

The design and NMR structure determination of yttrium-oligopeptide tags for recombinant proteins and antibodies

Marcela Múdra^a, Martin Breza^a, Lucia Lintnerová^b,
Juraj Filo^c, Jacob Bauer^d

^a*Faculty of Chemical and Food Technology, Slovak University of Technology,
Radlinského 9, SK-812 37 Bratislava, Slovakia*

^b*Department of Chemical Theory of Drugs, Faculty of Pharmacy, Comenius University,
Odbojárov 10, SK-832 32 Bratislava, Slovakia*

^c*Institute of Chemistry, Faculty of Natural Sciences, Comenius University,
Mlynská dolina CH-2, SK-842 15 Bratislava, Slovakia*

^d*Institute of Molecular Biology, Slovak Academy of Sciences,
Dúbravská cesta 21, SK-845 51 Bratislava, Slovakia
martin.breza@stuba.sk*

Abstract: A strategy for the design of new yttrium(III) tags consisting of sequences of naturally occurring amino acids is described. These tags are 4–6 amino acids in length and consist of aspartic and glutamic acids. The use of natural amino acids would allow these oligopeptides to be incorporated into recombinant proteins at the DNA level, enabling the protein to be Y(III)-labelled after protein isolation. This allows a radionuclide or heavy atom to be associated with the protein without the necessity of further synthetic modification. Suitable peptides able to chelate Y(III) in stable complexes were designed based on quantum-chemical calculations. The stability of complexes formed by these peptides was tested by isothermal titration calorimetry, giving dissociation constants in the micromolar range. The likely structure of the most tightly bound complex was inferred from a combination of NMR experiments and quantum-chemical calculations. This structure will serve as the basis for future optimizations.

Keywords: carboxylate ligands; density functional calculations; NMR spectroscopy; peptide design; yttrium tags

Introduction

Targeted radiotherapy is an appealing approach to cancer treatment because of its potential for delivering curative doses of radiation to tumor cells while sparing normal tissues. One of the most successful treatments makes use of the β -emitter ^{90}Y chelated to antibodies (Goodwin et al., 2001; Mather, 2007). In this treatment, the ^{90}Y atom is typically bound to the antibody through a completely synthetic bifunctional chelator which has been conjugated to the antibody. One of the best chelators is 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid or DOTA. The great advantage of these synthetic chelators is that they can be synthetically altered to optimize the stability and specificity of the radionuclide–chelator complex. Their great disadvantage is that the chelator must be chemically conjugated to the target antibody, sometimes under conditions which may either crosslink or degrade the target biomolecule. Furthermore, the resulting stoichiometry, of chelators per antibody after conjugation and complexes per antibody following radiolabeling, is unknown, though it is possible to determine the average number of complexes per

molecule. The actual location of the conjugated complex on the antibody is frequently unknown, the level of uncertainty depending on the conjugation method used (Liu et al., 2001).

An approach which might allow the benefits of synthetic chelators to be realized without their attendant disadvantages is to design chelators using only naturally occurring amino acids. The resulting tags could then be incorporated into recombinant biomacromolecules (e.g. antibodies) at the DNA level using molecular cloning. This would avoid the necessity of chemical conjugation and would provide clear antibody–radionuclide stoichiometry and would remove the uncertainty in the position of the complex. A happy side-effect of this would be that these tags could also be used in protein crystallography for binding a heavy atom to the protein for solving the phase problem.

Unfortunately, there are, at present, no examples in the Cambridge Structure Database of a Y or other Lanthanide in complex with a natural peptide ligand. Examples of oligopeptides which bind Lanthanides can be found in the structural biology literature, however, where they have been used to solve the phase problem in protein crystallography

(Nitz et al., 2004; Silvaggi et al., 2007; Martin et al., 2007) and to aid structure determination by NMR (Mustafi et al., 2006; Bentrop et al., 1997; Welch et al., 2003). There are also examples of Lanthanide-binding proteins designed with the ultimate goal of being improved MRI phase-contrast agents (Yang et al., 2005; Yang et al., 2008; Xue et al., 2014). These oligopeptide Lanthanide binding tags were all derived from calcium-binding protein structures. They are all relatively short (20–25 amino acids in length) and can be incorporated into recombinant proteins at the DNA level, but they are all too large for easy use in targeted radiotherapy applications. The Protein Data Bank (PDB) does contain a number of protein structures with bound Y(III) ions (72 as of April 2017), but for the most part, the Y(III) binding sites in these structures occur at the junction of two or more proteins (dubbed “crystal contacts” in the structural biology literature) and are rather poorly defined.

This paper reports the design of shorter tags with simpler tertiary structures which can still be incorporated into recombinant proteins by molecular cloning. Amino-acids with suitable donor atoms were selected based on hard-soft acid-base theory (Pearson, 1963; Pearson, 1968a; Pearson, 1968b). Quantum chemical modeling was then used to design peptides which were likely to possess suitable chelating properties. The dissociation constants of the resulting peptide–yttrium complexes were measured by isothermal titration calorimetry (ITC). The structures of the shortest peptide with the best dissociation constant was then determined. Unfortunately, crystallographic experiments were unsuccessful, so the structure of the resulting complex was modeled by density functional theory calculations and verified by confronting it with NMR experimental data.

Experimental

Materials

All peptides were obtained from Biomatik Corporation (Cambridge, ON, Canada) with 95 % declared purity. Elemental analysis of GDDDD.2TFA found C 34.09, H 4.53, N 9.59, in good agreement with the theoretical prediction of C 34.79, H 3.05, N 9.22, calculated based on a molecular formula of C₂₂H₂₃N₅O₁₈F₆ (759.43). Yttrium(III) chloride hexahydrate, 99.99 % purity, was purchased from Sigma-Aldrich (St. Louis, MO, USA). An 18 mg/ml (24 mM) aqueous solution of GDDDD was mixed with an equimolar amount of YCl₃(aq) and the pH of the resulting solution was adjusted to 6.5 with NaOH. Elemental analysis of the resulting H₃[YGDDDDTFAH₂O]·5H₂O complex

(C₂₀H₃₃N₅O₂₁Y₁F₃; 825.40) found C 29.20, H 5.05, N 8.40 versus the predicted C 29.10, H 4.03, N 8.48.

Calorimetry

Microcalorimetry experiments were performed at 25 °C using a VP-ITC titration calorimetric system (MicroCal, Northampton, USA). Reactant solutions were prepared in 50 mM cacodylate, MES or ACES buffer (pH 6.0). 0.02 mM peptide solutions (the exact peptide concentrations were determined by amino acid-analyses) were titrated with 8 µl aliquots of 0.3 mM yttrium(III) chloride to saturation. The control experiment of adding YCl₃ into buffer alone was also performed. Thermodynamic parameters, stoichiometry, and association and dissociation constants were determined using software developed by MicroCal and implemented in Origin 7.0 (MicroCal, Northampton, USA).

NMR spectroscopy

NMR spectra were measured with a Varian NMR System™ 600 NMR spectrometer working at 599.77 (¹H), 150.83 (¹³C) and 60.79 MHz (¹⁵N) equipped with a Varian 5 mm PFG OneNMR Probe. Measurements were done in 400 µl D₂O; TSP-d₄ (Sodium 3-(trimethylsilyl)propionate-2,2,3,3-d₄) served as an internal standard for ¹H and ¹³C. Nitromethane was used as a reference for the ¹⁵N chemical shift. NMR experiments were performed with 1 and 9 mg/ml solutions of GDDDD and its yttrium complex, respectively. The pH of both solutions was adjusted to 6.5 with NaOH. All pulse sequences were used as supplied with VNMRJ 2.3A. More specifically, the presaturation NOESY experiment was carried out using a 16 ppm sweep width, a 4 s acquisition time, a 142 ms mixing time, a 1 s relaxation delay, and a 3.0 saturation delay; 38400 complex data points were acquired in each transient and 256 transients were summed. Semi-selective TOCSY spectra were acquired using an excitation bandwidth of 862 Hz centered at 2.71 ppm (excitation band of aspartate β protons) using selective pulses generated automatically. The width of the acquisition window was set to 8.648 ppm and centered at 6.23 ppm. 256 increments were used in the evolution period and the acquisition time was set to 1.0 s. The acquired data was Fourier transformed with a linear prediction to 2048 points in the F1 domain and zero filling of 8192 points in the F2 domain. Semi-selective ROESY spectra were acquired using an excitation bandwidth of 1057.2 Hz, centered at 3.27 ppm (excitation band of aspartate α protons) using automatically generated selective pulses. The acquisition window width was 16 ppm, centered at 5.73 ppm. The evolution

period had 200 increments and the acquisition time was set to 1.0 s. The acquired data was Fourier transformed with a linear prediction to 2048 points in the F1 domain and zero filling of 8192 points in the F2 domain.

