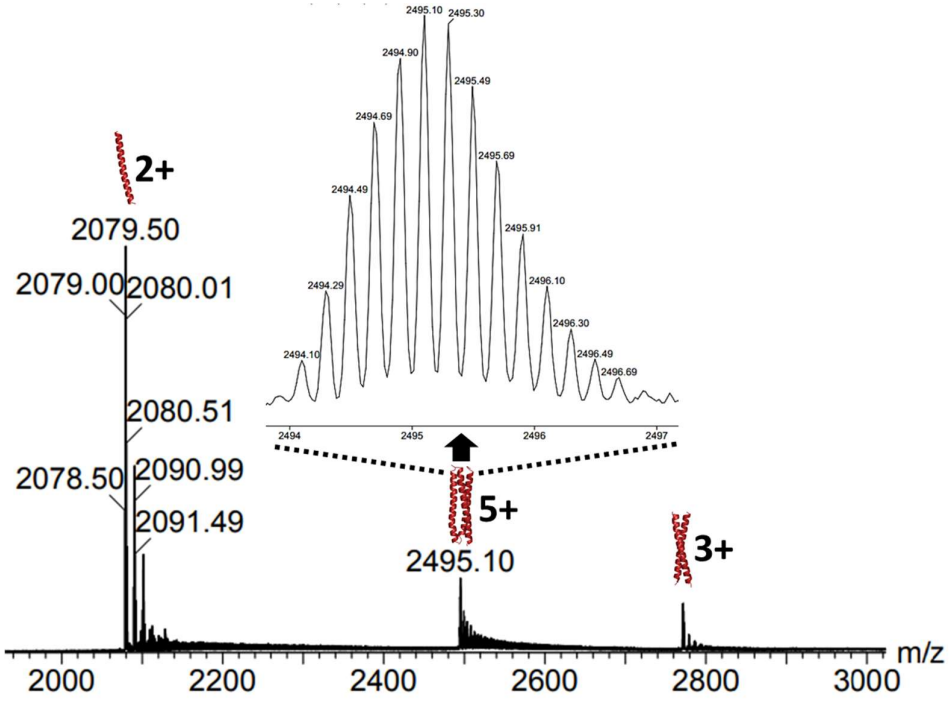

Figure 27: The mass spectrum of HC0C indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC0C (9 µM) in ammonium acetate buffer (10 mM, pH 6.4).

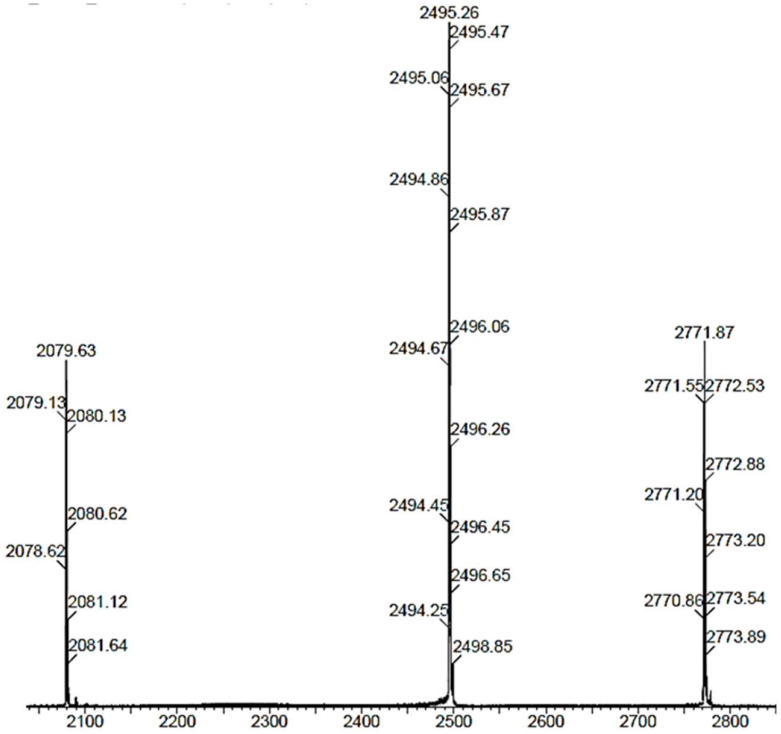

Figure 28: The tandem mass spectrum for the locked +5 charge peak from a trimer species of HC0C illustrating the presence of the degradation products being a dimer and a monomer.

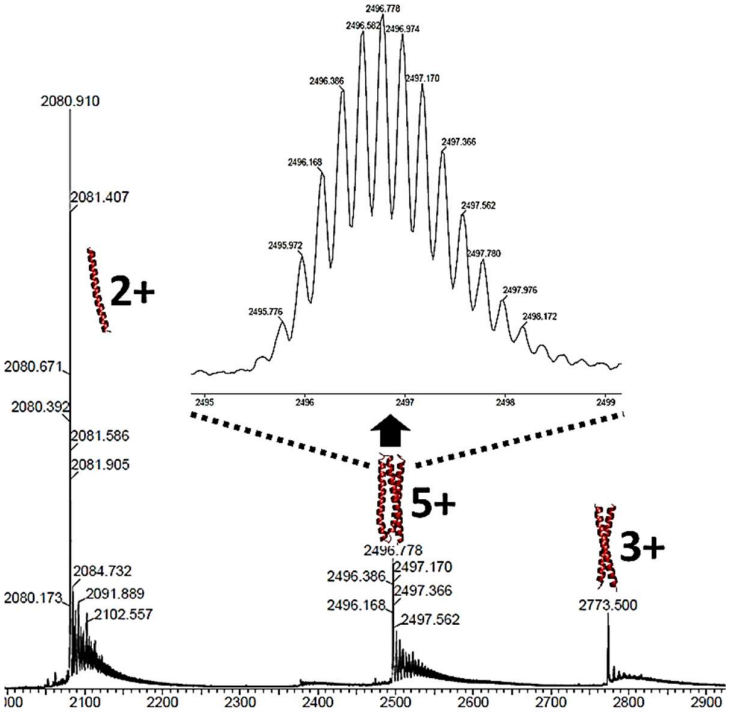

Figure 29: The mass spectrum of native HC01 indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC01 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4).

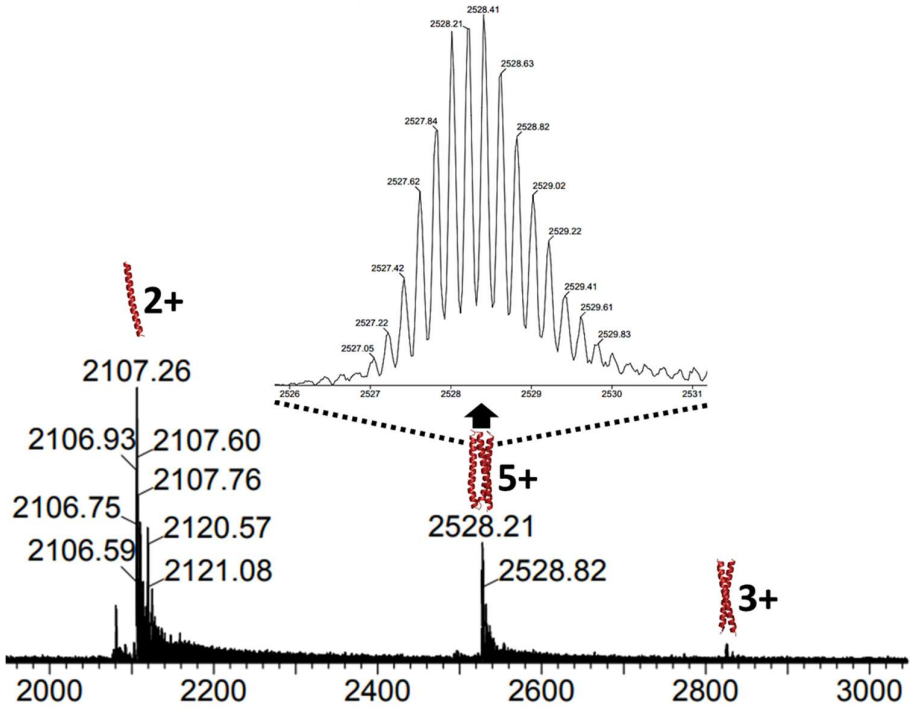

Figure 30: The mass spectrum of native HC01 with Tb(III) indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC01 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4) and Tb(III) ions at 1 equivalence to the trimer.

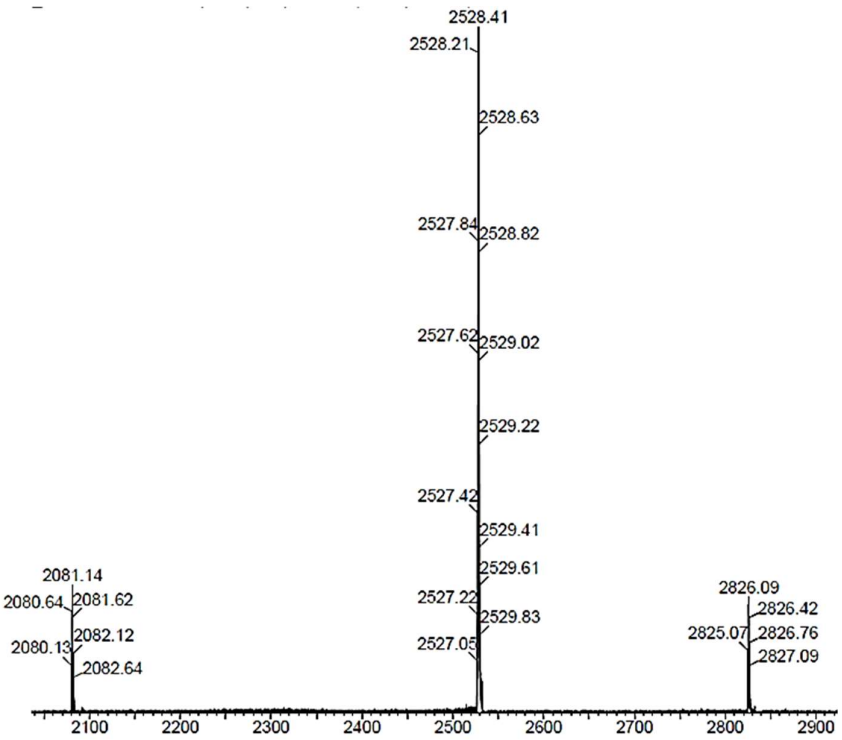

Figure 31: The tandem mass spectrum for the locked +5 charge peak from a trimer species of HC01 with Tb(III) ions illustrating the presence of the degradation products being a dimer with the metal ion and a monomer.

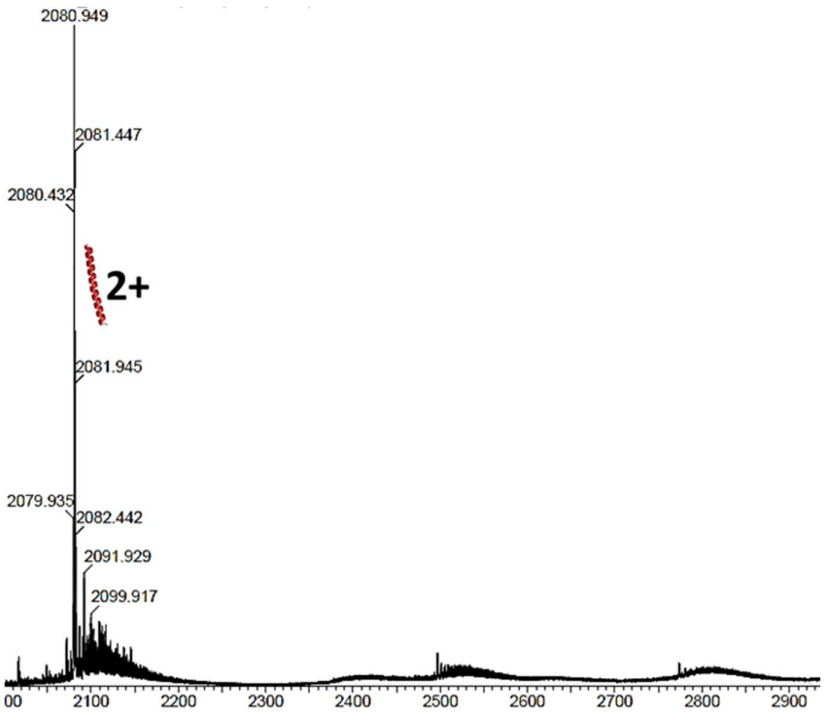

Figure 32: The mass spectrum of native HC02 indicating the different charged species resulting from different oligomeric states. The sample consisted of HC02 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4).

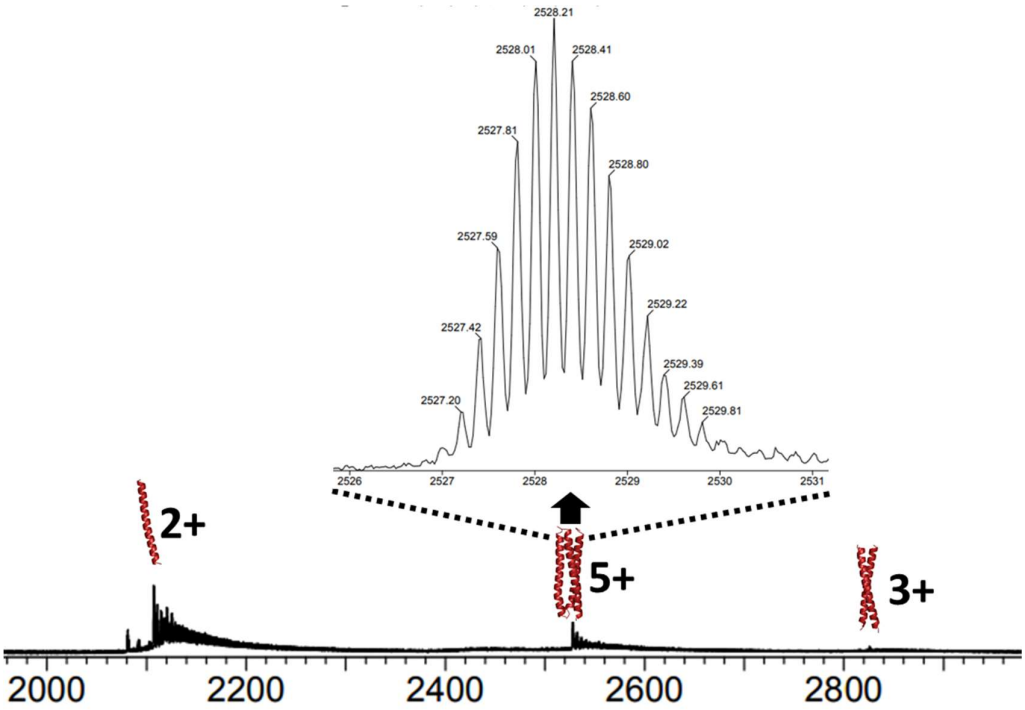

Figure 33: The mass spectrum of native HC02 with Tb(III) indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC02 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4) and Tb(III) ions at 1 equivalence to the trimer.

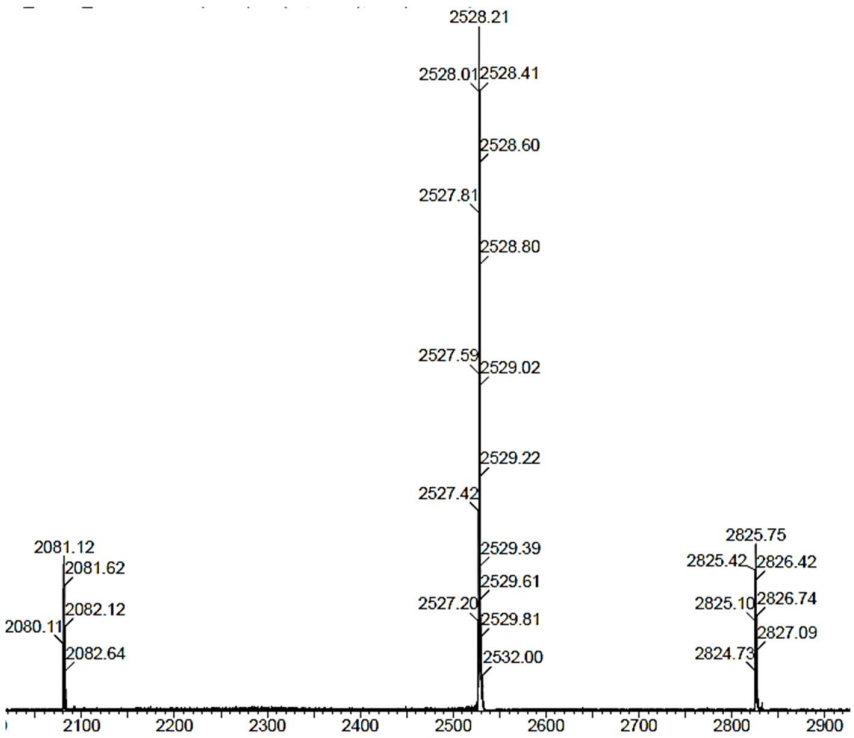

Figure 34: The tandem mass spectrum for the locked +5 charge peak from a trimer species of HC02 with Tb(III) ions illustrating the presence of the degradation products being a dimer with the metal ion and a monomer.

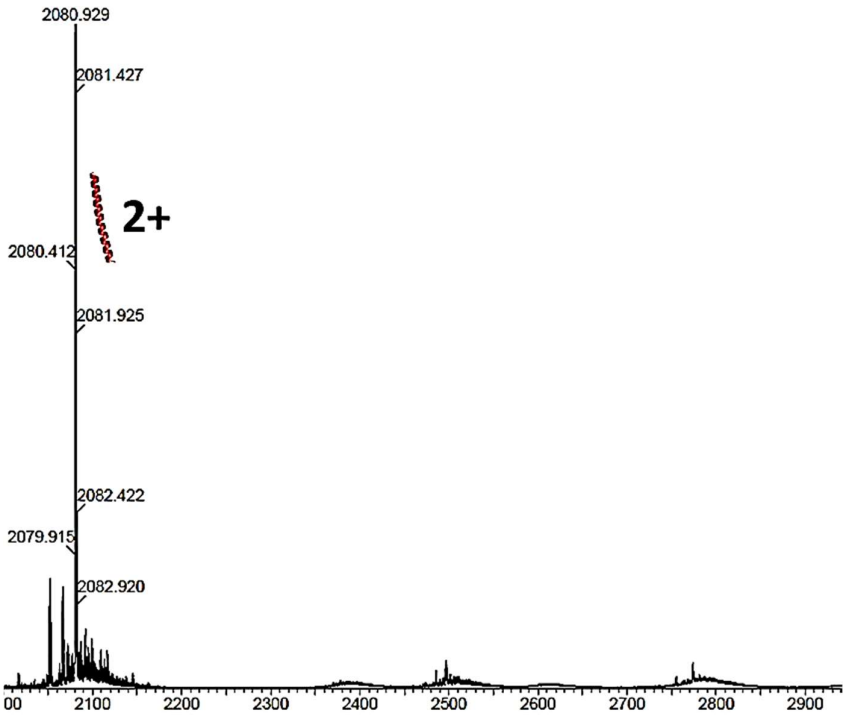

Figure 35: The mass spectrum of native HC03 indicating the different charged species resulting from different oligomeric states. The sample consisted of HC03 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4).

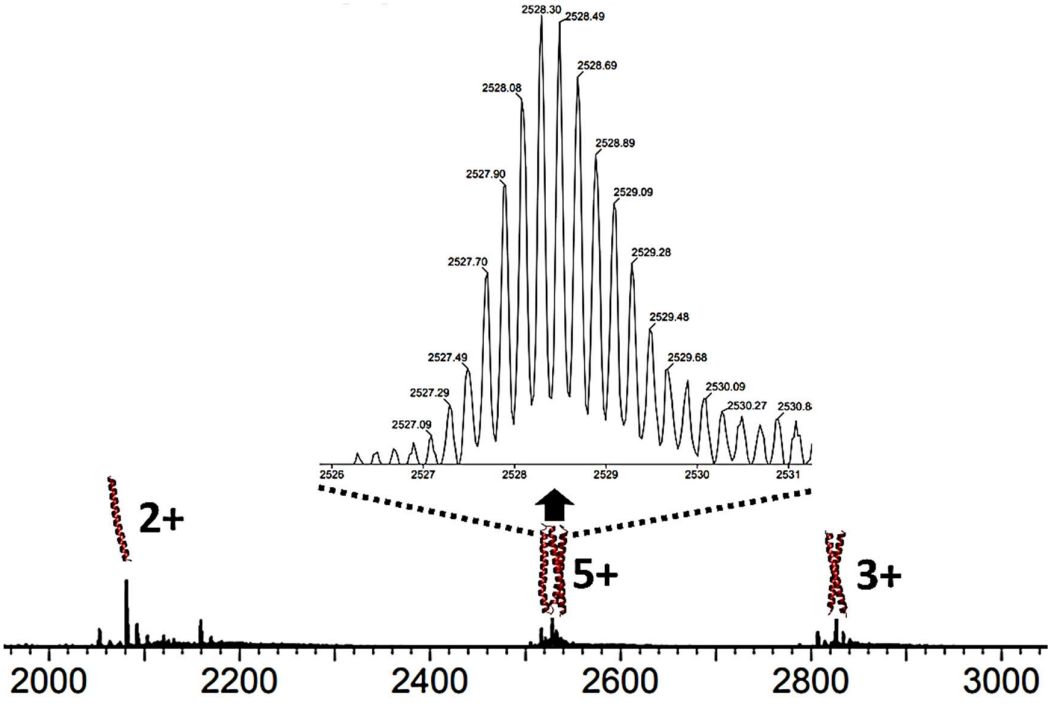

Figure 36: The mass spectrum of native HC03 with Tb(III) indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC03 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4) and Tb(III) ions at 1 equivalence to the trimer.

