Studies in immobilized metal ion affinity chromatography of proteins

A thesis presented for the degree of

Doctor of Philosophy (Ph.D.)

by

Gabriel Stephan Alexander Leone Widakowich

Centre for Green Chemistry
Monash University
Clayton, Victoria
Australia

September 2010

DECLARATION

To the best of my knowledge and belief this thesis contains no material which has been accepted for the award of any degree or diploma in any university or other institution, nor any material previously published or written by any other person except where due reference is made in the text.

Notice 1

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

Notice 2

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Gabriel Widakowich

…….

Acknowledgements

First I would like to thank my supervisor, Professor Milton T. W. Hearn for giving me the opportunity of performing these studies and for his supervision and support throughout my studies.

I would also like to thank Dr Chunfang Zhang for acting as a co-supervisor and being helpful and supportive from the first day I stepped into the laboratory. The help and co-supervision of Dr Dale Fredericks is also greatly appreciated. I am also grateful for all help I received from Tamsin Sawford with expression work and sample preparation.

I would also like to thank Eva Campi for the help with preparing and immobilizing ligands and for helpful discussions. The help from Kirk-Ski Troung, Tracy Warner and Khosse Mitri with protein purification is also greatly appreciated.

I would also like to thank Dr Simon Harris for help with MALDI-ToF experiments and for proofreading this thesis. The help received from Dr Reinhard Boysen with creating graphs in Sigma-Plot, and all relevant articles provided, is greatly appreciated.

I would like to thank Professor Stephen Bottomley and Dr Glenn Powers for assistance with the Circular Dichroism measurements. The help from Dr William Yang with electrospray mass spectrometry is greatly appreciated. I am also grateful for the help I received from Mrs Tina Hines with the AAS experiments.

I would also like to thank my fellow PhD students Agron Mataj, Anjali Bhagwat, Darant Longford, Ika Oktavianawati, Martin Petzold, Campbell Coghlan, Wing Wei Rong, Yididya Banti, and all other staff and members of the Centre for Green Chemistry for being great colleagues and friends.

To all my extended family and to my friends, thank you so much for your support and encouragement!

Finally, with all my heart I thank my fiancée Karin Gertow for being so understanding and encouraging during my time as a student. Karin – I could not have done this without your! I love you.

List of abbreviations

AAS	Atomic absorption spectroscopy
AHM	Aminohydroxamate
Alendronate	(4-amino-1-hydroxy-1-phosphono-butyl)phosphonic acid
AMP	Aminopeptidase M
APS	Ammonium persulfate
BSA	Bovine serum albumin
CD	Circular dichroism
CHES	N-Cyclohexyl-2-aminoethanesulfonic acid
CIP	Calf intestine phosphatase
CM-Asp	Carboxymethylated aspartic acid
CPA	Carboxypeptidase A
CPB	Carboxypeptidase B
CV	Column volumes
Cyclen	1,4,7,10-tetraazacyclododecane
5,5'-dibromo-BAPTA	5,5'-dibromo-1,2-bis(o-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid
DIT	Digital integration time
DO2A	1,4,7,10-tetraazacyclododecane-1,7-diacetic acid
DO3A	1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid
DOTA	1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid
DO2P	1,4,7,10-tetraazacycloddoecane-1,7-di(methanephosphonic acid)
DO3P	1,4,7,10-tetraazacyclododecane-1,4,7-tri(methanephopshonic acid)
DOTP	1,4,7,10-tetraazacyclododecane-1,4,7,10-tetra(methanephopshonic acid)
DPA	di(pyridylmethyl)amine
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EGTA	Ethylene glycol tetraacetic acid
ESI-MS	Electrospray ionization mass spectrometry

FP	Fluorescent protein
FPLC	Fast protein liquid chromatography
FRET	Fluorescence resonance energy transfer
GFP	Green fluorescent protein
GuHCl	Guanidine hydrochloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEWL	Hen egg white lysozyme
HIC	Hydrophobic interaction chromatography
HMYO	Horse skeletal muscle myoglobin
HPLC	High performance liquid chromatography
8-HQ	8-hydroxyquinoline
HRP	Horse reddish peroxidase
ICP-AES	Inductively coupled plasma atomic emission spectrometry
IEX	Ion exchange chromatography
IDA	Iminodiacetic acid
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl-β-D-1-thiogalactoside
LB	Lysogeny Broth
LRET	Luminescence resonance energy transfer
MALDI-ToF	Matrix-assisted laser desorption/ionization time of flight
MBP	Maltose Binding Protein
MIT	Metal ion transfer
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
MES	2-(N-morpholino)ethanesulfonic acid
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
NusA	N-utilization substance A
OPS	O-phosphoserine
Pamidronate	(3-amino-1-hydroxy-1-phosphonopropyl)phosphonic acid
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDC	Pentadentate chelator
Propyl-*bis*-tacn	*bis*(1,4,7-triazacyclonon-1-yl)propane
RPC	Reverse phase chromatography
RT	Room temperature
Sepharose 6 F F	Sepharose™ 6 Fast Flow

S. cerevisiae	*Saccharomyces cerevisiae*
SAP	Human serum amyloid P component
SEC	Size exclusion chromatography
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Tacn	1,4,7-triazacyclononane
TBAOH	Tetrabutyl ammonium hydroxide
TBS	Tris buffered saline
TED	Tris(carboxymethyl)ethylenediamine
TEMED	N,N,N',N'-tetramethylene-diamine
TEV	Tobacco etch virus
Tris	Tris(hydroxymethyl)aminomethane
λ	Wavelength
λ_{exc}	Excitation wavelength
λ_{emm}	Emission wavelength

Abbreviations of amino acids

Name	Three letter abbreviation	One letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Table of contents

Abstract

Immobilized metal ion affinity chromatography (IMAC) is a powerful tool for purification of proteins from complex mixtures. The technique is based on the interaction of electron donor groups on the surface of proteins with immobilized metal chelates. Currently, many IMAC systems employ borderline metal ions such as Ni^{2+}, Cu^{2+} and Co^{2+}, according to the classification of Pearson, chelated to linear ligands like iminodiacetic acid (IDA) or nitrilotriacetic acid (NTA). The stability constants of Ni^{2+} and Cu^{2+} chelates with IDA and NTA are relatively low (log β = 8-13). As a result, metal ions can leak during the chromatographic process. This is undesirable due to the toxicity of borderline metal ions.

The aim of this thesis was to contribute to the development of IMAC systems without leakage of borderline metal ions. Two strategies were employed – using borderline metal ions but minimizing their leakage, and employing the more benign metal ion Ca^{2+}.

Strategy 1 was pursued through studies of IMAC systems composed of the ligands 1,4,7-triazacyclononane (tacn) and *bis*(1,4,7-triazacyclonon-1-yl)propane (propyl-*bis*-tacn), forming stable metal complexes with Ni^{2+} and Cu^{2+} (log β ~ 16), and the affinity tag NT1A. To facilitate sequential purification, the NT1A tag was fused in sequence with the putative hard metal binding tag HIT2. Two studies were performed to test the suitability of the affinity tags NT1A and HIT2.

A. The arrangement of the NT1A and HIT2 tags with respect to the target protein was studied. Previously observed differences in expression levels of N-terminally tagged enhanced green fluorescent protein (EGFP) were not observed herein, when the tags were instead fused to the C-terminus of the protein. In IMAC experiments, C-terminally tagged EGFP's were however eluted at a slightly lower imidazole concentration compared to their N-terminally tagged counterparts.

B. The influence of the tags NT1A and HIT2 on the structure and stability of EGFP was studied by guanidine hydrochloride induced unfolding experiments. NT1A and HIT2 were not found to affect the structure and stability of EGFP, confirming their suitability as affinity tags.

Strategy 2 was pursued by studying two fundamental aspects of Ca^{2+} based IMAC.

A. Immobilized Ca^{2+} chelates were studied by packing ligands immobilized to Sepharose™ 6 Fast Flow into columns and subjecting them to chromatography experiments. The macrocyclic ligands 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid (DO3A) and 1,4,7,10-tetraazacyclododecane-1,4,7-tri(methanephopshonic acid) (DO3P) were compared with the linear ligands (4-amino-1-hydroxy-1-phosphono-butyl)phosphonic acid (alendronate)

and IDA. NaCl was found to displace bound Ca^{2+} from all ligands investigated. For Ca^{2+}-DO3P, pH values below 7.5 also caused displacement of Ca^{2+} ions.

B. The Ca^{2+} affinity to putative Ca^{2+} binding tags was studied. Novel disulfide bridge constrained cyclic tags fused to EGFP were successfully developed and purified, but their expression levels were low. One cyclic tag (previously developed by Imperiali *et al*.) bound the Ca^{2+} analogue Tb^{3+} ($\log \beta = 5.2$). As the Tb^{3+} affinity decreased dramatically when the disulfide bridge was broken by a reducing agent, cyclic tags present promising candidates as Ca^{2+} binding tags. In conclusion, both strategy 1 and 2 generated results of relevance to the overall aim of the thesis.

Chapter 1. Introduction

1.1 General Introduction

In the field of proteomics, protein of high purity is required in a number of biochemical applications such as protein identification, enzymology, biophysical analyses and three dimensional structure determinations. In addition, highly pure protein in large quantities is in demand in the pharmaceutical industry. Purifying proteins from complex mixtures such as homogenized tissue or cell lysates can be a technically challenging, costly and material-consuming multi-step procedure. Hence, there is a need for rapid, cost efficient methods to obtain proteins of high purity.

A number of chromatographic principles can be employed to purify proteins based on their intrinsic physical properties. In ion exchange chromatography (IEX), charged groups on the surface of proteins interact with oppositely charged groups on the stationary phase. In hydrophobic interaction chromatography (HIC) and reverse phase chromatography (RPC), hydrophobic patches of proteins interact with a hydrophobic stationary phase. In size exclusion chromatography (SEC), proteins are included or excluded from pores of the stationary phase and hence separated based on their size[1].

Many proteins in a complex mixture can have similar properties like size and/or hydrophobicity and/or charge distribution on their surface. As a consequence the number of sequential purification steps undertaken to obtain an individual protein of high purity can be large if IEX, HIC, RPC and SEC methods are used.

Affinity chromatography is on the other hand based on highly specific interactions such as antibody-antigen recognition. A protein of high purity can therefore often be obtained by a single affinity chromatography step[2].

Immobilized metal ion affinity chromatography (IMAC) is a subclass of affinity chromatography. The technique is based on interactions between metal ions immobilized onto a solid support using a chelating ligand and electron donating groups of proteins or peptides. The principle of using immobilized metal ions to separate metalloproteins was first adapted by Everson and Parker[3] in 1974. Porath *et al*.[4] introduced the concept of IMAC in 1975 under the name of metal chelate affinity chromatography and the technique has since gained popularity. The immobilized metal ion, commonly a borderline metal ion such as Ni^{2+}, Cu^{2+}, Co^{2+} or Zn^{2+} according to the definition of Pearson[5], acts as an electron acceptor, forming covalent coordination bonds with electron donor groups of proteins or peptides. Side chains of a range of amino acids such as His, Glu, Asp, Tyr, Cys, Arg, Lys and Met can act as electron donor groups in IMAC[6]. However, the nitrogen on the imidazole ring of the histidine (His)

side chain form the strongest bonds with immobilized borderline metals ions. Free reduced Cys residues that could contribute to binding to immobilized metal ions are rarely found on the surface of proteins[6; 7]. Early work was focused on proteins containing metal binding residues on their surface.

A significant development was the concept of introducing a polypeptide sequence containing metal binding residues such as His at the N- or C-terminus of a protein. These polypeptides, commonly referred to as affinity tags can easily be fused to the protein using recombinant DNA technology[8]. One of the most commonly used affinity tags in IMAC is the hexa His tag[9]. The affinity tag approach has the advantage of offering generic purification protocols. Many different proteins can be purified using the same affinity tag employing similar chromatographic purification protocols. IMAC is therefore the preferred purification strategy by many laboratories today[10]. After purification of the target protein, the affinity tag can be removed by a number of methods. Chemical treatment with for example hydroxylamine or cyanogens bromide can be employed, though this treatment is harsh. Enzymatic cleavage is milder and more specific. Removing the tag is usually desirable as it can affect the biochemical properties such as folding, conformation, solubility, enzymatic and biological activity of the target protein[11].

1.2 An overview of IMAC systems

An IMAC system consists of five main parts: a solid support, a linker or spacer, the chelating ligand, the metal ion and the protein surface or affinity tag. A schematic of these components is shown in Figure 1.1.

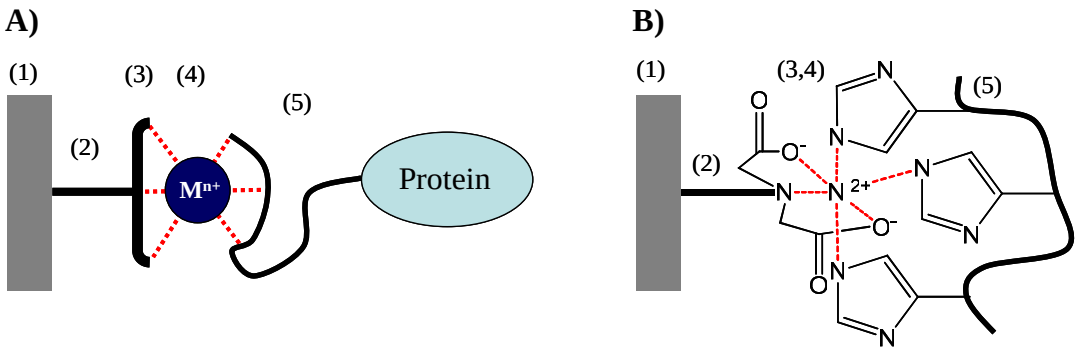


Figure 1.1A) Schematic drawing of the components of an IMAC system. **B)** Ni^{2+}-IDA as an example of an immobilized metal-chelate complex. (1) The solid support. (2) The spacer arm. (3) The ligand. (4) The metal ion M^{2+}. (5) A protein or an affinity tag attached to a protein. In schematic a) a tagged protein is depicted. Electron donating groups can also constitute amino acid side chains on the surface of a protein.

The ligand is coupled to a solid support such as Sepharose 6 F F via a spacer. When the ligand has been attached the metal of choice is chelated. This is achieved by passing an aqueous solution of the corresponding metal salt (commonly $M^{n+}Cl_n$, $M^{n+}(NO_3)_n$ or $M^{2n+}(SO_4)_n$) through the stationary phase. The metal ion will then displace any water or buffer molecules bound to the ligand. The immobilized ligand charged with the metal ion can be referred to as an immobilized metal chelate. When chelated, the metal ion will form coordination bonds both with the ligand and with surrounding solvent molecules or ions. The number of coordination bonds formed between the metal ion and the ligand depends on the number of electron donor groups provided by the ligand. When a target protein is introduced into the IMAC system, electron donor groups (usually from amino acid side chains) will form coordination bonds to the immobilized metal ion. The coordinated solvent molecules will then be displaced. The number of coordination bonds between the metal ion and the ligand should generally exceed the number of bonds between the metal ion and the electron donor groups on the protein. If this is not the case the metal ion will be coordinated more strongly by the protein than by the ligand. The metal can then be displaced from the ligand by the protein.

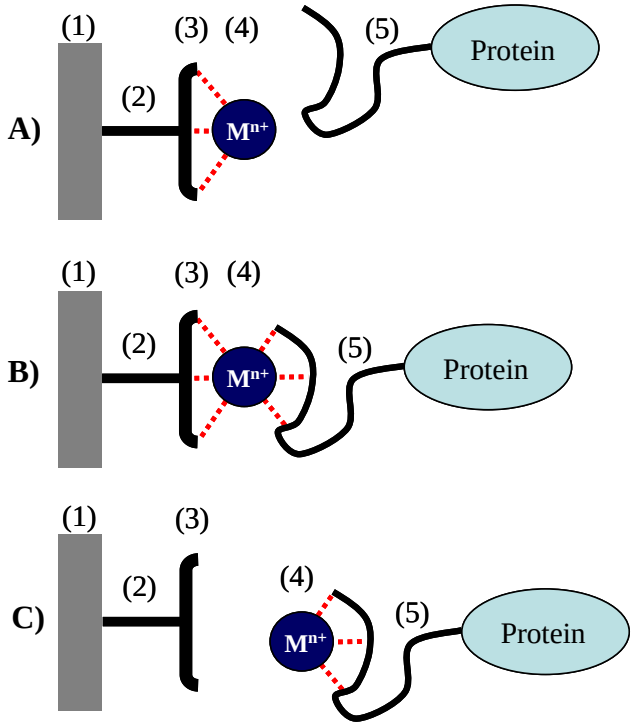


Figure 1.2. Schematic diagram representing MIT. (1) The solid support. (2) The spacer arm. (3) The ligand. (4) The metal ion M^{2+}. (5) The affinity tag attached to a protein. **A)** A tagged protein is introduced onto a chromatography system containing an immobilized metal chelate. **B)** The tagged protein binds to the metal chelate. **C)** The metal is displaced from the ligand by the tagged protein.

This phenomenon is known as Metal Ion Transfer (MIT). A schematic drawing of MIT is shown in Figure 1.2.

The example shown in Figure 1.2 depicts a process occurring with a tagged protein. It should be stressed that MIT can occur both with tagged and untagged proteins. Elution of the protein can be achieved in several ways. The pH can be lowered until the donor groups of the protein are protonated. By taking advantage of the protic equilibrium the coordination bonds between the donor groups and the metal ion will be broken. Elution of proteins bound to immobilized metal chelates *via* His residues can also be achieved by ligand exchange with imidazole, an analogue of the His side chain. A metal chelating agent such as ethylenediaminetetraacetic acid (EDTA) can also be employed. The chelator will form a complex with the metal ion. This will lead to elution of the protein.

As in other modes of chromatography secondary interactions can occur in IMAC. The ligand, the spacer arm or the solid support can interact with proteins and peptides in unpredictable ways through hydrophobic, polar or ionic interactions. Secondary interactions are generally considered as undesirable in chromatography and the goal is often to minimize their contribution.

1.2.1 The solid support

A wide range of solid supports have been developed and used for different chromatographic applications. Solid supports can either be spherical particles or monoliths, the former being more widely used today. Some chromatography supports used in IMAC are listed in Table 1.1.

One commonly used matrix is cross linked agarose, a polymer of the disaccharide agarobiose consisting of D-galactose and 3,6-anhydro-L-galactose sold under the trade name[12] Sepharose™. Several properties are of importance when considering the choice of the solid support. The choice of solid support is dictated by the application for which it is intended.

The mechanical stability will dictate the pressure range over which the material can be used. Materials with high mechanical stability are therefore used for high performance liquid chromatography (HPLC) applications. Examples include silica- and polystyrene divinylbenzene polymer-based materials. For low pressure applications the mechanical requirements of the solid support are lower. Materials such as agarose can therefore be used in low pressure applications like fast protein liquid chromatography (FPLC).

The chemical stability of the solid support will determine which mobile phase compositions can be used. This will dictate both the choice of eluents employed in the chromatography experiment and the choice of solvents employed for cleaning and regeneration of the solid support.

The particle size influences the mass transfer between the flow stream in the column and the interior of the particle. Small particles promote fast mass transfer and hence higher resolution.

Table 1.1. Examples of solid supports commonly used in chromatography.

Base material	Advantage(s)	Disadvantage(s)	Examples of commercially available brands
Agarose	Relatively pH stable (pH 3-12)	Mechanically relatively soft	Sepharose™ (GE Healthcare) Affi-Gel® (Bio-Rad) NTA-Agarose (Qiagen)
Silica	High mechanical stability, narrow pore size distribution	Dissolves at basic pH (unless modified). Reactive silanol groups can bind biological molecules irreversibly	Bakerbond™ (J.T. Baker) ProteinPak™ (Waters) TSK-Gel® (Toso Haas)
Poly(styrene divinyl benzene) (PSDVB)	High pH stability (pH 1-14)	Hydrophobic (unless surface modified)	ProPac® (Dionex)
Polymethacrylate	Relatively pH stable (pH 1-12)		TSK-Gel® (Toso Haas) Fractogel® (Merck/Novagen®)

The disadvantage of small particles is that they give rise to high back pressures. Particle sizes can range from below 2 to 400 µm and are chosen based on both the pressure capabilities of the chromatographic system (pumps, tubing, detector etc) and the scale of the separation[2]. Solid supports are often porous in order to increase the surface area of the material. The size of the pore should be sufficiently large to facilitate access for the molecule(s) to be separated. On the other hand larger pore sizes result in a lower surface area and hence lower capacity of the stationary phase. Pores should as a rule of thumb have a diameter five times the diameter of the solute. A typical protein has a diameter of 60 Å. Pore sizes of 300 Å or above are therefore commonly used in separation of biological molecules with chromatographic adsorbents[2].

1.2.2 The spacer

Ligands can sometimes not be linked directly to solid supports. In these cases the solid support can be "activated" by covalent attachment of a molecule which in turn contains

a reactive centre with the ability to form a bond with the ligand. The attachment of the ligand onto a solid support via an activating compound serves several purposes.

Firstly, a means of coupling the free ligand to the solid support is provided. The activating compound should be chosen so that it forms stable bonds with the gel and with the ligand, the attachment reaction is preceded with high yield and no groups providing potential non-specific interactions are introduced.

Secondly, the activating compound can crosslink the polymer chains, thereby improving the properties of the base material[13]. Sundberg and Porath employed 1,4-butanediol-diglycidyl ether to crosslink agarose gel and introduce reactive oxirane groups for ligand immobilization[13]. Agarose was also cross linked with divinyl sulphone yielding highly stable gel[14]. Epichlorohydrin (chloromethyloxirane) employed by Axen and Porath[15], has been widely used in IMAC to activate gels for ligand immobilization. A schematic diagram of the activation of Sepharose™ gel with epichlorohydrin and immobilization of a ligand via a secondary amine is shown in Figure 1.3. This immobilization diagram represents the procedure employed in our laboratory for activation of solid supports (such as Sepharose 6 F F) and immobilization of ligands.

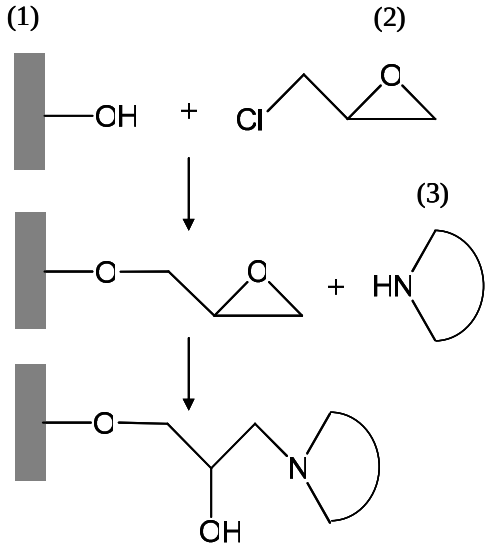


Figure 1.3. Schematic representation of the activation of Sepharose™ 6 Fast Flow (1) with epichlorohydrin (2) and the immobilization of a ligand containing a secondary amine (3).

Thirdly, the spacer arm can increase the steric availability of the metal chelate so that it can interact with a protein surface or affinity tag more readily. As an example, IDA was attached to a membrane with and without 1, 4-butanediol-diglycidyl ether. The capacity of both membranes was 2.25 mg lysozyme. The dynamic adsorption was determined using breakthrough curves. Dynamic adsorption of lysozyme was 0.42 mg higher for Cu^{2+}-IDA coupled to the membrane via a 1,4-butanediol-diglycidyl ether spacer compared to Cu^{2+}-IDA attached directly onto the membrane[16].

1.2.3 The metal ion

1.2.3.1 The concept of hard, borderline and soft metal ions

Metal ions are commonly classified as hard, borderline (also referred to as intermediate) or soft according to the classification of Pearson[5]. The terms hardness and softness refer to the properties of atoms or molecules to interact as Lewis acids and bases. The classification was first mentioned in the literature[17] in 1963. Pearson defined the terms hardness and softness as follows. Hard acids have high positive charge, small radius and no easily polarizable outer electrons whereas soft acids have low positive charge, large size and polarizable outer electrons. Hard bases have high electronegativity, are hard to oxidize and have low polarizability and only high energy empty orbitals. Soft bases have low electronegativity, are easily oxidized and highly polarizable, with low-lying empty orbitals[5]. Hence, when forming donor-acceptor complexes, hard metal ions form coordination bonds with ionic characteristics whereas soft metal ions form bonds of more covalent characteristics[5]. Coordination bonds formed between hard acids and bases are therefore more charge dependent then bonds formed between soft acids and bases. Pearson further developed the concept by attempting to quantify relative hardness and softness and correlating the quantities to molecular orbital theory[18]. The softness parameter σ_p was tabulated by Ahrland[19]. The parameter was defined by Pearson and Mawby[20] according to Equation 1.1 where E denotes the coordinate bond energy between the metal ion and a fluoride or iodide ion. The value of σ_p decreases with decreasing hardness (or increasing softness).

Equation 1.1 $$\sigma_p = (E_{fluoride} - E_{iodide})/E_{fluoride}$$

It has been found that hard metal ions such as Ca^{2+}, Fe^{3+}, Mg^{2+} and Al^{3+} show a preference for binding to oxygen atoms. Borderline metal ions such as Cu^{2+}, Ni^{2+}, Zn^{2+} and Co^{2+} tend to favor binding to nitrogen atoms. Soft metal ions like Cu^{+}, Hg^{+}, Ag^{+} have a preference to binding to sulfur atoms[7].

It should be stressed that factors other than hardness and softness, such as charge, will affect the coordination bond strength between a metal ion and a ligand. Furthermore, in the case of the ligand, the relative hardness or softness of different N, O or S atoms can be very difficult to assess. At an atom scale the relative hardness or softness is local and will vary over the molecule, which complicates the situation particularly for larger molecules[21].

1.2.3.2 Toxicity of metal ions

The toxicity of metal ions is a key consideration when designing a novel IMAC system with a Green Chemistry perspective. Several correlations between the

hardness/softness of metal ions and their toxicity have been made. For example, Jones *et al*. studied the correlation between softness and toxicity of metal ions on animals[22] whilst McCloskey *et al*. performed correlation studies using the commercially available and well established Microtox® assay[23]. A brief summary of the latter study is presented in Table 1.2, with metal ions of relevance in this thesis.

Table 1.2. Toxicity and relative softness, σ_p, of metal ions described as EC50 of free metal (μM/L) using the Microtox® assay, data from ref[23].

Metal ion	EC50 (±SD)	σ_p	Classification
K^+	468931 (±26 147)	0.232	Hard
Na^+	368920 (±7895)	0.211	Hard
Ca^{2+}	56821 (±9748)	0.181	Hard
Mg^{2+}	54090 (±3820)	0.167	Hard
La^{3+}	322 (±31)	0.171	Hard
Fe^{3+}	22 (±1)	0.103	Hard
Zn^{2+}	28 (±6)	0.115	Borderline
Co^{2+}	761 (±73)	0.130	Borderline
Ni^{2+}	436 (±39)	0.126	Borderline
Cu^{2+}	1.23 (±0.10)	0.104	Borderline

Some common ions such as Na^+ and K^+ are included for comparison. Both studies found correlations between toxicity and softness concluding that hard metals are less toxic than soft metal ions. It can be seen in Table 1.2 that the toxicity of Ca^{2+}, though higher than Na^+ and K^+, still is low compared to the borderline metal ions, Cu^{2+} in particular with an EC50 value of 1.23. In addition, Ni^{2+} and Co^{2+} are known allergens[24] and Ni^{2+} compounds are human carcinogens[25].

Most metal ions can form complexes with chelating ligands. Hence, in principle a wide range of metal ions could be used in IMAC. Currently borderline metal ions such as Ni^{2+}, Cu^{2+}, Zn^{2+} and Co^{2+} and the hard metal ion Fe^{3+} are most widely used[6]. Hence, there is a significant scope to develop IMAC systems based on metal ions not currently used.

1.2.4 Ligands used in IMAC

1.2.4.1 Basic concepts

The term chelation refers to a ligand forming two or more bonds with one metal ion. An important property of a ligand is its denticity – that is the number of potential bonds it can form with a metal ion. Metal chelates are more stable than the corresponding metal complex

formed with a monodentate ligand. This is due to the so called *chelate effect* and can be understood in terms of entropy effects. Simply viewed, the electron donating groups are held in place by the scaffold provided by the ligand and easily accommodate the metal ion. If one of the bonds would break the metal would still be held in place by the remaining bonds and the broken bond could reform. For the metal ion to dissociate, all bonds provided by the ligands would need to break simultaneously. If the bonds were formed between a metal ion and several monodentate ligands, these ligands could dissociate independently. An extension of the chelate effect is the *macrocyclic effect*, stating that a macrocycle will form more stable complexes with a metal ion as compared to the corresponding linear ligand of the same denticity[26]. The electron donor groups in a macrocycle are more rigid in space as compared to a non macrocyclic ligand.

Metal to ligand affinities are commonly expressed as the equilibrium constants for the formation of the metal-ligand complex in solution. For a metal ion M^+ forming a complex with a monodentate ligand L, the formation constants constant $K_{M:L}$ (M^{-1}) for the equilibrium is shown in Equation 1.2. If the metal ion forms a complex with n ligands the stepwise stability constants are defined as in Equations 1.2 – 1.5. The overall stability constant β can then be defined as in Equation 1.6. Hence, log β is used to in this thesis describe the formation of metal ion chelates with polydentate ligands. For clarity the complex to which the stability constant refers to is indicated as a suffix. The unit of log β is omitted for simplicity.

Equation 1.2	$[M] + [L] \rightleftharpoons [ML]$	$K_{M:L} = [ML] / [M]\,[L]$
Equation 1.3	$[ML] + [L] \rightleftharpoons [ML_2]$	$K_{M:L2} = [ML_2] / [ML]\,[L]$
Equation 1.4	$[ML_2] + [L] \rightleftharpoons [ML_3]$	$K_{M:L3} = [ML_3] / [ML_2]\,[L]$
Equation 1.5	$[ML_{(n-1)}] + [L] \rightleftharpoons [ML_n]$	$K_{M:Ln} = [ML_n] / [ML_{(n-1)}]\,[L]$
Equation 1.6	$[M] + n\,[L] \rightleftharpoons [ML_n]$	$\beta_{M:Ln} = [ML_n] / [M]\,[L]^n$

The expression is widely used and a vast amount of published data is available. An excellent compilation of critically selected stability constants by Martell is available in book format[27] and also as a database[28]. On the contrary, the dissociation constant, K_d is used to describe protein – metal binding equilibria. For a protein P binding to a metal ion M, the dissociation constant K_d (M) for the equilibrium is shown in Equation 1.7.

Equation 1.7	$[PM] \rightleftharpoons [P] + [M]$	$K_d = [P]\,[M] / [PM]$

Therefore stability and dissociation constants can be converted using Equation 1.8.

Equation 1.8 $\beta = 1 / K_d$

For consistency the logarithm of the overall stability constant of a metal M and a ligand L ($\log \beta_{M:L}$) is used to describe affinities herein. Values of K_d from the literature are converted to β using Equation 1.8 and logarithmized. In cases when the ligand exist in several protonation states the fully deprotonated ligand is referred to as L. The ligand protonated with n protons is referred to as LHn.

The stability constant of a metal and its respective ligand(s) is based on the assumption that the reaction is at or near equilibrium. The rates k_1 and k_{-1} of the association and dissociation reactions occurring for a metal – ligand complex (Equation 1.9) are described by the term *lability*. Labile complexes have high association and dissociation rates whereas the reverse prevails with kinetically *inert* complexes.

Equation 1.9
$$[M] + [L] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [ML]$$

Stability and lability are related but not proportional. It is thus possible for a metal-ligand complex to have a high stability constant but still be labile[29]. In this sense the term 'stability constant' can be somewhat misleading since it only describes how thermodynamically stable, but not kinetically inert a complex is.

1.2.4.2 Some ligands used in IMAC

A vast number of chelating ligands have been developed for use in IMAC systems since the technique was first established. The structures of some IMAC ligands are shown in Figure 1.4. Porath *et al.* immobilized the tridentate ligand iminodiacetic acid (IDA) and the pentadentate ligand tris(carboxymethyl)ethylenediamine (TED) onto Sepharose™. Their properties as IMAC ligands were investigated using Ni^{2+} and Fe^{3+} for fractionation of proteins[30]. To avoid the problem of metal ion leakage from IDA, Hochuli *et al.* introduced the tetradentate ligand nitrilotriacetic acid (NTA). NTA was shown to chelate Ni^{2+} and Cu^{2+} more strongly than IDA due to the additional coordination site provided by the ligand[31]. Another commonly used ligand employed in IMAC is carboxymethylated aspartic acid (CM-Asp). The ligand was introduced by Chaga *et al.*[32] and is currently sold under the tradename TALON® for use with immobilized Co^{2+}.

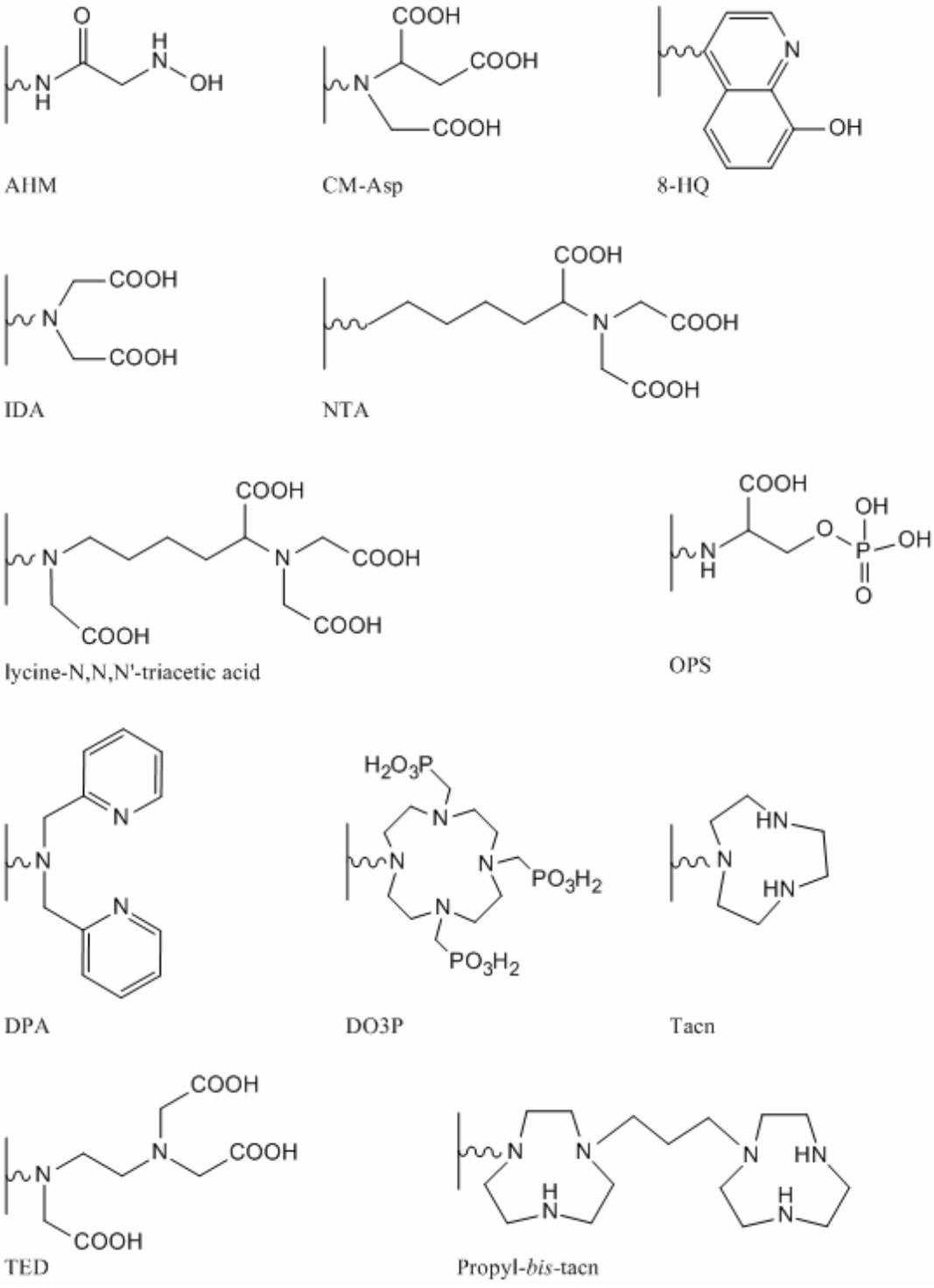


Figure 1.4. Chemical structures of some ligands employed in IMAC attached to a solid support (such as Sepharose) via a spacer arm (curved line).