Quantum-chemical calculations

In the first step, the model systems were optimized by molecular mechanics treatment using the MM+ forcefield using HyperChem (HyperChem, 2002). In the next step, these singlet-spin state geometries of the systems under study were optimized using density functional theory (DFT) calculations employing the B3LYP hybrid functional (Becke, 1993), which is preferred for metal complex studies and NMR spectra interpretation. The LANL2DZ effective core potential and basis set were used for the central yttrium atom (Hay et al., 1985) and the standard 6-31G* basis sets from the Gaussian03 library (Frisch et al., 2003) were employed for the remaining atoms. The stability of the optimized structures was tested by vibrational analysis; no imaginary vibrations were found. For the evaluation of NMR chemical shifts, the Individual Gauges for Atoms in Molecules (IGAIM) method (Keith et al., 1992; Keith et al., 1993) and B3LYP hybrid functional were used, and the basis sets for the non-yttrium atoms were improved to 6-311++G** (Krishnan et al., 1980). The calculated ^{13}C and ^{15}N shifts were referenced to tetramethylsilane $\text{Si}(\text{CH}_3)_4$ and nitromethane CH_3NO_2 , respectively. All calculations were performed using Gaussian03 (Frisch et al., 2003).

Results and discussion

Design of Yttrium(III) complexes

As noted above, there are no structures of short, lanthanide-binding peptides in either the Cambridge Structural Database (CSD) or the Protein Data Bank (PDB). It was therefore necessary to design our peptides from first principles, though a few insights could be gleaned from the Y(III)-binding proteins in the PDB. First, since the Y(III) ion (medium-hard to hard Lewis acid) and a carboxylic acid oxygen atom (hard base) are expected to form a hard–hard acid–base pair (Pearson, 1963; Pearson, 1968a; Pearson, 1968b), we chose to construct our initial tags using predominantly aspartate and glutamate residues with glycine and alanine residues used respectively to extend the peptide N-terminus and as spacers. This choice is supported by the PDB, which overwhelmingly features aspartate and glutamate residues in its Y(III) binding sites. Next, we chose to give our initial tags a length of 4–6 residues; the primary loop of the

calcium-binding motif is 12 residues long, but six of these are spacing residues or are otherwise involved in the structural scaffolding of the overall protein while only four are responsible for the majority of the coordination bonds.

A series of peptides 4–6 residues in length, containing a mix of aspartate and glutamate amino acids with the addition of glycine spacers was modeled by molecular mechanics treatment using the MM+ forcefield (HyperChem, 2002). For each peptide, a Y(III)–peptide complex model was constructed. These models were then geometry optimized using Density Functional Theory (DFT) computations (for details see Materials and Methods). The six peptides which contained the greatest number of coordination bonds after geometry optimization were selected for actual experimental analysis.

The Y(III) binding constants of the peptides were determined by isothermal titration calorimetry (ITC, see below). We had initially planned to determine the structure of each complex by X-ray crystallography, but no diffracting crystals could be grown of any complex. We therefore chose to determine the structures of these complexes using DFT calculations supported by NMR. Because of the high computational demand of this approach, we explored only the GDDDD peptide, the shorter of the two complexes with the lowest dissociation constants (Fig. 1). As explained below, NMR alone could not be used because of the high degree of peak overlap in the spectra produced by these compounds.

Two assumptions were made in both the initial peptide design and in the NMR interpretation, which should be noted here. The first is that the central yttrium atom would likely have 7–9 coordinate geometry (yttrium like the lanthanides, typically has 8–12-coordinate geometry). The second is that aspartate, when coordinated through its side-chain,

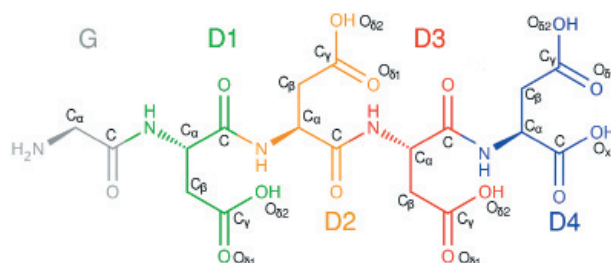


Fig. 1. A model of the GDDDD peptide in an extended conformation. The nomenclature shown here will be used to refer to individual atoms and residues in subsequent sections. G – glycine (gray), D1 – aspartate 1 (green), D2 – aspartate 2 (orange), D3 – aspartate 3 (red), D4 – aspartate 4 (violet).

would be only monodentate. The latter assumption is based on geometrical considerations. The total diameter of these peptide–Y(III) complexes is predicted to be less than 8 Å. Aspartate, which has only one methylene group between its carboxylic acid and the peptide backbone, lacks the flexibility to adopt a conformation which would allow it to form two coordination bonds in such a small complex. Glutamate, which has an extra methylene group between its carboxylate side-chain functional group and the peptide backbone, is not so constrained. These assumptions are consistent with the examples from the structural biology literature, many of which feature C=O···M(III) complexes and none of which feature bidentate aspartates in such tight complexes (Nitz et al., 2004 et al., Silvaggi et al., 2007 et al., Martin et al., 2007; Mustafi et al., 2006; Bentrop et al., 1997; Welch et al., 2003). Bidentate aspartates do occur, but only when the aspartate residue is anchored in a distal part of the protein chain and therefore has flexibility beyond what its side chain would ordinarily provide.

In modelling these complexes, we also considered the resulting charge of the entire complex. A final consideration is the possibility that the peptide amide nitrogen atom might become deprotonated, as observed in triglycine–Cu²⁺ complexes (Nakao et al., 1981), which could lead to coordination complex bonds between Y(III) and nitrogen, forming structures similar to those reported by Inoue et al. (Inoue et al., 1993).

Calorimetry

Initially, six different peptides were constructed containing different sequences of aspartate (D) and glutamate (E) residues. The abilities of these peptides to form Y(III) complexes was then tested using isothermal titration calorimetry (Table 1). All peptides had dissociation constants in the micromolar range; the GDDDD and EEDDEE peptides had the lowest dissociation constants, being nearly half the magnitude of the next best peptide. It can be seen that elongation from a 4-amino-acid peptide base improved the dissociation constant, but that further elongation either did not improve it or made it worse. Regardless, in all cases the yttrium–peptide stoichiometry was 1:1. Thus, at least 5, that is more than half, of the possible coordination sites of yttrium must have been occupied by the peptide.

The dissociation constants of these yttrium–peptide complexes are comparable to those of the Eu³⁺–P3W complex (K_d , 4.2 μM) derived from the calcium-binding EF-hand motif (Welch et al., 2003). Other modifications to the EF hand have led to nM dissociation constants for Tb³⁺ complexes (Martin et al., 2007). Thus our initial, unoptimized

Tab. 1. Association (K_a) and dissociation (K_d) constants, measured by ITC, for complexes of the indicated peptide with Y(III).

Peptide	K_a [M ⁻¹]	K_d [μM]
DDEE	$(1.68 \pm 0.13) \times 10^5$	5.95 ± 0.46
EEDDEE	$(7.24 \pm 0.46) \times 10^5$	1.38 ± 0.09
GDDDD	$(7.18 \pm 0.22) \times 10^5$	1.39 ± 0.04
GDDEE	$(2.58 \pm 0.15) \times 10^5$	3.88 ± 0.23
GDEED	$(3.65 \pm 0.21) \times 10^5$	2.74 ± 0.16
GDDEED	$(3.80 \pm 0.18) \times 10^5$	2.63 ± 0.13

trial peptides have K_d at least as good as (optimized) structures reported previously. The nM dissociation constants for the Tb³⁺ complexes suggest that K_d is improvable. A μM dissociation constant also suggests that the peptide does not coordinate the Y(III) atom to the maximum possible extent (8-coordinate).