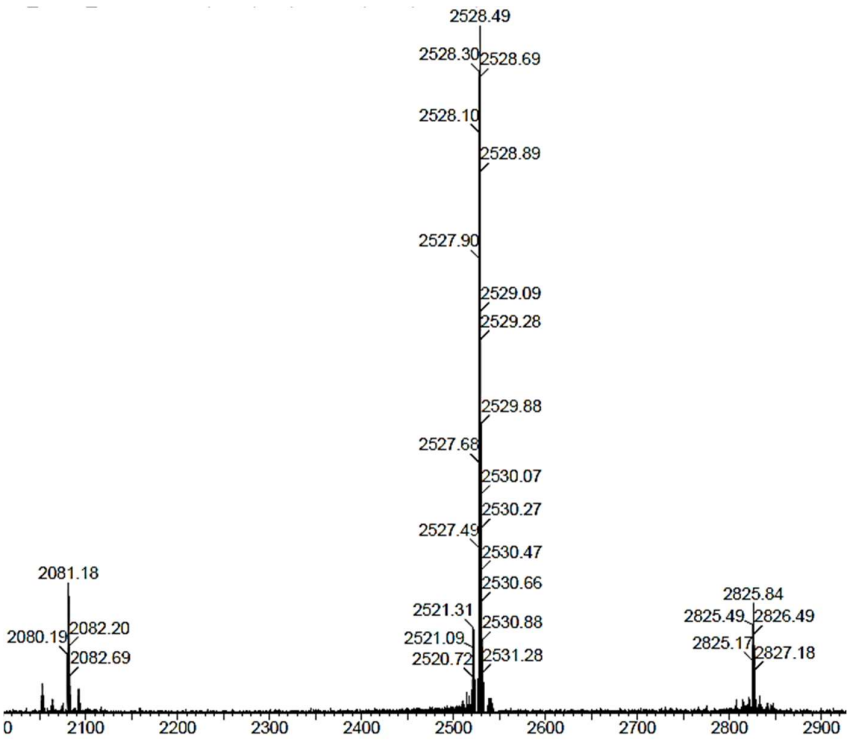

Figure 37: The tandem mass spectrum for the locked +5 charge peak from a trimer species of HC03 with Tb(III) ions illustrating the presence of the degradation products being a dimer with the metal ion and a monomer.

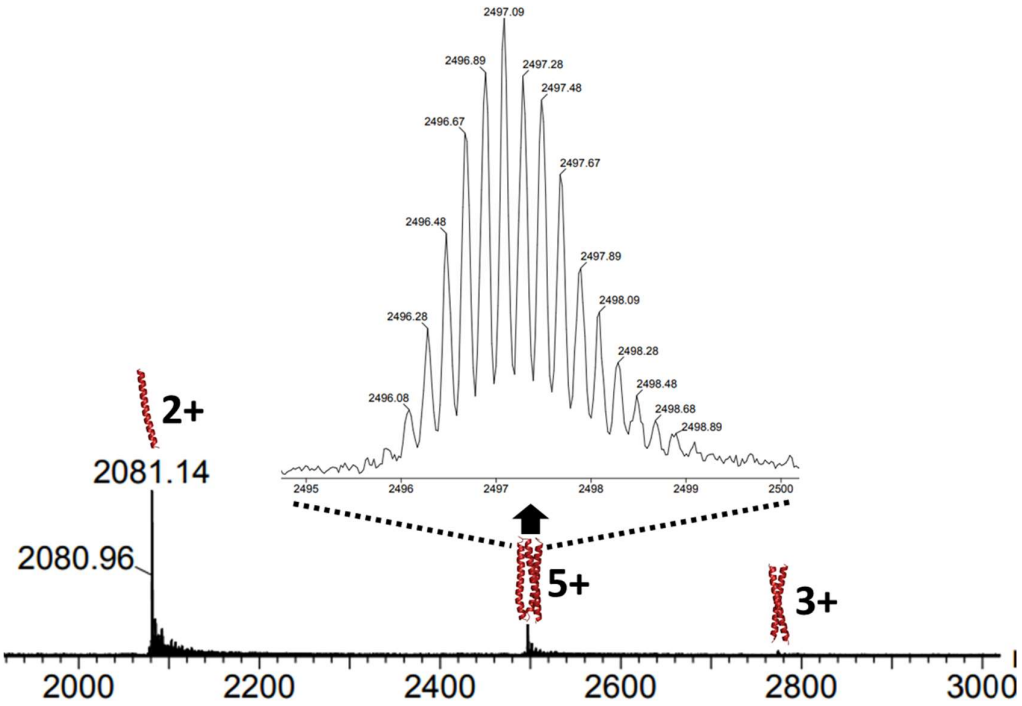

Figure 38: The mass spectrum of native HC04 with Tb(III) indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC04 (9 µM) in ammonium acetate buffer (10 mM, pH 6.4) and Tb(III) ions at 1 equivalence to the trimer.

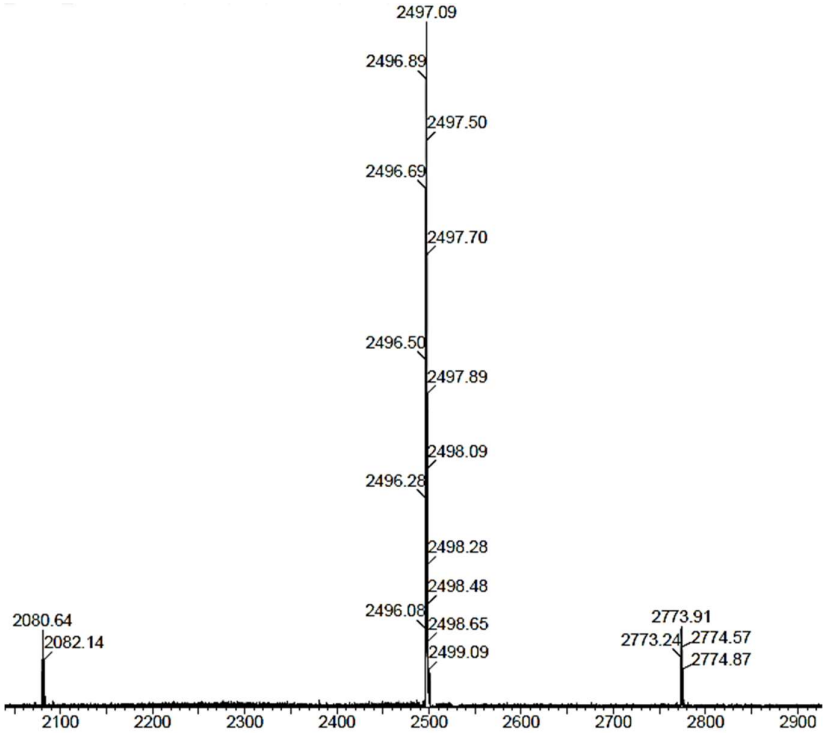

Figure 39: The tandem mass spectrum for the locked +5 charge peak from a trimer species of HC04 with Tb(III) ions illustrating the presence of the degradation products being a dimer and a monomer.

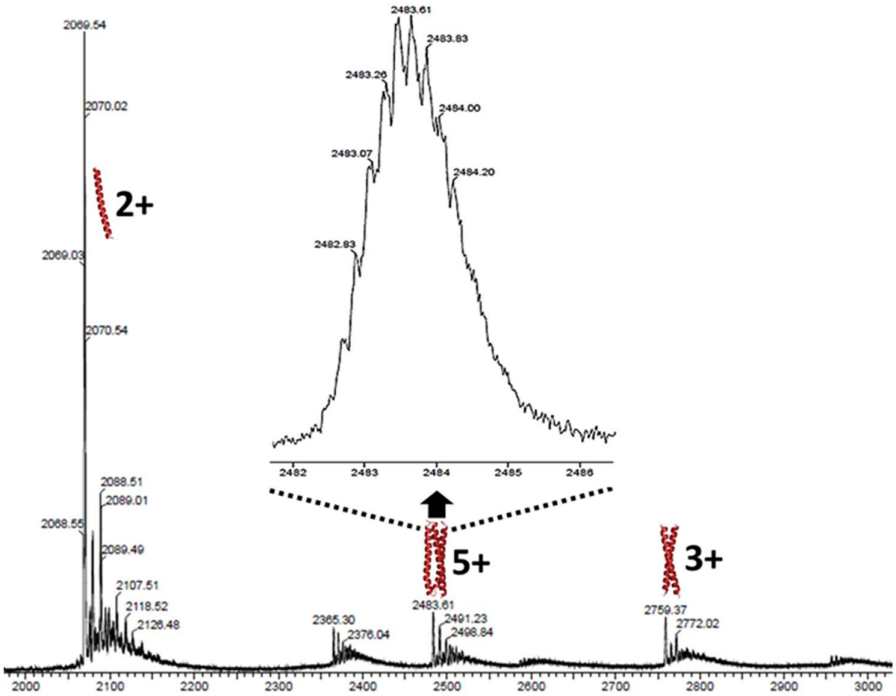

Figure 40: The mass spectrum of native HC01-2Y indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC01-2Y (9 µM) in ammonium acetate buffer (10 mM, pH 6.4). The lower resolution of this spectrum is due to damage to the instrument.

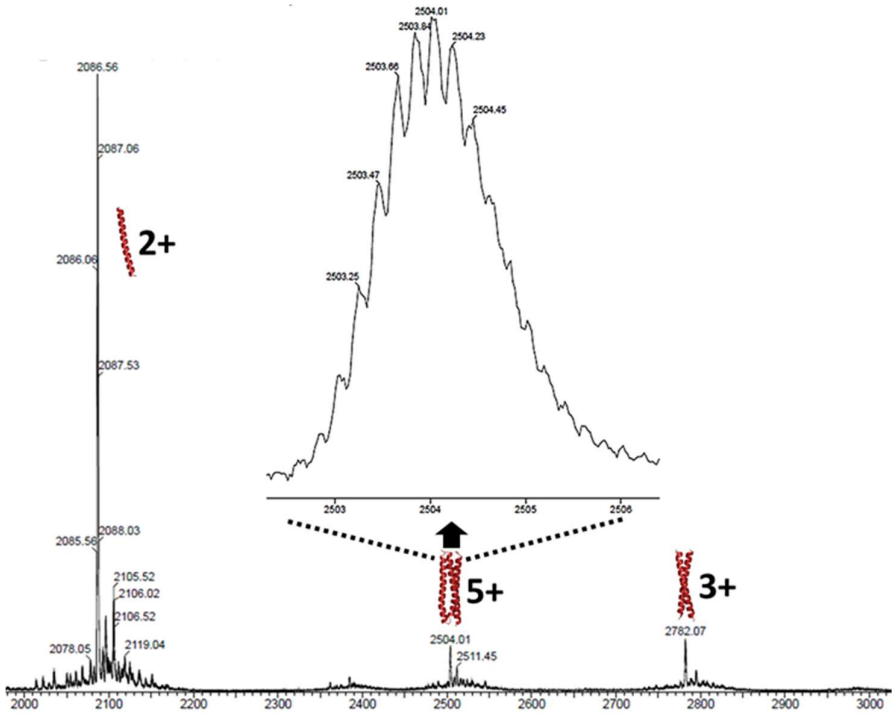

Figure 41: The mass spectrum of native HC01-2Nap indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC01-2Nap (9 µM) in ammonium acetate buffer (10 mM, pH 6.4). The lower resolution of this spectrum is due to damage to the instrument.

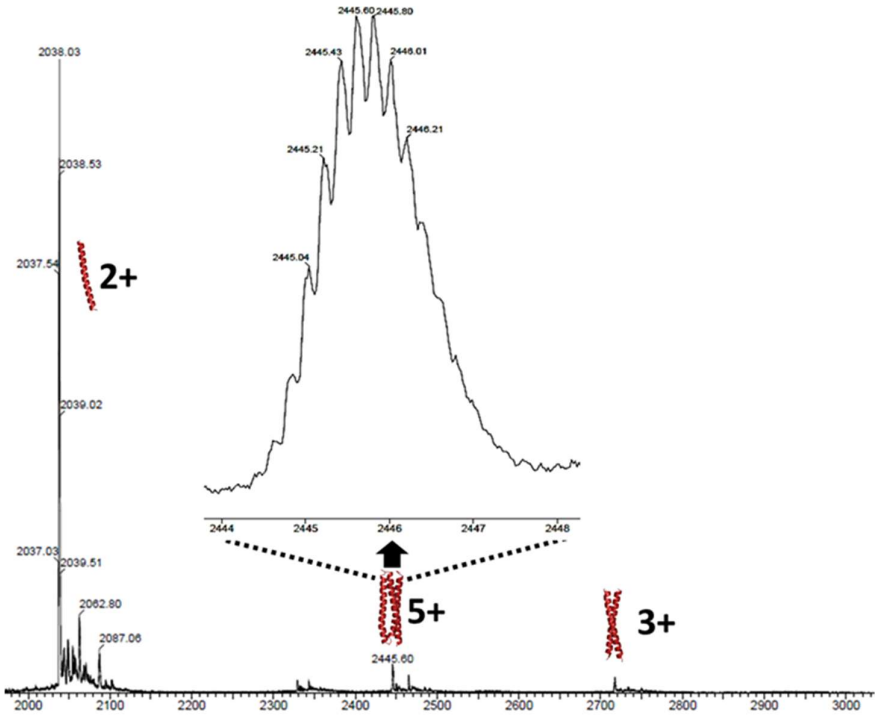

Figure 42: The mass spectrum of native HC01-2Cou indicating the different charged species resulting from different oligomeric states. The insert shows the distribution of the +5 charge peak cluster resulting from a trimeric species. The sample consisted of HC01-2Cou (9 µM) in ammonium acetate buffer (10 mM, pH 6.4). The lower resolution of this spectrum is due to damage to the instrument.

9.4 TERBIUM AND ZINC X-RAY ABSORPTION EDGE PROFILE

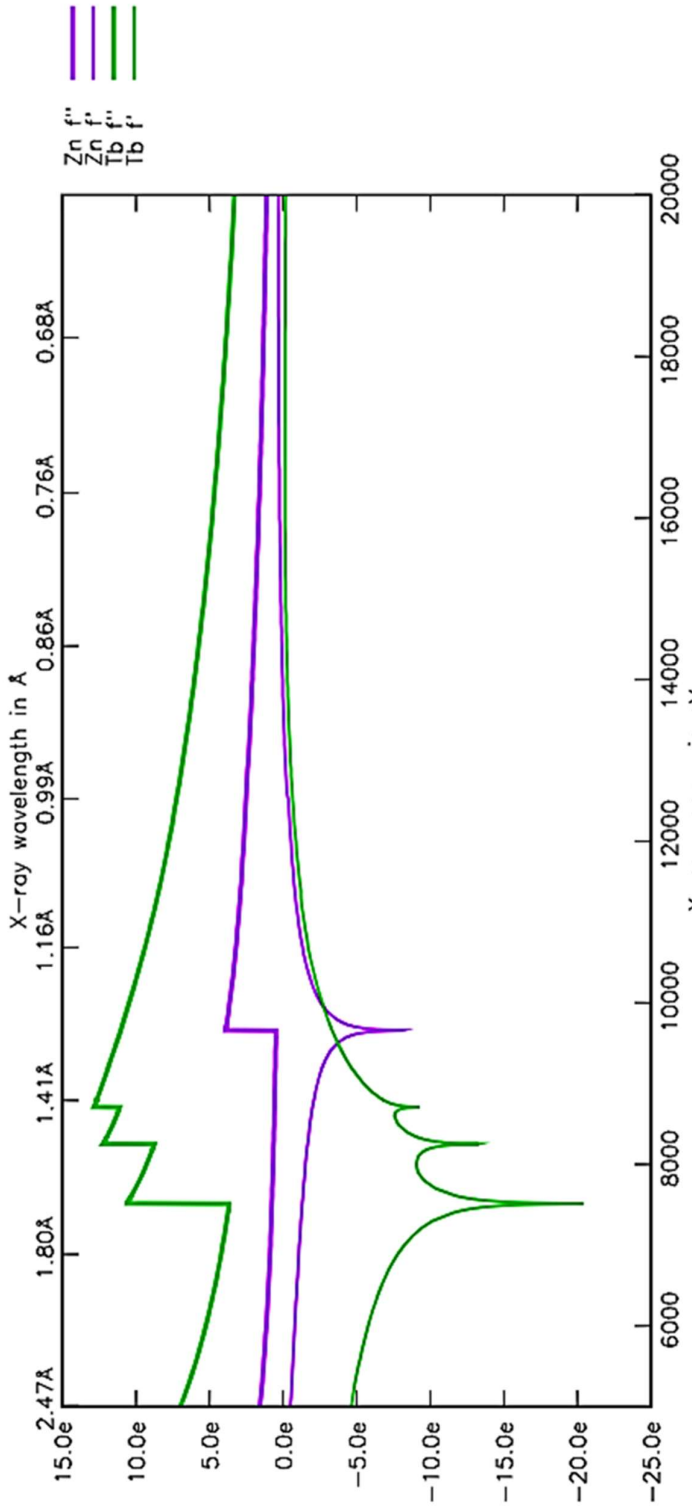

Figure 43: The predicted absorption edge profiles for both terbium and zinc generated from an open source website.[1]

9.5 RAMACHANDRAN PLOTS AND BINDING SITE DETAILS

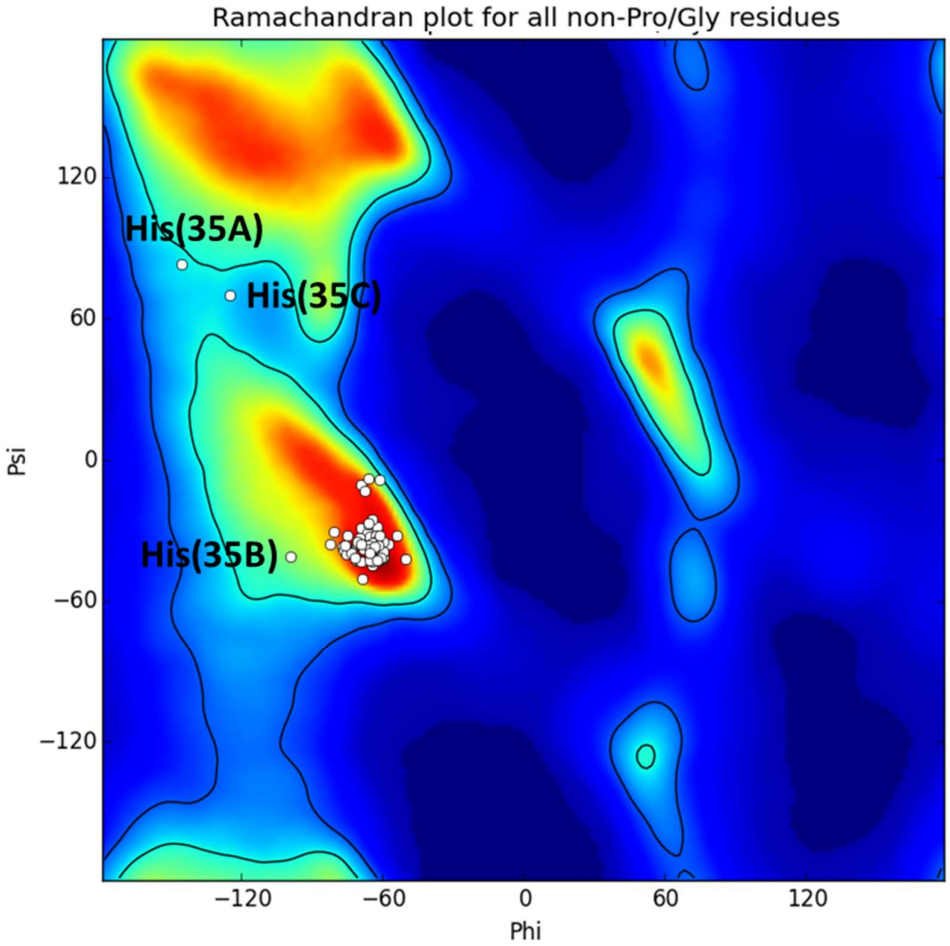

Figure 44: The Ramachandran plot for the crystallographically determined structure of HC0C.

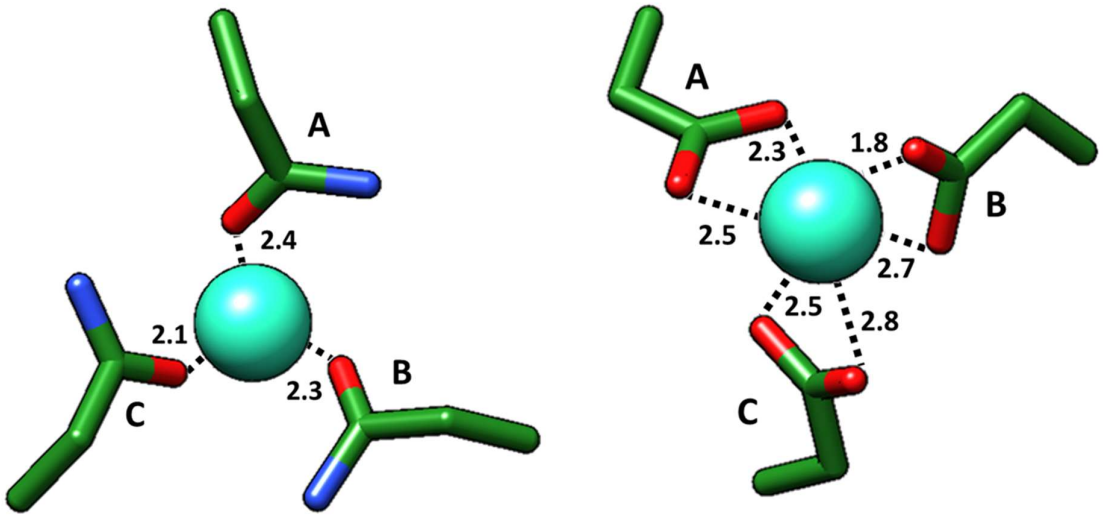

Figure 45: The crystallographically determined core binding site of HC02 with Tb(III) present, showing the interactions from both the asparagine and aspartate residues with their respective oxygen-terbium distances.