Other ligands investigated in IMAC include *N*-(2-pyridylmethyl)aminoacetate[33] used to chelate Cu^{2+}, *O*-phosphoserine (OPS)[34] and 8-hydroxyquinoline (8-HQ)[35; 36] for use with the hard metal ions Al^{3+}, Fe^{3+}, Yb^{3+} and Ca^{2+}. OPS and IDA formed chelates with all metals investigated. The loading of Al^{3+} was higher for OPS then for IDA[34].

Metal ion leakage has been reported with commonly used IMAC ligands such as IDA and NTA[37; 38; 39]. For Ni^{2+}-NTA-Agarose, displacement of 2 mol Ni^{2+} per mol of His_6 tagged protein was reported[37]. As can be seen in Table 1.3 the stability constants of the Ni^{2+}-IDA and Ni^{2+}-His complexes are in the same range.

It is therefore not surprising that this metal is stripped from the ligand by His containing protein surfaces or affinity tags. This can be problematic due to the toxicity of several borderline metal ions (Section 1.2.3.2). Attention has therefore been drawn to macrocyclic molecules such as 1,4,7-triazacyclononane (tacn) and its derivatives as IMAC ligands. Several tacn derivatives have been synthesised and employed as IMAC ligands at the Monash University Centre for Green Chemistry. *Bis*-tacn compounds in which two tacn are linked together by a spacer, for instance a butyl chain, has been of particular interest. Tacn and *bis*-tacn ligands have very high stability constants[40; 41; 42] with many borderline metals, in the range of $16 < \log \beta < 30$, due to the macrocyclic effect previously discussed. The difference in stability constant between classic IMAC ligands and tacn is clearly observed in Table 1.3.

Table 1.3. Critical stability constants ($\log \beta_{M:L}$) of a selection of ligand-metal complexes.

Ligand \ Metal	Ca^{2+}	Mg^{2+}	La^{3+}	Fe^{3+}	Zn^{2+}	Ni^{2+}	Cu^{2+}
NTA	6.3	5.5	10.47	16.00	10.45	11.51	12.7
IDA	2.6	2.98	5.88	10.72 [(a)]	7.15	8.3	10.56
tacn (ref[44])	n.a.	n.a.	n.a.	n.a.	11.62	16.24	15.52
His (H)	1.21	n.a	n.a.	4.7 [(b)]	6.51	8.66	10.16
Aspartic acid (D)	1.7	2.4	4.84 [(c)]	11.4 [(d)]	5.87	7.16	8.89
Glutamic acid (E)	1.43	1.9	n.a.	12.1 [(d)]	4.49 [(a)]	5.61	8.32
Lysine (K)	1.05	n.a	n.a.	n.a.	4.11	4.9	7.65
Glycine (G)	1.09	1.66	3.1	8.57	4.96	5.74	8.19

Values determined at 25˚C, ionic strength $\mu = 0.1$ M if not otherwise stated. Data from NIST stability database[28] if not otherwise stated. [(a)] T = 25 ˚C, $\mu = 0.5$ M, [(b)] T = 20 C, $\mu = 0.1$ M, [(c)] T= 30 ˚C, $\mu = 0.1$, [(d)] T = 20 ˚C, $\mu = 1.0$.

Several investigators have reported using tacn and *bis*-tacn as IMAC ligands. Jiang *et al*. studied the chromatographic behavior of hen egg white lysozyme (HEWL), horse heart cytochrome *c*, horse skeletal muscle myoglobin (HMYO) and human serum proteins on tacn-Sepharose^{TM} CL-6B. The metals used were Cu^{2+}, Ni^{2+}, Cr^{3+}, Co^{2+}, Mn^{2+} and Zn^{2+}. Cu^{2+} was found to have the highest binding affinity to HMYO in batch equilibrium binding studies and the highest binding capacity to the human serum proteins in chromatography studies[43].

The binding behavior of short peptides containing one or two His residues with Cu^{2+} complexes of tacn and its derivatives was also studied[45]. Tacn and the *bis*-tacn compounds of 1,4-bis(1,4,7-triazacyclonon-1-yl)butane, 1,2-bis(1,4,7-triazacyclonon-1-ylmethyl)benzene, 1,3-bis(1,4,7-triazacyclonon-1-ylmethyl)benzene and 1,4-bis(1,4,7-triazacyclonon-1-ylmethyl)benzene were immobilized to Sepharose^{TM} CL-4B.

The aim was to develop a two point binding where two His residues in the peptide would bind to the two Cu^{2+} ions immobilized onto the *bis*-tacn. Peptides containing two His residues were found to bind more strongly to both Cu^{2+}-tacn and Cu^{2+}-*bis*-tacn than peptides containing one His residue. Varying the positions of the His residues in the dihistidine peptides did not alter the binding significantly[45]. HEWL, horse heart cytochrome c and HMYO were however found to bind more strongly to xylenyl bridged Cu^{2+}-*bis*-tacn as compared to Cu^{2+}-tacn complexes[46] demonstrating that higher binding is achieved with these binuclear ligands. Further studies of the ligands tacn, *bis*(1,4,7-triazacyclonon-1-yl)propane (propyl-*bis*-tacn) and *bis*(1,4,7-triazacyclononyl)ethane immobilized onto Sepharose^{TM} CL-6B achieved high purity isolation of the fusion protein GST-dATPase-His_6 from crude cell lysates. For the metal ions investigated, Cu^{2+}, Ni^{2+} and Zn^{2+} were successful but Co^{2+}, Mn^{2+} and Cr^{3+} were not[47]. As can be seen in Table 1.3 trends can be observed for the stability constants between different metals and ligands. Divalent hard metal ions such as Ca^{2+} form relatively weak complexes with most ligands whereas trivalent hard metal ions such as La^{3+} and Fe^{3+} form more stable complexes. This can be explained by the charge difference between the ions. Trends can also be observed for borderline and hard metal ions. The borderline metal ions Zn^{2+}, Ni^{2+} and Cu^{2+} form more stable complexes with most ligands as compared to the hard metal ions Ca^{2+} and Mg^{2+}. The stability constants of complexes with La^{3+} are also lower as compared to the borderline metal ions. It therefore appears that the hardness of a metal ion will affect the strength of the bonds formed with a ligand. Several of the ligands described here are commercially available either as uncharged gels or precharged with metals. Depending on the intended application ligands have been immobilized onto different base matrices. A few examples are shown in Table 1.4. Gels can be obtained both loose and packed in a variety of column formats.

Table 1.4. Some common commercially available immobilized IMAC ligands.

Ligand	Base resin	Chelated metal ion	Brand name(s) and supplier in brackets
Proprietary	Agarose (Sepharose™)	Ni^{2+}	Ni Sepharose™ (GE Healthcare)
Proprietary	Agarose (Sepharose™)	Uncharged	IMAC Sepharose™ (GE Healthcare)
IDA	Polymethacrylate (crosslinked)	Ni^{2+}, Co^{2+} or uncharged	Fractogel® (Merck/Novagen®)
IDA	Polystyrene divinylbenzene	Na^{+}	Chelex® (Bio-Rad)
NTA	Agarose	Ni^{2+}	Ni-NTA His●Bind® (Merck)
NTA	Agarose	Ni^{2+}	Ni-NTA Agarose (Qiagen)
NTA	Agarose	Ni^{2+}	Ni-NTA Agarose (Invitrogen)
CM-Asp	Agarose (Sepharose™ and other crosslinking)	Co^{2+}	Talon® (Clontech)
PDC	Agarose (Sepharose™)	Ni^{2+}, Cu^{2+}, Co^{2+}, Zn^{2+}	PDC (Affiland)

IDA is supplied as an IMAC resin both uncharged and precharged with Co^{2+} and Ni^{2+} (Table 1.4). In addition, IDA is supplied as a chelating ion exchange resin. The Chelex® resin constitutes IDA immobilized to polystyrenedivinylbenzene. Metals chelated by the Chelex® resin will displace the Na^{+} originally chelated. NTA is available from several sources (such as Qiagen and Merck/Novagen®) as the immobilized metal free ligand or charged with a metal ion. Examples of other commercially available ligands include the pentadentate chelator (PDC) from Affiland (Belgium) and TALON® (immobilized Co^{2+}-CM-Asp).

1.2.5 Affinity tags

The concept of affinity tags is quite broad, ranging from small peptides to whole proteins. They can be used as affinity tags facilitating protein purification through a range of separation modes such as affinity chromatography, IEX, IMAC or HIC. Tags can also function as solubility tags promoting the solubility of the expressed fusion proteins[8; 48; 49]. Example of solubility tags include N-utilization substance A (NusA)[50], the SET tag[51] and Maltose Binding Protein (MBP) , the later functioning both as an affinity and a solubility tag[52]. Examples of affinity tags consisting of whole proteins are shown in Table 1.5. Some common peptide affinity tags for different modes of chromatography are shown in Table 1.6.

One of the first His containing tags for IMAC was the dipeptide tag HW. The dipeptide was fused to proinsulin resulting in enhanced retention of this construct on Ni^{2+}-

IDA as compared to untagged proinsulin[53]. The tag $(AHGHRP)_n$ where n = 4 or 8 was fused to β-galactosidase and IgG-binding protein. The proteins were purified on immobilized Zn^{2+}-IDA[54]. The hexa His (His_6) tag introduced by Hochuli *et al*.[9] for use with borderline metal IMAC systems is currently by far the most commonly used affinity tag for high throughput protein purification[48; 55]. Several other His containing affinity tags for borderline metal IMAC have also been developed. The HAT™ (Histidine Affinity Tag) is commercially available from Clontech. This is a naturally occurring amino acid sequence derived from chicken lactate dehydrogenase[56]. The $heli_{M14}$-tag was derived from the metal-binding site of *Helicobacter pylori* ATPase 439. It was found to increase the yield of EGFP by 20 % as compared to the His_6 tag under similar elution conditions[57]. A library of helix-loop-helix motif tags were designed to decrease the flexibility of the tag by introduction of a secondary structure motif[58]. Other tags employed in IMAC include VYIHPFHL[59] and HQHHH[60].

Table 1.5. Some common protein affinity tags and the solid support they are used in conjunction with.

Tag	Size (kDa)	Matrix	Chromatography mode	Reference
Glutathione S-Transferase (GST)	26	Glutathione (GSH) Sepharose™	Affinity	61
Maltose binding protein (MBP)	40	Maltose, Amylose	Affinity	52
β-Galactosidase	116	APTG or anti-galactosidase binding	Affinity	62
Staphylococcal protein A (SPA)	14-31	IA (on IgG)	Affinity	63; 64; 65
Calmodulin	16.7	Propyl, butyl or Phenyl	Hydrophobic	66

The NT tags[67] were developed for IMAC systems employing tacn and *bis*-tacn compounds as the chelating ligands. The NT1A tag[68] with the amino acid sequence HHHNSWDHDANR studied in this thesis belongs to the NT family of tags. The scope for further development of affinity tags for IMAC is huge, both for novel IMAC ligands with borderline metals and with hard metals not currently widely employed in IMAC.

Tandem purification strategies where two affinity tags of different kinds are attached to the target protein, either one at each side or in sequence at the N- or C-terminus of the protein, are a way of further enhancing the selectivity of the separation procedure. Examples of fusion tags include His_6-calmodulin for sequential IMAC and HIC[66], and His_6-MBP which is used at several large research centres[48]. Both of these fusion tags are large, containing a

whole protein in conjunction with a peptide tag and there is therefore scope for the development of smaller fusion tags.

Following purification it is usually desired to remove the affinity tag so that the properties of the target protein are not altered. This is of particular importance if the protein is intended for therapeutic use. A number of enzymes can facilitate the cleavage of the affinity tag from the target protein. These can belong to the group endopeptidases, cleaving specific sequences or exopeptidases cleaving two or three amino acids at a time.

Table 1.6. Overview of peptide tags employed in different modes of chromatography.

Tag	Sequence	Chromatography mode	Ligand	Reference
His_6	HHHHHH	IMAC	Ni^{2+}-NTA	[9]
$heli_{M14}$	HNHRYGCGCC	IMAC	Ni^{2+}-NTA	[57]
NT1A	HHHNSWDHDANR	IMAC	Cu^{2+}-tacn (or tacn derivatives)	[68]
FLAG-tag	DYKDDDDK	Affinity	Anti-flag monoclonal antibody	[70]
Strep-tag II	WSHPQFEK	Affinity	Strep-tactin	[71; 72]
Myc	EQKLISEEDL	Affinity	Monoclonal antibody	[73]
S	KETAAAKFERQHMDS	Affinity	S-fragment of RNaseA	[74; 75]
HAT (natural histidine affinity tag)	KDHLIHNVHKEFHAH AHNK	IMAC	TALON®	[56]
Calmodulin binding peptide	KRRWKKNFIAVSAAN RFKKISSSGAL	Affinity	Calmodulin	[76; 77; 78; 79]
Poly Arg tags	R_n where n = 5-15 amino acids	IEX	Anionic resins	[80; 81]
Poly Asp tags	D_n where n = 5-16 amino acids	IEX	Anionic resins	[82]

Examples of endopeptidases include enterokinase, tobacco etch virus (TEV) protease, thrombin and factor X_a. Enterokinase cleaves specifically at the carboxyl site of lysin in the

D-D-D-D-K (D_4K) motif and is therefore often used for N-terminally tagged fusion proteins. The TEV protease recognizes the sequence E-X-X-Y-X-Q-S and cleaves between glutamine and serine[69], with the optimal cleavage site being E-N-L-Y-F-Q-S.

Another commonly used enzyme for tag cleavage is thrombin which has several recognition sites with the structure X_1-X_2-P-R-[cleavage site]-X_3-X_4. X_1 and X_2 are hydrophobic amino acids and X_3 and X_4 are non-acidic amino acids. Thrombin leaves two amino acid residues at the C-terminus of the cleavage site. Common recognition sites include L-V-P-R-G-S, L-V-P-R-G-F and M-Y-P-R-G-N. Factor X_a cleaves at the carboxyl site of the sequence I-Y_1-G-R-Y_2 where Y_1 is aspartic acid or glutamic acid and Y_2 is any amino acid except proline and arginine[8].

Examples of exopeptidases include aminopeptidase M (APM), carboxypeptidase A and B (CPA and CPB) and DAPase[11]. APM is extracted from porcine kidney and cleaves one amino acid sequentially at the N-terminus of a protein[11]. CPA and CPB are obtained from bovine pancreas and cleave amino acids at the C-terminus of a protein. DAPase is contained in the TAGZyme system (commercially available from e.g. Qiagen). It is a recombinant rat dipeptidyl aminopeptidase I which cleaves two amino acids from the N-terminus of a protein until a stop position is encountered. Stop signals include P at position 2 or 3 in the protein and R or K at position 1 and are contained in many proteins. In cases where a stop signal is absent a Q is added at the end of the affinity tag. Upon cleavage by DAPase in the presence of Qcyclase, a pyroglutamyl is formed at the N-terminus of the protein. The pyroglutamyl can be removed by pGAPase[11; 83]. As the recombinant DAPase contains a His tag it can be removed from the cleavage reaction by IMAC[84].

Another strategy for tag removal is chemical cleavage. This can be achieved using substances such as cyanogen bromide, formic acid or hydroxylamine. Drawbacks of these methods include unspecific cleavage and risk of denaturing the target protein. An interesting way to remove an affinity tag is the use of photocleavable tags, which are removed when the recombinant proteins are subjected to UV radiation[49].

1.3 IMAC using hard metals such as Ca^{2+}

As opposed to borderline metal IMAC systems which are well characterized and commonly used, IMAC using hard metal ions such as Ca^{2+}, Mg^{2+} and Fe^{3+} remains poorly exploited. Only a handful papers have been published to date describing protein purification by means of Ca^{2+}-based IMAC. In 1981 and 1984 Borrebaeck *et al.* reported using Ca^{2+} chelated to IDA to purify a seed lectin[85] and a carbohydrate binding protein from the fungus *Arthrobotrys oligospora*[86]. The proteins were loaded in 50 mM Tris-acetate, 0.5 M NaCl, pH 8.2 and eluted by supplementing the buffer with 10 mM EDTA. In 1989, Porath *et al.*

reported using Ca^{2+} chelated on CM-Asp-agarose to purify human serum amyloid P component (SAP) from human serum[87]. In 1991 the same research group reported purifying blood coagulation factor VIII:c from rat liver cell culture medium[88], fibrinogen and calmodulin[89] using Ca^{2+}-CM-Asp-agarose. In 1996 Chaga *et al*. studied the fractionation of several Ca^{2+} binding proteins using Mg^{2+}, La^{3+}, Nd^{3+} and Eu^{3+} chelated to TED. These metals were used instead of Ca^{2+} to avoid MIT caused by the target protein having higher Ca^{2+} affinity as compared to the immobilized ligand[90]. Considering that CM-Asp is currently commercially used as a Co^{2+} chelate (sold under the tradename TALON®) and the lack of recent publications describing the Ca^{2+}-IDA and Ca^{2+}-CM-Asp systems, such systems remain poorly exploited. We attempted to repeat the procedures described in these papers using Calmodulin in our laboratory but without success (unpublished data).

Al^{3+}-IDA and Al^{3+}-NTA have also been used in IMAC to purify Al^{3+} binding components in lysates of the fungus *Pisolithus tinctorius*. Elution was not achieved employing classic IMAC elution components such as EDTA or imidazole. Instead the pH was raised using ammonium hydroxide which displaced the Al^{3+} ions from the gel thereby eluting the bound material[91]. The above studies describe the purification of Ca^{2+}-binding proteins, thus exploiting the natural metal binding properties of the gel.

As mentioned above, Zachariou *et al*. used the ligand 8-HQ immobilized onto Sepharose™-Cl-4B with the hard metal ions Ca^{2+}, Fe^{3+}, Yb^{3+} and Al^{3+}. Cu^{2+} was included as a borderline metal ion control. Binding of tuna heart cytochrome *c* (THCC) was studied. This protein is rich in surface exposed aspartic and glutamic acid residues. The oxygen in these amino acid side chains can interact with immobilized hard metal ions. Binding was achieved with Fe^{3+}, Yb^{3+} and Al^{3+}. The binding with Cu^{2+} and Ca^{2+} were however of similar magnitude as for metal free gel[35]. These investigators also studied the ligand OPS immobilized onto Sepharose™ Cl-4B using the same metals. Ca^{2+} bound only weakly to this ligand and to IDA-Sepharose™ Cl-4B after the gel had been washed with acidic buffer. The binding of THCC, HMYO and HEWL was screened between pH 5.5 and 8.0 for all metals. Fe^{3+}-OPS-Sepharose™ Cl-4B bound THCC under all conditions and HMYO between pH 5.5-7.0. No significant binding was observed for the other hard metal ions[34] Ca^{2+}, Fe^{3+}, Yb^{3+} and Al^{3+}. Studies with different species varieties of cytochrome *c*, sheep skeletal muscle myoglobin and HEWL showed binding of THCC, horse heart cytochrome *c* and dog heart cytochrome *c* to Fe^{3+}-IDA-Sepharose™ CL-4B at acidic pH. No binding was however observed[92] for any of the proteins to Ca^{2+}-IDA-Sepharose™ CL-4B. Employing buffers of lower ionic strength, Zachariou and Hearn observed pseudo cation exchange behaviour with the above hard metal IMAC systems. This effect was attributed to interactions both with the metal ion and with the ligand[93]. Employing human serum proteins, the higher binding of proteins at lower ionic

strength remained[94]. At low ionic strength the binding behaviour of the Ca^{2+} chelates was very similar to those of the uncharged ligands.

Employing affinity tags for Ca^{2+} IMAC was investigated by Hearn *et al*.[95] These affinity tags were linear peptide sequences based on the so-called EF-hand of Ca^{2+} binding proteins (Chapter 6, Figure 6.1) and were used with macrocyclic ligands. As mentioned above, macrocyclic ligands represent attractive candidates as IMAC ligands due to the macrocyclic effect. Derivatives of 1,4,7,10-tetraazacyclododecane (cyclen) have been of particular interest due to their ability to form stable complexes with hard metal ions such as Ca^{2+}. Examples of such derivatives include 1,4,7,10-tetraazacyclododecane-1,4,7-tri(methanephosphonic acid) (DO3P), 1,4,7,10-tetraazacycloddoecane -1,7-di(methanephosphonic acid) (DO2P), 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid (DO3A) and 1,4,7,10-tetraazacyclododecane-1,7-diacetic acid (DO2A). The structure of DO3P is shown in Figure 1.4.

Protocols for purification of e.g. calmodulin are based on HIC where binding is achieved when hydrophobic parts of the protein become exposed upon Ca^{2+} binding. Elution is induced by chelating the Ca^{2+} using ethylene glycol tetraacetic acid (EGTA) or EDTA. It is therefore possible that some early reports on Ca^{2+} based IMAC do not reflect actual interaction between the protein and chelated Ca^{2+}, but rather secondary hydrophobic interactions between the protein and the chromatography resin. The field of Ca^{2+} based IMAC clearly needs to be further investigated.

1.4 Considerations when developing IMAC systems

A stability constant of log β = 4 - 8 between antibody and antigen is considered suitable in affinity chromatography[96]. A log β < 4 will be too weak to facilitate binding and a log β > 8 will make elution difficult[97].

Since the Ni-NTA His_6 system is well characterized, it can serve as a guide as to the required characteristics of the interactions between the metal ion and the ligand on the one hand and between the affinity tag and the metal chelate on the other hand. As can be seen in Table 1.3 the stability constant of the Ni^{2+}-NTA chelate is log β = 11.51. The affinity between the His_6 tag and free Ni^{2+} was determined[98] to log β = 11.79. As mentioned in Section 1.2.4 the similar magnitude of the affinity between the metal-ligand and metal-tag complexes can explain the occurrence of MIT in these systems. The affinity of a His_6 tag to Ni^{2+}-NTA modified surfaces has been estimated to be log β = 6 by Biacore experiments. The binding with one His_6 tag was found to be labile with high rates of association and dissociation[99]. The corresponding affinity between a His_6 tag and a single Ni^{2+}-NTA complex was estimated to be log β = 5 using a solution based assay employing fluorescent Ni^{2+}-NTA conjugates[100]. The

lower affinity of the tag to the metal chelate compared to the free metal can be attributed to less numbers of coordination bonds available to the tag when the metal is chelated by the ligand. The freedom of the metal ion will also be decreased when chelated as compared to when free in solution.

Based on these literature values it is reasonable to assume that an affinity of log $\beta \sim 5$ - 6 would be suitable for the interaction between the affinity tag(s) and the metal-ligand complex in a novel IMAC system. A suitable value for the metal chelate would be log $\beta \sim 10$.

The selectivity of the immobilized metal chelate to the tagged recombinant protein with respect to the other proteins present in the sample is another important parameter[101]. If the binding is not selective contaminant proteins may bind to the immobilized metal chelate. This can result in coelution of these contaminant proteins. The selectivity of IMAC systems using poly-His tags is based on the fact that the density of His residues in these tags is higher than on the surface of many proteins[6].

1.5 Aim of thesis

The overall aim of this thesis is to contribute to the understanding and development of more effective and environmentally sustainable protein separation technologies based on the principles of IMAC. The main disadvantage with the current IMAC systems from an environmental perspective is the leakage[37; 38; 39] of metal ions such as Ni^{2+}, Cu^{2+} and Co^{2+} due to their toxicity[22; 23], carcinogenity[24] and allergenicity[23]. This drawback can be overcome either by minimizing the leakage of the toxic metal ions or by employing non toxic metal ions. The thesis is thus divided into two parts addressing these two strategies.

Part 1: The stability constants of Ni^{2+} and Cu^{2+} chelates of tacn and its derivatives are higher[40; 41; 42] ($16 < \log \beta < 30$) compared to non macrocyclic ligands like IDA and NTA[28] ($8 < \log \beta < 13$). Characterization of borderline metal IMAC systems based on the tacn motif, developed at the Monash University Centre for Green Chemistry[67; 68], is described in the first part of this thesis. These systems were not yet fully characterized in regards to the arrangement of the affinity tags NT1A and HIT2 and their effect on the target protein.

Employing two affinity tags in sequence can facilitate tandem IMAC purification strategies. Hence recombinant fusion proteins containing the borderline metal binding tag NT1A and the putative hard metal binding tag HIT2 were created. The expression level of the recombinant fusion protein NT1A-HIT2-EGFP was lower than for HIT2-NT1A-EGFP in *E. coli* BL-21. The effect of different arrangements of the affinity tags fused to the C-terminus of EGFP is therefore studied in Chapter 3. The expression level and chromatographic properties on immobilized Ni^{2+} and Cu^{2+} chelates of tacn and propyl-*bis*-tacn is investigated.

All affinity tags can potentially affect their fusion partner[8; 11; 48]. Investigating if novel affinity tags affect their fusion partner is therefore important. The effect of the NT1A and HIT2 tags on the structure and stability of EGFP is therefore studied in Chapter 4.

Part 2: The second part of this thesis describes a contribution towards the development of a Ca^{2+} based IMAC system. In such a system metal ion leakage would not pose a problem due to the low toxicity of Ca^{2+}. Since the field of Ca^{2+} based IMAC is less mature compared to that of IMAC based on borderline metal ions, the emphasis is to understand the fundamentals of such a system. An IMAC system can be divided into two parts; the immobilized metal chelate and the affinity tag.

Many Ca^{2+}-chelates of free ligands are well characterized[28]. Their properties as immobilized Ca^{2+}-chelates have however received less attention. Immobilized Ca^{2+}-chelates are therefore investigated under authentic chromatographic conditions in Chapter 5.

The affinity of tags to free metal ions can serve as a guide to the affinity of the tag to immobilized metal chelates. The affinity of free Ca^{2+} to different novel putative cyclic Ca^{2+} affinity tags fused to EGFP is therefore studied in Chapter 6.

Chapter 2. Materials and Methods

2.1 Materials

2.1.1 Chemicals

Glacial acetic acid, hydrochloric acid, iron nitrate, magnesium chloride, magnesium sulphate, sodium iodide and sodium thiosulphate were from Ajax Finechem (Taren Point, NSW, Australia). Sodium chloride and tris(hydroxymethyl)aminomethane (Tris) were from Amresco (Solon, OH, USA). Acrylamide solution (4K, 30 %), ampicillin and 2-(N-morpholino)ethanesulfonic acid (MES) were from AppliChem (Cheshire, CT, USA). Dithiothreitol (DTT) was from Astral Scientific (Caringbah, NSW, Australia). Ammonium persulphate and N,N,N',N'-tetramethylene-diamine (TEMED) were from Bio-Rad (Hercules, CA, USA). Calcium chloride was from May & Baker (Dagenham, United Kingdom). Ammonium chloride, calcium carbonate, calcium nitrate, dipotassium hydrogen phosphate, ethylenediaminetetraacetic acid (EDTA), formaldehyde, glycine, ethanol, ethidium bromide, glycerol, methanol, potassium chloride, potassium dihydrogen phosphate, potassium sulphate, sodium dodecyl sulphate, silver nitrate, and sodium carbonate were from Merck (Kilsyth, Vic, Australia). 5,5'-dibromo-1,2-bis(o-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid (5,5'-Br_2-BAPTA) was from Molecular Probes (Eugene, OR, USA). Cesium chloride was from NovaChem (Collingwood VIC, Australia). Isopropyl-β-D-1-thiogalactoside (IPTG) was from Promega (Madison, WI, USA). Ammonium persulfate (APS), Bromophenol blue, N-Cyclohexyl-2-aminoethanesulfonic acid (CHES), coomassie brilliant blue R-250, ethylene glycol tetraacetic acid (EGTA), guanidine hydrochloride (GuHCl), reduced glutathione, glutaraldehyde, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), imidazole, iminodiacetic acid (IDA), nickel nitrate, potassium nitrate, tryptone and yeast extract were from Sigma-Aldrich (Castle Hill, NSW, Australia). Ammonium sulfate was from Scharlau (Sentmenat, Barcelona, Spain). 1,4,7,10-tetraazacyclododecane (cyclen) was from Stem Chemicals (Newburyport, MA, USA).

2.1.2 Enzymes for DNA and protein manipulation and antibodies

MassRuler™ DNA ladder mix and Page Ruler™ protein ladder mix were from Fermentas (Glen Murnie, MD, USA). Calf intestine phosphatase (CIP), DynoZyme 2 Taq polymerase, 10 x Taq polymerase reaction buffer and $MgCl_2$ (50 mM) were from Finnzymes (Espoo, Finland). Bovine serum albumin (BSA), endonucleases EcoRI and EcoRV, ClaI,

HindIII and T4 DNA ligase were from New England Biolabs (Ipswich, MA, USA). Nuclease free water and the pGEM®-T Easy vector system were purchased from Promega (Madison, WI, USA). Plasmid Preparation and DNA gel extraction kits were from Macherey-Nagel (Düren, Germany), Eppendorf (New South Wales, Australia) and Sigma-Aldrich (Castle Hill, NSW, Australia). Calmodulin (cat.no P2227) was from Sigma-Aldrich. Polyclonal Mouse anti-GFP antibody (cat.no 11814460001) was from Roche (Basel, Switzerland). Rabbit anti-Mouse antibody conjugated with Horse Reddish Peroxidase (HRP) was from Dako (Glostrup, Denmark). EnterokinaseMax™ (EKMax™) was from Invitrogen (Carlsbad, CA, USA).

2.1.3 Oligonucleotides for gene cloning

The following oligonucleotides (Table 2.1) were synthesized by Micromon (Victoria, Australia): HindIII-GFP-r coding the C-terminus of EGFP, Primer A coding His_6-D_4K and being complementary to the start of the *egfp* gene and Primers B-K encoding the cyclic tags B-K and being complementary to the His_6 sequence.

Table 2.1. Primers employed in cloning of cyclic tags B-K and His_6-D_4K-EGFP.

HindIII-GFP-r	5'-CAAGCTTCTATTTGTATAGTTC-3'
Primer A	5'-CTCATGAAACATCACCATCACCATCACGATGACGATGACAAAATGAG-3'
Primer B	5'-CACATGTGTGATAAAGATGGCGATGGCTATATTACCACCAAAGAACTGGGCACCTGTCATCACCATCACCATCACG-3'
Primer C	5'- CACATGTGTGATAAAGATGGCGATGGCACCATTACCACCAAAGAACTGGGCACCTGTCATCACCATCACCATCACG-3'
Primer D	5'- CACATGTGTGATATTGATGGCGATGGCCATATTAACTATGAAGAACTGGGCACCTGTCATCACCATCACCATCACG-3'
Primer E	5'- CACATGTGTGATCAGGATAAATCTGGCTTTATTGAAGAAGATGAACTGGGCACCTGTCATCACCATCACCATCACG 3'
Primer F	5'- CACATGTGTGCGGATTATAACAAAGATGGCTGGTATGAAGAACTGGAATGTCATCACCATCACCATCACG-3'
Primer G	5'- CACATGTGTGCGGATAAAGATGGCGATGGCACCATTACCACCAAAGAATGTCATCACCATCACCATCACG-3'
Primer H	5'- CACATGTGTGCGGATATTGATGGCGATGGCCATATTAACTATGAAGAATGTCATCACCATCACCATCACG-3'
Primer I	5'- CACATGTGTGCGGATCAGGATAAATCTGGCTTTATTGAAGAAGATGAATGTCATCACCATCACCATCACG-3'
Primer J	5'- CACATGTGTGATAAAGATGGCGATGGCACCATTACCACCAAAGAATGTCATCACCATCACCATCACG-3'
Primer K	5'- CACATGTGTGATCAGGATAAATCTGGCTTTATTGAAGAAGATGAATGTCATCACCATCACCATCACG-3'

2.1.4 Buffers and solutions

All water used was of reverse osmosis (RO) grade unless otherwise stated.

2.1.4.1 Media, general buffers and solutions employed

- Lysogeny Broth (LB-medium): 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl in water.

- 2 x YT-medium: 16 g L^{-1} tryptone, 10 g L^{-1} yeast extract, 5 g L^{-1} NaCl in water.
- Phosphate buffered saline (PBS): 150 mM NaCl, 10 mM sodium phosphate, pH 7.6.
- Tris-Acetate-EDTA (TAE) buffer: 40 mM Tris-HCl, 0.1142 % (v/v) glacial acetic acid, 1 mM EDTA, pH 8.5.
- Tris buffered saline (TBS): 25 mM Tris, 120 mM NaCl, 2.5 mM KCl, pH 7.5.

All buffers used on FPLC and other liquid chromatography systems were filtered using 0.2 or 0.45 µm filters fitted to a disposable syringe.

2.1.4.2 Buffers and solutions for Atomic Absorption Spectroscopy (AAS) and Inductively Coupled Plasma Atomic Emission Spectrometry (ICP)

Interference suppressor solution for atomic absorption spectroscopy (AAS) was prepared by dissolving lanthanum chloride and cesium chloride in deionized water to a final concentration of 6.67 g L^{-1} lanthanum and 6.67 g L^{-1} cesium, and was provided ready-made by the Monash University Water Studies Centre (Clayton, VIC, Australia). Calcium standard stock solution and Cesium Chloride were from BDH Laboratory Supplies (VWR International, West Chester, PA, USA). Lanthanum Chloride was from Ajax Finechem (Greenmount, Auckland, New Zealand). Standard stock solutions for ICP were from Accu Standard (New Haven, CT, USA).

2.1.4.3 Buffers for SDS-PAGE

- Gel tank running buffer: 192 mM Glycine, 0.1 % SDS, 25 mM Tris-HCl, pH 8.3.
- 5 x Sample loading buffer: 0.1 g mL^{-1} SDS, 0.4 mg mL^{-1} Bromophenol Blue, 50 % (v/v) glycerol, 300 mM Tris-HCl, pH 6.8.
- 5 x Reducing loading buffer: 0.1 g mL^{-1} SDS, 0.4 mg mL^{-1} Bromophenol Blue, 50 % (v/v) glycerol, 20 % (v/v) β-mercaptoethanol, 300 mM Tris-HCl, pH 6.8.
- Coomassie Brilliant Blue staining solution: Coomassie Brilliant Blue R250 (250 mg mL^{-1}), 10 % (v/v) glacial acetic acid, 40 % (v/v) methanol.
- Coomassie Brilliant Blue Destaining Solution: 10 % (v/v) glacial acetic acid, 30 % (v/v) methanol.
- Fix Solution: 40 % Ethanol, 10 % acetic acid.
- Fix/sensitize solution: 400 µL glutaraldehyde, 108 µL formaldehyde, 400 mL 40 % ethanol.
- Sodium thiosulphate solution: 0.02 % w/v sodium thiosulphate.
- Silver nitrate solution: 0.1 % w/v silver nitrate.

- Development solution: 20 g sodiumcarbonate, 320 µL formaldehyde in 800 mL water.

2.1.4.4 Buffers used for western blotting

- Transfer buffer: 10 % methanol, 0.25 M Tris, 150 mM glycine in water.
- TBS-T: 10 mM Tris, 100 mM NaCl, 0.1 % Tween-20.
- Chemiluminescence solution was from Pierce (Rockford, Il, USA).

2.1.5 Accessories and material

- Econo-columns, TrUView cuvettes and DNA gel equipment were from BioRad (Hercules, CA, USA).
- Dark 96 well plates were from Perkin Elmer (Waltham, MA, USA).
- ZipTips were from Millipore (Billerica, MA, USA).
- BCA assay kits containing all reagents and solutions were from Pierce.
- 0.2 and 0.45 µm syringe filters and Viva Spin Protein concentrators were from Satorius (Goettingen, Germany).
- Dialysis tubing was from Spectrum Laboratories Inc (Rancho Dominguez, CA, USA).
- Sepharose 6 F F, Phenyl-Sepharose 6 F F (Low Sub), Q-Sepharose 6 F F and Glutathione Sepharose™ 4B resins, HisTrap™ F F, HiPrep™ 26/10 desalting, MonoQ™ 5/50, Superdex™ 75 10/300 and Superdex™ 75 16/60 prepacked columns, 1 mL Tricorn™ 5/50 glass columns, 5 mL Mono P™ and 50 mL XK 45 glass columns were from GE Healthcare Bio-Sciences (Uppsala, Sweden).
- The IMAC ligands (4-amino-1-hydroxy-1-phosphono-butyl)phosphonic acid (alendronate), cyclen, IDA, 1,4,7,10-tetraazacyclododecane-1,7-diacetic acid (DO2A), 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid (DO3A), 1,4,7,10-tetraazacycloddoecane-1,7-di(methanephosphonic acid) (DO2P), 1,4,7,10-tetraazacyclododecane-1,4,7-tri(methanephopshonic acid) (DO3P), 1,4,7-triazacyclononane (tacn), *bis*(1,4,7-triazacyclonon-1-yl)propane (propyl-*bis*-tacn) immobilized onto Sepharose 6 F F were provided by Ms Eva Campi. A brief description of the synthesis and immobilization of the ligands can be found in Section 2.18.
- Ni-NTA His●Bind® resin was from Merck.
- Glass plates, gel tanks and mounting racks for gel electrophoresis and Nitrocellulose membranes for western blotting were from Amersham (Buckinghamshire, United Kingdom).

