

**DETERMINANTS OF CIS-TRANS ISOMERISM OF  
THE AROMATIC-PROLYL AMIDE BOND  
AND  
DESIGN OF LANTHANIDE-BINDING PEPTIDES**

by

Hai Yun Meng

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Chemistry and Biochemistry

Winter 2006

Copyright 2006 Hai Yun Meng  
All Rights Reserved

UMI Number: 1432291



---

UMI Microform 1432291

Copyright 2006 by ProQuest Information and Learning Company.  
All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

**DETERMINANTS OF CIS-TRANS ISOMERISM OF  
THE AROMATIC-PROLYL AMIDE BOND  
AND  
DESIGN OF LANTHANIDE-BINDING PEPTIDES**

by

Hai Yun Meng

Approved: \_\_\_\_\_

Neal J. Zondlo, Ph. D.  
Professor in charge of thesis on behalf of the Advisory Committee

Approved: \_\_\_\_\_

Charles G. Riordan, Ph. D.  
Chair of the Department of Chemistry and Biochemistry

Approved: \_\_\_\_\_

Thomas M. Apple, Ph. D.  
Dean of the College of Arts and Sciences

Approved: \_\_\_\_\_

Conrado M. Gempesaw II, Ph. D.  
Vice Provost for Academic and International Programs

## ACKNOWLEDGMENTS

First and foremost, I thank my research advisor, Neal Zondlo, for his continuous guidance, encouragement and academic support necessary for the completion of this research.

I also must thank my friends, Rong Yang, Matthew Lee, and Silvio Le Rose for their emotional support for all the time I spent in the graduate school.

My labmates, also known as my friends, especially Shalini and Devan, have been helping me since day one I started my research in the lab. I want to thank them for all of the training, sharing and amusement.

Chris and Mao did some earlier valuable work on lanthanide finger peptide synthesis and analysis. I would not have been able to complete chapter 2 without them.

Finally, my family has always been wonderful and supportive. I want to thank them for the unconditional love, support and trust from my great parents, brothers and sisters.

## TABLE OF CONTENTS

LIST OF TABLES .....	vi
LIST OF FIGURES .....	vii
ABSTRACT .....	ix

### Chapter

1	EFFECTS OF $i$ AND $i+3$ RESIDUE IDENTITIES ON CIS-TRANS ISOMERISM OF THE AROMATIC <sub><math>i+1</math></sub> -PROLYL <sub><math>i+2</math></sub> AMIDE BOND: IMPLICATIONS FOR TYPE VI $\beta$ -TURN FORMATION .....	1
	Introduction .....	1
	Experimental .....	4
	Peptide synthesis, purification and characterization .....	4
	NMR analysis .....	7
	Results and Discussion .....	9
	Determination of a general sequence XYPZ to promote cis amide bonds .....	9
	Determination of the effects of residue $i$ on cis-trans isomerism in the sequence XYPN .....	10
	Determination of the effects of residue $i+3$ on cis-trans isomerism in the sequence TYPZ .....	16
	Effect of multiple proline residues on cis-trans isomerism .....	19
	Conclusions .....	20
	References .....	32
2	DESIGN OF A LANTHANIDE FINGER PEPTIDE .....	39
	Introduction .....	39
	Experimental .....	41
	Materials .....	41
	Peptide synthesis, purification and characterization .....	41
	Circular dichroism spectroscopy .....	42
	Calculation of dissociation constants for peptide-metal complexes .....	43
	NMR spectroscopy .....	44
	Results and Discussion .....	45
	Conclusions .....	51
	References .....	88

3	DESIGN OF A KINASE-RESPONSIVE AND PHOSPHORYLATION-DEPENDENT PROTEIN DOMAIN .....	90
	Introduction .....	90
	Experimental.....	94
	Materials .....	94
	Peptide synthesis, purification and characterization .....	95
	Fluorescence experiments.....	97
	Dissociation constant determination.....	98
	Phosphorylation of HM series peptides by Protein Kinase A .....	98
	Results and Discussion .....	99
	HM1 analysis.....	99
	HM2 analysis.....	101
	HM3 analysis.....	102
	HM4 analysis.....	103
	HM5 analysis.....	104
	HM6 analysis.....	104
	HM7 analysis.....	105
	HM8 analysis.....	106
	HM9 analysis.....	107
	Conclusions .....	108
	References.....	123

## LIST OF TABLES

Table 1.1	Ideal characteristics of type VI $\beta$ -turn subtypes.....	22
Table 1.2	NMR-derived data for the peptides XYPZ.....	23
Table 1.3	NMR-derived data for the peptides XYPN.....	24
Table 1.4	Selected NMR data for the peptides XYPN.....	25
Table 1.5	NMR-derived data for the peptides TYPZ.....	26
Table 1.6	Selected NMR data for the peptides TYPZ.....	27
Table 2.1	Dissociation constants for peptide-metal complexes.....	87
Table 3.1	Characterization data for non-phosphorylated and phosphorylated peptides.....	110
Table 3.2	Dissociation constants of peptide-Tb <sup>3+</sup> complexes in aqueous solutions.....	111

## LIST OF FIGURES

Figure 1.1	Aromatic-backbone interactions stabilizing cis or trans amide bonds..	28
Figure 1.2	Amide region of the NMR spectra of the peptides TYPN, GYPN and FYPN.....	29
Figure 1.3	Amide region of the NMR spectra of the peptides PYPN, pivaloyl-PYPN, TYPP, and PYPP.....	30
Figure 1.4	Van't Hoff plot of the temperature dependence of $K_{trans/cis}$ for the peptides AYPN and YYPN .....	31
Figure 2.1	Structural schematic of a zinc finger .....	53
Figure 2.2	Iterative design of lanthanide fingers .....	54
Figure 2.3	CD spectra of lanthanide fingers .....	55
Figure 2.4	Metal titrations of lanthanide-binding peptides .....	56
Figure 2.5	Comparative NMR spectra of LF4 in the absence and presence of $Lu^{3+}$ .....	57
Figure 2.6	Visible fluorescence spectra of 10 $\mu$ M LF6 in the absence and presence of 90 $\mu$ M $Tb^{3+}$ .....	58
Figure 2.7	Full titration of LF1- $Eu^{3+}$ .....	59
Figure 2.8	Full titration of LF1- $La^{3+}$ .....	60
Figure 2.9	Metal titrations of $Yb^{3+}$ - LF1 binding .....	61
Figure 2.10	Metal titrations of $Yb^{3+}$ - LF2 binding .....	62
Figure 2.11	Full titration of LF2- $Eu^{3+}$ .....	63
Figure 2.12	Full titration of LF2- $La^{3+}$ .....	64
Figure 2.13	Metal titrations of $Yb^{3+}$ - LF3 binding .....	65
Figure 2.14	Full titration of LF3- $Eu^{3+}$ .....	66
Figure 2.15	Full titration of LF3- $La^{3+}$ .....	67
Figure 2.16	Full titration of LF4- $Er^{3+}$ .....	68
Figure 2.17	Full titration of LF4- $Ho^{3+}$ .....	69
Figure 2.18	Metal titrations of $Tb^{3+}$ - LF4 binding .....	70
Figure 2.19	Metal titrations of $Yb^{3+}$ - LF4 binding .....	71
Figure 2.20	Full titration of LF4- $Dy^{3+}$ .....	72
Figure 2.21	Full titration of LF4- $Eu^{3+}$ .....	73
Figure 2.22	Full titration of LF4- $Gd^{3+}$ .....	74
Figure 2.23	Metal titrations of $La^{3+}$ - LF4 binding .....	75
Figure 2.24	Metal titrations of $Yb^{3+}$ - LF5 binding .....	76
Figure 2.25	Full titration of LF5- $La^{3+}$ .....	77
Figure 2.26	Metal titrations of $Eu^{3+}$ - LF5 binding .....	78

Figure 2.27	Metal titrations of Yb <sup>3+</sup> -LF6 binding .....	79
Figure 2.28	Full titration of LF6-Tb <sup>3+</sup> .....	80
Figure 2.29	Full titration of LF6-La <sup>3+</sup> .....	81
Figure 2.30	Full titration of LF6-Eu <sup>3+</sup> .....	82
Figure 2.31	Full fluorescence titration of LF6-Tb <sup>3+</sup> .....	83
Figure 2.32	Metal titrations of Yb <sup>3+</sup> -LF7 binding .....	84
Figure 2.33	Full titration of LF7-Eu <sup>3+</sup> .....	85
Figure 2.34	Full titration of LF7-La <sup>3+</sup> .....	86
Figure 3.1	Structure of a single EF hand calcium-binding loop .....	112
Figure 3.2	Design of a phosphorylation-dependent motif .....	113
Figure 3.3	Titration of 10 μM non-phosphorylated HM1 and phosphorylated HM1 with Tb <sup>3+</sup> .....	114
Figure 3.4	Titration of 10 μM non-phosphorylated HM2 and phosphorylated HM2 with Tb <sup>3+</sup> .....	115
Figure 3.5	Titration of 10 μM non-phosphorylated HM3 and phosphorylated HM3 with Tb <sup>3+</sup> .....	116
Figure 3.6	Titration of 10 μM non-phosphorylated HM4 and phosphorylated HM4 with Tb <sup>3+</sup> .....	117
Figure 3.7	Titration of 10 μM non-phosphorylated HM5 and phosphorylated HM5 with Tb <sup>3+</sup> .....	118
Figure 3.8	Titration of 10 μM non-phosphorylated HM6 and phosphorylated HM6 with Tb <sup>3+</sup> .....	119
Figure 3.9	Titration of 10 μM non-phosphorylated HM7 and phosphorylated HM7 with Tb <sup>3+</sup> .....	120
Figure 3.10	Titration of 10 μM non-phosphorylated HM8 and phosphorylated HM8 with Tb <sup>3+</sup> .....	121
Figure 3.11	Titration of 10 μM non-phosphorylated HM9 and phosphorylated HM9 with Tb <sup>3+</sup> .....	122

## ABSTRACT

Cis-trans isomerization of amide bonds plays critical roles in protein molecular recognition, protein folding, protein misfolding and disease. Aromatic-proline sequences are particularly prone to exhibit cis amide bonds. The roles of residues adjacent to a tyrosine-proline residue pair on cis-trans isomerism were examined. A series of peptides XYPN was synthesized and the following effects were observed: (a) aromatic residues immediately preceding Tyr-Pro disfavor cis amide bonds; (b) proline residues preceding Tyr-Pro lead to multiple species; and (c) other residues exhibit similar values of  $K_{\text{trans/cis}}$ . In addition, the effect of the  $i+3$  residue was examined in a limited series of peptides TYPZ. NMR data indicated that aromatic residues, Pro, Asn, Ala and Val at the  $i+3$  residue all favor cis amide bonds.

The zinc finger is a compact metal-binding motif, consisting of an N-terminal  $\beta$ -hairpin and a C-terminal  $\alpha$ -helix connected by a loop. A new, general lanthanide-binding motif inspired by zinc fingers was designed, which involved optimization of nearly all residues of the protein. A series of peptides was synthesized and analyzed by circular dichroism, and the dissociation constants were determined for each peptide. The optimized peptide design, LF4, bound Terbium (III) with a dissociation constant of 6.8  $\mu\text{M}$  and adopted a metal-bound structure similar to a zinc finger.

Many human proteins exhibit function which is dependent on phosphorylation. We are trying to understand the effects of protein phosphorylation on protein structure within native proteins by the design of proteins whose structures are dependent on their phosphorylation state. Nine peptides were synthesized and the fluorescence analyzed in the non-phosphorylated and phosphorylated states. In addition, the dissociation constants were determined for both phosphorylated and non-phosphorylated forms, and the rate of phosphorylation by Protein Kinase A (PKA) analyzed. Two peptides were developed that are rapidly phosphorylated by PKA and that show a significant increase in fluorescence on phosphorylation, enabling their use as genetically encoded protein kinase sensors.

## Chapter 1

# EFFECTS OF $i$ AND $i+3$ RESIDUE IDENTITIES ON CIS-TRANS ISOMERISM OF THE AROMATIC $_{i+1}$ -PROLYL $_{i+2}$ AMIDE BOND: IMPLICATIONS FOR TYPE VI $\beta$ -TURN FORMATION

## Introduction

Cis-trans isomerization plays critical roles in protein folding, protein misfolding, and protein-protein interactions.<sup>1-5</sup> Cis-trans isomerization is frequently implicated as the slow step in protein folding or unfolding pathways, and its importance in protein folding rates is emphasized by the conservation of prolyl isomerases from prokaryotes to eukaryotes.<sup>6-10</sup> Cis amide bonds are most frequently observed in Xaa-Pro amide bonds. Overall, 5.7% of Xaa-Pro amide bonds are observed in the cis conformation, and a majority of the proteins in the PDB contain at least one cis amide bond.<sup>11-15</sup> The increased cis population of Xaa-Pro amide bonds is primarily, though not<sup>16</sup> entirely, due to the reduced steric differentiation between the cis and trans conformations of proline residues, relative to non-proline residues.<sup>17-20</sup> Identification of sites exhibiting cis amide bonds in proteins is of fundamental

importance in understanding protein folding and protein function, but remains extremely challenging in the absence of crystallographic data.

Aromatic-prolyl residue pairs are relatively more likely to exhibit cis amide bonds, with favorable interactions between the aromatic and proline rings stabilizing the cis conformation (Figure 1a).<sup>11,14,18,20-35</sup> In the PDB, 9.7% of tyrosine-proline residue pairs exhibit a cis amide bond, indicating that both local and global interactions are important in determining cis-trans isomerization state.

Aromatic-prolyl interactions can form the basis of very stable local structures.<sup>18,21-23,25,31</sup> Aromatic-prolyl interactions have been observed to stabilize type VI  $\beta$ -turns, structures in which a cis amide bond exists between the  $i+1$  and  $i+2$  (Pro) residues of the  $\beta$ -turn.<sup>12,14,19,24,26,36-40</sup> Type VI  $\beta$ -turns, which are subdivided into three subtypes (Table 1), are observed in globular proteins, and have been implicated as recognition elements for a number of protein-protein interfaces. Examples of protein-protein interactions that may be mediated by type VI  $\beta$ -turns include binding of the HIV-1<sub>IIIb</sub> gp120 V3 loop to the CXCR4 chemokine receptor; binding of the hormones morphiceptin and endomorphin-2 to  $\mu$  opioid receptors; the agonist and antagonist receptor interactions of oxytocin; the binding of the post-synaptic metabotropic glutamate receptor 5 to the adaptor protein Homer; and the binding of Bowman-Birk protease inhibitors to serine proteases.<sup>41-48</sup> More generally, type VI  $\beta$ -turns have been implicated as important conformations in the ligands of G protein-coupled receptors (GPCRs).<sup>49-51</sup>

The prevalence of type VI  $\beta$ -turns in protein molecular recognition has stimulated the development of simple peptidic and non-peptidic motifs with this secondary structure.<sup>19,21-23,26,35,40,52-116</sup> The determination of parameters to understand the local effects on cis-trans isomerization and type VI  $\beta$ -turn formation would allow improved prediction, design and interpretation of sequences containing these elements. In the definitive work in this area, Dyson, Wright and coworkers observed a type VI  $\beta$ -turn in simple unblocked (charged termini) hexapeptides SYPFDV, which exhibited stacking of the Tyr and Phe aromatic rings on the opposite faces of the proline ring.<sup>21-23</sup> In these studies, the roles of residues 2-5 on cis-trans isomerization and peptide structure were thoroughly examined. However, the role of the residue preceding tyrosine, corresponding to the *i* residue of the type VI  $\beta$ -turn, was not significantly examined.

Lubell recently reported an analysis of the effects of the *i*, *i*+1, and *i*+3 residues on cis-trans isomerization in peptides with the type VI  $\beta$ -turn-inducing residue 5-*tert*-butylproline at the *i*+2 position.<sup>26</sup> Limited analysis of the effects of the *i* and *i*+3 residues on the fraction of cis isomer revealed that Ala > Val > Ser at the *i* residue, and that Ala > Leu > Phe > Val at the *i*+3 residue.

As aromatic-prolyl interactions may nucleate protein folding and structure, we sought to identify local sequences adjacent to Tyr-Pro residue pairs which favor cis amide bonds. As a starting point, we considered the frequencies of residues in type VI  $\beta$ -turns. Due to inadequate statistics, analysis of the PDB has revealed only a limited

number of trends for type VI  $\beta$ -turns.<sup>14,24,117,118</sup> The most significant residue preferences for type VI  $\beta$ -turns are (a) absolute conservation of Pro at the  $i+2$  residue, to promote cis amide bonds; and (b) a significant preference for aromatic residues at the  $i+1$  residue. As a step toward understanding local sequence effects on the stability of type VI  $\beta$ -turns, we have examined the effects of the  $i$  and  $i+3$  residues on cis-trans isomerism of the Tyr <sub>$i+1$</sub> -Pro <sub>$i+2$</sub>  amide bond.

## Experimental

### Peptide synthesis, purification and characterization

Peptides were synthesized on Rink amide resin on a Rainin PS3 or CEM Odyssey 12 peptide synthesizer using standard Fmoc synthesis protocols with HBTU as a coupling reagent. All peptides were acetylated on the N-terminus and contained C-terminal amides.

To synthesize the peptide phosphoSYPN, the peptide SYPN was synthesized with a trityl protecting group on the serine residue. After peptide synthesis, the trityl group was selectively removed with 2% TFA/ 5% triethylsilane (TES) in CH<sub>2</sub>Cl<sub>2</sub> (3 × 1 min). SYPN was chemically phosphorylated on resin by a two-step global phosphorylation protocol. Phosphitylation was performed under nitrogen by addition to resin of tetrazole (1.35 mmol; 3 mL of 3% tetrazole in acetonitrile) (Transgenomics) and O,O-dibenzyl-N,N-diisopropylphosphoramidite (500  $\mu$ L, 1.52

mmol) (Fluka), and allowed to react for 5 h with mixing. The solution was removed and the resin was washed with DMF (3×) and CH<sub>2</sub>Cl<sub>2</sub> (3×). Oxidation was performed with *t*-butyl hydroperoxide (2 mL of a 3 M solution in CH<sub>2</sub>Cl<sub>2</sub>) and allowed to react with mixing for 1 hour. The solution was removed and the resin washed with DMF (3×), MeOH (3×), and CH<sub>2</sub>Cl<sub>2</sub> (3×).

Peptides were cleaved from the resin and deprotected for 2 h under standard conditions (92.5% TFA/2.5% TES/5% H<sub>2</sub>O for TYPD, TYPF, TYPN, TYPP, TYPW; 84% TFA/4% each of H<sub>2</sub>O/phenol/thioanisole/ethanedithiol for all other peptides). TFA was removed by evaporation and the peptides were dissolved in water. All peptides were purified to homogeneity using reverse phase HPLC on a Vydac C18 semi-preparative column (10 × 250 mm, 5-10 μm particle, 300 Å pore). Peptide purity was determined by analytical HPLC reinjection on a Rainin Microsorb MV analytical C18 column (4.6 × 250 mm, 3-5 μm particle, 100 Å pore). Unless otherwise noted, peptides were purified using an isocratic gradient of 100% buffer A (98% water, 2% acetonitrile, 0.06% TFA) for 10 minutes, followed by a linear gradient of 0-15% buffer B (20% water, 80% acetonitrile, 0.05% TFA) in buffer A over 60 minutes. The peptides GYPA, DYPR, SYPS, SYPN, TYPN, YFPN and PYPP were purified using a linear gradient of 0-20% buffer B in buffer A over 60 minutes. The peptides WYPN, TYPW, pivaloyl-AYPN, and pivaloyl-PYPN were purified using a linear gradient of 0-30% buffer B in buffer A over 60 minutes. The peptides TYPF and TYPP were purified using a linear gradient of 0-40% buffer B in buffer A over 60 minutes. The

peptide WWPN was purified using a linear gradient of 0-50% buffer B in buffer A over 60 minutes. The peptides TYPG and TYPV were purified using a linear gradient of 0-10% buffer B in buffer A over 100 minutes. The peptide TYP(D-Ala) was purified using a linear gradient of 0-20% buffer B in buffer A over 100 minutes.

Peptide identity was characterized by ESI MS (positive ion mode) on an LCQ Advantage (Finnigan) mass spectrometer. Analytical data for the peptides: GYPA ( $t_R$  = 21.9 min, mass expected 447.2, mass observed (M+Na) 470.3); DYPR ( $t_R$  20.4 min, exp. 591.3, obs. 591.5 (M+H)); SYPS ( $t_R$  18.0 min, exp. 493.2, obs. 493.9); SYPN ( $t_R$  16.5 min, exp. 520.2, obs. 543.3 (M+Na)); TYPN ( $t_R$  22.8 min, exp. 534.2, obs. 557.3 (M+Na)); AYPN ( $t_R$  20.5 min, exp. 504.2, obs. 527.4 (M+Na)); pivaloyl-AYPN ( $t_R$  28.6 min, exp. 546.3, obs. 569.5 (M+Na)); CYPN ( $t_R$  24.5 min, exp. 536.2, obs. 559.4 (M+Na)); DYPN ( $t_R$  18.1 min, exp. 548.2, obs. 571.4 (M+Na)); EYPN ( $t_R$  22.2 min, exp. 562.2, obs. 585.4 (M+Na)); FYPN ( $t_R$  43.7 min, exp. 580.3, obs. 603.5 (M+Na)); GYPN ( $t_R$  21.1 min, exp. 490.2, obs. 513.4 (M+Na)); HYPN ( $t_R$  21.3 min, exp. 570.3, obs. 571.3 (M+H)); IYPN ( $t_R$  44.2 min, exp. 546.3, obs. 547.3 (M+H)); KYPN ( $t_R$  13.1 min, exp. 561.3, obs. 562.4 (M+H)); LYPN ( $t_R$  48.6 min, exp. 546.3, obs. 547.3 (M+H)); MYPN ( $t_R$  37.3 min, exp. 564.2, obs. 587.4 (M+Na)); NYPN ( $t_R$  17.4 min, exp. 547.2, obs. 570.4 (M+Na)); PYPN ( $t_R$  26.1 min, exp. 530.2, obs. 553.4 (M+Na)); pivaloyl-PYPN ( $t_R$  51.6 min, exp. 572.3, obs. 595.5 (M+Na)); QYPN ( $t_R$  22.9 min, exp. 561.3, obs. 584.4 (M+Na)); RYPN ( $t_R$  22.1 min, exp. 589.3, obs. 590.4 (M+H)); phospho-SYPN ( $t_R$  18.7 min, exp. 600.2, obs. 599.4 (M-H); analyzed in negative ion mode); VYPN ( $t_R$  34.4 min, exp. 532.3, obs. 555.4 (M+Na)); WYPN ( $t_R$  49.8 min,

exp. 619.3, obs. 642.5 (M+Na)); YYPN ( $t_R$  36.3 min, exp. 596.3, obs. 619.4 (M+Na)); YFPN ( $t_R$  48.2 min, exp. 580.3, obs. 603.5 (M+Na)); WWPN ( $t_R$  45.3 min, exp. 642.3, obs. 665.5 (M+Na)); TYPA ( $t_R$  24.4 min, exp. 491.2, obs. 514.4 (M+Na)); TYPD ( $t_R$  30.1 min, exp. 535.2, obs. 558.3 (M+Na)); TYPF ( $t_R$  44.3 min, exp. 567.3, obs. 590.3 (M+Na)); TYPG ( $t_R$  40.0 min, exp. 477.2, obs. 500.4 (M+Na)); TYPP ( $t_R$  15.2 min, exp. 517.3, obs. 540.4 (M+Na)); TYPV ( $t_R$  68.7 min, exp. 519.3, obs. 542.5 (M+Na)); TYPW ( $t_R$  31.3 min, exp. 606.3, obs. 629.5 (M+Na)); TYP(D-Ala) ( $t_R$  45.2 min, exp. 491.2, obs. 514.5 (M+Na)); PYPP ( $t_R$  33.0 min, exp. 513.3, obs. 536.5 (M+Na)). Peptide purity and identity were further confirmed by NMR spectroscopy (see below).

### **NMR analysis**

Peptides were dissolved in buffer containing 5 mM potassium phosphate (pH 4, or as indicated), 25 mM (Table 2) or 250 mM (other data) NaCl, 90% H<sub>2</sub>O/10% D<sub>2</sub>O, and 100  $\mu$ M TSP. Peptide concentrations were 200  $\mu$ M (YYPN) or 2 mM (AYPN) for temperature-dependent NMR experiments, and 100 - 500  $\mu$ M for all other experiments. NMR spectra were recorded at 296 K on a Bruker DRX 400 MHz NMR spectrometer with a QNP probe or on a Bruker AVC 600 MHz NMR spectrometer with a triple-resonance cryoprobe. 1-D NMR spectra were recorded using a gradient watergate pulse sequence and a 3-5 s relaxation delay. Watergate TOCSY spectra were recorded for all peptides for resonance assignment. Temperature-dependent NMR spectra were recorded at 285, 293, 301, 309, 316 and 323 K. Data (128-2048

scans) were collected 60-180 min after thermal stabilization, to assure the establishment of equilibrium between the cis and trans isomers at a given temperature.

The identification of the minor species as the isomer containing a cis Tyr-Pro amide bond was made by ROESY experiments on the peptide TYPN, using a gradient ROESY experiment with water suppression by excitation sculpting and with a 160 ms mixing time. Strong ROEs were observed between the Tyr  $H_\alpha$  and the Pro  $H_\delta$  protons in the major species, consistent with a trans Tyr-Pro conformation; and a strong ROE was observed between the Tyr  $H_\alpha$  and the Pro  $H_\alpha$  in the minor species, consistent with a cis conformation. In other XYPN peptides, the assignment of the major species as the species containing a trans Tyr-Pro bond was made by comparison of the chemical shifts of the Tyr, Pro and Asn  $H_\alpha$  with those in TYPN containing a trans or cis Tyr-Pro bond. In the peptides TYPZ, the assignment of the major species as the species containing a trans Tyr-Pro bond was made by comparison of the chemical shifts of the Thr  $H_\alpha$ ,  $H_\beta$  and  $H_\gamma$  and the Pro  $H_\alpha$  with those in TYPN containing a trans or cis Tyr-Pro bond. The observation that Pro<sub>trans</sub>  $H_\alpha$  is considerably downfield of Pro<sub>cis</sub>  $H_\alpha$  is consistent with previous analyses of peptides containing an aromatic-proline residue pair.<sup>20-22,25,26,33,34</sup>

NMR spectra were integrated after baseline correction using pairs of well-resolved peaks, comprising cis and trans isomers, of a single proton. Errors in  $K_{\text{trans/cis}}$  are estimated to be  $\leq 5\%$  of  $K_{\text{trans/cis}}$ .  $^3J_{\alpha\text{N}}$  was determined directly from the 1-D spectra. Errors  $^3J_{\alpha\text{N}}$  in are estimated to be  $\leq 0.2$  Hz.  $\phi$  was calculated based on the

parametrized Karplus equation,  ${}^3J_{\alpha N} = 6.51 \cos^2(\phi - 60) - 1.76 \cos(\phi - 60) + 1.6$ .<sup>119</sup>  $\Delta G$  indicates the energetic preference for the trans over the cis isomer, and was calculated via  $\Delta G = -RT \ln K_{\text{trans/cis}}$ .  $\Delta\Delta G$  was calculated via  $\Delta G_{\text{XYPN}} - \Delta G_{\text{HYPN}}$ .

## Results and Discussion

### Determination of a general sequence XYPZ to promote cis amide bonds

Cis-trans isomerism was initially examined within the sequence XYPZ, with X and Z chosen based on residues with the highest frequencies at these positions in type VI  $\beta$ -turns,<sup>14,24,117</sup> and a Tyr-Pro residue pair chosen due to the favorable aromatic-Pro interaction, which promotes cis amide bonds. In considering possible X and Z residues, attention was given to residues with the possibility of favorable hydrogen bonding or electrostatic interactions. In addition, residue Z ( $i+3$  residue) was chosen also considering the data of Dyson and Wright.<sup>21-23</sup> Aromatic residues, which were previously shown to favor cis amide bonds when following Pro, were not initially examined at the  $i+3$  position.

This initial series (Table 2) indicated that Asn at the  $i+3$  position stabilized cis amide bonds and promoted reduced values of  ${}^3J_{\alpha N}$  (corresponding to reduced  $\phi$ ) in the Tyr<sub>cis</sub> conformation. Other residues exhibited larger  $K_{\text{trans/cis}}$  and/or larger values of  ${}^3J_{\alpha N}$  for Tyr<sub>cis</sub>. These results are consistent with the results of Dyson and Wright, who

observed that Asn and protonated Asp were the most effective non-aromatic residues at the  $i+3$  position in promoting cis amide bonds and reduced  $\phi$  in Tyr<sub>cis</sub>.<sup>22</sup>

In peptides containing N-terminal and C-terminal amides, cis-trans isomerization of aromatic-Pro sequences has been previously examined in the contexts GYPG and AYPK, with the peptide GYPG exhibiting a lower cis population and a random coil-like 7.0 Hz coupling constant for Tyr<sub>cis</sub>.<sup>20,25</sup> Our data similarly indicate that Ala at the  $i+3$  (Z) residue promotes cis amide bond formation compared to Gly (Table 2). The peptides SYPN and TYPN also exhibited significant cis population, as well as a reduced  $^3J_{\alpha N}$  relative to other peptides examined in this initial series. In sum, these results suggest that Asn at residue  $i+3$  promotes cis amide bonds and restricted  $\phi$  in Tyr<sub>cis</sub>, both of which are required for type VIa1  $\beta$ -turns. Based on these results, we chose to examine the effects of the  $i$  residue on Tyr-Pro cis-trans isomerization in the context XYPN.

### **Determination of the effects of residue $i$ on cis-trans isomerism in the sequence XYPN**

Peptides of the sequence XYPN, with X = each of the canonical amino acids, were synthesized and analyzed by NMR (Table 3, Table 4, and Figure 2). Most peptides exhibited relatively small differences in  $K_{\text{trans/cis}}$  and  $^3J_{\alpha N}$  for Tyr<sub>cis</sub>. In addition, for most peptides there was little observed variation in the H<sub>N</sub>, H <sub>$\alpha$</sub>  and H <sub>$\beta$</sub>  chemical shifts of the more remote Asn residue, indicating few or no long-range

interactions. In sum, these results indicate that the *i* residue has only a modest effect on the cis-trans isomerization equilibrium in this context.

The highest populations of cis isomer for a canonical amino acid X ( $X \neq \text{Pro}$ ) in the sequence XYPN were observed for Thr and protonated His. As was observed in the initial series, most peptides in the XYPN series exhibited moderately reduced  $^3J_{\alpha\text{N}}$  (average  $^3J_{\alpha\text{N}} = 6.3 \text{ Hz}$ ) for  $\text{Tyr}_{\text{cis}}$ , relative to that observed in the random coil sequence GYPG, indicating a somewhat more compact conformation. Pro, Cys, Gly, Ile, Val, Gln and Lys also exhibited modestly lower values of  $K_{\text{trans/cis}}$ . In general,  $\beta$ -branched, short polar and Gly residues at the *i* position were the most favorable for inducing a cis amide bond, although the effects were very modest ( $\leq 0.2 \text{ kcal mol}^{-1}$ ).

The observation of higher populations of cis amide bonds for Gly, short polar, and  $\beta$ -branched residues at the *i* position is consistent with the statistical analyses of the *i* residue of type VI  $\beta$ -turns, where these residues were observed to be preferred.<sup>14</sup> Statistical analyses also showed reduced frequencies of Ala, Asn, and Arg at the *i* position of type VIa1  $\beta$ -turns, and reduced frequencies of Glu at the *i* position of type VIb  $\beta$ -turns. The data presented here are not inconsistent with the statistical analyses: Ala, Asn and Arg exhibit relatively larger values of  $K_{\text{trans/cis}}$  and/or  $^3J_{\alpha\text{N}}$  for  $\text{Tyr}_{\text{cis}}$ , both of which disfavor type VIa1  $\beta$ -turn formation; similarly, Glu exhibits a relatively larger value of  $K_{\text{trans/cis}}$  and a reduced value of  $^3J_{\alpha\text{N}}$  for  $\text{Tyr}_{\text{cis}}$ , both of which are inconsistent with type VIb  $\beta$ -turn formation. In sum, these data suggest that the observed residue preferences of the *i* residue, though modest in this disordered

context, are applicable to residue preferences in a more complex folded environment, where the energetic differences are expected to be more significant. Moreover, these data indicate that, in the unfolded state, extremely local interactions are in general insufficient to define cis-trans isomerization state, and that nascent secondary or tertiary structure is necessary to specify a cis amide bond, consistent with previous data on peptides and the observation of cis-trans isomerization as a slow step in protein folding.