NMR spectroscopy

The dissociation constants measured by ITC for the GDDDD and EEDDEE peptides are promising and may be sufficient for binding heavy atoms for solving the crystallographic phase problem. Nevertheless, further optimization is necessary for carrier-free labeling of radionuclides for radioimmunotherapy. The first logical step is to determine the structures of the two best complexes. Initially, crystallographic experiments were carried out to determine the structure of these yttrium–peptide complexes by X-ray analysis. Unfortunately, diffracting crystals were not produced by any of the many experiments carried out. Consequently, we resorted to the combination of density functional theory (DFT) computations and NMR spectroscopy to gain realistic structure of the complex. Because of the high computational demand of this approach, we explored only the shorter GDDDD peptide (Fig. 1)

One-dimensional ¹H and ¹³C spectra showed changes in chemical shifts of signals when comparing the free and complexed forms of the peptide (Table 2), suggesting a clear change in conformation. We also compared the ROESY spectrum of the Y(III)-complexed peptide (Fig. 2) and the NOESY spectrum of the free peptide (Fig. 3). Both contained very similar cross peaks and no new cross peaks were found in the spectrum of the complex. The most significant cross peaks between amide NH groups allowed us to assign them to specific amino acids (Table 2). The NH group in the first aspartic acid, which follows glycine in the chain (D1), also has a cross peak 8.6/3.9 ppm with the two Cα hydrogens of the glycine and the D4 NH group interacts with D4 Cα hydrogen represented

by the cross peak 7.9/4.4 ppm. These cross peaks are visible in both spectra. Also NOESY/ROESY provided the cross peaks between the amide NH groups and the Asp C β H₂ atoms, but the overlap between the C β H₂ groups was too large to allow clear associations to be made. Here a difference between the free ligand and complex can be seen, since in ¹H NMR free ligand spectrum the Asp C β H₂ hydrogens form two multiplets, each with 4 hydrogens,

and ¹H NMR complex spectrum these hydrogens are presented in one multiplet of 8 hydrogens. The cross peaks in NOESY consisted of doublets, while in the ROESY of the complex normal singlet signals were visible. The explanation of this feature is obvious – complexation with Y(III) simply changed the chemical shifts of the C β H₂ signals. Hindrance of conformational changes due to the bonds to the Y(III) ion may also play a role in this behavior. The

Tab. 2. Comparison of chemical shifts of signals in ¹H NMR spectra of the Y(III)-GDDDD complex and the free ligand GDDDD.

Proton label	Complex		Free ligand	
	δ (ppm)	signal multiplicity	δ (ppm)	signal multiplicity
G C α H ₂	3.78–3.84	m, 2H	3.87–3.95	m, 2H
D1 NH	8.56	d, 1H	8.86	d, 1H
D1 C α H	4.62–4.68	m, overlapped with H ₂ O	4.61–4.67	m, overlapped with H ₂ O
D1 C β H ₂	2.65–2.79	m, 8H	2.56–2.65 and 2.67–2.79	m, 4H and m, 4H
D2 NH	8.44	d, 1H	8.61	d, 1H
D2 C α H	4.62–4.68	m, overlapped with H ₂ O	4.61–4.67	m, overlapped with H ₂ O
D2 C β H ₂	2.65–2.79	m, 8H	2.56–2.65 and 2.67–2.79	m, 4H and m, 4H
D3 NH	8.17	d, 1H	8.16	d, 1H
D3 C α H	4.62–4.68	m, overlapped with H ₂ O	4.61–4.67	m, overlapped with H ₂ O
D3 C β H ₂	2.65–2.79	m, 8H	2.56–2.65 and 2.67–2.79	m, 4H and m, 4H
D4 NH	7.88	d, 1H	8.03	d, 1H
D4 C α H	4.42–4.46	m, 1H	4.41–4.45	m, 1H
D4 C β H ₂	2.65–2.79	m, 8H	2.56–2.65 and 2.67–2.79	m, 4H and m, 4H

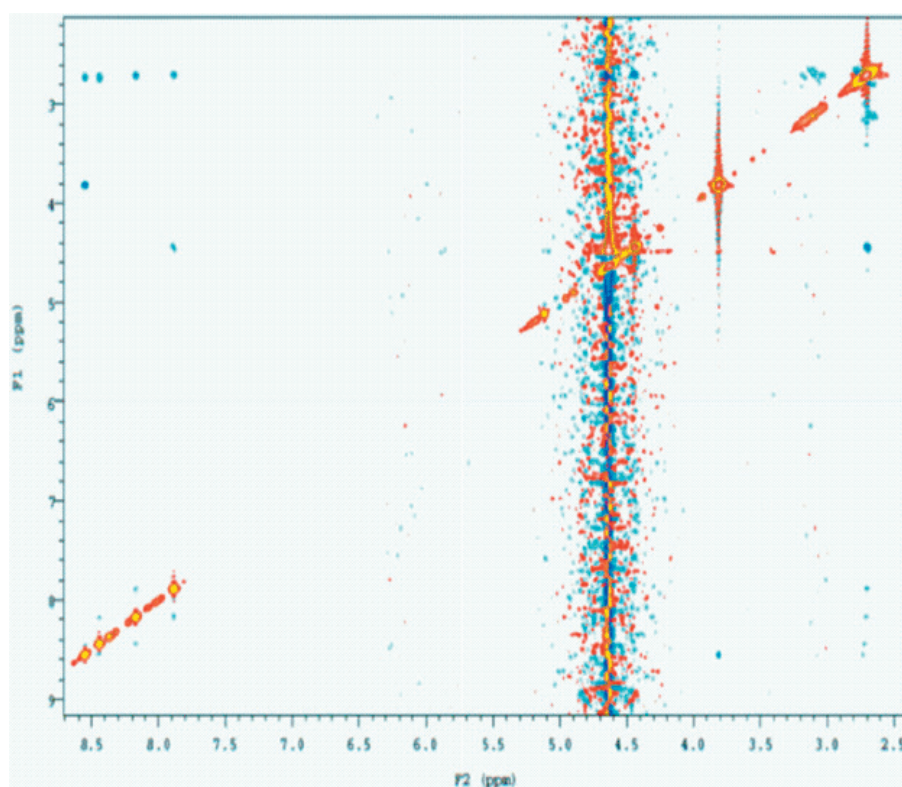


Fig. 2. ROESY spectrum of Y(III)-complexed GDDDD.

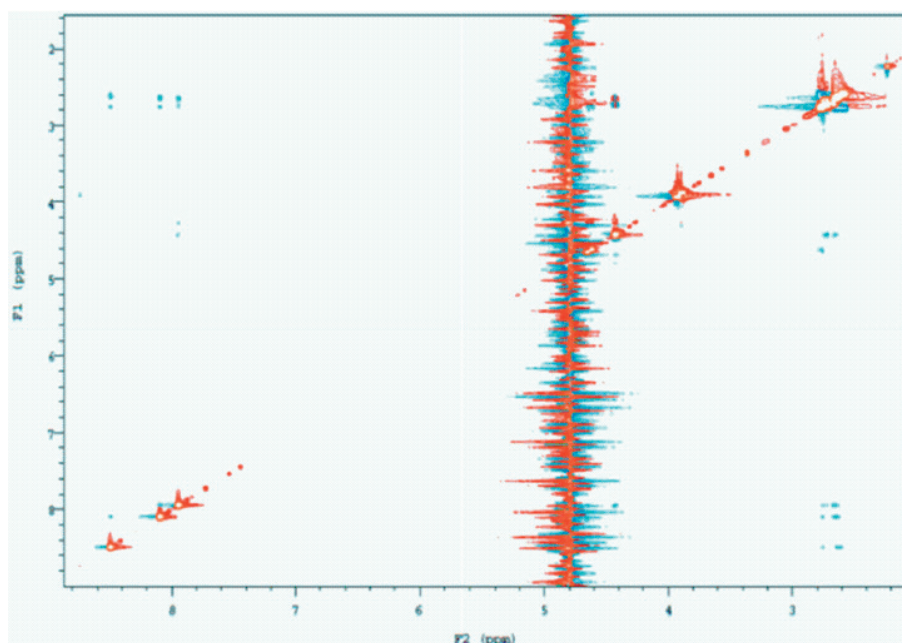


Fig. 3. NOESY spectrum of free GDDDD.

reason why two different types of through-space correlation spectra (ROESY for free ligand and NOESY for complex) were used, was that the NOESY spectrum of the Y(III)-complexed peptide did not provide any cross peaks at all (because of the re-

duced mobility of the complex). This confirms that in the presence of YCl_3 , the Y(III)-peptide complex, which has a higher molecular weight is formed.