- A custom ordered 500 µL CD matched fluorescence cell (product number 28-F/Q/10/CD) was obtained from Starna (Baulkham Hills BC, NSW, Australia).
- An ISE25Ca Ca^{2+} selective electrode (part number E91M001, E34M001) and a Ref201 reference electrode (part number E21M009) were from Radiometer Analytical (Lyon, France).

2.1.6 Machines and apparatus

- A PHM201 pH/mV reader was from Radiometer Analytical.
- A Voyager DE-STR Matrix-Assisted Laser Desorption/Ionization Time of Flight (MALDI-ToF) mass spectrometer was from Applied Biosystems (Foster City, CA, USA).
- An Agilent CE capillary electrophoresis system coupled to an Agilent 1100 series LC/MSD-SL ion trap mass spectrometer through an Agilent G1607A orthogonal electrospray interface, an Agilent 1100 series isocratic LC pump delivering shesh liquid and the system control and data acquisitions softwares ChemStation and MSD Trap Control were from Agilent Technologies (Waldbronn, Germany).
- A SmartSpec™ Spectrophotometer, a Model 3550 plate reader, PowerPac Basic™ power aggregates for running protein and DNA gels, A BioRad BioLogic DuoFlow system fitted with a BioFrac fraction collector using the BioLogic DuoFlow V 5.00 Build 12 software and a MyCycler™ thermal cycler for Polymerase Chain Reaction (PCR) were from BioRad.
- An ÄKTAexplorer™, an ÄKTApurifier™ and an ÄKTAxpress™ system were from GE Healthcare.
- An 1100 Atomic Absorption Spectometer was from Perkin Elmer (Waltham, MA, USA).
- An Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES) Liberty 200 ICP Emission Spectometer was from Varian (Palo Alto, CA, USA).
- A Jasco J-815 CD spectrometer fitted with an FMO-427S/L fluorescence monochroamator, controlled by Spectra Manager™ II was from Jasco (Hachioji, Tokyo, Japan).
- An Infinite® 200 plate reader controlled by the i-control™ software was from Tecan (Männedorf, Switzerland).
- A Misonix Sonicator 3000 was from Misonix (Farmingdale, NY, USA).

2.2 Gene Cloning

2.2.1 Liquid culture growth

Escherichia coli (*E. coli*) liquid cultures were grown in LB-media containing 100 µg mL^{-1} ampicillin at 37 °C for 16 h on a rotating shaker at a speed of 220 rpm.

2.2.2 Plasmid DNA preparation

Plasmid DNA was prepared from liquid culture using the plasmid preparation mini or maxi kits following the manufacturer's instructions.

Briefly, bacteria cells were centrifuged and the pellet reconstituted and plasmid DNA liberated from the cells by SDS/Alkaline lysis. The solution was then neutralized and centrifuged, which pelleted the SDS and cell debris while plasmid DNA remained in solution. Plasmid DNA was purified by loading the supernatant onto a spin column and washing off impurities such as salt, metabolites, soluble macromolecular cellular components with an ethanolic buffer. The plasmid DNA was eluted with nuclease free water when using the mini kits and the provided elution buffer when using the maxi kits.

2.2.3 Transformation

One µL plasmid DNA was added to 50 µL chemically competent *E. coli* DH5-α or BL-21 bacteria cells. The mixture was incubated on ice for 1 h. The cells were heat shocked at 42 °C for 90 s and 1 mL LB-medium was added. The culture was incubated at 37 °C for 1 h and centrifuged at 4000 x g for 30 s. The supernatant was removed and the cells were spread onto an LB plate containing ampicillin (100 µg mL^{-1}). The plate was incubated at 37 °C for 16 h.

2.2.4 Separation of DNA on agarose gels

Agarose gels (1-3 %) were prepared by dissolving agarose in TAE buffer and adding ethidium bromide to a final concentration of 0.5 µg mL^{-1}. Gels were run at 100 V for typically 20 min – 1 h depending on the size of the fragment to be visualized.

Small DNA fragments can sometimes migrate with the same rate as the colour component of the loading dye. Therefore glycerol (5 % v/v final concentration) was used instead of loading dye when small DNA fragments were visualized or isolated. Gels were photographed on a Chemidoc XRS system (BioRad, Hercules, CA, USA).

2.2.5 Extraction of DNA from agarose gels

Bands containing the DNA fragments of interest were identified under UV-light. The gel slices containing the bands were cut out of the agarose gel using a scalpel. The DNA was purified from the gel slices following the manufacturer's instructions of the DNA gel extraction kit. Briefly the gel slices were incubated in lysis buffer at 50 °C until the gel was dissolved (5-10 min). The solution was loaded onto a silica column binding the DNA. The column was washed with an ethanolic solution to remove impurities such as salts and soluble macromolecules. The DNA was eluted in 30 µL nuclease free water.

2.2.6 Plasmid digestion

Plasmid DNA was digested by setting up reactions containing > 3 units of restriction enzyme per µg of plasmid DNA to be digested in the appropriate reaction buffers. The reaction was then incubated at 37 °C for 2 h.

2.2.7 Dephosphorylation

Linear DNA molecules were dephosphorylated by adding 10-20 units of CIP to the reaction mixture and incubating at 37 °C for 1 h.

2.2.8 Ligation

DNA molecules were ligated in the appropriate proportions between the vector and the insert in a 10 µL reaction containing 1 µL DNA ligase and 1 µL 10 x DNA liagase buffer and incubated at 4 °C overnight.

2.2.9 Polymerase Chain Reaction (PCR)

The following reactions were set up: 400 nM of each primer, 1 µL template plasmid DNA, 200 µM deoxyribunucleotides (dNTP), 2 units Taq Polymerase, 1 x reaction buffer and 1.5 µM $MgCl_2$ in a total volume of 50 µL.

The following program was programmed on a MyCycler™ thermal cycler.

- 94 °C for 5 min.
- 94 °C for 30 sec., 55 °C for 30 sec., 72 °C for 90 sec. for 30 cycles.
- 72 °C for 5 min.

2.3 Protein expression and cell lysis

2.3.1 Protein expression

Individual bacterial colonies were inoculated into 2-50 mL LB medium containing 100 µg mL^{-1} ampicillin and incubated at 37 ˚C for 16 h at 220 rpm. These starter cultures were expanded in LB-medium containing 100 µg mL^{-1} ampicillin. The volumetric ratio of starter culture : LB-medium was 1 : 10. The expanded cultures were incubated for a further 1 h at 37˚C at 220 rpm. Protein expression was induced by adding IPTG to a final concentration of 1 mM and then incubating at 37 ˚C for 5 h at 190 or 220 rpm. The cells were harvested by centrifuging 10 mL, 50 mL, 250 mL or 500 mL aliquots at 3220 x g for 10 min. The supernatant was removed, the pellet resuspended in 10 mL, 20 mL or 40 mL water and centrifuged at 3220 x g for 10 min. The supernatant was discarded and the pellets were stored at -20 ˚C.

2.3.2 Cell lysis

The cell pellets were resuspended in lysis buffer. Cells corresponding to 50 mL, 250 mL, 500 mL and 2.5 L liquid bacteria culture were lysed in 2 mL, 4 mL, 8 mL and 70 mL buffer respectively. In cases when the cell lysates were to be used as starting materials for chromatographic experiments, the loading buffer in the chromatography experiment was used as lysis buffer.

The cells were lysed on a Misonix Sonicator using the following program: 12 x (30 s burst at output level 6, 2 min pause). Lysis of 2 mL resuspenced cells was performed using a micro probe in 15 mL falcon tubes. Larger volumes were lysed using a regular probe in 50 mL falcon tubes. When protein yield was of priority, the lysis cycle was repeated for improved yield. The cell lysate was centrifuged at 16060 x g for 20 min. The supernatant was filtered through a 0.2 µm syringe filter and used for protein purification.

2.4 Chromatographic procedures

2.4.1 Methodological considerations

Chromatographic procedures were performed for two purposes. One purpose was to characterize different immobilized metal chelates and novel affinity tags. In these FPLC IMAC experiments, lysates of cell pellets corresponding to 50 mL liquid cell culture was used in each experiment. These experiments are referred to as analytical IMAC experiments.

The other purpose was to prepare pure protein for downstream experiments. Larger quantities of cells were used, pellets commonly corresponding to 250 mL or 500 mL liquid cell culture. These experiments are referred to as preparative purification experiments.

2.4.2 Bench column purification of Glutathione-S-transferase

Cell pellets were lysed in 10 mL 1 x PBS, pH 8.0. Glutathione Sepharose™ 4B gel slurry in 20 % ethanol (1 mL) was packed in a 3 mL Econocolumn. The column was equilibrated with 10 mL 1 x PBS, pH 8.0. The cell lysate was applied and the column washed with 3 x 10 mL 1 x PBS, pH 8.0. The protein was eluted in 3 x 1 mL 10 mM reduced glutathione in 50 mM Tris-HCl at pH 8.0.

2.4.3 Bench column purification of Calmodulin using hydrophobic interaction chromatography

Hydrophobic interaction chromatography (HIC) was employed according to the protocol for purification of calmodulin from *Saccharomyces cerevisiae* (*S. cerevisiae*) by Ohya *et al*[102]. Cell pellets were lysed in 2 mL 50 mM Tris-HCl, pH 7.5. Phenyl-Sepharose 6 F F (Low Sub) gel slurry in 20 % ethanol (1 mL) was packed in a 3 mL Econocolumn. The gel was washed with 3 x 1 mL water and equilibrated with 3 x 1 mL 1 M $(NH_4)_2SO_4$, 5 mM $CaCl_2$, 50 mM Tris-HCl, pH 7.5. The cell lysate was applied and the column washed with 4 x 1 mL 100 mM $(NH_4)_2SO_4$, 1 mM $CaCl_2$, 50 mM Tris-HCl, pH 7.5. Calmodulin was eluted in 4 x 1 mL 100 mM $(NH_4)_2SO_4$, 5 mM EGTA, 50 mM Tris-HCl, pH 7.5. Starting material, wash fractions and elution fractions were analysed on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

2.4.4 FPLC Column Packing and metal loading of IMAC resins

The gel was packed in a Tricorn™ 5/50 column. The metal ion was loaded onto the column by passing 20 mL of aqueous metal salt solution through the column and then incubating the column for 1 h at RT in the metal salt solution. Non-bound metal ions were removed from the gel by washing with 20 mL water. The following metal salt solutions were used: 50 mM $Cu(NO_3)_2$, 50 mM $Ni(NO_3)_2$ and 50 mM $Ca(NO_3)_2$.

2.4.5 General FPLC procedures and system maintainance

Prior to each experiment the FPLC system was flushed first with 10 mL water, then with 10 mL of the stronger mobile phase (eluent B) followed by 10 mL of the weaker mobile phase, commonly the loading and washing buffer (eluent A). The sample loop was flushed with > 2 times its volume of eluent A and the injection valve was backflushed with 5 mL eluent A. The column was installed and equilibrated with 10 mL eluent A. After use the

sample loop was flushed with > 2 times its volume of 20 % ethanol, the injection valve was backflushed with 5 mL 20 % ethanol and the system was flushed with 10 mL 20 % ethanol through line A and 10 mL through line B.

2.4.6 Analytical IMAC experiments of NT1A-EGFP fusion proteins

Analytical FPLC IMAC experiments were performed on a BioRad BioLogic DuoFlow system fitted with a BioFrac fraction collector using the BioLogic DuoFlow V 5.00 Build 12 software. Ni^{2+} and Cu^{2+} chelates of tacn and propyl-*bis*-tacn immobilized onto Sepharose 6 F F were packed in 1 mL Tricorn™ 5/50 glass columns.

The mobile phases employed were:

Eluent A: 10 mM sodium phosphate buffer containing 650 mM NaCl, 10 mM imidazole, pH 7.6.

Eluent B: 10 mM sodium phosphate buffer containing 650 mM NaCl, 300 mM imidazole, pH 7.6.

Samples were introduced in a 2 mL injection loop. The following program was employed at the flowrate 1 mL min^{-1}.

1-2 min: Sample injection in 2 mL 100 % eluent A.

3-9 min: Wash using 100 % eluent A.

10-19 min: Elution employing a linear gradient from 100 % eluent A to 100 % eluent B.

20-24 min: Second wash step using 100 % eluent B.

Fractions (1 mL) were collected. *E. coli* BL-21 cell lysates, one or several flow-through fractions and one or several elution fractions were analysed by SDS-PAGE as described below.

2.4.7 Ca^{2+}-based chromatography experiments

2.4.7.1 Analytical FPLC experiments.

Analytical FPLC experiments were performed on a BioRad BioLogic DuoFlow system fitted with a BioFrac fraction collector using the BioLogic DuoFlow V 5.00 Build 12 software. Ca^{2+} chelates of DO3P immobilized onto Sepharose 6 F F were packed in 1 mL Tricorn™ 5/50 glass columns.

The mobile phases employed were:

Eluent A: 10 mM HEPES, pH 7.5.

Eluent B: 10 mM HEPES, 300 mM NaCl, pH 7.5.

Samples were introduced in a 2 mL injection loop. The following program was employed at the flowrate 1 mL min^{-1}.

1-2 min: Sample injection in 2 mL 100 % eluent A.

3-9 min: Wash using 100 % eluent A.

10-19 min: Elution employing a linear gradient from 100 % eluent A to 100 % eluent B.

20-24 min: Second wash step using 100 % eluent B.

Fractions (1 mL) were collected. *E. coli* BL-21 cell lysates, one or several flow-through fractions and one or several elution fractions were analysed by SDS-PAGE as described below.

2.4.7.2 Analytical chromatography experiments

Analytical chromatography experiments were performed on an ÄKTAexplorer™ system. Ca^{2+}-DO3P-Sepharose 6 F F, Ca^{2+}-DO3A-Sepharose 6 F F, Ca^{2+}-Alendronate-Sepharose 6 F F and Ca^{2+}-IDA-Sepharose 6 F F were packed in 1 mL Tricorn™ 5/50 glass columns.

Isocratic chromatography experiments were performed at a flowrate of 1 mL min^{-1}. The mobile phases employed were as follows.

10 mM HEPES, pH 7.5 and 10 mM HEPES, 50 mM $CaCl_2$, pH 7.5.

The EGFP specific absorption at 488 nm was measured. Fractions (1 mL) were collected.

2.4.8 Preparative purification of EGFP-His_6

Preparative purification of EGFP-His_6 was carried out on an ÄKTAxpress™ system. A three-step purification method was employed: IMAC on a HisTrap™ FF column, then desalting on a HiPrep™ 26/10 desalting column followed by an ion exchange chromatography (IEX) polishing step on a prepacked MonoQ™ 5/50 column.

The mobile phases employed were:

Eluent A: 10 mM sodium phosphate buffer containing 500 mM NaCl, 30 mM imidazole, pH 7.6.

Eluent B: 10 mM sodium phosphate buffer containing 500 mM NaCl, 500 mM imidazole, pH 7.6.

Eluent C: 20 mM Tris-HCl, pH 8.0.

Eluent D: 20 mM Tris-HCl, 1 M NaCl, pH 8.0.

Step 1. IMAC on a 1 mL HisTrap™ F F column

The flowrate was set at 1 mL min^{-1} throughout. The program contained the following segments.

Segment 1: Injection of sample, typically 20 column volumes (CV) filtered *E. coli* cell lysates.

Segment 2: 100 % eluent A for 10 CV was applied, the column was washed with 10 mL eluent A and the following gradient was programmed.

Segment 3: 0 to 5 % eluent B in eluent A 1 column volume (CV).

Segment 4: 5 % eluent B in eluent A for 5 CV.

Segment 5: 5 % to 100 % eluent B in eluent A 0.5 CV.

Segment 6: 100 % eluent B for 10 CV.

The EGFP fusion protein was collected in one 7.5 mL internal system loop. Collection parameters were set as follows. Start collection: Watch level greater than = 75 mAU, watch slope greater than = 200 mAU min^{-1}. Stop collection: Peak max factor = 0.5, watch level less than 50 mAU, stable time = 0.5 min, delta plateau = 5 mAU.

Step 2. Desalting on a HiPrep™ 26/10 desalting column

The flowrate was set at 10 mL min^{-1}. The EGFP fusion protein contained in a 7.5 mL loop in the ÄKTAxpress™ (automatically collected from step 1) was applied and eluted in 1.3 column volumes of eluent C. The protein was collected in several 7.5 mL loops and used as starting material in step 3.

Step 3. IEX on a MonoQ™ 5/50 column

The flowrate was set at 1 mL min^{-1}. The program contained the following segments.

Segment 1: A linear gradient of 0-25 % eluent D in eluent C in 5 CV

Segment 2: A linear gradient of 25-50 % eluent D in eluent C 15 CV.

Segment 3: A linear gradient of 50-100 % eluent D in eluent C in 5 CV.

Segment 4: 100 % eluent D in 10 CV.

Peak fractionation parameters were set as follows: Fixed volume = 1 mL, Peak fractionation size = 0.2 mL, Start level = 50 mAU, Start slope = 20 mAU min^{-1}, Peak max factor = 0.5, minimum peak width = 0.5 min, end level = 30 mAU, end slope = 50 mAU min^{-1}.

Fractions were analysed by SDS-PAGE and by OD_{488}/OD_{280}.

2.4.9 Preparative purification of NT1A-EGFP and EGFP-NT1A

Preparative purification of EGFP-NT1A and NT1A-EGFP was carried out on an ÄKTAxpress™ system by a three-step purification method - IMAC on a prepacked HisTrap™ FF column, then by desalting on a HiPrep™ 26/10 desalting column followed by an IEX polishing step on a prepacked MonoQ™ 5/50 column.

The mobile phases employed were:

Eluent A: 10 mM sodium phosphate buffer containing 500 mM NaCl, 10 mM imidazole, pH 7.6.

Eluent B: 10 mM sodium phosphate buffer containing 500 mM NaCl, 500 mM imidazole, pH 7.6.

Eluent C: 20 mM Tris-HCl, pH 8.0.

Eluent D: 20 mM Tris-HCl, 1 M NaCl, pH 8.0.

Step 1. IMAC on 1 mL HisTrap™ FF

The flowrate was set at 1 mL min^{-1} throughout. The program contained the following segments.

Segment 1: Injection of sample, typically 20 column volumes (CV) filtered *E. coli* cell lysates.

Segment 2: 100 % eluent A for 10 CV was applied, the column was washed with 10 mL eluent A and the following gradient was programmed.

Segment 3: 0 to 5 % eluent B in eluent A 1 column volume (CV).

Segment 4: 5 % eluent B in eluent A for 5 CV.

Segment 5: 5 % to 100 % eluent B in eluent A 0.5 CV.

Segment 6: 100 % eluent B for 10 CV.

The EGFP fusion protein was collected in a 7.5 mL internal system loop. Collection parameters were set as follows. Start collection: Watch level greater than 75 mAU, watch slope greater than 200 mAU min^{-1}. Stop collection: Peak max factor = 0.5, watch level less than 50 mAU, stable time = 0.5 min, delta plateau = 5 mAU.

Step 2. Desalting on a HiPrep™ 26/10 desalting column

The flowrate was set at 10 mL min^{-1}. The EGFP fusion protein contained in a 7.5 mL loop from step 1 was applied and eluted in 1.3 CV eluent C. The EGFP fusion protein was collected in several 7.5 mL loops.

Step 3. IEX on a MonoQ™ 5/50 column

The flowrate was set at 1 mL min^{-1}. The program contained the following segments.

Segment 1: 0-25 % eluent D in eluent C in 5 CV.

Segment 2: 25-50 % eluent D in eluent C in 15 CV.

Segment 3: 50-100 % eluent D in eluent C 5 CV.

Segment 4: 100 % eluent D in 10 CV.

Peak fractionation parameters were set as follows: Fixed volume = 1 mL, Peak fractionation size = 0.2 mL, Start level = 50 mAU, Start slope = 20 mAU min^{-1}, Peak max factor = 0.5, minimum peak width = 0.5 min, end level = 30 mAU,

end slope = 50 mAU min^{-1}.

Fractions were analysed by SDS-PAGE and by OD_{488}/OD_{280}.

2.4.10 Preparative purification of untagged EGFP

Preparative purification of EGFP by FPLC was achieved by a six step procedure. An initial $(NH_4)_2SO_4$ precipitation step was followed by HIC using Phenyl-Sepharose 6 F F (Low Sub) followed by IEX on a 5 mL Mono P™ column packed with Q-Sepharose 6 F F on an ÄKTAexplorer™ system. The next step was desalting followed by a high resolution IEX step on a MonoQ™ 5/50 column on an ÄKTAxpress™ system. A final SEC and buffer exchange step was performed using a HighLoad™ 16/16 Superdex™ column.

The mobile phases employed were:

Eluent A : 1.2 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0.

Eluent B: 20 mM Tris-HCl, pH 8.0.

Eluent C: 20 mM Tris-HCl, 1 M NaCl, pH 8.0.

Eluent D: PBS, pH 7.6 or TBS, pH 7.5.

Step 1. $(NH_4)_2SO_4$ precipitation

Cell pellets were lysed in 1 x PBS, pH 8.0 and centrifuged as previously described. $(NH_4)_2SO_4$ in Tris-HCl was added to a final concentration of 1.2 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0. The samples were incubated on ice for 30 min and centrifuged at 16060 x g for 20 min at 4 °C. The supernatant was filtered through a 0.2 µm CE syringe filter.

Step 2. HIC on an XK 45 glass column manually packed with 20 mL Phenyl Sepharose 6 F F (Low Sub).

The flowrate was 5 mL min^{-1} throughout. A program containing the following segments was employed.

Segment 1: Sample loading.

Segment 2: 100 % eluent A for two CV.

Segment 3: 70 % eluent B, 30 % eluent A for three CV.

Segment 4: 100 % eluent B for five CV.

Step 3. IEX on a 5 mL Mono P™ column packed with Q-Sepharose 6 F F.

The flowrate was set at 1 mL min^{-1}. A program containing the following segments was employed.

Segment 1: Sample loading.

Segment 2: 100 % eluent B for five CV.

Segment 3: 15 % eluent C, 85 % eluent B for three CV.

Segment 4: 100 % C for five CV.

Step 4. Desalting

The protein was desalted as described above (2.4.8, step 2).

Step 5. IEX on MonoQ™ 5/50.

High resolution IEX was performed as described above (2.4.8, step 3).

Step 6. SEC on Superdex™ HiLoad™ 75 16/60.

The flowrate was 1 mL min^{-1}. Eluent D was used.

2.4.11 Preparative purification of EGFP-HIT2

EGFP-HIT2 was purified as untagged EGFP (Section 2.4.10) with the following exception: In **segment 3**, the mobile phase composition was 25 % eluent B, 75 % eluent C.

2.4.12 Preparative purification of forced loops F, H and I

E. coli BL-21 cell pellet (30 g wet weight), corresponding to 2.5 L liquid culture, was lysed in 2 x 35 mL 20 mM Tris-HCl, 150 mM NaCl, 5 mM EDTA, pH 8.0. The cell lysate was cleared by centrifugation as described in Section 2.3.2. The supernatant was divided into 10 mL aliquots in 50 mL centrifuge tubes and diluted with lysis buffer if needed. Each 10 mL aliquot was treated as follows.

Three mL 5 M NaCl, 23.3 mL 4 M $(NH_4)_2SO_4$ and 12 mL 96 % etahanol were added. The mixture was shaken vigorously for 30 s and centrifuged for 7 min, 4° C at 3000 x g. The green top ethanolic layer was recovered and 3 mL n-butanol was added. The mixture was shaken vigorously for 30 s and centrifuged for 7 min, 4° C at 3000 x g. The bottom aqueous layer was filtered and further purified by HIC as described above (2.4.10, step 2) followed by SEC as described above (2.4.11, step 6).

2.5 Protein concentration measurements

2.5.1 Bicinchoninic acid (BCA) assay

The BCA assay is an indirect colorimetric assay measuring the purple colored compound formed by Cu^+ and BCA. Since proteins reduce Cu^{2+} to Cu^+ the colorimetric response is dependent on the protein concentration. The BCA assay employs BSA as a standard. Since the reduction rate of Cu^{2+} to Cu^+ depends on both the number of peptide bonds and the abundance of the amino acids cysteine, cystine, tryptophan and tyrosine, the

rate will differ for different proteins, introducing an error. The manufacturer states a coefficient of variance of 14.7 % for 14 different proteins investigated (http://www.piercenet.com/Objects/View.cfm?type=Page&ID=14DFF656-0B71-4F53-8CA1-96C130756EF4 on 2010-06-30). BCA measurements in 96 well formats were performed according to the manufacturer's instructions with the following exception. The absorbance was measured at a wavelength (λ) of 595 nm instead of the recommended 562 nm. 25 µL of all standards and samples were measured in triplicate. Linear standard curves were obtained using Microsoft Excel.

2.5.2 Optical density (OD) measurements

Absorbance measurements are based on Lambert-Beer's law, as shown in Equation 2.1. The absorbance (A), alternatively referred to as the optical density (OD) is a dimensionless unit. l is the pathlength of the light and is commonly 1 cm. ε is the extinction coefficient and c is the concentration of the material being measured. ε can be a molar extinction coefficient with the unit M^{-1} cm^{-1} or a mass extinction coefficient with the unit L g^{-1} cm^{-1}. If a molar extinction coefficient is used c will have the unit M and if a mass extinction coefficient is used c will have the unit g L^{-1}.

Equation 2.1. $$A = \varepsilon * l * c$$

In cases when the extinction coefficient of a protein was unknown, or when measuring mixtures of proteins, a mass extinction coefficient of 1 L g^{-1} cm^{-1} at 280 nm was assumed. This introduces a potential error in the determination of the protein concentration. For the following proteins their published extinction coefficients were used: EGFP, extinction coefficient 55000 cm^{-1} M^{-1} at 488 nm[103]; Calmodulin, extinction coefficient 3006 cm^{-1} M^{-1} at 276 nm[104].

2.6 Buffer exchange and protein concentration

2.6.1 Methodological considerations

Buffer exchange was achieved either by dialysis or by repeatedly concentrating the sample in Vivaspin columns, adding the desired buffer each time. Dialysis is the most effective technique whereas buffer exchange using spin columns commonly gives a 99 % buffer exchange over three spin cycles (according to manufacturer's instructions). Sample loss can be an issue in both techniques. In dialysis collecting the entire sample from the bag

can be difficult. In buffer exchange using spin columns protein can be adsorbed onto the filter and also precipitate.

2.6.2 Dialysis

Samples were dialyzed at 4 ˚C for more than 8 hours in a dialysis buffer more than 100 times the volume of the sample. The dialysis buffer was replaced with more than 100 times the volume of the sample of fresh buffer and the dialysis continued for a further 8 hours. This achieved a total dilution factor of more than 1:10000. The molecular weight cutoff was chosen to be < 50 % of the molecular weight of the sample.

2.6.3 Protein concentration

All protein concentration and buffer exchange procedures were performed using 5000 MWCO spin columns. All spin cycles were performed at 3220 x g for the 50 mL Vivaspin columns and 2415 x g for the 15 mL Vivaspin columns since the filter in smaller columns are more fragile, sometimes breaking at 3220 x g. The spin columns were washed two spin cycles of water followed with two spin cycles of buffer. The samples were applied and the column centrifuged until sufficient concentration was achieved.

2.7 Affinity tag cleavage using enterokinase

Protein (75 ng μL^{-1}) was reacted with 0.03 units μL^{-1} enterokinase EKMax^{TM} in the presence of 1 x EKMax^{TM} reaction buffer at 37 °C for 16 h.

2.8 Protein gel electrophoresis

2.8.1 Methodological considerations

The protein content of samples was assessed using 10 %, 12.5 % or 15 % SDS-PAGE, stained either by silverstaining according to the protocol of Swain and Ross[105] or coomassie blue staining. Silverstaining, being a more sensitive staining technique, was employed as the standard staining procedure for protein identification. Coomassie staining was emplyed to assess the relative abundance of different protein species in a sample.

2.8.2 Experimental procedures

Samples (15 µL) were prepared using 3 µL 5 x loading dye and a maximum of 12 µL protein. When needed PBS or TBS was added to reach the final volume. Samples were incubated at 100 ˚C for 90 s prior to loading on the gel.

All glass plates were cleaned with detergent and 100 % ethanol prior to gel assembly. The two outer lanes were commonly run as blank lanes containing loading buffer but no protein. The gels, prepared as described below, were run at 20 mA/gel until the dye front reached the bottom of the gel.

2.8.2.1 Gel preparation

Gels were 1 mm thick and contained a resolving gel and a stacking gel. The gels were prepared by mixing the solutions listed below, and adding the TEMED and APS last. To prepare resolving gels with lower acrylamide content, smaller amounts of acrylamide solution was used. The total volume was constant by increasing the amount water used.

	Resolving gel (15%)	**Stacking gel**
Acrylamide solution (4K, 30%)	10 mL	0.88 mL
1.5 M Tris-HCl, pH 8.8	5 mL	-
0.5 M Tris-HCl, pH 6.8	-	1.66 mL
10 % SDS	0.2 mL	66 µL
Water	4.7 mL	4 mL
10 % APS	100 µL	33.4 µL
TEMED	14 µL	7 µL

2.8.2.2 Silver Staining

Silver staining was performed according to the protocol by Swain and Ross[105]. Gels were incubated in the following solutions (volumes and masses are for two gels) on a rocking plate.

Time (min)	Solution
10	Fix solution
10	Water
5	Fix/sensitize solution
20	40 % ethanol
20	Water
1	Sodium thiosulphate solution
0	Water
1	Water
20	Silver Nitrate solution
1	Water
As required	Development solution. The gels were developed in ~ 50 % of the solution until the solution turned yellowish. The solution was then replaced by the remaining

fresh development solution. The reaction was quenced by additing ~ 5 % (v/v) glacial acetic acid.

2.8.2.3 Coomassie Brilliant Blue Staining

Gels were incubated in Coomassie Brilliant Blue staining solution for > 1 h at a shaker. The gels were then destained by sequential incubation in 100 mL destaining solution for 1 h until bands were clearly visible on the gels.

Wet gels were photographed on a Chemidoc XRS system (BioRad, Hercules, CA, USA). Dry gels were scanned using a conventional paper scanner.

2.9 Western blot

Two SDS-PAGE gels were run with the same loading plan. One was silverstained or coomassie stained and the other one was used for Western blot. Page Ruler prestained protein ladder was used as standard in the Wester Blot gel. The gel was placed in a transferring cassette and transferred to a Nitropure-supported nitrocellulose membrane Hybond C for 16 h at 50 V or for 2 h at 100 V in transfer buffer. The nitrocellulose membranes were incubated for 2 h in 5 g skimmed milk powder dissolved in 100 mL TBS. The membranes were incubated in 10 mL 0.8 µg mL^{-1} mouse anti-GFP antibody solution in a sealed plastic bag on a rotating disc for 16 h. The plastic bags were pre-washed with ethanol. The membranes were washed 3 x 10 min in TBS-T. The membranes were incubated in 20 mL TBS-T containing 40 µL Rabbit anti-Mouse secondary antibody conjugated with HRP on a shaking board for 80 min and washed 3 x 10 min in TBS-T. Chemiluminescence solution (12 mL) was applied to the membranes. The membranes were photographed in a Chemidoc XRS system. Images were recorded in custom-superchem mode.

2.10 Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-ToF MS)

2.10.1 Sample preparation

Protein samples were desalted using ZipTips by the following pipetting procedure where one cycle refers to one aspiration and one dispensing step with a manual pipette.

The ZipTips were equilibrated by three cycles of ACN/H_2O (75/25 v/v) containing 0.1 % formic acid followed by three cycles of H_2O/0.1 % formic acid. The protein sample was aspired onto the ZipTip by 10 cycles, dispensing to waste in the last cycle. The ZipTip was

washed by three cycles of H_2O/0.1 % formic acid with dispensing to waste. The sample was eluted by three cycles into a suitable volume of ACN/H_2O (75/25 v/v) containing 0.1 % formic acid. The volume was determined by the desired final protein concentration.

2.10.2 Spectra acquisition

Spectra were acquired on a Voyager DE-STR spectrometer in reflection mode at an acceleration voltage of 20 kV. Photon irradiation was performed with a 337 nm pulsed nitrogen laser with 3 ns duration and a maximum firing rate of 20 Hz. Spectra were acquired in the positive ion mode.

2.11 Electrospray ionization mass spectrometry

2.11.1 Sample preparation

Protein samples (40 – 100 µM) were in 2 mM Tris-HCl, pH 7.5 or in water to minimize suppression of the ionization. 1-2 µL $CaCl_2$ stock solution was added to the samples to obtain final protein : Ca^{2+} molar ratios of between 1 : 2 and 1 : 2000. The Ca^{2+} concentration was increased directly between measurements without incubation of the samples.

2.11.2 ESI-MS conditions

Spectra were acquired on a an Agilent CE capillary electrophoresis system coupled to an Agilent 1100 series LC/MSD-SL ion trap mass spectrometer through an Agilent G1607A orthogonal electrospray interface. Electrical contact at the electrospray needle tip was established via a liquid sheath flow delivered by an Agilent 1100 series isocratic LC pump. All system control and data acquisitions were conducted with Agilent ChemStation and MSD Trap Control software.

Samples were infused through a 68 cm x 50 µm ID bare fused silica capillary at a pressure of 50 mbar. The capillary was conditioned by flushing with water for 3 min, 0.1 N NaOH for 3 min, water again for 3 min, and separation buffer electrolyte for 10 min. Between runs, the capillary was flushed with fresh separation buffer electrolyte for 2 min. During all these flushing steps, the mass spectrometer was set to the “standby” mode. All the experiments were conducted under positive polarity mode at 25 kV (with the anode at the CE inlet and cathode at the MS side). The sheath liquid consisted of 0.1 % formic acid, 50 % isopropanol (v/v) and was delivered at a flowrate of 4 µL min^{-1}.

ESI-MS analysis was carried out in the positive ion mode. The electrospray voltage was 3 kV and the mass range was 400-1600 m/z. Averages were obtained from five spectra.

The nebulizing gas (N_2) pressure, the drying gas (N_2 flow rate) and the drying gas temperature were set at 8.00 psi, 5.0 L min^{-1} and 200 °C respectively. The target mass was set at 1000 *m*/*z*, the compound stability was set at 80 %, and the trap drive level was set at 100 %.

Spectra were averages of two minute collections. Between three and five spectra were acquired for each protein at each Ca^{2+} concentration. Ca^{2+} concentrations were increased until the quality of the data did not permit deconvolution of the spectra.

2.12 Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES)

2.12.1 Sample preparation

Gels were packed in 1 mL Tricorn™ columns and loaded with Ca^{2+} as described in Section 2.4.4. The columns were washed with 20 mL of various buffers at 1 mL min^{-1} on a FPLC system. Alternatively IMAC experiments were performed as described in 2.4.7. The gels were transferred to pre-weighed 15 mL falcon tubes, freeze-dried over night and weighed again to obtain the dry weight. In order to avoid sample carry-over and contaminations during the freeze-drying process the tubes were covered with parafilm punctured with a needle. In addition the tubes were sealed with a paper tissue using a rubber band.

Samples were digested in 10 mL 5 M HCl at 50˚C over night. The digested samples were diluted with water to 50 mL using a volumetric flask. The samples were filtered through a 0.45 µm CE syringe filter.

2.12.2 Measurements

Measurements were carried out on a Liberty 200 ICP Emission Spectometer. Concentration standards were prepared in the range 1 to 400 ppm Ca^{2+} and Na^{+}. Ca^{2+} was measured at 396.847 or 422.673 nm. Na^{+} was measured at 589.592 nm. All measurements were performed in triplicate. Measurements were performed employing the following parameter settings: 400 V PTM, 1.20 kW power, 12.0 L min^{-1} plasma, 1.50 L min^{-1} auxiliary, 15 rpm pump speed, 150 kPa nebulizer pressure, 20 s stabilizing time, 10 s rinse time and 30 s sample delay time.

2.13 Atomic Absorption Spectrometry (AAS)

2.13.1 Sample preparation

Samples were prepared as for ICP. Interference suppressor solution was added to a final concentration of 20 % v/v.

2.13.2 Measurements

Ca^{2+} and Na^{+} were measured at 422.7 nm and 589.2 nm respectively. All samples were measured in triplicate.

2.13.3 Data analysis

The calcium content of the gels was calculated using Equation 2.2 where Ca^{2+} refers to the signal obtained from the AAS or ICP experiment, V is the volume of sample, m_{gel} is the mass of freeze dried gel and Mw refers to the molecular weight of the element being measured.

Equation 2.2

$$Ca^{2+}[mol/g] = \frac{Ca^{2+}[g/L] * V[L]}{m_{gel}[g] * Mw_{Ca^{2+}}[g/mol]}$$

The mass of 1 mL gel was estimated to 82.5 mg based on the average of repeated freeze drying experiments of gels containing no salts and no chelated metal ions.