The largest effects on cis-trans isomerization were observed for X = aromatic residues. Aromatic residues at the *i* position significantly stabilized the trans isomer relative to the cis isomer, with Trp > Tyr > Phe in stabilizing the trans isomer. In addition, an increased  $^3J_{\alpha N}$  for Tyr<sub>cis</sub> was also observed for the peptide YYPN. The trans conformation was stabilized in peptides with X = aromatic by 0.40 - 0.60 kcal mol<sup>-1</sup> relative to X = protonated His, with the largest trans stabilization induced by the most electron-rich aromatic residue, Trp. The effect of aromatic residues on increasing  $K_{\text{trans/cis}}$  is presumably due to the competitive interaction of the *i* residue aromatic ring with the aromatic ring of Tyr<sub>*i*+1</sub>, which otherwise interacts with the proline ring to stabilize the cis conformation.<sup>120-123</sup> By reducing the relative stability of the aromatic-prolyl interaction via the competitive trans-promoting aromatic-aromatic interaction,  $K_{\text{trans/cis}}$  is increased.

Consistent with this interpretation, the H<sub>α</sub> chemical shift of Y<sub>trans</sub> is shifted 0.13 - 0.30 ppm upfield in FYPN, YYPN and WYPN, relative to TYPN, consistent with ring current effects expected for a transient aromatic-aromatic interaction (Table

4). The effect of aromatic *i* residues is negligible on  $Y_{\text{cis}}$  (0.00 - 0.04 ppm upfield shift in  $H_{\alpha}$ ). Similarly, significant upfield shifts in  $H_{\beta}$  were observed in  $Y_{\text{trans}}$ ; in  $Y_{\text{cis}}$ , the shifts observed were stereospecific (one upfield shift, one downfield shift, relative to  $Y_{\text{cis}}$  in TYPN). Moreover, WYPN, YYPN and FYPN exhibited the only significant effects on the amide chemical shifts of  $N_{\text{trans}}$ . Similar effects on  $K_{\text{trans/cis}}$  were also observed for aromatic-aromatic pairs in related peptides: YFPN significantly favors trans ( $\Delta\Delta G = -0.47 \text{ kcal mol}^{-1}$ ) compared to TFPN, and WWPN significantly favors trans ( $\Delta\Delta G = -0.55 \text{ kcal mol}^{-1}$ ) compared to TWPN (data not shown). In total, these data suggest that aromatic-aromatic-Pro sequences may be identified as aromatic-Pro sequences with a relative preference for trans amide bonds.

In most cases, side chain protonation state had no significant effect on  $K_{\text{trans/cis}}$ . In contrast, histidine exhibited pH-dependent effects on  $K_{\text{trans/cis}}$ . Protonated His (pH 4) was the most favorable residue for inducing a cis conformation. The peptide containing neutral His, on the other hand, exhibited a significantly larger  $K_{\text{trans/cis}}$ , with a 0.21 kcal/mol preference for trans compared to protonated His. These effects may be related to protonation-mediated changes in the electronics of the aromatic imidazole ring: neutral His exhibited similar (though smaller) effects on chemical shifts as were described for Trp, Tyr, and Phe at the *i* residue, suggesting that neutral His-Tyr aromatic-aromatic interactions stabilize the trans conformation.

The peptide PYPN exhibited four species, due to cis-trans isomerism of either X-Pro amide bond (Ac-Pro or Tyr-Pro) (Figure 3a). The major species, whose NMR

data are consistent with the all-trans isomer, represented only 55% of the total population ( $K_{\text{major}/(\Sigma \text{ minor})} = 1.2$ ), with the other three species representing 22%, 15% and 8% of the total population, respectively. Based on the consistent  $H_N$   $\delta$  of  $\text{Asn}_{\text{cis}}$  among all peptides in the XYPN series, the 22% species was tentatively assigned to the trans(Ac-Pro)-cis(Tyr-Pro) isomer, yielding  $K_{(\text{trans-trans})/(\text{trans-cis})} = 2.5$ , and indicating that Pro at the  $i$  residue most strongly favors a cis Tyr-Pro bond. The effect of Pro on Tyr-Pro isomerism was independently addressed by replacement of the acetyl group on the N-terminus with the pivaloyl group, which is sterically demanding and lacks alpha protons. Pivaloylation enforced a trans isomer for the Piv-Pro amide bond, allowing determination of the effect of Pro at the  $i$  position on Tyr-Pro cis-trans isomerism. The peptide Piv-PYPN exhibited  $K_{\text{trans/cis}} = 3.2$  and  $^3J_{\alpha N}$  for  $\text{Tyr}_{\text{cis}} = 6.2$  Hz, indicating that Pro is comparable to Thr or protonated His as the most cis-favoring  $i$  residue (Figure 3b). As a control for the effect of the pivaloyl group on  $K_{\text{trans/cis}}$ , we observed that the pivaloyl group minimally affected  $K_{\text{trans/cis}}$  for AYPN. These data are consistent with a high frequency of Pro at the  $i$  position of type VI  $\beta$ -turns.<sup>14,24,117,118</sup> Moreover, these data emphasize the importance of the residue preceding  $\text{Pro}_i$  on cis-trans isomerism of the  $\text{Xaa}_{i-1}$ - $\text{Pro}_i$  amide bond, in which a cis conformation is stabilized by a favorable interaction between the aromatic ring of  $\text{Tyr}_{i+1}$  and  $H_\alpha$  of the  $i-1$  residue.<sup>22,28-32</sup>

The temperature dependence of cis-trans isomerism was examined for the peptides AYPN and YYPN (Figure 4). For AYPN,  $K_{\text{trans/cis}}$  exhibited a very modest

temperature dependence between 285 K and 323 K. The trans conformation is primarily favored entropically for AYPN ( $\Delta H = -0.22 \text{ kcal mol}^{-1}$ ,  $\Delta S = +1.8 \text{ cal mol}^{-1} \text{ K}^{-1}$ ). In contrast, the peptide YYPN exhibited a larger temperature dependence for  $K_{\text{trans/cis}}$ , indicating primarily enthalpic stabilization of the trans conformation for YYPN ( $\Delta H = -1.01 \text{ kcal mol}^{-1}$ ,  $\Delta S = +0.4 \text{ cal mol}^{-1} \text{ K}^{-1}$ ), similar to that observed in Ac-GP-OMe ( $\Delta H = -1.27 \text{ kcal mol}^{-1}$ ).<sup>124</sup> Wu and Raleigh observed that the aromatic-Pro interaction reduces the enthalpic driving force for the trans conformation, as was also observed here in AYPN.<sup>25</sup> The results with YYPN indicate that the aromatic-aromatic interaction provides additional enthalpic stabilization of the trans conformation, reestablishing the usual favorable enthalpy of adopting the trans conformation in an Xaa-Pro bond.<sup>125-127</sup>

Protein phosphorylation commonly occurs in disordered and surface regions of proteins, including proline-rich sequences.<sup>128</sup> As phosphorylation could stabilize local structures, we compared SYPN to phosphoSYPN (Table 3). Phosphorylation modestly increased the population of the cis isomer, with a larger increase for the dianionic state ( $K_{\text{trans/cis}} = 3.0$ ) than the monoanionic state ( $K_{\text{trans/cis}} = 3.2$ ). The increase in cis content was accompanied by a phosphorylation-dependent decrease in  $^3J_{\alpha\text{N}}$  of Tyr<sub>cis</sub> (monoanionic phosphoSYPN,  $^3J_{\alpha\text{N}}$  of Tyr<sub>cis</sub> = 5.7 Hz, versus 6.5 Hz in SYPN; the amide protons were not observable in dianionic phosphoSYPN due to rapid exchange). Interestingly, phosphoserine induced larger changes in the chemical shifts of adjacent residues than did any residue other than Pro or aromatics (Table 4). These

data suggest that phosphorylation may induce local conformational changes and affect cis-trans equilibria in proteins.<sup>10,129</sup> Indeed, the complex of the adaptor protein 14-3-3 $\zeta$  with the binding motif from polyoma middle T is mediated by a type VIb  $\beta$ -turn with the closely related sequence phosphoSer-Tyr-cisPro-Ala.<sup>130,131</sup>

### **Determination of the effects of residue $i+3$ on cis-trans isomerism in the sequence TYPZ**

The effects of the  $i+3$  residue on cis-trans isomerism were analyzed within peptides of the sequence TYPZ (Table 5 and Table 6). Dyson and Wright previously thoroughly examined the effects of  $i+3$  residue identity on cis-trans isomerism within a related SYPXDV peptide context, with charged N- and C-termini. We examined a limited number of peptides in the TYPZ series to determine the effects of  $i+3$  residue identity on conformational equilibria within a peptide context with uncharged termini, as significant termini effects were observed for SYPYDV. The residues Z examined were chosen primarily based on high % cis observed by Dyson and Wright.

The peptide TYPA exhibited  $K_{\text{trans/cis}} = 3.1$ , similar to that observed for TYPN. In contrast, TYPG exhibited a larger  $K_{\text{trans/cis}} = 4.3$ . These data are consistent with previous data indicating that Ala at the  $i+3$  position favors cis amide bonds.<sup>20,22,26,40</sup> Interestingly, peptides with Ala at the  $i+3$  position exhibited larger  $^3J_{\alpha\text{N}}$  for Tyr<sub>cis</sub> than did peptides with Asn at the  $i+3$  position. These data suggest that Ala may be favored in type VIa2 or typeVIb  $\beta$ -turns, which have more extended values of  $\phi$  at residue  $i+1$ .

Indeed, type VI  $\beta$ -turn potentials indicate that Ala is the most common residue at the  $i+3$  position of typeVIb  $\beta$ -turns, but is rarely observed at the  $i+3$  position in typeVIa1  $\beta$ -turns.<sup>14,24,117</sup>

Aromatic residues at the  $i+3$  position most strongly favored the cis isomer and a reduced  $^3J_{\alpha N}$  for Tyr<sub>cis</sub>, as was previously observed by Dyson and Wright. Phe was somewhat more effective than Trp in favoring cis amide bonds, suggesting a role for aromatic electronics in stabilizing both the trans and cis isomers. NMR data provide evidence for aromatic-aromatic interactions in the trans conformations of TYPW and TYPF (Table 6). In TYPW, the chemical shifts of all Tyr protons are shifted upfield in the trans conformation, relative to other peptides TYPZ. Of particular note, the diastereotopic Tyr<sub>trans</sub> beta protons are both shifted significantly upfield ( $\delta = 2.42$  ppm) and are rendered degenerate in TYPW. Moreover, the amide proton of Trp<sub>trans</sub> in TYPW is shifted significantly upfield, compared to Phe<sub>trans</sub> or Thr<sub>trans</sub>, consistent with aromatic-aromatic and/or aromatic-amide interactions stabilizing the trans conformation.<sup>28-31,132-135</sup> In the cis isomers, the NMR data are consistent with stabilization of the cis conformation by stacking of the aromatic rings against the opposite faces of the proline ring, as is generally observed in X-aromatic-Pro-aromatic sequences in type VI  $\beta$ -turns: in TYPF and TYPW, the four Pro<sub>cis</sub> beta and gamma protons are all well-resolved, and are shifted significantly upfield compared to other TYPZ peptides (TYPW: 1.70, 1.52, 1.36, 1.13 ppm; TYPF 1.78, 1.62, 1.45, 1.33 ppm; TYPA 1.95, 1.86, 1.74, 1.74 ppm); upfield shifts in the Tyr<sub>cis</sub> alpha protons were also

observed (Table 6). In total, these data suggest that, in addition to the cis-favoring interaction between the  $i+3$  aromatic residue and the preceding ( $i+2$ ) proline residue, an interaction between the aromatic rings of the  $i+1$  and  $i+3$  residues may stabilize the trans conformer, with the Tyr-Trp interaction more trans-stabilizing than Tyr-Phe, as was also observed in the XYPN series.

In the SYPZDV series of Dyson and Wright, Asp was observed to be the most cis-favoring non-aromatic residue at the  $i+3$  position, exhibiting the highest %cis and lowest  $^3J_{\alpha N}$  for Tyr<sub>cis</sub>. Those data were collected at pH 4.1 in the context of a peptide with free termini. Data on the peptide TYPD indicate that neutral (protonated) Asp favors trans comparably to Asn, with a somewhat larger  $^3J_{\alpha N}$  for Tyr<sub>cis</sub>. On the other hand, the more physiologically relevant anionic Asp at the  $i+3$  position relatively disfavors the cis amide conformation.

The peptide TYPP exhibited a mixture of four species, with cis and trans Tyr-Pro and Pro-Pro bonds (Figure 3c).<sup>136</sup> Ignoring the relatively minor (14%) contributions of species with Pro-Pro cis amide bonds, Pro at the  $i+3$  residue was observed to modestly favor cis Tyr-Pro amide bonds compared to Asn or Ala ( $K_{(trans-trans)/(cis-trans)} = 2.6$  for TYPP). In this peptide, the normally trans-stabilizing interaction between the aromatic ring of Tyr <sub>$i+1$</sub>  and the amide proton of the  $i+3$  residue (Figure 1c) is not possible.<sup>30,31,132-135</sup> Notably, the  $^3J_{\alpha N}$  for Tyr<sub>cis</sub> = 8.7 Hz, indicating that Pro strongly favors type VIa2 or type VIb  $\beta$ -turns (Table 1), which exhibit an extended conformation for the  $i+1$  residue. These data are consistent with  $\beta$ -turn potentials,

which indicate that Pro is relatively disfavored at the  $i+3$  position in type VIa1  $\beta$ -turns, but strongly favored at the  $i+3$  position in type VIb  $\beta$ -turns.<sup>14,24,117</sup> In addition, the critical type VIb  $\beta$ -turn observed in the Bowman-Birk trypsin inhibitors contains a conserved Pro at the  $i+3$  residue of the  $\beta$ -turn.<sup>46,47</sup>

As was observed previously, Gly at the  $i+3$  residue disfavored cis amide bond formation, consistent with data on type VI  $\beta$ -turn potentials.<sup>22,25,40</sup> Moreover, D-Ala particularly disfavored the cis conformation. Pro-Gly, Pro-D-Ala, and Pro-D-Xaa sequences are known to strongly promote type II  $\beta$ -turns centered on the Pro-Xaa bond, and to reduce the population of the cis-Pro isomer.<sup>40,137-142</sup> These data are consistent with the interpretation that the reduced cis population observed in peptides TYPZ containing Gly or D-Ala at the Z position may be due to competitive nascent type II  $\beta$ -turn formation, which enforces a trans proline amide bond.

### **Effect of multiple proline residues on cis-trans isomerism**

Statistical analyses indicate that proline is observed with increased frequencies at all 4 positions of type VI  $\beta$ -turns.<sup>14,24,117,118</sup> However, the data above present a complex picture of the effects of aromatic residues on cis-trans isomerism when multiple proline residues are present (Figure 3). In order to further examine cis-trans isomerism of tyrosine-containing peptides in a proline-rich environment, the peptide PYPP was synthesized. PYPP, in which proline residues were incorporated at each non-Tyr residue of the sequence, exhibited a mixture of 3 major species (35%, 21%,

and 18% of the total population), 2 minor species (12% and 8%), and 2 very minor species (3% and 2%), accounting for 7 of the 8 possible isomers with combinations of cis and trans Ac-Pro, Tyr-Pro and Pro-Pro bonds (Figure 3d).

In this proline-rich context, tyrosine promotes cis amide bond formation of both the Ac-Pro and Tyr-Pro amide bonds (Figures 1a and 1b). With Pro at the  $i+3$  position, the trans-stabilizing  $\text{Tyr}_{i+1}\text{H}^{\text{N}}_{i+3}$  interaction (Figure 1c) is not possible, further promoting cis amide bond formation. In addition, Pro-Pro cis-trans isomerism is expected, as 10% of Pro-Pro amide bonds in the PDB are cis. Although a complete conformational analysis of PYPP was not conducted, the data indicate that, at most, only 35% of the total peptide population is the all-trans isomer. The complex picture of cis-trans isomerism in peptides containing aromatic residues in proline-rich sequences suggests that aromatic residues should be strongly disfavored in proline-rich domains of proteins due to significant cis-trans isomerism, consistent with data recently reported on the polyproline helix propensities of aromatic residues in the context GPPXPPGY.<sup>143</sup>

## Conclusions

The results herein are the first extensive analysis of the effects of the  $i$  residue on cis-trans isomerism of a  $\text{Tyr}_{i+1}\text{-Pro}_{i+2}$  sequence. We have determined by NMR spectroscopy the effect of each of the 20 canonical amino acids on cis-trans isomerism of the subsequent Tyr-Pro amide bond. We observed that aromatic residues in the  $i$

position disfavor cis amide bonds; that proline in the  $i$  position promotes cis-trans isomerism of both Xaa-Pro bonds; and that most residues at the  $i$  position have only a modest effect on cis-trans isomerization. Pro, Thr, protonated His, and phosphorylated serine at the  $i$  position most strongly promote cis amide bond formation. At the  $i+3$  position, we observed that aromatic residues, Asn, Ala, Val and Pro promote cis amide bonds, with aromatic residues and Asn favoring conformations more consistent with type VIa1  $\beta$ -turns, and Ala and Pro favoring conformations more consistent with type VIb  $\beta$ -turns. D-Ala at the  $i+3$  position particularly disfavored cis amide bond formation. Proline-rich peptides, containing Pro at the  $i$  and/or  $i+3$  positions, exhibited complex mixtures of species due to cis-trans isomerization of each X-Pro bond.

At both the  $i$  and  $i+3$  positions, most residues had only a minor effect on cis-trans isomerism within the disordered context of a small linear peptide, although the modest residue preferences observed are consistent with analyses of residue preferences in type VI  $\beta$ -turns. These data provide a context for understanding local sequence effects on cis-trans isomerism, in predicting sites of cis-trans isomerism in native proteins, and in the design of novel sequences to adopt type VI  $\beta$ -turns. Indeed, we recently analyzed a series of peptides derived from TYPN, which was identified above as a sequence favoring a cis Tyr-Pro bond.<sup>35</sup> Substitution of the Pro residue with 4-*cis*-fluoroproline generated a simple tetrapeptide with 40% cis amide bond ( $K_{\text{trans/cis}} = 1.5$ ) and a reduced (4.9 Hz)  $^3J_{\alpha\text{N}}$  coupling constant on Tyr<sub>cis</sub>, suggestive of a type

Vla1  $\beta$ -turn. Further efforts to apply the data herein to develop simple peptides adopting type VI  $\beta$ -turns are ongoing.

Note: This chapter was published in *Biopolymers (Peptide Science)*:  
doi 10.1002/bip.20382

© 2005 Wiley Periodicals, Inc.

## Figures and Tables

Type VI $\beta$ -turn subtype	$\phi (i+1)$	$\psi (i+1)$	$\phi (i+2)$	$\psi (i+2)$	${}^3J_{\alpha N}$ $(i+1)_{\text{cis}}$	$i_{\text{C=O}} \rightarrow$ $i+3_{\text{N-H}}$ hydrogen bond
Vla1	-60	+120	-90	0	4 Hz	Y/N
Vla2	-120	+120	-60	0	10 Hz	N
Vlb	-135	+135	-75	+160	9 Hz	N

**Table 1.1.** Ideal characteristics of type VI  $\beta$ -turn subtypes.<sup>24,36,38,39</sup> All subtypes exhibit a cis amide bond ( $\omega = 0^\circ$ ) between the  $i+1$  and  $i+2$  residues. Expected  ${}^3J_{\alpha N}$  values are calculated based on a parametrized Karplus equation.<sup>119</sup>

peptide	$K_{\text{trans/cis}}$	$\Delta G$ , kcal mol <sup>-1</sup>	$^3J_{\alpha\text{N}}(\text{Tyr}_{\text{cis}})$ , Hz
DYPR	3.8	- 0.78	6.9
SYPS	3.8	- 0.78	6.8
GYPA	3.2	- 0.68	n.d.
SYPN	3.2	- 0.68	6.3
TYPN	2.7	- 0.58	6.1
GYPG ( <sup>25</sup> )	4.0	- 0.82	7.0
AYPAK ( <sup>20</sup> )	3.2	- 0.68	n.d.

**Table 1.2.** NMR-derived data for the peptides XYPZ. Experiments were conducted at 23 °C in 5 mM phosphate buffer, 25 mM NaCl, pH 6.5 (DYPR) or 4.0 (other peptides). n.d. = not determined. (Data from Krista M. Thomas.)

XYPN, X =	$K_{\text{trans/cis}}$	$\Delta G_{\text{trans/cis}}$ kcal mol <sup>-1</sup>	$\Delta\Delta G$ kcal mol <sup>-1</sup>	$Y_{\text{cis}} \text{}^3J_{\text{HN}\alpha}$ Hz
H (pH 4)	2.9	- 0.63	0.00	n.d.
T	3.0	- 0.65	- 0.03	6.3
pivaloyl-P	3.2	- 0.68	- 0.05	6.2
C (pH 9)	3.2	- 0.68	- 0.06	n.d.
C (pH 4)	3.2	- 0.69	- 0.07	6.0
G	3.3	- 0.70	- 0.08	n.d.
I	3.4	- 0.71	- 0.08	6.1
V	3.4	- 0.71	- 0.09	6.3
Q	3.4	- 0.72	- 0.09	6.3
K	3.4	- 0.72	- 0.10	6.1
S	3.6	- 0.75	- 0.12	6.5
R	3.6	- 0.76	- 0.13	6.4
A	3.7	- 0.76	- 0.14	6.5
pivaloyl-A	3.7	- 0.77	- 0.14	6.3
E (pH 3)	3.7	- 0.78	- 0.15	6.0
E (pH 6.5)	3.9	- 0.79	- 0.17	5.7
M	3.9	- 0.80	- 0.17	6.3
L	4.0	- 0.81	- 0.18	6.3
D (pH 6.5)	4.0	- 0.82	- 0.19	n.d.
D (pH 2.5)	4.1	- 0.82	- 0.20	6.1
N	4.1	- 0.83	- 0.20	n.d.
H (pH 8.5)	4.2	- 0.84	- 0.21	n.d.
F	5.7	- 1.02	- 0.40	6.4
Y	6.7	- 1.12	- 0.50	6.7
W	8.0	- 1.22	- 0.60	n.d.
phosphoS (pH 5.5)	3.2	- 0.68	- 0.06	5.7
phosphoS (pH 8.5)	3.0	- 0.64	- 0.01	n.d.

**Table 1.3.** NMR-derived data for the peptides XYPN. Experiments were conducted at 23 °C in 5 mM phosphate buffer, 250 mM NaCl, pH 4.0 or as indicated. Peptides are listed in order of increasing  $K_{\text{trans/cis}}$ .  $\Delta\Delta G = \Delta G_{\text{XYPN}} - \Delta G_{\text{HYPN}}$ . n.d. = not determined because of spectral overlap.

XYPN, X =	$\delta, H_N$	$\delta, H_N$	$\delta, H_\alpha$	$\delta, H_{\beta A}$	$\delta, H_{\beta B}$	$\delta, H_{\beta A}$	$\delta, H_{\beta B}$	$\delta, H_{\delta B}$	$\delta, H_N$
	$Y_{trans}$	$Y_{cis}$	$Y_{trans}$	$Y_{trans}$	$Y_{trans}$	$Y_{cis}$	$Y_{cis}$	$P_{trans}$	$N_{trans}$
H (pH 4)	8.17	8.34	4.88	3.14	2.83	2.97	2.91	3.65	8.46
T	8.31	8.39	4.90	3.12	2.86	3.03	2.91	3.62	8.43
pivaloyl-P	7.93	8.15	n.d.	3.09	2.93	3.08	2.89	3.50	8.31
C (pH 9)	8.35	8.28	4.91	3.13	2.86	3.01	2.93	3.63	8.46
C (pH 4)	8.34	8.40	4.91	3.13	2.85	3.01	2.90	3.65	8.45
G	8.07	8.18	4.85	3.11	2.86	3.02	2.89	3.60	8.43
I	8.29	8.32	4.94	3.12	2.84	3.01	2.91	3.67	8.43
V	8.33	8.36	4.92	3.12	2.84	3.02	2.92	3.64	8.41
Q	8.26	8.37	4.90	3.12	2.84	3.00	2.91	3.64	8.46
K	8.23	8.32	4.91	3.13	2.84	3.00	2.90	3.65	8.45
S	8.26	8.34	4.89	3.12	2.86	3.03	2.91	3.61	8.44
R	8.25	8.33	4.92	3.13	2.84	2.99	2.91	3.66	8.45
A	8.17	8.27	4.85	3.11	2.87	3.03	2.90	3.60	8.42
pivaloyl-A	8.05	8.18	4.85	3.11	2.86	3.02	2.88	3.60	8.41
E (pH 3)	8.28	8.38	4.89	3.12	2.85	3.02	2.93	3.64	8.46
E (pH 6.5)	8.29	8.40	4.88	3.11	2.85	3.03	2.92	3.62	8.45
M	8.22	8.30	4.92	3.13	2.84	3.01	2.91	3.65	8.47
L	8.19	8.23	4.91	3.11	2.84	3.00	2.90	3.65	8.46
D (pH 6.5)	8.13	8.25	4.86	3.11	2.86	3.03	2.91	3.62	8.45
D (pH 2.5)	8.15	8.19	4.88	3.14	2.85	3.02	2.92	3.66	8.46
N	8.17	8.24	4.88	3.13	2.84	3.00	2.91	3.65	8.46
H (pH 8.5)	n.d.	n.d.	4.85	3.06	2.82	n.d.	n.d.	3.58	n.d.
F	8.15	8.22	4.60	3.05	2.78	2.98	2.90	3.46	8.40
Y	8.09	8.17	4.77	3.02	2.75	3.15	2.89	3.32	8.38
W	7.72	8.05	4.62	2.94	2.66	3.18	2.82	3.19	8.31
phosphoS (pH 5.5)	8.34	8.40	4.87	3.10	2.89	3.09	2.91	3.55	8.42
phosphoS (pH 8.5)	8.36	8.43	4.87	3.10	2.91	3.12	2.89	3.53	8.40

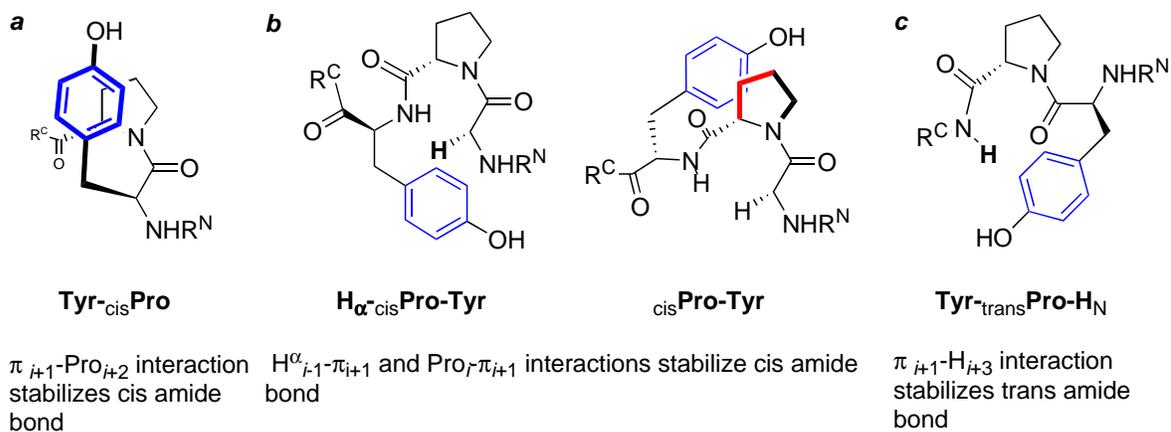
**Table 1.4.** Selected NMR data for the peptides XYPN. Peptides are listed in order of increasing  $K_{trans/cis}$ . Tyr  $H_{\beta A}$  was assigned as the most downfield tyrosine  $\beta$ -proton resonance. Pro  $H_{\delta B}$  was assigned as the most upfield proline  $\delta$ -proton resonance.

TYPZ, Z =	$K_{\text{trans/cis}}$	$\Delta G_{\text{trans/cis}}$	$\Delta\Delta G$	$Y_{\text{cis}} \text{}^3J_{\text{HN}\alpha}$
		kcal mol <sup>-1</sup>	kcal mol <sup>-1</sup>	Hz
F	2.0	- 0.41	+ 0.22	5.7
W	2.6	- 0.56	+ 0.06	5.5
P	2.6	- 0.56	+ 0.06	8.7
N	3.0	- 0.65	- 0.03	6.3
A	3.1	- 0.67	- 0.04	7.7
D (pH 2.5)	3.1	- 0.67	- 0.04	6.8
V	3.2	- 0.68	- 0.06	n.d.
D (pH 6.5)	3.8	- 0.79	- 0.16	n.d.
G	4.3	- 0.86	- 0.23	7.1
D-Ala	5.3	- 0.98	- 0.36	6.6

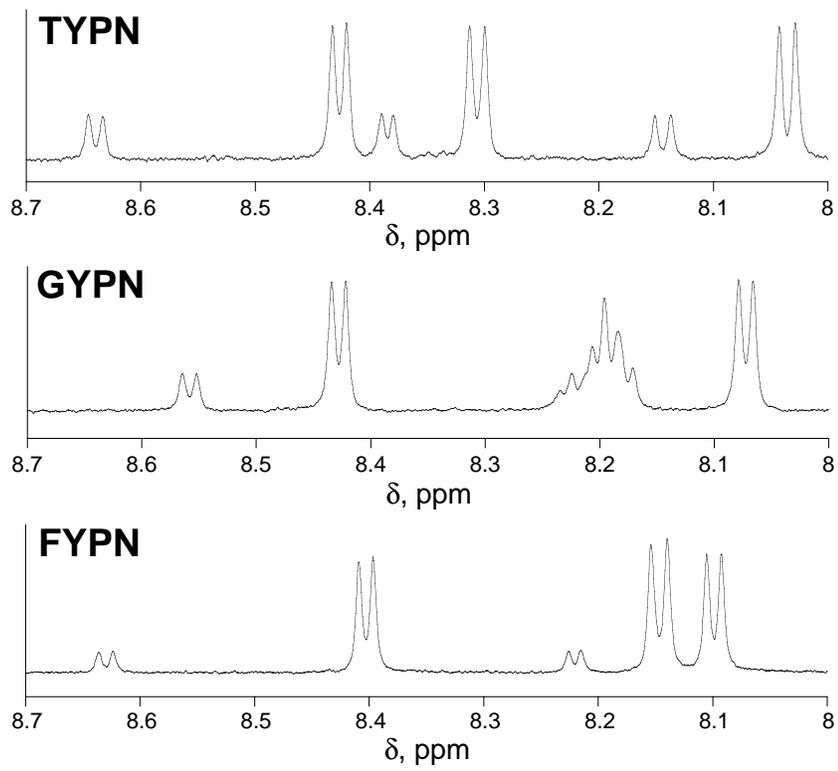
**Table 1.5.** NMR-derived data for the peptides TYPZ. Experiments were conducted at 23 °C in 5 mM phosphate buffer, 250 mM NaCl, pH 4.0 or as indicated. Peptides are listed in order of increasing  $K_{\text{trans/cis}}$ .  $\Delta\Delta G = \Delta G_{\text{TYPZ}} - \Delta G_{\text{HYPN}}$ . n.d. = not determined because of spectral overlap.  $K_{\text{trans/cis}}$  for TYPP reflects only species with a trans Pro-Pro amide bond. (TYPG, TYPV, and D-Ala data from Aaron E. Lee.)