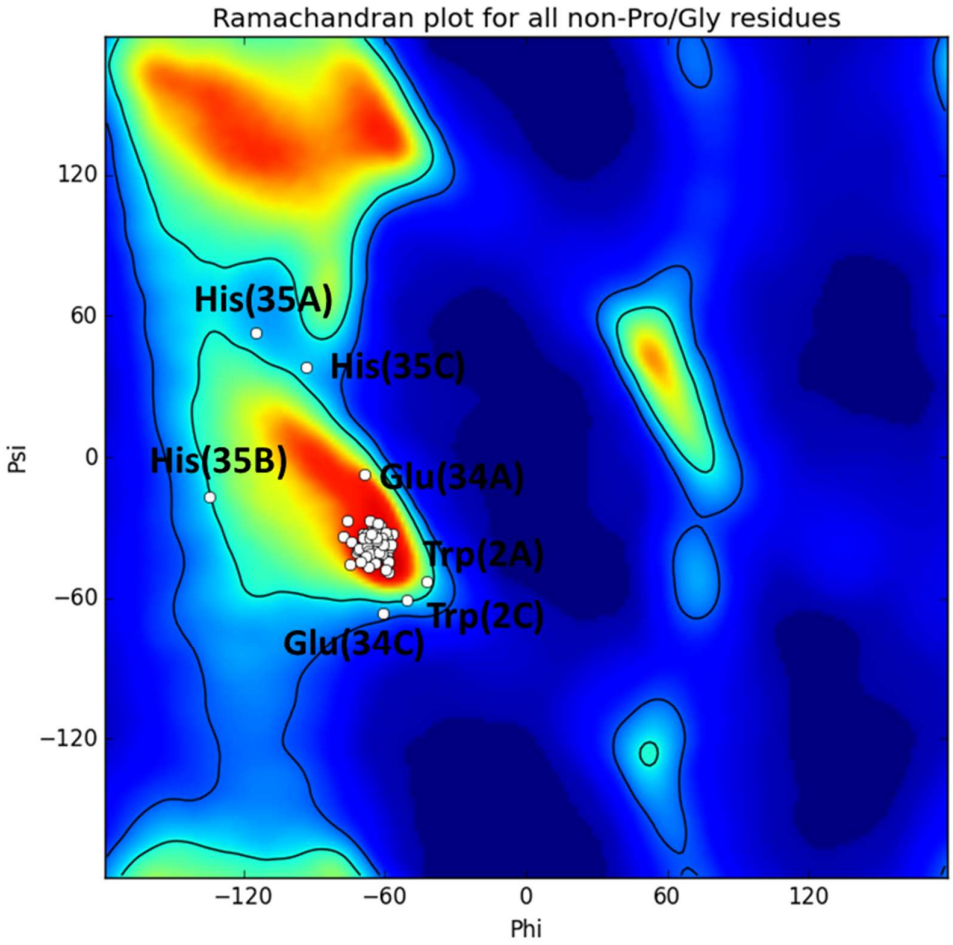

Figure 46: The Ramachandran plot for the crystallographically determined structure of HC02.

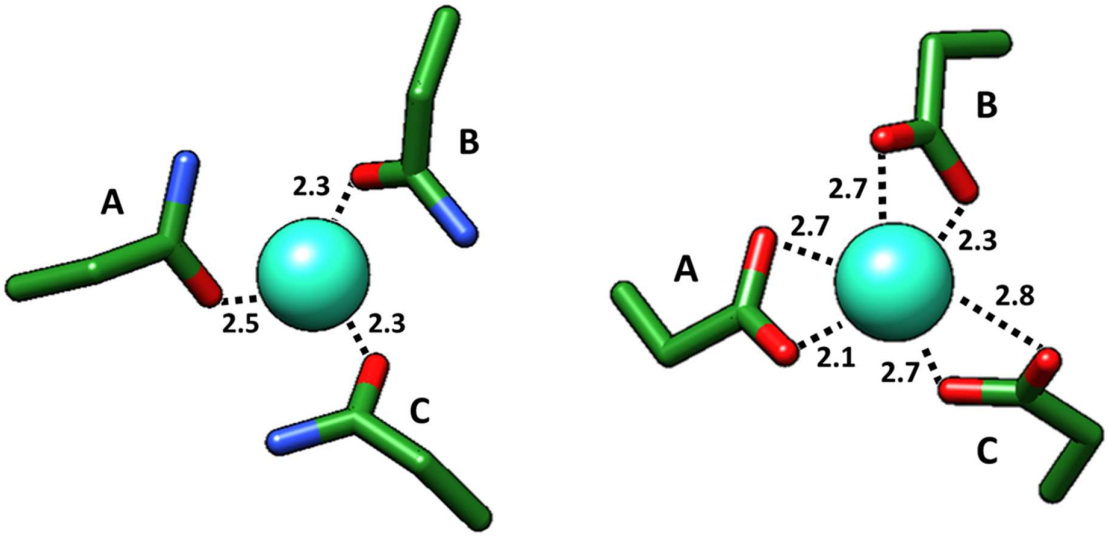

Figure 47: The crystallographically determined core binding site of HC03 with Tb(III) present, showing the interactions from both the asparagine and aspartate residues with their respective oxygen-terbium distances.

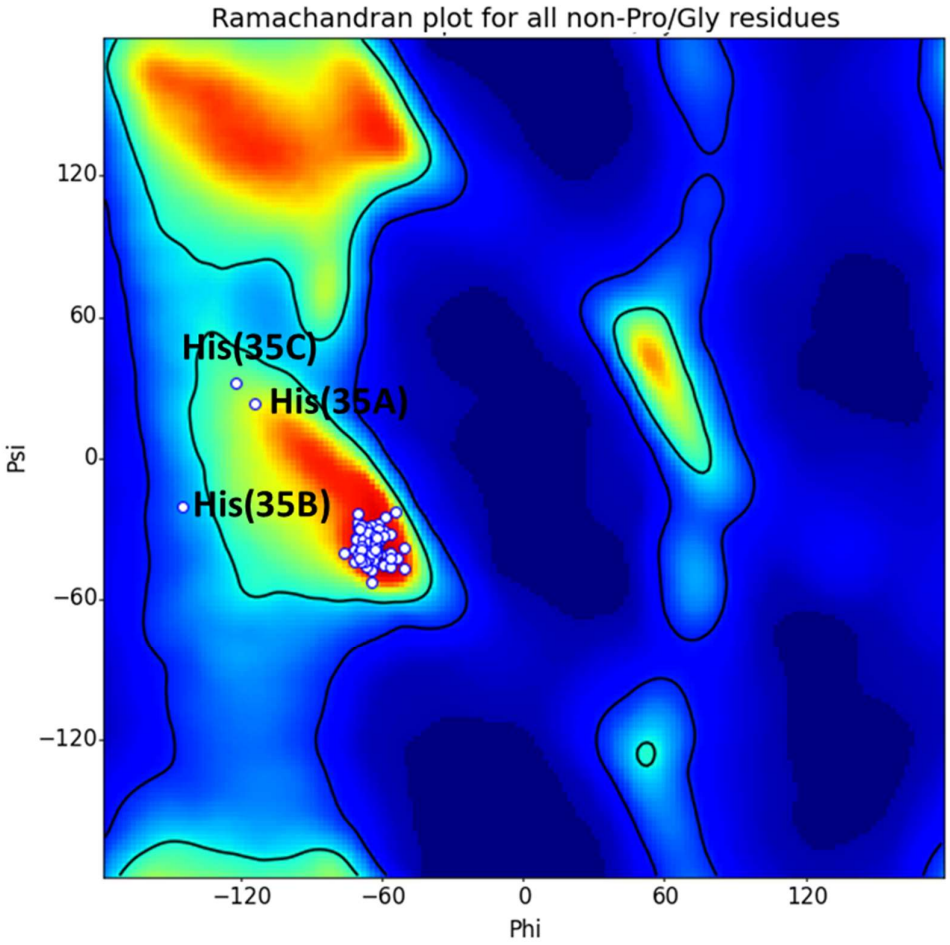

Figure 48: The Ramachandran plot for the crystallographically determined structure of HC03.

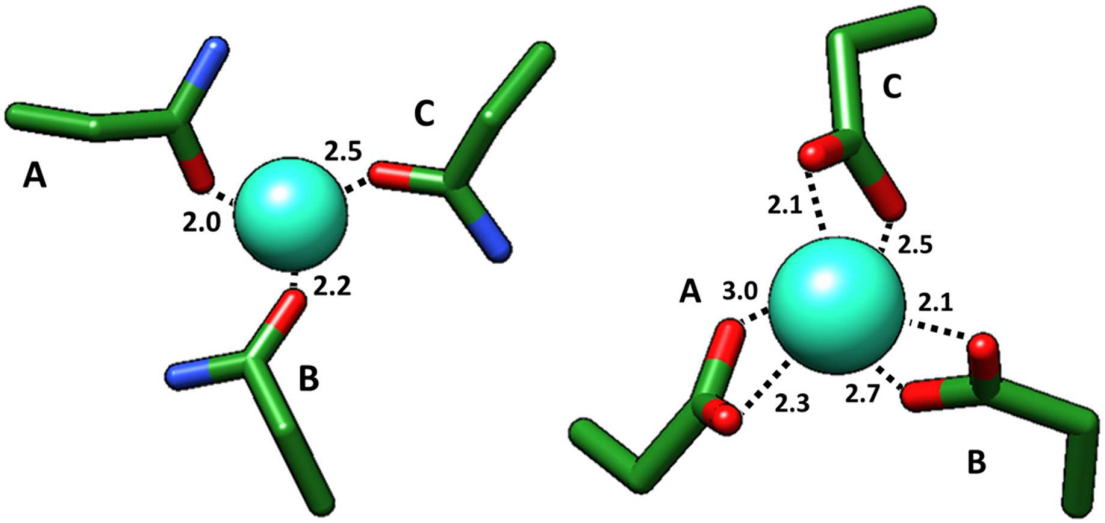

Figure 49: The crystallographically determined core binding site of HC04 with Tb(III) present, showing the interactions from both the asparagine and aspartate residues with their respective oxygen-terbium distances.

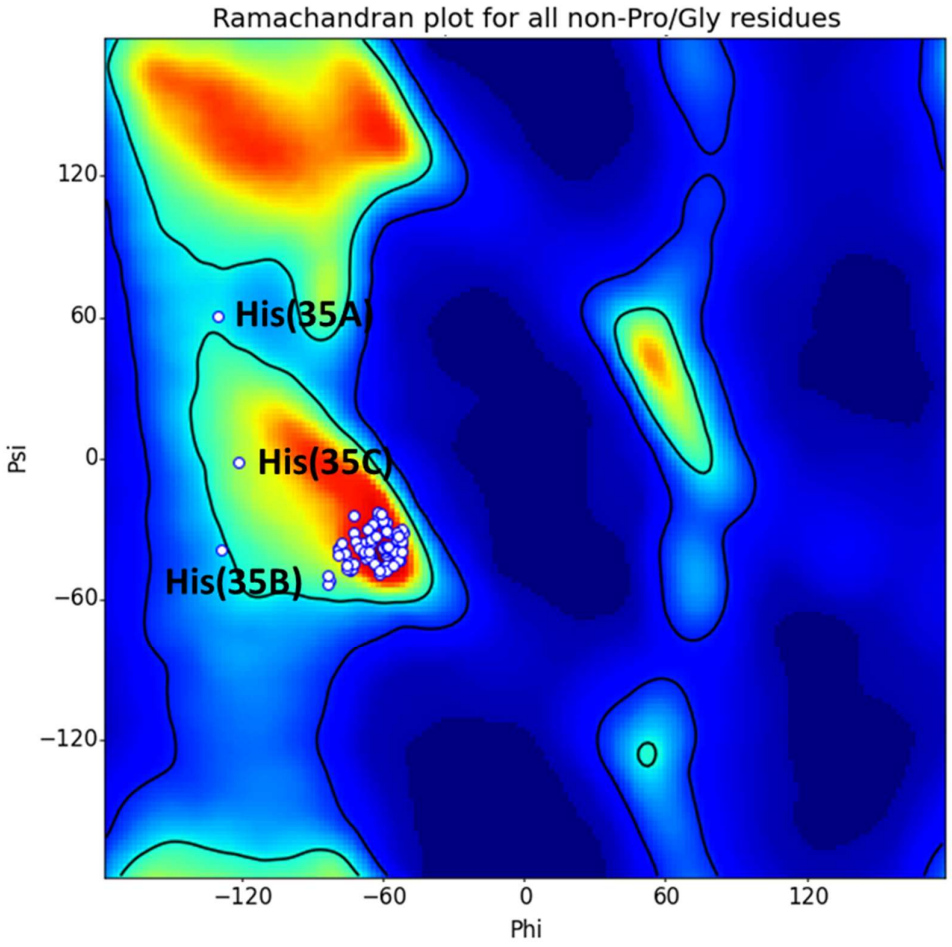

Figure 50: The Ramachandran plot for the crystallographically determined structure of HC04.

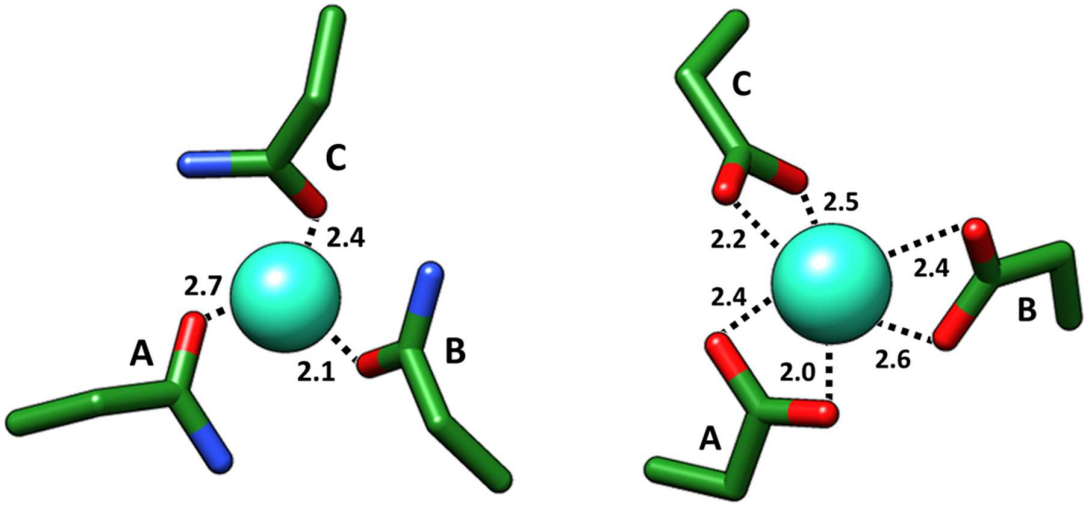

Figure 51: The crystallographically determined core binding site of HC01-2Y with Tb(III) present, showing the interactions from both the asparagine and aspartate residues with respective oxygen-terbium distances.

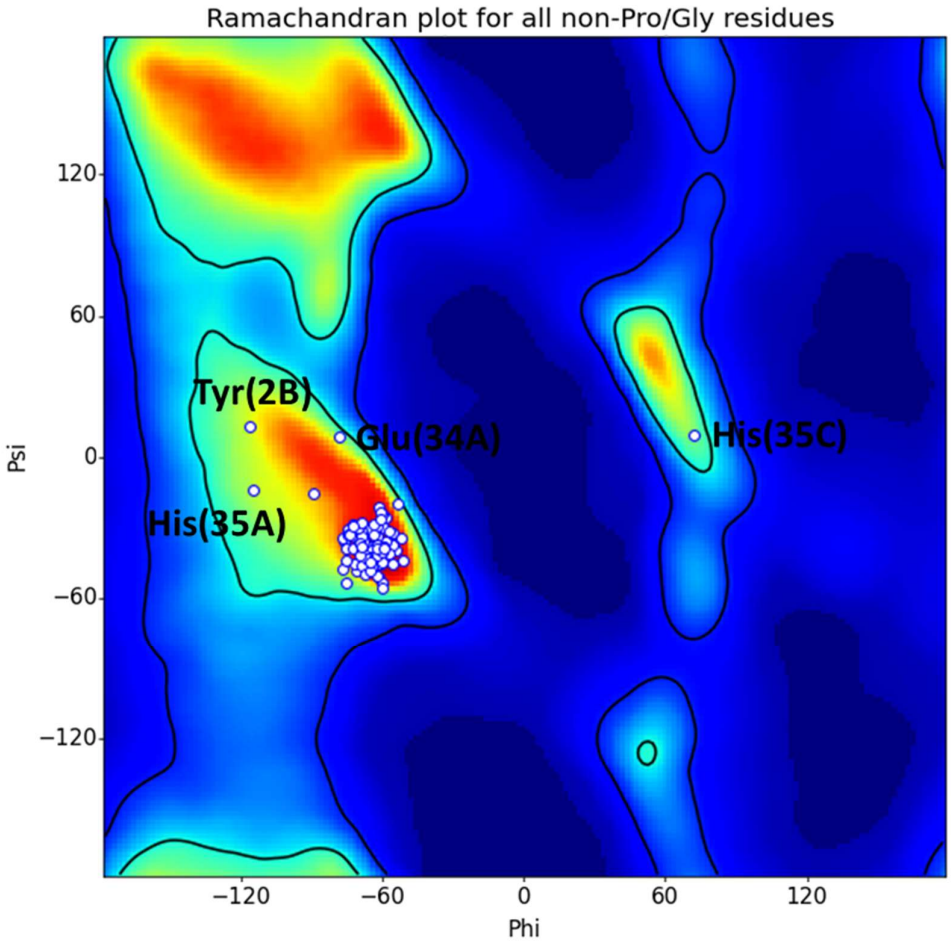

Figure 52: The Ramachandran plot for the crystallographically determined structure of HC01-2Y.

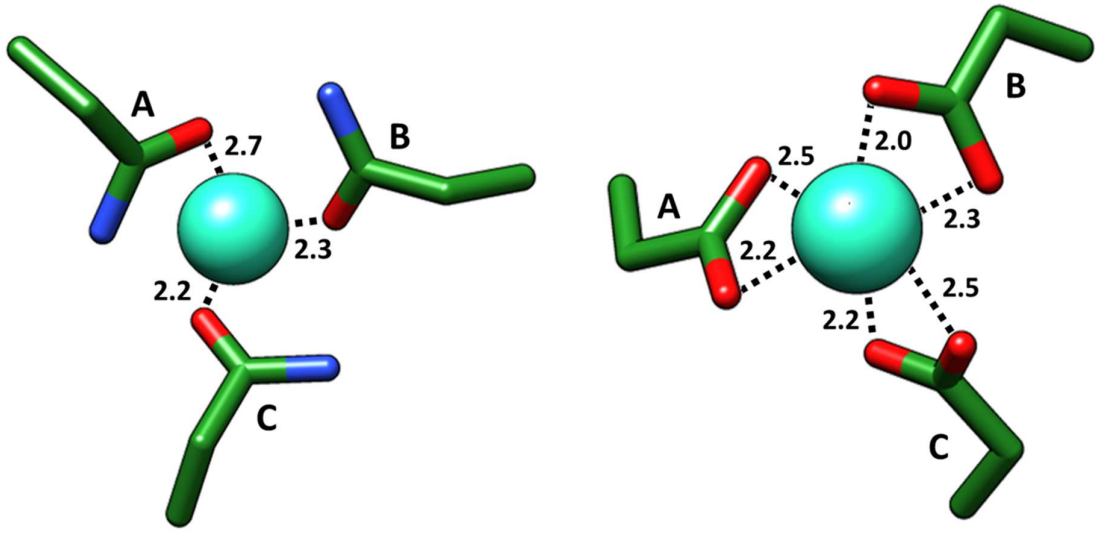

Figure 53: The crystallographically determined core binding site of HC01-2Nap with Tb(III) present, showing the interactions from both the asparagine and aspartate residues with respective oxygen-terbium distances.

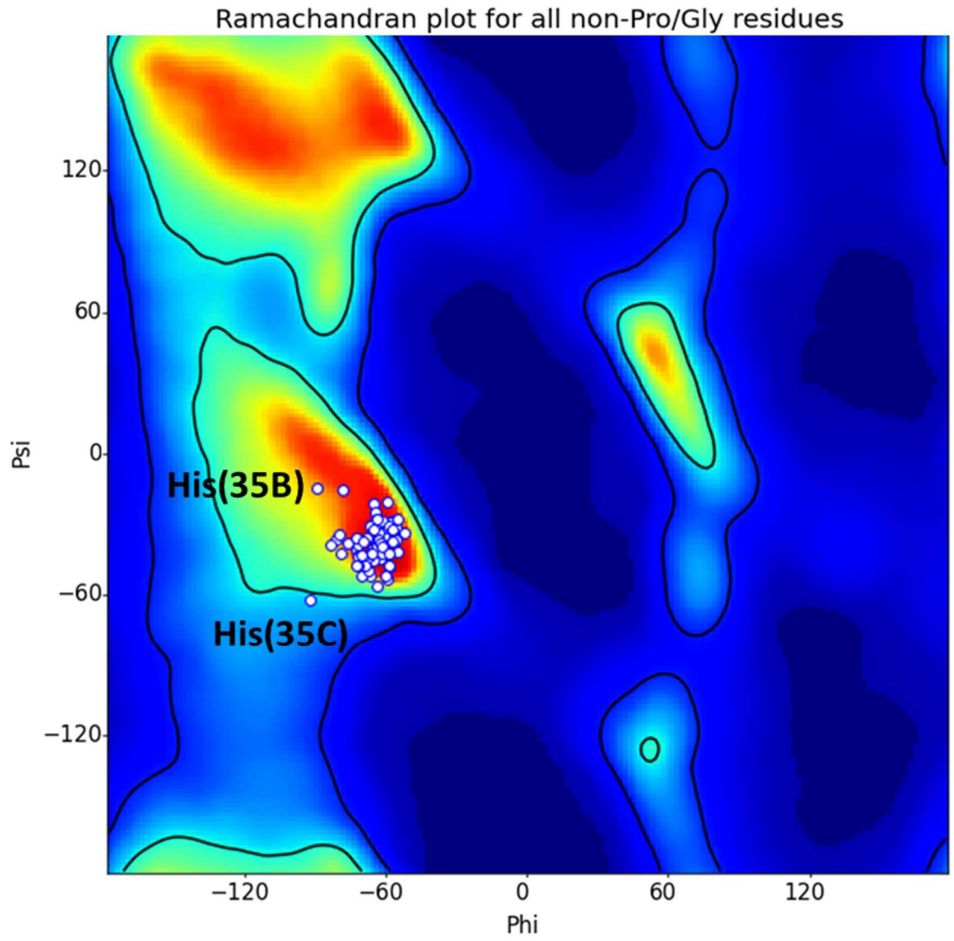

Figure 54: The Ramachandran plot for the crystallographically determined structure of HC01-2Nap.