2.14 Competitive chelator titrations

All solutions were prepared using MilliQ water. Beakers for buffer preparation and cuvettes were incubated in EDTA and rinsed extensively with MilliQ water. The instructions provided by Melanie Nelson and Anders Malmedal[106] were followed and involved the following steps.

2.14.1 Sample preparation

Stock solutions of ~ 56 µM 5,5'-Br_2-BAPTA were prepared in 2 mM Tris-HCl, pH 7.5, containing various concentrations of KCl. The concentration of the 5,5'-Br_2-BAPTA stock solutions was determined by measuring the absorbance of the chelator solution in the presence of Ca^{2+} at 239 nm using the extinction coefficient 16000 M^{-1} cm^{-1}. One mL samples containing 28 µM protein and 28 µM 5,5'-Br_2-BAPTA in 2 mM Tris-HCl, pH 7.5 with various KCl concentrations were prepared directly in the cuvette.

2.14.2 Determination of residual Ca^{2+} in the chelator solution

The residual Ca^{2+}-concentration in the chelator stock solutions was determined by measuring the absorbance at 263 nm of 28 µM 5,5'-Br_2-BAPTA in 2 mM Tris-HCl, pH 7.5 with various KCl concentrations: Initial (A1), after adding 2 µL 0.1 M EDTA (A0) and after adding 2 µL 1 M $CaCl_2$ (A2). The Ca^{2+}-concentration was calculated using the Equation 2.3 where [C] denotes the concentration of the chelator 5,5'-Br_2-BAPTA.

.

Equation 2.3 $$[Ca^{2+}] = \frac{A1 - A0}{A2 - A0} * [C]$$

2.14.3 Ca^{2+}-titration

Aliquots of 1 µL of 1 mM $CaCl_2$ were added and the absorbance was monitored at 263 nm. When the decrease in absorbance after adding one aliquot 1 mM $CaCl_2$ was less than 0.015 absorbance units, one µL aliquots of stronger $CaCl_2$ solutions (3 mM, 10 mM, 30 mM, 100 mM and 1 M) were added until the absorbance was constant.

2.14.4 Determination of residual calcium in the protein solution

The residual calcium content of the protein solution was calculated using Equation 2.4. AS is the initial absorbance before any $CaCl_2$ is added, AE is the final absorbance after the last $CaCl_2$ is added in the calcium titration (Section 2.14.3.) and C is the concentration of the chelator. The values of A0 and A2 are those determined in Section 2.14.2.

Equation 2.4 $$[Ca^{2+}] = \left(1 - \frac{AS - AE}{A0 - A2}\right) * C$$

2.14.5 Data analysis

Titration data were analyzed using the program[107] Caligator v. 1.07. The temperature was assumed to be 25° C for all measurements. log $\beta_{Ca(II):L}$ for 5,5'-Br_2-BAPTA was determined at different KCl concentrations using model 1 (1 Calcium site). log $\beta_{Ca(II):L}$ was the only parameter fitted to the dataset. The obtained log $\beta_{Ca(II):L}$ values for 5,5'-Br_2-BAPTA at different KCl concentrations were used as input parameters when fitting the data from the protein titration experiments. For calmodulin the best fit was obtained using model 1 (4 Calcium sites + chelator). For HIT2-tagged GST and EGFP fusion proteins the best fit was obtained using the model 2 (2 Calcium sites + chelator) and log β for one of the binding sites

was set to 0.1. The obtained values of log β were however inconsistent between different experiments. No reliable results could be obtained for HIT2-tagged GST and EGFP fusion proteins.

2.15 Titrations with a Ca^{2+}-selective electrode

2.15.1 Preparation of Ca^{2+} free samples and solutions

In order to remove residual Ca^{2+}, buffers and solutions were incubated at 4 °C for several days in plastic containers containing a dialysis bag filled with 5 mL IDA-Sepharose 6 F F. Protein solutions were passed through Econo columns packed with 2-3 mL IDA-Sepharose 6 F F for five cycles or more. Previously decalcified buffers and solutions were used to equilibrate the IDA-Sepharose 6 F F resin and to elute the protein. Plastic containers and MilliQ water were used throughout.

2.15.2 Ca^{2+} electrode measurement conditions

Measurements were carried out using an ISE25Ca Ca^{2+} selective electrode and a Ref201 reference electrode coupled to a PHM201 pH/mV reader. The filling solutions of the Ca^{2+} selective electrode and reference electrode were changed each week and the function of the electrode was validated by recording the mV response of 0.01 µM $CaCl_2$ in 100 mM KCl and of 10 mM $CaCl_2$ in 100 mM KCl. According to the manufacturer the difference in the readings should exceed 55 mV. If this was not the case the electrode was dissembled and reassembled. All experiments were performed in 50 mM KCl, 2 mM Tris-HCl, pH 7.5 to maintain a constant ionic strength between samples. Prior to each measurement the response was recorded for EGTA/Ca^{2+} solutions in 2 mM Tris-HCl, 50 mM KCl, pH 7.5 calculated according to the EGTA calculator http://brneurosci.org/egta.html to have a free Ca^{2+} concentration of 9.1 nM (1.5 mM $CaCl_2$, 5 mM EGTA), 0.10 µM (4.15 mM $CaCl_2$, 5 mM EGTA), 1.0 µM (4.9 mM $CaCl_2$, 5 mM EGTA) and 1.6 µM (4.9375 mM $CaCl_2$, 5 mM EGTA). These readings were used to validate the electrode response between different measurements. All measurements were performed in 10 mL plastic beakers with a magnetic stirrer. Prior to each measurement, the electrodes were equilibrated in 5 mL protein solution and the measurement was performed using 5 mL fresh solution.

2.15.3 Ca^{2+} titrations

Standard curves were obtained by dispensing 6 mL 50 mM KCl, 2 mM Tris-HCl, pH 7.5 in a beaker, retaining 1 mL of this solution for determination of initial Ca^{2+} concentration using 5,5'-Br_2-BAPTA (Section 2.14.4). To the remaining 5 mL solution, aliquots of 2 µL $CaCl_2$ (1 mM, 3 mM, 10 mM, 30 mM, 100 mM, 300 mM and 1 M) in 50 mM KCl, 2 mM

Tris-HCl, pH 7.5 were added and the electrode response was recorded. The procedure was repeated immediately with samples containing 8 µM decalcified protein solution in 50 mM KCl, 2 mM Tris-HCl, pH 7.5.

2.15.4 Data analysis

The initial Ca^{2+} concentrations determined by the 5,5'-Br_2-BAPTA method (Section 2.14.4) fluctuated between 0 and 3 µM. An initial Ca^{2+} concentration of 1.0 µM was therefore assumed for all standards. For each experiment, a standard curve was generated by plotting the logarithm of the concentration of free Ca^{2+}, log[Ca^{2+}_{free}] *versus* the electrode output in millivolt (mV). Excluding the lowest readings, a line was fitted to the plot, yielding the first order Equation 2.5.

Equation 2.5 $mV = log[Ca^{2+}_{free}] + k$

For each of the sample calibration points, [Ca^{2+}_{free}] was calculated using Equation 2.4. For concentrations below 3 µM, [Ca^{2+}_{free}] was determined by manual interpolation. [Ca^{2+}_{bound}] was calculated by subtracting [Ca^{2+}_{free}] from the total amount of Ca^{2+} added. The ratio *r* was defined as [Ca^{2+}_{bound}]/[protein]. r was then plotted *versus* [Ca^{2+}_{free}].

For calmodulin the data were fitted to a binding equation for one-site binding with a Hill slope (Equation 2.6) where K_d is the dissociation constant and h is the Hill slope.

Equation 2.6 $$r = \frac{r_{\max} \cdot [Ca^{2+}]^h}{K_d{}^h + [Ca^{2+}]^h}$$

2.16 Protein unfolding studies

2.16.1 Sample preparation

All buffers were filtered through a 0.22 um syringe filter. The guanidine hydrochloride (GuHCl) concentration in stock solutions was determined by refractive index measurements using an R-5000 hand refractometer (Atago). The concentration was calculated using Equation 2.7 where ΔN is the difference in refractive index between the GuHCl solution and the buffer in which it was prepared (10 mM phosphate buffer, 0.65 M NaCl, pH 7.6) and *c* is the concentration of GuHCl.

Equation 2.7 $$c = 57.147\Delta N + 38.68(\Delta N)^2 - 91.60(\Delta N)^3$$

Protein samples were prepared in 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6 and with concentrations of GuHCl varying between 0 and 6 M. For the simultaneous CD and fluorescence measurements, 1 mL samples were prepared containing 200 mg mL^{-1} protein. For the fluorescence measurements performed in a platereader, 50 – 100 µL samples were loaded in each well of black 96 well plates. The protein concentration was 1 µM.

2.16.2 CD measurement of unfolding curves

Spectra were acquired on a Jasco J-815 CD spectrometer fitted with an FMO-427S/L fluorescence monochroamator, controlled by Spectra Manager™ II. Spectra were averages of three acquisitions from 260 to 190 nm using a digital integration time (DIT) of 4 seconds, a band width of 1 nm, and 0.1 nm data pitch at a scanning speed of 100 nm min^{-1}. The start mode was set to immediate, the scanning mode was set to continuous, the sensitivity was set at standard and no baseline correction was applied.

2.16.3 Fluorescence measurements in CD spectrometer

Simultaneous circular dichroism (CD) and fluorescence measurements were performed on a Jasco J-815 CD spectrometer fitted with an FMO-427S/L fluorescence monochroamator, controlled by Spectra Manager™ II. EGFP fluorophore fluorescence was measured by exciting at 488 nm and detecting between 450 and 600 nm (maximum emission was observed between 510 and 512 nm). The excitation bandwidth was set to 5 nm, the data pitch to 1 nm and DIT to 4 sec. The sensitivity was set to 450 V.

Tryptophan fluorescence was measured by exciting at 295 nm and detecting between 300 and 450 nm. The sensitivity was set to 700 V. All other settings were as above. Tyrosin/Tryptophan fluorescence was measured by exciting at 280 nm and detecting between 300 and 450 nm. The sensitivity was set to 700 V. All other settings were as above.

2.16.4 Unfolding curves in platereader

Measurements were carried out in an Infinite® 200 plate reader. The excitation bandwidth was 9 nm and the emission bandwidth was 20 nm. The number of flashes was set to 50. The integration time was 20 µs. The lag time was 20 µs. To avoid overlap between excitation and emission, EGFP was excitated at 480 nm instead of 488 nm and the emission was measured at 512 nm. The gain was adjusted for each plate.

2.16.5 Kinetic measurements in platereader

Measurements were carried out on a Infinite® 200 plate reader as described in 2.16.4. Data were acquired every five minutes.

2.16.6 Data analysis of unfolding curves

CD and fluorescence data from the GuHCl unfolding experiments were fitted to a two-state model[108] (Equation 2.8) accounting for non-linear baselines on both sides of the transition.

Equation 2.8
$$Y = \frac{(F_n + M_n x) + (F_u + M_u x)\exp\left(\left(-\Delta G \cdot M_g x\right)/RT\right)}{1 + \exp\left(-\left(\Delta G - M_g x\right)/RT\right)}$$

Y is the observed intensity of optical property (CD signal or fluorescence), x is the denaturant concentration, F_n and F_u are the observed intensity before and after the transition respectively, M_n and M_u are the slopes of the baseline before and after the transition respectively and M_g is the slope of the transition. F_n, M_n, and M_g were shared between the data sets. Data filling of points with zero fluorescence was applied between 4 and 5 M GuHCl for the plate reader experiments to improve the curve fitting. The fluorescence had decreased to zero at a GuHCl concentration of less than 4 M.

2.16.7 Kinetic data of kinetic measuremenst

The fluorescence decays as a function of time were fitted to a one-phase decay curve using Equation 2.9. Y is the fluorescence signal as a function of time, Y(0) is the fluorescence at time 0, k is the rate constant and t is the time.

Equation 2.9 $Y(t) = Y(0)\, e^{-kt}$

2.17 Terbium binding experiments

2.17.1 Sample preparation

The concentration of the $TbCl_3$ stock employed was determined by elemental analysis by Campbell Microanalytical Laboratory, University of Otago (Dunedin, New Zealand). Two µM protein samples containing varying concentrations of $TbCl_3$ were prepared in 2 mM Tris-HCl, pH 7.5. 100 µL samples were measured in black 96 well plates.

2.17.2 Luminescence resonance energy transfer (LRET) measurements

Measurements were carried out on an Infinite® 200 plate reader. The excitation bandwidth was 9 nm and the emission bandwidth was 20 nm. The number of flashes was set

to 50 and the delay time was 50 µs. The integration time was 200 µs. The excitation wavelength (λ_{exc}) was at 280 nm and the emission wavelength (λ_{emm}) was 512 nm. The gain was kept constant for all plates.

2.17.3 Fluorescence measurements

Measurements were performed as in 2.16.4 with the exception that the gain was set manually to be equal for all plates.

2.17.4 Data analysis

The data were fitted to the model "one-site fit total and nonspecific binding" contained in the software package Graphpad Prism 5.0. The calculation employs Equations 2.10 where $B_{Specific}$ represents the specific binding, B represents the nonspecific binding, K_d is the dissociation constant, BGR represents the background signal, NS is the slope of the nonspecific binding, Y_{max} represents the signal at saturation and x is the concentration of ligand (Tb^{3+} here). $B_{specific}$ + B was then fitted to a dataset representing total binding. Simultaneously B was fitted to a dataset obtained for nonspecific binding. Here the dataset obtained for untagged EGFP was employed to represent the nonspecific binding (B).

Equations 2.10

$$\begin{cases} B_{Specific} = \dfrac{Y_{\max} \cdot x}{x + K_d} \\ B = NS \cdot x + BGR \end{cases}$$

2.18 Ligand synthesis and immobilization

2.18.1 Ligand synthesis

Alendronate was prepared following the method of Kieczykowski *et al.*[109] DO2A and DO2P were both prepared following the method of Kovacs and Sherry[110; 111] from 1,7-bis(Cbz)cyclen, prepared by the improved method of De León-Rodriguez *et al.*[112] DO3A was prepared following the method of Dischino *et al.*[113] DO3P was synthesised by a method similar to the one described by Sun *et al.*[114] Tacn was synthesised based on the method of Richman and Atkins[115; 116]. Propyl-*bis*-tacn was synthesised using tacn orthoamide by modifying the method of Zhang *et al.*[42]

2.18.2 Ligand immobilization

The ligands were immobilized onto Sepharose 6 F F by modifying methods previously described[67; 94; 95]. Briefly, ligand solution was adjusted to pH 12 using NaOH.

Suction-dried epichlorohydrin-activated Sepharose 6 F F was added. The volumetric ratio of igand solution to suction-dried activated Sepahrose 6 F F was 1 : 1. The mixture was shaked on an AXYOS orbital shaker or mixed on a Ratek suspension mixer. The concentration of ligand solution, times and temperatures used for the different ligands are listed below.

Alendronate:	0.2 M ligand, 23 ºC for five days.
Cyclen, DO2P, tacn, propyl-*bis*-tacn:	0.2 M ligand, 28 ºC for 4.5 days.
IDA:	0.38 M ligand, 28 ºC for 3.5 days.
DO2A, DO3A:	0.2 M ligand, 28 ºC for 21 h.
DO3P:	0.2 M ligand, 28 ºC for 21 h or 4.5 days.

Chapter 3. Comparison of N- and C-terminally tagged EGFP on borderline metal IMAC systems

3.1 Introduction

3.1.1 Borderline metal IMAC systems based on macrocyclic ligands

As mentioned in the Introduction, macrocyclic compounds form more stable complexes with metal ions as compared to non macrocyclic ligands with the same donor groups and denticity. This is due to the so called macrocyclic effect[26]. Studies of novel borderline metal IMAC systems based on tacn[43] and tacn derivatives[45; 46; 47] (Figure 3.1) were therefore undertaken in the laboratories of my supervisors of the Monash University Centre for Green Chemistry. Such systems would have the advantage of less metal displacement from the ligand due to the higher stability constant[40; 41; 42] of the metal chelate compared to non macrocyclic ligands like IDA and NTA[28]. This would lead to less generation of potentially toxic waste products. It would also require less frequent metal re-loading of the resin(s). These studies are described in detail the Introduction (Section 1.2.4.2).

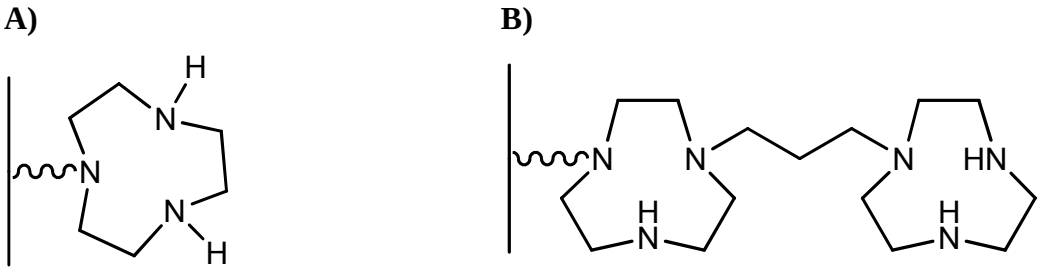


Figure 3.1. Chemical structures of the IMAC ligands **A)** tacn and **B)** propyl-*bis*-tacn used in this work. The spacer arm is schematically drawn as a wavy bond.

The affinity tag NT1A with the amino acid sequence HHHNSWDHDANR was developed in my supervisor's laboratory based on phage display library screening with the metal chelates of tacn derivatives[68]. Similar peptide tags for use with IMAC were also patented[67]. The NT1A tag contains His residues at positions 1, 2, 3 and 8. A two-point

binding with the two metal centres of *bis*-tacn metal chelates could therefore occur. This could result in high affinities between bis-tacn metal chelates and recombinant fusion proteins containing the NT1A tag.

3.1.2 Differences of N- and C-terminally tagged proteins

Previous studies with immobilized metal chelates of tacn and propyl-*bis*-tacn and recombinant fusion proteins containing the NT1A tag were performed in my supervisor's laboratory by the graduate student Darant Longford and his postdoctoral fellows, Dr Dale Fredericks and Dr Chunfang Zhang. The NT1A tag was fused to the N-terminus of amylase, xylanase and EGFP. The NT1A tag was used alone or in sequence with the putative hard metal ion binding HIT2 tag[95] with the amino acid sequence DIDGDGHINYEE. The two tags NT1A and HIT2 were employed in sequence to facilitate sequential IMAC protocols. Such protocols could contain a borderline IMAC step followed by a hard IMAC step or *vice versa*.

These studies revealed that NT1A tagged fusion proteins were retained on Cu^{2+}-tacn-Sepharose 6 F F, Ni^{2+}-tacn-Sepharose 6 F F, Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F and Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. When *E. coli* BL-21 was employed as expression host, differences in protein expression levels were observed for different arrangements of the HIT2 and NT1A tags. The arrangement HIT2-NT1A-protein gave higher expression levels as compared to the arrangement NT1A-HIT2-protein. This pattern was clearly observed when the tags were fused to EGFP and xylanase but was not as pronounced when the tags were fused to Amylase (unpublished data).

3.2 Aim

The aim of the studies undertaken in this Chapter was to investigate if the differences in expression levels observed for N-terminally tagged fusion proteins would remain if the tags were instead fused to the C-terminus of the protein. Fusing affinity tags to the C- instead of the N-terminus of the protein can potentially affect the chromatographic behaviour of the fusion protein. As a consequence the chromatographic behaviour of N- and C-terminally tagged fusion proteins needed to be investigated.

3.3 Results

To simplify the experimental work, we chose EGFP as a model protein (see Chapter 4 for details). NT1A and HIT2 were fused to the C-terminus of EGFP as single tags and in sequence, creating the fusion proteins EGFP-HIT2-NT1A, EGFP-NT1A-HIT2, EGFP-NT1A and EGFP-HIT2. The fusion protein EGFP-His_6 was included as a positive control for the

NT1A tag. Genes for the N-terminally tagged fusion proteins HIT2-NT1A-EGFP, NT1A-HIT2-EGFP and NT1A-EGFP had previously been created by Mr Darant Longford, Dr Dale Friedericks and Dr Chunfang Zhang (unpublished data).

3.3.1 Creating C-terminally tagged EGFP constructs

Five C-terminally tagged recombinant fusion proteins EGFP-HIT2-NT1A, EGFP-NT1A-HIT2, EGFP-NT1A, EGFP-HIT2 and EGFP-His6 were cloned and expressed by Dr Chunfang Zhang. The amino acid sequences of the tags attached to EGFP are listed in Table 3.1. The double tag fusion proteins contain a five glycine spacer between the NT1A and HIT2 tags. The D_4K enterokinase cleavage site present in the N-terminally tagged fusion proteins was omitted when the tags were fused to the C-terminus of EGFP as the D_4K sequence only can be completely removed when located at the N-terminus of a protein (as discussed in Section 1.2.5).

Table 3.1. Amino acid sequences of the tags attached to EGFP used in this chapter. The tags NT1A and HIT2 are displayed in bold. EGFP is displayed in italic.

Fusion protein	Amino acid sequence	Cloned by
EGFP-HIT2-NT1A	*EGFP*-**DIDGDGHINYEE**FGGGGG**HHHNSWDHDANR**	Zhang and Hearn
EGFP-NT1A-HIT2	*EGFP*-**HHHNSWDHDANR**GGGGG**DIDGDGHINYEE**F	Zhang and Hearn
EGFP-NT1A	*EGFP*-**HHHNSWDHDANR**	Zhang and Hearn
EGFP-HIT2	*EGFP*-**DIDGDGHINYEE**F	Zhang and Hearn
EGFP-His_6	*EGFP*-**HHHHHH**	Zhang and Hearn
NT1A-HIT2-EGFP	MK**HHHNSWDHDANR**GGGGG**DIDGDGHINYEE**FGGGGGDDDDK-*EGFP*	Zhang, Fredericks and Hearn
HIT2-NT1A-EGFP	MK**DIDGDGDINYEE**FGGGGG**HHHNSWDHDANR**GGGGGDDDDK-*EGFP*	Zhang, Fredericks and Hearn
NT1A-EGFP	MK**HHHNSWDHDANR**VDQ-*EGFP*	Longford and Hearn

3.3.2 Characterization of C-terminally tagged EGFP

The migration pattern of the C-terminally tagged EGFP's in cell lysates on a 15 % SDS-PAGE gel was investigated by anti-GFP western blotting (Figure 3.2). The N-terminally tagged proteins NT1A-EGFP, NT1A-HIT2-EGFP and HIT2-NT1A-EGFP were included for

comparison. It is worth noting that NT1A-HIT2-EGFP and HIT2-NT1A-EGFP migrate as if they were ~ 8 kDa bigger than their C-terminal counterparts (Figure 3.2). This difference can not be explained by the mass of the D_4K cleavage site present in the N-terminal tags and absent in the C-terminal tags.

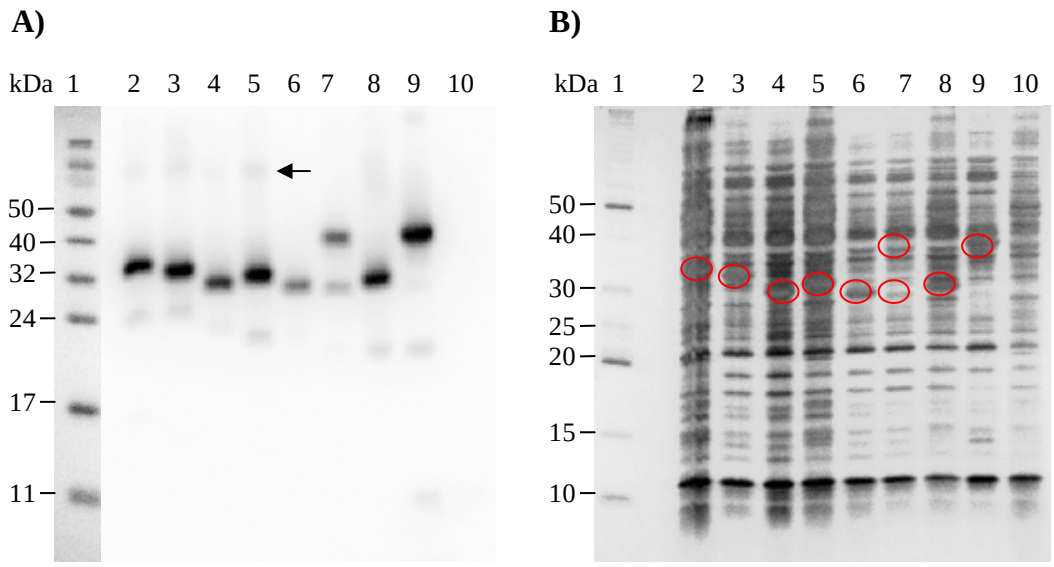


Figure 3.2.A) Anti-GFP Western Blot and **B)** silverstained 15 % SDS-PAGE. Lane 1: size marker, lane 2: EGFP-HIT2-NT1A, lane 3: EGFP-NT1A-HIT2, lane 4: EGFP-NT1A, lane 5: EGFP-HIT2, lane 6: EGFP-His_6, lane 7: NT1A-HIT2-EGFP, lane 8: NT1A-EGFP, lane 9: HIT2-NT1A-EGFP, lane 10: *E. coli* BL-21 (negative control for Western Blot). Bands in SDS-PAGE gel corresponding to signals in Western Blot are circled. The signal corresponding to the EGFP dimer is indicated with an arrow.

It could be due to differences in unfolding when the protein is subjected to SDS in the gel electrophoresis. It could also be due to differences in steric accessibility of the tags between the N- and C-terminally tagged fusion proteins. High molecular weight signals in the western blot (marked with an arrow in Figure 3.2) can be due to dimer formation, which is known to occur for EGFP[103]. The lower molecular weight signals at ~ 20 kDa are thought to be a degradation product of EGFP.

The molecular weights of the fusion proteins were determined by MALDI-ToF MS. Protein samples analysed were purified by a single step IMAC purification yielding semi-pure protein as described in Section 3.3.4. An example of a mass spectrum is shown in Figure 3.3. The signal at 30 059.71 Da correspond to the monoisotopic weight of EGFP-NT1A-HIT2. A summary of the mass spectrometry results for all C-terminally tagged fusion proteins and their molecular weights assessed by western blot (Figure 3.2) are listed in Table 3.2. NT1A-EGFP and HIT2-NT1A-EGFP were included as positive controls since their masses had been

determined previously by MALDI-ToF-MS by Longford, Harris and Hearn (unpublished data). The masses obtained for NT1A-EGFP and HIT2-NT1A-EGFP were slightly higher than the theoretical monoisotopic molecular weights, but within the precision of the instrument. The masses of the C-terminally tagged fusion proteins were on the other hand all between 121 and 135 Da lower than their theoretical masses. This mass difference was too great to merely be due to errors associated with the precision of the instrument.

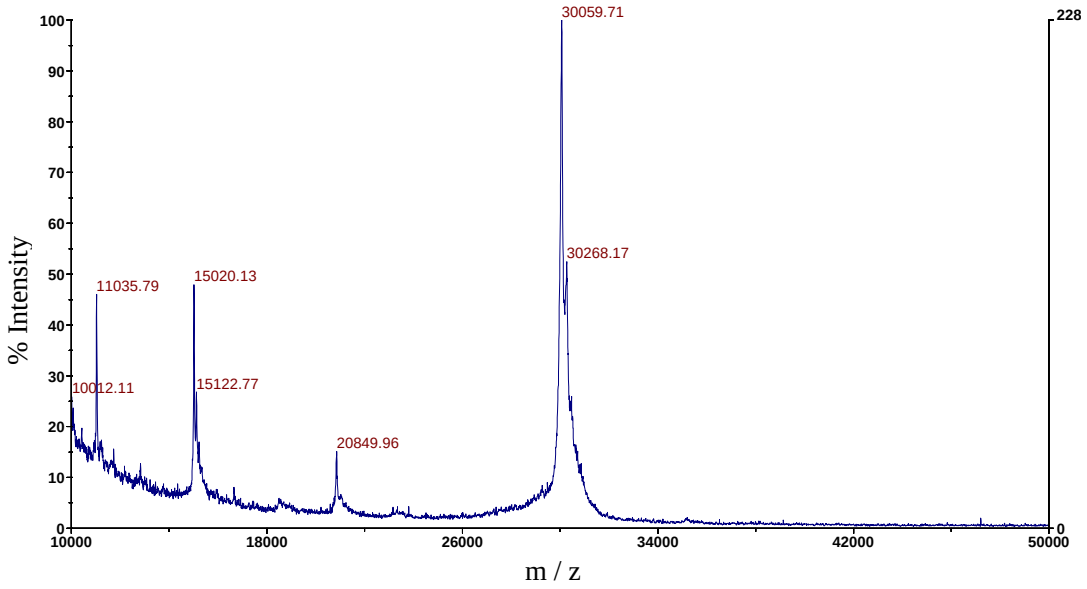


Figure 3.3. MALDI-ToF MS of EGFP-NT1A-HIT2 showing the monoisotopic molecular weight 30059.71 Da of the construct.

The molecular weight of two proteins, untagged EGFP and EGFP-HIT2, were therefore determined by ESI-MS and found to be 26746 and 28251 Da, respectively. Both masses were 131 Da lower than the theoretical monoisotopic masses. Methionine has a molecular weight of 146 Da. Adjusting for the dehydration occurring in the peptide bond formation the loss of the starting methionine would result in a mass change of 131 Da.

Cleavage of N-terminal methionine has previously been observed both in eukaryotes such as *Saccharomyces cerviciae*[117] and in several strains of *E. coli*[118; 119] including the strain BL-21 used in this work[120]. Cleavage has been correlated with the gyration radius[121] of the neighbouring amino acid sidechain. A gyraton radius of < 1.29 Å resulted in 90-100 % processing of the starting methionine[122] of amino terminal extended human growth hormone expressed in *E. coli*[123] MC1061. The second amino acid in EGFP is a serine residue with a gyration radius[121] of 1.08 Å. The second amino acid in the N-terminally tagged constructs NT1A-EGFP, NT1A-HIT2-EGFP and HIT2-NT1A-GFP is lysine with a gyration radius[121] of

2.08 Å. Hence, it is reasonable to assume that the mass difference observed for the C-terminally tagged recombinant proteins was caused by processing of the starting methionine.

Table 3.2. Summary of molecular weights (Mw) of C-and N-terminally tagged EGFP constructs determined by western blot and MALDI-ToF MS compared to their theoretical values determined by Expacy "Compute pI/Mw tool" (http://ca.expasy.org/tools/pi_tool.html). All numbers are in Dalton. The number in brackets represent the difference between the theoretical molecular weight and the molecular weight obtained from the MALDI-ToF-MS experiments. n.d: not determined.

Construct	Theoretical Mw (average)	Mw by western blot	Theoretical Mw (monoisotopic)	Mw by MALDI-TOF MS
EGFP-HIT2-G5-NT1A	30192.67	33000	30173.77	30039 (-135)
EGFP-NT1A-G5-HIT2	30192.67	32000	30173.77	30060 (-114)
EGFP-NT1A	28401.89	29000	28384.06	28260 (-124)
EGFP-HIT2	28399.90	30000	28382.04	28261 (-121)
EGFP-His_6	27717.23	29000	27699.79	27577 (-122)
NT1A-G5-HIT2-D_4K-EGFP	31 325.83	38 000	31 306.22	n.d.
HIT2-G5-NT1A-D_4K-EGFP	31325.83	38000	31306.22	31336 (+30)
NT1A-EGFP	29003.61	30000	28985.35	29019 (+37)

3.3.3 Comparing expression of N- and C-terminally tagged EGFP

The previously observed difference in expression of the N-terminal double-tagged proteins HIT2-NT1A-EGFP and NT1A-HIT2-EGFP was semi-quantified and compared to their C-terminal counterparts EGFP-HIT2-NT1A and EGFP-NT1A-HIT2 and to untagged EGFP. The expression of three individual colonies of *E. coli* BL-21 containing the genes coding for each recombinant protein was studied. Culturing and expression were carried out as described in detail in the Materials and Methods Section 2.3.1. Briefly, cultures were grown in 2 mL LB-medium over night. The OD_{600} was measured using LB-medium as a blank. A linear relationship was assumed between OD_{600} and the cell density of each culture. Based on the OD_{600} reading, identical numbers of cells of each culture were expanded to 11

mL by adjusting the amount of starter culture and LB-medium. Following expression, cell pellets of each culture were lysed in 10 mM phosphate buffer, 150 mM NaCl, pH 7.6 as described in the Materials and Methods Section 2.3.2 and cleared by centrifugation. The amount of expressed recombinant protein in the cleared cell lysate of each culture was determined by measuring the EGFP fluorescence at 512 nm (λ_{exc} = 480 nm). The concentration was calculated against a standard curve of EGFP of known concentration, purified as described in the Materials and Methods Section 2.4.10. The absorbance at 280 nm was used to estimate of the total protein concentration in each sample as described in Materials and Methods Section 2.5.2. A mass extinction coefficient of 1 L g^{-1} cm^{-1} was used. The proportion of fluorescent recombinant EGFP proteins in the total protein content of the cleared cell lysate could hence be calculated. The results are shown in Table 3.3. In the cleared lysates of *E. coli* BL-21 cells expressing EGFP, HIT2-NT1A-EGFP, EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A 0.7-1.4 % of the total protein was tagged EGFP. The corresponding values for NT1A-HIT2-EGFP were 0.2-0.3 %. No fluorescence was detected in lysates of *E. coli* BL-21 expressing no EGFP. An Anova analysis with a Tukey post test showed that the difference in expression levels were significant between NT1A-HIT2-EGFP and the other four proteins. The difference in expression levels was not significantly different between HIT2-NT1A-EGFP, EGFP, EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A.

Table 3.3. Content of tagged and untagged recombinant EGFP in lysates of *E. coli* BL-21. Values are given as % fluorescent protein of the total amount of protein in cleared cell lysates.

Fusion protein	Colony 1	Colony 2	Colony 3	Mean ± Stdv
EGFP	1.46	0.987	0.794	1.1 ± 0.3
EGFP-HIT2-NT1A	0.943	1.20	0.832	1.0 ± 0.2
EGFP-NT1A-HIT2	0.963	0.915	0.963	0.95 ± 0.03
NT1A-HIT2-EGFP	0.269	0.298	0.233	0.27 ± 0.03
HIT2-NT1A-EGFP	0.911	0.886	1.22	1.0 ± 0.2

Affinity tags can affect the structure and stability of the protein to which it is attached[124; 125; 126; 127; 128] as discussed in greater detail in Chaper 4. The lower fluorescence of the cleared lysates of *E. coli* BL-21 cells expressing NT1A-HIT2-EGFP could therefore potentially be caused by misfolding of the protein, rather than a low expression level. This could potentially result in expression of EGFP molecules with lower or no fluorescence. The EGFP content of the cell lysates was therefore analysed by western blotting. It is evident that

the signal corresponding to NT1A-HIT2-EGFP marked with an arrow in Figure 3.4 is weaker as compared to the other proteins. This pattern is observed for cell lysates originating from three different colonies. It can therefore be concluded that the lower fluorescence of cleared lysates of *E. coli* BL-21 cells expressing NT1A-HIT2-EGFP was due to lower abundance of this protein in the cell lysate. Hence, the results in Table 3.3 confirm that the expression of soluble NT1A-HIT2-EGFP is lower compared to untagged EGFP. The expression levels of soluble HIT2-NT1A-EGFP, EGFP-HIT2-NT1A, EGFP-NT1A-HIT2 and untagged EGFP were on the other hand not significantly different. Hence, the arrangement of the affinity tags NT1A and HIT2 did not affect the expression of EGFP fusion proteins when the tags were fused to the C-terminus of the protein.

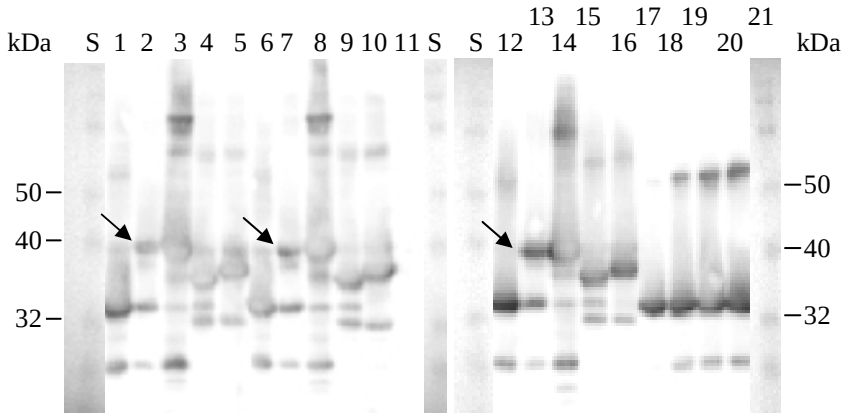


Figure 3.4. Anti-GFP Western Blots of cell lysates. Total protein (57 µg) was loaded in each lane, corresponding to 160-830 ng EGFP. S: Size marker. Three cell lines were analysed for each recombinant protein. Lanes 1, 6 and 12: EGFP. Lanes 2, 7 and 13: NT1A-HIT2-EGFP. Lanes 3, 8 and 14: HIT2-NT1A-EGFP. Lanes 4, 9 and 15: EGFP-NT1A-HIT2, Lanes 5, 10 and 16: EGFP-HIT2-NT1A. Lane 11: Cell lysate of *E. coli* BL-21 containing no genes coding for EGFP (negative control). Lane 18: EGFP, 0.3 µg. Lane 19: EGFP, 0.6 µg. Lane 20: EGFP, 1.2 µg. Lane 21: EGFP, 1.8 µg.