TYPZ, Z =	$\delta, H_N$	$\delta, H_\alpha$	$\delta, H_\alpha$	$\delta, H_{\beta A}$	$\Delta\delta_{\beta A-\beta B}$	$\Delta\delta_{\beta A-\beta B}$	$\delta, H_\delta$	$\delta, H_{\delta B}$	$\delta, H_\alpha$
	Z <sub>trans</sub>	Y <sub>trans</sub>	Y <sub>cis</sub>	Y <sub>trans</sub>	Y <sub>trans</sub>	Y <sub>cis</sub>	Y <sub>trans</sub>	P <sub>trans</sub>	P <sub>cis</sub>
F	7.84	4.86	4.18	3.00	0.17	0.13	7.14	3.47	3.56
W	7.24	4.71	4.17	2.42	0.00	0.11	7.01	3.22	3.51
P	n.a.	4.87	4.43	3.12	0.30	0.00	7.17	3.64	3.91
N	8.43	4.90	4.55	3.12	0.26	0.11	7.17	3.62	3.84
A	8.35	4.89	4.59	3.11	0.26	0.00	7.17	3.59	3.91
D (pH 2.5)	8.43	4.91	4.72	3.12	0.25	0.09	7.18	3.62	3.85
V	8.21	4.90	4.60	3.08	0.19	0.08	7.16	3.58	3.99
D (pH 6.5)	8.33	4.90	4.58	3.14	0.28	n.d.	7.18	3.65	3.85
G	8.21	4.90	4.60	3.08	0.19	0.08	7.16	3.58	3.84
D-Ala	8.58	4.89	4.56	3.08	0.22	0.10	7.15	3.56	3.60

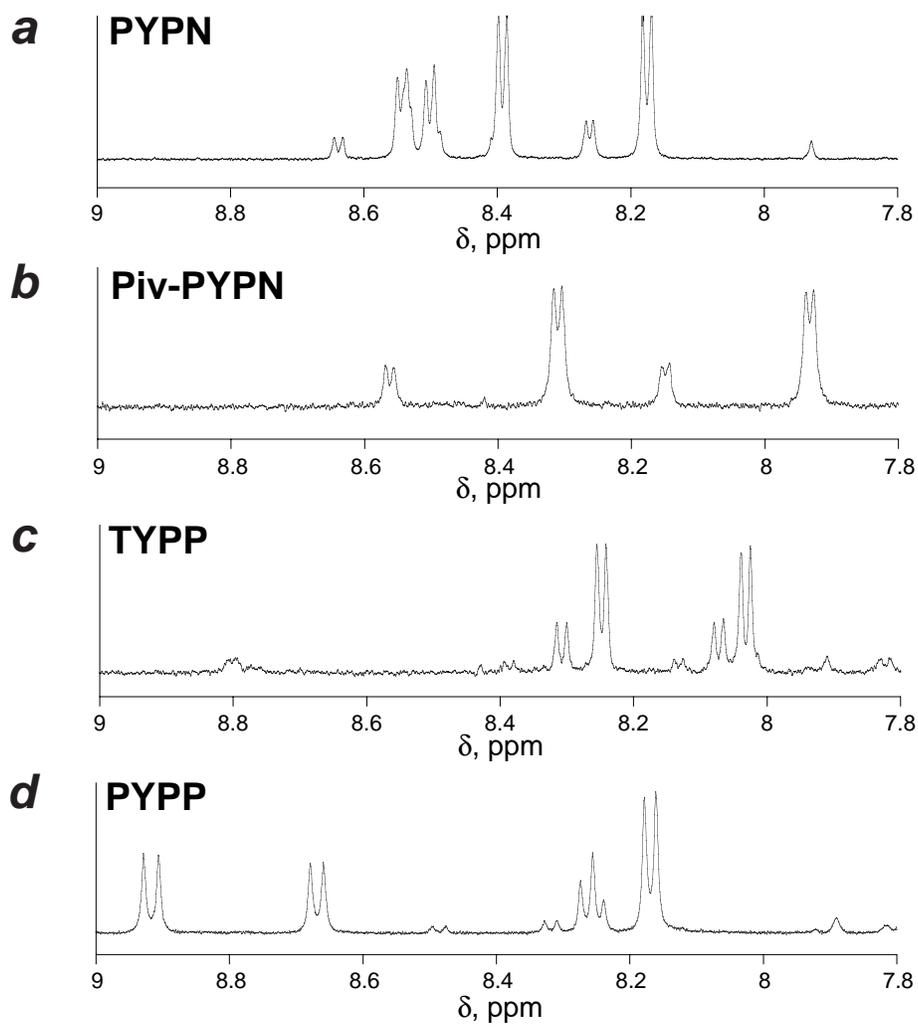
**Table 1.6.** Selected NMR data for the peptides TYPZ. Peptides are listed in order of increasing  $K_{\text{trans/cis}}$ . Tyr  $H_{\beta A}$  was assigned as the most downfield tyrosine  $\beta$ -proton resonance.  $\Delta\delta_{\beta A-\beta B}$  indicates the measured difference in chemical shifts between the diastereotopic  $\beta$ -protons ( $\delta(H_{\beta A}) - \delta(H_{\beta B})$ ) of Tyr in the indicated isomeric state. Pro  $H_{\delta B}$  was assigned as the most upfield proline  $\delta$ -proton resonance. (TYPG, TYPV, and D-Ala data from Aaron E. Lee.)



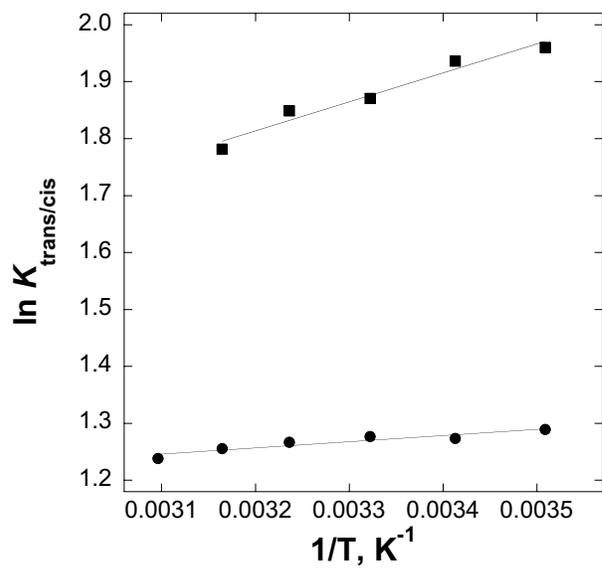
**Figure 1.1.** Aromatic-backbone interactions stabilizing cis (aromatic-prolyl: a, b) or trans (aromatic-amide: c) amide bonds. (a) Tyr<sup>-cis</sup>Pro; (b) H<sub>α</sub>(Xaa)<sup>-cis</sup>Pro-Tyr and cisPro-Tyr; (c) Tyr<sup>-trans</sup>Pro-H<sub>N</sub>(Zaa). Type c interactions may include any residue at the *i*+2 position in place of Pro.



**Figure 1.2.** Amide region of the NMR spectra of the peptides TYPN, GYPN and FYPN.



**Figure 1.3.** Amide region of the NMR spectra of the peptides PYPN, pivaloyl-PYPN, TYPP, and PYPP.



**Figure 1.4.** Van't Hoff plot of the temperature dependence of  $K_{\text{trans/cis}}$  for the peptides AYPN (circles) and YYPN (squares).

## References

1. Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. *Biochemistry* 2002, 41, 14637-14644.
2. Reimer, U.; Fischer, G. *Biophys Chem* 2002, 96, 203-212.
3. Dugave, C.; Demange, L. *Chem Rev* 2003, 103, 2475-2532.
4. Andreotti, A. H. *Biochemistry* 2003, 42, 9515-9524.
5. Lorenzen, S.; Peters, B.; Goede, A.; Preissner, R.; Frommel, C. *Proteins Struct Funct Bioinform* 2005, 58, 589-595.
6. Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* 1975, 14, 4953-4963.
7. Fischer, G. *Angew Chem Int Ed* 1994, 33, 1415-1436.
8. Schmid, F. X. *Adv Protein Chem* 2002, 59, 243-282.
9. Houry, W. A.; Scheraga, H. A. *Biochemistry* 1996, 35, 11719-11733.
10. Yaffe, M. B.; Schutkowski, M.; Shen, M. H.; Zhou, X. Z.; Stukenberg, P. T.; Rahfeld, J. U.; Xu, J.; Kuang, J.; Kirschner, M. W.; Fischer, G.; Cantley, L. C.; Lu, K. *P. Science* 1997, 278, 1957-1960.
11. Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J Mol Biol* 1990, 214, 253-260.
12. MacArthur, M. W.; Thornton, J. M. *J Mol Biol* 1991, 218, 397-412.
13. Milner-White, E. J.; Bell, L. H.; Maccallum, P. H. *J Mol Biol* 1992, 228, 725-734.
14. Pal, D.; Chakrabarti, P. *J Mol Biol* 1999, 294, 271-288.
15. Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Sci* 2001, 10, 2627-2632.
16. Hinderaker, M. P.; Raines, R. T. *Protein Sci* 2003, 12, 1188-1194.
17. Ramachandran, G. N.; Mitra, A. K. *J Mol Biol* 1976, 107, 85-92.
18. Grathwohl, C.; Wüthrich, K. *Biopolymers* 1976, 15, 2025-2041.
19. Beausoleil, E.; Lubell, W. D. *J Am Chem Soc* 1996, 118, 12902-12908.
20. Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J Mol Biol* 1998, 279, 449-460.
21. Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J Mol Biol* 1988, 201, 161-200.
22. Yao, J.; Feher, V. A.; Espejo, B. F.; Reymond, M. T.; Wright, P. E.; Dyson, H. J. *J Mol Biol* 1994, 243, 736-753.
23. Yao, J.; Dyson, H. J.; Wright, P. E. *J Mol Biol* 1994, 243, 754-766.
24. Hutchinson, E. G.; Thornton, J. M. *Protein Sci* 1994, 3, 2207-2216.
25. Wu, W.-J.; Raleigh, D. P. *Biopolymers* 1998, 45, 381-394.
26. Halab, L.; Lubell, W. D. *J Am Chem Soc* 2002, 124, 2474-2484.
27. Bhattacharyya, R.; Chakrabarti, P. *J Mol Biol* 2003, 331, 925-940.
28. Oka, M.; Montelione, G. T.; Scheraga, H. A. *J Am Chem Soc* 1984, 106, 7959-7969.

29. Montelione, G. T.; Arnold, E.; Meinwald, Y. C.; Stimson, E. R.; Denton, J. B.; Huang, S. G.; Clardy, J.; Scheraga, H. A. *J Am Chem Soc* 1984, 106, 7946-7958.
30. Kemmink, J.; Vanmierlo, C. P. M.; Scheek, R. M.; Creighton, T. E. *J Mol Biol* 1993, 230, 312-322.
31. Kemmink, J.; Creighton, T. E. *J Mol Biol* 1995, 245, 251-260.
32. Nardi, F.; Kemmink, J.; Sattler, M.; Wade, R. C. *J Biomol NMR* 2000, 17, 63-77.
33. Taylor, C. M.; Hardre, R.; Edwards, P. J. B.; Park, J. H. *Org Lett* 2003, 5, 4413-4416.
34. Taylor, C. M.; Hardre, R.; Edwards, P. J. B. *J Org Chem* 2005, 70, 1306-1315.
35. Thomas, K. M.; Naduthambi, D.; Tririya, G.; Zondlo, N. J. *Org Lett* 2005, 7, 2397-2400.
36. Richardson, J. S. *Adv Protein Chem* 1981, 34, 167-339.
37. Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv Protein Chem* 1985, 37, 1-109.
38. Wilmot, C. M.; Thornton, J. M. *J Mol Biol* 1988, 203, 221-232.
39. Wilmot, C. M.; Thornton, J. M. *Protein Eng* 1990, 3, 479-493.
40. Müller, G.; Gurrath, M.; Kurz, M.; Kessler, H. *Proteins Struct Funct Gen* 1993, 15, 235-251.
41. Tugarinov, V.; Zvi, A.; Levy, R.; Amglistter, J. *Nature Struct Biol* 1999, 6, 331-335.
42. Wittelsberger, A.; Keller, M.; Scarpellino, L.; Patiny, L.; Acha-Orbea, H.; Mutter, M. *Angew Chem Int Ed* 2000, 39, 1111-+.
43. Keller, M.; Boissard, C.; Patiny, L.; Chung, N. N.; Lemieux, C.; Mutter, M.; Schiller, P. W. *J Med Chem* 2001, 44, 3896-3903.
44. Bélec, L.; Slaninova, J.; Lubell, W. D. *J Med Chem* 2000, 43, 1448-1455.
45. Beneken, J.; Tu, J. C.; Xiao, B.; Nuriya, M.; Yuan, J. P.; Worley, P. F.; Leahy, D. J. *Neuron* 2000, 26, 143-154.
46. Lockett, S.; Santiago Garcia, R.; Barker, J. J.; Konarev, A. V.; Shewry, P. R.; Clarke, A. R.; Brady, R. L. *J Mol Biol* 1999, 290, 525-533.
47. Brauer, A. B. E.; Domingo, G. J.; Cooke, R. M.; Matthews, S. J.; Leatherbarrow, R. J. *Biochemistry* 2002, 41, 10608-10615.
48. McNulty, J. C.; Jackson, P. J.; Thompson, D. A.; Chai, B. X.; Gantz, I.; Barsh, G. S.; Dawson, P. E.; Millhauser, G. L. *J Mol Biol* 2005, 346, 1059-1070.
49. Marshall, G. R. *Biopolymers* 2001, 60, 246-277.
50. Tyndall, J. D. A.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. *Chem Rev* 2005, 105, 793-826.
51. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* 2000, 289, 739-745.
52. Galardy, R. E.; Alger, J. R.; Liakopouloukyriakides, M. *Int J Peptide Protein Res* 1982, 19, 123-132.
53. Delaney, N. G.; Madison, V. *Int J Peptide Protein Res* 1982, 19, 543-548.
54. Delaney, N. G.; Madison, V. *J Am Chem Soc* 1982, 104, 6635-6641.

55. Yu, K. L.; Johnson, R. L. *J Org Chem* 1987, 52, 2051-2059.
56. Zabrocki, J.; Smith, G. D.; Dunbar, J. B.; Ijima, H.; Marshall, G. R. *J Am Chem Soc* 1988, 110, 5875-5880.
57. Elseviers, M.; Van Der Auwera, L.; Pepermans, H.; Tourwé, D.; Van Binst, G. *Biochem Biophys Res Comm* 1988, 154, 515-521.
58. Smith, G. D.; Zabrocki, J.; Flak, T. A.; Marshall, G. R. *Int J Peptide Protein Res* 1991, 37, 191-197.
59. Sukumaran, D. K.; Prorok, M.; Lawrence, D. S. *J Am Chem Soc* 1991, 113, 706-707.
60. Huang, Z. W.; He, Y. B.; Raynor, K.; Tallent, M.; Reisine, T.; Goodman, M. J. *J Am Chem Soc* 1992, 114, 9390-9401.
61. Paul, P. K. C.; Burney, P. A.; Campbell, M. M.; Osguthorpe, D. J. *Bioorg Med Chem Lett* 1992, 2, 141-144.
62. Abell, A. D.; Hoult, D. A.; Jamieson, E. J. *Tetrahedron Lett* 1992, 33, 5831-5832.
63. Zabrocki, J.; Dunbar, J. B.; Marshall, K. W.; Toth, M. V.; Marshall, G. R. *J Org Chem* 1992, 57, 202-209.
64. Lecoq, A.; Boussard, G.; Marraud, M.; Aubry, A. *Biopolymers* 1993, 33, 1051-1059.
65. Brady, S. F.; Paleveda, W. J.; Arison, B. H.; Saperstein, R.; Brady, E. J.; Raynor, K.; Reisine, T.; Veber, D. F.; Freidinger, R. M. *Tetrahedron* 1993, 49, 3449-3466.
66. Boteju, L. W.; Hruby, V. J. *Tetrahedron Lett* 1993, 34, 1757-1760.
67. Magaard, V. W.; Sanchez, R. M.; Bean, J. W.; Moore, M. L. *Tetrahedron Lett* 1993, 34, 381-384.
68. Cumberbatch, S.; North, M.; Zagotto, G. *Tetrahedron* 1993, 49, 9049-9066.
69. Dumas, J. P.; Germanas, J. P. *Tetrahedron Lett* 1994, 35, 1493-1496.
70. Gramberg, D.; Robinson, J. A. *Tetrahedron Lett* 1994, 35, 861-864.
71. Boros, L. G.; Decorte, B.; Gimi, R. H.; Welch, J. T.; Wu, Y.; Handschumacher, R. E. *Tetrahedron Lett* 1994, 35, 6033-6036.
72. Beusen, D. D.; Zabrocki, J.; Slomczynska, U.; Head, R. D.; Kao, J. L. F.; Marshall, G. R. *Biopolymers* 1995, 36, 181-200.
73. Curran, T. P.; McEnaney, P. M. *Tetrahedron Lett* 1995, 36, 191-194.
74. Weisshoff, H.; Frost, K.; Brandt, W.; Henklein, P.; Mugge, C.; Frommel, C. *FEBS Lett* 1995, 372, 203-209.
75. Tourwé, D. *Lett Pept Sci* 1995, 2, 182-186.
76. Gramberg, D.; Weber, C.; Beeli, R.; Inglis, J.; Bruns, C.; Robinson, J. A. *Helv Chim Acta* 1995, 78, 1588-1606.
77. Osapay, G.; Zhu, Q.; Shao, H.; Chadha, R. K.; Goodman, M. *Int J Peptide Protein Res* 1995, 46, 290-301.
78. Lenman, M. M.; Ingham, S. L.; Gani, D. *Chem Comm* 1996, 85-87.
79. Kim, K. H.; Dumas, J. P.; Germanas, J. P. *J Org Chem* 1996, 61, 3138-3144.

80. Kern, D.; Schutkowski, M.; Drakenberg, T. *J Am Chem Soc* 1997, 119, 8403-8408.
81. Kim, K.; Germanas, J. P. *J Org Chem* 1997, 62, 2847-2852.
82. Kim, K.; Germanas, J. P. *J Org Chem* 1997, 62, 2853-2860.
83. Dumy, P.; Keller, M.; Ryan, D. E.; Rohwedder, B.; Wöhr, T.; Mutter, M. *J Am Chem Soc* 1997, 119, 918-925.
84. Abell, A. D.; Foulds, G. J. *J Chem Soc Perkin 1* 1997, 2475-2482.
85. Didierjean, C.; DelDuca, V.; Benedetti, E.; Aubry, A.; Zouikri, M.; Marraud, M.; Boussard, G. *J Peptide Res* 1997, 50, 451-457.
86. Keller, M.; Sager, C.; Dumy, P.; Schutkowski, M.; Fischer, G. S.; Mutter, M. *J Am Chem Soc* 1998, 120, 2714-2720.
87. Tran, T.-A.; Mattern, R.-H.; Afargan, M.; Amitay, O.; Ziv, O.; Morgan, B. A.; Taylor, J. E.; Hoyer, D.; Goodman, M. *J Med Chem* 1998, 41, 2679-2685.
88. Takeuchi, Y.; Marshall, G. R. *J Am Chem Soc* 1998, 120, 5363-5372.
89. Zouikri, M.; Vicherat, A.; Aubry, A.; Marraud, M.; Boussard, G. *J Peptide Res* 1998, 52, 19-26.
90. Halab, L.; Lubell, W. D. *J Org Chem* 1999, 64, 3312-3321.
91. An, S. S. A.; Lester, C. C.; Peng, J.-L.; Li, Y.-J.; Rothwarf, D. M.; Welker, E.; Thannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J Am Chem Soc* 1999, 121, 11558-11566.
92. De Luca, L.; Falorni, M.; Giacomelli, G.; Porcheddu, A. *Tetrahedron Lett* 1999, 40, 8701-8704.
93. Derrer, S.; Davies, J. E.; Holmes, A. B. *J Chem Soc Perkin 1* 2000, 2957-2967.
94. Kaczmarek, K.; Kaleta, M.; Chung, N. N.; Schiller, P. W.; Zabrocki, J. *Acta Biochimica Polonica* 2001, 48, 1159-1163.
95. Hoffmann, T.; Lanig, H.; Waibel, R.; Gmeiner, P. *Angew Chem Int Ed* 2001, 40, 3361-3364.
96. Nachman, R. J.; Zabrocki, J.; Olczak, J.; Williams, H. J.; Moyna, G.; Scott, A. I.; Coast, G. M. *Peptides* 2002, 23, 709-716.
97. Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. *J Am Chem Soc* 2002, 124, 8522-8523.
98. Descours, A.; Moehle, K.; Renard, A.; Robinson, J. A. *Chembiochem* 2002, 3, 318-323.
99. Cavelier, F.; Vivet, B.; Martinez, J.; Aubry, A.; Didierjean, C.; Vicherat, A.; Marraud, M. *J Am Chem Soc* 2002, 124, 2917-2923.
100. Hitosuyanagi, Y.; Motegi, S.; Fukaya, H.; Takeya, K. *J Org Chem* 2002, 67, 3266-3271.
101. Kee, S.; Jois, S. D. S. *Current Pharm Design* 2003, 9, 1209-1224.
102. Bruns, K.; Fossen, T.; Wray, V.; Henklein, P.; Tessmer, U.; Schubert, U. *J Biol Chem* 2003, 278, 43188-43201.
103. Chakraborty, T. K.; Ghosh, A.; Kumar, S. K.; Kunwar, A. C. *J Org Chem* 2003, 68, 6459-6462.

104. Boruah, A.; Rao, I. N.; Nandy, J. P.; Kumar, S. K.; Kunwar, A. C.; Iqbal, J. J. *J Org Chem* 2003, 68, 5006-5008.
105. Rombouts, F. J. R.; De Borggraeve, W. M.; Delaere, D.; Froeyen, M.; Toppet, S. M.; Compennolle, F.; Hoornaert, G. J. *Eur J Org Chem* 2003, 1868-1878.
106. Zhang, W. J.; Berglund, A.; Kao, J. L. F.; Couty, J. P.; Gershengom, M. C.; Marshall, G. R. *J Am Chem Soc* 2003, 125, 1221-1235.
107. Harris, P. W. R.; Brimble, M. A.; Gluckman, P. D. *Org Lett* 2003, 5, 1847-1850.
108. Melendez, R. E.; Lubell, W. D. *J Am Chem Soc* 2004, 126, 6759-6764.
109. Che, Y.; Marshall, G. R. *J Org Chem* 2004, 69, 9030-9042.
110. De Borggraeve, W. M.; Verbist, B. M. P.; Rombouts, F. J. R.; Pawar, V. G.; Smets, W. J.; Kamoune, L.; Alen, J.; Van der Eycken, E. V.; Compennolle, F.; Hoornaert, G. J. *Tetrahedron* 2004, 60, 11597-11612.
111. Creighton, C. J.; Du, Y. M.; Reitz, A. B. *Bioorg Med Chem Lett* 2004, 12, 4375-4385.
112. Wang, X. D. J.; Xu, B. L.; Mullins, A. B.; Neiler, F. K.; Etkorn, F. A. *J Am Chem Soc* 2004, 126, 15533-15542.
113. Nachman, R. J.; Kaczmarek, K.; Williams, H. J.; Coast, G. M.; Zabrocki, J. *Biopolymers* 2004, 75, 412-419.
114. Hartley, O.; Gaertner, H.; Wilken, J.; Thompson, D.; Fish, R.; Ramos, A.; Pastore, C.; Dufour, B.; Cerini, F.; Melotti, A.; Heveker, N.; Picard, L.; Alizon, M.; Mosier, D.; Kent, S.; Offord, R. *Proc Natl Acad Sci USA* 2004, 101, 16460-16465.
115. Kaul, R.; Surprenant, S.; Lubell, W. D. *J Org Chem* 2005, 70, 3838-3844.
116. Sukopp, M.; Schwab, R.; Marinelli, L.; Biron, E.; Heller, M.; Várkonyi, E.; Pap, Á.; Novellino, E.; Kéri, G.; Kessler, H. *J Med Chem* 2005, 48, 2916-2926.
117. Guruprasad, K.; Rajkumar, S. *J Biosciences* 2000, 25, 143-156.
118. Balamurugan, B.; Samaya Mohan, K.; Ramesh, J.; Roshan, M. N. A. M.; Sumathi, K.; Sekar, K. *Acta Cryst D* 2005, 61, 634-636.
119. Vuister, G. W.; Bax, A. *J Am Chem Soc* 1993, 115, 7772-7777.
120. Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew Chem Int Ed* 2003, 42, 1210-1250.
121. Bhattacharyya, R.; Samanta, U.; Chakrabarti, P. *Protein Eng* 2002, 15, 91-100.
122. Waters, M. L. *Curr Opin Chem Biol* 2002, 6, 736-741.
123. Meurisse, R.; Brasseur, R.; Thomas, A. *Proteins Struct Funct Gen* 2004, 54, 478-490.
124. Eberhardt, E. S.; Loh, S. N.; Raines, R. T. *Tetrahedron Lett* 1993, 34, 3055-3056.
125. Butterfield, S. M.; Patel, P. R.; Waters, M. L. *J Am Chem Soc* 2002, 124, 9751-9755.
126. Tatko, C. D.; Waters, M. L. *J Am Chem Soc* 2002, 124, 9372-9373.
127. Tatko, C. D.; Waters, M. L. *Org Lett* 2004, 6, 3969-3972.
128. Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.; Obradovic, Z.; Dunker, A. K. *Nucl Acids Res* 2004, 32, 1037-1049.

129. Wong, S. E.; Bernacki, K.; Jacobson, M. *J Phys Chem B* 2005, 109, 5249-5258.
130. Yaffe, M. B.; Rittinger, K.; Volinia, S.; Caron, P. R.; Aitken, A.; Leffers, H.; Gamblin, S. J.; Smerdon, S. J.; Cantley, L. C. *Cell* 1997, 91, 961-971.
131. Rittinger, K.; Budman, J.; Xu, J.; Volinia, S.; Cantley, L. C.; Smerdon, S. J.; Gamblin, S. J.; Yaffe, M. B. *Molecular Cell* 1999, 4, 153-166.
132. Toth, G.; Murphy, R. F.; Lovas, S. *Protein Eng* 2001, 14, 543-547.
133. Toth, G.; Watts, C. R.; Murphy, R. F.; Lovas, S. *Proteins Struct Funct Gen* 2001, 43, 373-381.
134. Toth, G.; Murphy, R. F.; Lovas, S. *J Am Chem Soc* 2001, 123, 11782-11790.
135. Toth, G.; Kover, K. E.; Murphy, R. F.; Lovas, S. *J Phys Chem B* 2004, 108, 9287-9296.
136. Nagarajaram, H. A.; Paul, P. K. C.; Ramanarayanan, K.; Soman, K. V.; Ramakrishnan, C. *Int J Peptide Protein Res* 1992, 40, 383-394.
137. Kopple, K. D.; Schamper, T. J.; Go, A. *J Am Chem Soc* 1974, 96, 2597-2605.
138. Kopple, K. D.; Go, A.; Pilipauskas, D. R. *J Am Chem Soc* 1975, 97, 6830-6838.
139. Kostanek, E. C.; Lipscomb, W. N.; Thiessen, W. E. *J Am Chem Soc* 1979, 101, 834-837.
140. Kostanek, E. C.; Thiessen, W. E.; Schomburg, D.; Lipscomb, W. N. *J Am Chem Soc* 1979, 101, 5811-5815.
141. Gierasch, L. M.; Deber, C. M.; Madison, V.; Niu, C.-H.; Blout, E. R. *Biochemistry* 1981, 20, 4730-4738.
142. Kartha, G.; Bhandry, K. K.; Kopple, K. D.; Go, A.; Zhu, P.-P. *J Am Chem Soc* 1984, 106, 3844-3850.
143. Brown, A. M.; Zondlo, N. J. submitted.

## Chapter 2

### DESIGN OF A LANTHANIDE FINGER PEPTIDE

#### Introduction

Lanthanides are versatile metals that have found wide application in materials, luminescence, biomedical imaging, and catalysis.<sup>1-6</sup> Incorporation of lanthanides within designed proteins provides a functional handle for long-lifetime fluorescence, luminescence resonance energy transfer over large distances, orientation in magnetic fields, and hydrolytic activity.<sup>7-16</sup>

Although lanthanides are not known to exist natively within proteins, lanthanides have long been used as spectroscopic probes of calcium-binding proteins, binding to calcium-binding sites due to similarities in ionic radius and hardness.<sup>17</sup> EF-hand proteins represent the smallest natural calcium-binding proteins, a 70 amino acid motif which binds two calcium ions.<sup>18-20</sup> The EF hand can be reduced, with significant loss in affinity, to a simple metal-binding loop of 12-20 amino acids.<sup>21</sup> EF hand loops, however, contain no controllable secondary structure elements, adopt structures which are significantly dependent on the ionic radius of the incorporated metal, and have been observed to readily dimerize at low micromolar concentrations, thus representing

a problematic scaffold for precise molecular engineering to use lanthanides to probe structure and function of complex protein mixtures.

Despite the utility of lanthanides, few convenient protein-based ligands for lanthanides exist. Indeed, EF hand proteins, including Imperiali's protein design and selection of potent lanthanide-binding tags, represent the only small protein-based lanthanide-binding motif known.<sup>11,12,14</sup> Lanthanide-binding loops have been incorporated by design into globular proteins, but rely on the tertiary structure of the globular protein for stability.<sup>22-24</sup> Moreover, other known calcium-binding proteins are significantly larger than EF hands (usually > 25 kD), making them less attractive as lanthanide-binding motifs.

We sought to develop, via protein design, a small lanthanide-binding motif that consists of regular secondary structure elements and whose structure renders it readily adaptable, via steric and electronic tuning, to many potential functional applications of lanthanides.<sup>25</sup> Among compact metal-binding motifs, zinc fingers have been the subject of considerable study, both for their biological function and as the basis of protein design.<sup>26-28</sup> Zinc fingers are a well-studied platform in protein design, including the design of zinc fingers motifs with high affinity for lead, the design of chimeric zinc fingers that may be used as metal-dependent switches, and the design of zinc finger motifs that do not require metal binding to fold.<sup>26-34</sup> Herein we report the iterative design of lanthanide fingers, a new, general lanthanide-binding motif inspired by zinc fingers.

## Experimental

### Materials

Fmoc-L-amino acids were purchased from Novabiochem (San Diego, CA) or Bachem (San Carlos, CA). Rink amide resin was purchased from Novabiochem. Acetonitrile was purchased from J.T. Baker. O-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) was purchased from Senn Chemicals (San Diego, CA). Lanthanide (III) chlorides were purchased from Acros or Aldrich. All other chemicals were purchased from Aldrich or Acros. Deionized water was purified by a Millipore Synergy 185 water purification system with a Simpapak2 cartridge. All compounds were used as purchased with no additional purification.

### Peptide synthesis, purification and characterization

Peptides were synthesized by standard solid phase peptide synthesis on a CEM Liberty microwave peptide synthesizer or on a Rainin PS3 peptide synthesizer. All peptides were purified to homogeneity using reverse phase HPLC on a Vydac C18 semi-preparative column (10 × 250 mm, 5-10 μm particle, 300 Å pore). Peptide purity was determined by analytical HPLC reinjection on a Rainin Microsorb MV analytical C18 column (4.6 × 250 mm, 3-5 μm particle, 100 Å pore).

The peptides LF1, LF5, LF6, and LF7 were purified using a linear gradient of 0-75% buffer B in buffer A over 60 minutes. The peptides LF2, LF3, and LF4 were

purified using a linear gradient of 0-100% buffer B (20% H<sub>2</sub>O, 80% acetonitrile, 0.05% TFA) in buffer A (98% H<sub>2</sub>O, 2% acetonitrile, 0.06% TFA) over 60 minutes.

Peptide identity was confirmed by ESI-MS (positive ion mode) on an LCQ Advantage (Finnigan) mass spectrometer. Analytical data for the peptides: LF1 ( $t_R$  = 27.6 min, mass expected 2758.0, mass observed 920.4 (M+3H)<sup>3+</sup>); LF2 ( $t_R$  35.9 min, exp. 2840.1, obs. 947.8 (M+3H)<sup>3+</sup>); LF3 ( $t_R$  28.7 min, exp. 3061.4, obs. 1021.6 (M+3H)<sup>3+</sup>); LF4 ( $t_R$  31.3 min, exp. 2832.1, obs. 1016.8 (M+2H)<sup>2+</sup>); LF5 ( $t_R$  40.1 min, exp. 2862.1, obs. 1431.7 (M+2H)<sup>2+</sup>); LF6 ( $t_R$  41.9 min, exp. 2901.18, obs. 1451.0 (M+2H)<sup>2+</sup>); LF7 ( $t_R$  37.5 min, exp. 2831.1, obs. 1416.1 (M+2H)<sup>2+</sup>).

### **Circular dichroism spectroscopy**

Metal binding titrations were conducted by circular dichroism spectroscopy. CD spectra were recorded on an AVIV Circular Spectrometer Model 215 using a Microlab 500 series automatic titrator and a 10 mm cell. The solution was stirred continuously during metal addition. The data were collected at 222 nm with an averaging time of 6 s and stirring time of 6 s for each metal addition by autotitration. The concentration of the peptide was decreased as the titrant was added into the titration solution. Both the concentration of the peptide and its corresponding concentration of metal were recorded and used to calculate final mean residue ellipticity (MRE). Data were the average of at least three independent trials.

Solutions for binding titrations were prepared in aqueous buffer containing 10 mM NaCl and 10 mM HEPES (pH 7.5). The peptide concentration prior to metal

addition was 20  $\mu\text{M}$  for all solutions except the LF4 titrations with  $\text{EuCl}_3$ ,  $\text{DyCl}_3$ ,  $\text{HoCl}_3$ ,  $\text{ErCl}_3$ ,  $\text{GdCl}_3$ ,  $\text{TmCl}_3$ , and  $\text{TbCl}_3$ , where the initial peptide concentration was 10  $\mu\text{M}$ . The metal solutions were prepared in buffer containing 10 mM NaCl, 10 mM HEPES (pH 7.5), and 20 mM metal ( $\text{YbCl}_3$ ,  $\text{LaCl}_3$ ,  $\text{ScCl}_3$ ,  $\text{EuCl}_3$ ,  $\text{DyCl}_3$ ,  $\text{HoCl}_3$ ,  $\text{ErCl}_3$ ,  $\text{GdCl}_3$ ,  $\text{TmCl}_3$ , and  $\text{TbCl}_3$ ). Metal concentrations were determined by titration against a 10 mM EDTA solution using xylenol orange as the indicator in the presence of acetate buffer at pH 5.8. The pH of the metal stock solution was adjusted to pH 5.8. The final pH of the solution containing the metal-peptide complex was pH 7.2.