9.6 LANTHANIDE ABSORPTION PROFILES

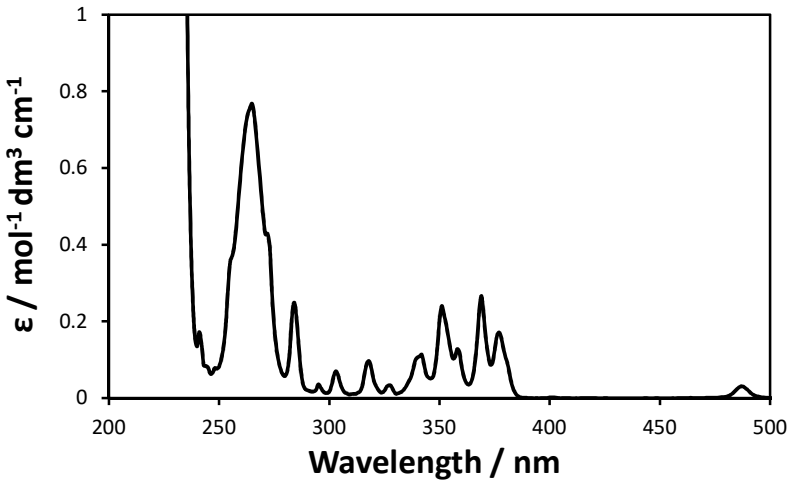

Figure 55: The absorption profile of Tb(III) ions in aqueous solution. Sample consisted of $TbCl_3$ (0.1 M) in water.

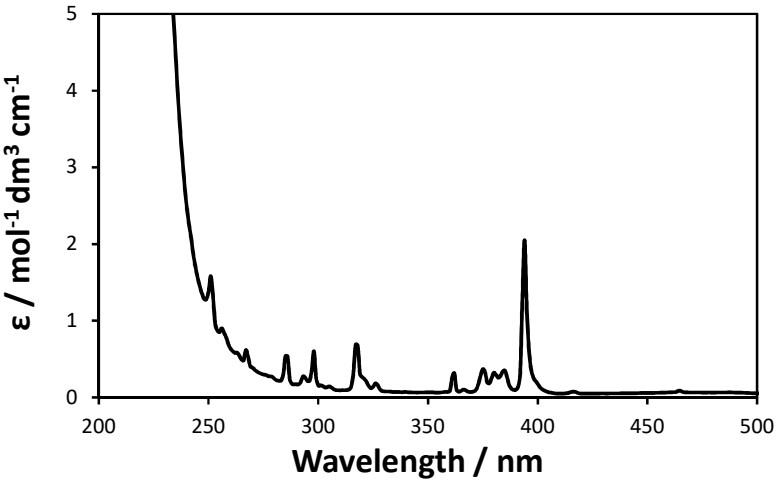

Figure 56: The absorption profile of Eu(III) ions in aqueous solution. Sample consisted of $EuCl_3$ (0.1 M) in water.

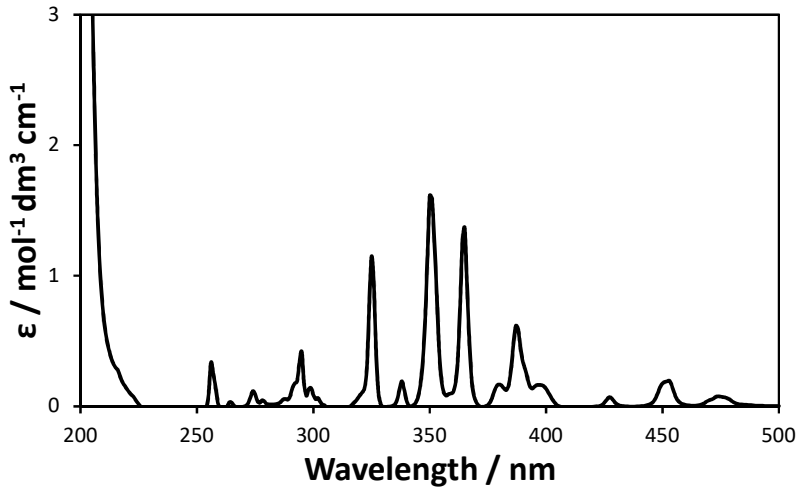

Figure 57: The absorption profile of Dy(III) ions in aqueous solution. Sample consisted of $DyCl_3$ (0.1 M) in water.

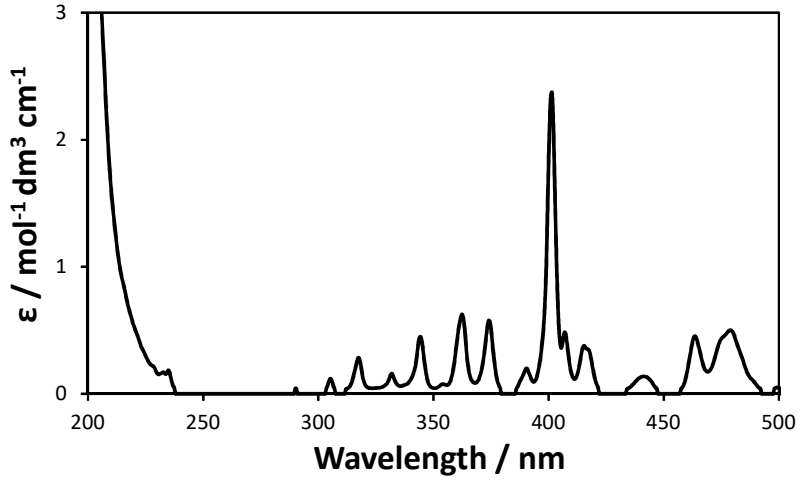

Figure 58: The absorption profile of Sm(III) ions in aqueous solution. Sample consisted of $SmCl_3$ (0.1 M) in water.

9.7 LBT DYSPROSIUM AND SAMARIUM EMISSION

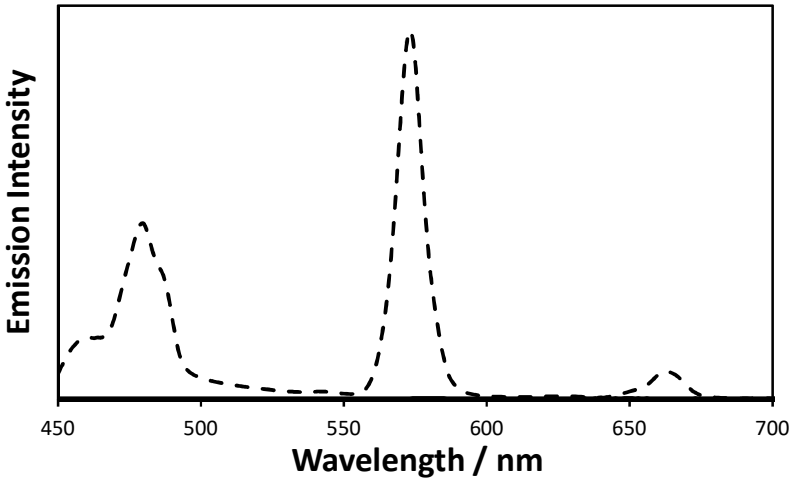

Figure 59: The emission spectrum of the LBT peptide with Dy(III) bound, focusing on the lanthanide emission region. The sample consisted of LBT (90 μM), aqueous HEPES (10 mM, pH 7.0) and Dy(III) at 1 equivalence.

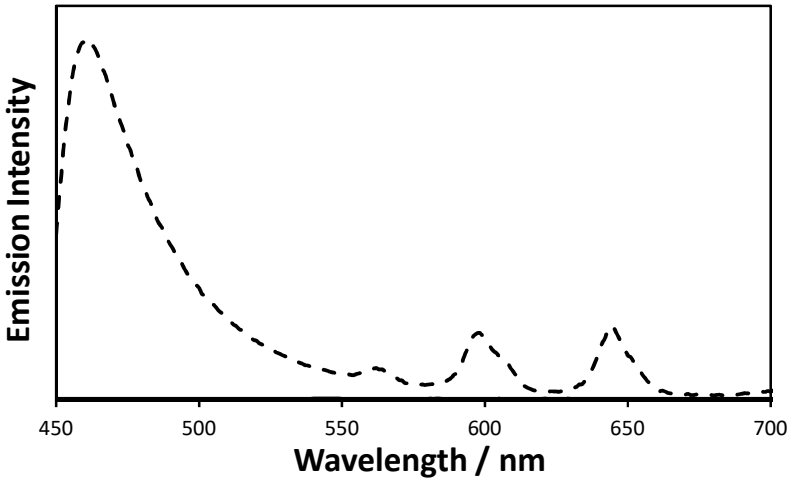

Figure 60: The emission spectrum of the LBT peptide with Sm(III) bound, focusing on the lanthanide emission region. The sample consisted of LBT (90 μM), aqueous HEPES (10 mM, pH 7.0) and Sm(III) at 1 equivalence.

9.8 MB1-2 DYSPROSIUM EMISSION

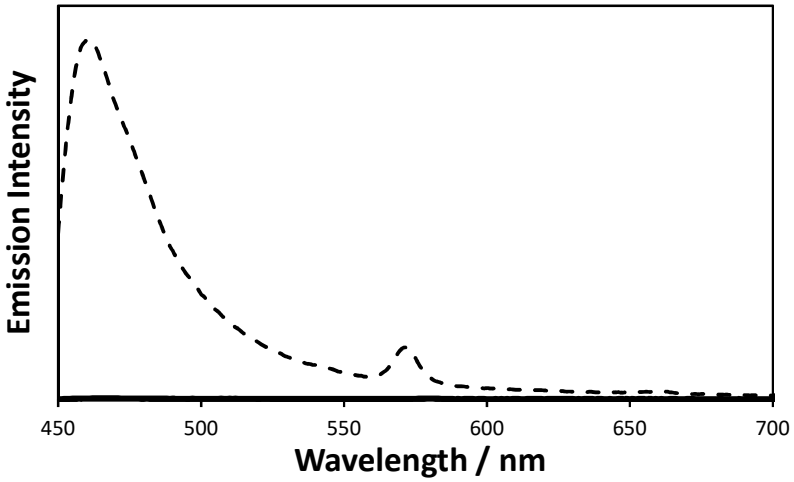

Figure 61: The emission spectrum of the MB1-2 peptide with Dy(III) bound, focusing on the lanthanide emission region. The sample consisted of MB1-2 (90 μM), aqueous HEPES (10 mM, pH 7.0) and Dy(III) at 1 equivalence per trimer unit.

9.9 HC01-HC04 EXCITATION SPECTRA AND LIFETIMES

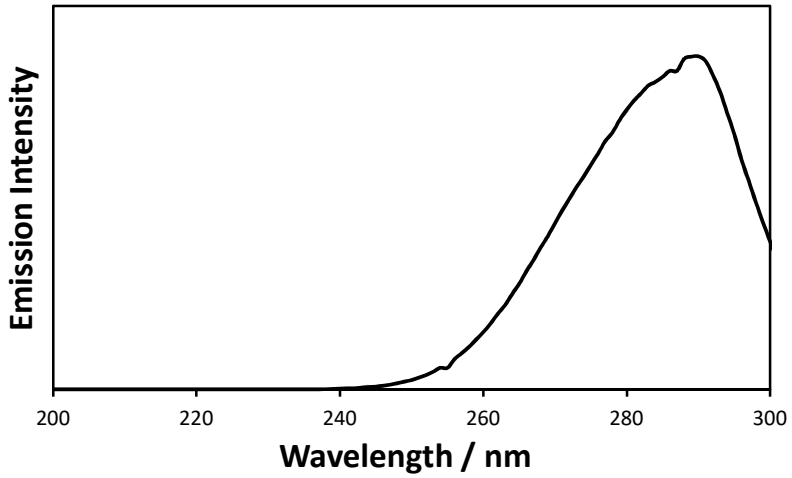

Figure 62: The excitation profile from monitoring the Tb(III) emission at 545 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01 (90 μM), aqueous HEPES (10 mM, pH 7.0) and Tb(III) at 1 equivalence to the trimer unit.

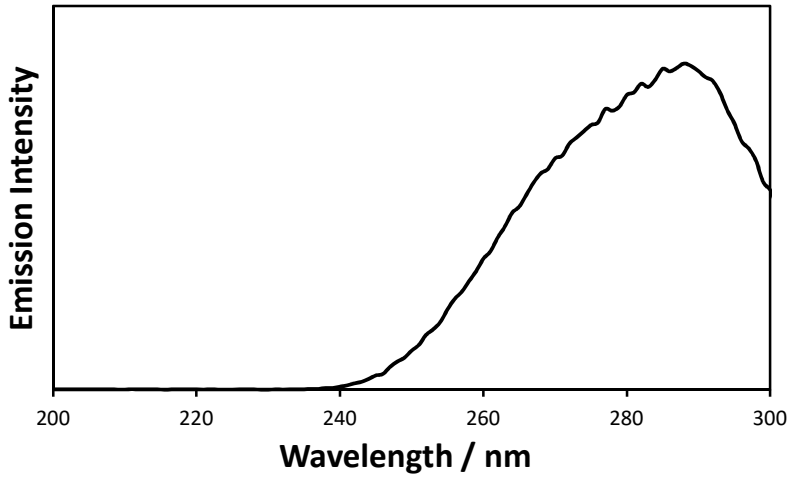

Figure 63: The excitation profile from monitoring the Eu(III) emission at 616 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01 (90 μM), aqueous HEPES (10 mM, pH 7.0) and Eu(III) at 1 equivalence to the trimer unit.

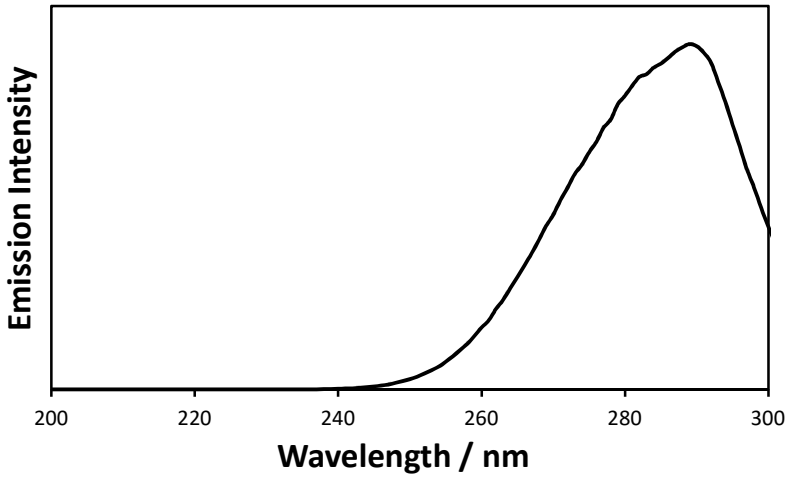

Figure 64: The excitation profile from monitoring the Dy(III) emission at 571 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01 (90 μM), aqueous HEPES (10 mM, pH 7.0) and Dy(III) at 1 equivalence to the trimer unit.

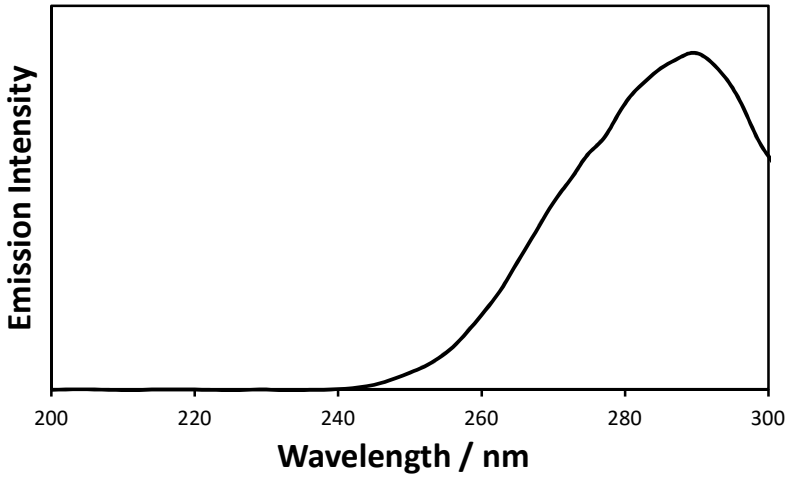

Figure 65: The excitation profile from monitoring the Sm(III) emission at 643 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01 (90 μM), aqueous HEPES (10 mM, pH 7.0) and Sm(III) at 1 equivalence to the trimer unit.

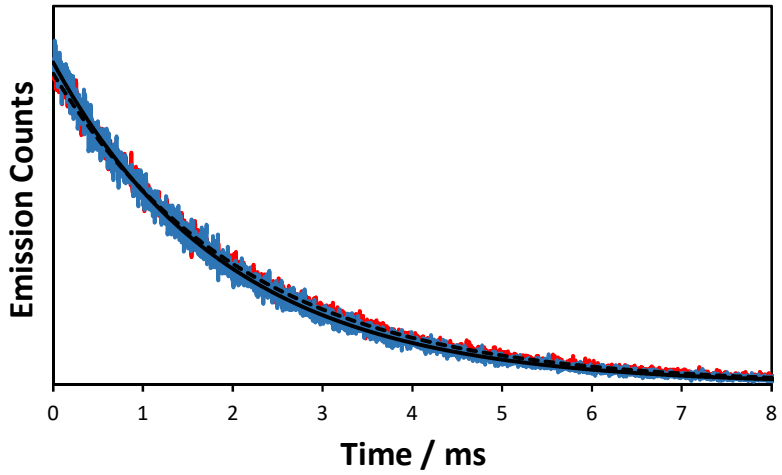

Figure 66: The lifetime profiles for HC01 with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC01 (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 1: The respective lifetimes associated with the emission from Tb(III) bound in the HC01 system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O *Run 1*	1956.8	1201.9	**1.190**
Run 2	1942.0	10606.6	**1.222**
Run 3	1942.4	10667.0	**1.228**
Average	**1947.1 ± 8.4**		
D2O *Run 1*	2117.3	1454.7	**1.152**
Run 2	2346.4	11292.8	**1.248**
Run 3	2343.3	10940.0	**1.223**
Average	**2269.0 ± 131.4**		

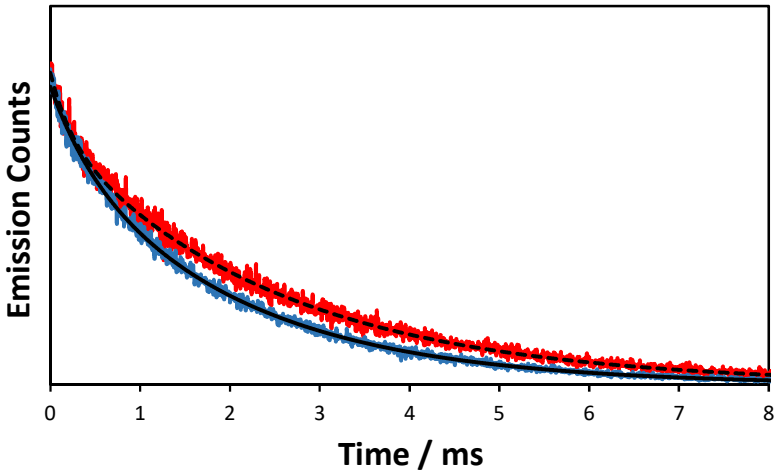

Figure 67: The lifetime profiles for HC02 with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC02 (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 2: The respective lifetimes associated with the emission from Tb(III) bound in the HC02 system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O *Run 1*	1971.8	1513.7	**1.010**
Run 2	2007.6	1483.4	**1.055**
Run 3	2014.7	1382.2	**1.052**
Average	**1998.3 ± 23.1**		
D2O *Run 1*	2476.7	847.6	**1.135**
Run 2	2470.9	778.5	**1.094**
Run 3	2418.0	371.5	**1.037**
Average	**2455.3 ± 32.5**		

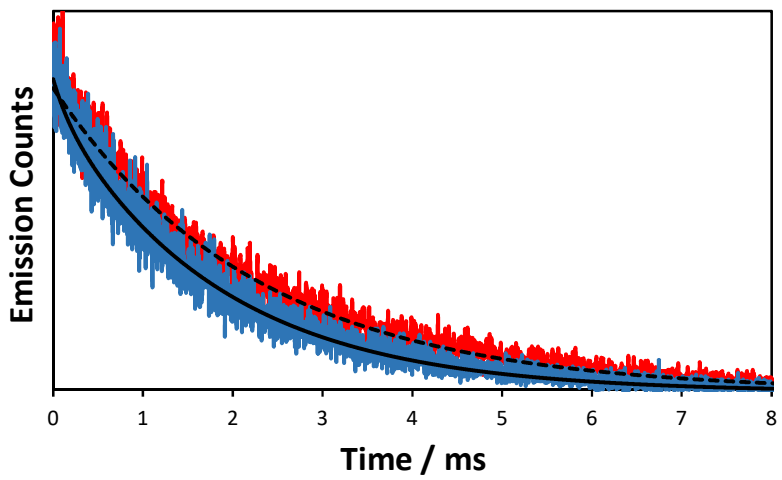

Figure 68: The lifetime profiles for HC03 with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC03 (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 3: The respective lifetimes associated with the emission from Tb(III) bound in the HC03 system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O *Run 1*	1897.6	196.2	**1.177**
Run 2	1821.8	213.1	**1.207**
Run 3	1816.6	307.3	**1.211**
Average	**1845.7 ± 45.4**		
D2O *Run 1*	2202.8	256.3	**1.102**
Run 2	2260.5	238.6	**1.042**
Run 3	2291.6	611.5	**1.432**
Average	**2251.3 ± 45.2**		