The heteronuclear 2D NMR experiments HMBC (Heteronuclear multiple-bond correlation spectro-

Tab. 3. Comparison of ^{13}C NMR shifts (in ppm) with those calculated for various possible model geometries. Atom and residue notation is given in Fig. 1. A, B, and C are signals from D1, D2 and D3 amino-acids grouped by HMBC and HSQC. NP refers to the model with deprotonized amide nitrogens (Fig. 5), $O_{1,3}P$ belongs to the model with coordinated peptide bond oxygens (Fig. 6) and W is a water molecule. The predicted model chemical shifts were obtained from DFT calculations.

Experiment		Predicted model chemical shifts							
^{13}C NMR		[Y-NP] $^{5-}$		[Y- $O_{1,3}P$ -W] $^{2-}$		[Y-OP- W_2] $^{2-}$.W		[Y-OP-TFA-W] $^{3-}$.W	
180.3	D4C/ C_{γ}	186.2	D4 C	182.4	D1 C_{γ}	180.9	D1 C_{γ}	183.8	D4 C
179.9	B	181.9	D4 C_{γ}	178.9	G C	180.7	D1 C	178.0	D2 C_{γ}
178.9	C	179.3	D2 C	178.6	D2 C	178.4	D2 C	179.2	D1 C
178.5	A	179.0	D3 C	177.1	D4 C	177.0	D2 C_{γ}	177.3	D3 C
178.3	D4C/ C_{γ}	177.9	D2 C_{γ}	173.5	D2 C_{γ}	176.8	D4 C	177.3	G C
173.6	B	177.1	D3 C_{γ}	173.4	D1 C	175.7	D4 C_{γ}	171.3	D1 C_{γ}
173.5	C	174.6	D1 C	173.2	D4 C_{γ}	175.6	D3 C_{γ}	171.3	D4 C_{γ}
172.7	A	173.2	D1 C_{γ}	172.0	D3 C	173.5	G C	171.0	D2 C
167.0	G C	172.4	G C	169.4	D3 C_{γ}	168.9	D3 C	170.6	D3 C_{γ}
52.4	D4 C_{α}	69.1	D2 C_{α}	63.1	D3 C_{α}	60.3	D3 C_{α}	58.8	D2 C_{α}
51.7	C	65.9	D3 C_{α}	61.4	D1 C_{α}	52.5	D4 C_{α}	56.4	D4 C_{α}
51.4	B	63.7	D4 C_{α}	59.5	D4 C_{α}	52.1	D1 C_{α}	55.7	D3 C_{α}
51.0	A	63.3	D1 C_{α}	54.2	D2 C_{α}	51.6	G C_{α}	53.1	D1 C_{α}
40.8	G C_{α}	55.7	G C_{α}	47.5	G C_{α}	51.1	D2 C_{α}	45.6	G C_{α}
39.0	D4 C_{β}	44.3	D4 C_{β}	44.1	D3 C_{β}	45.5	D2 C_{β}	42.1	D4 C_{β}
38.2	A	43.6	D1 C_{β}	42.6	D2 C_{β}	42.3	D4 C_{β}	39.9	D3 C_{β}
38.1	B	41.8	D3 C_{β}	38.6	D1 C_{β}	38.6	D1 C_{β}	37.8	D1 C_{β}
38.0	C	40.5	D2 C_{β}	37.5	D4 C_{β}	34.0	D2 C_{β}	36.8	D2 C_{β}

Tab. 4. Comparison of ^{15}N NMR chemical shifts (in ppm) from the amide nitrogen atoms of the Y–GDDDD peptide complex with those calculated for various possible model geometries. NP refers to the model with deprotonated amide nitrogens (Fig. 5), $\text{O}_{1,3}\text{P}$ belongs to the model with coordinated peptide bond oxygens (Fig. 6) and W is a water molecule. The assignment of the signals was made by study of ROESY spectrum of the complex. The predicted model chemical shifts were obtained from DFT calculations.

Experiment ^{15}N NMR	Assignment	Predicted model chemical shift			
		$[\text{Y-NP}]^{5-}$	$[\text{Y-O}_{1,3}\text{P-W}]^{2-}$	$[\text{Y-OP-W}_2]^{2-}\cdot\text{W}$	$[\text{Y-OP-TFA-W}]^{3-}\cdot\text{W}$
-260.2	D1-NH	-236.6	-258.9	-269.3	-263.0
-259.7	D2-NH	-178.9	-255.6	-265.8	-260.0
-259.9	D3-NH	-178.7	-248.4	-264.1	-259.3
-255.8	D4-NH	-174.5	-237.1	-261.1	-254.8

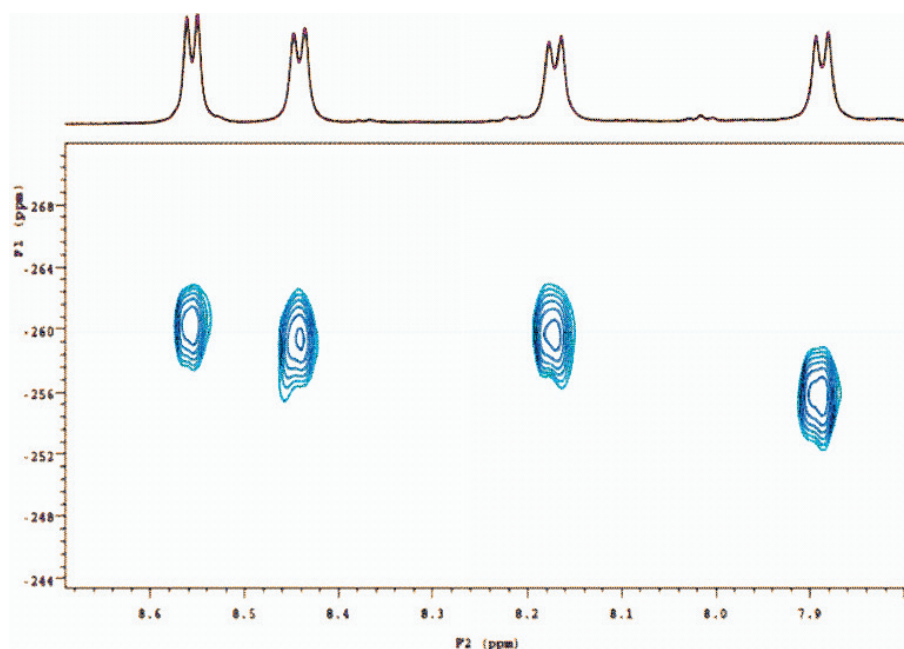


Fig. 4. Example ^{15}N HSQCAD spectrum of Y(III)-complexed GDDDD.

scopy) and HSQC (Heteronuclear single-quantum correlation spectroscopy) together with TOCSY (Total correlation spectroscopy) allowed us to unambiguously assign the CH_2 group of the N-terminal glycine residue and to assign the signals belonging to the aspartate residue at the C-terminal fourth position (D4). We were moreover able to group together signals belonging to atoms of amino acids D1, D2, and D3, but were not able to determine which signals came from which amino acid (Table 3). This suggests that the three middle aspartates D1, D2, and D3 must be in chemically very similar environments with similar conformations while the two terminal residues are different.

These NMR experiments also revealed the presence of trifluoroacetic acid (TFA) in the reaction mixtures (the characteristic chemical shifts of TFA may be seen in the ^{13}C spectra where they appear at 117 and 115 ppm in the spectra of both the complex and free

peptide). TFA was most likely introduced during the peptide synthesis where it is a common reagent. This TFA might easily influence the coordination sphere of Y(III) and alter the hydrogen-bonding system of the complex. As will be seen below, the best complex model did, in fact, have TFA as one of the Y(III) ligands.