CD spectra were collected at 296 K Jasco J-810 Spetropolarimeter in a 1 mm cell with 50  $\mu\text{M}$  peptide in 5 mM HEPES buffer (pH 7.5) and 10 mM NaCl. Data are the average of at least three independent trials. Individual spectra were collected every nm with an averaging time of 3 s and three accumulations. Data were background corrected but not smoothed.

### **Calculation of dissociation constants for peptide-metal complexes**

Dissociation constants for peptide-metal complexes were determined by fitting circular dichroism data as a function of lanthanide concentration to equation 1 using a non-linear least squares fitting algorithm (KaleidaGraph version 4.0, Synergy Software), where  $\Theta_{222}$  = observed mean residue ellipticity at 222 nm ( $\Theta_{222}$ ),  $\Theta_o = \Theta_{222}$  of the apo-peptide,  $\Theta_c = \Theta_{222}$  of the peptide-metal complex,  $M_t$  = total metal concentration,  $K_d$  = dissociation constant and  $P_t$  = total peptide concentration. Non-

linear least squares fits were fit to  $K_d$  and  $\Theta_c$ . Dissociation constants were calculated based on at least 3 independent trials.

Equation (1):

$$\Theta_{222} = \Theta_o + \frac{(\Theta_c - \Theta_o) [(M_t + K_d + P_t) - \sqrt{[(M_t + P_t + K_d)^2 - 4(P_t M_t)}]}{2 P_t}$$

### **NMR spectroscopy**

Peptides were dissolved in buffer containing 10 mM NaCl, 90% H<sub>2</sub>O/10% D<sub>2</sub>O, and 100  $\mu$ M TSP. The pH of the solution was adjusted to pH 6.5. Peptide concentrations were 500  $\mu$ M- 1 mM. NMR experiments were conducted in the absence and presence of 1 equivalent of diamagnetic metal Lu<sup>3+</sup> (LuCl<sub>3</sub>) to evaluate changes in structure due to metal-peptide complex formation. NMR spectra were collected on a Bruker AVX 600 MHz NMR spectrometer equipped with a triple-resonance cryoprobe. 1-D spectra and TOCSY spectra were collected with water suppression using a Watergate w5 pulse sequence with gradients (Bruker pulse programs zgpgw5 and mlevgpww5, respectively). NOESY spectra were collected with watergate water suppression and mixing times from 30 ms to 300 ms. <sup>1</sup>H-<sup>15</sup>N HSQC experiments were conducted with watergate water suppression (Bruker pulse program hsqcwf3gpwwg). 1-D spectra were collected with a sweep width of 6009 Hz, 8192 data points, and a relaxation delay of 2 s. TOCSY spectra were acquired with sweep widths of 6009 Hz in  $t_1$  and  $t_2$ , with 600  $\times$  4096 complex data points, 4 scans per  $t_1$  increment, a relaxation delay of 1.5 s, and a TOCSY mixing time of 60

ms. NOESY spectra were acquired with sweep widths of 6009 Hz in  $t_1$  and  $t_2$ ,  $600 \times 2048$  complex data points, 16 scans per  $t_1$  increment, and a relaxation delay of 1.5 s.  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were conducted using the peptides with natural abundance  $^{15}\text{N}$ , with sweep widths of 2432 Hz in  $t_1$  and 5387 Hz in  $t_2$ , with  $128 \times 1024$  complex data complex data points, 128 scans per  $t_1$  increment, and a relaxation delay of 1.8 s.

## Results and Discussion

The zinc finger is a compact metal-binding motif, consisting of an N-terminal  $\beta$ -hairpin and a C-terminal  $\alpha$ -helix connected by a loop, that binds numerous soft metals, including zinc(II), and cobalt(II) (Figure 2.1).<sup>26-28</sup> Moreover, by incorporating additional thiolate ligands, designed artificial zinc fingers may bind lead with high affinity.<sup>35</sup> Alternatively, replacement of the metal-binding residues with hydrophobic residues, combined with additional structural optimization, can lead to metal-independent miniproteins with a zinc finger fold.<sup>29,30,36</sup>

Zinc finger 2 of Zif268 was used as a starting point in the design.<sup>37,38</sup> In a classical  $\text{Cys}_2\text{His}_2$  zinc finger, four side chain ligands, two each from the  $\beta$ -hairpin ( $\text{Cys}_2$ ) and  $\alpha$ -helix ( $\text{His}_2$ ) of the protein, coordinate zinc.<sup>26</sup> Soft thiolate and imidazole ligands poorly coordinate hard lanthanide ions. Therefore, the coordination sphere of the zinc finger was redesigned to introduce residues favoring lanthanide binding. In the  $\alpha$ -helix, both His residues were replaced by the approximately isosteric residue Glu, with the hard carboxylate group designed to bind metal. In the  $\beta$ -hairpin, the Cys

residues were replaced by Asp residues, which are larger than Cys but contain a carboxylate group for tight lanthanide coordination.

Zinc binding by a zinc finger is thermodynamically driven by strong zinc-thiolate bonds. As lanthanide-carboxylate coordination is weaker than zinc-thiolate, a series of additional design elements were incorporated in the peptide sequence to stabilize the lanthanide finger fold (Figure 2.2): (a) solvent-exposed residues within the  $\alpha$ -helix were changed to residues with high  $\alpha$ -helical propensity (Ala, Lys, Leu); (b) solvent-exposed residues in the  $\beta$ -hairpin were changed to Thr to maximize  $\beta$ -sheet propensity and solubility; (c) the loop sequence from the zinc finger consensus peptide CP-1 was incorporated; (d) The N-terminal PF sequence was changed to AY to introduce a Tyr residue for concentration determination and the Pro changed to Ala to prevent cis-trans isomerization;<sup>39</sup> and (e) a type II'  $\beta$ -turn was introduced into the  $\beta$ -hairpin, as was used by Imperiali in the design of metal-free zinc finger proteins.<sup>29,30</sup> Peptides lacking these modifications bound lanthanides very poorly in water, exhibiting evidence of metal binding only in aqueous-organic media (M. Ye and N. J. Zondlo, data not shown).

The resultant peptide, LF1, was analyzed for metal binding by circular dichroism (Figure 2.3a, Table 2.1). LF1 bound  $\text{Yb}^{3+}$  modestly in competitive aqueous environment (10 mM HEPES, 10 mM NaCl, Figure 2.9). The LF1- $\text{Yb}^{3+}$  complex showed an increase in  $\alpha$ -helicity compared to the apo peptide, consistent with metal-induced folding of the  $\alpha$ -helix, as is seen in zinc fingers, although the LF1- $\text{Yb}^{3+}$  complex was less  $\alpha$ -helical than is observed for a well-folded zinc finger.

Encouragingly, LF1 showed size selectivity in lanthanide binding, binding the smaller  $\text{Yb}^{3+}$  significantly better than the larger  $\text{Eu}^{3+}$  or  $\text{La}^{3+}$ , consistent with formation of a specific complex rather than non-specific binding (Figure 2.7, 2.8)

In order to increase metal affinity, LF2 was designed to incorporate an  $\alpha$ -helical N-cap motif.<sup>40</sup> The most frequently observed residues from each position of an N<sup>n</sup>→N4 Big Box N-capping motif were employed, with Phe the N<sup>n</sup> residue. The required dihedral angle of a Big Box N-cap are almost identical to those observed at these residues in zinc fingers. In small proteins, helix capping can provide critical structural stabilization. Interestingly, in zinc fingers helix N-capping by target DNA or RNA is an important component to binding affinity and specificity.

If a zinc finger fold is adopted by lanthanide fingers, an increase in  $\alpha$ -helicity due to the N-cap should increase overall metal affinity, by preorganizing the metal-binding Glu residues. LF2 exhibited a significant increase in lanthanide affinity (Figure 2.3c, Table 2.1), binding  $\text{Eu}^{3+}$  with  $K_d = 205$  nM (Figure 2.11). Notably, LF2 was significantly more  $\alpha$ -helical than LF1, both in the absence and in the presence of metal, consistent with the importance of N-cap motifs in stabilizing  $\alpha$ -helices. The CD spectrum of the LF2-lanthanide complexes were consistent with previously observed zinc finger CD spectra, exhibiting 30%  $\alpha$ -helix (Figure 2.10, 2.12).

In order to further increase metal affinity, LF3 was designed to incorporate an  $\alpha$ -helical C-cap motif.<sup>40</sup> The most frequently observed residues from a Schellman motif were employed. In zinc fingers, DNA-induced C-capping by the TGEKP linker

sequence between successive fingers significantly increases the stability of zinc finger-DNA complexes.<sup>41</sup>

LF3 was significantly more  $\alpha$ -helical than LF2 in the absence of metal (Figure 2.3c). Metal titration experiments indicated that LF3 bound the smaller lanthanide Yb<sup>3+</sup> (ionic radius 1.01 Å) with 2-fold improvement in affinity over the LF2-Yb<sup>3+</sup> complex (Figure 2.13, Table 2.1). In contrast, complexes of LF3 with larger lanthanides (Eu<sup>3+</sup> (ionic radius 1.09 Å, Figure 2.14) or La<sup>3+</sup> (1.17 Å, Figure 2.15)) displayed affinities comparable to those observed with LF2, suggesting that the metal-binding site was well-packed and sterically congested, as would be expected for replacement of smaller ligands (Cys) with larger ligands (Asp) in a metal-binding site. Consistent with this interpretation, LF3 bound the smaller metal Sc<sup>3+</sup> (ionic radius 0.89 Å) with significantly greater affinity ( $K_d = 20 \mu\text{M}$ ) than was observed for binding to any lanthanide.

The design of LF4 removed 2 residues from the  $\beta$ -hairpin sequence in order to provide a larger binding site for lanthanides. This approach was previously used by Imperiali to accommodate the sterically congested phenanthoalanine (Fen) residue in her metal-free zinc finger design, resulting in a more open structure.<sup>29,30</sup>

LF4 is a general lanthanide-binding peptide (Figure 2.3, Figure 2.4, and Table 2.1). The highest affinities to LF4 were observed for the lanthanides Tb<sup>3+</sup> (Figure 2.18) Ho<sup>3+</sup> (Figure 2.17) and Er<sup>3+</sup> (ionic radius 1.03-1.06 Å, Figure 2.16), with good affinities also observed for binding to smaller and larger lanthanides (Figure 2.19; 2.20; 2.21; 2.22; 2.23). Metal-bound LF4 exhibited a characteristic CD spectrum for a

metal-bound zinc finger, with over 50%  $\alpha$  helicity, as expected based on a zinc finger structure.

Metal binding to LF4 was further examined by NMR (Figure 2.5). Comparison of the TOCSY and  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of LF4 in the absence of metal and in the presence of the diamagnetic lanthanide  $\text{Lu}^{3+}$  reveals a large induction of structure throughout the peptide on addition of  $\text{Lu}^{3+}$ : nearly every resonance undergoes a significant chemical shift change on metal binding. These data are consistent with the adoption of an ordered structure involving folding of residues from the entire peptide upon metal binding, and are consistent with metal-binding involving residues from both the  $\alpha$ -helical and  $\beta$ -hairpin secondary structure elements, as designed.

We sought to design an expressible lanthanide finger motif, in which the D-Pro residue is replaced by a canonical amino acid. In LF5, D-Pro-Ser was replaced with a Thr-Ile sequence, incorporating  $\beta$ -branched residues observed in zinc fingers. LF5 is a general, expressible lanthanide-binding sequence, binding smaller and larger lanthanides with mid-micromolar affinity (Figure 2.24; 2.25; 2.26). Interestingly, apo-LF5 is less  $\alpha$ -helical than apo-LF4, despite no changes in the residues of the  $\alpha$ -helix, suggesting that tertiary structure is an important contributor stabilizing the LF4  $\alpha$ -helix. Similarly, metal-bound LF5 was less  $\alpha$ -helical than metal-bound LF4, consistent with less effective packing and an important role for turn geometry in complex formation, as was also observed by Imperiali.<sup>29,30</sup>

Long-lifetime terbium(III) luminescence is one of the most broadly utilized applications of lanthanides.<sup>1,2,11,16</sup> Terbium luminescence requires both a donor

chromophores located within 10 Å of the metal and low water accessibility.

Tryptophan is the most effective terbium luminescence donor among the canonical amino acids. In Zif268, the conserved phenylalanine residue (Phe11, numbering from the N-terminus) at the C-terminus of the  $\beta$ -hairpin comprises the core of the protein folding nucleus. Phe11 projects toward the metal-binding site, with less than 4 Å separation between the aromatic ring and the zinc center, stabilizing metal binding as the secondary coordination sphere which interacts with and orients the metal-binding residues.<sup>42</sup> Therefore, we expected that replacement of Phe11 with Trp should generate a protein with Trp located near the metal-binding site. The additional size of Trp was expected to be well-tolerated in the more open structure expected in LF4 and LF5.<sup>29,30</sup>

Phe from LF5 was replaced with Trp to generate LF6. In this peptide, the Trp indole side chain should be positioned near the metal-binding site, permitting tryptophan to serve as a donor for lanthanide luminescence. Indeed, LF6 shows strong terbium luminescence, indicating that Trp must be located near the metal-binding site and consistent with LF6-Tb<sup>3+</sup> adopting a zinc finger structure (Figure 2.6; 2.31). Circular dichroism and binding constant determinations (Figure 2.3h, Table 2.1) indicated that LF6 bound lanthanides with comparable or better affinity than LF5 (Figure 2.27; 2.28; 2.29; 2.30). In addition, apo and metal-bound LF6 were more  $\alpha$ -helical than LF5, consistent with Trp effectively substituting for Phe in the hydrophobic core of the lanthanide finger.

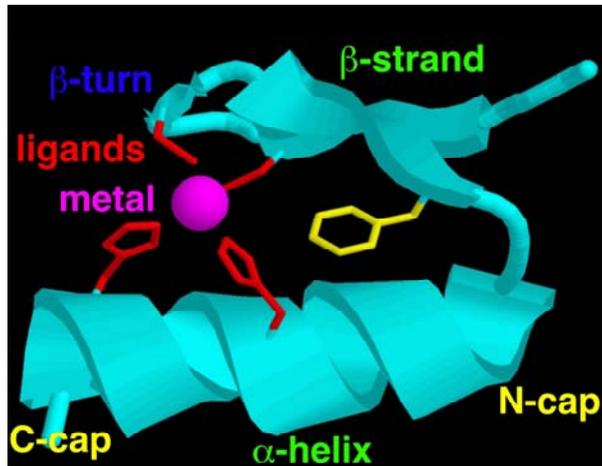
Lanthanides generally exhibit 7-9 liganding atoms in their coordination spheres. The observation of terbium luminescence indicates that the terbium metal has relatively low water accessibility (0- 2 water molecules bound). The observed terbium luminescence suggests that 5 or more liganding atoms are provided by the peptide, therefore providing further evidence that the metal site is coordinated by multiple ligands from both the  $\alpha$ -helical and  $\beta$ -hairpin portion of the peptide.

The properties of lanthanides are dependent on the local ligand environment and are tunable by changes in the ligand sphere. In order to examine whether lanthanide fingers are electronically tunable, we synthesized the peptide LF7, in which one of the liganding Asp residues from LF4 was replaced by the less electronic rich residue Asn (Figure 2.2). Metal binding of LF4 was examined (Figure 2.3i, Table 2.1) by circular dichroism and revealed that LF7 bound lanthanides only modestly worse than LF4 (Figure 2.32; 2.33; 2.34). These data are consistent with the ability of the -3 charge in the ligand environment to effectively bind the +3-charged metal and of the ability of the amide group of the Asn side chain to potentially act as a ligand via the carbonyl.

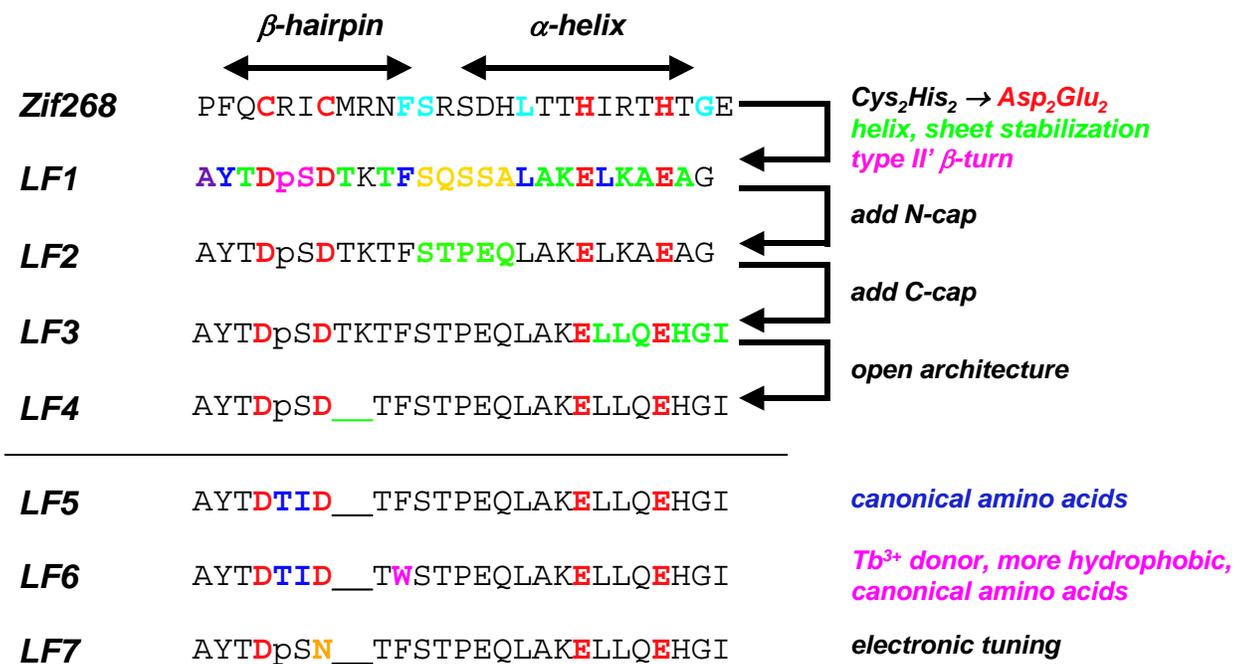
## **Conclusions**

We have developed, via iterative protein design, a new lanthanide-binding motif, lanthanide fingers. The design employed the following principles: (a) reengineering a zinc finger motif to bind lanthanides by changing the ligands to hard

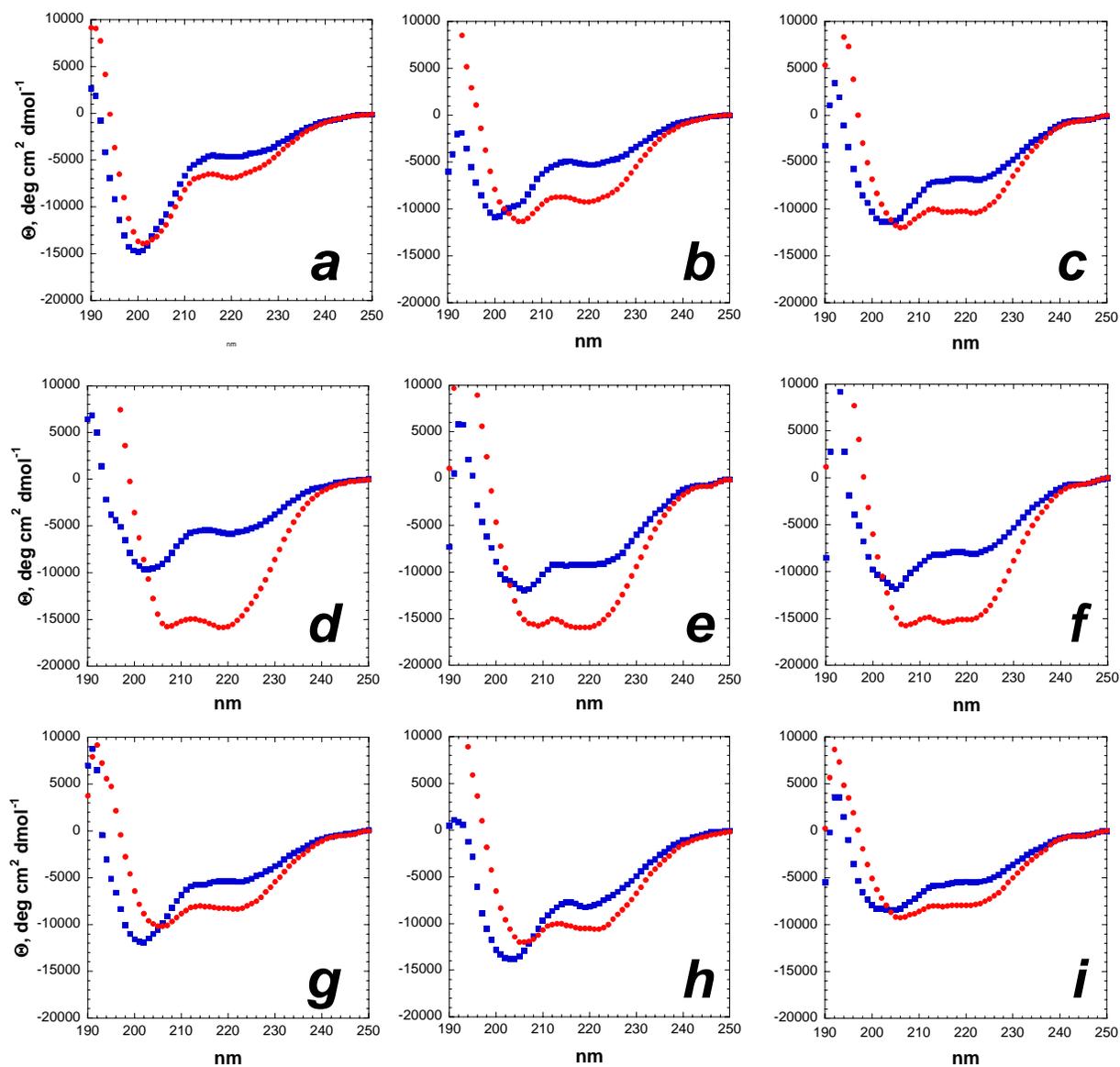
carboxylate ligands (Asp<sub>2</sub>Glu<sub>2</sub>) commonly observed as ligands for lanthanides; (b) optimize folding of the  $\alpha$ -helix by incorporation of optimized N-cap and C-cap motifs; (c) inclusion of residues with high  $\beta$ -sheet propensity in the  $\beta$ -hairpin and high  $\alpha$ -helix propensity in the  $\alpha$ -helix; and (d) the incorporation of an open environment to generate a general lanthanide-binding motif. The design involved optimization of nearly all residues of the protein, with LF4 retaining only 4 residues from the zinc finger Zif268, which served as the inspiration of the protein design. Lanthanide fingers are a new, general, protein-based lanthanide-binding motif that should find broad application in the development of new functional proteins based on the diverse chemical properties of lanthanides.



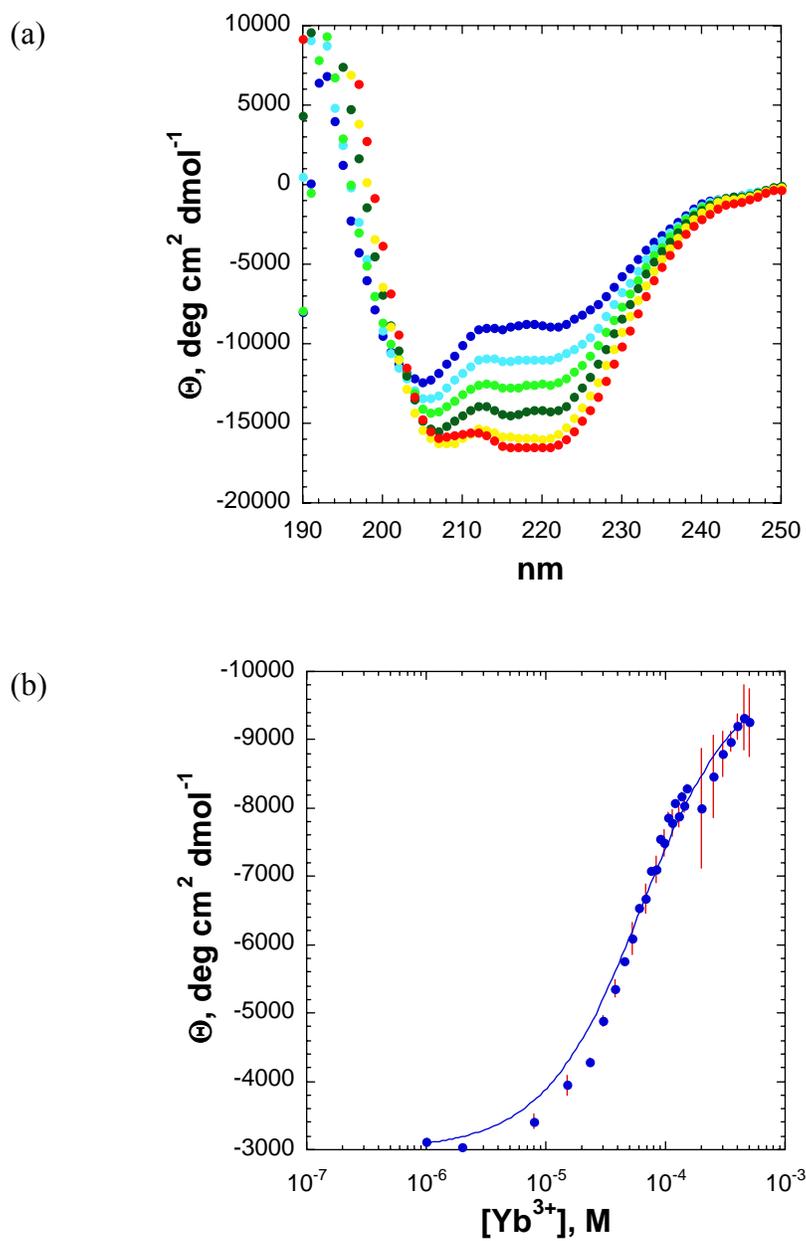
**Figure 2.1.** Structural schematic of a zinc finger (finger 2 of Zif268, pdb 1zaa), with the key structural elements labeled.



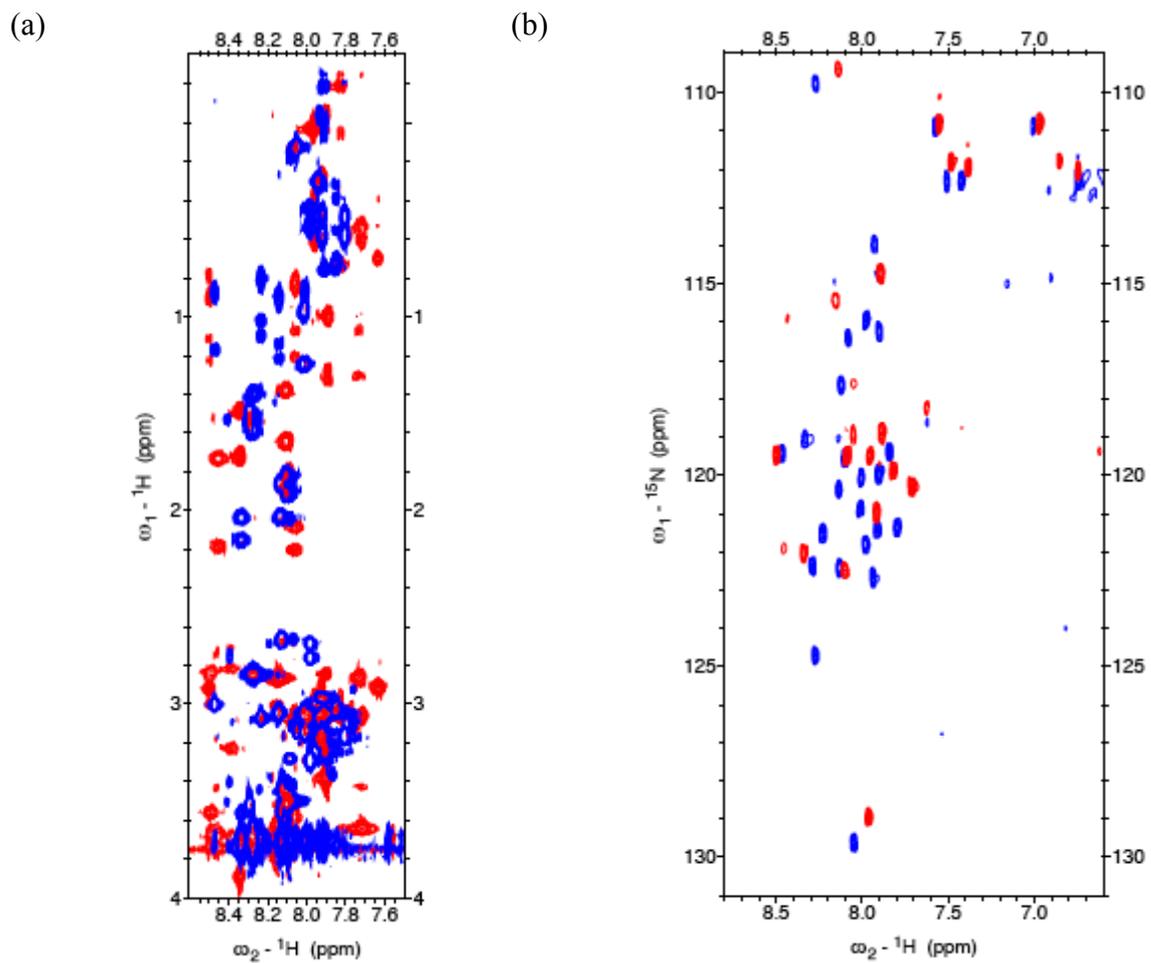
**Figure 2.2.** Iterative design of lanthanide fingers. Sequences of finger 2 of Zif268 and of the designed peptides. Metal-binding residues are indicated in red boldface. Modified residues from the previous peptide generation are indicated by color. Lowercase p indicates D-proline. All peptides were acetylated on the N-terminus and contained C-terminal amides.



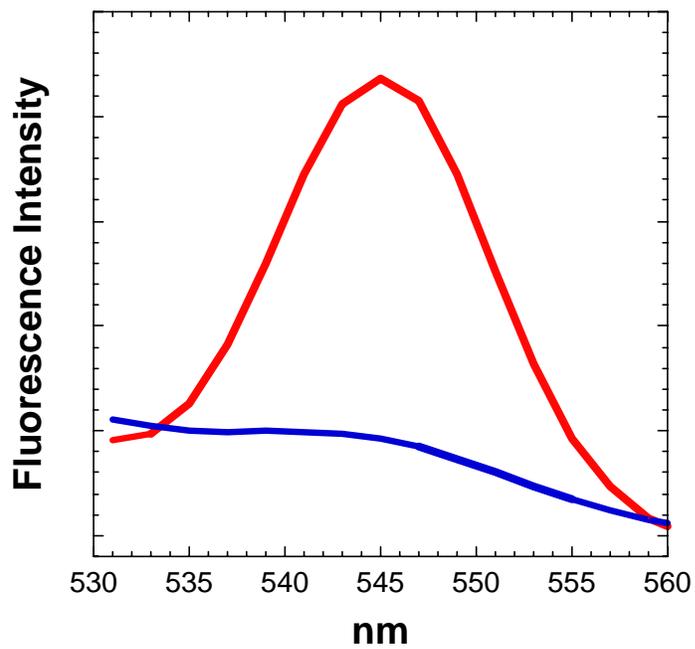
**Figure 2.3.** CD spectra of lanthanide fingers. CD spectra are shown for the peptide (50  $\mu\text{M}$ ) in the absence of lanthanide (blue squares) and in the presence of saturating lanthanide (red circles). (a) LF1  $\pm$  5 mM  $\text{Yb}^{3+}$ ; (b) LF2  $\pm$  800  $\mu\text{M}$   $\text{Yb}^{3+}$ ; (c) LF3  $\pm$  500  $\mu\text{M}$   $\text{Yb}^{3+}$ ; (d) LF4  $\pm$  200  $\mu\text{M}$   $\text{Yb}^{3+}$ ; (e) LF4  $\pm$  100  $\mu\text{M}$   $\text{Tb}^{3+}$ ; (f) LF4  $\pm$  400  $\mu\text{M}$   $\text{La}^{3+}$ ; (g) LF5  $\pm$  400  $\mu\text{M}$   $\text{Yb}^{3+}$ ; (h) LF6  $\pm$  300  $\mu\text{M}$   $\text{Yb}^{3+}$ ; (i) LF7  $\pm$  800  $\mu\text{M}$   $\text{Yb}^{3+}$ .