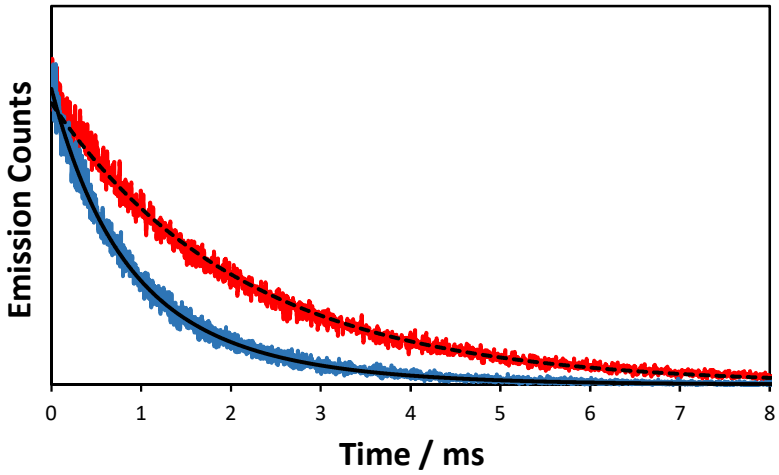

Figure 69: The lifetime profiles for HC04 with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC04 (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 4: The respective lifetimes associated with the emission from Tb(III) bound in the HC04 system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ_1 / μs	τ_2 / μs	B_1	B_2	χ^2
H2O *Run 1*	1425.4 (41.4)	625.9 (58.6)	979.5	3153.6	**1.171**
Run 2	1394.5 (44.9)	613.3 (55.1)	807.3	2256.5	**1.099**
Run 3	1350.5 (41.7)	610.2 (58.3)	778.6	2406.3	**1.116**
Average	**1390.3 ± 37.2**	**616.3 ± 8.5**			
D2O *Run 1*	2134.6	-	759.2	-	**1.209**
Run 2	2223.3	-	829.1	-	**1.258**
Run 3	2188.4	-	841.6	-	**1.174**
Average	**2182.1 ± 44.3**				

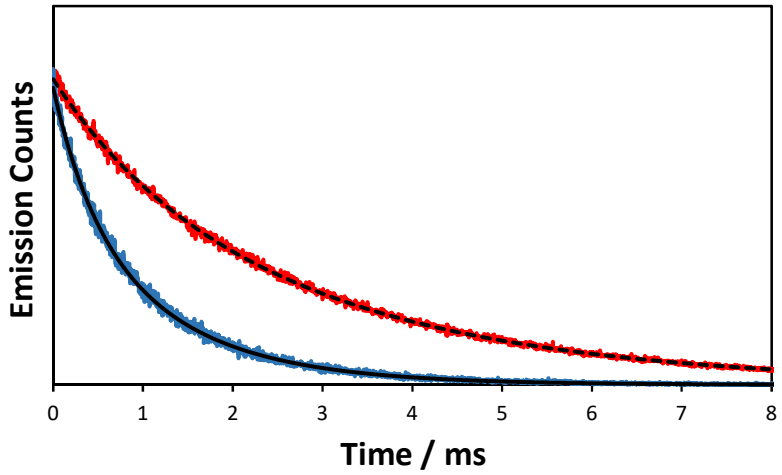

Figure 70: The lifetime profiles for HC0C with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC0C (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 5: The respective lifetimes associated with the emission from Tb(III) bound in the HC0C system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ_1 / μs	τ_2 / μs	B_1	B_2	χ^2
H2O *Run 1*	1204.4 (70.8)	444.0 (29.2)	1163.0	1300.6	**1.137**
Run 2	1243.6 (76.6)	439.6 (23.4)	1201.6	1038.4	**1.148**
Run 3	1251.1 (74.7)	429.2 (25.3)	917.0	904.3	**1.120**
Average	**1233.0 ± 25.1**	**437.6 ± 7.6**			
D2O *Run 1*	2834.4	-	4138.2	-	**1.152**
Run 2	2892.4	-	3836.7	-	**1.153**
Run 3	2885.8	-	3830.6	-	**1.171**
Average	**2870.9 ± 31.8**				

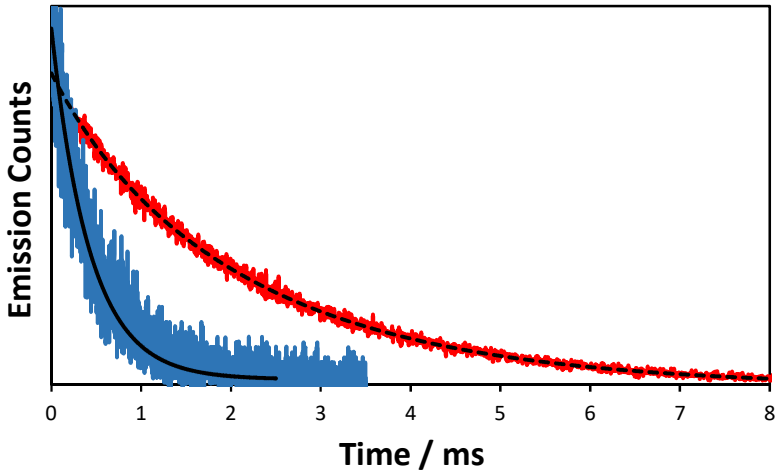

Figure 71: The lifetime profiles for Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of $TbCl_3$ (30 μM), HEPES buffer (10 mM, pH 7.0) in water or deuterium oxide.

Table 6: The respective lifetimes associated with the emission from Tb(III) in both water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O	430.1	176.2	**1.269**
D2O	2255.8	1275.5	**1.313**

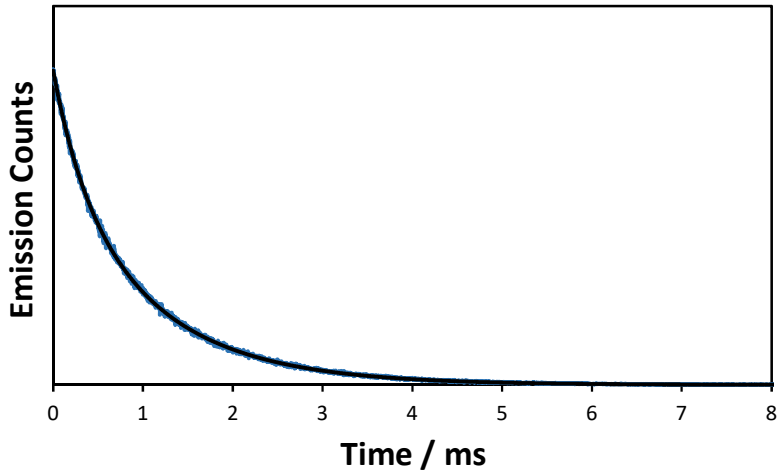

Figure 72: A lifetime profile for HC01 with Eu(III) monitoring the emission at 616 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01 (90 μM), HEPES buffer (10 mM, pH 7.0) and Eu(III) in a third equivalence to trimer unit.

Table 7: The respective lifetimes associated with the emission from Eu(III) bound in the HC01 system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O ***Run 1***	1101.0	4239.5	**1.128**
Run 2	1102.2	4235.3	**1.132**
Run 3	1095.6	4293.4	**1.165**
Average	**1099.6 ± 3.5**		

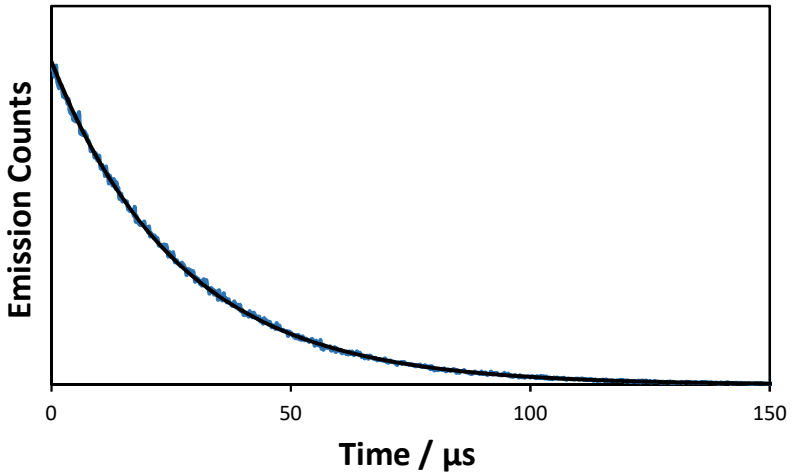

Figure 73: A lifetime profile for HC01 with Dy(III) monitoring the emission at 571 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01 (90 µM), HEPES buffer (10 mM, pH 7.0) and Dy(III) in a third equivalence to trimer unit.

Table 8: The respective lifetimes associated with the emission from Dy(III) bound in the HC01 system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / µs	B	χ^2
H2O *Run 1*	27.05	4755.6	**1.136**
Run 2	27.09	4761.9	**1.033**
Run 3	27.05	4787.1	**1.067**
Average	**27.06 ± 0.02**		

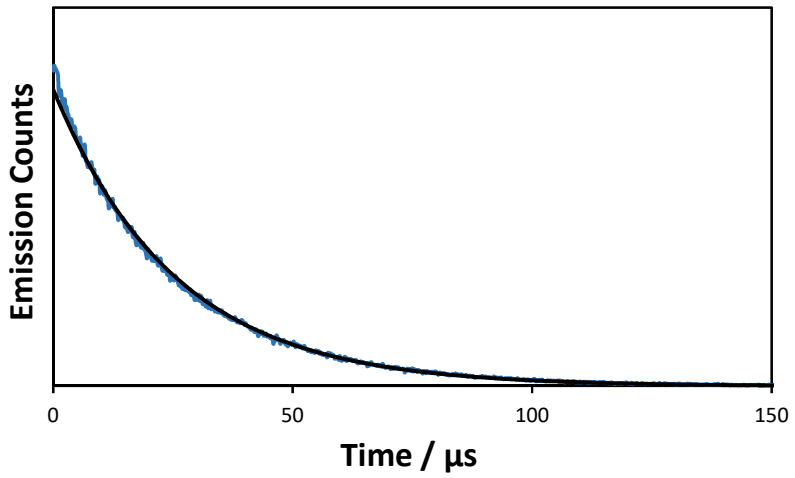

Figure 74: A lifetime profile for HC01 with Sm(III) monitoring the emission at 643 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01 (90 µM), HEPES buffer (10 mM, pH 7.0) and Sm(III) in a third equivalence to trimer unit.

Table 9: The respective lifetimes associated with the emission from Sm(III) bound in the HC01 system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / µs	B	χ^2
H2O ***Run 1***	25.48	4587.6	**0.930**
Run 2	25.46	5170.6	**0.985**
Run 3	25.44	4745.5	**1.097**
Average	**25.46 ± 0.02**		

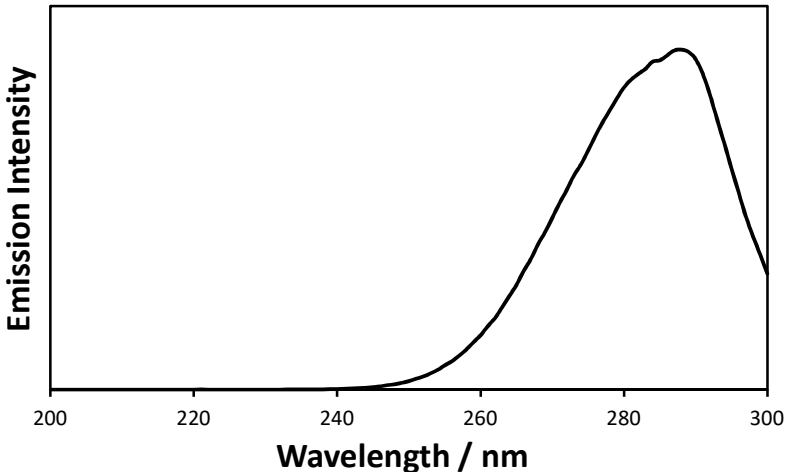

Figure 75: The excitation profile from monitoring the HC01 tryptophan emission at 350 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01 (90 µM) and aqueous HEPES (10 mM, pH 7.0).

9.10 HC01-2Y EXCITATION SPECTRA AND LIFETIMES

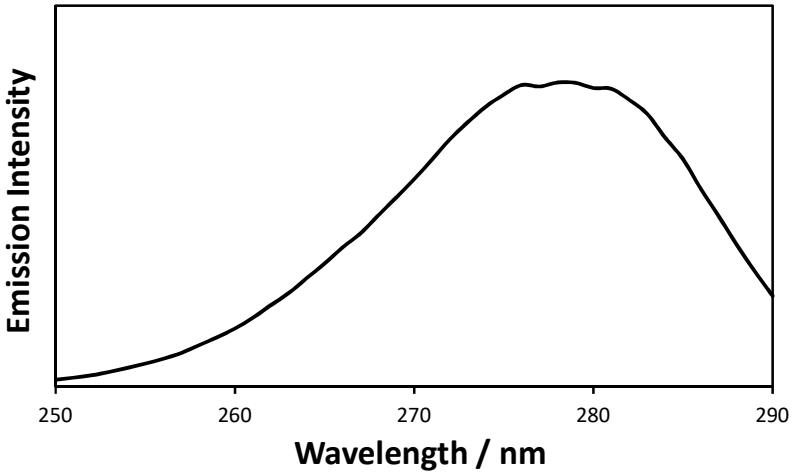

Figure 76: The excitation profile from monitoring the Tb(III) emission at 545 nm while scanning the excitation wavelength from 250-290 nm. The sample consisted of HC01-2Y (90 μM), aqueous HEPES (10 mM, pH 7.0) and Tb(III) at 1 equivalence to the trimer unit.

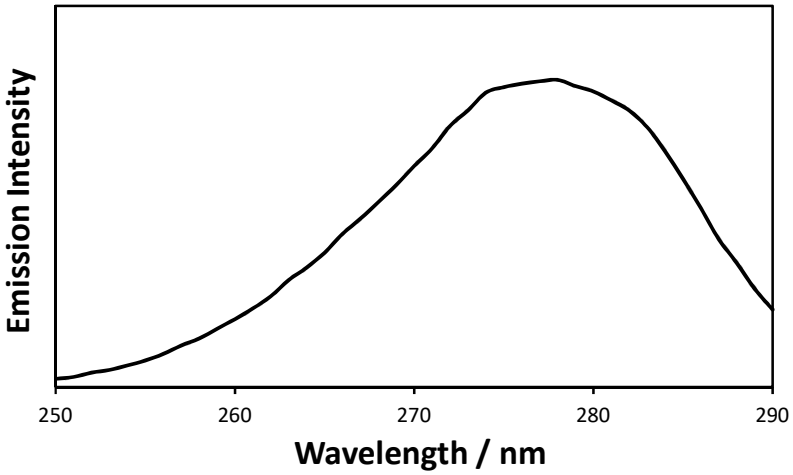

Figure 77: The excitation profile from monitoring the Eu(III) emission at 616 nm while scanning the excitation wavelength from 250-290 nm. The sample consisted of HC01-2Y (90 μM), aqueous HEPES (10 mM, pH 7.0) and Eu(III) at 1 equivalence to the trimer unit.

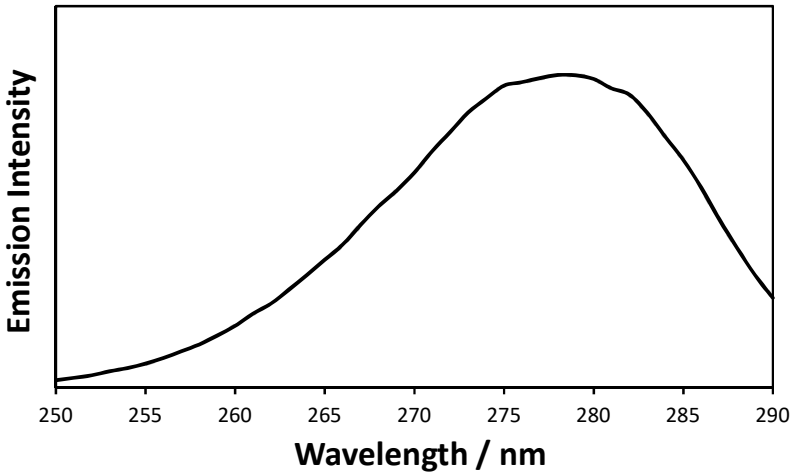

Figure 78: The excitation profile from monitoring the Dy(III) emission at 571 nm while scanning the excitation wavelength from 250-290 nm. The sample consisted of HC01-2Y (90 µM), aqueous HEPES (10 mM, pH 7.0) and Dy(III) at 1 equivalence to the trimer unit.

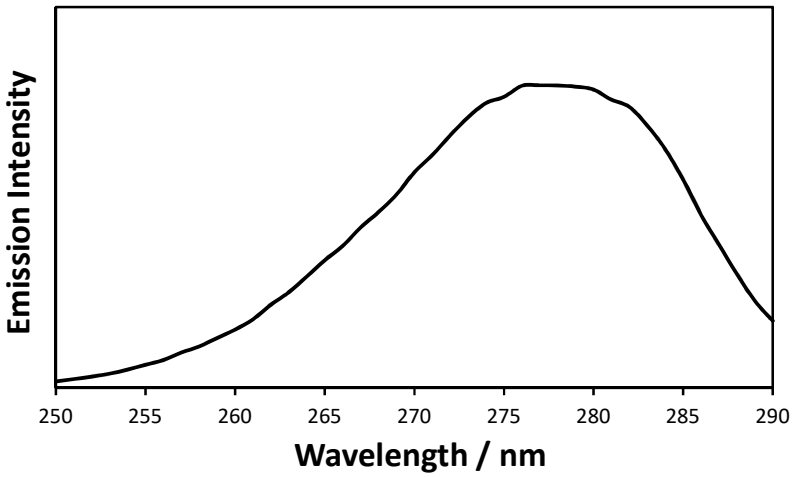

Figure 79: The excitation profile from monitoring the tyrosine emission of native HC01-2Y at 303 nm while scanning the excitation wavelength from 250-290 nm. The sample consisted of HC01-2Y (90 µM) and aqueous HEPES (10 mM, pH 7.0).

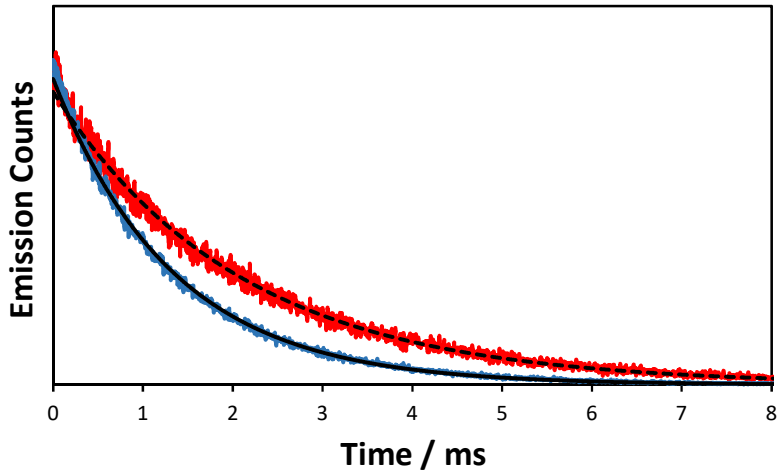

Figure 80: The lifetime profiles for HC01-2Y with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential tail fits as a solid black line and a dashed black line respectively. Samples consisted of HC01-2Y (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 10: The respective lifetimes associated with the emission from Tb(III) bound in the HC01-2Y system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O *Run 1*	1335.2	2696.4	**1.014**
Run 2	1338.4	2782.2	**0.953**
Run 3	1307.7	2945.2	**1.018**
Average	**1327 ± 17**		
D2O *Run 1*	2070.9	1098.4	**1.165**
Run 2	2082.2	972.9	**1.126**
Run 3	2207.3	823.0	**1.234**
Average	**2120 ± 76**		

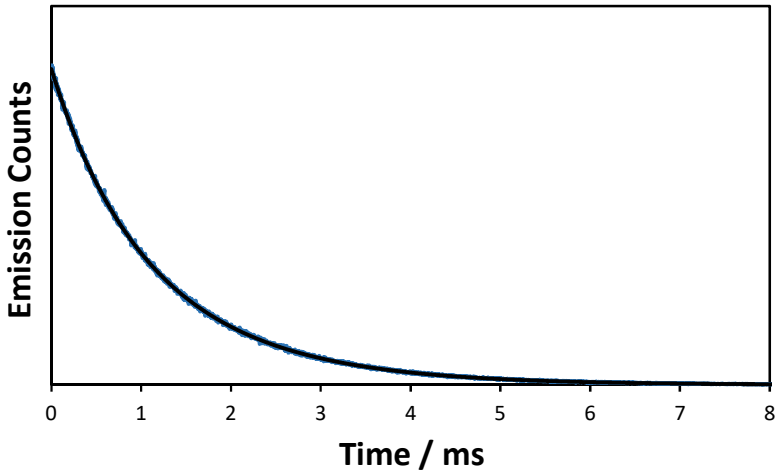

Figure 81: A lifetime profile for HC01-2Y with Eu(III) monitoring the emission at 616 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01-2Y (90 μM), HEPES buffer (10 mM, pH 7.0) and Eu(III) in a third equivalence to trimer unit.