We also performed ^{15}N HSQCAD NMR experiments (Fig. 4) because the four N atoms which are part of the peptide bonds joining the amino-acid residues together are exquisitely sensitive to the details of the coordination complex geometry. Because of the π delocalization of electrons across the C–N amide bond, the ^{15}N NMR chemical shifts of peptide bonds are typically quite sensitive to hydrogen bond or coordination bond formation and to shifts in torsion angles around this bond. We were able to assign the NH signals to specific amide groups, using a combination of ROESY/HSQC spectra. The NH group

from D1 can be safely assigned thanks to the cross-peak (8.56/3.8 ppm) from the ROESY spectrum. The signal of this NH group has a cross peak with the closest NH group from D2 (8.56/8.44 ppm). There is another cross-peak between the D2 NH group and the D3 NH group (8.44/8.17 ppm) and the last cross-peak in this region is between the D3 and D4 NH groups (8.17/7.88 ppm). The D4 NH group also has a cross-peak with the D4 α -CH group (4.44/7.88 ppm) which confirmed the D4 assignment. Using this knowledge, we could assign the signals also in the HSQC spectrum (Table 4). The glycine amine N atom was not visible (solvent exchange). Ultimately, these ^{15}N chemical shifts were the most useful for distinguishing amongst the many possible structures.

Modeling and DFT Optimization

Alone, these NMR measurements are not sufficient to determine the structure of the complexes. They can be used to sort amongst several candidate structures to select the most probable one, however. We constructed more than 20 different possible structures of the Y-GDDDD complex using DFT computations. The ^{13}C and ^{15}N NMR spectra of each structure were then calculated and compared to the experimental spectra (Tables 3 and 4). Several hypotheses of structural composition were explored. The three most important were involvement of the amide N atoms in yttrium coordination, coordination by both amide carbonyl oxygen and side-chain carboxylate oxygen atoms, and coordination by the side-chain carboxylate groups alone. The best structure for each of these hypotheses is described below along with the optimization of the most successful structure. To ensure that the models we produced were physically plausible, we also compared them to the Y(DTPA-benzylamine₂)(CH₃OH) complex described by Hsieh and Liu (Hsieh et al., 2004)

to exclude improper and incorrect models. (This particular structure was chosen because it features a central Y(III) atom coordinated by carboxylate oxygens, amide carbonyl oxygens, and amine nitrogen atoms.) Selected bond lengths and angles involving the central Y(III) ion can be seen in Tables 5 and 6. Our first hypothesis was that a complex similar to that observed for triglycine-Cu²⁺ complexes (Nakao et al., 1981) might be formed. In these complexes, one or more of the amide nitrogen atoms from one or more of the peptide bonds might become deprotonated and therefore participate in binding the Y(III) ion (the [Y-NP]⁵⁻ complex). An octadentate complex involving three coordination-complex bonds from peptide nitrogen atoms can be seen in Fig. 5. In this model, the deprotonated amide nitrogen atoms from the peptide bonds joining D1 with D2, D2 with D3, and D3 with D4 coordinate the central Y(III) along with the side-chain carboxylate groups of D2, D3, and D4, and the C-terminal carboxylate group attached to D4.

The NMR chemical shifts calculated from this structure were consistent with the ^{13}C NMR measurements, but not with the ^{15}N measurements, which were all predicted to fall between 25–80 ppm downfield of where they were observed. Thus the [Y-NP]⁵⁻ complex likely does not feature coordination bonds involving the peptide N atoms. Given that there are no common Y(III) redox reactions which would normally lead to amide nitrogen deprotonation, this result is not unexpected.

Next, complex structures involving the carbonyl oxygen atoms were explored. It was readily apparent early on that complexes involving all main-chain oxygens and all side-chain carbonyl oxygen atoms are sterically impossible. Consequently, we first focused on structures containing a mixture of side-chain and main-chain carbonyls. Since the amino-acid side chains have a higher degree of

Tab. 5. Selected bond lengths (in Å) calculated by DFT for the [Y-NP]⁵⁻, [Y-O_{1,3}P-W]²⁻, [Y-OP-W₂]²⁻.W, and [Y-OP-TFA-W]³⁻.W models. O_{δ1} is the side-chain carboxylate atom coordinated to yttrium. O_{αt} refers to the carboxylate group at the peptide C-terminus.

Bond	[Y-NP] ⁵⁻	[Y-O _{1,3} P-W] ²⁻	[Y-OP-W ₂] ²⁻ .W	[Y-OP-TFA-W] ³⁻ .W
Y—O _{δ1} (D1)	2.397	2.444	2.399	2.251
Y—O _{δ1} (D2)	2.393	-	2.281	2.375
Y—O _{δ1} (D3)	2.503	2.193	2.232	2.355
Y—O _{δ1} (D4)	2.804	2.207	2.306	2.250
Y—O _{αt} (D4)	2.418	2.375	2.398	2.421
Y—O(D1)	-	2.501	-	-
Y—O(D3)	-	2.349	-	-
Y—N(D2)	2.651	-	-	-
Y—N(D3)	2.360	-	-	-
Y—N(D4)	2.421	-	-	-

Tab. 6. Selected bond angles (in degrees) calculated by DFT for the $[Y-NP]^{5-}$, $[Y-O_{1,3}P-W]^{2-}$, $[Y-OP-W_2]^{2-}.W$, and $[Y-OP-TFA-W]^{3-}.W$ models. $O_{\delta 1}$ indicates the side chain carboxylate oxygen atom coordinated to yttrium, O_{xt} refers to the additional O atom which forms the C-terminus of the peptide at D4.

Angle	$[Y-NP]^{5-}$	$[Y-O_{1,3}P-W]^{2-}$	$[Y-OP-W_2]^{2-}.W$	$[Y-OP-TFA-W]^{3-}.W$
N(D2)—Y—N(D4)	117.2	-	-	-
N(D2)—Y—O $_{\delta 1}$ (D4)	82.9	-	-	-
N(D2)—Y—O $_{\delta 1}$ (D1)	77.1	-	-	-
N(D2)—Y—O $_{\delta 1}$ (D2)	70.7	-	-	-
N(D2)—Y—O $_{\delta 1}$ (D3)	127.2	-	-	-
N(D2)—Y—O $_{xt}$ (D4)	153.2	-	-	-
N(D2)—Y—N(D3)	66.6	-	-	-
N(D3)—Y—N(D4)	67.0	-	-	-
N(D3)—Y—O $_{\delta 1}$ (D4)	100.4	-	-	-
N(D3)—Y—O $_{\delta 1}$ (D1)	143.7	-	-	-
N(D3)—Y—O $_{\delta 1}$ (D2)	86.6	-	-	-
N(D3)—Y—O $_{\delta 1}$ (D3)	74.1	-	-	-
N(D3)—Y—O $_{xt}$ (D4)	131.5	-	-	-
N(D4)—Y—O $_{xt}$ (D4)	67.1	-	-	-
N(D4)—Y—O $_{\delta 1}$ (D1)	134.4	-	-	-
N(D4)—Y—O $_{\delta 1}$ (D2)	142.7	-	-	-
N(D4)—Y—O $_{\delta 1}$ (D3)	74.5	-	-	-
N(D4)—Y—O $_{\delta 1}$ (D4)	66.7	-	-	-
O $_{xt}$ (D4)—Y—O $_{\delta 1}$ (D1)	82.4	127.5	142.8	77.9
O $_{xt}$ (D4)—Y—O $_{\delta 1}$ (D2)	123.5	-	72.1	70.7
O $_{xt}$ (D4)—Y—O $_{\delta 1}$ (D3)	79.6	126.8	109.4	85.0
O $_{xt}$ (D4)—Y—O $_{\delta 1}$ (D4)	74.7	77.1	71.7	76.8
O $_{\delta 1}$ (D1)—Y—O $_{\delta 1}$ (D2)	82.3	-	73.1	87.9
O $_{\delta 1}$ (D1)—Y—O $_{\delta 1}$ (D3)	133.7	86.8	84.2	158.8
O $_{\delta 1}$ (D1)—Y—O $_{\delta 1}$ (D4)	73.3	149.8	145.5	107.3
O $_{\delta 1}$ (D2)—Y—O $_{\delta 1}$ (D3)	73.2	-	91.9	74.7
O $_{\delta 1}$ (D2)—Y—O $_{\delta 1}$ (D4)	147.4	-	137.3	140.2
O $_{\delta 1}$ (D3)—Y—O $_{\delta 1}$ (D4)	139.4	90.1	79.6	80.3
O(D1)—Y—O $_{\delta 1}$ (D1)	-	68.0	-	-
O(D1)—Y—O(D3)	-	91.2	-	-
O(D1)—Y—O $_{\delta 1}$ (D3)	-	154.4	-	-
O(D1)—Y—O $_{\delta 1}$ (D4)	-	113.7	-	-
O(D1)—Y—O(D4)	-	70.5	-	-
O(D3)—Y—O $_{\delta 1}$ (D1)	-	76.7	-	-
O(D3)—Y—O $_{\delta 1}$ (D3)	-	78.4	-	-
O(D3)—Y—O $_{\delta 1}$ (D4)	-	131.9	-	-
O(D3)—Y—O(D4)	-	73.5	-	-

flexibility than the peptide backbone and since the side-chain carboxylic acid groups are more nucleophilic than the amide carbonyl oxygen atoms, the complex geometries favoring side-chain—Y(III) interactions were favored over those with more main-chain—Y(III) interactions. The best of these is the $[Y-O_{1,3}P]^{2-}$ complex shown in Fig. 6, in which the central Y(III) is bound by four side-chain carboxylate groups and two amide carbonyl oxygen atoms.