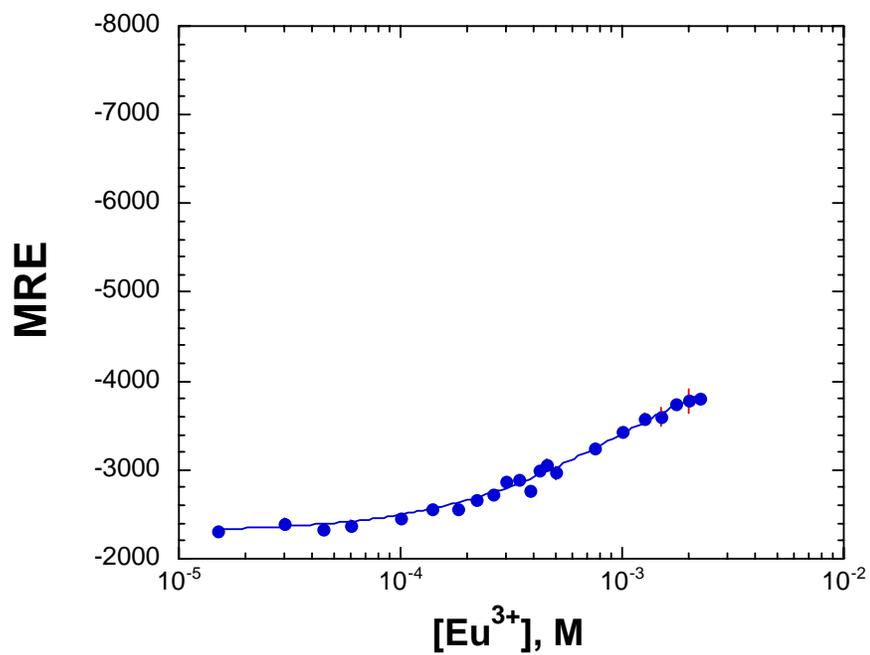
**Figure 2.4.** Metal titrations of  $\text{Gd}^{3+}$ -binding peptides (LF4). (a) CD spectra of  $50 \mu\text{M}$  LF4 in the presence of  $0 \mu\text{M}$  (blue),  $12.5 \mu\text{M}$  (cyan),  $25 \mu\text{M}$  (light green),  $50 \mu\text{M}$  (dark green),  $100 \mu\text{M}$  (yellow) and  $200 \mu\text{M}$  (ref)  $\text{Gd}^{3+}$ ; (b) full titration of LF4- $\text{Yb}^{3+}$ , with the curve fit to the binding equation (1).



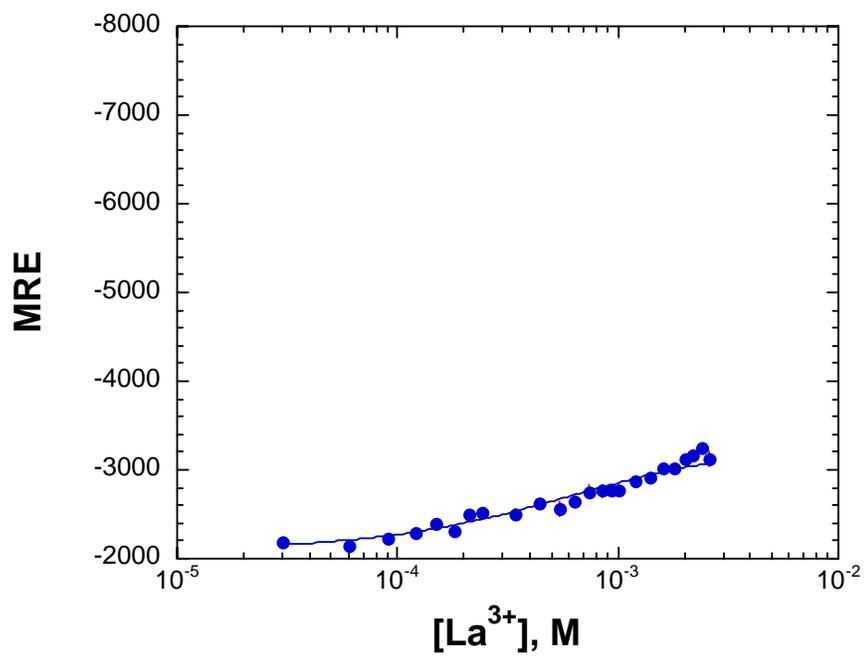
**Figure 2.5.** Comparative NMR spectra of LF4 in the absence (cyan) and presence (red) of  $\text{Lu}^{3+}$ . (a) Fingerprint region of the TOCSY spectra. (b)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra.



**Figure 2.6.** Visible fluorescence spectra of 10  $\mu\text{M}$  LF6 in the absence (blue) and presence (red) of 90  $\mu\text{M}$   $\text{Tb}^{3+}$ .

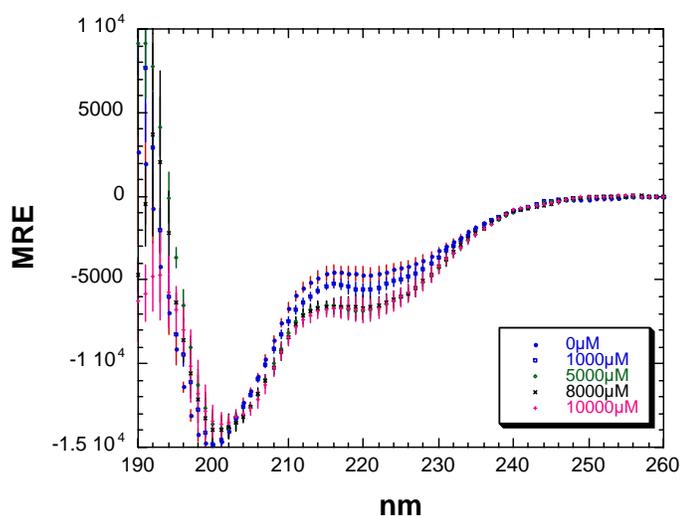


**Figure 2.7.** Full titration of LF1- $\text{Eu}^{3+}$ , with the curve fit to the binding equation (1).

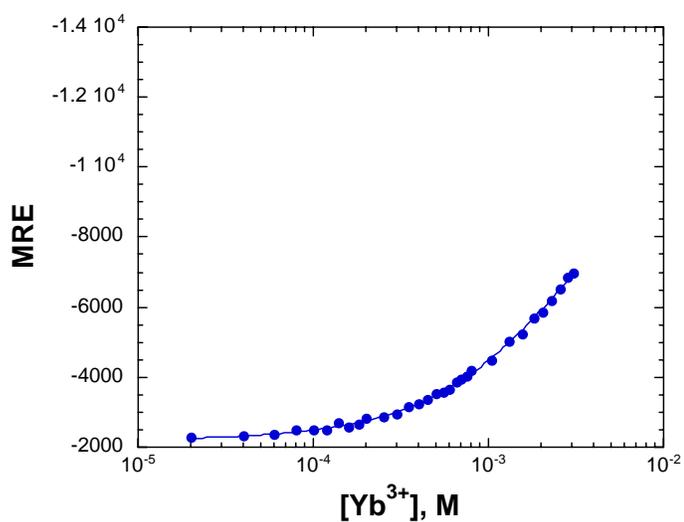


**Figure 2.8.** Full titration of LF1-La<sup>3+</sup>, with the curve fit to the binding equation (1).

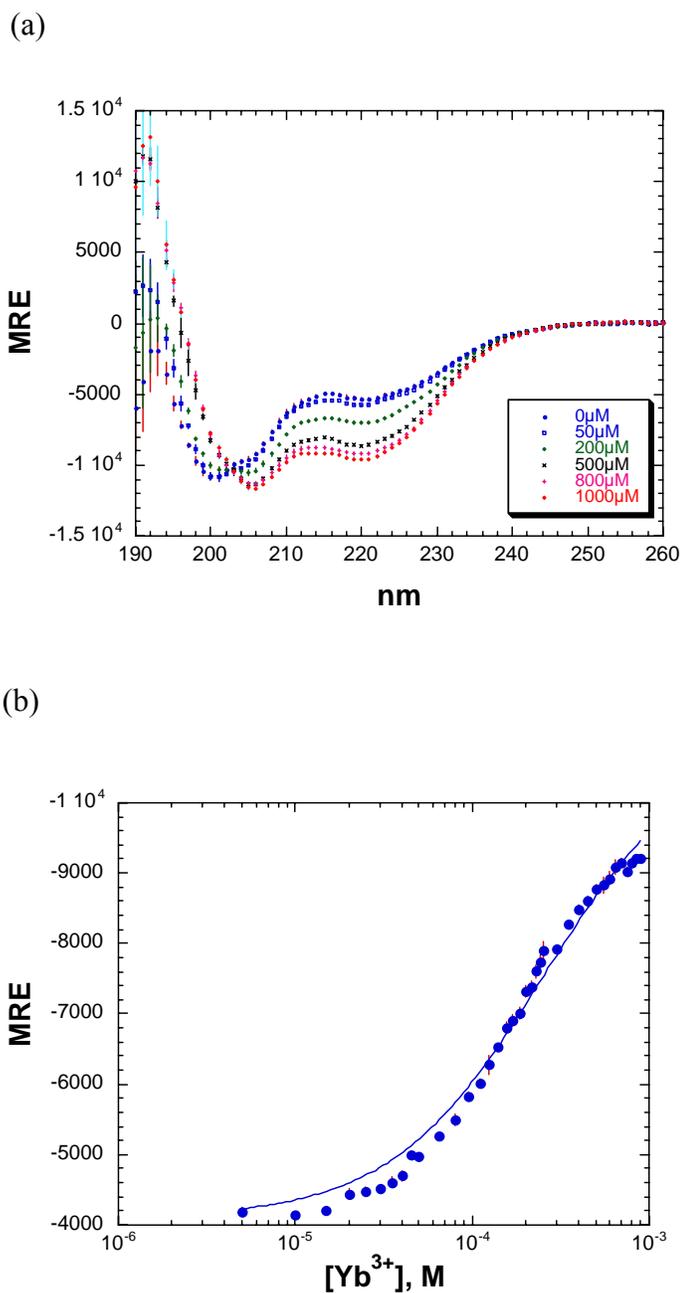
(a)



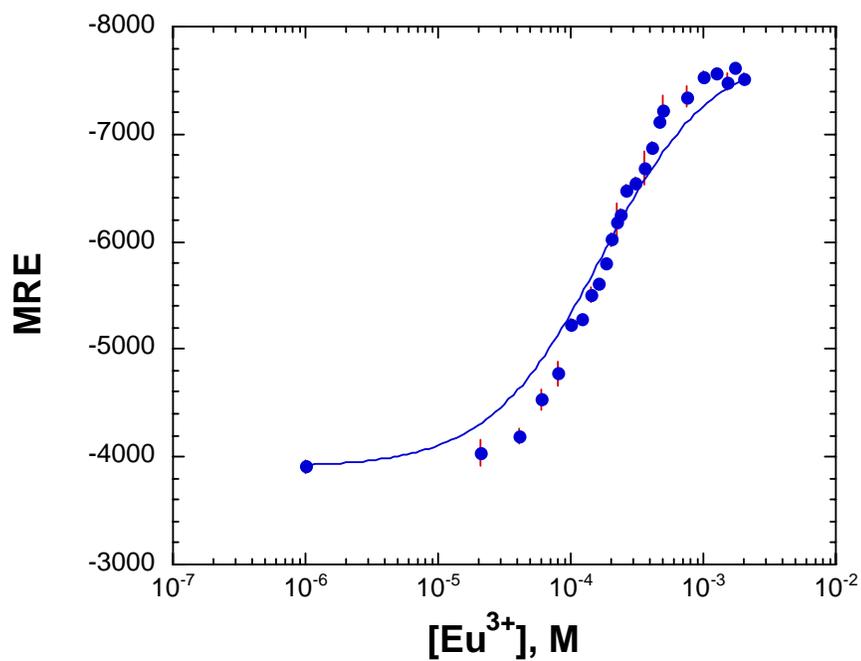
(b)



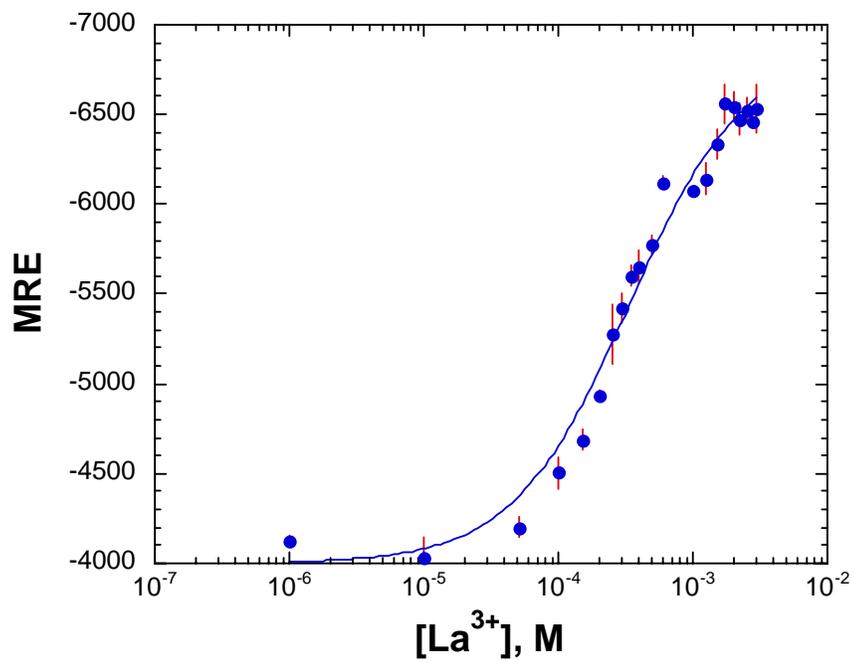
**Figure 2.9.** Metal titrations of Yb<sup>3+</sup>-LF1 binding. (a) CD spectra of 50 μM LF1 in the presence of Yb<sup>3+</sup>. Concentration of Yb<sup>3+</sup> is shown in the legend; (b) full titration of LF1-Yb<sup>3+</sup>, with the curve fit to the binding equation (1).



**Figure 2.10.** Metal titrations of  $\text{Yb}^{3+}$ - LF2 binding. (a) CD spectra of 50  $\mu\text{M}$  LF2 in the presence of  $\text{Yb}^{3+}$ . Concentration of  $\text{Yb}^{3+}$  is shown in the legend; (b) full titration of LF2- $\text{Yb}^{3+}$ , with the curve fit to the binding equation (1).

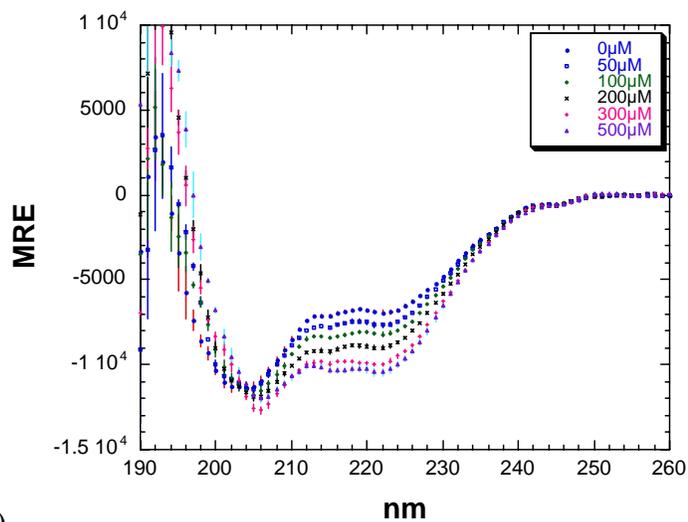


**Figure 2.11.** Full titration of LF2- $Eu^{3+}$ , with the curve fit to the binding equation (1).

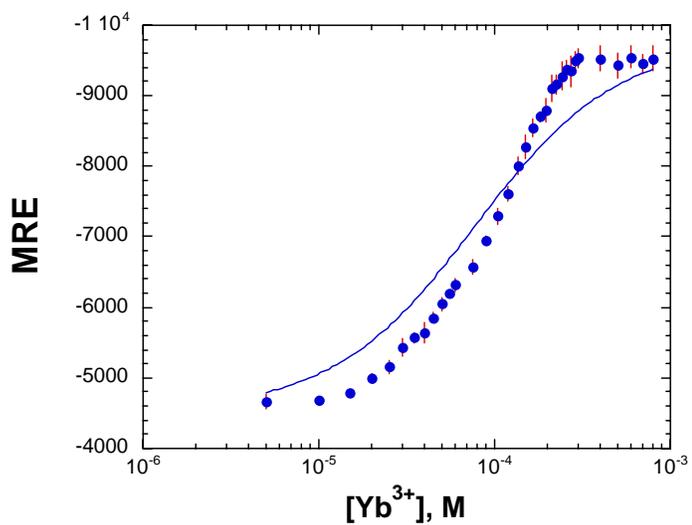


**Figure 2.12.** Full titration of LF2-La<sup>3+</sup>, with the curve fit to the binding equation (1).

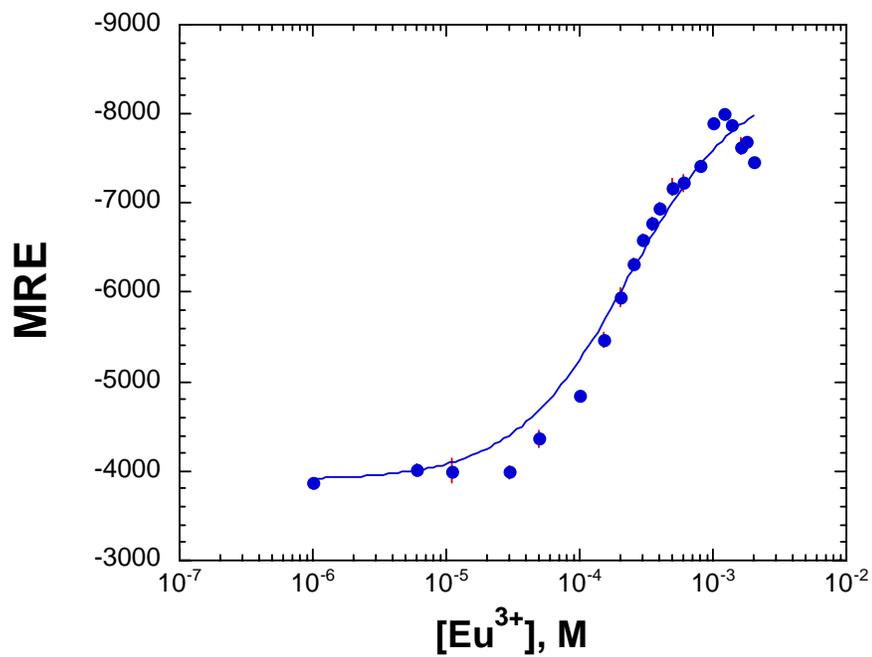
(a)



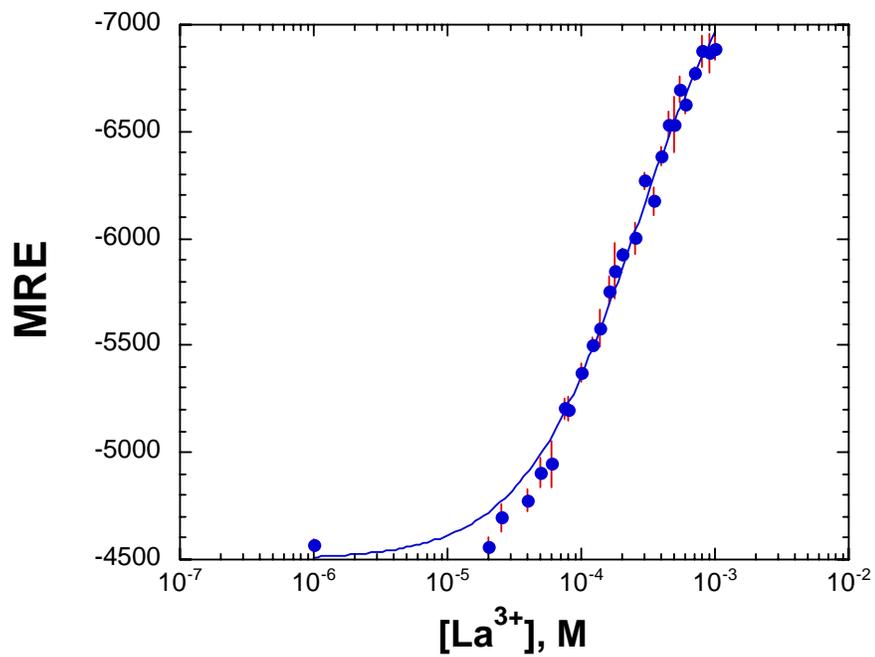
(b)



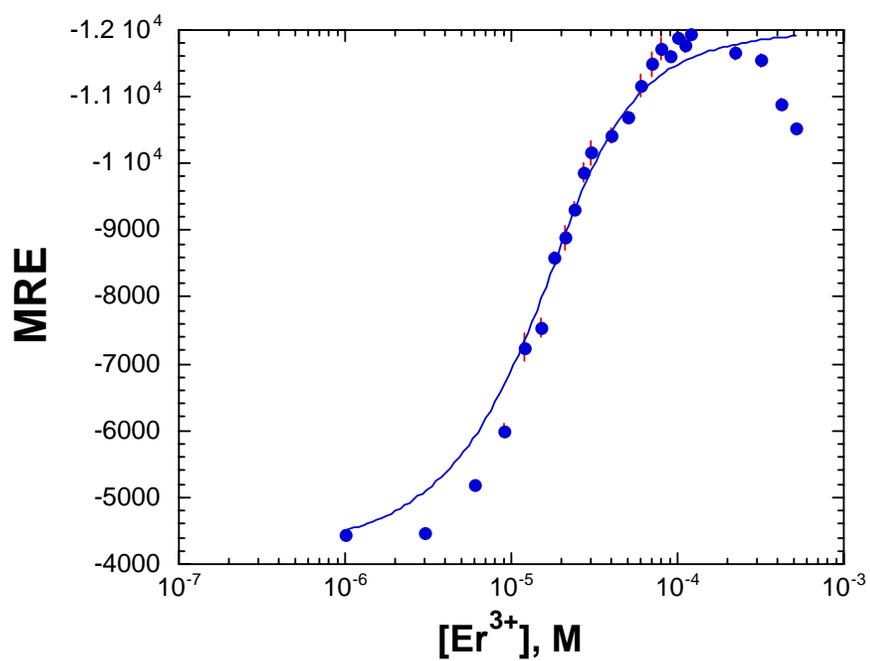
**Figure 2.13.** Metal titrations of Yb<sup>3+</sup>- LF3 binding. (a) CD spectra of 50 μM LF3 in the presence of Yb<sup>3+</sup>. Concentration of Yb<sup>3+</sup> is shown in the legend; (b) full titration of LF3-Yb<sup>3+</sup>, with the curve fit to the binding equation (1).



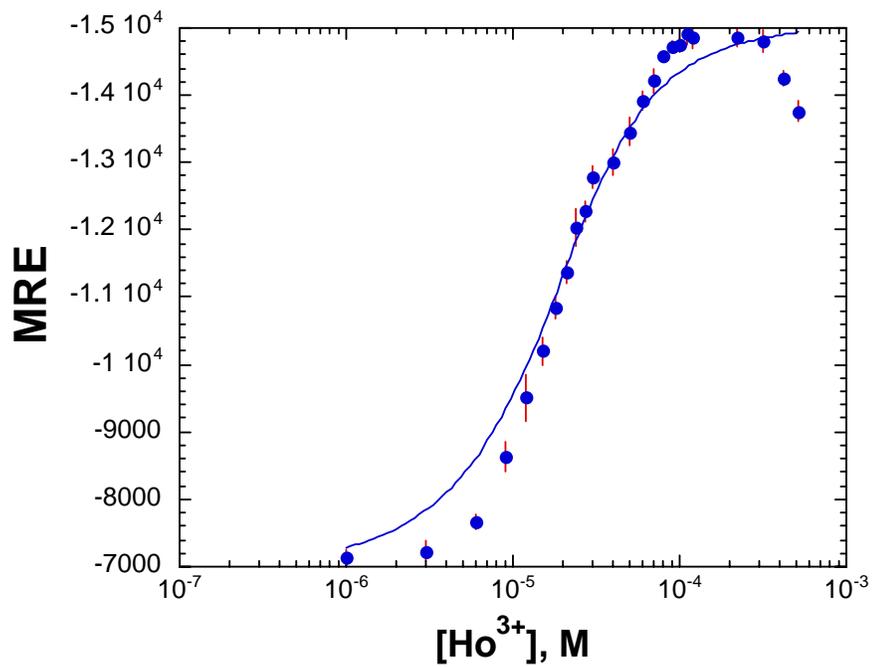
**Figure 2.14.** Full titration of LF3- $\text{Eu}^{3+}$ , with the curve fit to the binding equation (1).



**Figure 2.15.** Full titration of LF3- $La^{3+}$ , with the curve fit to the binding equation (1).

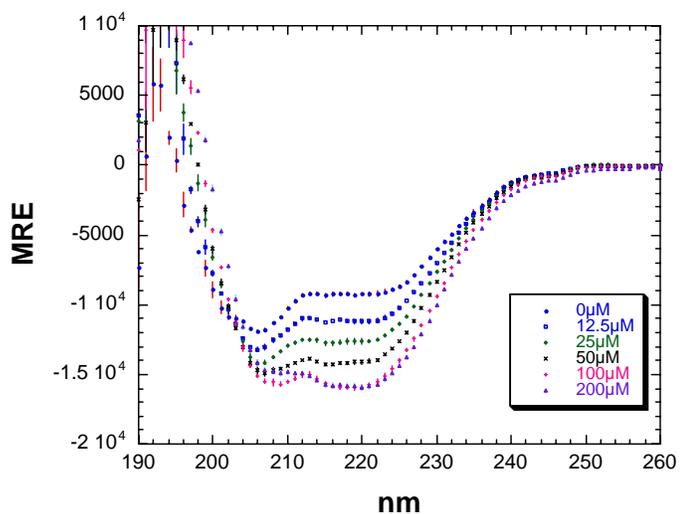


**Figure 2.16.** Full titration of LF4-Er<sup>3+</sup>, with the curve fit to the binding equation (1).

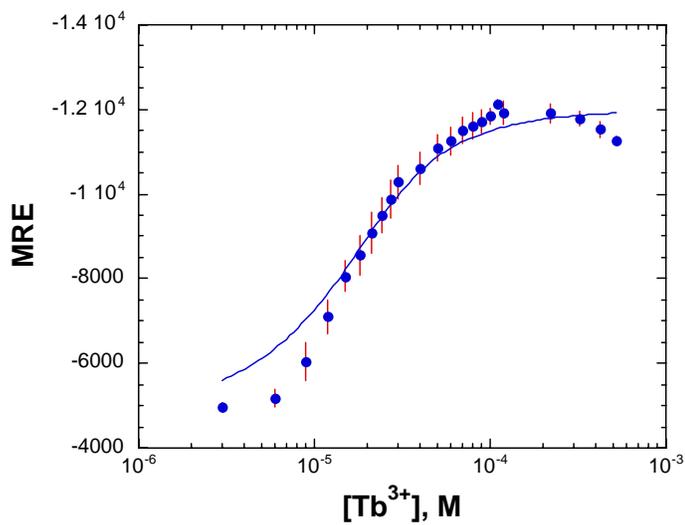


**Figure 2.17.** Full titration of LF4- $Ho^{3+}$ , with the curve fit to the binding equation(1).

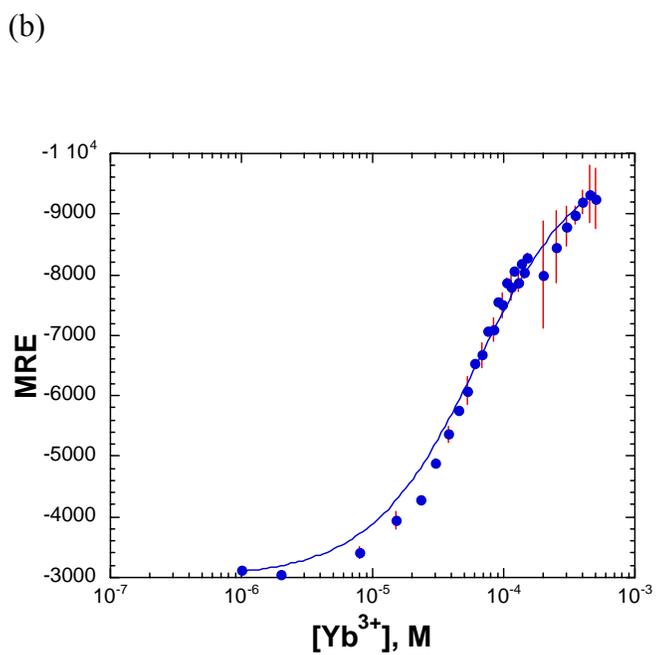
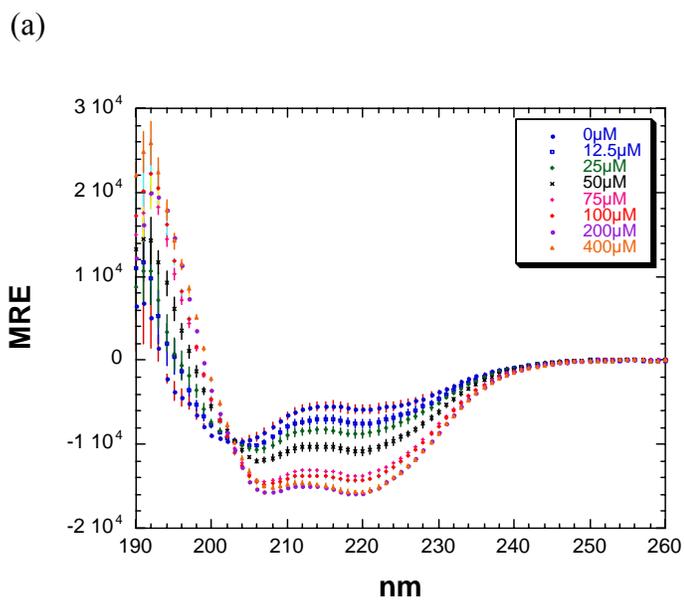
(a)



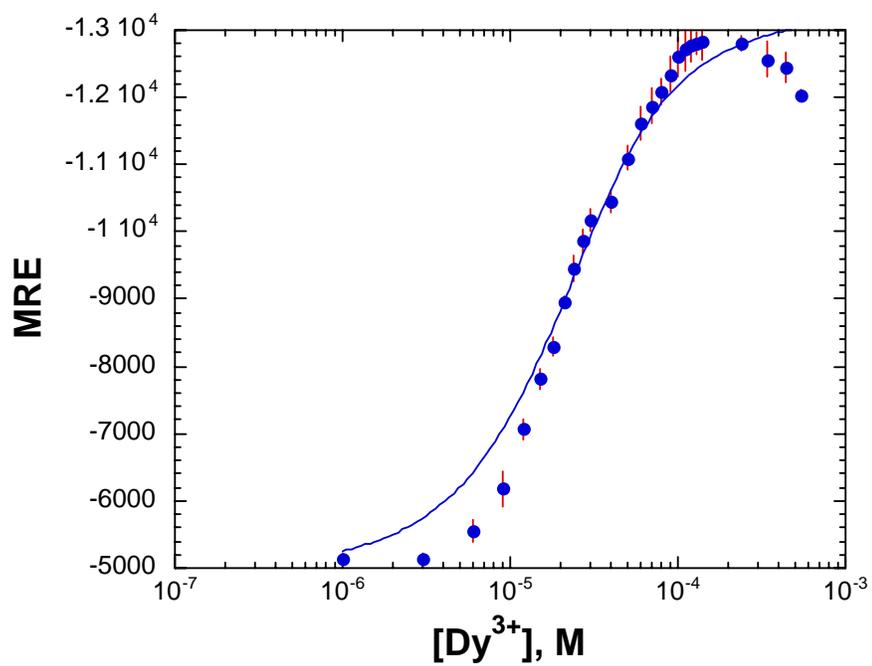
(b)



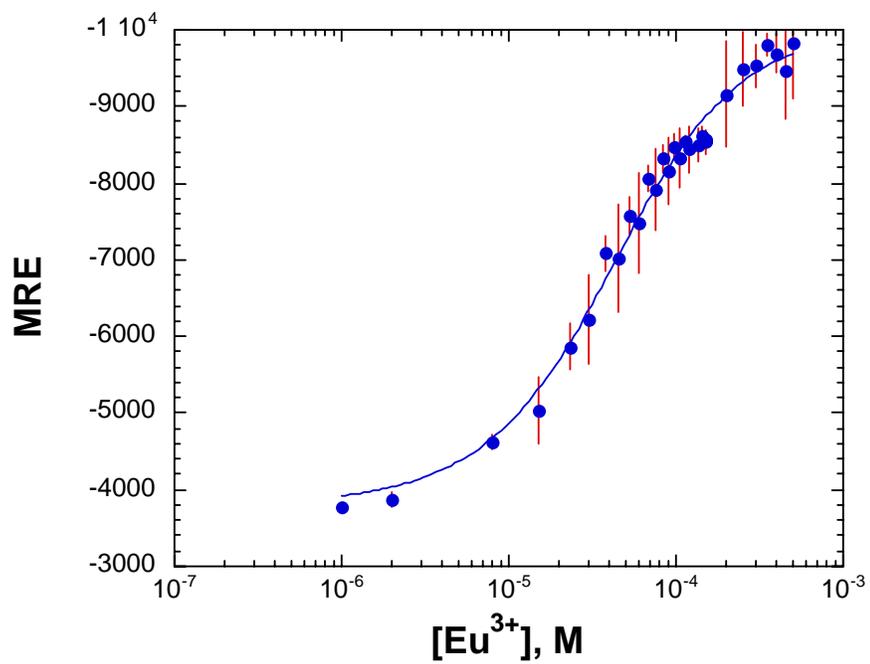
**Figure 2.18.** Metal titrations of  $\text{Tb}^{3+}$  - LF4 binding. (a) CD spectra of 50  $\mu\text{M}$  LF4 in the presence of  $\text{Tb}^{3+}$ . Concentration of  $\text{Tb}^{3+}$  is shown in the legend; (b) full titration of LF4- $\text{Tb}^{3+}$ , with the curve fit to the binding equation (1)



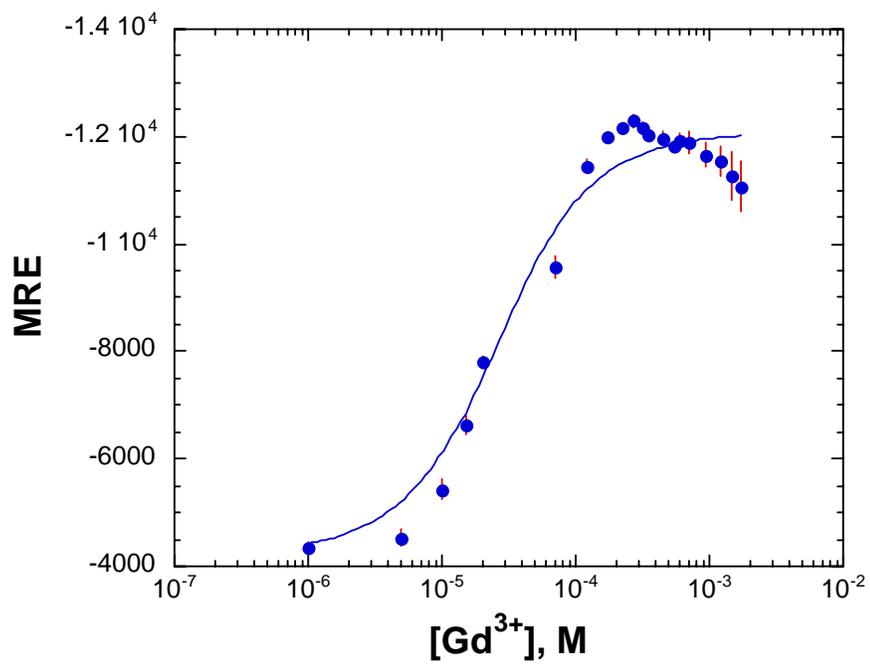
**Figure 2.19.** Metal titrations of  $\text{Yb}^{3+}$  - LF4 binding. (a) CD spectra of 50  $\mu\text{M}$  LF4 in the presence of  $\text{Yb}^{3+}$ . Concentration of  $\text{Yb}^{3+}$  is shown in the legend; (b) full titration of LF4- $\text{Yb}^{3+}$ , with the curve fit to the binding equation (1).