The expression levels of around 1 % soluble EGFP of the total proteins obtained in these studies (Table 3.3) were calculated assuming a mass extinction coefficient of 1 g^{-1} L^{-1} at 280 nm for the total protein contents in the cell lysates. As a comparison, the expression level of GFP_{uv} in *E. coli* DH5-α was found to be 20 µg mg^{-1} total proteins[129] (total protein was measured against a standard curve of BSA). This corresponds to 2 % of the total proteins. McRae *et al*. obtained expression level of EGFP in *E. coli* DH5-α of 16 % of the total protein[103]. These investigators employed the Bradford assay with BSA as a standard to estimate the total proteins in the cell lysate. Hence the expression levels obtained here appear lower as compared to some values from the literature. This could reflect different methods in calculating the expression levels or lower actual expression.

3.3.4 IMAC on borderline metal ion systems

To investigate if fusing the tags to the C-terminus of EGFP would alter the chromatographic behaviour of the fusion proteins, FPLC IMAC experiments were performed. All IMAC experiments were performed on manually packed 1 mL columns with identical mobile phase conditions as described in detail in the Materials and Methods Section (2.4.6). Cell lysate corresponding to 50 mL liquid cell culture of *E. coli* BL-21 expressing the respective recombinant EGFP proteins was used in each experiment. The pellet was lysed in 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. The sample was loaded and the column was washed with 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Elution was achieved employing a 10 mL linear gradient from 10 to 300 mM imidazole in 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Fractions of 1 mL were collected throughout the entire chromatography experiment. Due to the fluorescent properties of EGFP, binding of the protein to the IMAC resins could be visually observed throughout the experiment. Fractions for subsequent protein analysis by SDS-PAGE could be selected based on their green fluorescence. IMAC experiments were performed using Ni^{2+} and Cu^{2+} chelates of tacn-Sepharose 6 F F and propyl-*bis*-tacn-Sepharose 6 F F. Commercially obtained Ni^{2+}-NTA-Agarose (Ni-NTA His●Bind®, Merck, product number 70666) was used as a positive control for the tacn systems. The His_6 tag was used as a positive control for the borderline metal tag NT1A. Negative control experiments were performed with untagged EGFP.

3.3.4.1 Ni^{2+}-NTA

EGFP-His_6 and Ni^{2+}-NTA-Agarose were used as a positive control for the borderline metal IMAC systems. A chromatogram and corresponding 15 % SDS-PAGE gel are shown in Figure 3.5. The chromatogram shows a high UV response in the flow through fractions (fractions 3 to 5) corresponding to the majority of the *E. coli* host proteins not being retained on the column. It is evident from the SDS-PAGE (Figure 3.5) that the target protein was bound on the column during the load and wash steps. The band at 28 kDa corresponding to EGFP-His_6 is absent in the wash fractions but present in the cell lysate and in the elution fraction 14 – 19 (Figure 3.5). The wash step removed almost all unbound proteins from the resin. This can be seen in fraction 11 where the absorbance at 280 nm in the chromatogram is close to zero and the protein content in the corresponding lane of the SDS-PAGE (Figure 3.5) is very low. Two incompletely resolved elution peaks can be observed in the chromatogram. Fraction 13 (Figure 3.5) belongs to the first of the unresolved peaks and contains some *E. coli* host proteins, possibly containing surface His groups, but no EGFP-His_6. These proteins were eluted at a range in the gradient corresponding to an imidazole concentration range of 10-130

mM. This range corresponds to fractions 11-14 (Figure 3.5). As the imidazole concentration increases EGFP-His_6 is eluted, which can be observed in fractions 15-19. These fractions correspond to an imidazole concentration of 130-270 mM. Fractions 16-18 show high abundance of EGFP-His_6 on SDS-PAGE (Figure 3.5) but also additional bands. The EGFP dimer[103] and degradation product observed on anti-GFP western blot (Figure 3.2, page 65) can be identified at ~ 60 kDa and at ~ 20 kDa, respectively. Besides these relatively strong bands the abundance of other individual contaminants does not appear to be significant as judged by SDS-PAGE (Figure 3.5, fractions 16-18).

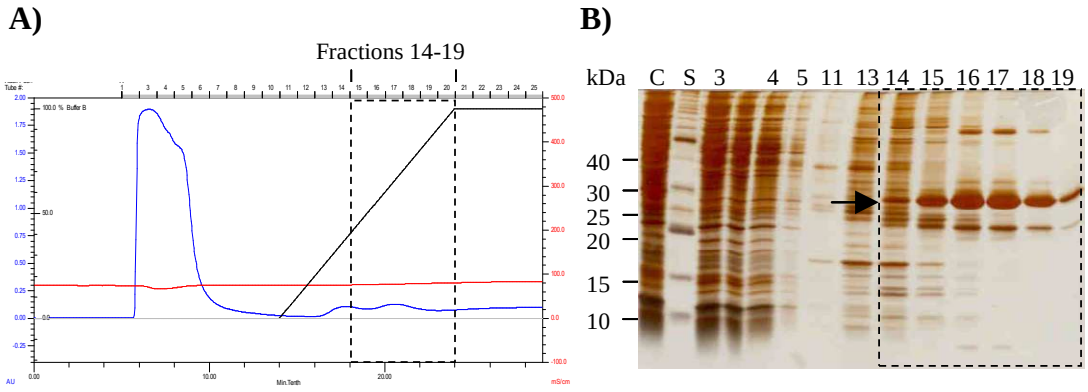


Figure 3.5A) Chromatogram of EGFP-His_6 on a 1 mL column packed with Ni^{2+}- NTA-Agarose. The flowrate was 1 mL min^{-1} throughout. Eluent A was 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Eluent B was 300 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Black trace: Per cent eluent B in mixture. Blue trace: absorbance at 280 nm. Red trace: conductivity. Fractions of 1 mL were collected as indicated at the top of the chromatogram. **B)** 15 % SDS-PAGE of selected fractions from the chromatography experiment. C: Starting material (cleared cell lysate) diluted 1/24; S: size marker. Flow through fractions 3, 4 and 5 were diluted 1/16. Elution fractions were loaded undiluted. 15 µL was loaded in each lane. GFP-His_6 bands are marked with an arrow. Visually green fractions are marked with a dashed box in the chromatogram and the SDS-PAGE.

EGFP fused to the NT1A tag alone, or the NT1A tag in sequence with the HIT2 tag, also bound to Ni^{2+}- NTA-Agarose. All recombinant proteins investigated (NT1A-EGFP, HIT2-NT1A-EGFP, NT1A-HIT2-EGFP, EGFP-NT1A, EGFP-HIT2-NT1A and EGFP-NT1A-HIT2) were eluted at a range in the gradient corresponding to an imidazole concentration of 70-180 mM (Figure 3.6). They were thus eluted at a lower imidazole concentration as compared to EGFP-His_6. The chromatograms contained one elution peak (Figure 3.6A-F). The imidazole concentration (70-180 mM) at which the NT1A tagged proteins were eluted overlaps with the imidazole concentration where only *E. coli* host protein was eluted (10-140 mM) in the EGFP-His_6 experiment (Figure 3.5). It also overlaps with the

elution peak at 100-180 mM imidazole observed for untagged EGFP (Figure 3.6A). Moreover the fluorescent fractions of the NT1A tagged proteins analysed by SDS-PAGE (Figure 3.6B-F) contained contaminating bands. It can thus be concluded that EGFP tagged with NT1A is retained to a lesser extent on Ni^{2+}-NTA-Sepahrose as compared to EGFP-His_6. The lesser retention leads to coelution of *E. coli* host protein under these buffer conditions.

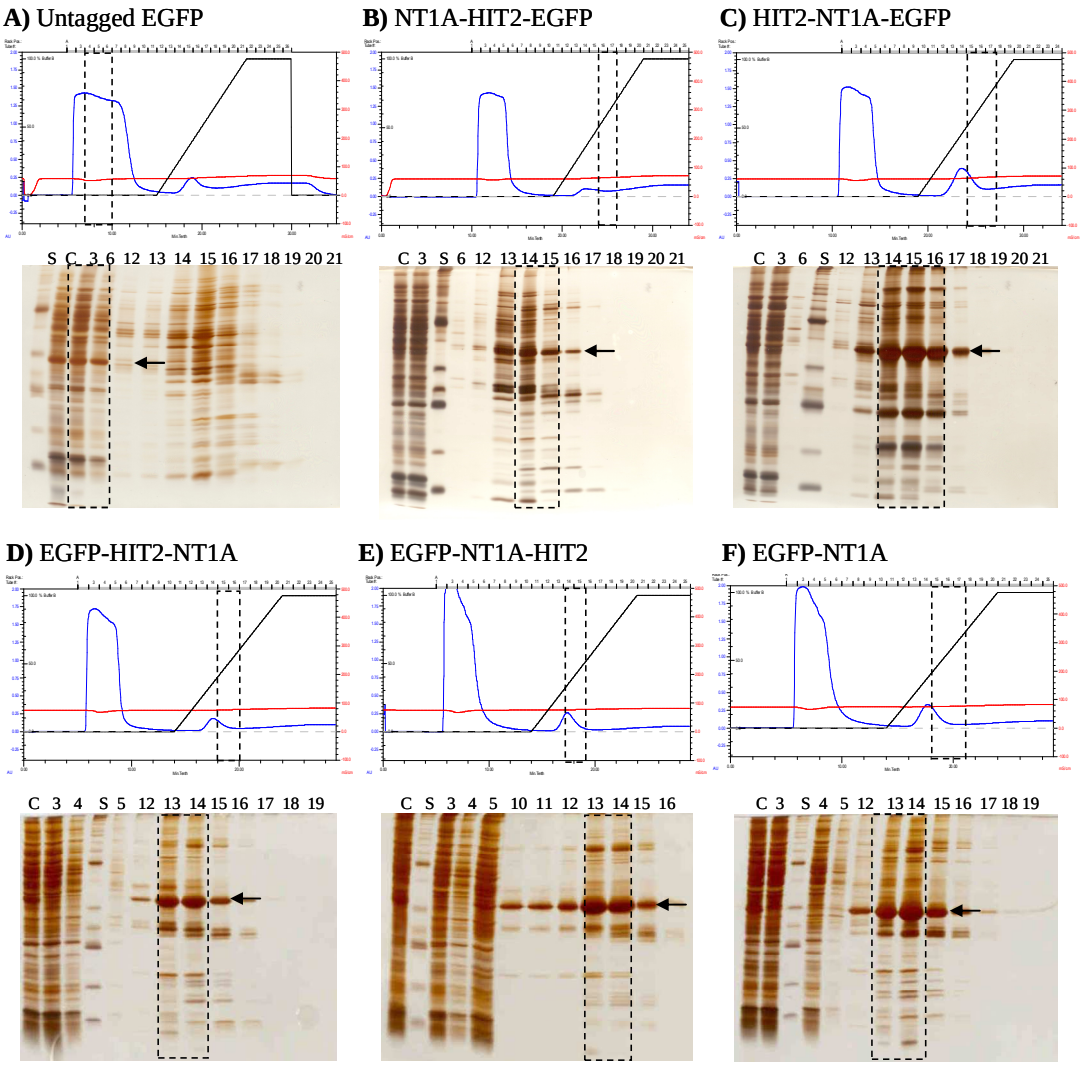


Figure 3.6. Chromatograms and 15 % SDS-PAGE of fractions from FPLC IMAC experiments on a 1 mL column packed with Ni^{2+}-NTA-Sepharose 6 F F. The flowrate was 1 mL min^{-1} throughout. Eluent A was 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Eluent B was 300 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Black trace: Per cent eluent B mixture. Blue trace: absorbance at 280 nm. Red trace: conductivity. 1 mL fractions were collected and analysed on SDS-PAGE. C: Starting material (cleared cell lysate) diluted 1/24; S: size marker. Flow through fractions 3 to 6 were diluted 1/16 with the exception of e) where fraction 5 was loaded undiluted. Elution fractions were loaded undiluted. 15 µL was loaded in each lane. EGFP bands are marked with an arrow. Visually green fractions are marked with a dashed box in the chromatogram and the SDS-PAGE.

Ni^{2+}-NTA has been shown to selectively bind neighbouring His residues[6; 31]. Furthermore, the probability of two neighbouring His residues to be in the correct orientation to interact with Ni^{2+}-NTA has been found to increase if five to six His residues are arranged in a sequence[96]. The weaker binding of NT1A tagged EGFP as compared to EGFP-His_6 can therefore be due to the difference in His density and content of the tags. The NT1A tag contains only four His residues as compared to six in the His_6 tag. Furthermore three of the His residues are consecutive and the fourth is located four residues downstream. The HIT2 tag contains one histidine residue which is not adjacent to any of the histidine residues in the NT1A tag in any of the double tagged proteins (Table 3.1). Despite the total five His residues in the fusion proteins containing both the NT1A and the HIT2 tag they showed identical chromatographic behavior as compared to EGFP-NT1A. This could also be due to the low density of the histidine residues in the double tags.

NT1A-HIT2-EGFP and HIT2-NT1A-EGFP eluted at an imidazole concentration of 100-160 mM and 100-180 mM respectively (Figure 3.6B, C). The elution of HIT2-NT1A-EGFP in a larger range of the gradient compared to NT1A-HIT2-EGFP could be due to the lower expression of NT1A-HIT2-EGFP (Section 3.3.3). EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A eluted at an imidazole concentration of 70-130 mM (Figure 3.6D, E). The N-terminally tagged fusion proteins were hence retained slightly more compared to their C-terminally tagged counterparts on Ni^{2+}-NTA-Agarose.

3.3.4.2 Ni^{2+}-tacn

As observed for Ni^{2+}-NTA (Section 3.3.4.1), His_6 tagged EGFP bound stronger than NT1A tagged EGFP to Ni^{2+}-tacn-Sepharose 6 F F (Figure 3.7B). The elution peak for EGFP-His_6 was broad. The front of the elution peak at 100 mM imidazole containined mainly *E. coli* host protein as judged by SDS-PAGE (Figure 3.7B). The later fractions at 130-210 mM imidazole contained EGFP-His_6 (Figure 3.7B). The EGFP proteins tagged with NT1A alone or with NT1A and HIT2 in sequence were not completely retained on the column. As evident from the SDS-PAGE gels (Figures 3.7D-H) NT1A-EGFP, EGFP-NT1A, HIT2-NT1A-EGFP, EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A were eluted from the column already in the first wash step prior to the gradient. The exception was NT1A-HIT2-EGFP (Figure 3.7C) where only two vaguely green fractions were observed at 100-160 mM imidazole. The lack of green wash fractions could be due to the lower expression of this protein (see Section 3.3.3) and might not reflect a different retention pattern on Ni^{2+}-tacn-Sepharose 6 F F. An imidazole concentration of 10 mM in 10 mM phosphate buffer, 650 mM NaCl pH 7.6 was hence sufficient to elute some EGFP tagged with NT1A alone or NT1A and HIT2 in sequence from Ni^{2+}-tacn-Sepharose 6 F F.

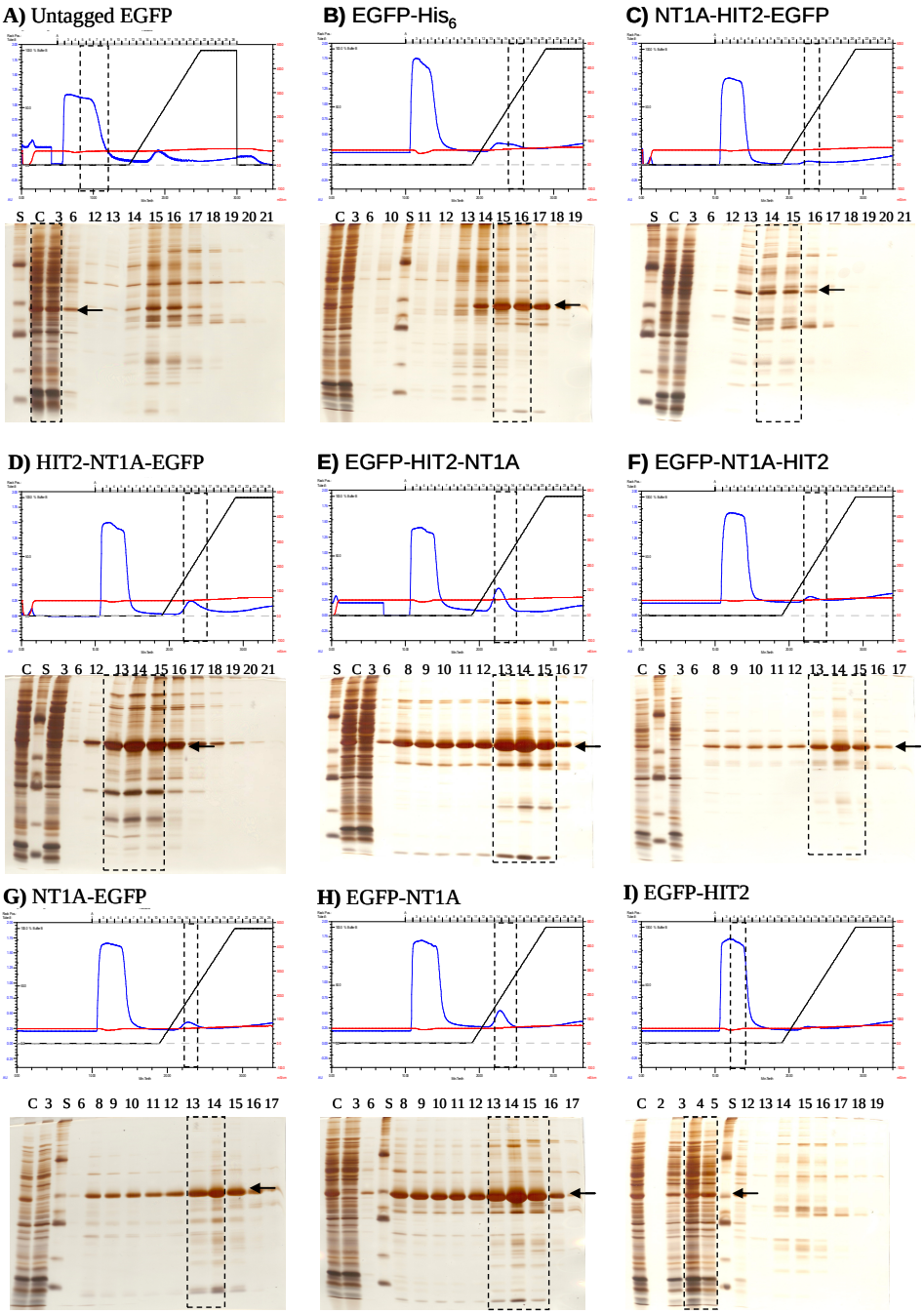


Figure 3.7. Chromatograms and 15 % SDS-PAGE of fractions from FPLC IMAC experiments on a 1 mL column packed with Ni^{2+}-tacn-Sepharose 6 F F. The flowrate was 1 mL min^{-1} throughout. Eluent A was 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Eluent B was 300 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Black trace: Per cent eluent B in mixture. Blue trace: absorbance at 280 nm. Red trace: conductivity. 1 mL fractions were collected and analysed on SDS-PAGE. C: Starting material (cleared cell lysate) diluted 1/24; S: size marker. Flow through fractions 3-6 were diluted 1/12. Elution fractions were loaded undiluted. 15 µL was loaded in each lane. EGFP bands are marked with an arrow. Visually green fractions are marked with a dashed box in the chromatogram and the SDS-PAGE.

A significant amount of EGFP tagged with NT1A alone or NT1A and HIT2 in sequence was however bound to the column during the wash step and was eluted in the gradient at an imidazole concentration range of 70-160 mM. This is the same range of imidazole concentrations as where NT1A tagged EGFP was found to elute from Ni^{2+}-NTA-Agarose (Section 3.3.4.1).

Untagged EGFP was not retained on Ni^{2+}-tacn-Sepharose 6 F F. This was evident from the fluorescence of the flow-through fractions and the lack of fluorescence of the elution fractions. The band corresponding to EGFP in flow through fraction 3 is marked with an arrow (Figure 3.7A). The band corresponding to EGFP-HIT2 in the SDS-PAGE is clearly visible in the flow through lanes 3 and 5 and also in elution fraction 12 as marked with an arrow in Figure 3.7I. Some EGFP-HIT2 was hence retained on Ni^{2+}-tacn-Sepharose 6 F F during the flow through and wash step but was eluted at a point in the gradient corresponding to 70 mM imidazole. EGFP tagged with NT1A alone (Figure 3.7G, H) was eluted at the same imidazole concentration as EGFP tagged with NT1A and HIT2 in sequence (Figure 3.7C-F). The HIT2 tag hence did not contribute significantly to the retention of HIT2 tagged EGFP on Ni^{2+}-tacn-Sepharose 6 F F under the experimental conditions employed.

NT1A-HIT2-EGFP and HIT2-NT1A-EGFP were eluted at 100-160 and 70-160 mM imidazole. EGFP-HIT2-NT1A and EGFP-NT1A-HIT2 were eluted at 70-160 mM imidazole. No significant difference in retention could hence be observed between the N- and C-terminally tagged fusion proteins on Ni^{2+}-tacn-Sepharose 6 F F.

3.3.4.3 Ni^{2+}-propyl-bis-tacn

All EGFP fusion proteins containing a borderline metal binding tag (NT1A or His_6) were retained on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. Baseline separation was achieved between the flow-through and elution peaks for both the N-terminally and C-terminally tagged fusion proteins (Figure 3.8). As an example the retention of HIT2-NT1A-EGFP is evident by the absence of the target protein in the lane corresponding to the flow-through fraction on SDS-PAGE (Figure 3.8D, lane 3 in SDS-PAGE).

EGFP-His_6 and NT1A-HIT2-EGFP (Figure 3.8B, C) were eluted at 160-240 mM imidazole. HIT2-NT1A-EGFP was eluted at 160-270 mM imidazole (Figure 3.8D). In contrast to Ni^{2+}- NTA-Agarose and Ni^{2+}-tacn-Sepharose 6 F F, EGFP-His_6 was not eluted at a higher imidazole concentration than the NT1A tagged proteins on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F.

EGFP-HIT2-NT1A and EGFP-NT1A-HIT2 (Figure 3.8E, F) were eluted at 130-240 and 100-210 mM imidazole, respectively. Hence EGFP fused with NT1A and HIT2 at the N-terminus were retained to a greater extent as compared to the corresponding C-terminally tagged proteins on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F.

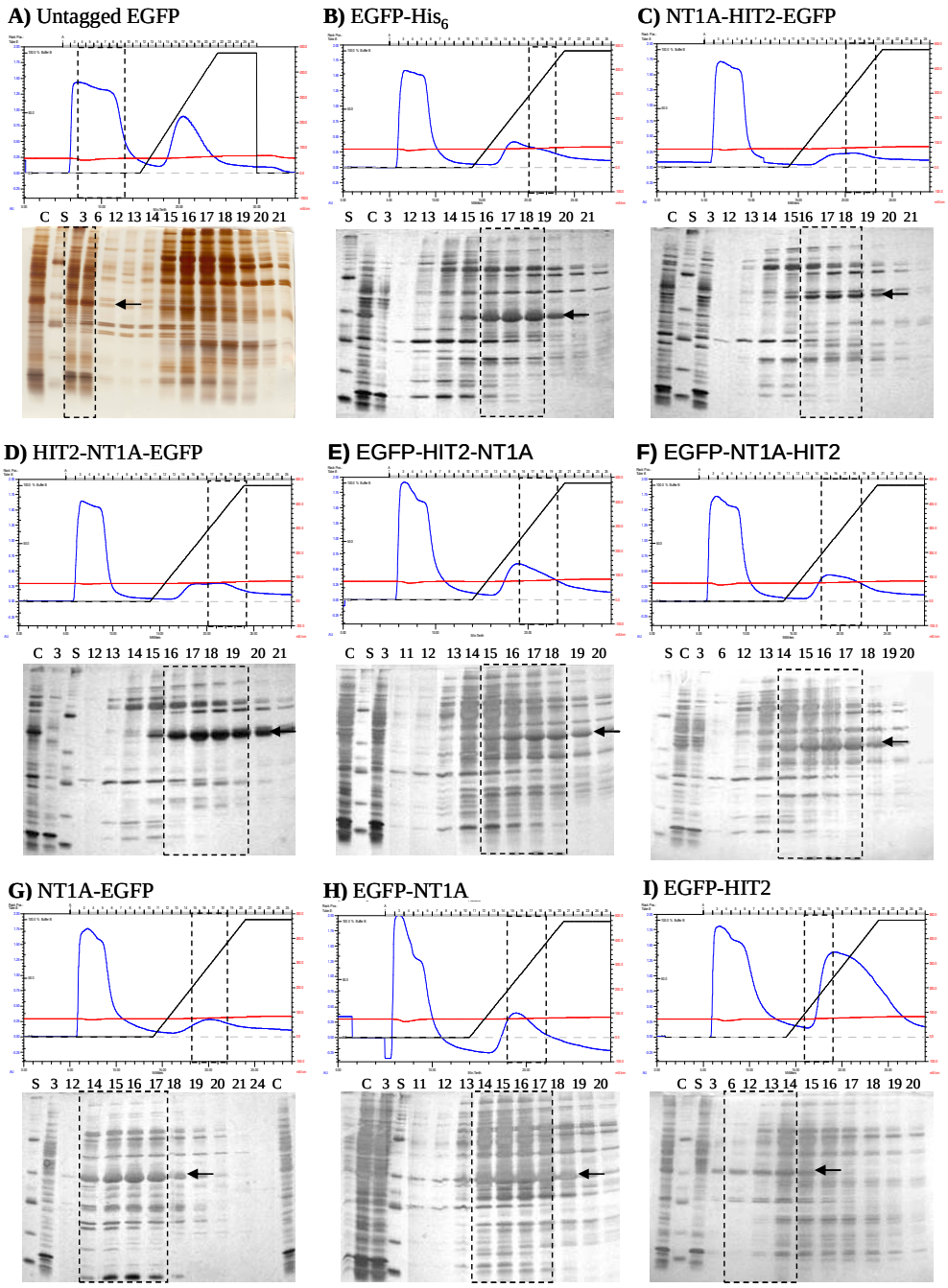


Figure 3.8. Chromatograms and 15 % SDS-PAGE of fractions from FPLC IMAC experiments on a 1 mL column packed with Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. The flowrate was 1 mL min^{-1} throughout. Eluent A was 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Eluent B was 300 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Black trace: Per cent eluent B in mixture. Blue trace: absorbance at 280 nm. Red trace: conductivity. 1 mL fractions were collected and analysed on SDS-PAGE. C: Starting material (cleared cell lysate) diluted 1 / 24; S: size marker. Flow through fractions 3-6 were diluted 1/16 with the exception of a) where fractions 3-6 were diluted 1/12. Elution fractions were loaded undiluted. 15 µL was loaded in each lane. EGFP bands are marked with an arrow. Visually green fractions are marked with a dashed box in the chromatogram and the SDS-PAGE.

The elution peaks were larger compared to the corresponding IMAC experiments with Ni^{2+}-NTA-Agarose (Figure 3.6) and Ni^{2+}-tacn-Sepharose 6 F F (Figure 3.7). The green elution fractions contained significant amounts of *E. coli* host protein as judged by SDS-PAGE (Figure 3.8B-F). The amount of *E. coli* host protein eluted in the imidazole gradient with untagged EGFP (Figure 3.8A) was also greater compared to the corresponding experiments with untagged EGFP on Ni^{2+}-NTA-Agarose (Figure 3.6A) and Ni^{2+}-tacn-Sepharose 6 F F (Figure 3.7A). It is therefore reasonable to assume that the larger elution peaks observed with Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F were due to greater retention of *E. coli* host protein.

Furthermore, the imidazole concentration (100-270 mM) at which the tagged EGFP proteins were eluted on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F was higher compared to 70-180 mM for Ni^{2+}-NTA-Agarose (Section 3.3.4.1) and Ni^{2+}-tacn-Sepharose 6 F F (Section 3.3.4.2).

In addition, EGFP-HIT2 was retained on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. This is evident from the EGFP-HIT2 bands in fractions 12-14 marked with an arrow in the SDS-PAGE (Figure 3.8I). These fractions correspond to an imidazole concentration of 40-130 mM. Moreover NT1A-EGFP and EGFP-NT1A were eluted at 100-210 mM imidazole (Figure 3.8G, H) while HIT2-NT1A-EGFP, NT1A-HIT2-EGFP and EGFP-HIT2-NT1A were eluted at 160-270, 160-240 and 130-240 mM imidazole respectively (Figure 3.8C-E).

Hence, Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F appeared to be more retentive as compared to Ni^{2+}-NTA-Agarose and Ni^{2+}-tacn-Sepharose 6 F F. With the exception of EGFP-NT1A-HIT2, proteins tagged with NT1A and HIT2 in sequence were more strongly retained as compared to proteins only tagged with NT1A.

3.3.4.4 Cu^{2+}-tacn

All EGFP proteins fused to a borderline metal binding tag (NT1A or His_6) were retained on Cu^{2+}-tacn-Sepharose 6 F F (Figure 3.9). EGFP-His_6 was eluted at 160-210 mM imidazole at the tail of the elution peak (Figure 9.3B). The purity of the elution fractions was high as judged by SDS-PAGE (Figure 9.3B, lanes 16-18).

In addition to the flow-through and elution peaks, a third peak was observed for EGFP-His_6, EGFP-NT1A and EGFP-HIT2 (Figure 3.9B, H, I). For EGFP-His_6 this peak was in the first wash step with 10 mM imidazole, 10 mM phosphate buffer, 650 mM NaCl, pH 7.6 prior to the gradient. Fraction 10 at the centre of the second peak had low content of protein as judged by SDS-PAGE (Figure 3.9B, lane 10). The molar extinction coefficient of the $Cu(NO_3)_2$ used herein was determined at 280 nm using water as blank and was found to be 11.5 M^{-1} cm^{-1}. The peak could hence represent loosely bound Cu^{2+} ions being washed off.

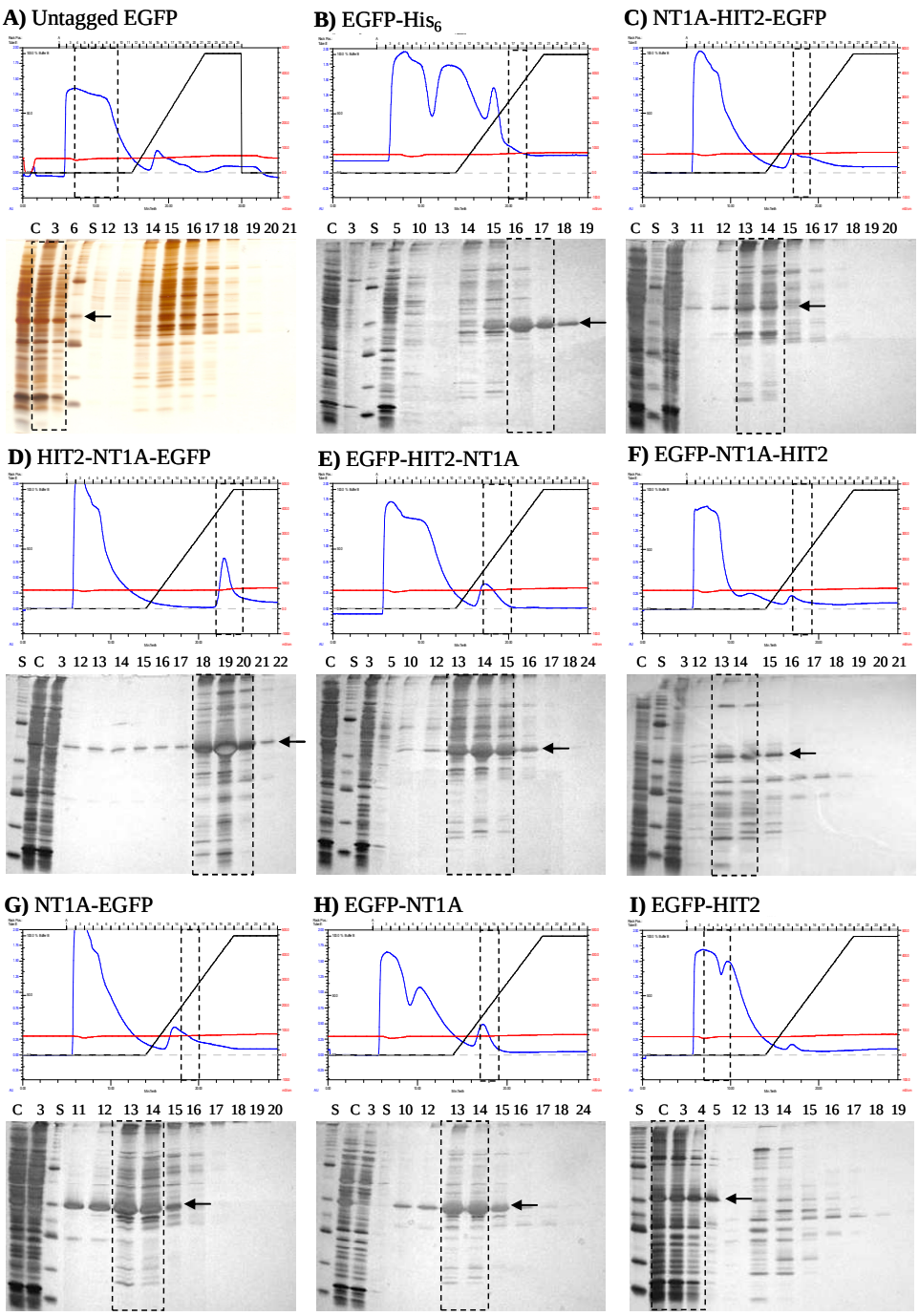


Figure 3.9. Chromatograms and 15% SDS-PAGE of fractions from FPLC IMAC experiments on a 1 mL column packed with Cu^{2+}-tacn-Sepharose 6 F F. The flowrate was 1 mL min^{-1} throughout. Eluent A was 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Eluent B was 300 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Black trace: Per cent eluent B in mixture. Blue trace: absorbance at 280 nm. Red trace: conductivity. 1 mL fractions were collected and analysed on SDS-PAGE. C: Starting material (cleared cell lysate) diluted 1/24; S: size marker. Flow through fractions 3-6 were diluted 1/16 with the exception of a) where fractions 3-6 were diluted 1/12. Elution fractions were loaded undiluted. 15 µL was loaded in each lane. EGFP bands are marked with an arrow. Visually green fractions are marked with a dashed box in the chromatogram and the SDS-PAGE.