In this model, the carbonyl oxygen atoms from the G (residue 1) and D2 (residue 3) residues coordinate the Y(III) atom along with the side-chain carboxylate groups of D2, D3, and D4, and the C-terminal carboxylate group. A water molecule occupies a seventh coordination position. During geometry optimization, a coordination bond between Y(III) and the D2 aspartate side-chain atom O $_{\delta 1}$ became disconnected and, consequently, cannot be seen in

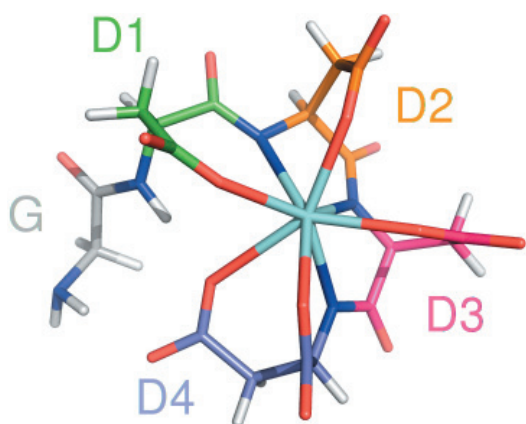


Fig. 5. The $[Y-NP]^{5-}$ model. In this model, deprotonated amide nitrogen atoms from the peptide bonds joining D1, D2, D3, and D4 coordinate the central Y(III) along with the side-chain carboxylate groups of D2, D3, and D4, and the C-terminal carboxylate group attached to D4. Following standard conventions from the structural biology literature, all oxygens are red and all nitrogens are blue.

the model shown in Fig. 6. This disconnection and the poor agreement with the ^{15}N measurements suggest that this complex is also not likely to be the correct one.

The models with the best agreement with the NMR measurements had the side-chain carboxyl groups coordinated monodentately to Y(III) with no amide oxygen atoms involved. Typically, the coordination sphere was completed by one or more water molecules. One of the best of these structures is the $[Y-OP-W_2]^{2-} \cdot W$ complex shown in Fig. 7, which features Y(III) coordinated to five side-chain carbonyl groups and two water molecules (originally three water molecules, but one was excluded from the yttrium coordination sphere during geometry optimization). In this model, all side-chain carboxylate groups, along with the C-terminal carboxylate group, form coordination bonds to the central Y(III) atom. Two additional water molecules give a total coordination sphere of seven. A third water molecule was excluded from the coordination sphere during geometry optimization but remained associated with the complex through hydrogen-bonds to the main-chain carbonyl O atoms of D3 and D4.

Given the relative success of the $[Y-OP-W_2]^{2-} \cdot W$ complex in predicting the NMR data, we attempted to optimize this complex to produce the best possible match. In doing this optimization, we primarily adjusted the torsion angles surrounding the amide peptide bonds (the so-called φ , ψ , and ω angles) (Branden et al., 1999). One of the most important

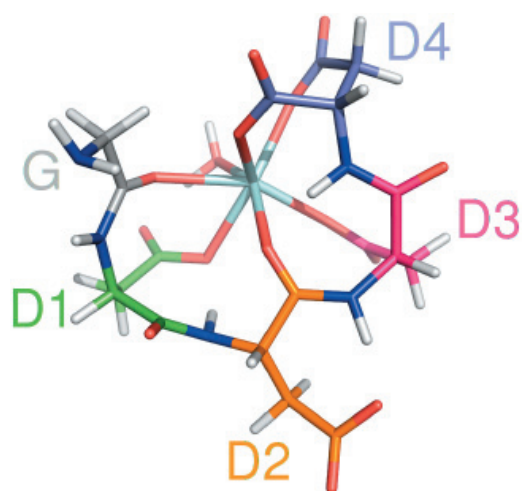


Fig. 6. The $[Y-O_{1,3}P]^{2-}$ model. In this model, the carbonyl oxygen atoms from the G and D2 amide residues coordinate the Y(III) atom along with the side-chain carboxylate groups of D1, D3, and D4, and the C-terminal carboxylate group. A water molecule completes the coordination sphere of Y(III). An eighth coordination bond, between D2 $O_{\delta 1}$ and the Y(III) atom, was unstable during DFT geometry optimization.

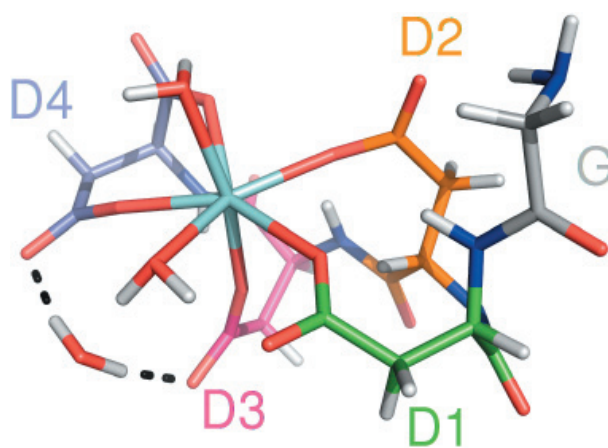


Fig. 7. The $[Y-OP-W_2]^{2-} \cdot W$ model. In this model, all side-chain carboxylate groups, along with the C-terminal carboxylate group, form coordination bonds to the central Y(III) atom. Two additional water molecules give a total coordination sphere of seven. A third water molecule was excluded from the coordination sphere during DFT geometry optimization but remained associated with the complex through hydrogen-bonds to the main-chain carbonyl O atoms.

changes was to correct the ω angle of the bond joining D1 and D2 ($O(D1)-C(D1)-N(D2)-H$) from 8.6° , almost never observed in peptides, to a more reasonable 171.1° (ω angles are normally $180 \pm 10^\circ$). A water molecule was added to the complex con-

Tab. 7. Selected ω torsion angles (in degrees) calculated by DFT for the $[\text{Y-NP}]^{5-}$, $[\text{Y-O}_{1,3}\text{P-W}]^{2-}$, $[\text{Y-OP-W}_2]^{2-}$.W, and $[\text{Y-OP-TFA-W}]^{3-}$.W models.

Angle	$[\text{Y-O}_{1,3}\text{P-W}]^{2-}$	$[\text{Y-OP-W}_2]^{2-}$.W	$[\text{Y-OP-TFA-W}]^{3-}$.W
O(G)—C(G)—N(D1)—H	173.0	169.2	-176.4
O(D1)—C(D1)—N(D2)—H	163.0	8.6	171.1
O(D2)—C(D2)—N(D2)—H	172.7	173.9	175.1
O(D3)—C(D3)—N(D4)—H	171.7	177.8	-176.2

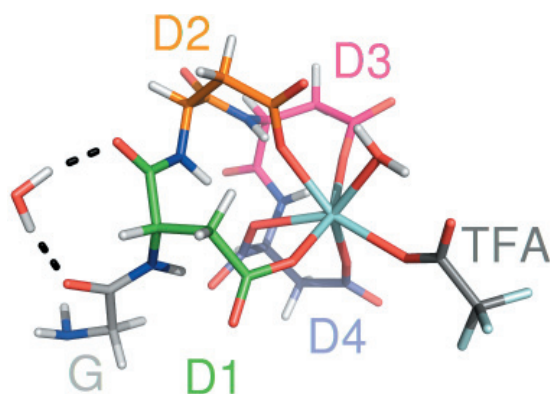


Fig. 8. The $[\text{Y-OP-TFA-W}]^{3-}$.W model. This model is similar to the $[\text{Y-OP-W}_2]^{2-}$.W model shown in Fig. 7, but with TFA replacing one of the water molecules, while the other water now bridges the main-chain carbonyl O atoms of G and D1. The most important conformational difference between this model and $[\text{Y-OP-W}_2]^{2-}$.W is the approximately 180° rotation about the peptide bond joining D1 and D2.

necting the amide oxygen atoms of G and D1. This caused the ψ torsion angle of G (N—CH₂—C—O) to shift from 89.5° to 33.0° , improving the stability of the structure (Table 7). Finally, we added a trifluoroacetic acid (TFA) molecule to the complex since one was detected by NMR and confirmed by elemental analysis to give the complex $[\text{Y-OP-TFA-W}]^{3-}$.W (Fig. 8). The initial complex had an 8-coordinate Y(III) with a pentadentate peptide, monodentate TFA, and

two water molecules. During geometry optimization, one of these was pushed out of the coordination sphere, giving a 7-coordinate Y(III) in the final complex. This complex had the best agreement with the ¹⁵N data. Moreover, the dissociation constant and stoichiometry given by ITC is also consistent with a pentadentate peptide ligand.