Table 11: The respective lifetimes associated with the emission from Eu(III) bound in the HC01-2Y system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O ***Run 1***	728.5	3212.4	**1.001**
Run 2	744.7	1005.9	**0.896**
Run 3	743.2	1432.1	**0.950**
Average	**738.8 ± 9.0**		

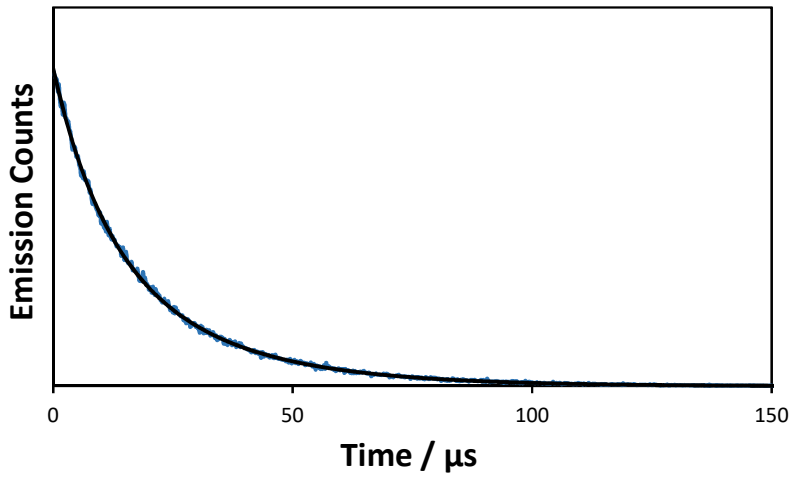

Figure 82: A lifetime profile for HC01-2Y with Dy(III) monitoring the emission at 571 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01-2Y (90 µM), HEPES buffer (10 mM, pH 7.0) and Dy(III) in a third equivalence to trimer unit.

Table 12: The respective lifetimes associated with the emission from Dy(III) bound in the HC01-2Y system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / µs	B	χ^2
H2O ***Run 1***	25.22	2172.7	**0.735**
Run 2	25.14	2258.6	**0.697**
Run 3	25.66	2915.6	**0.753**
Average	**25.34 ± 0.28**		

9.11 HC01-2NAP EXCITATION SPECTRA AND LIFETIMES

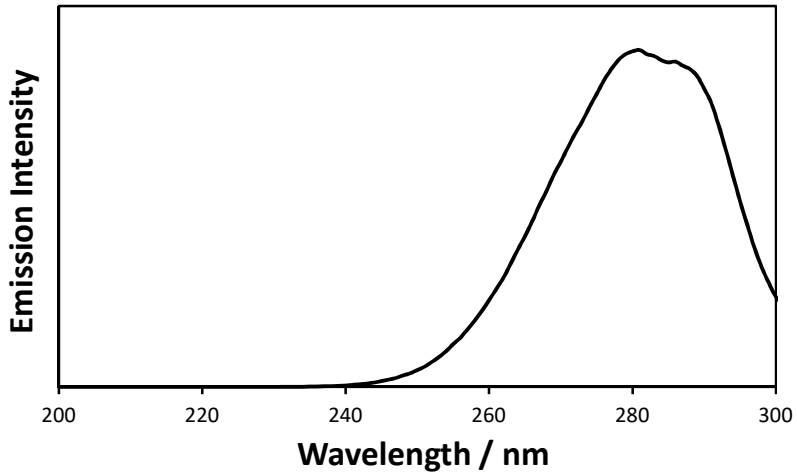

Figure 83: The excitation profile from monitoring the Tb(III) emission at 545 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01-2Nap (90 μM), aqueous HEPES (10 mM, pH 7.0) and Tb(III) at 1 equivalence to the trimer unit.

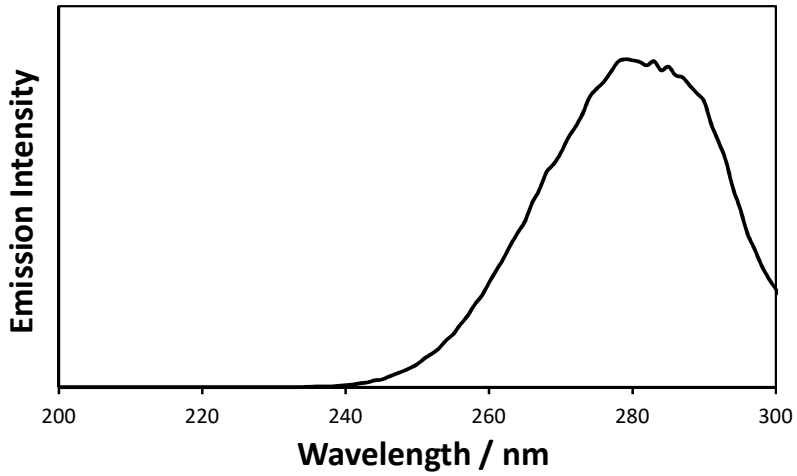

Figure 84: The excitation profile from monitoring the Eu(III) emission at 616 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01-2Nap (90 μM), aqueous HEPES (10 mM, pH 7.0) and Eu(III) at 1 equivalence to the trimer unit.

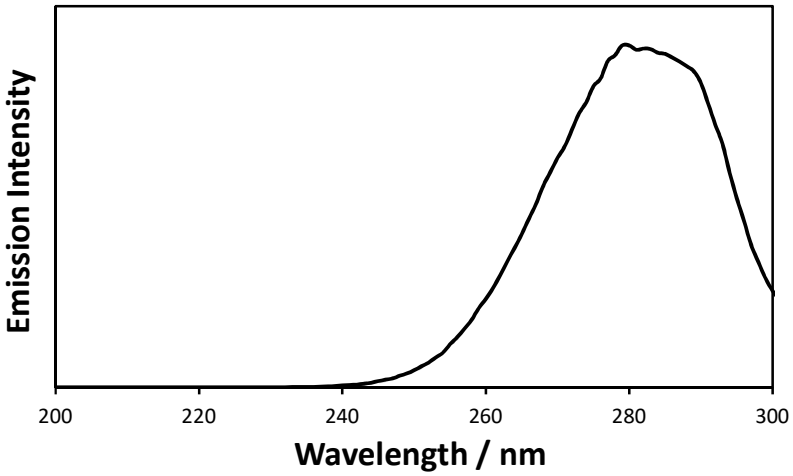

Figure 85: The excitation profile from monitoring the Dy(III) emission at 616 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01-2Nap (90 μM), aqueous HEPES (10 mM, pH 7.0) and Dy(III) at 1 equivalence to the trimer unit.

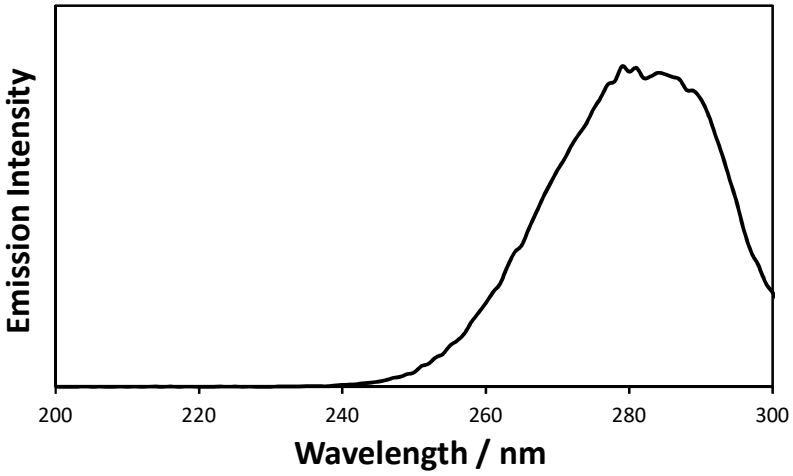

Figure 86: The excitation profile from monitoring the Sm(III) emission at 616 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01-2Nap (90 uM), HEPES (10 mM, pH 7.0) and Sm(III) at 1 equivalence to the trimer unit.

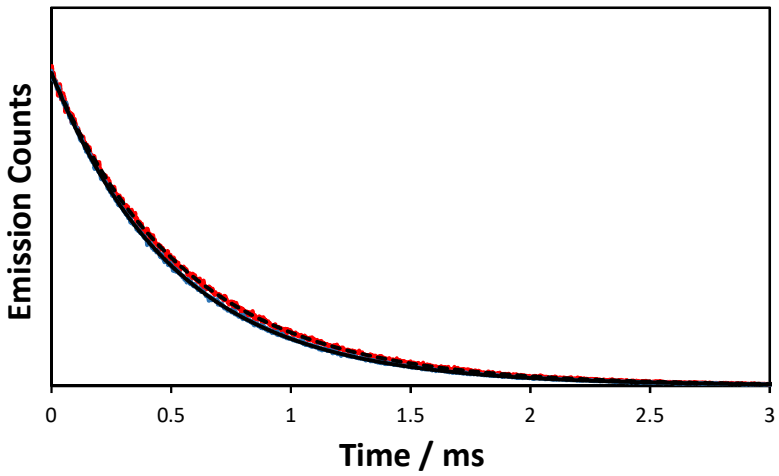

Figure 87: The lifetime profiles for HC01-2Nap with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC01-2Nap (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 13: The respective lifetimes associated with the emission from Tb(III) bound in the HC01-2Nap system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O *Run 1*	509.0	23603.1	**1.223**
Run 2	506.2	41665.4	**1.242**
Run 3	508.5	21772.8	**1.184**
Average	**507.9 ± 1.5**		
D2O *Run 1*	549.2	10835.8	**1.298**
Run 2	546.9	9801.5	**1.240**
Run 3	551.0	14298.0	**1.250**
Average	**549.0 ± 2.0**		

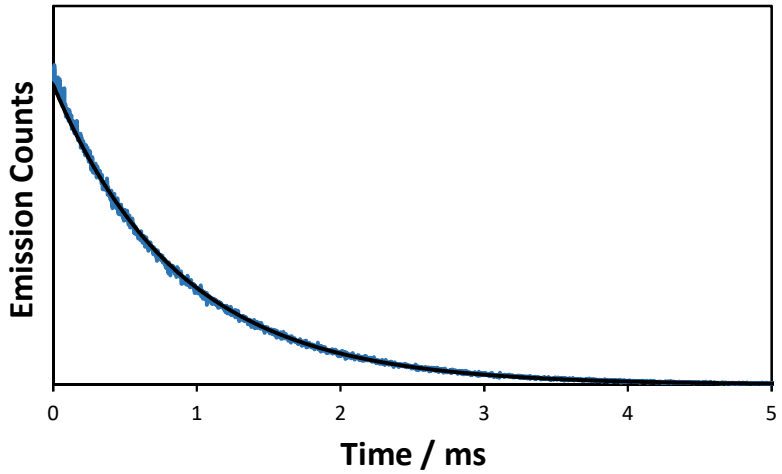

Figure 88: A lifetime profile for HC01-2Nap with Eu(III) monitoring the emission at 616 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01-2Nap (90 µM), HEPES buffer (10 mM, pH 7.0) and Eu(III) in a third equivalence to trimer unit.

Table 14: The respective lifetimes associated with the emission from Eu(III) bound in the HC01-2Nap system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / µs	B	χ^2
H2O ***Run 1***	884.8	4777.0	**1.101**
Run 2	886.9	4953.8	**1.099**
Run 3	887.2	4979.0	**1.120**
Average	**886.3 ± 1.3**		

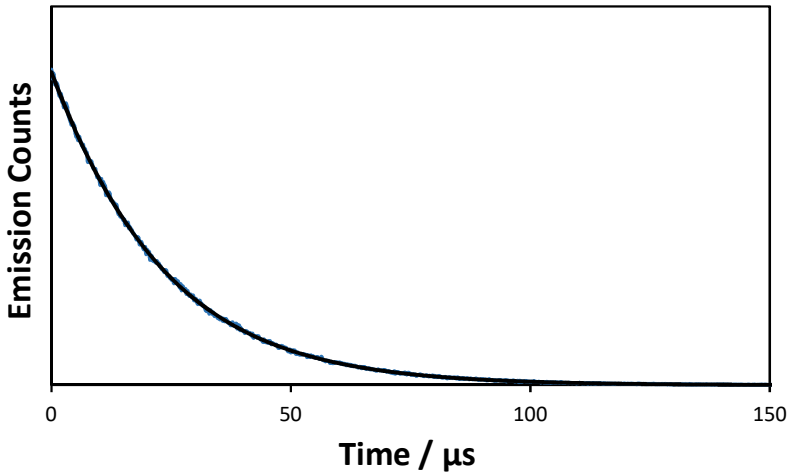

Figure 89: A lifetime profile for HC01-2Nap with Dy(III) monitoring the emission at 571 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01-2Nap (90 µM), HEPES buffer (10 mM, pH 7.0) and Dy(III) in a third equivalence to trimer unit.

Table 15: The respective lifetimes associated with the emission from Dy(III) bound in the HC01-2Nap system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / µs	B	χ^2
H2O *Run 1*	21.50	15580.0	**0.740**
Run 2	22.33	9782.9	**1.212**
Run 3	21.77	13044.8	**1.094**
Average	**21.87 ± 0.42**		

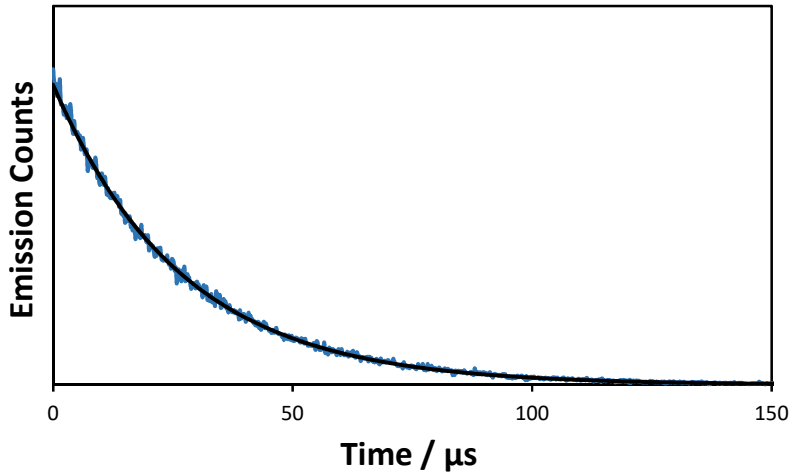

Figure 90: A lifetime profile for HC01-2Nap with Sm(III) monitoring the emission at 643 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01-2Nap (90 µM), HEPES buffer (10 mM, pH 7.0) and Sm(III) in a third equivalence to trimer unit.

Table 16: The respective lifetimes associated with the emission from Sm(III) bound in the HC01-2Nap system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / µs	B	χ^2
H2O ***Run 1***	26.96	7960.4	**0.872**
Run 2	26.85	1926.9	**1.066**
Run 3	27.09	2400.1	**1.046**
Average	**26.97 ± 0.12**		

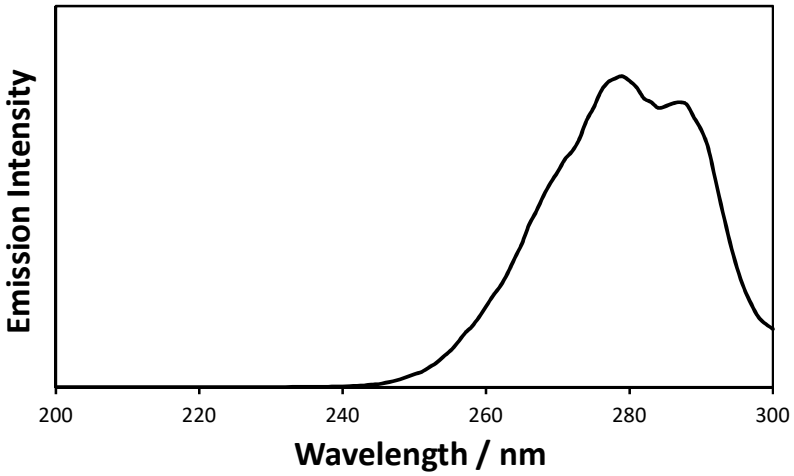

Figure 91: The excitation profile from monitoring the naphthyl- emission of native HC01-2Nap at 333 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01-2Nap (90 μM) and aqueous HEPES (10 mM, pH 7.0).

9.12 HC01-2COU EXCITATION SPECTRA AND LIFETIMES

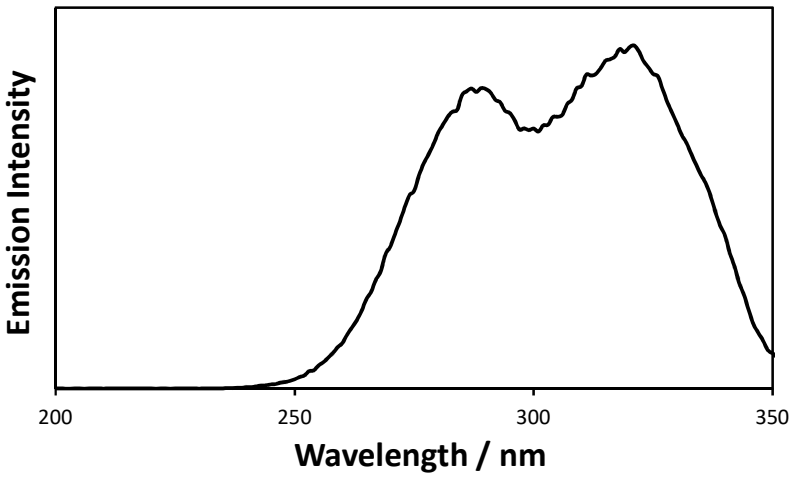

Figure 92: The excitation profile from monitoring the Tb(III) emission at 545 nm while scanning the excitation wavelength from 200-350 nm. The sample consisted of HC01-2Cou (90 μM), aqueous HEPES (10 mM, pH 7.0) and Tb(III) at 1 equivalence to the trimer unit.

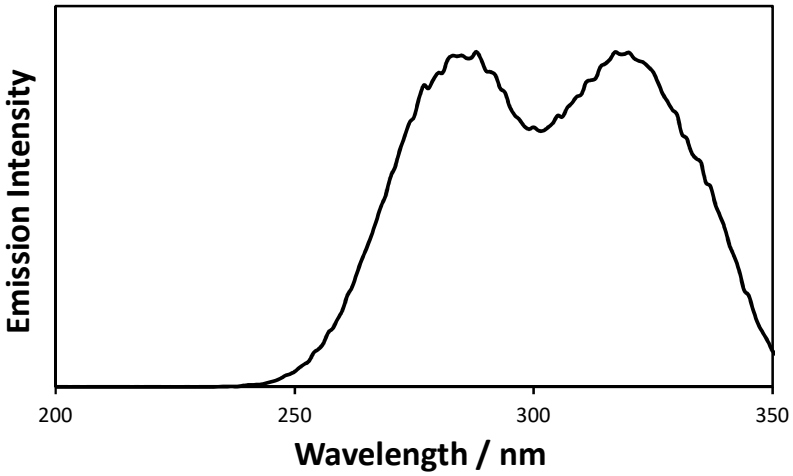

Figure 93: The excitation profile from monitoring the Eu(III) emission at 545 nm while scanning the excitation wavelength from 200-350 nm. The sample consisted of HC01-2Cou (90 μM), aqueous HEPES (10 mM, pH 7.0) and Eu(III) at 1 equivalence to the trimer unit.

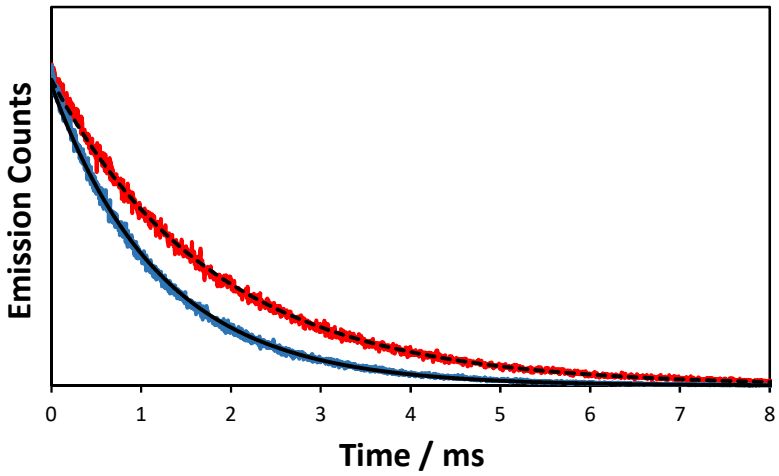

Figure 94: The lifetime profiles for HC01-2Cou with Tb(III) monitoring the emission at 545 nm in both water (blue) and deuterium oxide (red) with their respective exponential fits as a solid black line and a dashed black line respectively. Samples consisted of HC01-2Cou (90 μM), HEPES buffer (10 mM, pH 7.0) and Tb(III) in a third equivalence to trimer unit.

Table 17: The respective lifetimes associated with the emission from Tb(III) bound in the HC01-2Cou system both in water and deuterium oxide. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O *Run 1*	1174.1	2003.0	**1.257**
Run 2	1217.8	3308.6	**1.465**
Run 3	1234.5	1231.7	**1.242**
Average	**1208.8 ± 31.2**		
D2O *Run 1*	1784.0	3025.2	**1.187**
Run 2	1812.9	2471.4	**1.199**
Run 3	1817.2	2440.3	**1.177**
Average	**1804.7 ± 18.0**		

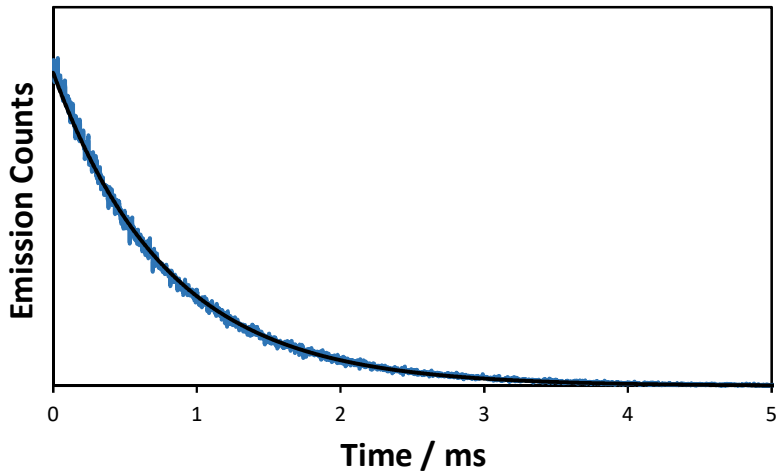

Figure 95: A lifetime profile for HC01-2Cou with Eu(III) monitoring the emission at 616 nm in water (blue) with the exponential tail fit as a solid black line. Samples consisted of HC01-2Cou (90 μM), HEPES buffer (10 mM, pH 7.0) and Eu(III) in a third equivalence to trimer unit.