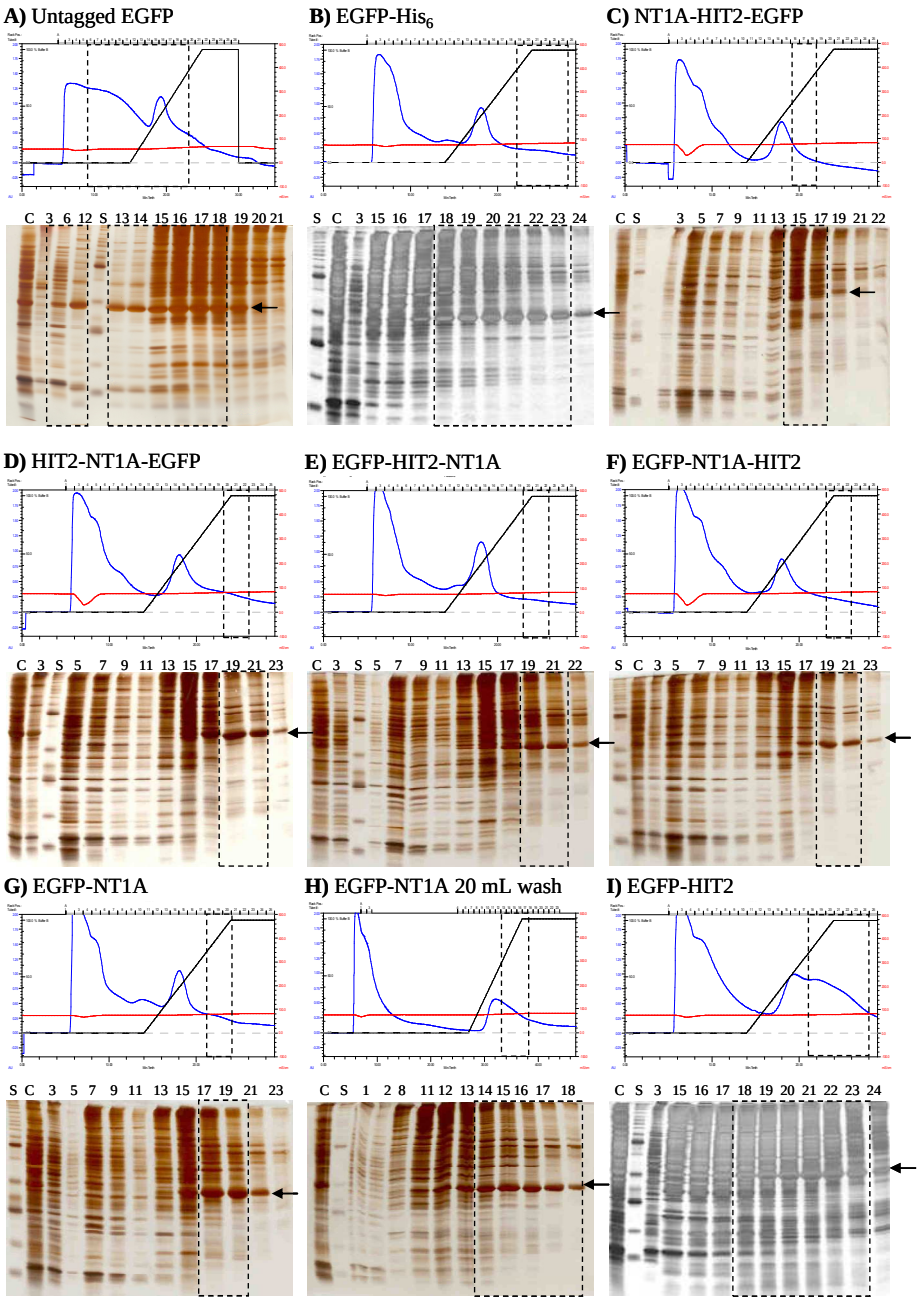


Figure 3.10. Chromatograms and 15% SDS-PAGE of fractions from FPLC IMAC experiments on a 1 mL column packed with Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. The flowrate was 1 mL min^{-1} throughout. Eluent A was 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Eluent B was 300 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Black trace: Per cent eluent B in mixture. Blue trace: absorbance at 280 nm. Red trace: conductivity. 1 mL fractions were collected and analysed on SDS-PAGE. C: Starting material (cleared cell lysate) diluted 1/24; S: size marker. Flow through fractions 3-6 were diluted 1/16 with the exception of a) where fractions 3-6 were diluted 1/12. Elution fractions were loaded undiluted. 15 µL was loaded in each lane. EGFP bands are marked with an arrow. Visually green fractions are marked with a dashed box in the chromatogram and the SDS-PAGE.

EGFP-NT1A, NT1A-EGFP, EGFP-NT1A-HIT2 and NT1A-HIT2-EGFP were eluted at 70-130 mM imidazole (Figure 9.3C, F-H). EGFP-HIT2-NT1A was eluted at 70-160 mM imidazole (Figure 9.3E). These imidazole concentrations were in the same range as observed for Ni^{2+}-NTA-Agarose (Section 3.3.4.1) and Ni^{2+}-tacn-Sepharose 6 F F (Section 3.3.4.2).

HIT2-NT1A-EGFP was eluted at 210-300 mM imidazole with fractions 18-20 containing the protein as judged by SDS-PAGE (Figure 3.9D). Untagged EGFP and EGFP-HIT2 were not retained on Cu^{2+}-tacn-Sepharose 6 F F but were eluted in the flow-through fractions. This is evident by the flow-through fractions containing the proteins and by the absence of the proteins in the elution fractions as judged by SDS-PAGE (Figure 3.9A, I).

3.3.4.5 Cu^{2+}-propyl-bis-tacn

Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F was the only immobilized metal chelate for which no baseline separation was achieved between the flow-through and elution peak for any protein investigated (Figure 3.10). All EGFP proteins tagged with a borderline metal binding tag (NT1A or His_6) were retained on Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F and were eluted at 180-300 mM imidazole in the gradient. The elution fractions were rich in *E. coli* host proteins as judged by SDS-PAGE (Figure 3.10). Retention was also observed for untagged EGFP. This is evident as flow through fraction 3 does not contain EGFP as judged by SDS-PAGE (Figure 3.10A, lane 3). EGFP was eluted gradually up to an imidazole concentration of 270 mM (Figure 3.10A, lanes 12-19). EGFP-HIT2 was also retained (Figure 3.10I). To achieve higher protein purity the first wash step was increased from 7 mL to 20 mL 10 mM imidazole, 10 mM phosphate buffer, 650 mM NaCl, pH 7.6. A 20 mL wash step gave near baseline separation for EGFP-NT1A (Figure 3.10H). The purity of the elution fractions containing EGFP-NT1A was however not significantly improved by the increased length of the wash step as seen in Figures 3.10G and H.

3.4 Discussion

3.4.1 Differences in expression levels between N- and C-terminally tagged EGFP

EGFP expressed in *E. coli* is known to form inclusion bodies containing both fluorescent protein[130] and improperly folded non fluorescent protein[131]. It is therefore possible that the lower abundance of NT1A-HIT2-EGFP in the cell lysate reflected a greater tendency of this protein to form inclusion bodies. This aspect was however not investigated in this thesis.

The lower content of NT1A-HIT2-EGFP in the cell lysate could also reflect a lower expression level of this fusion protein as compared to untagged EGFP, HIT2-NT1A-EGFP, EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A. Low expression levels of recombinant proteins in *E. coli* can be due to a number of factors including mRNA stability, secondary structure formation near the 5' end of the mRNA, rare codon usage and a weak Shine Dalgarno sequence[132]. The Shine Dalgarno sequence was identical for all constructs. The difference in expression levels could hence not be related to the Shine Dalgarno sequences of the constructs. The 5' end of the mRNA is coding for the N-terminus of the protein[133]. The 5' end of the mRNA molecules coding for NT1A-HIT2-EGFP and HIT2-NT1A-EGFP would therefore differ from the mRNA's coding for EGFP, EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A. A low stability or secondary structure formation near the 5' end of the mRNA coding for NT1A-HIT2-EGFP could hence explain the low expression of this fusion protein.

3.4.2 Trends of IMAC experiments

The results of the IMAC experiments in Section 3.3.4 were summarized by assessing the imidazole concentrations at which the EGFP proteins were eluted. The assessments were performed by visually inspecting the SDS-PAGE gels in Figures 3.5 – 3.10. The summary of these assessments is shown in Table 3.4. As expected, all recombinant proteins fused to a borderline metal ion binding tag (NT1A or His_6) were retained on all the immobilized metal chelates investigated during the loading and wash step.

Variations in retention between N- and C-terminally tagged fusion proteins were observed. NT1A-HIT2-EGFP and HIT2-NT1A-EGFP were eluted at equal or higher imidazole concentrations as compared to EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A (Table 3.4). The difference was most pronounced for HIT2-NT1A-EGFP on Cu^{2+}-tacn-Sepharose 6 F F (Figure 3.9D). Trends could also be observed for the elution of the recombinant proteins between the different immobilized metal chelates. Ni^{2+}-NTA-Agarose, Ni^{2+}-tacn-Sepharose 6 F F and Cu^{2+}-tacn-Sepharose 6 F F were less retentive than Ni^{+}-propyl-*bis*-tacn-Sepharose 6 F F and Cu^{+}-propyl-*bis*-tacn-Sepharose 6 F F. With the exception of HIT2-NT1A-EGFP on Cu^{2+}-tacn-Sepharose 6 F F, all NT1A and His_6 tagged proteins were eluted between 70 and 180 mM imidazole on these immobilized metal chelates (Table 3.4). Untagged EGFP and EGFP-HIT2 were not retained. All NT1A and His_6 tagged proteins were eluted at an imidazole concentration range of 100-270 mM on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. In addition, EGFP-HIT2 was retained (Table 3.4). For Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F all proteins including untagged EGFP were retained. NT1A and His_6 tagged proteins were eluted at an imidazole concentration range of 180-300 mM (Table 3.4).

Table 3.4. Overview of FPLC IMAC experiments of *E. coli* cell lysate containing recombinant EGFP fusion proteins on 1 mL columns packed with immobilized Ni^{2+} and Cu^{2+} chelates of tacn and propyl-*bis*-tacn immobilized on Sepharose 6 F F. *E. coli* cell lysate was loaded in 10 mM imidazole, 10 mM phosphate buffer, 0.65 M NaCl, pH 7.6. Elution was achieved with a linear gradient from 10 to 300 mM imidazole. The imidazole concentration range in mM at which the EGFP fusion proteins were eluted is shown.

Metal chelate / **Fusion protein**	**Ni^{2+}-NTA-Agarose**	**Ni^{2+}-tacn**	**Ni^{2+}-propyl-*bis*-tacn**	**Cu^{2+}-tacn**	**Cu^{2+}-propyl-*bis*-tacn**
EGFP-His_6	130-270	130-210	160-240	160-210	210-300
EGFP-NT1A	70-160	70-160	100-210	70-130	180-270
EGFP-NT1A-HIT2	70-130	70-160	100-210	70-130	210-300
EGFP-HIT2-NT1A	70-130	70-160	130-240	70-160	210-300
NT1A-EGFP	Not determined	70-160	100-210	70-130	Not determined
HIT2-NT1A-EGFP	100-180	70-160	160-270	210-300	210-300
NT1A-HIT2-EGFP	100-160	100-160	160-240	70-130	130-210
EGFP-HIT2	Not retained	Not retained	10-130	Not retained	210-300
EGFP	Not retained	Not retained	Not retained	Not retained	10-270

Moreover the elution fractions obtained with Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F and Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F were richer in *E. coli* host proteins compared to Ni^{2+}-NTA-Agarose, Ni^{2+}-tacn-Sepharose 6 F F and Cu^{2+}-tacn-Sepharose 6 F F as judged by SDS-PAGE (Figure 3.5-3.10). The more retentive properties of Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F as compared to Cu^{2+}-tacn-Sepharose 6 F F found here are in agreement with the results reported by Graham, Spiccia and Hearn[46]. These investigators found that immobilized Cu^{2+} complexes of xylylene bridged *bis*-tacn ligands were more retentive as compared to immobilized Cu^{2+}-tacn.

Hydrophobic interactions have been proposed as a mechanism of *E. coli* host protein retention on the actual base resin in IMAC resins such as[134] Ni^{2+}-NTA-Sepharose™. The more retentive properties of the propyl-*bis*-tacn chelates compared to the tacn chelates could hence be due to the hydrophobicity of the propyl bridge.

The imidazole concentration at which the EGFP fusion proteins were eluted was not significantly different between Ni^{2+}-tacn-Sepharose 6 F F and Cu^{2+}-tacn-Sepharose 6 F F but

was higher for Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F as compared to Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F (Table 3.4). Hence the metal ion affected the retention behaviour for immobilized propyl-*bis*-tacn but not for immobilized tacn. The retention of proteins on immobilized IDA chelates followed the trend[6] $Cu^{2+} > Ni^{2+} > Zn^{2+} \leq Co^{2+}$. A single histidine residue[6] can be sufficient to retain a protein on Cu^{2+}-IDA. This is in agreement with the trend observed for immobilized propyl-*bis*-tacn. Even untagged EGFP containing one surface accessible histidine is retained on Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F but not on Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F.

For the various metal chelates investigated, the elution fractions of the highest purity were obtained for EGFP-His_6 on Cu^{2+}-tacn-Sepharose 6 F F and Ni^{2+}-NTA-Agarose as judged by SDS-PAGE (Figures 3.5 and 3.9B). For these immobilized metal chelates different retention patterns were observed for the NT1A and His_6 tagged proteins (Figures 3.5, 3.6 and 3.9). It therefore appears that the mononuclear chelates required a higher density of His residues in the tag. No such difference was observed for Ni^{2+}-propyl-*bis*-tacn and Cu^{2+}-propyl-*bis*-tacn. This could be because the two His sites in the NT1A tag can interact with the two metal centres provided by the ligand. The distance between the two metal centres in Cu^{2+}-propyl-*bis*-tacn has been determined to be[42] 8.6 Å. It is hence reasonable to assume that two His residues in both the His_6 and NT1A tag could interact simultaneously with the two metal centres.

The purification achieved with the metal chelates investigated was Cu^{2+}-tacn > Ni^{2+}-tacn > Ni^{2+}-propyl-*bis*-tacn > Cu^{2+}-propyl-*bis*-tacn, under the conditions used. Different results could be obtained by varying buffer components and/or other chromatographic parameters such as flowrate, mobile phase gradients and temperature.

3.5 Conclusions

In conclusion, the expression level of soluble NT1A-HIT2-EGFP was ~ 30 % compared to that of HIT2-NT1A-EGFP in *E. coli* BL-21. No significant difference in expression levels was however observed for the C-terminally tagged counterparts EGFP-NT1A-HIT2 and EGFP-HIT2-NT1A. The C-terminally tagged fusion proteins were retained on the immobilized metal chelates Ni^{2+}-NTA-Agarose, Ni^{2+}-tacn-Sepharose 6 F F, Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F, Cu^{2+}-tacn-Sepharose 6 F F and Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F. The C-terminally tagged fusion proteins were however eluted at slightly lower imidazole concentrations compared to their N-terminal counterparts.

Chapter 4. Effects of affinity tags on protein stability

4.1 Introduction

4.1.1 Affinity tags and protein stability

An important aspect in IMAC is the effect of the affinity tag on their fusion partner. Ideally, affinity tags should not affect properties such as structure and stability of the proteins to which they are fused. All affinity tags can however potentially affect their fusion partners[8; 11; 48]. Numerous examples when this occurs have been reported. The hexa His (His_6) tag has been shown to produce one active and one inactive form of L-lactate dehydrogenase when fused to the C-terminus of the protein. This is possibly due to misfolding of the protein[124]. The His_6 tag also altered the structure of the DNA binding protein AreA[135]. Furthermore, His_6 tagged hydantoinase was shown to be inactive when eluted from TALON® either by decreased pH or by imidazole[125]. Insertion of a His_6 tag at the C-terminus of the so-called 33 kDa protein subunit of higher plant photosystem II was reported to cause a conformational change of the protein[126]. When fused to aminopeptidase B, the His_6 tag was shown to partially inhibit the exopeptidase activity of the enzyme[136]. The activity of the trimeric cytokine tumor necrosis factor alpha decreased by 16 % and 7 % when a His_{10} and a His_7 tag respectively was fused to the N-terminus of the protein[137]. On the other hand, affinity tags have also been reported to affect proteins advantageously by increasing expression, solubility and stability[8; 11; 49; 138] and to promote proper folding[48; 138].

Hence the effect of tags on their fusion partner is an important aspect to consider when developing novel affinity tags. The effect of the novel affinity tags NT1A and HIT2 on the structure and stability of their fusion partner EGFP was therefore studied in this Chapter.

4.1.2 GFP's

The green fluorescent protein (GFP) is a widely studied and exploited[139] fluorescent protein used for an ever increasing number of applications[140]. Examples include its use as a reporter for gene expression and localization[141], as an *in situ* tag for fusion proteins[142], as a biosensor[143] and as a probe for protein-protein interactions[144]. The importance of GFP in various fields was demonstrated by the award of the Nobel Prize in Chemistry 2008 to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien for the discovery and development of the

protein. Shimumora *et al*. isolated the protein from the jelly fish *Aequorea victoria* in 1962 and demonstrated that it could fluoresce when irradiated with ultraviolet light[145]. The gene coding for GFP was cloned in 1992 by Prascher *et al*.[146] and the protein was expressed in *E. coli* by Chalfie *et al*.[147]

The structure and function of GFP have been extensively studied in several laboratories including that of Roger Tsien[140]. Wild type GFP (GFP_{wt}) is a 238 residue, 27 kDa protein forming an 11 stranded ß-barrel wrapped around one α-helix[148; 149] (Figure 4.1). The fluorophore, *p*-hydroxybenzylideneimidazolinone, is formed[148] auto catalytically by cyclization and oxidation of residues Ser65, Tyr66 and Gly67. The fluorophore presumably exists in one predominant neutral and one minor ionized form causing two absorption maxima at 395 and 475 nm respectively with emission of green light at 508 nm[150]. Hence the ß-barrel and two capping α-helical segments protect the fluorophore from the surrounding solvent. When the tertiary structure of GFP is altered the fluorophore is quenched by the surrounding solvent molecules[151].

Many mutants of GFP have been developed in order to increase its fluorescence and expression levels in different expression systems. Mutating Ser65 to Thr (S65T) simplifies the absorption spectra to a single maximum at 489 nm[152] since the extra methyl group in Thr stabilizes a hydrogen bond network which renders the fluorophore permanently ionized[150]. The GFP used in this work, enhanced GFP (EGFP), originally contained within the cloning cassette pGreenTIR[153], has been optimized for prokaryotic expression systems in my supervisor's laboratory. EGFP contains three mutations from GFP_{wt}; S65T as previously mentioned, F64L which increases protein folding and solubility[154] and Q80R which was originally introduced by accident but is known not to alter the protein properties[148]. The intensity of the fluorescence of EGFP is 30 times higher than for the wild type protein.

Another GFP mutant is GFP_{UV} (also referred to as the cycle3 mutant) containing the mutatations F99S, M153T and V163A. The fluorescence signal of cells expressing GFP_{UV} was 45 times greater[155] compared to the GFP_{wt} expressed from the Clontech plasmid pGFP (http://www.clontech.com/products/detail.asp?product_id=10441&tabno=2 on 2010-06-16). The ability of GFP_{UV} to mature more efficiently in vivo at 37 ºC is thought to be due to its surface being less hydrophobic compared to GFP_{wt}, thus preventing aggregation[156].

Other fluorescent proteins (FP's) emitting red[157; 158], yellow[159; 160], blue[161] and cyan[162] light as well as FP's with increased folding properties[163] have also been developed. The list of new variants of fluorescent proteins is constantly increasing[140].

GFP has been shown to dimerize at high protein concentrations and high ionic strength buffers[164]. At protein concentrations above 5 mg mL^{-1} bands corresponding to the monomeric and dimeric protein were observed on SDS-PAGE[103]. Only the monomer was

detected at concentrations[103] below 1 mg mL^{-1}. EGFP is also known to dimerize at high concentrations[140]. This was also observed in Chapter 3, Figure 3.2.

The stability of various GFP's has been extensively studied. His_6 tagged EGFP was thermally and chemically denatured and was found to unfold over the time course of days at elevated temperatures and at relatively high (> 2 M) guanidine hydrochloride (GuHCl) concentrations[165; 166]. Several studies imply that the folding and unfolding of GFP are multi-step processes involving several intermediate states. Mechanically induced unfolding of GFP_{UV} with an additional point mutation, S2G, was found to involve two intermediate states; one with the internal α-helix detached and the second with one of the strands detached from the ß-barrel[167]. Acid denaturing studies of GFP_{UV} revealed the presence of an equilibrium unfolding intermediate at pH 4, characterized by a molten globular state lacking fluorophore fluorescence. In addition, two kinetic folding intermediates were observed in the refolding process[151; 168]. Huang *et al.*[169] found an unfolding intermediate state during the chemical (GuHCl) denaturation of GFP_{UV}. This intermediate was also compact with preserved secondary and tertiary structure but with ß-strands 7-10 displaced, as evident from hydrogen/deuterium nuclear magnetic resonance (NMR) exchange rate data of amide protons and with reduced fluorescence. Equilibrium was however not reached even after incubations at 37° C for eight weeks[169] and the protein kept losing fluorescence at GuHCl concentrations below 4 M. The same intermediate state was observed by Xie *et al.*[170] The authors used a constant incubation time of 24 h for GuHCl induced equilibrium unfolding experiments of GFP_{UV} in the presence and absence of *E. coli* chaperone "trigger factor" (TF)[170].

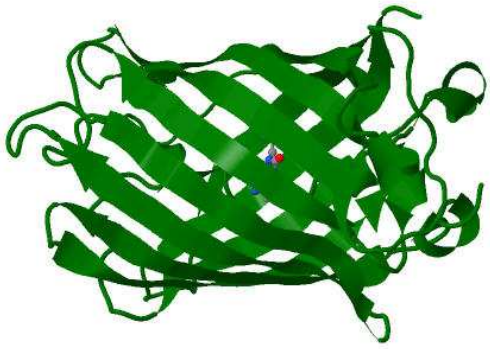


Figure 4.1. Cartoon representation of the green fluorescent protein GFP, pdb code: 1emb[150]. The fluorophore is displayed as "balls and sticks".

Due to the extensive knowledge obtained about the properties of fluorescent proteins, and the fact that the fluorophore fluorescence is a sensitive probe of the tertiary structure of GFP's[151], EGFP is well suited as a model for studying the effects of affinity tags on protein structure and stability.

4.2 Aim

The aim of the studies described in this Chapter was to investigate if the novel affinity tags NT1A and HIT2 would affect the structure and stability of the target protein to which they were fused. EGFP was used as a model protein since its strong fluorescence serves as a sensitive probe for the tertiary structure of the protein. As mentioned above, the stability of EGFP has been extensively studied and it is therefore a suitable model protein. The prescence of an affinity tag at the N- or C terminus of the target protein can potentially affect the target protein to different extents. N- and C-terminally tagged proteins were therefore compared. EGFP with the His_6 attached to the C-terminus, EGFP with NT1A attached to the N- and C-terminus and EGFP with HIT2 attached to the C-terminus was compared with untagged EGFP.

4.3 Results

4.3.1 Protein Purification

4.3.1.1 Purification of EGFP-His_6, NT1A-EGFP and EGFP-NT1A

The proteins with a borderline metal binding tag (EGFP-His_6, NT1A-EGFP and EGFP-NT1A) were purified on an Äkta XPress system as described in detail in the Materials and Methods Sections 2.4.8 and 2.4.9. All steps A-C were performed in sequence and the purity of the sample could therefore not be assessed after the individual steps.

IMAC using a 1 mL HisTrap™ F F column. *E. coli* BL-21 cell lysate containing EGFP-His_6 was loaded onto the column in 30 mM imidazole, 350 mM NaCl, 1 x PBS, pH 7.6 and eluted in 500 mM imidazole, 350 mM NaCl, 1 x PBS, pH 7.6. *E. coli* BL-21 cell lysate containing NT1A-EGFP or EGFP-NT1A was loaded in 10 mM imidazole, 350 mM NaCl, 1 x PBS, pH 7.6 and eluted in 500 mM imidazole, 350 mM NaCl, 1 x PBS, pH 7.6.

Buffer exchange. The eluted protein was buffer exchanged into 20 mM Tris-HCl, pH 8.0 by SEC.

High resolution IEX on a MonoQ™ 5/50 column. The protein was loaded in 20 mM Tris-HCl and eluted by a two step linear gradient from 20 mM Tris-HCl, pH 8.0 to 1 M NaCl, 20 mM Tris-HCl, pH 8.0.

EGFP-His_6 was strongly bound to the HisTrap™ column in the first chromatography step where the bulk of the *E. coli* host protein was washed off (Figure 4.2A). Some contaminating proteins were however co-eluted from the HisTrap™ column and removed by the final IEX step as can be seen in the chromatogram (Figure 4.2A) and the SDS-PAGE gel (Figure 4.2C, lane 3).

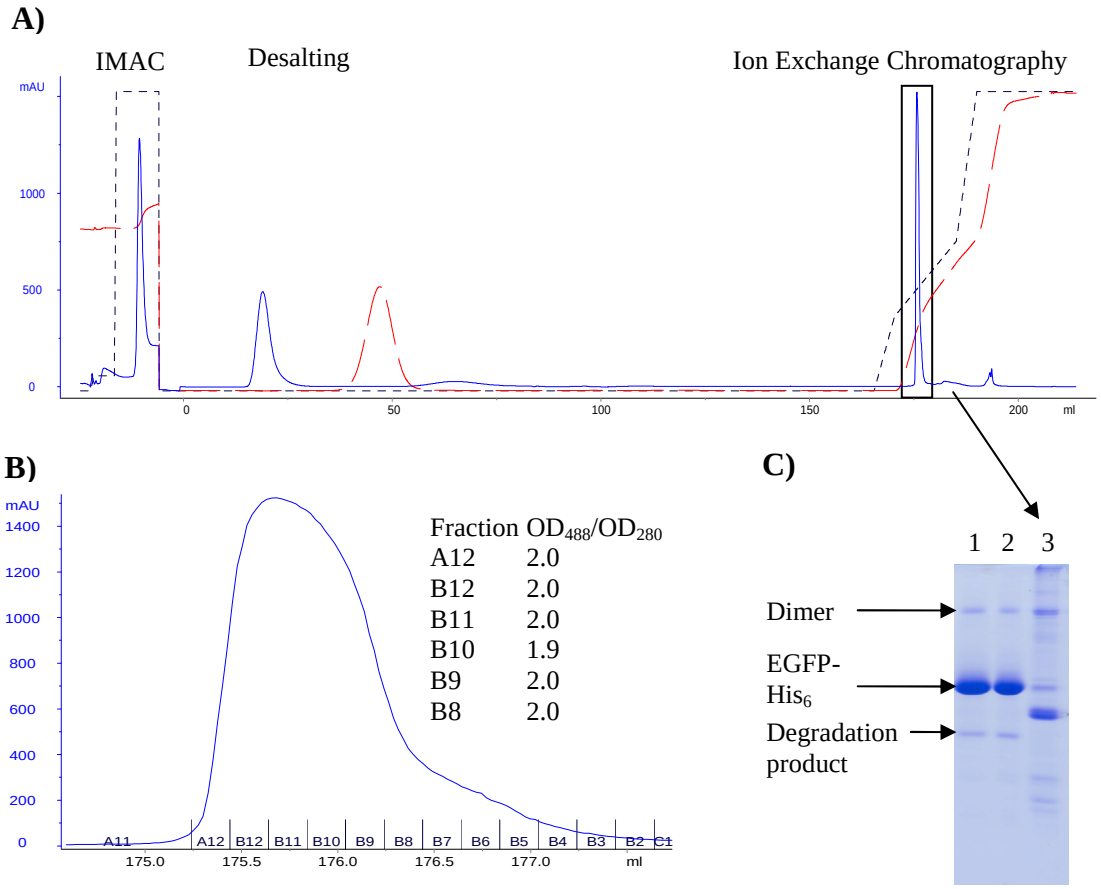


Figure 4.2A) Chromatogram representing a three step purification of EGFP-His_6 on an Äkta Xpress system. Solid trace: Absorbance at 280 nm. Dashed trace: Conductivity. Dotted trace: Gradient. **B)** Enlargement of final IEX elution step (boxed in top chromatogram) with fraction purity assessed by OD_{488}/OD_{280}. **C)** 15% SDS-PAGE of IEX elution fractions. Lane 1: pool of fractions B12-B9, lane 2: pool of fractions B8-B7, lane 3: contaminating proteins marked with arrow in chromatogram.

Xie *et al.*[170] assessed the purity of GFP_{UV} by comparing the absorbance of samples containing the protein at 397 nm and at 280 nm. The authors used a ratio of OD_{397}/OD_{280} > 1.1 as a threshold when assessing the purity of the protein[170] using this method. A similar method was used in this thesis for assessing the purity of EGFP. The absorbance maximum of EGFP (488 nm) was used instead of 397 nm (the absorbance maximum for GFP_{wt} and GFP_{UV}). Hence the ratio OD_{488}/OD_{280} was used in addition to SDS-PAGE to assess the purity of EGFP in samples. The elution fractions in the IEX elution peak were assessed by OD_{488}/OD_{280} and those with the highest ratios were analysed on SDS-PAGE. In Figure 4.2C, lane 1 represents pooled fractions from the centre of the elution peak (Figure 4.2B, fractions B12-B9) and lane 2 represents elution fractions from the side of the peak (Figure 4.2B, fractions B8 and B7).

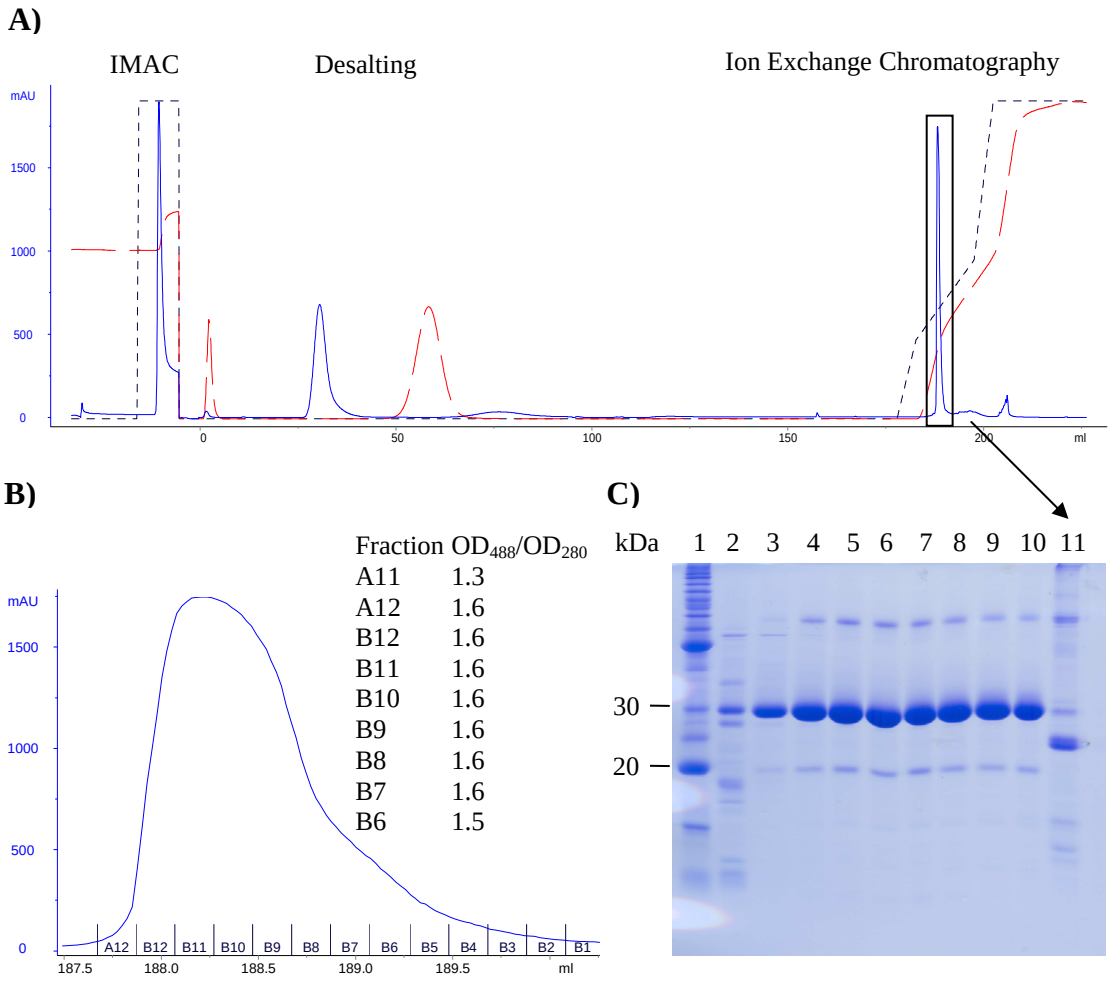


Figure 4.3A) Chromatogram representing a three step purification of NT1A-EGFP on an Äkta Xpress system. Solid trace: Absorbance at 280 nm. Dashed trace: Conductivity. Dotted trace: Gradient. **B)** Enlargement of final IEX elution step (boxed in top chromatogram). **C)** 15 % SDS-PAGE of ion exchange elution fractions. Lane 1: size marker, lanes 2-3: fractions A11-A12, lanes 4-10: fractions B12-B6, lane 11: Contaminating proteins (pool of fractions C7-C10) marked with arrow in chromatogram.

Both pools only contain fractions with OD_{488}/OD_{280} values 2.0 and appear to be of the same purity as judged by the SDS-PAGE (Figure 4.2C, lanes 1 and 2). Apart from the main band corresponding to EGFP-His_6 monomer, the dimer of the protein and a degradation product can also be observed. The degradation product was not successfully removed by any chromatographic methods and was present in all elution fractions, also in the presence of protease inhibitors.

Due to the lower content of His residues in the NT1A tag compared to the His_6 tag, NT1A-EGFP and EGFP-NT1A were loaded in a buffer containing 10 mM instead of 30 mM

imidazole. Under this condition, the purification profiles, as exemplified in Figure 4.3A, were nearly identical to EGFP-His_6 (Figure 4.2A). The OD_{488}/OD_{280} ratio for the elution fractions in the centre of the elution peak in the final IEX step (Figure 4.3B, fractions B11 – B9) was typically 1.6 for both NT1A-EGFP and EGFP-NT1A. The purity of the IEX elution fractions as judged by SDS-PAGE (Figure 4.3C, lanes 4-10) was similar as for EGFP-His_6 (Figure 4.2C, lanes 1 and 2). As observed for EGFP-His_6, some *E. coli* host proteins co-eluted with the NT1A tagged EGFP proteins in the IMAC step and were removed by the IEX step as seen in the chromatogram (Figure 4.3A) and in the SDS-PAGE (Figure 4.3C, lane 11).

4.3.1.2 Purification of untagged EGFP and EGFP-HIT2

Several strategies have been employed to purify untagged GFP's. A few examples are summarized in Figure 4.4. All are multistep procedures employing conventional chromatography procedures such as hydrophobic interaction chromatography (HIC), ion exchange chromatography (IEX) and size exclusion chromatography (SEC) in different orders. Xie *et al.*[170] used $(NH_4)_2SO_4$ precipitation followed by HIC on a Butyl-Sepharose™ column, SEC on a Sephadex G-75 column followed by IEX on DEAE-Sepharose™ (Figure 4.4A). Huang *et al.*[169] employed IEX on Q-Sepharose™ followed by SEC on a Superdex G-75 column. Fukuda *et al.*[156] used $(NH_4)_2SO_4$ precipitation, SEC on a Sephadex G-75 column, IEX on a DEAE-Sepharose™ column followed by a second SEC step on a Sephadex G-75 column. Yang *et al.*[149] purified GFP using IEX on a DEAE-Sepharose™ column followed by HIC, high performance liquid chromatography (HPLC) IEX on a Bio-Gel DEAE-5PW column followed by HPLC SEC on a BioGel SEC-125 column. Organic extraction followed by HIC has also been reported as a technique to purify fluorescent proteins such as EGFP[171; 172]. The organic extraction protocols involve GFP being dissolved in ethanol. They were avoided in this Chapter due to concerns of potentially destabilizing effects of the protein. This could potentially introduce an error in the unfolding studies.

The strategy employed to purify untagged EGFP and EGFP-HIT2 in this thesis (Figure 4.4F) was chosen based on a number of practical considerations. $(NH_4)_2SO_4$ precipitation in 1.2 M $(NH_4)_2SO_4$ was employed as a rapid primary sample preparation step prior to HIC since the loading buffer in the HIC step contains 1.2 M $(NH_4)_2SO_4$ and buffer exchange therefore could be avoided. The 20 mL Phenyl-Sepharose 6 F F (Low Sub) column employed in the HIC step had high capacity. Hence this step was suitable early in the purification when the total protein content of the sample was high. IEX on a 5 mL Q-Sepharose 6 F F column was chosen as the next step because the loading buffer (20 mM Tris-HCl, pH 8.0) was used as the elution buffer in the previous HIC step. Dilution of the sample in loading buffer was hence sufficient to lower the concentration of any remaining $(NH_4)_2SO_4$ to a level allowing binding of EGFP to the Q-Sepharose 6 F F column. A high resolution IEX

step on a MonoQ™ 5/50 column was introduced as the next step to remove any low abundance contaminants. Even if this step required buffer exchange into 20 mM Tris-HCl it was deemed necessary to achieve the desired level of protein purity. SEC on a HiLoad™ Superdex™ 75 prep grade 16/60 column was a suitable last step based on two considerations. Firstly the capacity of SEC is low.