The ¹³C chemical shifts of carboxylate groups can be greatly influenced by hydrogen bonding. In $[\text{Y-OP-W}_2]^{2-}$.W, the D1 carboxylate group is engaged in two hydrogen bonds, one to its own amide nitrogen backbone atom and one to a water molecule (Table 8). In $[\text{Y-OP-TFA-W}]^{3-}$.W, it is the D4 carboxylate which has two hydrogen bonds, on the other hand, this time to the amide nitrogen atoms of D1 and D2 (Table 9). For both complexes, the Y(III)—O₅₁ distances of the respective carboxylates are similar (Table 5). These hydrogen bonds are responsible for giving these two carboxylate groups chemical shifts of 180.9 ($[\text{Y-OP-W}_2]^{2-}$.W) and 183.8 ppm ($[\text{Y-OP-TFA-W}]^{3-}$.W) in contrast to a carboxylate group bound only to Y(III) (e.g C_v in $[\text{Y-OP-TFA-W}]^{3-}$.W with a chemical shift of 170.6 ppm). Because it was not possible to clearly separate the chemical shifts of the individual aspartate residues, it is possible that the 180.3 ppm signal observed in the ¹³C NMR spectra could belong to one of the other carboxylate groups, but the downfield shift of this peak relative to that predicted for a yttrium-only associated carboxylate strongly suggests that one of the carboxylate groups of the peptide does have two hydrogen bonds.

Tab. 8. Selected hydrogen bond lengths (in Å) and angles (in degrees) calculated by DFT for $[\text{Y-OP-W}_2]^{2-}$.W with A and D being the hydrogen acceptor and donor. O₆₁ is used for the side chain carboxylate atom bound to yttrium, O₆₂ for the side chain carboxylate atom not bound to yttrium, W1, W2, W3 refer to water molecules.

D—H...A	D—H (Å)	H...A (Å)	D...A (Å)	D—H...A (°)
N(G)—H...O ₆₂ (D2)	1.025	1.976	2.960	159.8
N(D1)—H...O ₆₁ (D1)	1.016	2.084	2.834	128.9
O(W2)—H1...O ₆₂ (D1)	0.998	1.720	2.690	162.8
N(D3)—H...O ₆₁ (D2)	1.021	1.959	2.917	155.0
O(W3)—H1...O ₆₂ (D3)	0.983	2.008	2.981	170.2
N(D4)—H...O ₆₁ (D3)	1.011	1.938	2.885	154.8
O(W3)—H2...O ₆₂ (D4)	0.986	1.924	2.898	168.6

Tab. 9. Selected hydrogen bond lengths (in Å) and angles (in degrees) calculated by DFT for [Y-OP-TFA-W]³⁻.W. O_{δ1} is used for the side chain carboxylate atom bound to yttrium, O_{δ2} for the side chain carboxylate atom not bound to yttrium, O_{xt} is the C-terminal carboxylate atom of D4. W_p refers to the water molecule bound to the peptide, and W_y refers to the water molecule bound to the central yttrium atom.

D—H...A	D—H (Å)	H...A (Å)	D...A (Å)	D—H...A (°)
O(WP)—H1...O(G)	0.979	1.962	2.932	170.1
N(D1)—H...O2(D4)	1.029	1.990	2.964	157.0
O(WP)—H2...O(D1)	0.980	1.978	2.934	164.6
N(D2)—H...O _{xt} (D4)	1.024	1.969	2.991	176.1
N(D3)—H...O _{δ1} (D2)	1.020	1.906	2.826	148.4
O(WY)—H1...O _{δ1} (D2)	0.989	1.823	2.774	160.1
N(D4)—H...O _{δ1} (D3)	1.023	1.827	2.824	162.8

Finally, we compared our predicted Y(III)—peptide complex models with the X-ray crystal structure of the Y(DTPA-Benzylamine₂)(CH₃OH) complex described by Hsieh and Liu (Hsieh et al., 2004). This complex features both Y—O and Y—N coordination bonds. This comparison allowed us to verify that our calculated bond lengths and angles, especially to the central metal atom, were reasonable. In the Y(DTPA-Benzylamine₂)(CH₃OH) structure, the Y—O distances were of the order of 2.36 Å, and the Y—N distances were in average of 2.68 Å. In our postulated structures, the Y—O and Y—N distances are all in the range of 2.2–2.8 Å. The smallest angle centering on the Y(III) ion in the experimental structure was 62.02° (O(9)—Y—N(11)), while the smallest angle in any of our structures was the N(D2)—Y—N(D3) angle in the [Y-NP]⁵⁻ complex of 66.6° (Table 6). Given that this is the least probable of the complex forms reported here, we may conclude that our calculated geometries are physically plausible.

Conclusions

In this study we designed new yttrium(III)-binding tags consisting of 4–6 natural amino acids (aspartate, glutamate, and glycine). These tags are relatively small, lack complicated tertiary structures, and bind Y(III) with a dissociation constant in the micromolar range. To determine the structure of the best of these, the GDDDD peptide, we measured its ¹H, ¹³C, and ¹⁵N NMR spectra and used these spectra as reference to choose several plausible geometries. The geometry of each suggested model was optimized by quantum chemical calculations and its calculated NMR spectra were compared with those measured. The [Y-OP-TFA-W]³⁻.W complex (Fig. 8) was found to have best correspondence with the data. This complex featured a pentadentate peptide and a 7-coordinate Y(III). This complex also possessed a

trifluoroacetic acid molecule, which both NMR and elemental analysis suggested was present.

Presently, none of the tags has a dissociation constant low enough to allow them to be used for radioimmuno-therapy, but in their present form, they may still be suitable for anomalous phasing applications in protein crystallography as well as for improvement of protein crystal quality. Yttrium or other metal ions may be essential for obtaining well-diffracting protein crystals, for example Ni²⁺ in Imidazole glycerol phosphate synthase (Chaudhuri et al., 2001) and Ca²⁺ in the dimer structure of proliferating cell nuclear antigen (Xing et al., 2007; Hlinkova et al., 2008). In this context, it is worth noting that these tags are also usable for coordinating other transition metals and lanthanides: e.g. the GDDDD peptide has roughly 10× higher affinity for Cu(II) than Y(III) (Mudra, unpublished results). The geometry inferred in this study can be used as the basis for further improvements in these tags, with the most immediate goal being to increase their coordination occupancy, thereby lowering the dissociation constants.

Acknowledgements

We would like to thank the HPC center at Slovak University of Technology in Bratislava, part of the Slovak Infrastructure of High Performance Computing (SIVVP project No. 26230120002, funded by European Regional Development Funds), for computing facilities. The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at the Jülich Supercomputing Centre (JSC). NMR measurements were made possible by Slovak Program Project No. 2003SP200280203. We thank Branislav Horváth for performing NMR experiments and Dr. Jozef Ševčík for his help and support. The ITC experiments were performed by Milan Kožíšek, IMG ASCR, Czech Republic. This study was supported by EUREKA E! 3537 ENGRAB, Scientific Grant Agency VEGA (contracts No. 1/0598/16 and

1/0500/15) and the Research and Development Agency of the Slovak Republic APVV (contract No. APVV-15-0079).