**Figure 2.20.** Full titration of LF4-Dy<sup>3+</sup>, with the curve fit to the binding equation (1).

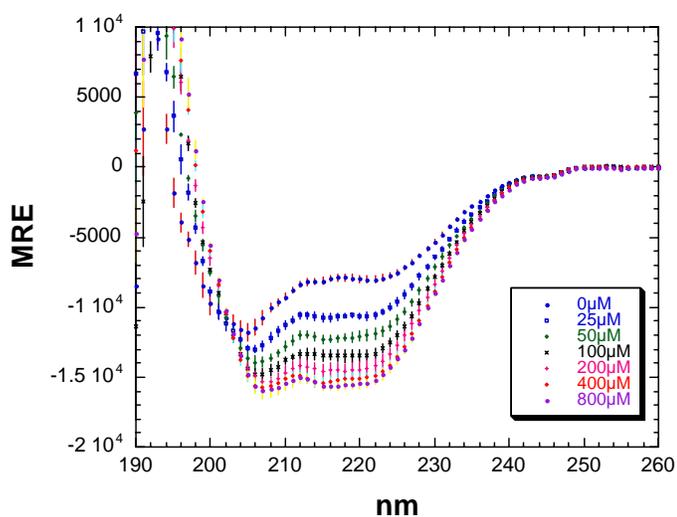


**Figure 2.21.** Full titration of LF4- $Eu^{3+}$ , with the curve fit to the binding equation (1).

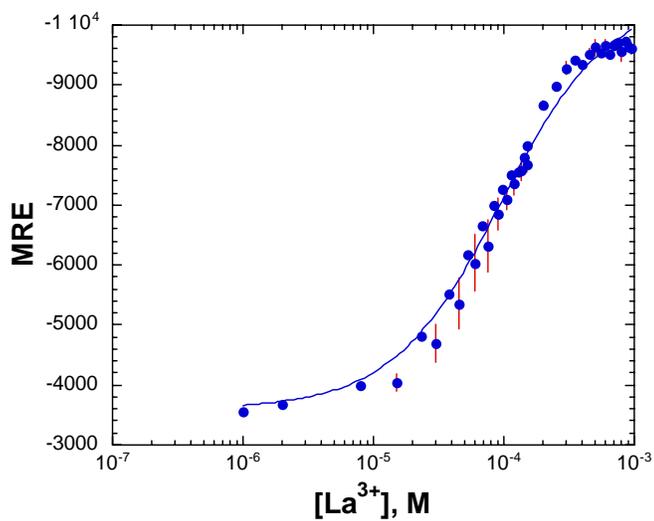


**Figure 2.22.** Full titration of LF4-Gd<sup>3+</sup>, with the curve fit to the binding equation (1).

(a)

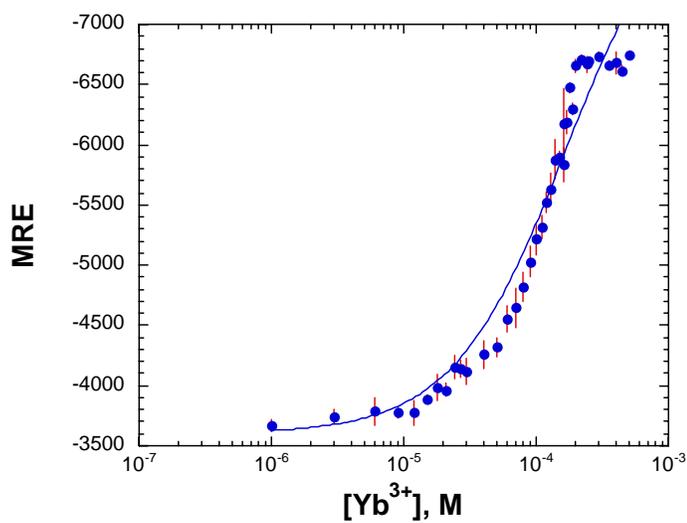


(b)

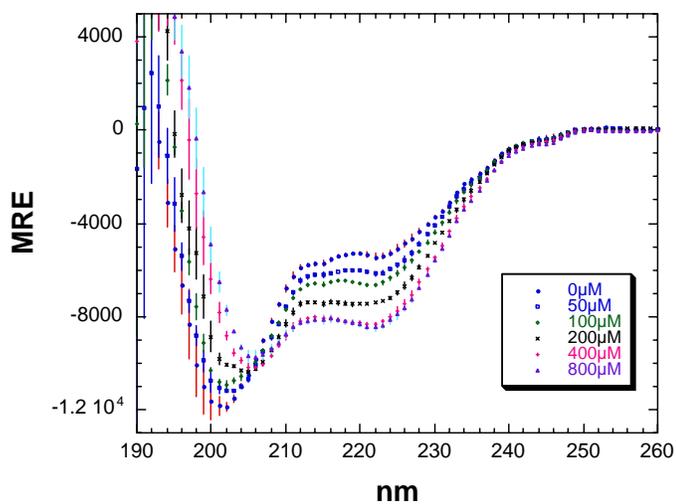


**Figure 2.23.** Metal titrations of La<sup>3+</sup> - LF4 binding. (a) CD spectra of 50 μM LF4 in the presence of La<sup>3+</sup>. Concentration of La<sup>3+</sup> is shown in the legend; (b) full titration of LF4-La<sup>3+</sup>, with the curve fit to the binding equation.

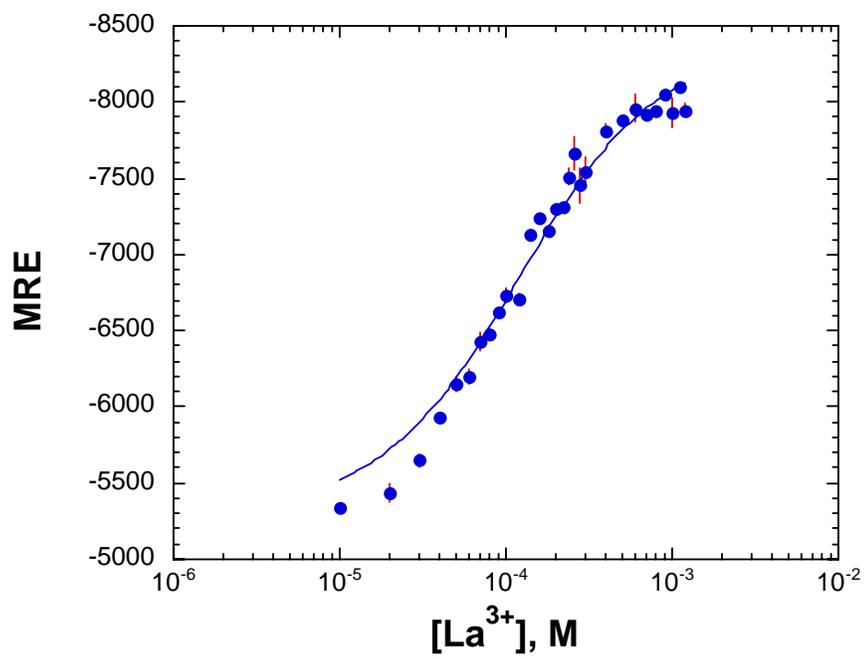
(a)



(b)

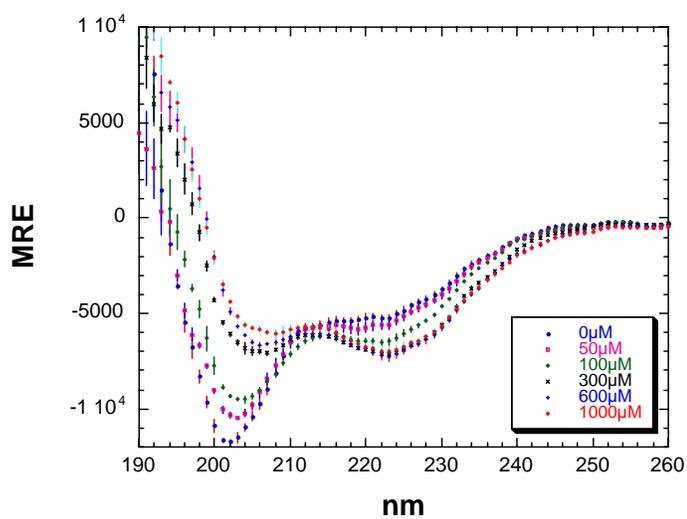


**Figure 2.24.** Metal titrations of  $\text{Yb}^{3+}$ - LF5 binding. (a) CD spectra of 50  $\mu\text{M}$  LF5 in the presence of  $\text{Yb}^{3+}$ . Concentration of  $\text{Yb}^{3+}$  is shown in the legend; (b) full titration of LF5- $\text{Yb}^{3+}$ , with the curve fit to the binding equation (1).

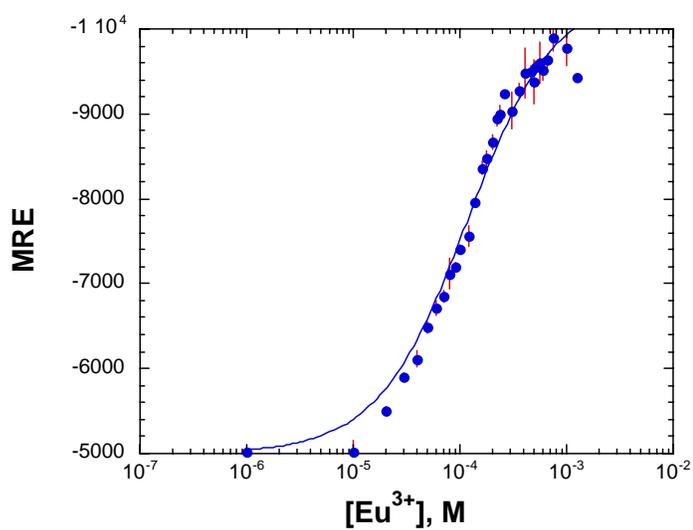


**Figure 2.25.** Full titration of LF5-La<sup>3+</sup>, with the curve fit to the binding equation(1).

(a)

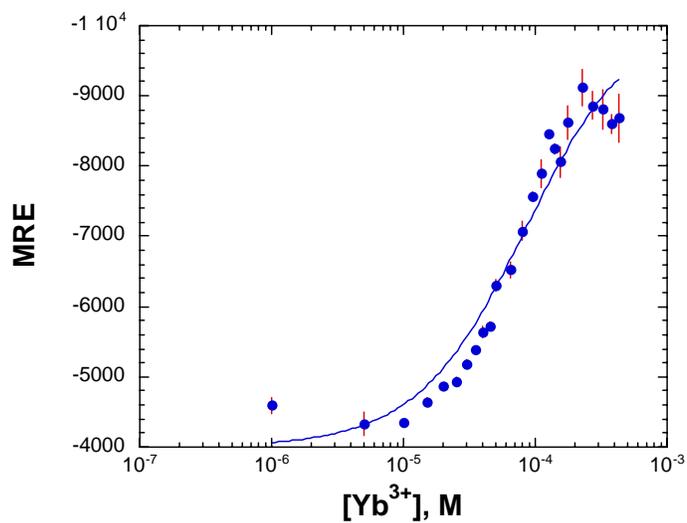
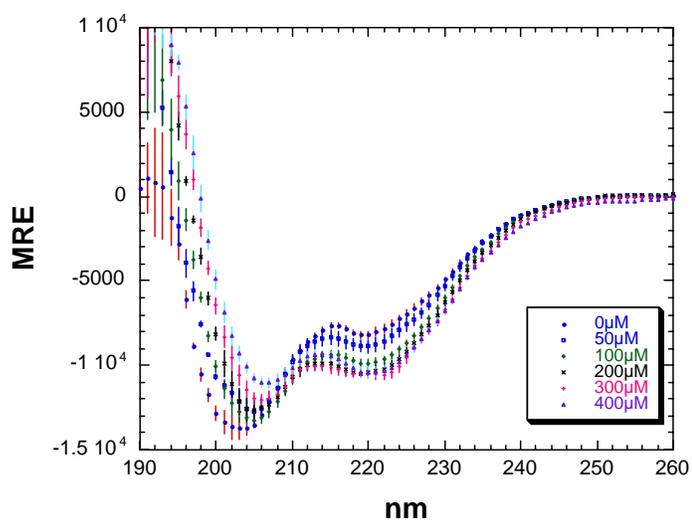


(b)



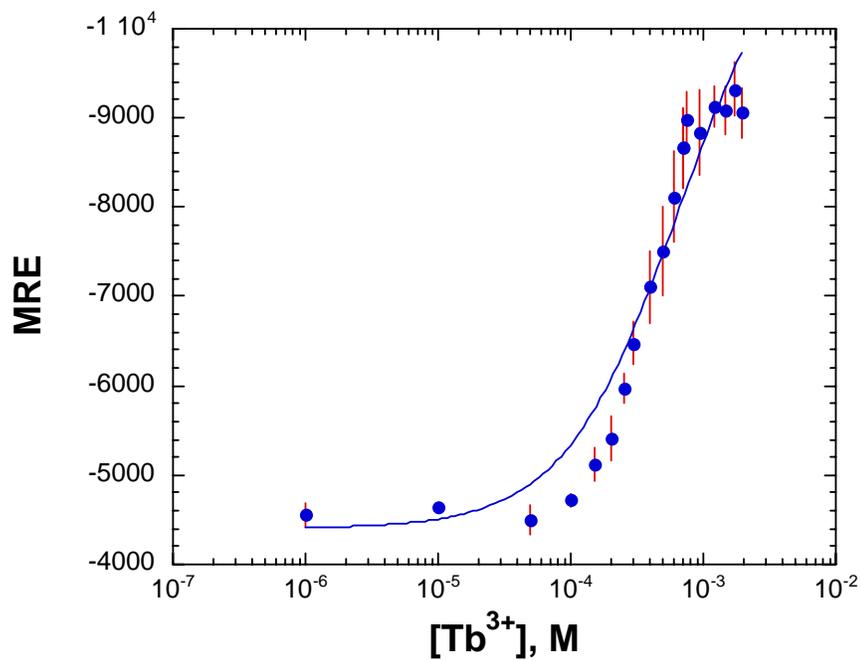
**Figure 2.26.** Metal titrations of  $\text{Eu}^{3+}$ -LF5 binding. (a) CD spectra of 50  $\mu\text{M}$  LF5 in the presence of  $\text{Eu}^{3+}$ . Concentration of  $\text{Eu}^{3+}$  is shown in the legend; (b) full titration of LF5- $\text{Eu}^{3+}$ , with the curve fit to the binding equation (1).

(a)

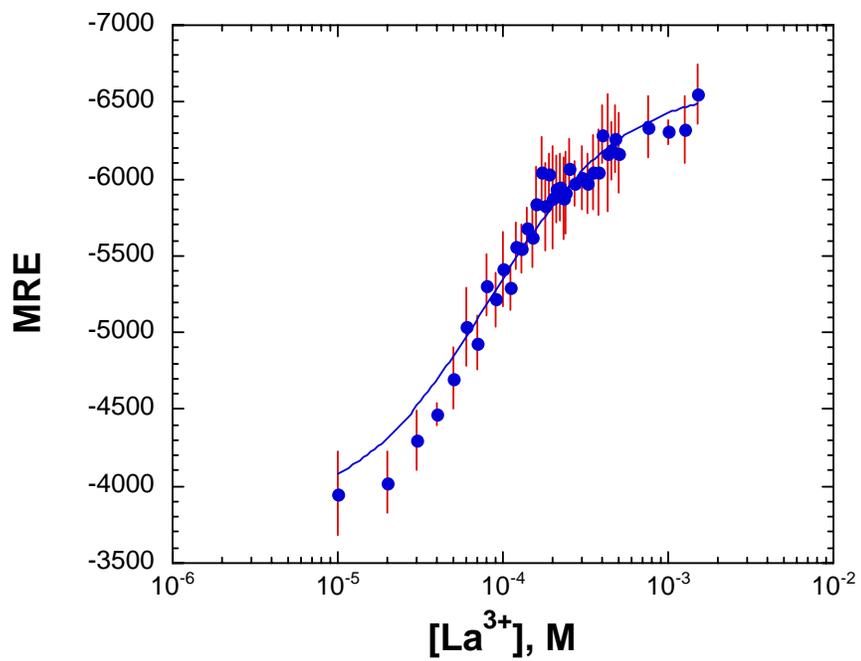


(b)

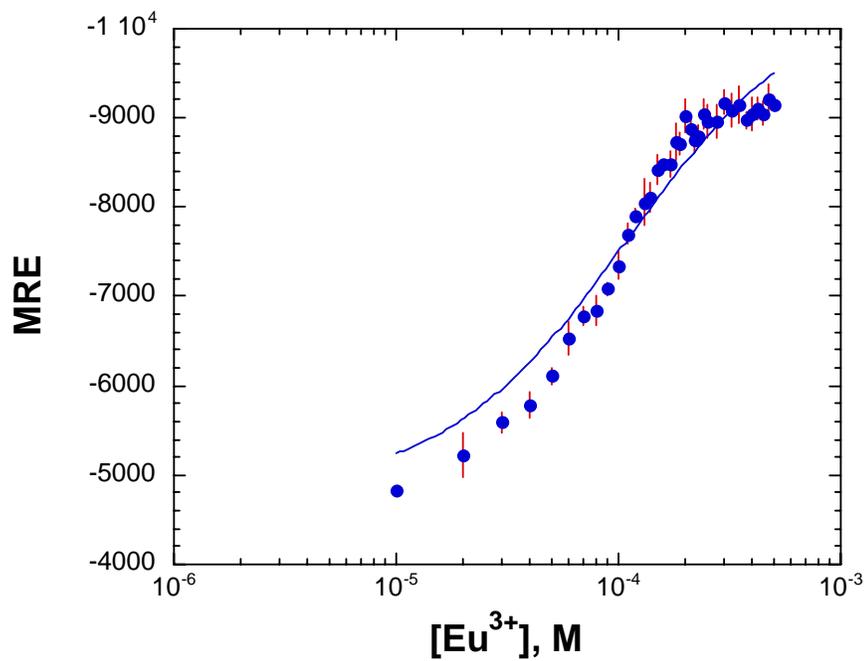
**Figure 2.27.** Metal titrations of Yb<sup>3+</sup>-LF6 binding. (a) CD spectra of 50 μM LF6 in the presence of Yb<sup>3+</sup>. Concentration of Yb<sup>3+</sup> is shown in the legend; (b) full titration of LF6-Yb<sup>3+</sup>, with the curve fit to the binding equation (1).



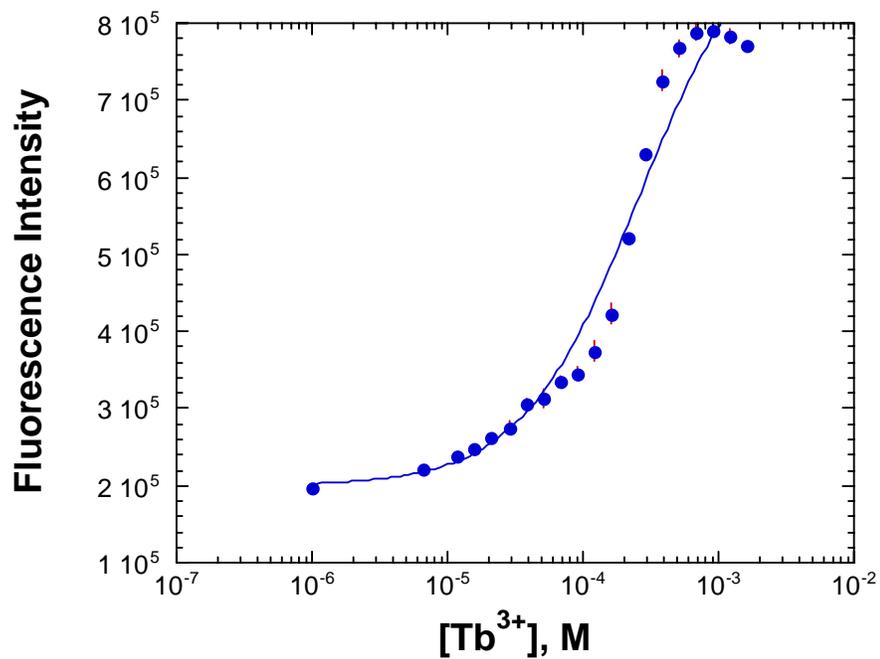
**Figure 2.28.** Full titration of LF6- $Tb^{3+}$ , with the curve fit to the binding equation (1).



**Figure 2.29.** Full titration of LF6- $La^{3+}$ , with the curve fit to the binding equation (1).

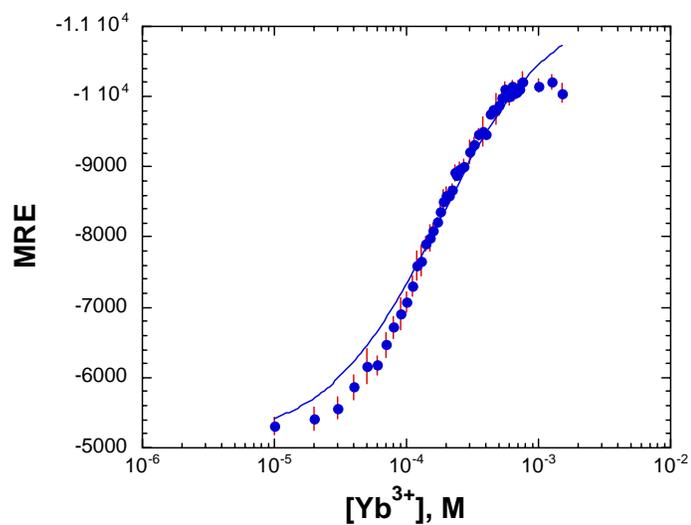


**Figure 2.30.** Full titration of LF6- $Eu^{3+}$ , with the curve fit to the binding equation (1).

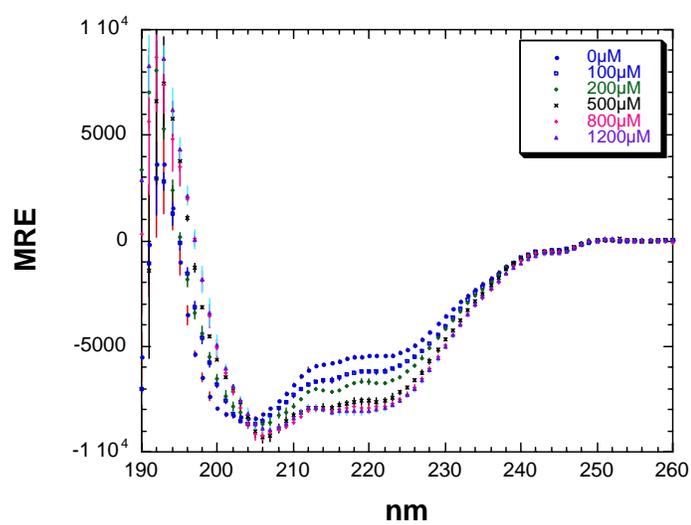


**Figure 2.31.** Full fluorescence titration of LF6- $Tb^{3+}$ , with the curve fit to the binding equation (1).

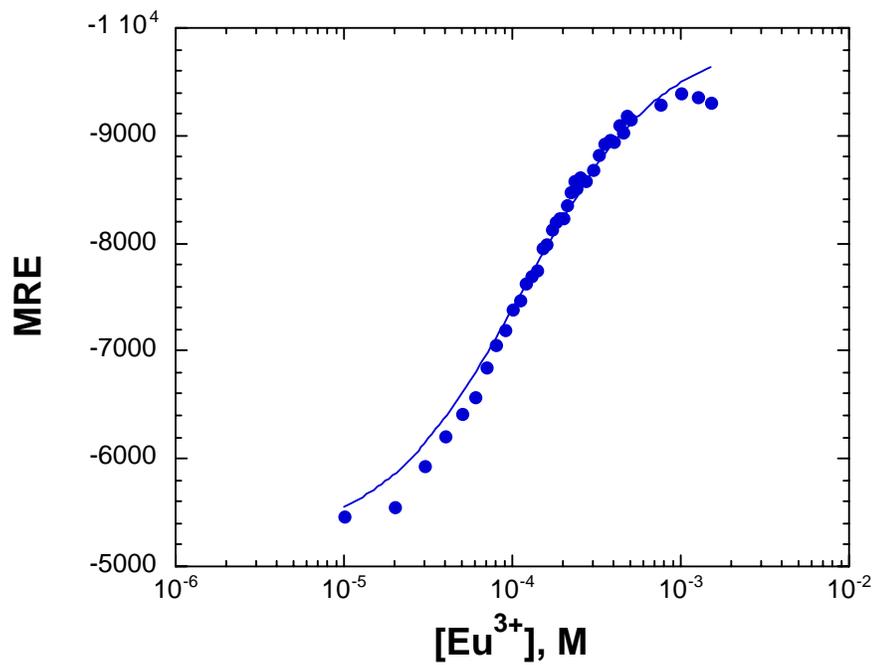
(a)



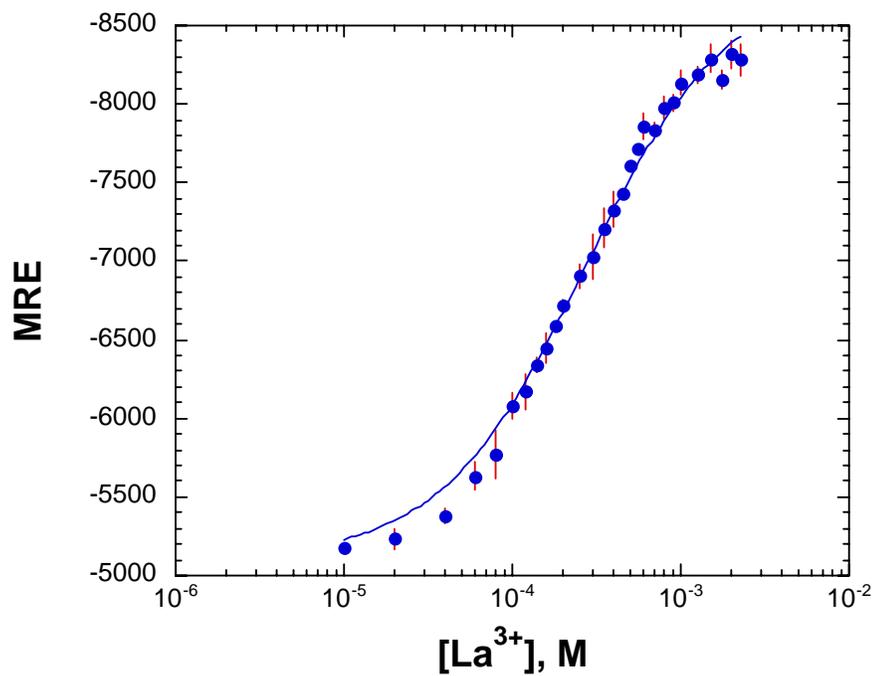
(b)



**Figure 2.32.** Metal titrations of  $Yb^{3+}$ - LF7 binding. (a) CD spectra of 50  $\mu$ M LF7 in the presence of  $Yb^{3+}$ . Concentration of  $Yb^{3+}$  is shown in the legend; (b) full titration of LF7- $Yb^{3+}$ , with the curve fit to the binding equation (1).