Table 18: The lifetime associated with the emission from Eu(III) bound in the HC01-2Cou system in water. The χ^2 value are given for the validity of the lifetime tail fits.

Sample	τ / μs	B	χ^2
H2O ***Run 1***	800.7	2055.1	**0.894**
Run 2	799.2	4982.6	**1.156**
Run 3	809.8	2432.3	**0.923**
Average	**803.3 ± 5.7**		

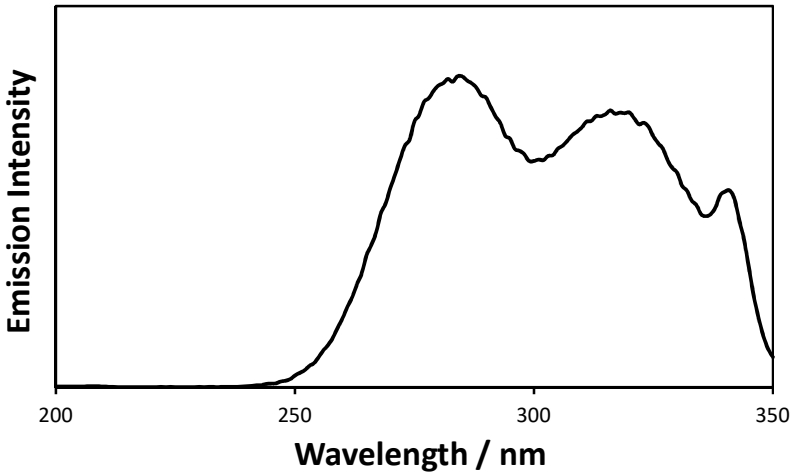

Figure 96: The excitation profile from monitoring the coumaryl- emission of native HC01-2Cou at 389 nm while scanning the excitation wavelength from 200-300 nm. The sample consisted of HC01-2Cou (90 μM) and aqueous HEPES (10 mM, pH 7.0).

9.13 MOLAR ABSORPTIVITY DETERMINATION

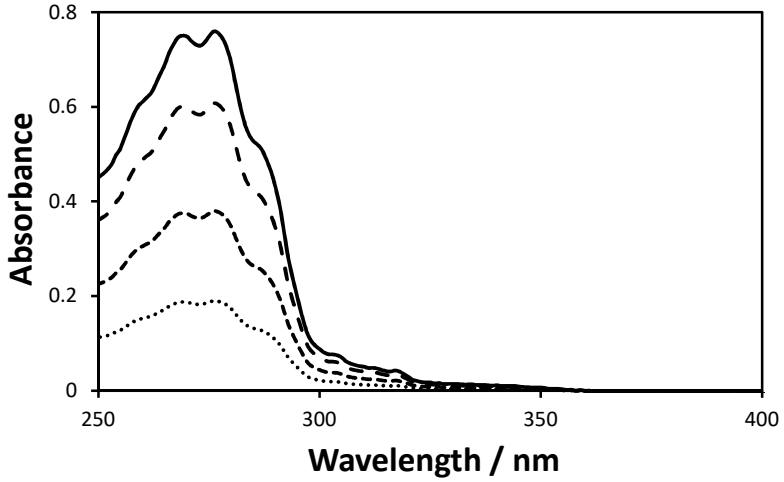

Figure 97: The absorption profile of HC01-2Nap at 120 μM (solid line), 90 μM (long dashed line), 67 μM (short dashed line) and 33 μM (dotted line). The sample solvent was urea (7 M).

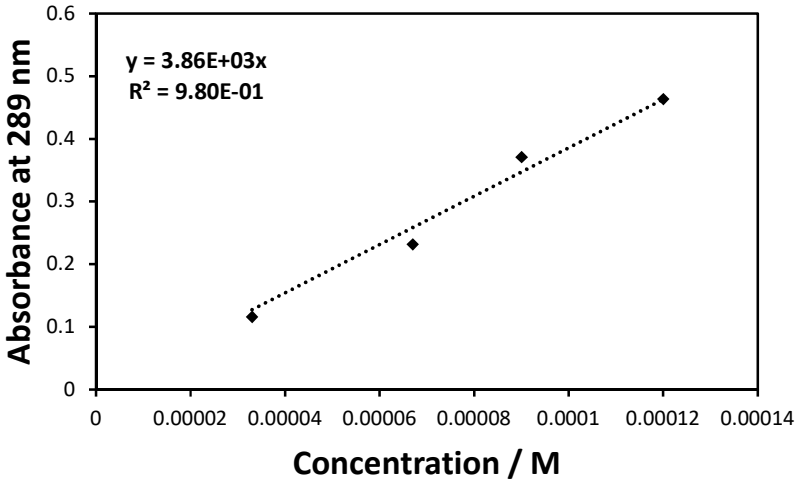

Figure 98: The absorbance of HC01-2Nap at 289 nm as a function of the concentration of the peptide. The gradient of the linear fit is equal to the molar absorptivity of HC01-2Nap at 289 nm.

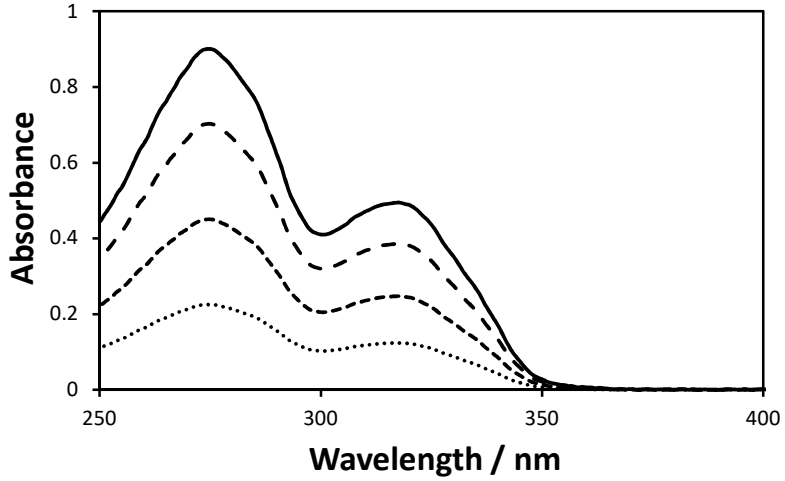

Figure 99: The absorption profile of HC01-2Cou at 120 μM (solid line), 90 μM (long dashed line), 60 μM (short dashed line) and 30 μM (dotted line). The sample solvent was urea (7 M).

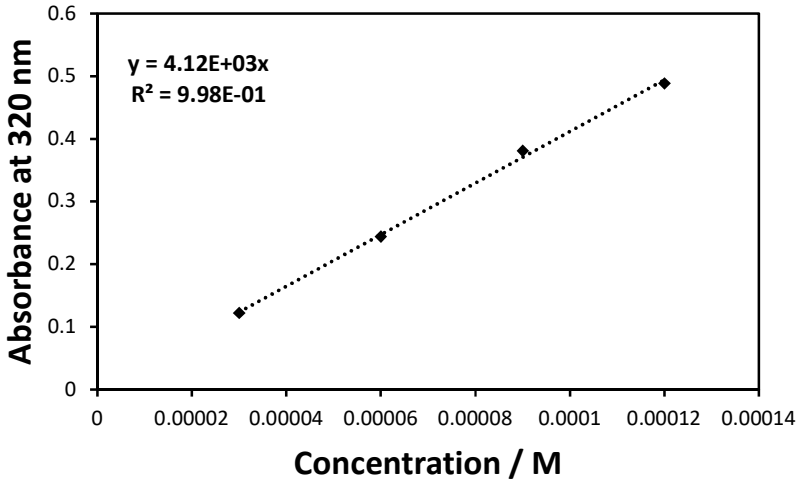

Figure 100: The absorbance of HC01-2Cou at 320 nm as a function of the concentration of the peptide. The gradient of the linear fit is equal to the molar absorptivity of HC01-2Nap at 320 nm.

9.14 SMALL MOLECULE CHARACTERISATION

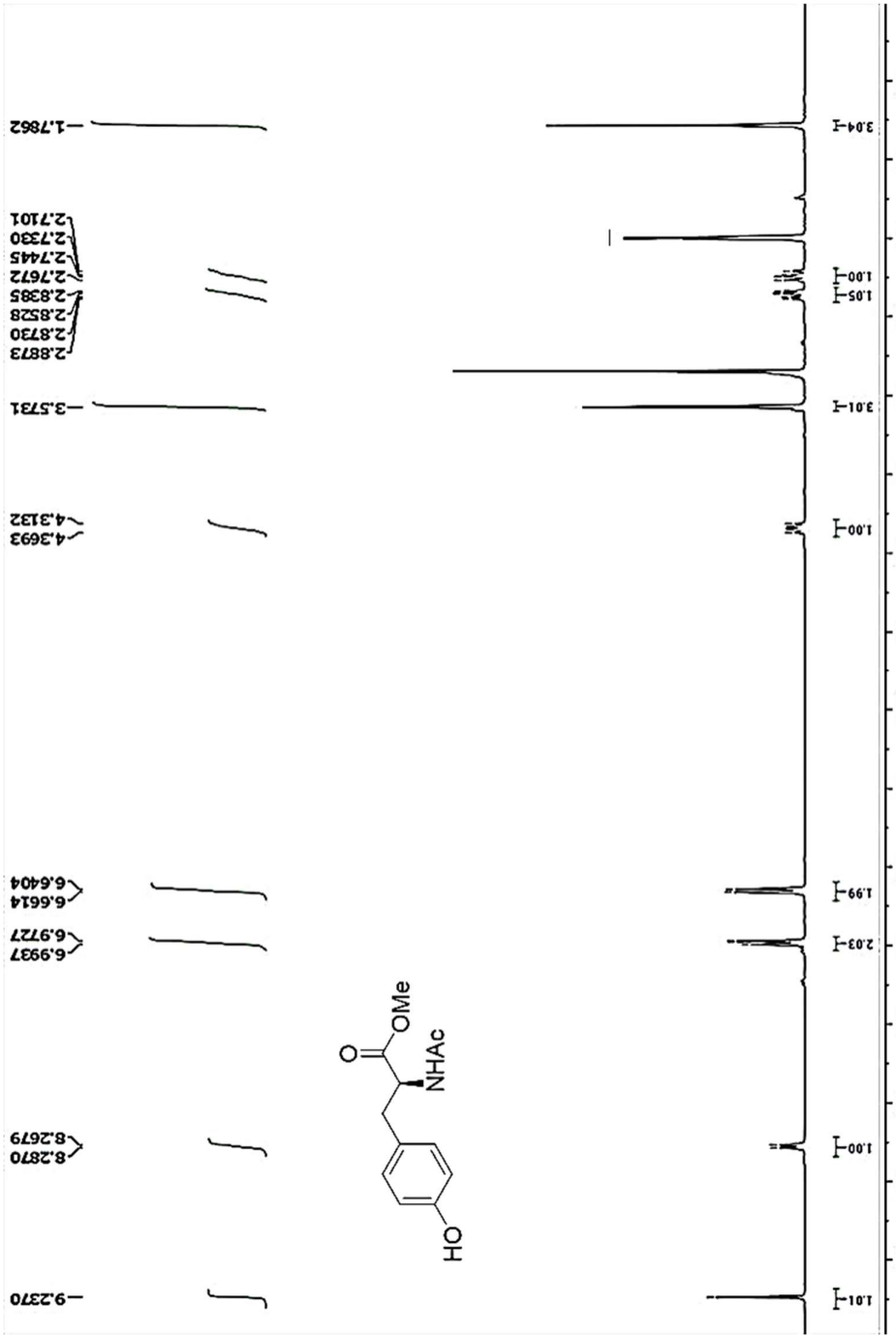

Figure 101: The proton nuclear magnetic resonance spectrum for N-acetyl-L-tyrosine methyl ester in DMSO (400 MHz).

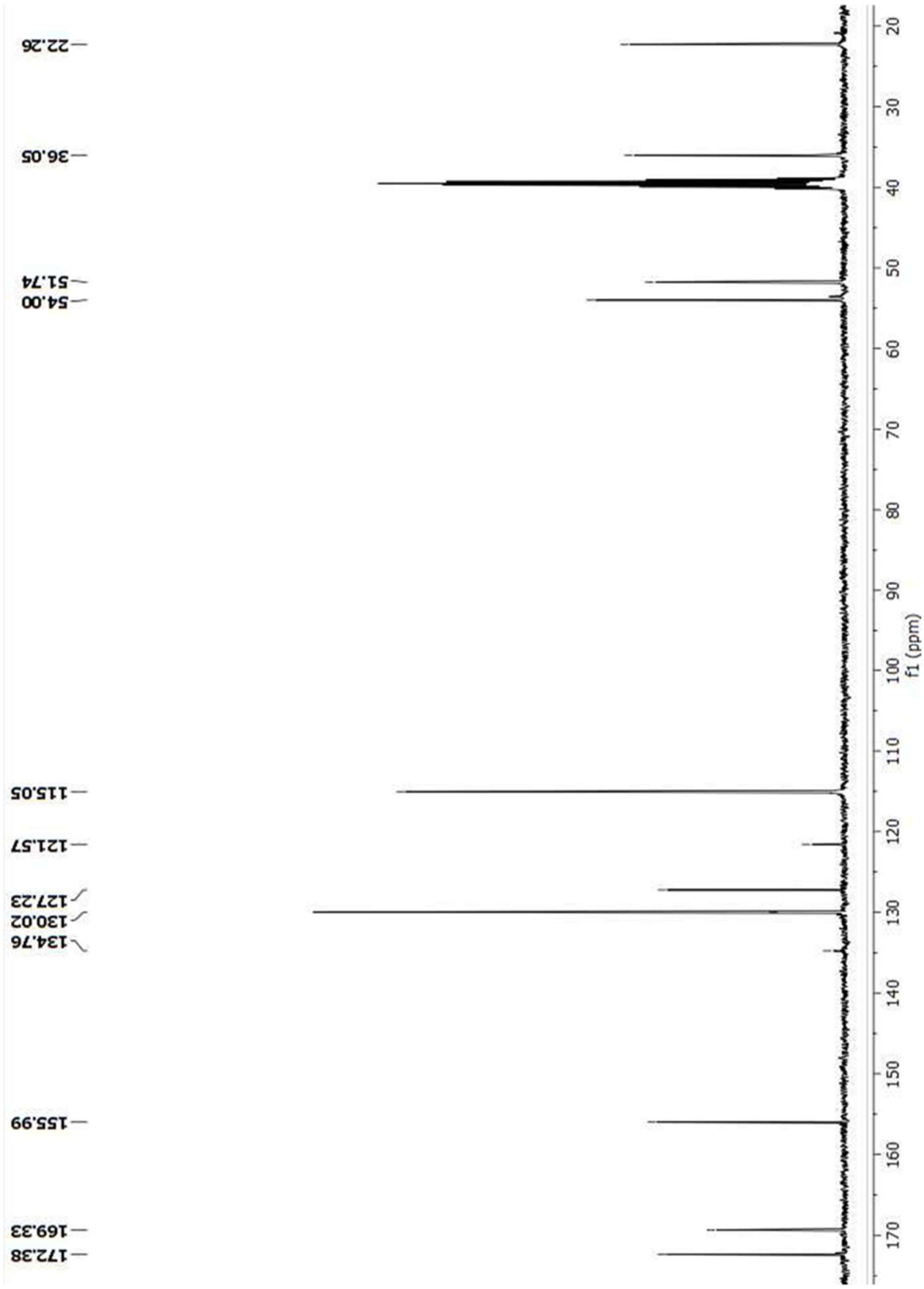

Figure 102: The carbon nuclear magnetic resonance spectrum for N-acetyl-L-tyrosine methyl ester in DMSO (400 MHz)

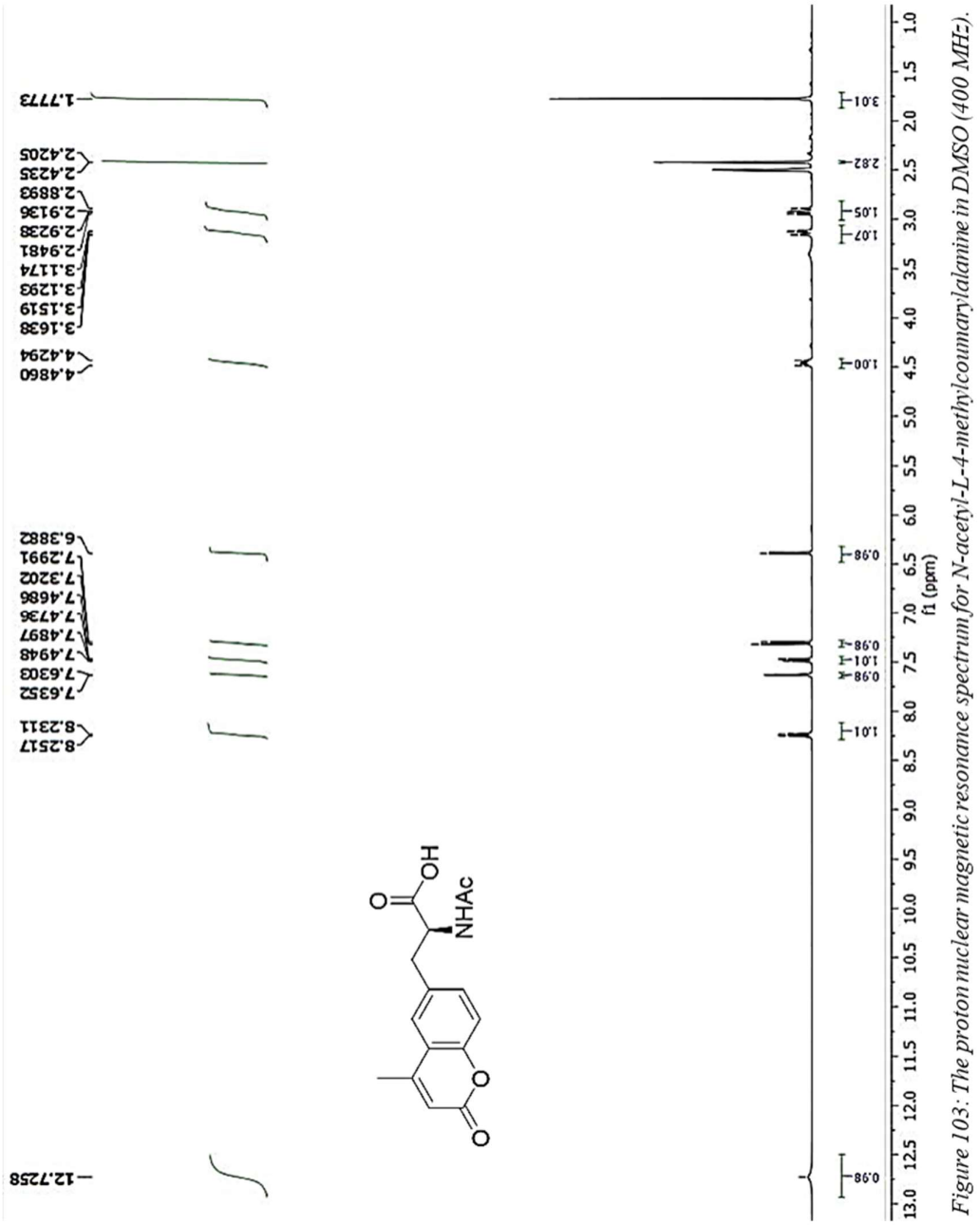

Figure 103: The proton nuclear magnetic resonance spectrum for N-acetyl-L-4-methylcoumarylalanine in DMSO (400 MHz).

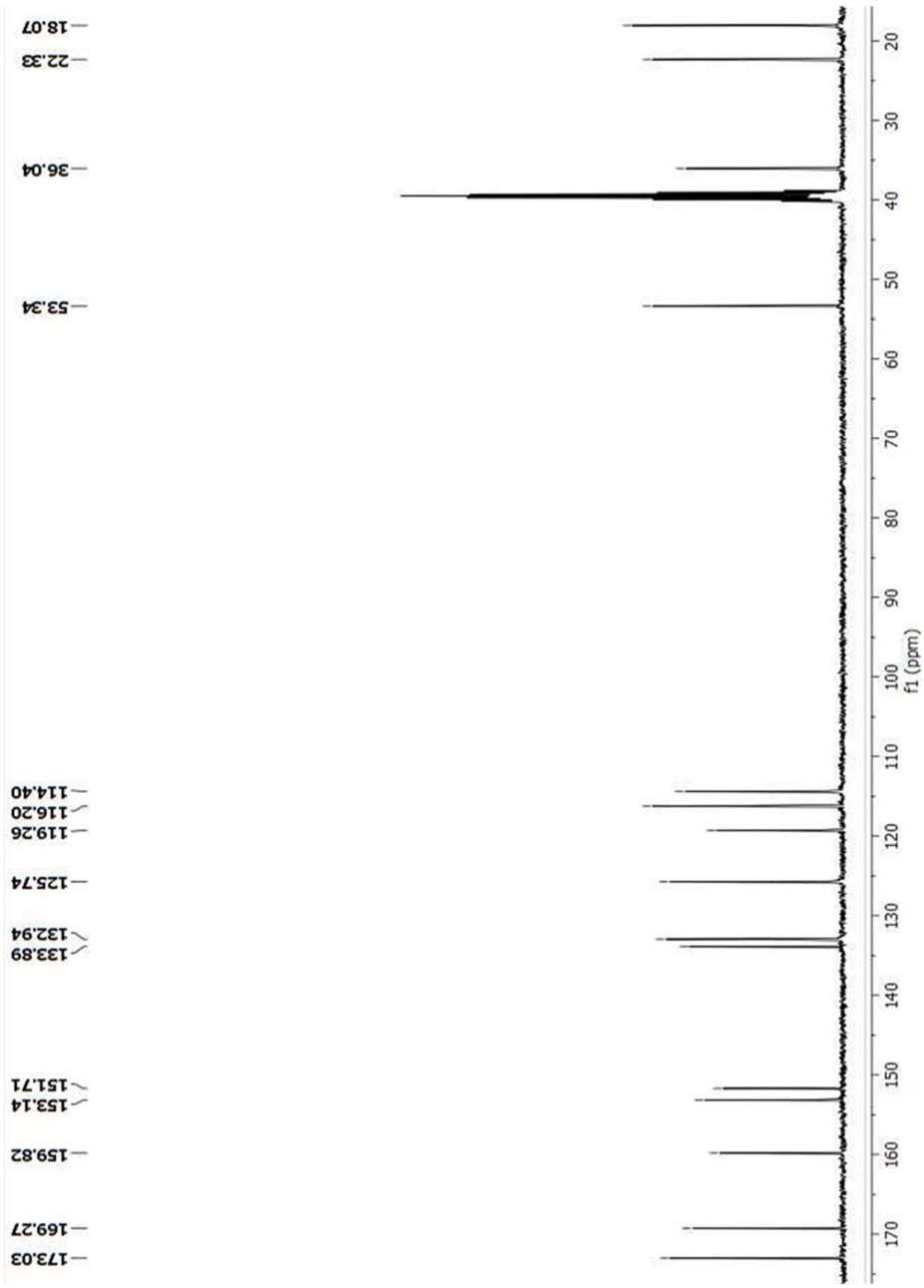

Figure 104: The carbon nuclear magnetic resonance spectrum for N-acetyl-L-4-methylcoumarylalanine in DMSO (400 MHz).