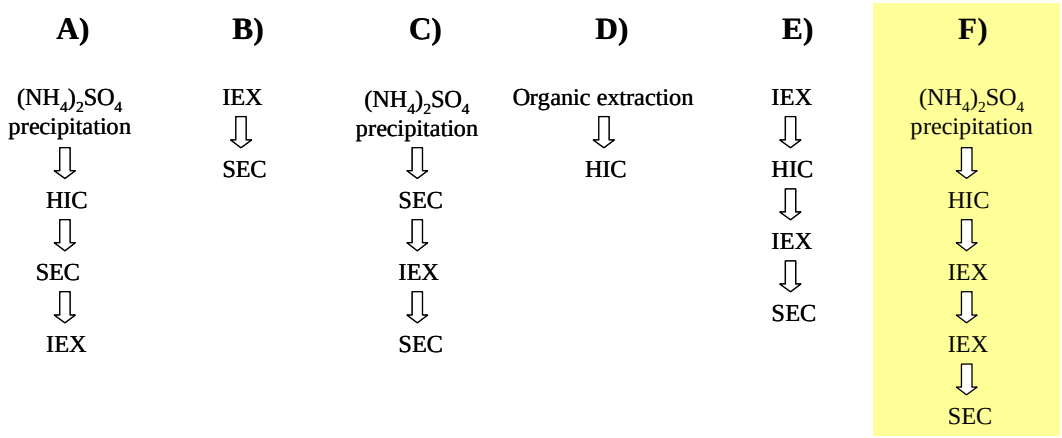


Figure 4.4. Schematic representations of some strategies previously employed for purification of untagged GFP's from *E. coli* cell lysates. **A)** Xie *et al.*[170]; **B)** Huang *et al.*[169]; **C)** Fukuda *et al.*[156]; **D)** Yakhnin *et al.*[171], Samarkina *et al.*[172]; **E)** Yang *et al.*[149]; **F)** Purification strategy employed in this thesis for the purification of EGFP and EGFP-HIT2. See the Materials and Methods Sections 2.4.10 and 2.4.11 for detailed protocols.

A SEC step is therefore unsuitable at earlier stages in a purification process when the total protein content of the sample is high due to high abundance of contaminants. Secondly the SEC step can be used as a buffer exchange step. The buffer employed could hence be chosen based on the demands of downstream assays.

The detector in the machines employed in several purification steps in the purification of untagged EGFP and EGFP-HIT2 could detect several wavelengths. Both the absorbance at 280 nm and the EGFP specific absorbance 485 nm could therefore be monitored simultaneously. The purity of the elution peaks of EGFP and EGFP-HIT2 could therefore be assessed by the relative intensity of absorbance at 280 nm and 485 nm.

The protocols for the preparative purification of EGFP and EGFP-HIT2 (Figure 4.4F) are described in detail in the Materials and Methods Sections 2.4.10 and 2.4.11. The $(NH_4)_2SO_4$ precipitation step removed some *E. coli* host protein, as could be observed by a new pellet forming after incubation of the cleared cell lysate in 1.2 M $(NH_4)_2SO_4$ on ice for 30 min. The majority of the contaminants still remained as judged by SDS-PAGE (Figure 4.5E, lane 3).

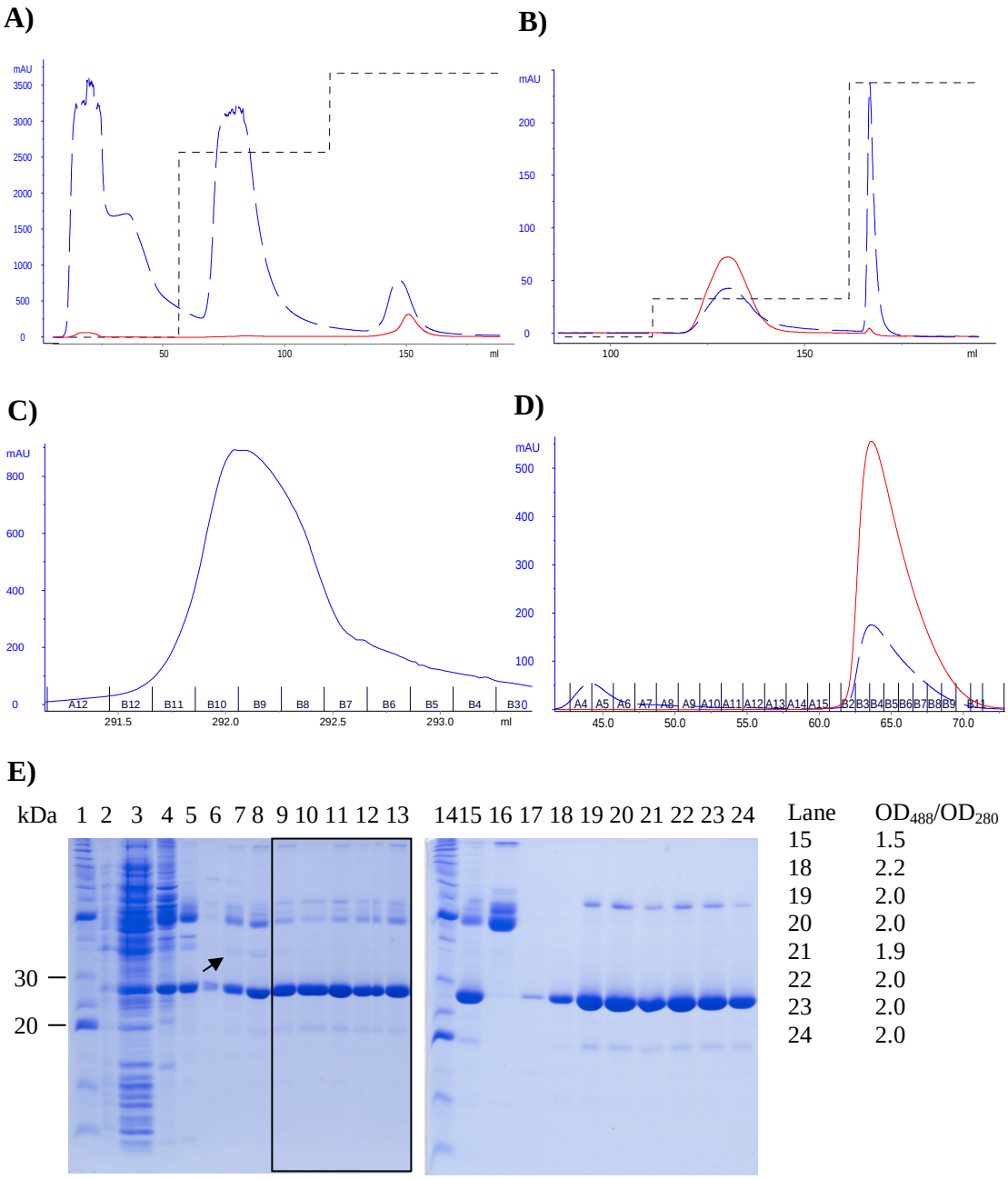


Figure 4.5. Chromatograms representing the purification of untagged EGFP. **A)** HIC on a 20 mL phenyl Sepharose 6 F F column. **B)** IEX on a Q-Sepharose™ column. **C)** IEX on a MonoQ™ 5/50 column. **D)** SEC on a Superdex 75 16/60 column. Solid trace: absorbance at 485 nm. Dashed trace: absorbance at 280 nm. Dotted trace: gradient. **E)** 15 % SDS-PAGE gels. Lanes 1 and 14: size marker, lane 2: cell lysate, lane 3: sample after AS precipitation, lane 4: HIC elution fraction, lane 5: Q-Sepharose elution fraction, lane 6: MonoQ™ elution fraction A12, lanes 7-13: MonoQ™ elution fractions B12-B6, lane 15: SEC starting material pooled MonoQ™ fractions (boxed), lane 16: pool of SEC fractions A4+A5, lanes 17-24: SEC fractions B1-B9.

The supernatant was loaded onto a Phenyl Sepharose 6 F F (Low Sub) column in 1.2 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0. A first wash step (1.2 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0) and a second wash step (0.36 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0) removed significant amounts of contaminant proteins as judged by the intense absorbance at 280 nm and the lack of absorbance at 485 nm (Figure 4.5A). EGFP was eluted in 20 mM Tris-HCl, pH 8.0. In the elution peak the protein purity could be assessed as low since absorbance at 280 nm was higher than the absorbance at 485 nm. Some contaminant proteins were however removed as judged by SDS-PAGE (Figure 4.5E, lane 4).

The eluate was diluted 10-fold in 20 mM Tris-HCl, pH 8.0 to decrease the concentration of $(NH_4)_2SO_4$ remaining from the HIC step. A 1/10 dilution was sufficient to facilitate binding on a 5 mL Q-Sepharose 6 F F column. Untagged EGFP was washed with 20 mM Tris-HCl, pH 8.0 and eluted in 150 mM NaCl, 20 mM Tris-HCl, pH 8.0 (Figure 4.5B). Remaining contaminant proteins bound to the Q-Sepharose 6 F F column were removed by 1 M NaCl, 20 mM Tris-HCl, pH 8.0 (Figure 4.5B). EGFP-HIT2 was eluted in 250 mM NaCl, 20 mM Tris-HCl, pH 8.0. The higher salt concentration required to elute EGFP-HIT2 as compared to untagged EGFP could be due to the HIT2 tag containing several acidic amino acid residues (three aspartic acid residues and two glutamic acid residues) providing additional ionic interactions to the Q-Sepharose 6 F F resin.

As can be seen in the chromatogram (Figure 4.5B) the absorbance at 485 nm was higher than the absorbance at 280 nm in the elution peak. This was not observed in the previous HIC step (Figure 4.5A). Hence the purity of the EGFP was significantly improved by the first IEX step. The higher purity was confirmed by SDS-PAGE (Figure 4.5E, lane 5) containing only one major contaminant protein at ~50 kDa as opposed to the starting material (Figure 4.5E, lane 4). The sample was subjected to buffer exchange on-line and another high resolution IEX step on a prepacked MonoQ™ column as described in experimental step 4.3.1.1.B (Figure 4.5C). This step further increased the purity of the protein as judged by SDS-PAGE analysis of the elution fractions (Figure 4.5E, lanes 6 - 13).

Some elution fractions contained contaminating proteins of low abundance (Figure 4.5E, lanes 7, 8). Elution fractions represented in lanes 9-13 were pooled due to the absence of the weak contaminant band at ~35 kDa (marked with an arrow in Figure 4.5E) since this contaminant was too close in size to untagged EGFP to be separated by the following SEC step. The pooled elution fractions from the Q-Sephaose™ step were used as starting material in the final SEC step. These samples are marked with a box in Figure 4.4E.

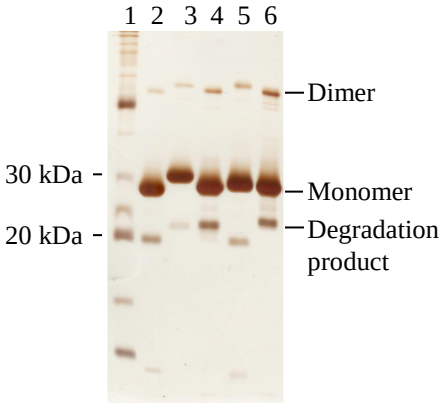
The pooled fractions from the MonoQ™ step contain high molecular weight contaminant proteins as judged by SDS-PAGE (Figure 4.5E, lane 15). These high molecular weight contaminants could be separated from the EGFP by the final SEC step (Figure 4.5E, lane 16). The high molecular weight contaminants were not dimers of EGFP. This was

evident by the lack of absorbance at 485 nm of the peak at 450 mL in the SEC chromatogram (Figure 4.5D). The removal of this high molecular weight contaminant increased the OD_{488}/OD_{280} ratio from 1.5 to 2.0.

4.3.1.3 Final protein purity assessment and yields

The purity of all proteins was analysed by SDS-PAGE and OD_{488}/OD_{280} as shown in Figure 4.6. The yield of each protein is also shown in Figure 4.6. EGFP-His_6, untagged EGFP and EGFP-HIT2 all had OD_{488}/OD_{280} ratios of 1.9 while the ratios for NT1A-EGFP and EGFP-NT1A were 1.6 and 1.5 respectively.

The absorbance of a protein solution at 280 nm depends on the sidechains of Trp, Tyr and cystine[173]. The extinction coefficients of these chromophores are dependent on their local environment[173]. Pace *et al*. determined the extinction coefficient ε of Trp, Tyr and cystine to 5520, 1457 and 173 M^{-1} cm^{-1} respectively based on the average molar absorption of 18 proteins[173]. The NT1A tag contains one Trp residue. This could lead to higher absorbance at 280 nm for the NT1A tagged proteins, which could explain the lower OD_{488}/OD_{280} ratios of NT1A-EGFP and EGFP-NT1A compared to EGFP, EGFP-His_6 and EGFP-HIT2 (Figure 4.6).



Lane	Protein	OD_{488}/OD_{280}	Yield*
1	Size marker 0.5 µL		
2	Untagged EGFP	1.9	16
3	EGFP-HIT2	1.9	30
4	EGFP-His_6	1.9	15
5	NT1A-EGFP	1.6	14
6	EGFP-NT1A	1.5	4.5

Figure 4.6. Silverstained 15 % SDS-PAGE of purified proteins and corresponding OD_{488}/OD_{280} values as an additional purity assessment. One µg protein was loaded in each lane. *Yields are in milligram pure protein (> 95 %) obtained per 10 L liquid cell culture.

Based on the SDS-PAGE (Figure 4.6) all proteins were considered > 95 % pure. The yields (stated as mg pure protein per 10 L liquid cell culture in Figure 4.6) were mainly dependent on the level of purification in each chromatography step. In some cases fractions containing the target protein were excluded due to the presence of contaminating proteins in low abundance. This led to a lower yield. The automated purification processes of EGFP-His_6, NT1A-EGFP and EGFP-NT1A on the Äkta XPress system gave lower yields compared to the manual purification processes for untagged EGFP and EGFP-HIT2. This was most

likely due to the automated wash steps performed by the Äkta XPress system in which protein was lost. The expression of protein could also differ between different batches of expression.

The low yield of EGFP-NT1A was caused by poor binding to the HisTrap™ resin. Different retention of N- and C-terminally tagged proteins was observed for fusion proteins containing the NT1A and HIT2 tag in sequence (Chapter 3). No such difference was however observed for NT1A-EGFP and EGFP-NT1A on immobilized Ni^{2+} and Cu^{2+} chelates of tacn and propyl-*bis*-tacn (Chapter 3). The lower binding of EGFP-NT1A to the HisTrap™ resin could also be due to decreasing performance of the column over time.

4.3.2 Methodological considerations for unfolding studies

The fluorescence of GFP in water is quenched if the tertiary structure of the protein is destroyed. The fluorescence is therefore a sensitive indicator of correct tertiary structure of the protein[151]. The secondary structure content of the protein can be measured using circular dichroism (CD) spectroscopy. In the case of GFP the CD spectrum at 190-260 nm is very similar to reference spectra of β-sheet structure due to the β-barrel of GFP. Loss of the signal indicates destruction of the β-barrel. Fluorescence spectroscopy and CD spectroscopy therefore complement each other well and can provide complementary information about the protein secondary and tertiary structure. The spectroscopy methods were therefore used in parallel to assess the relative stability of the variously tagged EGFP's. Intrinsic Tyr and Trp fluorescence has also been used to monitor EGFP unfolding[151; 168; 169; 170]. The emission maxima of these amino acids shift wavelength when buried in the protein or exposed to the surrounding solvent. The NT1A tag contains one Trp residue. Hence, two emission signals were observed for the EGFP-NT1A fusions. One corresponded to the Trp residue in the NT1A tag and one to Trp57 of EGFP. Hence the intrinsic fluorescence of NT1A tagged EGFP differed from untagged EGFP and EGFP-HIT2. As the aim was to compare the stability of the different proteins Trp and Tyr fluorescence was not employed.

Thermal unfolding was investigated as a means to measure the stability of the proteins. Due to the slow unfolding kinetics of EGFP this method was however not practically feasible. Chemical denaturation using GuHCl was therefore chosen as the method to unfold the protein.

4.3.3 Structural comparison of EGFP and EGFP fusion proteins

Two different purification strategies were employed to obtain pure proteins. EGFP-His_6, EGFP-NT1A and NT1A-EGFP were purified by IMAC followed by IEX while EGFP and EGFP-HIT2 were purified by $(NH_4)_2SO_4$ precipitation, HIC, two IEX steps and SEC.

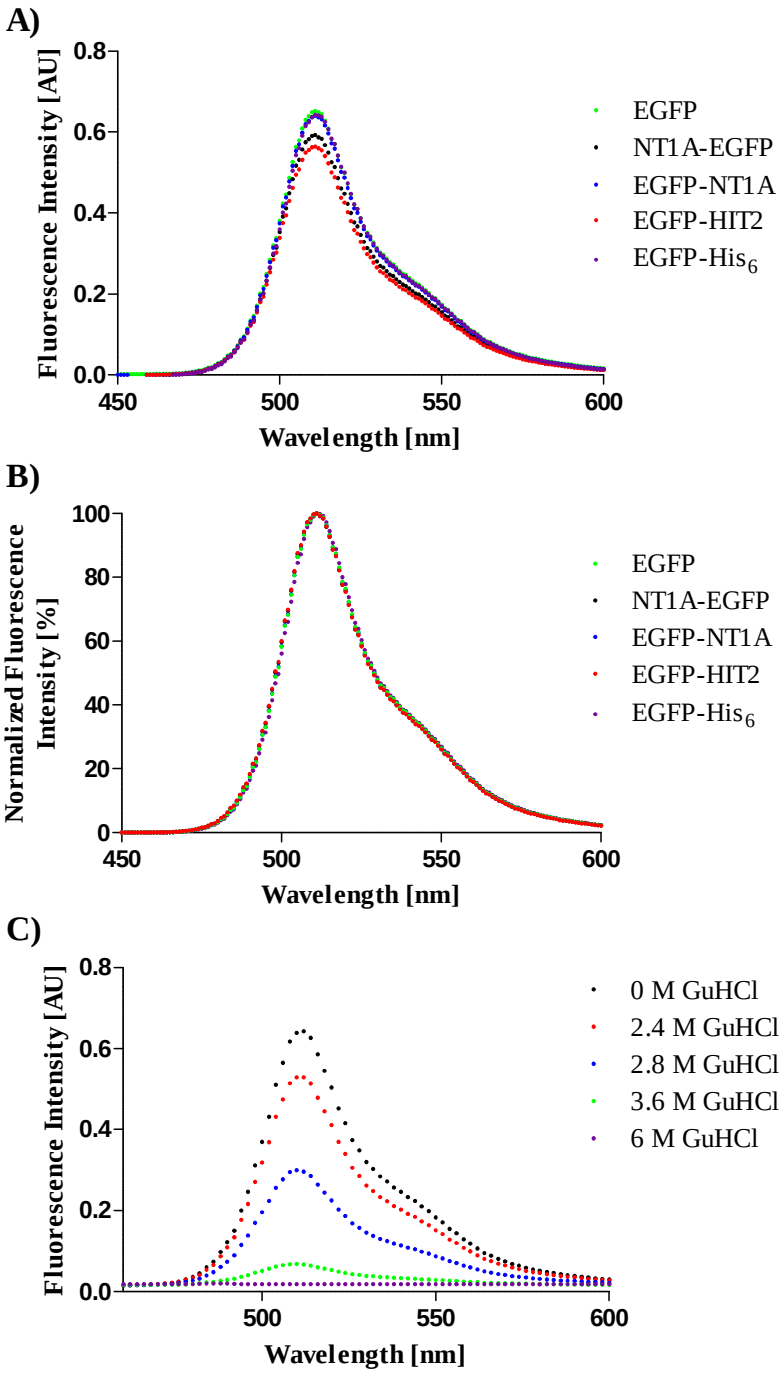


Figure 4.7A) Fluorescence emission spectra and **B)** normalized fluorescence spectra of 10 µM protein samples in 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6. Green: EGFP, black: NT1A-EGFP, blue: EGFP-NT1A, red: EGFP-HIT2, purple: EGFP-His_6. **C)** Fluorescence emission spectra of 10 µM EGFP-His_6 incubated in 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 supplemented with GuHCl for 24 h at RT. Black: 0 M GuHCl, red: 2.4 M GuHCl, Blue: 2.8 M GuHCl, green: 3.6 M GuHCl, purple: 6 M GuHCl.

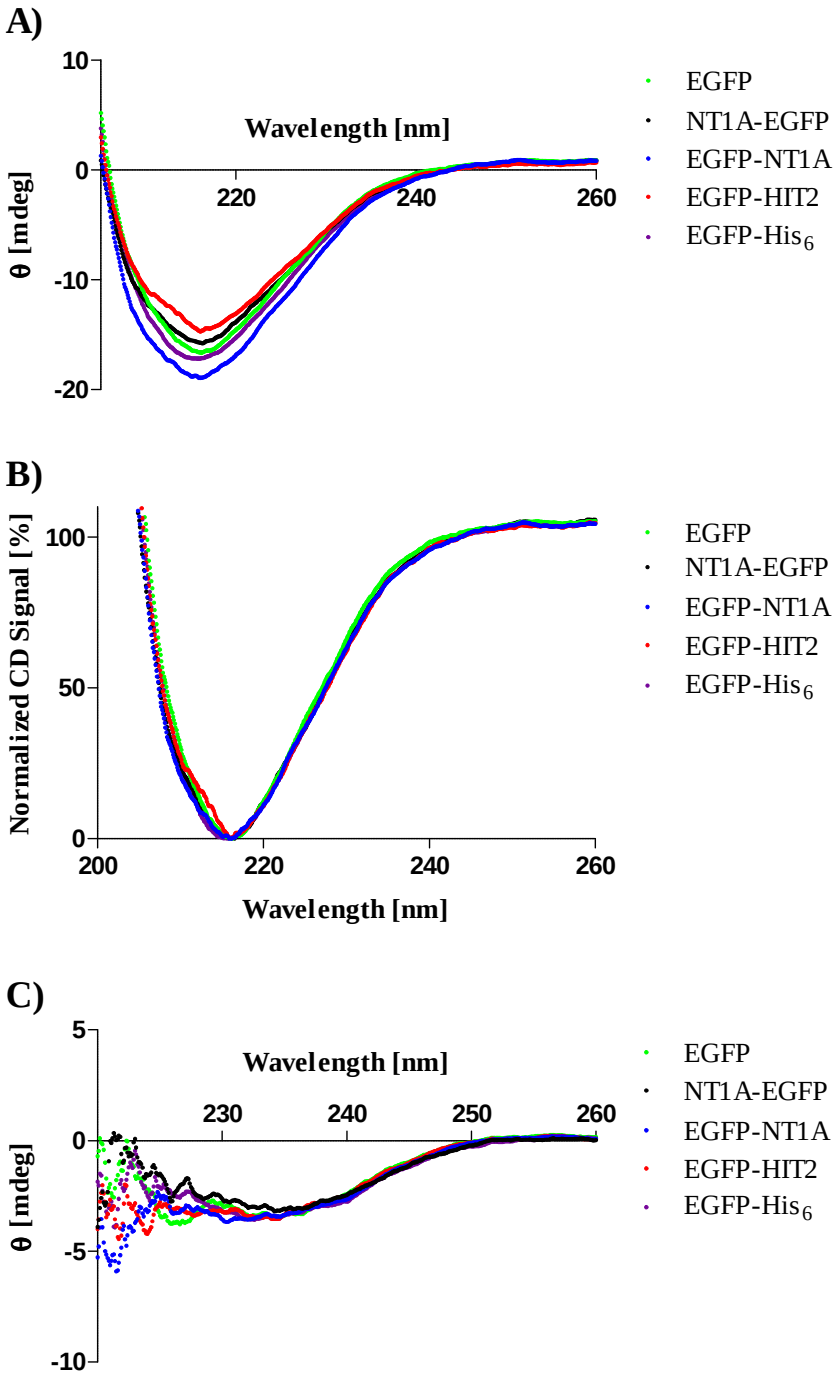


Figure 4.8A) Original CD spectra and **B)** normalized CD spectra of proteins in 10 µM protein samples in 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6. **C)** CD spectra of proteins denatured by incubation in 6 M GuHCl, 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 for 24 h at RT. Green: EGFP, black: NT1A-EGFP, blue: EGFP-NT1A, red: EGFP-HIT2, purple: EGFP-His_6.

From visual observation throughout the purification steps of all five expressed and purified proteins, they remained green, indicative of no gross structural distortion of the tertiary structures caused by the purification procedures.

The spectral properties of the proteins were also investigated. All recombinant proteins EGFP, EGFP-His_6, EGFP-HIT2, EGFP-NT1A and NT1A-GFP, displayed similar fluorescent emission spectra profiles (Figure 4.7A). When the fluorescence emission spectra were normalized against each other (Figure 4.7B) they overlay completely and can not be distinguished from each other. Hence, the slight differences of the intensities of the emission spectra observed in Figure 4.7A could be due to slight variations in the concentrations of the different proteins. The far-UV CD spectra of the recombinant proteins are shown in Figure 4.8A. A negative peak can be observed for all proteins at 205-240 nm (Figure 4.8A). The shape of these peaks (Figure 4.8A) correspond to that obtained for a protein with high β-sheet content[174]. This is in agreement with EGFP consisting of a β-barrel surrounding an α-helix[148; 149]. Furthermore the CD spectra of EGFP obtained here (Figure 4.8A) appear identical to that obtained by Stepanenko *et al.*[175] The CD spectra were normalized against each other at 216 nm to adjust for any minor differences in concentration between the different proteins. A slight difference in the shape of the CD spectrum for EGFP-HIT2 can be observed between 210-215 nm (Figure 4.8B).

All proteins were denatured when incubated in 6 M GuHCl, 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 for 24 h at RT. This was evident by loss of fluorescence of the samples as judged by eye and by the complete loss of fluorescence at 512 nm (Figure 4.9A). As an example the gradual loss of fluorescence of EGFP-His_6 incubated in increasing concentrations of GuHCl in 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 for 24 h at RT is shown in (Figure 4.7C).

The CD spectra of the denatured proteins, incubated in 6 M GuHCl, 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 for 24 h at RT for 24 h, are shown in Figure 4.8C. It is evident that the majority of secondary structure is lost for all proteins with the spectra of the proteins being nearly identical (Figure 4.8C). CD spectra of random structure generally have a weak positive signal at 210-230 nm[174]. It is thus noteworthy that a weak negative signal can be observed for all proteins (Figure 4.8C). Some secondary structure might hence still be preserved even after incubation of the proteins in 6 M GuHCl, 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 for 24 h. These observations are in agreement with previous studies showing that GFP contain a superstable core[176] and that the acid denatured state retain secondary structure[151].

Hence the spectral properties of EGFP obtained here (Figure 4.7 and 4.8) show that the affinity tags NT1A, HIT2 and His_6 do not affect the secondary or tertiary structure of EGFP. Moreover the NT1A tag did not affect the secondary or tertiary structure of EGFP when fused to the N- or C-terminus of the protein.

4.3.4 GuHCl induced quasi-equilibrium unfolding experiments

The conformational stability of the tagged and untagged EGFP's was studied by chemical denaturation using GuHCl. GFP unfolding is a slow process and chemical denaturation curves change with equilibration time[169]. GuHCl induced quasi-equilibrium unfolding experiments were therefore performed by incubating protein samples (10 μM) in various concentrations of GuHCl in 10 mM phosphate buffer, 150 mM NaCl, pH 7.6 for 18 h at 25 °C. Two such GuHCl induced quasi-equilibrium unfolding experiments were performed for each protein.

The fluorescence emission at 512 nm and the CD signal at 225 nm were measured simultaneously. The obtained fluorescence emission and CD signals were plotted as a function of GuHCl concentration (Figure 4.9A, B). The resulting fluorescence unfolding curves were normalized so that the fluorescence in the absence of denaturant represented 100 %. The fluorescence unfolding curves show a slow decrease in fluorescence intensity between 0 and 2 M GuHCl followed by a steep transition between 2 and 3 M GuHCl (Figure 4.9A). A similar fluorescence unfolding curve was obtained by Xie *et al*.[170]

The unfolding curves of the simultaneously obtained CD signal at 225 nm were normalized so that the signal obtained in the absence of denaturant represented 0 % and the highest signal obtained represented 100 %. The CD unfolding curves show only one transition between 2 and 4 M GuHCl (Figure 4.9B). The quality of the data at GuHCl concentrations above 4 M was low due to the contribution of noise in the CD signal originating from the GuHCl.

As mentioned in the introduction, several research groups have reported the presence of GFP unfolding intermediates characterized by preserved secondary structure but with distorted tertiary structure leading to decreased fluorescence[151; 167; 168; 169]. Hence, the different transitions of the fluorescence and CD unfolding curves observed here (Figure 4.9A, B) can be attributed to the presence of unfolding intermediates with decreased fluorescence but preserved secondary structure. The unfolding intermediate of GFP_{UV} characterized by Huang *et al*.[169] was studied by chemical denaturation using GuHCl. The unfolding intermediate had preserved secondary and tertiary structure. *ß*-strands 7-10 were however displaced and the fluorescence was reduced[169]. As the methods used by these authors were similar as those used in this thesis it is probable that the intermediate observed here had similar characteristics. Dietz *et al*.[167] studied the unfolding of GFP_{UV} with an additional S2G point mutation using mechanically induced unfolding. The authors detected two unfolding intermediates. One with the N-terminal α-helix detached but with the *β*-barrel intact and the other one with the N-terminal α-helix and one *β*-strand detached[167]. Enoki *et al*.[151; 168] studied the refolding of GFP_{UV} from the acid denatured state.

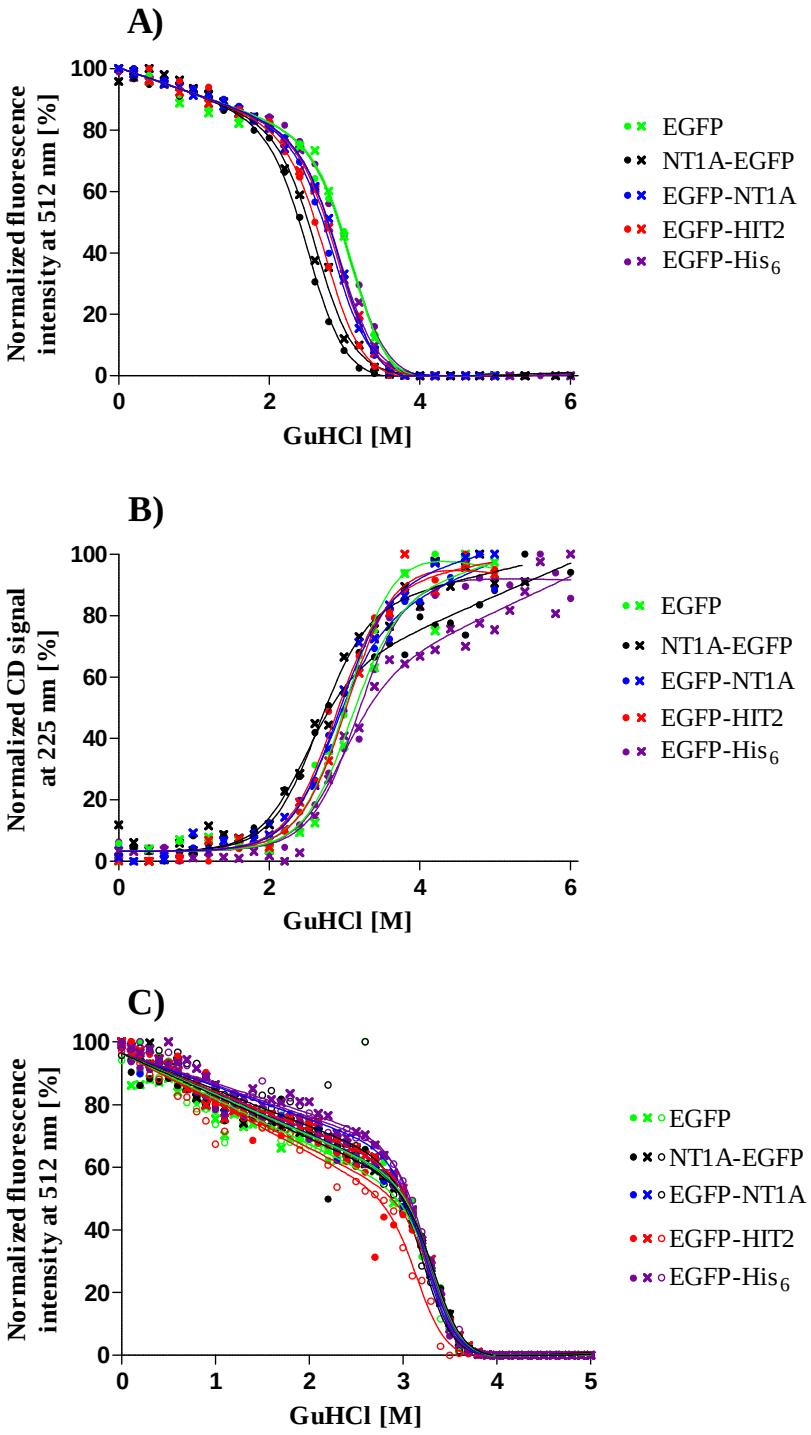


Figure 4.9. Unfolding curves of proteins incubated in 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6 supplemented with various concentrations of GuHCl. **A)** Fluorescence at 512 nm (λ_{exc} = 488 nm) and **B)** CD signal at 225 nm of 10 µM samples measured simultaneously in a CD spectrometer/fluorometer. **C)** Fluorescence at 512 nm (λ_{exc} = 480 nm) of 1 µM samples measured in a platereader. Green: EGFP, black: NT1A-EGFP, blue: EGFP-NT1A, red: EGFP-HIT2, purple: EGFP-His_6.

The authors identified two unfolding intermediates, one native-like and one exhibiting the characteristics of a molten globular state[168]. As the methods of unfolding employed by Dietz *et al.*[167] and Enoki *et al.*[151; 168] were different to the method employed in this Chapter, these intermediates could differ from the one(s) observed here.

Data analysis of unfolding curves containing unfolding intermediate states can be performed assuming a native state N, any number of intermediate states I_1, I_2,…I_n and a fully unfolded state U. This method would reflect the actual experimental observations most accurately. Alternatively a simplified two-state model containing only the two states N and U can be employed. Research groups have fitted GuHCl unfolding curves of different mutants of GFP to either a two state unfolding model[166] or both a two and a three state model[169; 170]. The aim in this Chapter was to explore whether any difference in stability occurred for the variously tagged proteins. The CD and fluorescence data from the GuHCl unfolding experiments were therefore fitted to a two-state model[108] (Equation 4.1).

Equation 4.1
$$Y = \frac{(F_n + M_n x) + (F_u + M_u x)\exp\left(\left(-\Delta G \cdot M_g x\right)/RT\right)}{1 + \exp\left(-\left(\Delta G - M_g x\right)/RT\right)}$$

Y is the observed intensity of optical property (CD signal or fluorescence), x is the denaturant concentration, F_n and F_u are the observed intensity of the native and unfolded state respectively, M_n and M_u are the slopes of the baseline before and after the transition respectively and M_g is the slope of the transition. The data were fitted globally; F_n, M_n, and M_g were shared between the data sets. The difference in free energy of the native state N and the unfolded state U, ΔG_{NU} (kcal mol^{-1}) and the midpoint in each unfolding curve $c_{1/2}$ (M GuHCl) could then be extracted.

Since the unfolding curves obtained in this Chapter had not reached equilibrium, ΔG_{UN} can not be determined accurately from this data. Hence the apparent free energy of unfolding ($\Delta G'_{UN}$) in kcal mol^{-1} was extracted from these quasi-equilibrium unfolding curves. The thermodynamic parameters obtained are summarized in Table 4.1.

The midpoint of transition appeared to be similar between the fluorescence and CD unfolding curves as judged by the $c_{1/2}$ values (Table 4.1). The values of $\Delta G'_{UN}$ (Table 4.1) were however consistently higher for the fluorescence unfolding curves (6.1 – 7.4 kcal mol^{-1}) compared to the CD unfolding curves (4.7 – 5.8 kcal mol^{-1}). This could be an effect of treating the three-state unfolding observed by fluorescence as a two-state unfolding process.

The midpoints of transition varied slightly between the proteins. This can be observed both visually (Figure 4.9A, B) and by the calculated $c_{1/2}$ values in Table 4.1. NT1A-EGFP had a notably low $c_{1/2}$ value of 2.5 ± 0.1 M GuHCl in the CD unfolding curve and 2.6 ± 0.1 M

GuHCl in the fluorescence unfolding curve. For all other proteins the $c_{1/2}$ values were clustered between 2.8 ± 0.1 and 3.1 ± 0.1 M GuHCl (Table 4.1). $\Delta G'_{UN}$ was also lower for NT1A-EGFP in the CD unfolding curves; 4.7 ± 0.3 kcal mol^{-1} compared to 5.3 ± 0.1 to 5.8 ± 0.1 kcal mol^{-1} for the other proteins (Table 4.1). The same pattern was observed for the fluorescence unfolding curves where $\Delta G'_{UN}$ was 6.1 ± 0.2 kcal mol^{-1} for NT1A-EGFP compared to 6.9 ± 0.2 to 7.4 ± 0.3 kcal mol^{-1} for the other proteins (Table 4.1).