**Figure 2.33.** Full titration of LF7- $Eu^{3+}$ , with the curve fit to the binding equation (1).



**Figure 2.34.** Full titration of LF7- $La^{3+}$ , with the curve fit to the binding equation (1).

Peptide	Dissociation constant ( $K_d$ ), $\mu\text{M}$							
	Yb(III)	Eu(III)	La(III)	Gd(III)	Tb(III)	Dy(III)	Ho(III)	Er(III)
LF1	3400	> 10,000	> 10,000					
LF2	237	205	324					
LF3	114	223	243					
LF4	52	37	86	16	6.8	12	8	5.8
LF5	154	107	109					
LF6	74	73	73					
LF7	167	107						

**Table 2.1.** Dissociation constants for peptide-metal complexes.

## References

1. Selvin, P. R. *Ann. Rev. Biophys. Biomol. Struct.* 2002, 31, 275-302.
2. Bunzli, J. C. G.; Piguet, C. *Chem. Soc. Rev.* 2005, 34, 1048-1077.
3. Bunzli, J. C. G.; Piguet, C. *Chem. Rev.* 2002, 102, 1897-1928.
4. Kido, J.; Okamoto, Y. *Chem. Rev.* 2002, 102, 2357-2368.
5. Shibasaki, M.; Yoshikawa, N. *Chem. Rev.* 2002, 102, 2187-2209.
6. Lim, S.; Franklin, S. J. *Cell. Mol. Life Sci.* 2004, 61, 2184-2188.
7. Lee, L.; Sykes, B. D. *Biochemistry* 1980, 19, 3208-3214.
8. Horrocks, W. D.; Collier, W. E. *J. Am. Chem. Soc.* 1981, 103, 2856-2862.
9. Allegrozzi, M.; Bertini, I.; Janik, M. B. L.; Lee, Y. M.; Lin, G. H.; Luchinat, C. *J. Am. Chem. Soc.* 2000, 122, 4154-4161.
10. Barbieri, R.; Bertini, I.; Cavallaro, G.; Lee, Y. M.; Luchinat, C.; Rosato, A. J. *Am. Chem. Soc.* 2002, 124, 5581-5587.
11. Franz, K. J.; Nitz, M.; Imperiali, B. *Chembiochem* 2003, 4, 265-271.
12. Nitz, M.; Franz, K. J.; Maglathlin, R. L.; Imperiali, B. *Chembiochem* 2003, 4, 272-276.
13. Wohnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. *J. Am. Chem. Soc.* 2003, 125, 13338-13339.
14. Nitz, M.; Sherawat, M.; Franz, K. J.; Peisach, E.; Allen, K. N.; Imperiali, B. *Angew. Chem. Int. Ed.* 2004, 43, 3682-3685.
15. Kovacic, R. T.; Welch, J. T.; Franklin, S. J. *J. Am. Chem. Soc.* 2003, 125, 6656-6662..
16. Balakrishnan, S.; Zondlo, N. J. Submitted.
17. Pidcock, E.; Moore, G. R. *J. Biol. Inorg. Chem.* 2001, 6, 479-489.
18. Babu, Y. S.; Sack, J. S.; Greenhough, T. J.; Bugg, C. E.; Means, A. R.; Cook, W. J. *Nature* 1985, 315, 37-40.
19. Ikura, M. *Trends Biochem. Sci.* 1996, 21, 14-17.
20. Yang, W.; Lee, H. W.; Hellinga, H.; Yang, J. J. *Proteins* 2002, 47, 344-356.
21. Gariepy, J.; Sykes, B. D.; Hodges, R. S. *Biochemistry* 1983, 22, 1765-1772.
22. Yang, W.; Jones, L. M.; Isley, L.; Ye, Y. M.; Lee, H. W.; Wilkins, A.; Liu, Z. R.; Hellinga, H. W.; Malchow, R.; Ghazi, M.; Yang, J. J. *J. Am. Chem. Soc.* 2003, 125, 6165-6171.
23. Yang, W.; Wilkins, A. L.; Ye, Y. M.; Liu, Z. R.; Li, S. Y.; Urbauer, J. L.; Hellinga, H. W.; Kearney, A.; van der Merwe, P. A.; Yang, J. J. *J. Am. Chem. Soc.* 2005, 127, 2085-2093.
24. Ye, Y. M.; Lee, H. W.; Yang, W.; Shealy, S.; Yang, J. J. *J. Am. Chem. Soc.* 2005, 127, 3743-3750.
25. Lu, Y.; Berry, S. M.; Pfister, T. D. *Chem. Rev.* 2001, 101, 3047-3080.
26. Berg, J. M.; Godwin, H. A. *Ann. Rev. Biophys. Biomol. Struct.* 1997, 26, 357-371.
27. Pabo, C. O.; Peisach, E.; Grant, R. A. *Ann. Rev. Biochem.* 2001, 70, 313-340.

28. Jantz, D.; Amann, B. T.; Gatto, G. J.; Berg, J. M. *Chem. Rev.* 2004, 104, 789-799.
29. Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* 1996, 271, 342-345.
30. Struthers, M. D.; Cheng, R. P.; Imperiali, B. J. *Am. Chem. Soc.* 1996, 118, 3073-3081.
31. Walkup, G. K.; Imperiali, B. J. *Am. Chem. Soc.* 1996, 118, 3053-3054.
32. Nomura, A.; Sugiura, Y. *J. Am. Chem. Soc.* 2004, 126, 15374-15375.
33. Cerasoli, E.; Sharpe, B. K.; Woolfson, D. N. *J. Am. Chem. Soc.* 2005, 127, 15008-15009. *ZiCo*:
34. Ambroggio, X. I.; Kuhlman, B. J. *Am. Chem. Soc.* 2006, ASAP.
35. Payne, J. C.; ter Horst, M. A.; Godwin, H. A. *J. Am. Chem. Soc.* 1999, 121, 6850-6855.
36. Dahiyat, B. I.; Mayo, S. L. *Science* 1997, 278, 82-87..
37. Pavletich, N. P.; Pabo, C. O. *Science* 1991, 252, 809-817.
38. Elrod-Erickson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* 1996, 4, 1171.
39. Meng, H. Y.; Thomas, K. M.; Lee, A. E.; Zondlo, N. J. *Biopolymers (Peptide Sci.)* 2005, In press, DOI: 10.1002/bip.20382.
40. Aurora, R.; Rose, G. D. *Protein Sci.* 1998, 7, 21-38.
41. Laity, J. H.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* 2000, 295, 719-727. nt of
42. Marino, S. F.; Regan, L. *Chem. Biol.* 1999, 6, 649-655.

**Chapter 3**  
**DESIGN OF A KINASE-RESPONSIVE AND**  
**PHOSPHORYLATION-DEPENDENT PROTEIN DOMAIN**

**Introduction**

Protein kinases play a central role in cellular signal transduction, and it has been found that misregulation of kinases is tightly associated with many human diseases, such as cancer<sup>1</sup>, heart disease and Alzheimer's disease. Changes in patterns of post-translational modifications of protein side chains, including phosphorylation, sulfation, glycosylation, methylation, and acetylation, are usually associated with human diseases. Among those post-translational modifications, the importance of phosphorylation is emphasized by the observation that human genome encodes for 518 protein kinases and 140 genes encode protein phosphatases, corresponding to 2.5% of all human genes<sup>2</sup>.

Many human proteins exhibit function which is dependent on phosphorylation. The design of proteins whose function differs depending on their phosphorylation state is a potentially powerful strategy to predict and responsively control cell activities. These proteins would be able to generate difference responses which are dependent on

external cell signaling. Fluorescence is commonly used as a tool for the analysis of protein function. Although some sensitive fluorescence-based sensing systems for protein kinase activity have been developed, those systems all have their own limitations. For example, Lawrence has developed very sensitive fluorescence-based sensing systems for protein kinase activity. Their systems displayed a 3-fold fluorescence difference between phosphorylated and non-phosphorylated proteins, and allowed real-time analysis of kinase activity.<sup>3,4,5</sup> However, they included external fluorophores in their systems, and thus can not be expressed within cells. Imperiali also developed kinase sensors. Their probes can show as much as a 5-fold fluorescence difference between phosphorylated and non-phosphorylated peptides.<sup>6,7</sup> However, they incorporate an unnatural amino acid in the systems and thus the sensor also cannot be expressed. The ability to be expressed within cells is very critical for developing a kinase sensor because the purpose for developing this sensor is to monitor the phosphorylation signaling pathways inside cells.

Tsien has developed an expressible sensor for protein kinase A.<sup>8,9,10</sup> The system included two large fluorescent proteins, and only displayed moderate (around 25%) fluorescence changes upon phosphorylation. This large kinase sensor might interfere with the conformation of the protein of interest if it is incorporated to the protein of interest and expressed in cells.

Given the limitations of all the existing kinase sensor systems, we sought to design a system combining large changes in fluorescence with the ability to be expressed within cells. The system can be used as an expressible tag attached to a

protein of interest. Because the kinase sensor tag is genetically encoded, it will be expressed along with protein of interest inside cells, and thus allow the determination of kinase activity within cells in real-time. The kinase sensor will be useful in understanding molecular signaling by detecting its kinase activity in the macromolecular scaffold complexes where kinases are localized.

In an early work, Balakrishnan and Zondlo successfully designed a series of kinase-inducible domains (pKID) based on EF hands. The structures of kinase-inducible domains are dependent on their phosphorylation state. In an EF-hand motif, the peptide binds to calcium and forms a loop structure (Figure 3.1). We substituted calcium with a lanthanide in our design because lanthanides are chemically similar to calcium (hardness, ionic radius), and have a lot of other interesting chemical properties, such as sensitized luminescence, long luminescence lifetimes ( $\mu\text{s}$  to  $\text{ms}$ ), and strong magnetic properties because of their large number of unpaired f electrons. We also included a natural chromophore, Trp, in our sequence. When Trp is close ( $<10 \text{ \AA}$ ) to lanthanides (especially  $\text{Tb}^{3+}$ ), and is excited by UV light at 280 nm, it will undergo energy transfer to lanthanides, and display a fluorescence emission band in the visible wavelength range. In an EF-hand motif, the peptide binds to calcium ion through six residues, as shown in Figure 3.1. The position 12 in an EF-hand is nearly always Glu, which binds the metal in a bidentate manner. Since phosphoserine has a similar structure to Glu (Figure 3.2), Glu is often used to mimic phosphoserine in biological studies. So in our design, we integrated phosphoserine in position 12, and

expected that phosphoserine will have the same function as Glu in EF hands, while non-phosphorylated Ser will not and will bind metal poorly.

One of the best phosphorylation-dependent kinase-inducible domains we developed in our initial studies used Protein Kinase A (PKA) as a kinase, and thus we called the peptide pKID-PKA. It is a 14 residue peptide. It includes a 12-residue EF-hand loop (Figure 3.1). In our design of pKID-PKA, Glu12 in the EF hand motif was replaced with Ser. A chromophore, Trp, was incorporated in the sequence to sensitize lanthanide luminescence. When  $Tb^{3+}$  is bound to pKID-PKA, forming a metal-peptide complex,<sup>11-14</sup> a close proximity will be provided. If Trp is excited by UV light at 280 nm, it will be able to transfer its emission energy to  $Tb^{3+}$  and undergo  $Tb^{3+}$  emission at 544 nm. In this system, we predicted and demonstrated that the phosphoserine-containing pKID-PKA can effectively bind to  $Tb^{3+}$ , undergo a significant conformational change to a helix-loop-helix structure, and show strong  $Tb^{3+}$  luminescence. In addition, non-phosphorylated pKID-PKA bound poorly to  $Tb^{3+}$ , and was not able to display high luminescence. This peptide showed complete dependence on phosphorylation for its structure and its fluorescence. However, the limitation for this peptide was the slow phosphorylation by Protein Kinase A (PKA). It required 18-24 hrs incubation with PKA to achieve 80-90% conversion to phosphorylated peptide. For an ideal kinase sensor system, a rapid phosphorylation by PKA (within the second to minute range) is required to allow real time readout. This initial system is too slow to be used as an effective kinase sensor.

In this study, we wanted to design a peptide which can exhibit a large fluorescence difference between phosphorylated and non-phosphorylated peptides, and at the same time is rapidly phosphorylated by PKA. Nine different peptides, HM1, HM2, HM3, HM4, HM5, HM6, HM7, HM8, and HM9, were synthesized and analyzed. For each of the peptide, the fluorescence was measured, and the dissociation constant was determined for both the phosphorylated (p-HM#) and non-phosphorylated peptides (HM#). In addition, the rates of phosphorylation by PKA were analyzed by analytical HPLC.

## **Experimental**

### **Materials**

Amino acids and resins for peptide synthesis were purchased from Novabiochem. HBTU was purchased from Senn Chemicals. Diisopropylethylamine (DIPEA) was purchased from Aldrich. Dimethyl formamide (DMF), methylene chloride (DCM), trifluoroacetic acid (TFA) were purchased from Acros. All other compounds were purchased from Acros unless otherwise indicated. Peptide synthesis was carried out on a Rainin PS3 automated peptide synthesizer. Peptide stock solutions were prepared in ultrapure water purified by a Millipore Synergy 185 water purification system with a Simpак® 2 cartridge. Peptide concentrations were

determined by UV/VIS based on tryptophan absorbance ( $\epsilon = 5400 \text{ M}^{-1} \text{ cm}^{-1}$  at 280 nm in water) on a PerkinElmer Lambda 25 spectrometer. Peptide stock solutions were stored at  $-10 \text{ }^\circ\text{C}$ . Postsynthetic modification reactions were performed in capped disposable fritted columns (Image Molding). All compounds were used as purchased with no additional purification. Terbium (III) chloride stock solutions were prepared from the hexahydrate salts (Aldrich). 500 mM HEPES buffer pH 7.5 was prepared from HEPES (Acros) and adjusted to pH 7.5 with a 1 M NaOH solution. NaCl (1 M) and  $\text{MgCl}_2$  (1 M) solutions were prepared from the corresponding salts (Acros). Protein Kinase A was purchased from New England Biolabs.

### **Peptide synthesis, purification and characterization**

Peptides were synthesized using standard Fmoc solid phase peptide synthesis with Rink amide resin (0.25 mmol). The resin was swelled in DMF (5 mins) prior to the start of the synthesis. Amino acid couplings were performed using Fmoc amino acids (1 mmol, 4 equiv) and HBTU (1 mmol, 4 equiv). The following steps were used for each cycle of peptide synthesis: (1) removal of the Fmoc group with 20% piperidine in DMF, 3 x 5 mins; (2) resin wash (DMF, 5 x 1 min); (3) amide coupling (amino acid, HBTU and 0.05 M DIPEA in DMF, 50 mins); (4) resin wash (DMF, 3 x 1 min.). Trityl-protected serine was incorporated at the phosphorylation site to allow for selective phosphorylation. After addition of the final residue, the N-terminal Fmoc group was removed (20% piperidine in DMF, 3 x 5 mins) and the amino terminus

acetylated (10% acetic anhydride in pyridine, 5 mins). The resin was washed with DMF (6 x) and DCM (3 x).

Non-phosphorylated peptides subjected to cleavage from resin and deprotection for 3 hrs using the TFA:thioanisole:ethanedithiol:phenol:water (84:4:4:4:4). The solutions were concentrated under nitrogen and precipitated with ether. The precipitate was dissolved in water and the resulting solution was filtered and purified by reverse phase HPLC (Vydac semipreparative C18, 10 mm x 25 cm, 5  $\mu$ m particle size, 300 Å pore) using a linear gradient of 0-50% buffer A (98% H<sub>2</sub>O, 2% MeCN, 0.06% TFA) in buffer B (80% MeCN, 20% H<sub>2</sub>O, 0.05% TFA) over 60 mins. Peptides were purified to homogeneity as indicated by the presence of a single peak on reinjection on analytical HPLC (Microsorb MV C18, 4.6 x 250 mm, 100 Å). Peptides were characterized by ESI-MS (negative ion mode) on an LCQ Advantage (Finnigan) mass spectrometer (Table 3.1).

Peptides were chemically phosphorylated on resin by the following procedure: (1) deprotection of the trityl group with 2% TFA/ 5% triethylsilane (TES)/ 93% DCM, 3 x 1 min or until the flow-through solution was clear; (2) phosphorylation was performed under nitrogen by addition to the resin of tetrazole (1.35 mmol; 3 mL of 3% tetrazole solution in MeCN) (Transgenomics) and O,O-dibenzyl-N,N-diisopropylphosphoramidite (500  $\mu$ L, 1.52 mmol) (Fluka), and allowed to react for 3 hrs with gentle mixing on a Barnstead-Thermoline Labquake rotary shaker. The solution was removed and the resin washed with DMF (3x) and DCM (3x); (3) oxidation was performed with t-butyl hydroperoxide (2 mL of a 3 M solution in DCM)

and allowed to react with mixing for 30 mins. The solution was removed and the resin washed with DMF (3x), MeOH (3x), and ether (3x).

The phosphorylated peptides were purified and characterized as described above for non-phosphorylated peptides (Table 3.1).

### **Fluorescence experiments**

Fluorescence spectra were collected on a Photon Technology International fluorescence system model QM-3/2003 with a CW source and a Hamamatsu R928 PMT. All experiments were conducted with an excitation wavelength of 280 nm. 10 nm excitation and emission slit widths were used unless otherwise indicated. All spectra were acquired at room temperature collecting data every 1 nm with a scan rate of 1 nm per second. At least 2 independent titrations were conducted for each peptide. All fluorescence experiments were conducted using a 495 nm highpass filter (model 495FG03-25 AM-53074; Andover Corporation, Andover, NH) on the emission monochromator.

Peptide solutions were prepared by dilution of stock solutions into 5 mM HEPES buffer (pH 7.5) with 100 mM NaCl. Spectra were acquired in 10 mm quartz fluorescence cells (Starna).  $Tb^{3+}$  titrations were conducted by dilution of a  $Tb^{3+}$  solution into a peptide solution as described above. Each individual emission spectrum represents an independently prepared solution of peptide and metal. The  $Tb^{3+}$  emission band at 544 nm was studied to evaluate metal binding.

### **Dissociation constant determination**

To determine dissociation constants for Tb<sup>3+</sup>-peptide complexes, titrations were conducted in solutions of 5 mM HEPES (pH 7.5), 2 mM Mg<sup>2+</sup>, and 100 mM NaCl, with 10 nm emission and excitation slit widths. The solutions were prepared as serial dilutions of Tb<sup>3+</sup>, from 2675 μM to 6.64 μM Tb<sup>3+</sup>. Data points at the emission maxima of 544 nm were plotted against Tb<sup>3+</sup> concentrations to calculate the final dissociation constant. These plots were fit to equation 1 using a non-linear least squares fitting algorithm (Kaleidagraph version 3.6, Synergy Software), where Q = fluorescence, Q<sub>o</sub> = observed fluorescence of the apo-peptide, Q<sub>c</sub> = fluorescence of the peptide-metal complex, M<sub>t</sub> = total metal concentration, K<sub>d</sub> = dissociation constant and P<sub>t</sub> = total peptide concentration. Non-linear least squares fits were fit to K<sub>d</sub> and Q<sub>c</sub>. Dissociation constants were calculated based on at least 2 independent trials (Table 3.2).

Equation (2):

$$Q = Q_o + (Q_c - Q_o) \frac{[(M_t + K_d + P_t) - \sqrt{[(M_t + P_t + K_d)^2 - 4(P_t M_t)}]]}{2 P_t}$$

### **Phosphorylation of HM series peptides by Protein Kinase A**

Protein Kinase A (PKA) was used to study the phosphorylation kinetics for all peptides. PKA enzyme and PKA reaction buffer used for the reaction were purchased from New England Biolabs. For HM1, HM2, HM3, HM4, HM5, and HM6, PKA reaction mixtures were prepared to a final volume of 100 μL as follows: stock

solutions were mixed to yield final concentrations of 600  $\mu\text{M}$  ATP (15  $\mu\text{L}$  of 4 mM ATP (Alexis Biochemicals)), 1x PKA buffer (10  $\mu\text{L}$  of 10x buffer), and 200 $\mu\text{M}$  peptides. The reaction mixture was incubated at 30 °C water bath for 5 mins, then 3  $\mu\text{L}$  of PKA enzyme solution (7.5 units) was added to start the reaction. For HM7, HM8 and HM9, PKA reaction mixtures were prepared to a final volume of 100  $\mu\text{L}$  as follows: stock solutions were mixed to yield final concentrations of 500  $\mu\text{M}$  ATP (25  $\mu\text{L}$  of 2 mM ATP (Alexis Biochemicals)), 1x PKA buffer (10  $\mu\text{L}$  of 10x buffer), and 50 $\mu\text{M}$  peptides. The reaction mixture was incubated at 30 °C water bath for 5 mins, then 0.6  $\mu\text{L}$  of PKA enzyme solution (1.5 units) was added to start the reaction. The reaction was interrupted by sitting in a 90 °C water bath for 5 mins. To study the phosphorylation conversion rate of the peptides, the resulting interrupted reaction mixture was analyzed by analytical HPLC (Microsorb MV C18, 4.6 x 250 mm, 100 Å) using a linear gradient of 20-50% buffer A (98%  $\text{H}_2\text{O}$ , 2% MeCN, 0.06% HFB) in buffer B (80% MeCN, 20%  $\text{H}_2\text{O}$ , 0.05% HFB) over 50 mins. TFA was replaced by Heptafluorobutyric acid (HFB) in the buffer system to produce a better separation between phosphorylated peptide and non-phosphorylated peptide.<sup>15</sup>

## **Results and Discussion**

### **HM1 analysis**

Cantley identified that sequence RRRSII is an optimized substrate for Protein Kinase A (PKA).<sup>16</sup> The Ser in bold is the residue that is phosphorylated by PKA.

HM1 incorporated the partial PKA recognition site, RRSII, between residues 10 and 14 from the optimized sequence of RRRSII. Arg was substituted by Asp at position 9 to improve  $Tb^{3+}$  affinity, as was present in pKID-PKA. In the EF-hand motif, Asp in position 9 was able to contact with metal through a water-mediated hydrogen bond, and thus promote  $Tb^{3+}$  binding, so we predicted that Asp in position 9 in our design would function similarly.

Compared to the fluorescence data on HM2 in which only one residue, position 9, was different from HM1, fluorescence data on HM1 showed that Asp in position 9 did improve  $Tb^{3+}$  binding. HM1 displayed as much as 4.0-fold difference in fluorescence between phosphorylated HM1 and non-phosphorylated HM1 in the presence of 100  $\mu M$   $Tb^{3+}$  (Figure 3.3). However, phosphorylation of HM1 by PKA was extremely slow. Only around 1% of the non-phosphorylated HM1 was converted into phosphorylated HM1 in 6 hrs.

The dissociation constants ( $K_d$ ) for non-phosphorylated HM1 and phosphorylated HM1 were 334  $\mu M$  and 102  $\mu M$  respectively. Compared to the  $K_d$  of pKID-PKA, which was 174  $\mu M$  for non-phosphorylated form and >5000  $\mu M$  for the phosphorylated form, the dissociation constant difference between two phosphorylation states of HM1 was small.

HM1 demonstrated the function of Asp in position 9, and also addressed the importance of position 9 to PKA enzyme recognition. However, HM1 cannot be used as a protein tag inside cells because the probe we try to develop needs to have fast

kinetics so that the kinase activities can be monitored in real time within a range between seconds to minutes.

### **HM2 analysis**

HM2 integrated a full optimized PKA recognition site, RRRSII, from residue 9 to residue 14 of the EF hand.<sup>12</sup> Three Arginine residues consecutively resulted in a significant positive charge near the metal-binding site, and also the loss of a water-mediated metal-binding contact by Asp of HM1 in position 9. However, this sequence should provide optimal kinase kinetics.

The PKA enzyme kinetics study showed that 83% of HM2 was converted into phosphorylated HM2 within only 5 mins. However, the phosphorylated HM2 and non-phosphorylated HM2 exhibit almost the same fluorescence intensity at same level of Tb<sup>3+</sup> concentration (Figure 3.4). Both peptides also displayed a significant loss in Tb<sup>3+</sup> binding affinity compared to pKID-PKA or HM1.

Non-phosphorylated HM2 and phosphorylated HM2 had almost the same dissociation constants. They were 456  $\mu$ M for non-phosphorylated HM2 and 549  $\mu$ M for phosphorylated HM2. Compared to their counterparts in HM1, the dissociation constants were both larger.

HM2 demonstrated that Arg in position 9 was necessary for better PKA recognition and kinetics. Due to the small fluorescence difference between phosphorylated and non-phosphorylated forms, HM2 cannot be used as a protein tag in cells.

### HM3 analysis

Based on the rapid PKA kinetics observed for HM2, we adopted the same recognition site, RRRSII, in HM3. In order to reduce the positive charge near the metal binding site, we replaced Lys2 and Asn3 with Ala and Asp, respectively. Ala provided a neutral charge to the peptide. Asp had similar side chain to Asn, and also offered a negative charge to the peptide. Asp is commonly observed at residue 3 in EF hand proteins, and it was predicted to be good for lanthanide binding because lanthanides have a +3 charge.

Fluorescence studies indicated that HM3 exhibited as much as a 3.8-fold difference in fluorescence between phosphorylated HM3 and non-phosphorylated HM3 in the presence of 100  $\mu\text{M}$   $\text{Tb}^{3+}$  (Figure 3.5). Fast enzyme kinetics was observed in HM3. 73% of non-phosphorylated HM3 was phosphorylated within 5 mins.

A 1.7-fold difference in dissociation constant was observed between phosphorylated and non-phosphorylated HM3. The  $K_d$  was 462  $\mu\text{M}$  for non-phosphorylated HM3, and 279  $\mu\text{M}$  for phosphorylated HM3. Compared to their counterparts in HM1, they were both larger, but the dissociation constant of phosphorylated HM3 was smaller than that of phosphorylated HM2.

HM3 combined the features of  $\text{Tb}^{3+}$  affinity in HM1 and kinetics in HM2. However, compared to pKID-PKA, which had more than 20-fold fluorescence difference between non-phosphorylated and phosphorylated forms at 100  $\mu\text{M}$   $\text{Tb}^{3+}$ , HM3 is less favorable as a protein tag. In order to make the kinase sensor system to

work, we need to have larger fluorescence changes from non-phosphorylated state to phosphorylated state.

#### **HM4 analysis**

Similar to HM2, we kept the optimized PKA recognition site, RRRSII, in HM4, but we replaced Lys with Ala in position 2 to reduce the positive charge in the peptide.

In the fluorescence experiments, compared to HM3, phosphorylated HM4 and non-phosphorylated HM4 displayed almost the same fluorescence intensity in the presence of 100 $\mu$ M Tb<sup>3+</sup>, but they showed a 1.5-fold difference in fluorescence in the presence of 1 mM Tb<sup>3+</sup> (Figure 3.6). The phosphorylation of HM4 by PKA required only 5 mins to achieve 80% conversion.

The dissociation constant of non-phosphorylated HM4 was almost double of that of phosphorylated HM4. They were 624  $\mu$ M and 317  $\mu$ M respectively. Compared to HM3, either of the dissociation constants in HM4 was bigger than their counterparts in HM3.

From the fluorescence comparison between HM3 and HM4, we confirmed that a negative charged residue, like Asp, in position 3 was helpful to metal binding. Again, like HM3, this peptide cannot be used as a kinase sensor due to its small fluorescence difference between two phosphorylation states.

## HM5 analysis

In HM5, we wanted to test how much impact the Arg in position 11 had on PKA kinetics. Although three consecutive arginine residues did improve the PKA kinetics, it also sacrificed the  $Tb^{3+}$  affinity due to the significant positive charge near metal binding site. So we kept the sequence the same as HM4 but only substituted Ala in position 11 for Arg.

Similar to HM4, phosphorylated HM5 and non-phosphorylated HM5 exhibited almost the same fluorescence intensity in the presence of  $100\mu M Tb^{3+}$ , and showed a 1.8-fold difference in fluorescence in the presence of  $1 mM Tb^{3+}$  (Figure 3.7). The PKA kinetics of HM5 was also similar to HM4. 72% of HM5 was phosphorylated by PKA within 5 mins.

The dissociation constants for phosphorylated HM5 and non-phosphorylated HM5 were competitive. They were  $277 \mu M$  and  $214 \mu M$  respectively. Compared to HM4, they were both smaller than their counterparts in HM4.

HM5 demonstrated that Ala substitution for Arg in position 11 did not significantly impact on the PKA kinetics. However, like HM4, HM5 cannot be used as a protein tag in cells due to its small fluorescence difference between two phosphorylation states.

## HM6 analysis

Cantley pointed out that Lys was a medium selection for PKA recognition compared to Arg.<sup>12</sup> In HM6, we wanted to evaluate how much impact it would have

on kinase activity if we replaced both of the Arg residues in HM5 with Lys. We also kept Asp in position 3 to reduce the positive charge of the peptide and improve metal binding.

Compared to HM4 and HM5, fluorescence experiments showed a little improvement on the difference between non-phosphorylated HM6 and phosphorylated HM6, showing a 2.3-fold difference at 100  $\mu\text{M}$   $\text{Tb}^{3+}$  and a 1.9-fold difference at 1 mM  $\text{Tb}^{3+}$  respectively (Figure 3.8). However, only 1% non-phosphorylated HM6 was converted into phosphorylated HM6 in 5 mins.

Compared to each of the peptides we analyzed before, from HM1 to HM5, the difference in dissociation constants between phosphorylated state and non-phosphorylated state in HM6 was the largest observed. The dissociation constants were 173  $\mu\text{M}$  for phosphorylated HM6 and 685  $\mu\text{M}$  for non-phosphorylated HM6, a 4-fold difference.

HM6 verified the necessity of two Arginine residues in the recognition site for rapid phosphorylation by PKA. HM6 cannot be used as a kinase sensor because of its slow PKA kinetics and small fluorescence difference between the two phosphorylation states.

### **HM7 analysis**

The design of HM7 was based on HM3 and HM5. From HM3 we knew that Asp in position 3 helped metal binding. From HM5 we knew that Ala substituted for

Arg in position 11 does not affect the PKA kinetics but would help reduce the positive charge in the peptide. HM7 combined these two features together.

A 2.8-fold difference in fluorescence intensity was observed at 100  $\mu\text{M}$   $\text{Tb}^{3+}$  level between phosphorylated HM7 and non-phosphorylated HM7 (Figure 3.9). The PKA kinetics was fast. 50% of HM7 was phosphorylated within 5mins.

Phosphorylated HM7 had the smallest dissociation constant, 70  $\mu\text{M}$ , of all the HM series peptides examined. Although the fluorescence intensity was almost the same for phosphorylated HM7 and non-phosphorylated HM7 at 1 mM  $\text{Tb}^{3+}$ , the dissociation constant for non-phosphorylated HM7 was 366  $\mu\text{M}$ , which is around 5.2 fold higher than that of phosphorylated HM7.

It is not worthy that the fluorescence difference between phosphorylated and non-phosphorylated HM7 was comparable to that observed by Lawrence and Imperiali, in a sequence that is expressible, like that of them.

### **HM8 analysis**

In our initial examination with pKID-PKA, Lys was used at position 2. pKID-PKA displayed a large fluorescence difference between the phosphorylated and non-phosphorylated states. In the design of HM8, we wanted to test whether Lys in position 2 would help metal binding.

The fluorescence data demonstrated that HM8 did have better  $\text{Tb}^{3+}$  affinity than HM7. It displayed a 3.8-fold difference in fluorescence intensity at 100  $\mu\text{M}$   $\text{Tb}^{3+}$  (Figure 3.10). HM8 was also rapidly phosphorylated by PKA. 53% of HM8 was

phosphorylated within 5 mins, 69% of that was converted within 10 mins, and was completely phosphorylated within 20 mins.

Similar to HM7, phosphorylated HM8 had a dissociation constant of 79  $\mu\text{M}$ , which was also below 100  $\mu\text{M}$ . The  $K_d$  for non-phosphorylated HM8 was 492  $\mu\text{M}$ . The  $K_d$  of phosphorylated HM8 was 6.2-fold lower than that of non-phosphorylated HM8. The  $K_d$  difference in HM8 was the biggest out of all HM series peptides we analyzed so far.

Like HM7, the fluorescence difference between phosphorylated and non-phosphorylated HM8 was comparable to that observed by Lawrence and Imperiali, in a sequence that is expressible, like that of them.

### **HM9 analysis**

Since two consecutive Arginine residues were necessary for PKA recognition, and three consecutive Arginine residues reduce the  $\text{Tb}^{3+}$  affinity, we tried to put a neutral Ser in position 9, and replaced the Ile with Arg in position 8. Like Asp, Ser has a hydroxyl group in its side chain, and it is commonly observed at residue 9 in EF hand proteins. Because Asp in position 9 is able to contact the metal through a water-mediated hydrogen bond, and thus promote  $\text{Tb}^{3+}$  binding, as observed in HM1, we expected that Ser can have the same function in HM9. When Ser occupies position 9, the peptide loses one of the two critical Arg recognition residues, position 9 and 10, for PKA. We replaced Ile in position 8 with Arg, and expected that this Arg will help PKA enzyme recognize the peptide.

Similar to HM7, fluorescence experiments showed a 3.3-fold fluorescence increase from the non-phosphorylated HM9 to phosphorylated HM9 in the presence of 100  $\mu\text{M}$   $\text{Tb}^{3+}$  (Figure 3.11). Compared to HM7 and HM8, the PKA kinetics was relatively slow. 46% of HM9 was converted to phosphorylated HM9 within 30 mins, and 59% of that was converted within 1 hr.

The dissociation constant was 122  $\mu\text{M}$  for phosphorylated HM9, and 534  $\mu\text{M}$  for non-phosphorylated HM9, a 4.4-fold difference. Compared to either HM7 or HM8, the difference in dissociation constants between phosphorylated state and non-phosphorylated state was smaller.

Compared to Tsien's system, which only displayed 25% fluorescence changes upon phosphorylation, HM9 was a better system, while compared to HM8, the fluorescence difference showed in HM9 was smaller, and compared to either HM7 or HM8, the kinetics for HM9 was slower too.