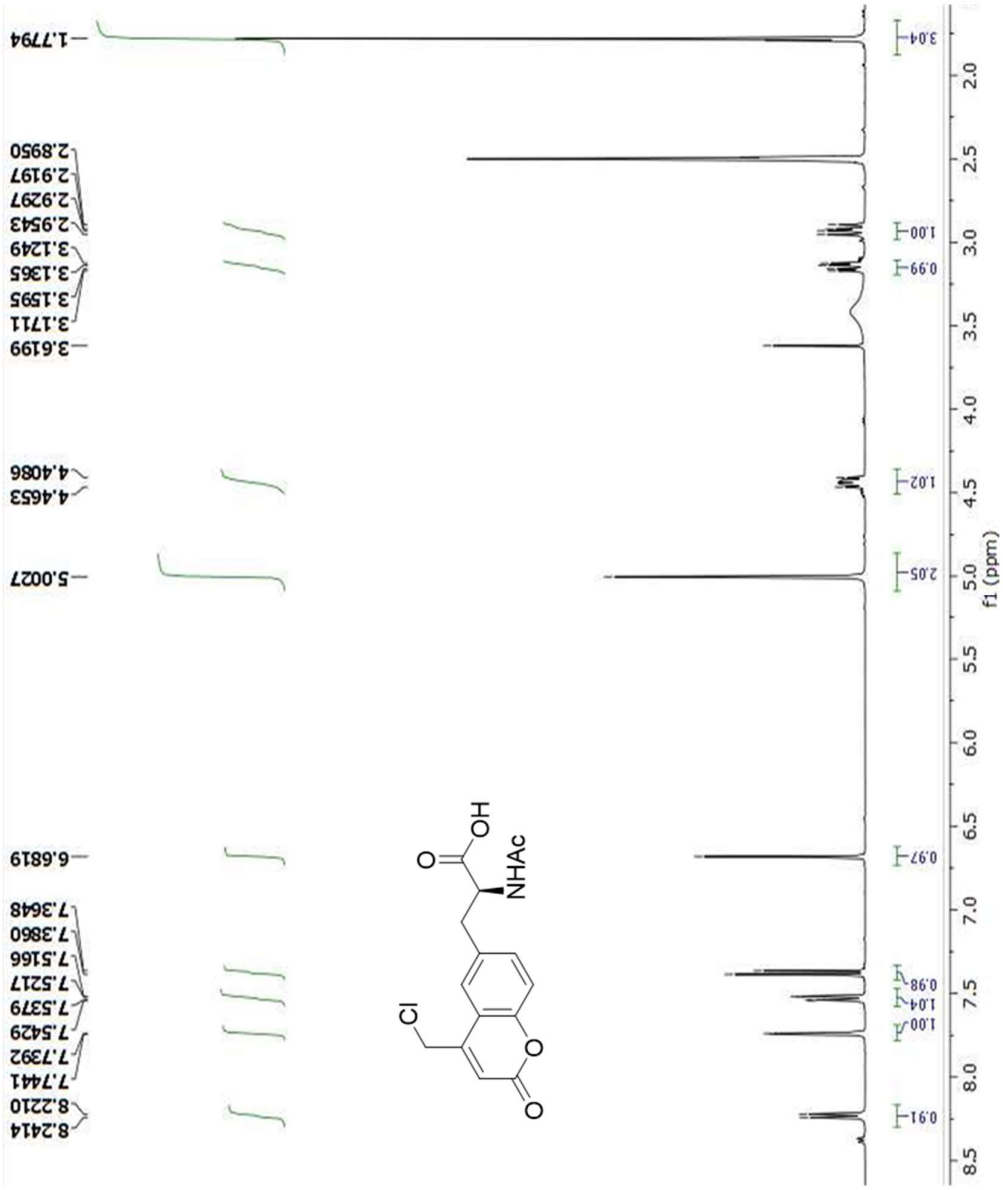

Figure 105: The proton nuclear magnetic resonance spectrum for N-acetyl-L-4-chloromethylcoumarylalanine in DMSO (400 MHz).

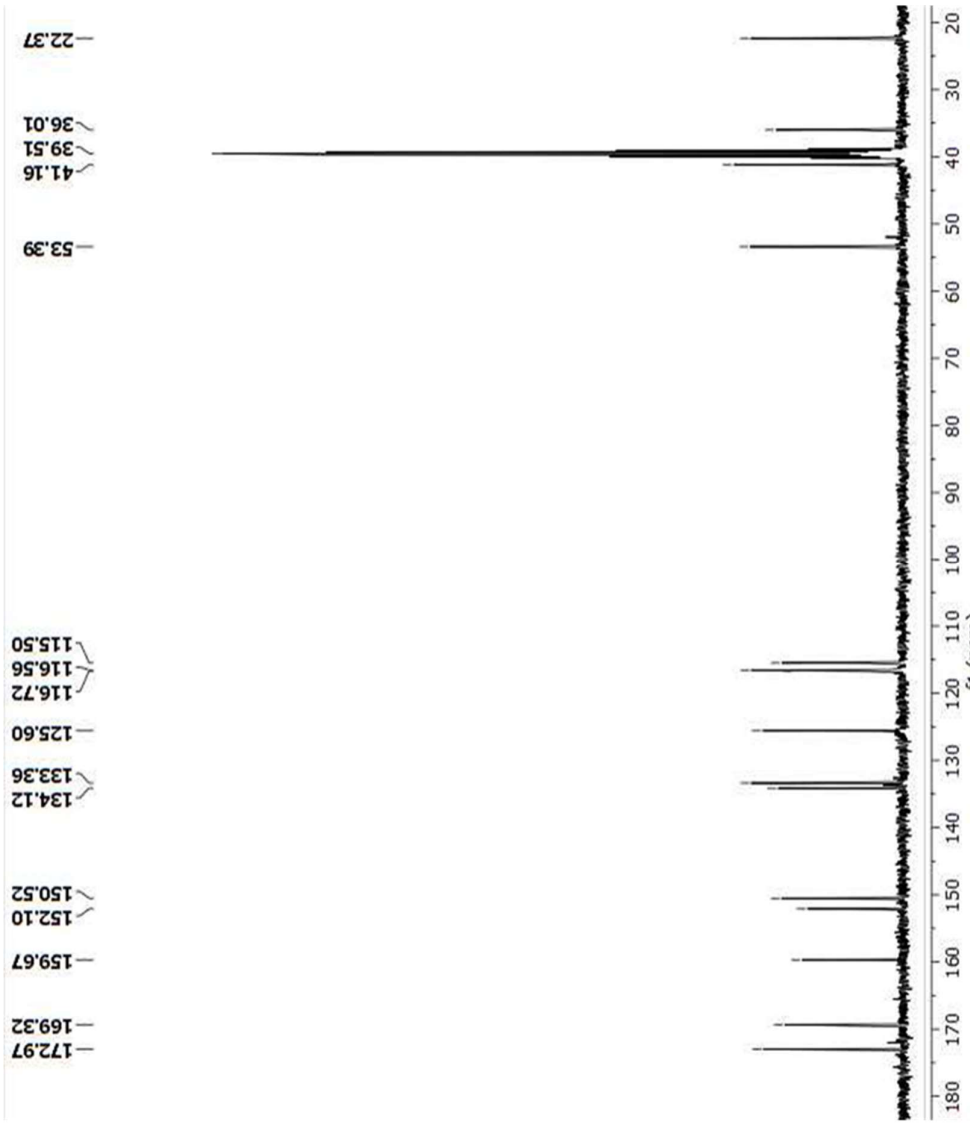

Figure 106: The carbon nuclear magnetic resonance spectrum for N-acetyl-L-4-chloromethylcoumarylalamine in DMSO (400 MHz).

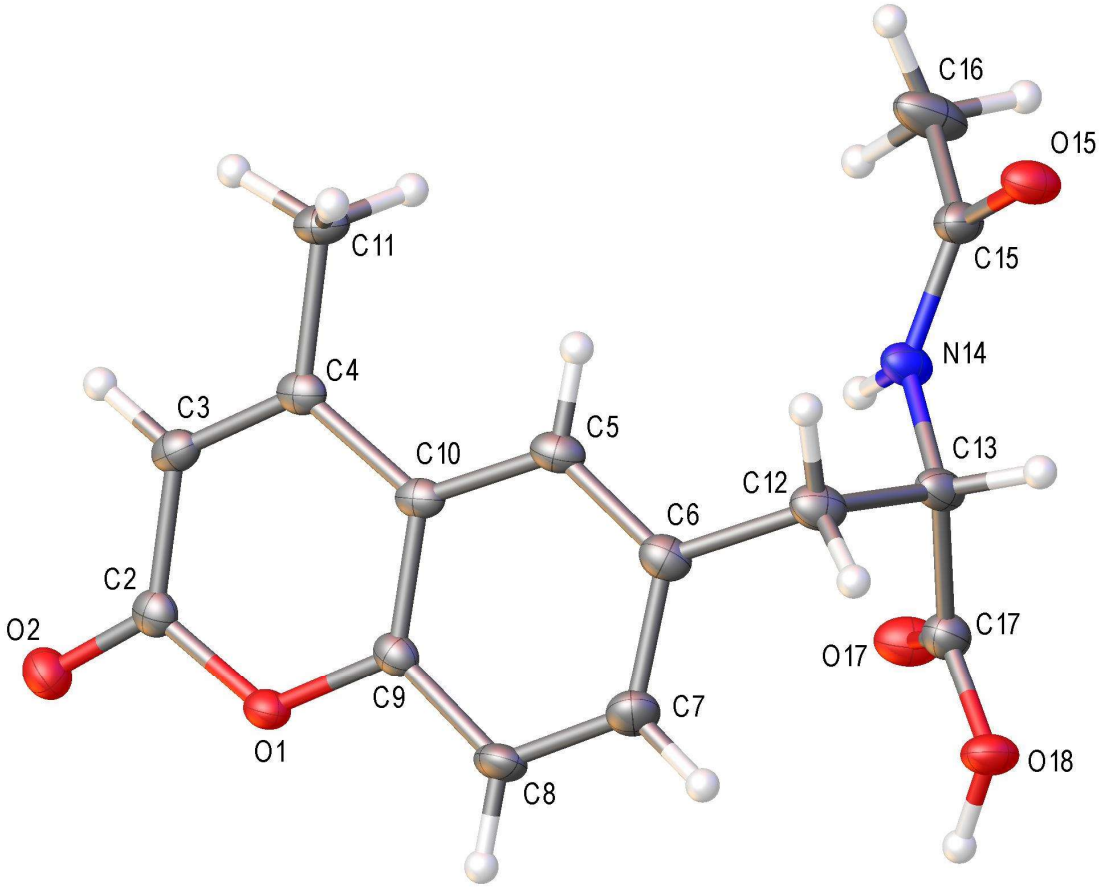

Figure 107: The crystal structure of N-acetyl-L-4-methylcoumarylalanine with ellipsoids drawn at 50% probability level. The atom numerical identities are shown and corresponds with the coordinate tables.

Crystal Notes: The space group is chiral and the Flack parameter is 0.09 (8). The error in the Flack parameter is too large for the absolute structure to have been determined from the diffraction data, but from the synthetic method we can conclude that C(13) is S.

The hydrogen atoms bonded to N(14) and O(18) were located in the electron density and their positions and thermal displacement parameters were refined freely. All other hydrogen atoms were fixed as riding models.

Table 19: Crystal data and structure refinement for *N*-acetyl-*L*-4-methylcoumarylalanine	
Identification code	C1-NAC-OH
Empirical formula	$C_{15}H_{15}NO_5$
Formula weight	289.28
Temperature/K	100.01(10)
Crystal system	orthorhombic
Space group	$P2_12_12_1$
a/Å	6.45398(12)
b/Å	7.55646(16)
c/Å	27.4103(5)
α/°	90
β/°	90
γ/°	90
Volume/Å^3	1336.78(5)
Z	4
ρ_{calc}g/cm^3	1.437
μ/mm^{-1}	0.913
F(000)	608.0
Crystal size/mm^3	0.265 × 0.224 × 0.051
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	12.918 to 144.188
Index ranges	$-7 \leq h \leq 7$, $-8 \leq k \leq 9$, $-32 \leq l \leq 33$
Reflections collected	7501
Independent reflections	2564 [R_{int} = 0.0207, R_{sigma} = 0.0207]
Data/restraints/parameters	2564/0/200
Goodness-of-fit on F^2	1.062
Final R indexes [I>=2σ (I)]	R_1 = 0.0319, wR_2 = 0.0828
Final R indexes [all data]	R_1 = 0.0344, wR_2 = 0.0857
Largest diff. peak/hole / e Å^{-3}	0.29/-0.18
Flack parameter	0.09(8)

Table 20: Fractional Atomic Coordinates (×104) and Equivalent Isotropic Displacement Parameters (Å2×103) for C1-NAC-OH. Ueq is defined as 1/3 of of the trace of the orthogonalised UIJ tensor.

Atom	*x*	*y*	*z*	U(eq)
C2	-2(3)	8916(3)	5332.8(6)	17.2(4)
C3	3(3)	7005(2)	5343.9(7)	17.4(4)
C4	15(3)	6012(2)	4934.2(7)	16.6(4)
C5	89(3)	6028(3)	4011.2(7)	18.2(4)
C6	63(3)	6946(3)	3574.7(7)	19.0(4)
C7	-87(3)	8798(3)	3589.7(7)	21.4(4)
C8	-125(3)	9706(3)	4025.6(7)	20.6(4)
C9	-46(3)	8750(2)	4458.2(7)	16.2(4)
C10	19(3)	6905(2)	4462.8(7)	15.5(4)
C11	16(3)	4029(2)	4951.4(7)	21.0(4)
C12	312(3)	6001(3)	3092.5(7)	21.2(4)
C13	2620(3)	5840(3)	2942.4(7)	19.9(4)
C15	4060(3)	2959(3)	3182.3(7)	20.4(4)
C16	5568(4)	2045(3)	3513.1(9)	34.7(6)
C17	3639(3)	7659(3)	2961.6(7)	21.2(4)
N14	3794(3)	4680(2)	3261.0(6)	20.5(4)
O1	-52(2)	9722.7(17)	4882.5(5)	17.5(3)
O2	19(2)	9898.4(18)	5683.2(5)	23.3(3)
O15	3161(2)	2166.3(19)	2845.2(5)	24.1(3)
O17	4985(3)	8029.0(18)	3250.1(5)	27.6(3)
O18	2796(2)	8780.7(19)	2648.8(5)	23.2(3)

Table 21: Anisotropic Displacement Parameters (Å2×103) for C1-NAC-OH. The Anisotropic displacement factor exponent takes the form: -2π2[h2a*2U11+2hka*b*U12+…].

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C2	12.4(8)	18.8(9)	20.5(9)	-0.4(7)	1.5(8)	0.7(9)
C3	13.0(8)	17.4(9)	21.9(9)	4.5(7)	1.1(8)	1.0(9)
C4	10.3(8)	15.8(9)	23.5(9)	1.0(7)	0.4(8)	0.2(8)
C5	15.8(8)	13.5(9)	25.4(9)	-1.1(7)	-3.3(9)	0.1(9)
C6	17.5(9)	17.3(9)	22.1(9)	-2.3(7)	-3.1(8)	0.5(9)
C7	22.5(9)	19.2(10)	22.4(9)	3.3(7)	-1.9(9)	4(1)
C8	22.3(9)	12.7(9)	26.8(10)	1.3(7)	2.0(9)	4.7(9)
C9	13.4(8)	15.1(9)	20.0(9)	-2.3(7)	0.7(9)	0.1(8)
C10	10.1(8)	13.6(9)	22.8(9)	1.2(7)	-0.4(8)	-0.3(8)
C11	22.6(9)	14.1(9)	26.3(9)	2.7(7)	-0.2(9)	-0.5(9)
C12	24.6(10)	17.4(9)	21.8(10)	-2.5(7)	-6.6(8)	-0.1(9)
C13	28.1(10)	15.2(9)	16.3(9)	-1.9(7)	-3.6(8)	3.2(8)
C15	28(1)	15.0(9)	18.2(9)	0.9(8)	-0.4(8)	1.1(8)
C16	52.5(15)	20.5(11)	31.0(11)	-5.1(9)	-16.4(10)	12.6(11)
C17	25.2(10)	18.9(10)	19.6(9)	0.0(8)	0.2(8)	2.9(9)
N14	29.0(9)	15.7(8)	16.8(8)	-2.4(6)	-4.9(7)	3.1(7)
O1	19.4(6)	12.9(6)	20.3(7)	-0.4(5)	1.6(6)	0.1(6)
O2	26.1(7)	21.6(7)	22.3(7)	-3.9(6)	2.6(6)	0.0(7)
O15	33.4(8)	15.2(6)	23.6(7)	-1.5(5)	-6.0(6)	1.4(7)
O17	29.7(7)	20.0(7)	33.1(8)	0.6(6)	-7.8(7)	-0.7(7)
O18	31.4(8)	14.7(7)	23.6(7)	2.2(6)	-3.3(6)	1.9(6)

Table 22: Bond Lengths for C1-NAC-OH.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
C2	C3	1.444(3)	C8	C9	1.389(3)
C2	O1	1.377(2)	C9	C10	1.395(2)
C2	O2	1.214(2)	C9	O1	1.376(2)
C3	C4	1.350(3)	C12	C13	1.550(3)
C4	C10	1.458(2)	C13	C17	1.524(3)
C4	C11	1.500(3)	C13	N14	1.451(3)
C5	C6	1.383(3)	C15	C16	1.498(3)
C5	C10	1.405(3)	C15	N14	1.329(3)
C6	C7	1.403(3)	C15	O15	1.245(2)
C6	C12	1.511(3)	C17	O17	1.207(2)
C7	C8	1.378(3)	C17	O18	1.323(2)

Table 23: Bond Angles for C1-NAC-OH.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
O1	C2	C3	117.48(16)	C5	C10	C4	124.27(17)
O2	C2	C3	126.50(18)	C9	C10	C4	118.04(17)
O2	C2	O1	116.02(17)	C9	C10	C5	117.69(17)
C4	C3	C2	122.53(17)	C6	C12	C13	111.79(16)
C3	C4	C10	118.72(17)	C17	C13	C12	109.57(17)
C3	C4	C11	121.93(17)	N14	C13	C12	112.93(17)
C10	C4	C11	119.35(17)	N14	C13	C17	107.37(16)
C6	C5	C10	121.67(17)	N14	C15	C16	115.88(18)
C5	C6	C7	118.40(18)	O15	C15	C16	122.01(18)
C5	C6	C12	121.21(17)	O15	C15	N14	122.07(18)
C7	C6	C12	120.30(18)	O17	C17	C13	122.81(18)
C8	C7	C6	121.54(18)	O17	C17	O18	124.88(19)
C7	C8	C9	118.73(18)	O18	C17	C13	112.19(16)
C8	C9	C10	121.89(18)	C15	N14	C13	124.13(17)
O1	C9	C8	116.33(16)	C9	O1	C2	121.40(15)
O1	C9	C10	121.78(17)				

Table 24: Torsion Angles for C1-NAC-OH.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
C2	C3	C4	C10	-0.1(3)	C10	C5	C6	C12	175.04(18)
C2	C3	C4	C11	-179.86(19)	C10	C9	O1	C2	-0.2(3)
C3	C2	O1	C9	-1.6(3)	C11	C4	C10	C5	-2.1(3)
C3	C4	C10	C5	178.14(19)	C11	C4	C10	C9	178.08(19)
C3	C4	C10	C9	-1.7(3)	C12	C6	C7	C8	-174.03(19)
C5	C6	C7	C8	2.5(4)	C12	C13	C17	O17	113.7(2)
C5	C6	C12	C13	-89.0(2)	C12	C13	C17	O18	-62.7(2)
C6	C5	C10	C4	179.17(19)	C12	C13	N14	C15	91.9(2)
C6	C5	C10	C9	-1.0(3)	C16	C15	N14	C13	171.6(2)
C6	C7	C8	C9	-1.0(4)	C17	C13	N14	C15	-147.20(19)
C6	C12	C13	C17	-52.3(2)	N14	C13	C17	O17	-9.3(3)
C6	C12	C13	N14	67.3(2)	N14	C13	C17	O18	174.32(15)
C7	C6	C12	C13	87.5(3)	O1	C2	C3	C4	1.7(3)
C7	C8	C9	C10	-1.7(3)	O1	C9	C10	C4	1.9(3)
C7	C8	C9	O1	178.88(19)	O1	C9	C10	C5	-177.94(17)
C8	C9	C10	C4	-177.55(18)	O2	C2	C3	C4	-178.81(19)
C8	C9	C10	C5	2.6(3)	O2	C2	O1	C9	178.93(18)
C8	C9	O1	C2	179.23(18)	O15	C15	N14	C13	-6.3(3)
C10	C5	C6	C7	-1.5(3)					

9.15 REFERENCES

[1] Edgeplots Web Tool, http://skuld.bmsc.washington.edu/scatter/AS_form.html, (accessed December 2018).