Appendix. Monodentate Aspartate

For any carboxylate group to form two coordination bonds to a metal ion, the carbon atom bridging the two carboxylate oxygen atoms must be a distance of

$$R = d \cos(\theta/2) + R_0 \quad (\text{A1})$$

away from the metal ion, where d is the length of the $\text{C}=\text{O}\cdots\text{M}$ bond, θ is the $\text{O}\cdots\text{M}\cdots\text{O}$ bond angle, and $R_0 = 0.60 \text{ \AA}$ is the distance from the gridding carbon to the plane containing the two oxygen atoms. For a Y(III) –carboxylate bidentate interaction, $d = 2.4 \text{ \AA}$ and $\theta = 60^\circ$ are the most common values for bidentate carboxylate groups in the PDB. This gives a $\text{C}\gamma\cdots\text{Y(III)}$ distance of 2.7 \AA . In a bidentate aspartate, the $\text{Y(III)}\cdots\text{C}\gamma\cdots\text{C}\beta$ angle is effectively fixed at 180° , meaning that the distance from the Y(III) ion to the first rotatable bond is about 4.2 \AA ($2.7 \text{ \AA} + 1.5 \text{ \AA}$, which is the length of a $\text{C}-\text{C}$ single bond). Since the maximum diameter of these complexes is 8 \AA , the Y(III) ion will be displaced from the center of the complex, likely resulting in either the loss of coordination bonds or in the disruption of the overall complex. Aspartate has only one rotatable bond outside the peptide backbone, so it cannot change its conformation to compensate for this shift. Glutamate has more than one rotatable bond outside of the backbone, so it should be able to adopt bidentate coordination without causing strain to the peptide backbone.

To verify that this assumption was reasonable, we examined every structure in the Protein Data Bank which had a bound Y(III) ion (48 as of September 2014). In nearly all cases, aspartate and glutamate side-chains were involved in the Y(III) -binding site. Most often, these residues were only mono-dentate rather than bidentate. Bidentate glutamates and aspartates were found, though glutamate was more likely to be bidentate. Aspartate was never bidentate when it occurred less than five residues away from another coordinating residue, however. Glutamate did not have this restriction, though it was still unusual for it to be bidentate in this situation. Thus, in the context of this small complex, less than 8 \AA across, the assumption of a monodentate aspartate side chain seems reasonable.

References

Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 98: 5648–5652.

Bentrop D, Bertini I, Cremonini MA, Forsen S, Luchinat C, Malmendal A (1997) Solution structure of the paramagnetic complex of the N-terminal domain of calmodulin with two Ce^{3+} ions by ^1H NMR. *Biochemistry* 36: 11605–11618.

Branden C, Tooze J (1999) Introduction to Protein Structure, 2nd ed. Garland Publishing, New York.

Chaudhuri BN, Lange SC, Myers RS, Chittur SV, Davisson VJ, Smith JL (2001) Crystal Structure of Imidazole Glycerol Phosphate Synthase: A Tunnel through a $(\beta/\alpha)_8$ Barrel Joins Two Active Sites. *Structure* 9: 987–997.

Frisch MJT, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr., Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2003) Gaussian 03, Revision C.1. Gaussian, Inc., Pittsburgh (PA).

Goodwin DA, Meares CF (2001) Advances in pretargeting biotechnology. *Biotechnol Adv* 19: 435–450.

Hay PJ, Wadt WR (1985) Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J Chem Phys* 82: 299–310.

Hlinkova V, Xing GX, Bauer J, Shin YJ, Dionne I, Rajashankar KR, Bell SD, Ling H (2008) Structures of monomeric, dimeric and trimeric PCNA: PCNA-ring assembly and opening. *Acta Crystallogr D Biol Crystallogr* 64: 941–949.

Hsieh WY, Liu S (2004) Synthesis, Characterization, and Structures of Indium $\text{In}(\text{DTPA-BA}_2)$ and Yttrium $\text{Y}(\text{DTPA-BA}_2)(\text{CH}_3\text{OH})$ Complexes ($\text{BA} = \text{Benzylamine}$): Models for ^{111}In - and ^{90}Y -Labeled DTPA-Biomolecule Conjugates. *Inorg Chem* 43: 6006–6014.

Hyperchem. Release 7.5 for Windows. (2002) HyperCube, Inc. Gainesville, FL.

Inoue MB, Inoue M, Munoz IC, Bruck MA, Fernando Q (1993) Syntheses of new 15-membered and 16-membered macrocyclic ligands with three pendant acetato groups and the structures of the gadolinium(III) complexes. *Inorg Chim Acta* 209: 29–34.

Keith TA, Bader RFW (1992) Calculation of magnetic response properties using atoms in molecules. *Chem Phys Lett* 194: 1–8.

Keith TA, Bader RFW (1993) Calculation of Magnetic Response Properties Using a Continuous Set of Gauge Transformations. *Chem Phys Lett* 210: 223–231.

Krishnan R, Binkley JS, Seeger R, Pople JA (1980) Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J Chem Phys* 72: 650.

- Liu S, Edwards DS (2001) Bifunctional Chelators for Therapeutic Lanthanide Radiopharmaceuticals. *Bioconjugate Chem* 12: 7–34.
- Martin LJ, Hahnke MJ, Nitz M, Wöhnert J, Silvaggi NR, Allen KN, Schwalbe H, Imperiali B (2007) Double-lanthanide-binding tags: Design, photophysical properties, and NMR applications. *J Am Chem Soc* 129: 7106–7113.
- Mather SJ (2007) Design of radiolabelled ligands for the imaging and treatment of cancer. *Mol Biosyst* 3: 30–35.
- Mustafi SM, Mukherjee S, Chary KVR, Cavallaro G (2006) Structural basis for the observed differential magnetic anisotropic tensorial values in calcium binding proteins. *Proteins* 65: 656–669.
- Nakao YM, Mori W, Sakurai T, Nakahara A (1981) The monomeric and dimeric copper(II) complexes containing imidazole and dipeptides. *Inorg Chim Acta* 55: 103–107.
- Nitz M, Sherawat M, Franz KJ, Peisach E, Allen KN, Imperiali B (2004) Structural Origin of the High Affinity of a Chemically Evolved Lanthanide-Binding Peptide. *Angew Chem Int Ed* 43: 3682–3685.
- Pearson RG (1963) Hard and Soft Acids and Bases. *J Am Chem Soc* 85: 3533–3539.
- Pearson RG (1968a) Hard and soft acids and bases, HSAB, part I: Fundamental principles. *J Chem Ed* 45: 581–586.
- Pearson RG (1968b) Hard and soft acids and bases, HSAB, part II: Underlying theories. *J Chem Ed* 45: 643–647.
- Silvaggi NR, Martin LJ, Schwalbe H, Imperiali B, Allen KN (2007) Double-Lanthanide-Binding Tags for Macromolecular Crystallographic Structure Determination. *J Am Chem Soc* 129: 7114–7120.
- Welch JT, Kearney WR, Franklin SJ (2003) Lanthanide-binding helix-turn-helix peptides: Solution structure of a designed metallonuclease. *Proc Natl Acad Sci USA* 100: 3725–3730.
- Xing GX, Hlinkova V, Ling H (2007) Purification, Crystallization and Preliminary Diffraction Studies of the *Sulfolobus solfataricus* PCNA Proteins in Different Oligomeric Forms. *Cryst Growth Des* 7: 2202–2205.
- Xue S, Qiao J, Jiang J, Hubbard K, White N, Wei L, Li S, Liu ZR, Yang JJ (2014) Design of ProCAs (protein-based Gd(3+) MRI contrast agents) with high dose efficiency and capability for molecular imaging of cancer biomarkers. *Med Res Rev* 34: 1070–1099.
- Yang W, Wilkins AL, Ye Y, Liu ZR, Li SY, Urbauer JL, Hellinga HW, Kearney A, van der Merwe PA, Yang JJ (2005) Design of a calcium-binding protein with desired structure in a cell adhesion molecule. *J Am Chem Soc* 127: 2085–2093.
- Yang, JJ, Yang J, Wei L, Zurkiya O, Yang W, Li S, Zou J, Zhou Y, Maniccia ALW, Mao H, Zhao F, Malchow R, Zhao S, Johnson J, Hu X, Krogstad E, Liu ZR (2008) Rational Design of Protein-Based MRI Contrast Agents. *J Am Chem Soc* 130: 9260–9267.