## **Conclusion**

The results shown here were the second generation designs for kinase inducible domains. In our first generation designs, phosphoserine-containing pKID-PKA can effectively bind to  $\text{Tb}^{3+}$ , and displayed high fluorescence difference between phosphorylated form and non-phosphorylated form even though the PKA kinetics for pKID-PKA was slow. In our second generation designs, both HM3 and HM8 exhibited a good combination between large fluorescence difference and rapid enzyme

kinetics. Both HM3 and HM8 exhibited a 3.8-fold difference in fluorescence between phosphorylated state and non-phosphorylated state, and the difference was the highest out of all HM series peptides analyzed. 73% of non-phosphorylated HM3 was phosphorylated within only 5 mins, 53% of HM8 was phosphorylated within 5 mins, and HM8 was completely phosphorylated by PKA within 20 mins.

The data shown here look promising. Compared to all those systems developed by Lawrence, Imperiali, or Tsien, our system definitely has far more advantages over each of theirs. In our system, all those peptides can be expressed within cells. Those peptides are all small, and thus can be used as a tag attached to a protein of interest without disturbing the conformation of the protein. A 3.8-fold fluorescence difference, and a rapid PKA kinetic present in both of HM8 and HM3, will allow the kinase sensor system to work better than any other system currently existing in the field.

Further efforts to apply the analysis and conclusion herein from each peptide to develop a better kinase sensor system are still ongoing in the lab.

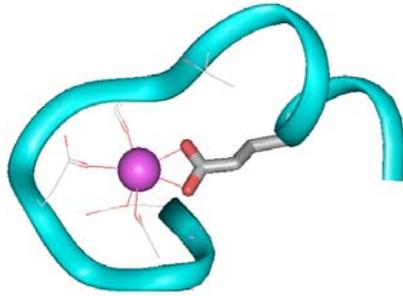
## Figures and Tables

Peptide	Sequences	Calculated Mass	[M+2H] <sup>2+</sup> observed
non-phosphorylated HM1	Ac-DKNADGWIDRRSIIAK-NH2	1899.1	950.4
phosphorylated HM1	Ac-DKNADGWIDRRpSIIAK-NH2	1979.1	990.4
non-phosphorylated HM2	Ac-DKNADGWIRRRSIIAK-NH2	1940.2	970.9
phosphorylated HM2	Ac-DKNADGWIRRRpSIIAK-NH2	2020.2	1010.8
non-phosphorylated HM3	Ac-DADADGWIRRRSIIAK-NH2	1884.1	943.0
phosphorylated HM3	Ac-DADADGWIRRRpSIIAK-NH2	1964.1	983.5
non-phosphorylated HM4	Ac-DANADGWIRRRSIIAK-NH2	1883.1	942.5
phosphorylated HM4	Ac-DANADGWIRRRpSIIAK-NH2	1963.1	982.6
non-phosphorylated HM5	Ac-DANADGWIRRASIIAK-NH2	1798.0	900.0
phosphorylated HM5	Ac-DANADGWIRRApSIIAK-NH2	1878.0	940.0
non-phosphorylated HM6	Ac-DADADGWIKKASIIAK-NH2	1743.0	872.5
phosphorylated HM6	Ac-DADADGWIKKApSIIAK-NH2	1823.0	912.6
non-phosphorylated HM7	Ac-DADADGWIRRASIIAK-NH2	1799.0	899.6
phosphorylated HM7	Ac-DADADGWIRRApSIIAK-NH2	1879.0	940.0
non-phosphorylated HM8	Ac-DKDADGWIRRASIIAK-NH2	1856.1	929.5
phosphorylated HM8	Ac-DKDADGWIRRApSIIAK-NH2	1936.1	968.9
non-phosphorylated HM9	Ac-DADADGWRSRRSIIAK-NH2	1858.0	929.6
phosphorylated HM9	Ac-DADADGWRSRRpSIIAK-NH2	1938.0	969.9

**Table 3.1.** Characterization data for non-phosphorylated and phosphorylated peptides. pS indicates phosphoserine.

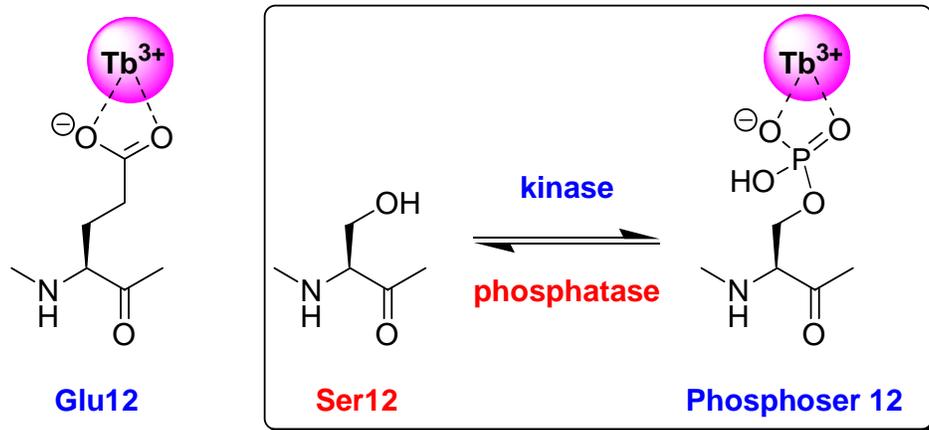
Peptide	Sequence	$K_d$ , $\mu\text{M}$	Standard Error ( $\pm$ )
non-phosphorylated HM1	Ac-DKNADGWIDRRSIIAK-NH2	334	29
phosphorylated HM1	Ac-DKNADGWIDRRpSIIAK-NH2	102	10
non-phosphorylated HM2	Ac-DKNADGWIRRRSIIAK-NH2	456	26
phosphorylated HM2	Ac-DKNADGWIRRRpSIIAK-NH2	549	3
non-phosphorylated HM3	Ac-DADADGWIRRRSIIAK-NH2	462	23
phosphorylated HM3	Ac-DADADGWIRRRpSIIAK-NH2	270	3
non-phosphorylated HM4	Ac-DANADGWIRRRSIIAK-NH2	624	71
phosphorylated HM4	Ac-DANADGWIRRRpSIIAK-NH2	317	19
non-phosphorylated HM5	Ac-DANADGWIRRASIIAK-NH2	214	10
phosphorylated HM5	Ac-DANADGWIRRApSIIAK-NH2	277	27
non-phosphorylated HM6	Ac-DADADGWIKKASIIAK-NH2	685	49
phosphorylated HM6	Ac-DADADGWIKKApSIIAK-NH2	173	25
non-phosphorylated HM7	Ac-DADADGWIRRASIIAK-NH2	366	35
phosphorylated HM7	Ac-DADADGWIRRApSIIAK-NH2	80	26
non-phosphorylated HM8	Ac-DKDADGWIRRASIIAK-NH2	492	37
phosphorylated HM8	Ac-DKDADGWIRRApSIIAK-NH2	79	44
non-phosphorylated HM9	Ac-DADADGWRSRRSIIAK-NH2	534	18
phosphorylated HM9	Ac-DADADGWRSRRpSIIAK-NH2	122	14

**Table 3.2.** Dissociation constants of peptide-Tb<sup>3+</sup> complexes in aqueous solutions of 5 mM HEPES (pH 7.5), 100 mM NaCl, and 2 mM MgCl<sub>2</sub>. pS indicates phosphoserine.

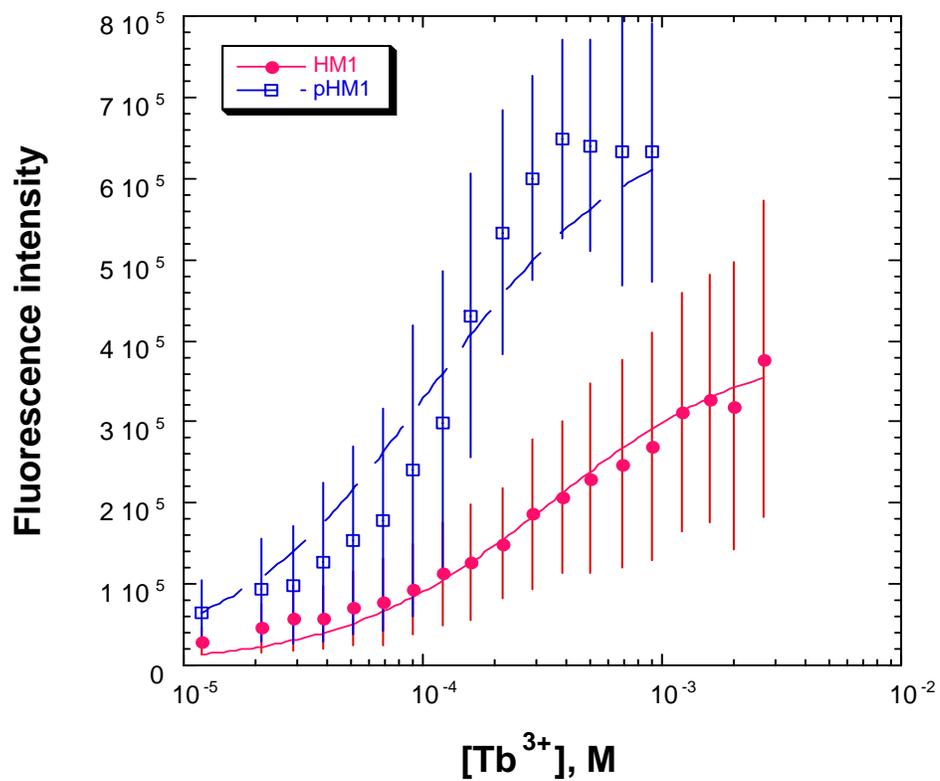


EF Hand	DKNADG Y IDAAELK
pKID-PKA	DKNADG W IDRASLA

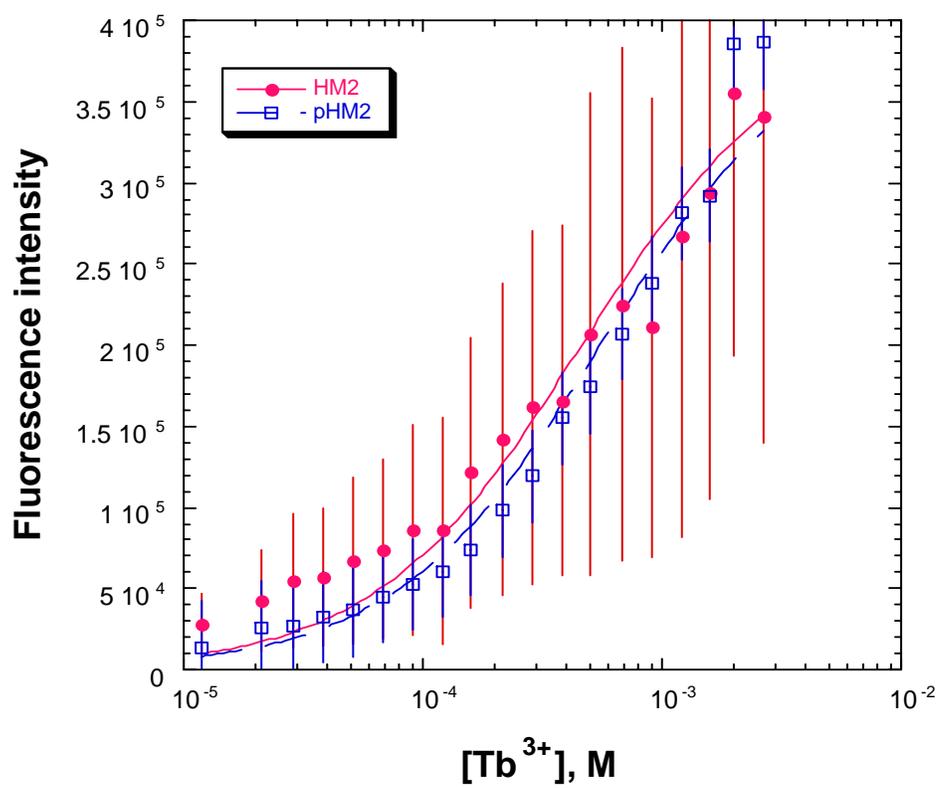
**Figure 3.1.** Left: Structure of a single EF hand calcium-binding loop (1cll), with Glu12 side chain, bound to  $\text{Ca}^{2+}$  in a bidentate manner, emphasized. Right: EF hand consensus sequence and pKID-PKA sequence. Side chains of residues in red contact metal. Trp (magenta) is used as the sensitizer of lanthanide luminescence. Glu12 in EF is replaced by Ser12 (green) in pKID-PKA. Ser12 may be phosphorylated or non-phosphorylated.



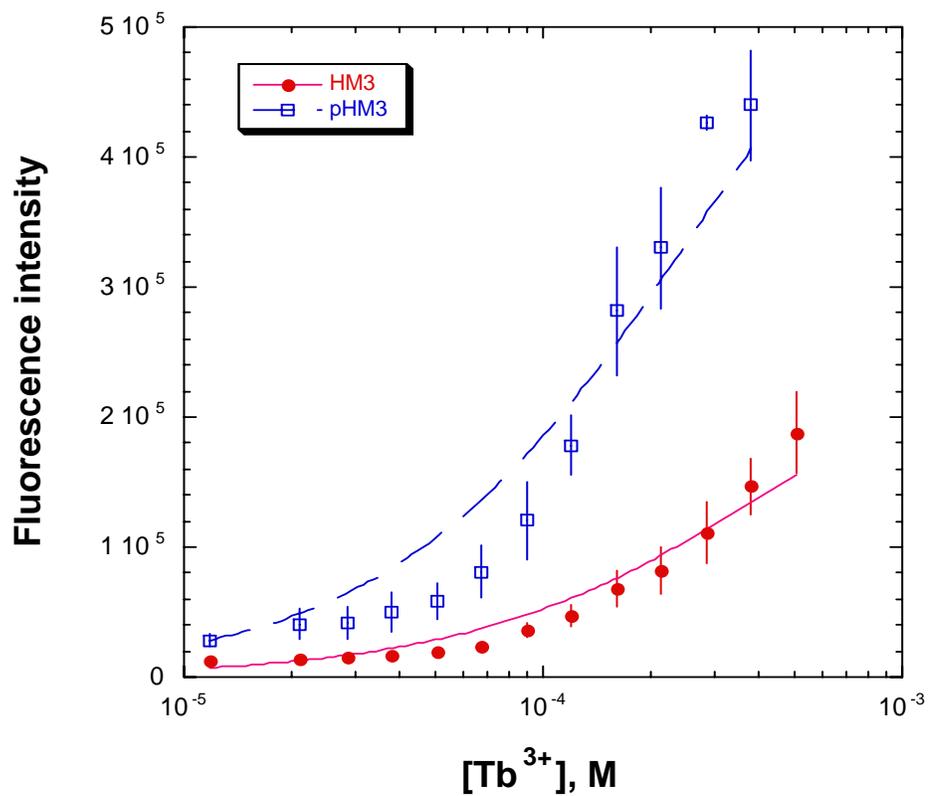
**Figure 3.2.** Design of a phosphorylation-dependent motif, indicating the roles of Glu, Ser, and phosphoserine residues.



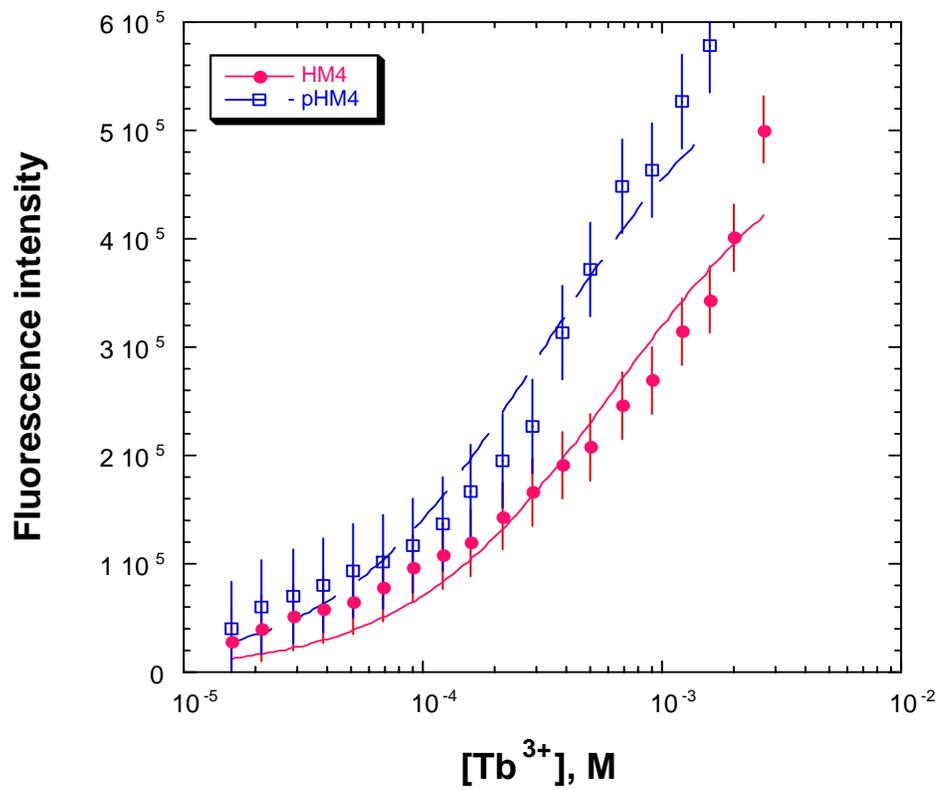
**Figure 3.3.** Titration of 10  $\mu\text{M}$  non-phosphorylated (red circle) HM1 and phosphorylated (blue square) HM1 with  $\text{Tb}^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



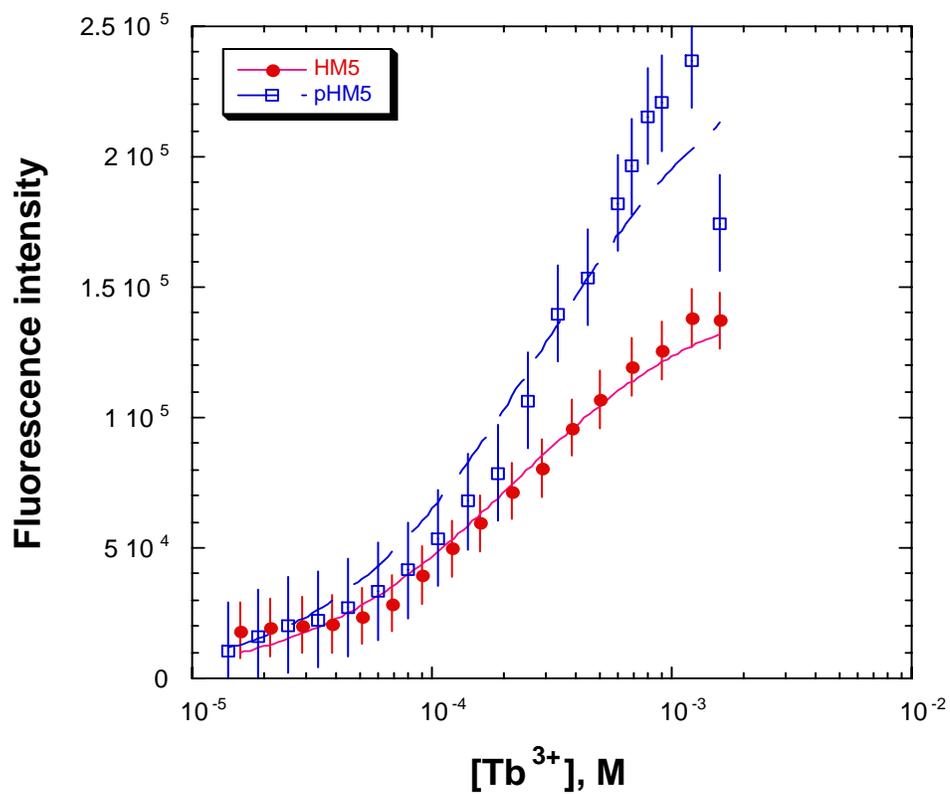
**Figure 3.4.** Titration of 10  $\mu\text{M}$  non-phosphorylated (red circle) HM2 and phosphorylated (blue square) HM2 with  $\text{Tb}^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



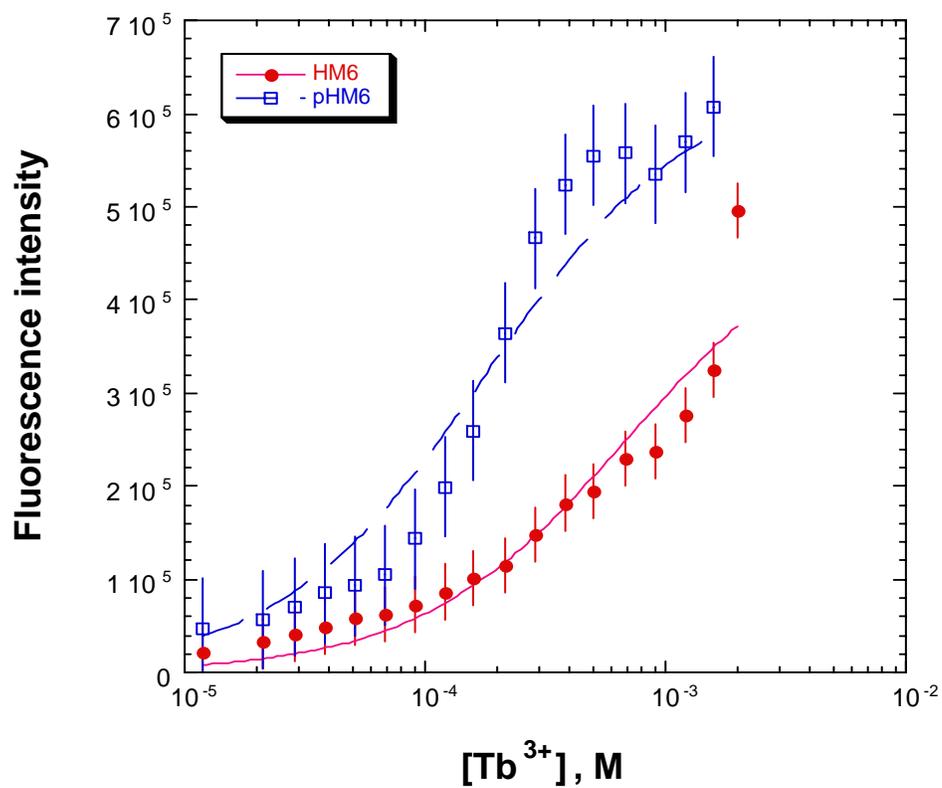
**Figure 3.5.** Titration of 10  $\mu\text{M}$  non-phosphorylated (red circle) HM3 and phosphorylated (blue square) HM3 with  $\text{Tb}^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



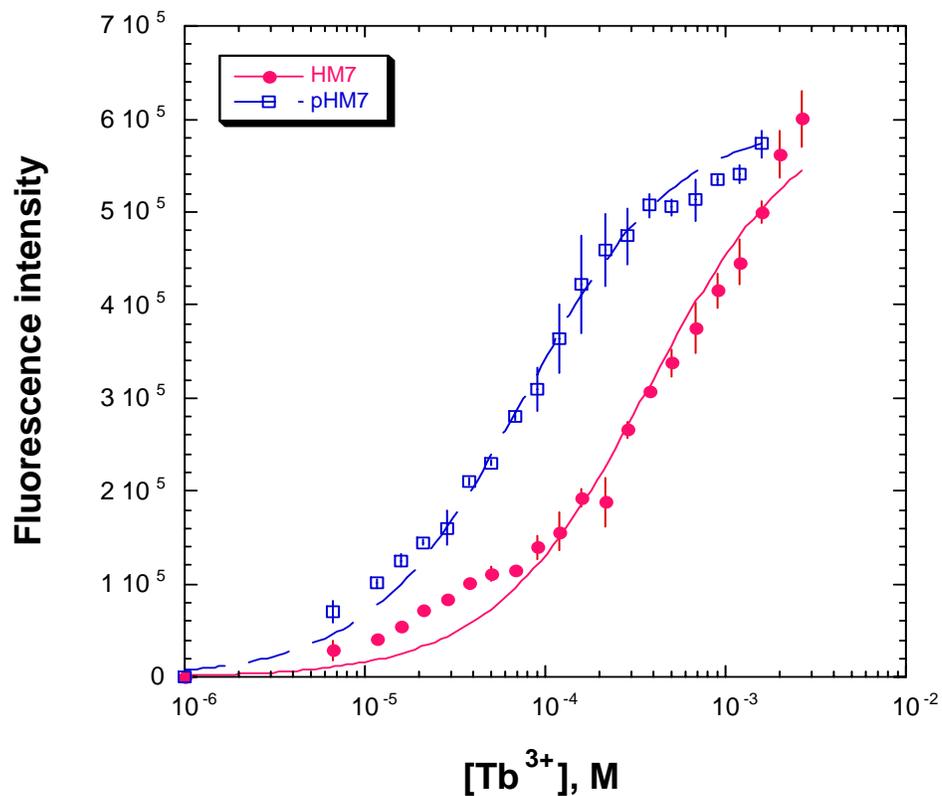
**Figure 3.6.** Titration of 10  $\mu M$  non-phosphorylated (red circle) HM4 and phosphorylated (blue square) HM4 with  $Tb^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



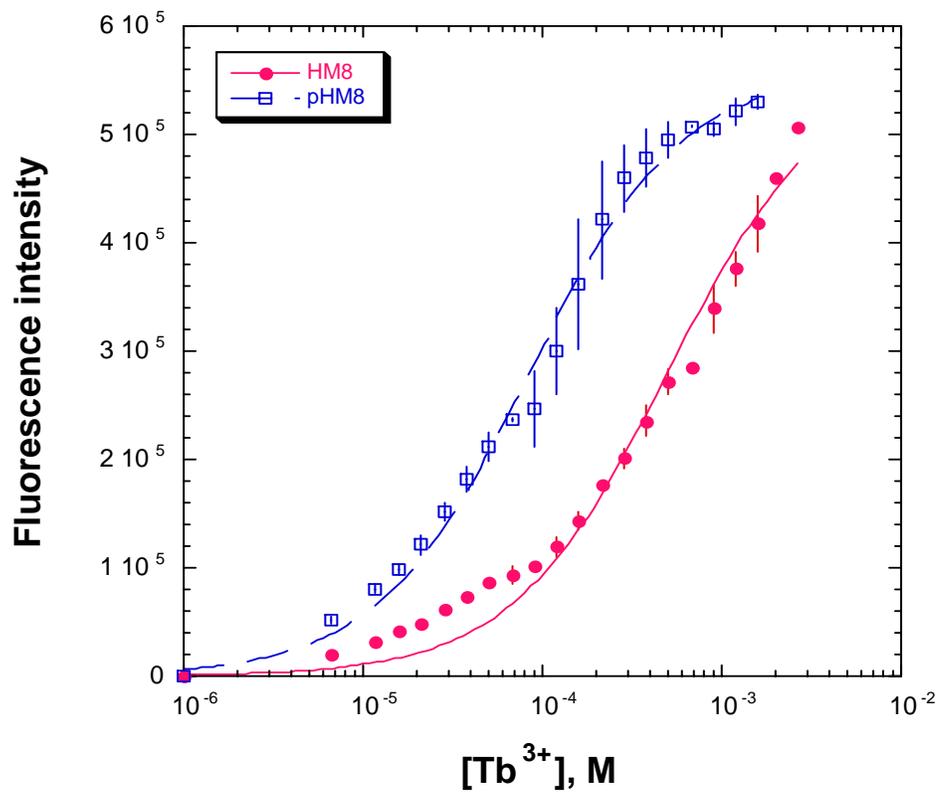
**Figure 3.7.** Titration of 10  $\mu$ M non-phosphorylated (red circle) HM5 and phosphorylated (blue square) HM5 with  $Tb^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



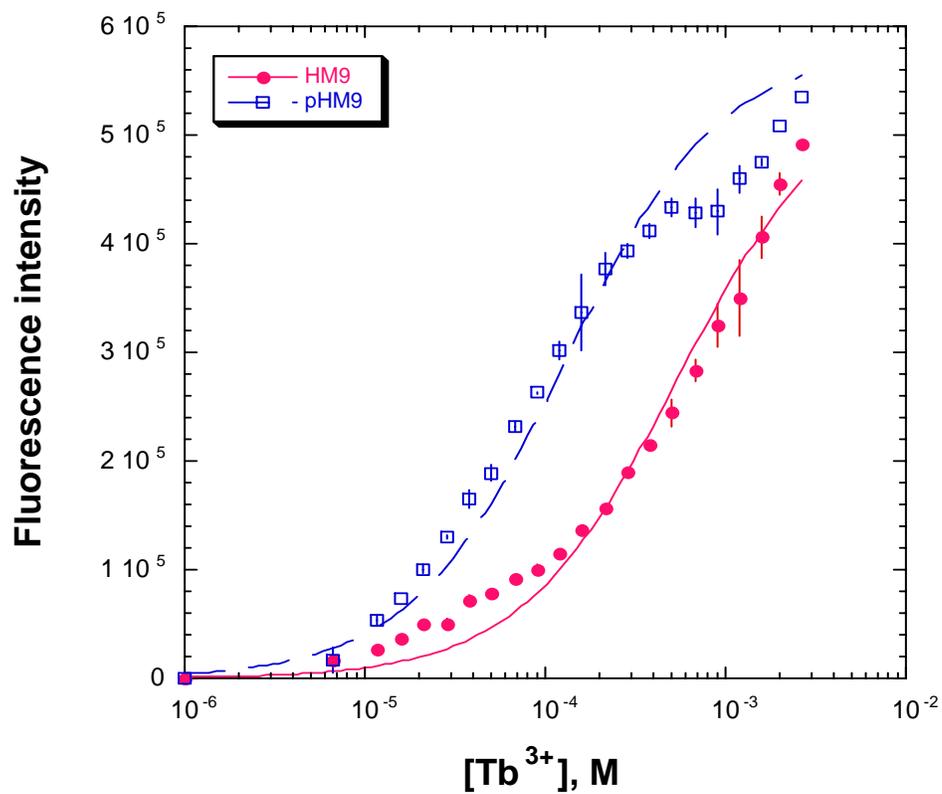
**Figure 3.8.** Titration of 10  $\mu\text{M}$  non-phosphorylated (red circle) HM6 and phosphorylated (blue square) HM6 with  $\text{Tb}^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



**Figure 3.9.** Titration of 10  $\mu M$  non-phosphorylated (red circle) HM7 and phosphorylated (blue square) HM7 with  $Tb^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



**Figure 3.10.** Titration of 10  $\mu$ M non-phosphorylated (red circle) HM8 and phosphorylated (blue square) HM8 with  $Tb^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.



**Figure 3.11.** Titration of 10  $\mu$ M non-phosphorylated (red circle) HM9 and phosphorylated (blue square) HM9 with  $Tb^{3+}$ . The curves show the fits to equation (2). Error bars indicate the standard error of corresponding points.

## References

1. Bridges, A.J. *Chem. Rev.* 2001, 101, 2541-2571..
2. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* 2002, 298, 1912-.
3. Chen, C.-A.; Yeh, R.-H.; Lawrence, D. S. *J. Am. Chem. Soc.* 2002, 124, 3840-3841.
4. Veldhuyzen, W. F.; Nguyen, Q.; McMaster, G.; Lawrence, D. S. *J. Am. Chem. Soc.* 2003, 125, 13358-13359.
5. Lawrence, D. S. *Accounts of Chemical Research* 2003, 36, 401-409.
6. Shults, M. D.; Imperiali, B. *J. Am. Chem. Soc.* 2003, 125, 14248-14249.
7. Vazquez, M. E.; Nitz, M.; Stehn, J.; Yaffe, M. B.; Imperiali, B. *J. Am. Chem. Soc.* 2003, 125, 10150-10151.
8. Zhang, J.; Ma, Y. L.; Taylor, S. S.; Tsien, R. Y. *Proc. Natl. Acad. Sci. USA* 2001, 98, 14997-15002.
9. Ting, A. Y.; Kain, K. H.; Klemke, R. L.; Tisen, R. Y. *Proc. Natl. Acad. Sci. USA* 2001, 98, 15003-15008.
10. Violin, J. D.; Zhang, J.; Tisen, R. Y.; Newton, A. C. *J cell Biol* 2003, 161, 899-909.
11. Franz, K. J.; Nitz, M.; Imperiali, B. *ChemBiochem* 2003, 4, 165-271.
12. Nitz, M.; Sherawat, M.; Franz, K. J.; Peisach, E.; Allen, K. N.; Imperiali, B. *Angew. Chem. Int. Ed.* 2004, 43, 3682-3685.
13. Nitz, M.; Franz, K. J.; Maglathlin, R. L. Imperiali, B.; *ChemBiochem* 2003, 4, 272-276.
14. Wohnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. J. *J. Am. Chem. Soc.* 2003, 125, 13338-13339.
15. Ohguro, H.; Palczewski, K. *FEBS Letters* 1995, 368, 452-454.
16. Zhou, S. Y.; Blenchner, S.; Hoagland, N.; Hoekstra, M. F.; Piwnicaworms, H.; Cantley, L. C. *Curr Biol* 1994, 4, 973-982.