To investigate if the affinity tags NT1A and HIT2 affected the stability of EGFP, the difference in free energy of unfolding ΔΔG between any two proteins P1 and P2 was calculated using Equation 4.2. If ΔΔG = 0 the affinity tags have no effect on the stability of EGFP.

Equation 4.2 $\Delta\Delta G(P1 - P2) = \Delta G'(P1) - \Delta G'(P2)$

The statistical significance ($P < 0.05$) of the ΔG' values was tested by a one-way ANOVA test with a Tukey post test. NT1A-EGFP had significantly lower ΔG' values compared to the other recombinant proteins. From the CD unfolding curves ΔΔG(EGFP-His_6 - NT1A-EGFP) = 1.0 ± 0.5 kcal mol^{-1} and ΔΔG(EGFP - NT1A-EGFP) = 1.1 ± 0.3 kcal mol^{-1}. From the fluorescence unfolding curves ΔΔG(EGFP-His_6 - NT1A-EGFP) = 1.3 ± 0.4 kcal mol^{-1} and ΔΔG(EGFP - NT1A-EGFP) = 1.3 ± 0.3 kcal mol^{-1}. Hence, it appeared as if the NT1A tag had a destabilizing effect on EGFP when fused to the N-terminus of the protein.

The unfolding curves obtained on the CD/fluorescence spectrometer had an experimental dead time of several hours. As the unfolding curves were not at equilibrium, variations in the dead time of the experiments could hence potentially affect the results. To determine whether the observed ΔΔG was due to variations in the length of incubation time, GuHCl induced quasi-equilibrium unfolding experiments were performed in a platereader. One μM protein solutions containing various concentrations of GuHCl in 10 mM phosphate buffer, 150 mM NaCl, pH 7.6 were incubated for 18 h at RT. Three quasi-equilibrium experiments were performed for each protein. The fluorescence emission intensity was measured at 512 nm in a platereader. The excitation wavelength was 480 nm to avoid overlap with the emission wavelength. The fluorescence intensity was plotted as a function of the GuHCl concentration (Figure 4.9C). The presence of an unfolding intermediate at low GuHCl concentrations can more clearly be observed from these experiments (Figure 4.9C) compared to the unfolding experiments performed in the CD/fluorescence spectrometer (Figure 4.9A). This probably reflects the higher density of data obtained in the platereader. The major transition is steeper (Figure 4.9C), possibly reflecting the difference in experimental dead time between the plate reader and spectrometer setups.

Table 4.1. Thermodynamic parameters obtained from GuHCl induced quasi-equilibrium unfolding experiments of tagged and untagged EGFP proteins.

Protein	Untagged EGFP	EGFP-His_6	EGFP-NT1A	NT1A-EGFP	EGFP-HIT2
Fluorescence measurements obtained in a CD/fluorescence spectrometer (Figure 4.9A). Average values and standard deviation of two experiments are shown.					
$\Delta G'_{NU}$ (kcal mol^{-1})	7.4 ± 0.1	7.4 ± 0.3	6.9 ± 0.2	6.1 ± 0.2	6.7 ± 0.3
$C_{1/2}$ (M GuHCl)	3.1 ± 0.1	3.0 ± 0.1	2.9 ± 0.1	2.6 ± 0.1	2.8 ± 0.1
CD measurements obtained in a CD/fluorescence spectrometer (Figure 4.9B). Average values and standard deviation of two experiments are shown.					
$\Delta G'_{NU}$ (kcal mol^{-1})	5.8 ± 0.1	5.7 ± 0.4	5.3 ± 0.1	4.7 ± 0.3	5.6 ± 0.3
$C_{1/2}$ (M GuHCl)	3.1 ± 0.1	3.0 ± 0.2	2.8 ± 0.1	2.5 ± 0.1	3.0 ± 0.1
Fluorescence measurements obtained in a plate reader (Figure 4.9C). Average values and standard deviation of three experiments are shown.					
$\Delta G'_{NU}$ (kcal mol^{-1})	12.2 ± 0.1	12.0 ± 0.2	12.0 ± 0.2	12.0 ± 0.2	12.0 ± 0.3
$C_{1/2}$ (M GuHCl)	3.28 ± 0.02	3.23 ± 0.04	3.23 ± 0.04	3.25 ± 0.06	3.23 ± 0.08

The data were fitted using Equation 4.1. Each unfolding curve was fitted individually to Equation 4.1 rather than taking the average of each point and fit the obtained curve of the average values. Hence the data analysis conformed to that of the unfolding curves obtained on the CD/fluorescence spectrometer. The unfolding curves obtained in the platereader were clustered more tightly in the transition of the curve but were more scattered in the region 0-3 M GuHCl (Figure 4.9C) compared to the unfolding curves obtained on the CD/fluorescence spectrometer (Figure 4.9A). This is reflected in the narrow range of $c_{1/2}$ values of 3.23 ± 0.04 to 3.28 ± 0.08 M GuHCl for all proteins (Table 4.1). This is slightly higher compared to the $c_{1/2}$ values of 2.6 ± 0.1 to 3.1 ± 0.1 M GuHCl of the fluorescence unfolding curves obtained in the CD/fluorescence spectrometer (Table 4.1). The value of $\Delta G'_{NU}$ was 12.0 ± 0.3 kcal mol^{-1} for all proteins (Table 4.1). This is dramatically different from the corresponding range of 6.1 ± 0.2 to 7.4 ± 0.3 kcal mol^{-1} obtained using the CD/fluorescence spectrometer (Table 4.1). This difference in $\Delta G'_{NU}$ from the different experiments could be due to different dead times in the experiments.

No significant difference in stability for any of the proteins could be observed from the data obtained from the plate reader. It is therefore reasonable to conclude that the obtained

difference in free energy of unfolding ($\Delta\Delta G \neq 0$) calculated from the unfolding curves obtained on the CD/fluorescence spectrometer was an effect of the method employed. Hence the results obtained here show that the affinity tags NT1A, HIT2 and His_6 do not affect the $\Delta G'_{NU}$ of EGFP. Moreover the terminus to which the NT1A tag was attached did not affect $\Delta G'_{NU}$ of EGFP.

4.3.5 Kinetics of EGFP unfolding

In addition to the quasi-equilibrium unfolding measurements described in Section 4.3.4, the fluorescent decay of EGFP, EGFP-NT1A, NT1A-EGFP, EGFP-His_6 and EGFP-HIT2 was studied as a function of time at different concentrations of GuHCl. The experiment was carried out in a plate reader using 1 µM protein solutions.

Control samples containing no GuHCl were included as comparison. Fluorescent emission at 512 nm was measured every five minutes (λ_{exc} = 480 nm). The decay of protein samples in 4 M GuHCl is shown in Figure 4.10.

The positive control experiments without denaturant show significant fluorescent decay over time for all proteins (Figure 4.10A). EGFP is known to be stable at the bench in physiological buffer and retains its fluorescence intensity for about a month[165]. EGFP is both reversible and irreversible photobleached[177]. It is therefore reasonable to assume that the observed fluorescence decay was due to photobleaching of the EGFP fusion proteins during the experiment. The decay curves were hence adjusted for the decay of the positive controls. The decay curves normalized so that the signal at the first measurement corresponds to 100 % and adjusted for the decay of the positive controls are shown in Figure 4.10B.

Fluorescence decay of GFP_{wt} and GFP_{UV} was shown to be a monoexponential function by Fukuda *et al.*[156] The fluorescence decays as a function of time were hence fitted to a one-phase decay curve using Equation 4.3. Y is the fluorescence signal as a function of time, Y(0) is the fluorescence at time 0, k is the rate constant and t is the time. The rate constants and half lives for the decay curves are tabulated in Table 4.2. The fluorescence of the unfolded EGFP's was assumed to be zero.

Equation 4.3 $Y(t) = Y(0)\, e^{-kt}$

It can be seen that the decay rate is very similar for all proteins with half lives of ~ 40 min and complete loss of fluorescence emission in 200 min (Figure 4.10). The half lives and rate constants were not significantly different between the proteins as judged by a one-way ANOVA generating a P value of 0.8211. As such no significant difference in the rate of unfolding could be determined between the variously tagged proteins.

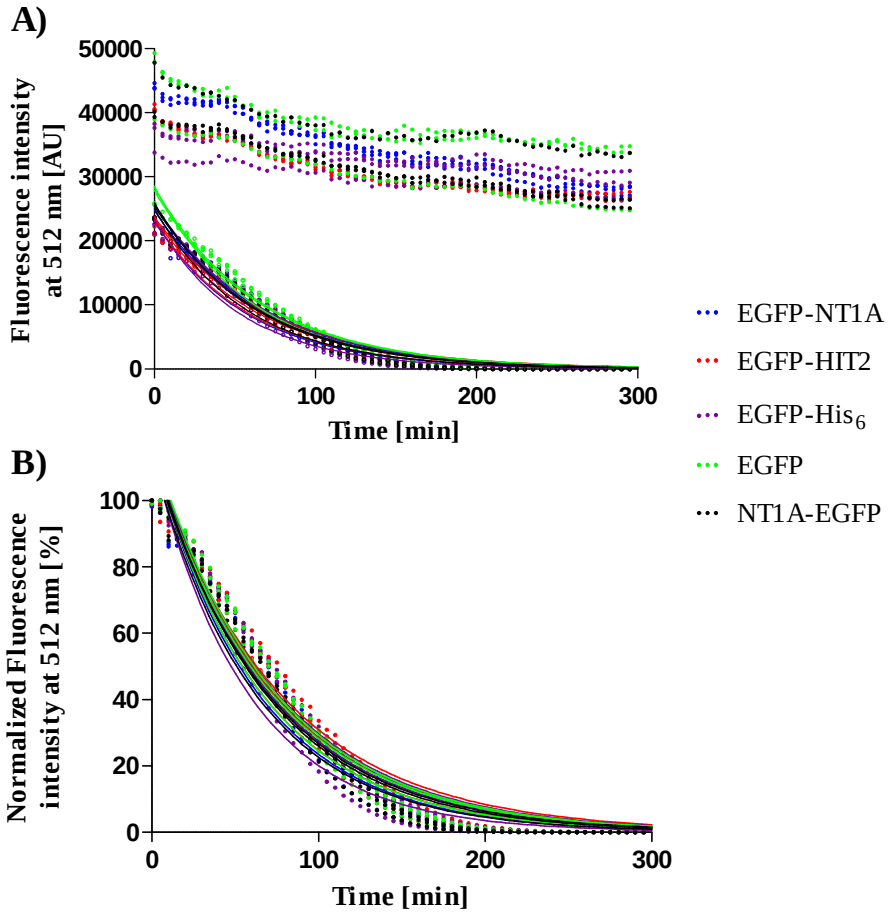


Figure 4.10. Fluorescence intensity at 512 nm (λ_{exc} = 480 nm) as a function of time of tagged and untagged EGFP proteins in 4 M GuHCl, 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6. The intensity was measured every 5 minutes. EGFP (green), EGFP-NT1A (blue), NT1A-EGFP (black), EGFP-His_6 (purple) and EGFP-HIT2 (blue) as a function of time. a) Raw data with positive controls consisting of proteins without GuHCl shown at the top of the graph. b) Decay curve normalized to 100 % at time = 0. The data was corrected by the decay observed for EGFP in 0 M GuHCl.

Table 4.2. Rate constants and half-lives for tagged and untagged EGFP proteins. Average values and standard deviation of three experiments are shown.

Protein	Untagged EGFP	EGFP-His_6	EGFP-NT1A	NT1A-EGFP	EGFP-HIT2
k [s^{1-}] x 10^{-3}	0.27 ± 0.02	0.27 ± 0.03	0.27 ± 0.02	0.26 ± 0.02	0.28 ± 0.01
Half-life [min]	44 ± 2	42 ± 5	44 ± 3	42 ± 2	45 ± 4
k adjusted [s^{1-}] x 10^{-3}	0.24 ± 0.01	0.25 ± 0.04	0.24 ± 0.02	0.25 ± 0.01	0.24 ± 0.02
Half-life adjusted [min]	48 ± 3	46 ± 6	47 ± 3	46 ± 2	49 ± 4

Hence the results obtained here show that the affinity tags NT1A, HIT2 and His_6 do not affect the rate of unfolding of EGFP in 4 M GuHCl, 10 mM $NaPO_4$, 150 mM NaCl, pH 7.6. Under the same conditions the terminus to which the NT1A tag was attached did not affect the rate of unfolding of EGFP.

4.4 Discussion

4.4.1 Protein purification

The advantage of employing affinity tags in protein purification was clearly illustrated by the difference in the numbers of chromatography steps required to purify the proteins to similar purity. The purification of EGFP and EGFP-HIT2 was a five-step procedure; $(NH_4)_2SO_4$ precipitation, HIC, two IEX steps and a final SEC step. EGFP-His_6, NT1A-EGFP and EGFP-NT1A were on the other hand purified by only two chromatography steps (IMAC followed by IEX). Considering the difference in the number of chromatography steps required to obtain protein of similar purity, it was noteworthy that the yields of EGFP and EGFP-HIT2 were higher as compared to the proteins fused to a His containing tag. This could reflect losses in the automated Äkta Xpress system originating from dead volumes in the system and losses due to pre-programmed automated wash steps.

4.4.2 Unfolding studies

The values of $\Delta G'_{NU}$ obtained here were in the range of 6 - 7 kcal mol^{-1} for the unfolding curves measured by CD and fluorescence in a spectrometer and 12 ± 0.3 kcal mol^{-1} for the unfolding curves obtained in the platereader (Table 4.1). Literature values of ΔG_{NU} range from[166] 5.16 ± 0.54 kcal mol^{-1} to approximatley[169] 16 kcal mol^{-1}. This range of values reflects the fact that the chemical denaturation unfolding curves do not reach equilibrium even after several months[166; 169], which makes it difficult to determine ΔG_{NU} accurately. Hence it can be concluded that the values of $\Delta G'_{NU}$ obtained here were in the range of those reported in the literature. Since the aim of the study was to determine if there was any difference in free energy of unfolding ($\Delta\Delta G$) caused by the affinity tags NT1A and HIT2, no further efforts were taken towards determining a "true" value of ΔG.

The rate constants and half lives of fluorescence of the EGFP's in 4 M GuHCl obtained here were ~ 0.3 x 10^{-3} s^{-1} and ~ 45 min (Table 4.2). This is not in agreement with the value of k = 1.5 x 10^{-3} s^{-1} and $t_{1/2}$ of 11 min for EGFP in 4 M GuHCl reported by Verkhusha *et al*.[165; 166] These inconsistencies could be due to differences in the methods employed.

The unfolding experiments in the platereader showed no difference in $\Delta G'_{NU}$ between the proteins (Table 4.1 and Figure 4.9C). A clear correlation could be observed between the values of $\Delta G'_{NU}$ obtained by measuring the CD and fluorescence *simultaneously* (Table 4.1).

Hence, the variations in $\Delta G'_{NU}$ between the proteins observed in the simultaneously obtained CD and fluorescence data could be related to the dead time of the experiment. This implies that the affinity tags NT1A and HIT2 do not affect the stability of their fusion partner EGFP.

4.4.3 Measuring the influence of affinity tags on their fusion partner using EGFP

Using EGFP (or other GFP's) as a probe to study the effect of novel affinity tags has both drawbacks and advantages. The biggest limitation is the relatively high stability of the protein secondary structure. This makes it difficult to detect any influence of the tag since the contribution of free energy of unfolding by the affinity tag is small compared to the free energy of unfolding of the β-barrel. Lower stability would also enable thermal unfolding studies to be undertaken. These experiments consume less protein and are less labour intensive as compared to chemical denaturation unfolding curves. This is because only one sample is used and a change in signal as a function of temperature is observed.

An obvious advantage using EGFP for studying the influence of affinity tags on protein stability is the high sensitivity of the fluorophore fluorescence. Small changes in the protein tertiary structure give rise to decreased fluorescence even if the secondary structure, consisting mainly of the β-barrel, is intact. As discussed in detain in Sections 4.1.2, 4.3.3 and 4.3.4, such intermediates have been observed by several research groups and are thought to be caused both by increased flexibility in the β-barrel[169] and detachment of the seven residue N-terminal α-helix[167] capping the β-barrel. Moreover the high fluorescence of the protein enables low concentrations to be used (1 µM as compared to 10 µM for CD measurements). The requirement of pure protein was therefore 500-1000 times lower for fluorescence emission studies in 96 or 384 well formats as compared to CD measurements in a spectrometer.

Assay like the one described in this Chapter is an important part of the process of developing and characterizing new affinity tags since it is of utmost importance that there are no inherent properties in the tag affecting the structure and stability of the target protein. Once this has been confirmed employing a suitable model protein such as EGFP, functional assays will need to be carried out using the protein(s) intended to be purified using the tag(s) in question.

4.5 Conclusions

In conclusion EGFP, EGFP-NT1A, NT1A-EGFP, EGFP-His_6 and EGFP-HIT2 could be purified to > 95 % purity as judged by SDS-PAGE (Figure 4.6) by two different

purification strategies. Spectroscopic measurements, quasi equilibrium unfolding and kinetic unfolding studies of the recombinant proteins showed that the affinity tags NT1A, HIT2 and His_6 do not affect the structure, thermodynamic stability or rate of unfolding of their fusion partner EGFP. Moreover these parameters were not affected by which terminus of EGFP the NT1A tag was fused to. Slightly different purification strategies did not affect the structure, thermodynamic stability and rate of unfolding of EGFP. The affinity tags NT1A and HIT2 therefore fulfil the critical criteria of not altering the function, stability and structure of the model protein EGFP to which they were fused. These findings support that NT1A and HIT2 can be used as affinity tags in IMAC. Quasi-equilibrium unfolding studies using EGFP in 96-well format proved to be a rapid primary screening method for the effect of affinity tags on protein stability.

Chapter 5. Dynamic binding studies of immobilized Ca^{2+} chelates

5.1 Introduction

One of the key parameters when developing a Ca^{2+}-based IMAC system is the nature of the Ca^{2+} chelate. As previously mentioned one of the requirements in an IMAC system is that the metal ion has a stronger affinity for the ligand than for the target protein or the affinity tag. If this is not the case, metal ion transfer (MIT), where the metal ion is displaced by the target protein[90] or affinity tag[37], can occur. From a green chemistry perspective this is a lesser concern if the metal ion is non toxic in the amounts and concentrations employed on a chromatographic scale. Typical ligand densities in IMAC are in the range of hundreds µmol g^{-1} of dry gel. This correspond to tens µmol mL^{-1} swollen gel. If a typical elution peak in a chromatogram is assumed to correspond to a few bed volumes and all Ca^{2+} would be displaced into the elution peak, the resulting Ca^{2+} contamination of the sample would be in the range of tens of µM range at most. As a comparison, the Ca^{2+} concentration in human serum range[178] between 2.2 and 2.7 mM. Considering the fundamentals of IMAC, displacement of the metal ion can however be a concern. It can lead to lack of retention of the protein on the stationary phase.

The metal ion can, in addition to the protein itself, potentially also be displaced by components in the eluents and in the sample. The stability of the Ca^{2+} chelate would therefore need to be higher than the stability of complexes potentially formed between Ca^{2+} and other species such as buffer components and Ca^{2+} chelating amino acids on protein surfaces or affinity tags. Chaga *et al.*[90] discuss the challenge of MIT for purification of Ca^{2+} binding proteins on Ca^{2+}-based IMAC systems. The authors argue that MIT occurs since many Ca^{2+} binding proteins have higher affinities to Ca^{2+} as compared to the immobilized ligands[90]. Nevertheless, the process of MIT in Ca^{2+}-based IMAC systems remains poorly characterized.

5.1.1 Ligands for Ca^{2+} IMAC

Some clear trends can be observed for the stability constants between different metals and ligands. Ca^{2+} form relatively weak complexes with many ligands as compared to trivalent hard metal ions such as La^{3+} and Fe^{3+} and borderline metal ions such as Zn^{2+}, Ni^{2+} and Cu^{2+} (Chapter 1, Table 1.3). For iminodiacetic acid (IDA), a ligand previously employed in Ca^{2+}-

based IMAC[85; 86], this trend is clear with[28] log $\beta_{Ca(II):L}$ = 2.6 < log $\beta_{La(III):L}$ = 5.88 < log $\beta_{Zn(II):L}$ = 7.15 < log $\beta_{Ni(II):L}$ = 8.3 < log $\beta_{Cu(II):L}$ = 10.56 ~ log $\beta_{Fe(III):L}$ = 10.72.

There are however several ligands with relatively high affinities to Ca^{2+}. Macrocyclic ligands based around 1,4,7,10-tetraazacyclododecane (cyclen) are attractive candidates as Ca^{2+} chelating ligands in IMAC. Several lanthanide complexes of cyclen derivatives have been used clinically as magnetic resonance imaging (MRI) contrast agents[179; 180] and nuclear magnetic resonance (NMR) shift reagents[181; 182]. Examples include the N-methyl-D-glucamine salt of $[Gd(DOTA)H_2O]^-$ which has been sold under the trade name DOTAREM® and $[Gd(DO3AHP)H_2O]$ which has been sold as PROHANCE®, both for use as MRI contrast agents[183]. For use as a ligand in IMAC, substituents or pendant arms can be attached to a maximum of three of the cyclen nitrogens since the fourth nitrogen will participate in a covalent bond attaching the ligand to the solid support.

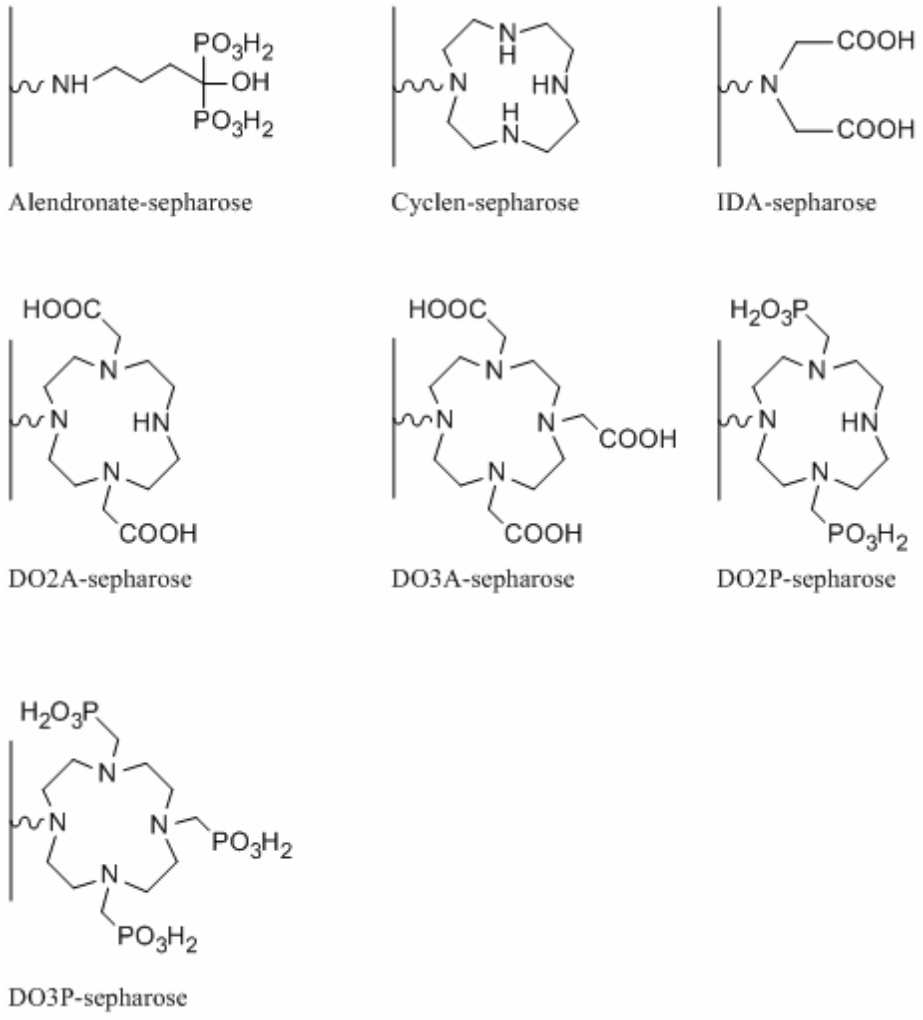


Figure 5.1. Chemical structures of putative Ca^{2+} chelating ligands studied in this Chapter immobilized onto Sepharose 6 F F.

The cyclen derivatives 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid (DO3A), 1,4,7,10-tetraazacycloddoecane-1,7-di(methanephosphonic acid) (DO2P) and 1,4,7,10-tetraazacyclododecane-1,4,7-tri(methanephosphonic acid) (DO3P) in particular have been immobilized to Sepharose 6 F F as depicted in Figure 5.1 and investigated as IMAC ligands for hard metal ions at the Monash Centre for Green Chemistry[95].

Another class of compounds of potential interest as ligands in Ca^{2+} based IMAC are non-macrocyclic bisphosphonates like (3-amino-1-hydroxy-1-phosphonopropyl)phosphonic acid (pamidronate) and (4-amino-1-hydroxy-1-phosphono-butyl)phosphonic acid (alendronate). These compounds are clinically employed for Ca^{2+}-related disorders like hypercalcemia of malignacy, Paget's disease, osteoporosis and tumor osteolysis. The structure of alendronate immobilized onto Sepharose 6 F F is shown in Figure 5.1.

5.1.2 Stability constants of Ca^{2+} to different ligands

The log $\beta_{Ca(II):L}$ of ligands with potential use in IMAC systems has been studied both at the Monash University Centre for Green Chemistry[184] and by other research groups[185; 186; 187; 188; 189; 190]. For several ligands their log $\beta_{Ca(II):L}$ has been compiled in the NIST Critically Selected Stability Constants of Metal Complexes Database[28]. These values were used in this Chapter when available. The log $\beta_{Ca(II):L}$ values are usually determined for the free ligand.

The results can be used as a guide for the log β of the complex formed by the metal and the immobilized ligand. Since many ligands have several protonation states, the stability of the ligands to metals will vary with pH. Table 5.1 show protonation constants for the ligands studied in this Chapter. Table 5.2 shows an overview of stability constants of Ca^{2+} complexes of these ligands in their different protonation states.

The difference in stability of the various Ca^{2+}-chelates (Table 5.2) is quite striking. DO3A and DO3P form stable complexes with[28] log $\beta_{Ca(II):DO3A}$ = 12.5 and[184] log $\beta_{Ca(II):DO3P}$ = 10.96. DO2P[189], 1,4,7,10-tetraazacyclododecane-1,7-diacetic acid (DO2A)[28] and alendronate[186] form complexes with log $\beta_{Ca(II):L}$ between 7 and 9. IDA and cyclen form the least stable complexes[28] with log $\beta_{Ca(II):IDA}$ = 2.6 and log $\beta_{Ca(II):cyclen}$ = 3.1, respectively. The increasing number of pendant arms on the cyclen ring clearly enhances the Ca^{2+} affinity of the ligand.

A general trend can be observed with the stability constants of the Ca^{2+}-chelates increasing with decreasing protonation states of the ligands. As can be seen in Table 5.1 DO3P has six protonation sites. Thus the protonation state of the ligand will vary with pH, as will its affinity to different metals. It should be noted that the high[184] log $\beta_{Ca:DO3P}$ = 10.96 only apply under highly basic conditions since the deprotonated form of the ligand only exist in low proportions at physiological pH. At pH 6 - 9 the predominant complexes are[184] log $\beta_{Ca(II):DO3PH}$ = 9.7, log $\beta_{Ca(II):DO3PH2}$ = 7.9 and log $\beta_{Ca(II):DO3PH3}$ = 7.5.

Table 5.1. Protonation constants of ligands used in this work and related compounds for reference. All values were from the NIST Critically Selected Stability Constants of Metal Complexes Database[28] unless otherwise stated. n.a: not available.

Ligand	log K_1^H	log K_2^H	log K_3^H	log K_4^H	log K_5^H	log K_6^H	log K_7^H	log K_8^H
Alendronate[186]	11.82	10.96	6.36	2.22	1.33	n.a	n.a	n.a
Cyclen	10.39-10.82	9.37-9.72	1.2-1.5	n.a	n.a	n.a	n.a	n.a
IDA	9.10-9.71	2.56-2.84	n.a	n.a	n.a	n.a	n.a	n.a
DO2A	10.9-11.4	9.5-9.58	4.09-3.92	2.59-3.2	n.a.	n.a.	n.a	n.a
DO3A	10.5-12.0	9.08-9.50	4.35-4.5	3.50	2.0	n.a.	n.a	n.a
DOTA	11.2-11.7	9.70-9.73	4.44	4.49	4.14-4.34	2.35	n.a	n.a
DO2P	12.7	10.88	8.48	6.39	1.3	n.a.	n.a	n.a
DO3P	12.9	11.4	8.69	7.09	5.35	1.4	n.a	n.a
DOTP	12.8	11.44-12.6	8.90-9.28	7.71-8.05	5.99-6.09	5.13-5.20	1.8	1.3

Table 5.2. Stability constants of ligands in their different protonation states to Ca^{2+}. All values were from the NIST Critically Selected Stability Constants of Metal Complexes Database[28] unless otherwise stated. n.a: not available.

Ligand	log $\beta_{Ca(II):L}$	log $\beta_{Ca(II):LH}$	log $\beta_{Ca(II):LH2}$	log $\beta_{Ca(II):LH3}$	log $\beta'_{Ca(II):L}$
Alendronate[186]	7.58 ± 0.11	n.a.	n.a.	n.a.	n.a.
Cyclen	3.1	n.a.	n.a.	n.a.	n.a.
IDA	2.60	n.a.	n.a.	n.a.	n.a.
DO2A	7.2	n.a.	n.a.	n.a.	n.a.
DO3A	12.5	6.4 Kreher, PhD thesis[191]	n.a.	n.a.	5.71[a)]
DOTA	16.4-17	8.68 Delgado *et al.*[187]	n.a.	n.a	10.26 [a)]
DO2P[189]	9.0	8.5	n.a.	n.a.	n.a.
DO3P[184]	10.97	9.7	7.9	7.5	n.a.
DOTP[188]	11.12	n.a	n.a	n.a	n.a.

[a)] Calculated stability constant at pH 7.4 by Kumar *et al.*[185]

These log β values are lower than 10.96 but are still in a range useful for the purpose of developing an IMAC system. Interestingly the complexes [CaH_4DO3P] and [Ca H_5DO3P]$^+$ have not been shown to form even though the species [H_4DO3P]$^{2-}$ and [H_5DO3P]$^-$ have been shown to exist. Similarly DO2P only form[189] the complexes [$CaDO2P$]$^{2-}$ and [$CaHDO2P$]$^-$.

Kumar *et al*.[185] calculated a conditional stability constant log $\beta'_{Ca(II):DO3A}$ = 5.71 at pH 7.4 accounting for competition of endogenous metal ions and hydrogen ions at this pH. This value is also lower compared to log $\beta_{Ca(II):DO3A}$ = 12.5 of the unprotonated ligand[28] (Table 5.2).

By definition log β is an equilibrium constant. Values are consequently obtained under equilibrium conditions. Chromatography experiments are on the other hand dynamic processes. Reactions between the stationary and mobile phase (in the case of IMAC the immobilized metal-ligand chelate and the eluent) may not reach equilibrium since the mobile phase constantly passes through the stationary phase. Therefore, stability constants of the free metal-chelating complex only serve as a guide to how well metal ions are chelated in a chromatography system. Other parameters such as the rate of the metal-ligand association and dissociation reactions also need to be taken into consideration.

5.2 Aim

As discussed above, several Ca^{2+} chelates with free ligands under equilibrium conditions have been extensively studied. Their properties as IMAC ligands, attached to a solid support and subjected to a flow of a mobile phase, have not received the same attention. The aim of the studies in this Chapter was therefore to investigate the properties of a range of immobilized Ca^{2+} chelates under chromatographic conditions. Specifically the extent of Ca^{2+} loading on different immobilized ligands and any occurance of MIT were considered important parameters to investigate.

5.3 Results

5.3.1 Methodological considerations

5.3.1.1 Experimental setup

In order to investigate the different immobilized Ca^{2+} chelates under chromatographic conditions, all experiments were carried out as follows (see Materials and Methods Sections 2.4.4, 2.12 and 2.13 for details). The gels were packed in 1 mL Tricorn^{TM} columns. The columns were washed with 20 mL water, charged with 50 mM $Ca(NO_3)_2$ and washed with another 20 mL water on a bench pump. The charged columns were then washed with 20 mL mobile phase at a flow rate of 1 mL min^{-1} on an FPLC system. By mimicking the conditions

used when performing IMAC experiments as closely as possible the data obtained had direct relevance for the chromatographic conditions. Complete chromatography experiments were also performed.

The Ca^{2+} content remaining on the immobilized ligands was determined by digesting the entire gel in 5 M HCl and measuring the Ca^{2+} concentration in the resulting solution by atomic absorption spectrometry AAS. An alternative approach would involve incubating the metal loaded gels in buffers, measuring the Ca^{2+} in the buffer and calculate the remaining Ca^{2+} on the gel by mass balance. This approach has two advantages. It is less time consuming then the experiments described here. Secondly the gels are not digested in HCl and can therefore be reused. The disadvantage is that the experimental conditions are at equilibrium. This does not reflect a chromatography experiment where a mobile phase is flowing through the gel.

5.3.1.2 Buffer preparation

Buffers were chosen by their respective buffer ranges. N-Cyclohexyl-2-aminoethanesulfonic acid (CHES) was employed for prepating buffers of pH 9.5. Tris(hydroxymethyl)aminomethane (tris) was employed for pH 8.5. 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was used to prepare buffers of pH 7.5. 2-(N-morpholino)ethanesulfonic acid (MES) was used for preparing buffers of pH 5.5 and 6.5.

Alkali earth metals can form complexes with macrocyclic ligands[188; 192; 193]. In potentiometric titrations tetrabutylammonium salts are therefore commonly employed as background electrolytes since they complex ligands weakly[188; 192; 193]. In order to avoid introducing Na^+ or other cation with the potential of complexing with DO3P, the pH in Na^+-free buffers was therefore adjusted using tetrabutylammonium hydroxide (TBAOH). The Na^+ containing buffers were adjusted with NaOH yielding higher Na^+ concentrations. The Na^+ contents of the different buffer compositions are shown in Table 5.3.

Table 5.3. Total Na^+ content (mM) in 10 mM buffers supplemented with various concentrations of NaCl and pH adjusted using NaOH.

	10 mM NaCl	**50 mM NaCl**	**100 mM NaCl**	**300 mM NaCl**	**600 mM NaCl**
CHES pH 9.5	15.8	55.0	105	304	303
HEPES pH 7.5	15.4	54.6	104	303	303
MES pH 6.5	17.2	57.0	107	306	305
MES pH 5.5	13.2	52.4	102	302	301

5.3.1.3 AAS versus ICP

Inductively coupled plasma atomic emission spectrometry (ICP-AES) is a sensitive technique for measuring the concentration of a wide range of metals. Practically, there are two advantages of ICP over AAS. Firstly the detection limit is lower. Secondly several metals can be detected simultaneously without changing any instrumental settings. The method is therefore much faster when measuring more then one metal. ICP-AES was initially employed for measuring Ca^{2+} and Na^{+} in samples simultaneously. AAS was however judged sensitive enough for determining the amount of Ca^{2+} on Ca^{2+}-DO3P-Sepharose 6 F F.

Na^{+} is known to interfere with Ca^{2+} in AAS and enhance the observed signal[194]. Interference suppression was achieved using the method of Schinkel[195] by supplementing all samples with lanthanum and cesium to a final concentration of 1.334 g L^{-1}. The effect of employing interference suppression was tested by analyzing samples containing 10 mg L^{-1} Ca^{2+} and 10 mg L^{-1} Ca^{2+} supplemented with 138 mg L^{-1} (6 mM) NaCl. This concentration of Na^{+} corresponds to the amount of Na^{+} in 1 mL gel when buffers containing 300 mM NaCl was used. In the absence of interference suppression the NaCl containing sample gave a 5 % higher response for Ca^{2+} as compared to the NaCl free sample. In the presence of interference suppression (20 % v/v of 6.67 g L^{-1} lanthanum and 6.67 g L^{-1} cesium) the response in the NaCl containing sample was only encanced 1 %. This enhancement was not considered significant as the % RDS for points in the standard curve varied between 0.7 and 1 % at 8 to 20 g L^{-1} Ca^{2+}.

5.3.2 Supporting experimental section

5.3.2.1 Difference between ICP and AAS

Analysis of the same set of samples by AAS and ICP gave comparable results as shown in Figure 5.2. All experiments were carried out in triplicate. The following samples were analysed.

A) Unactivated Sepharose 6 F F.

B) Uncharged DO3P-Sepharose 6 F F.

C) Ca^{2+} charged DO3P-Sepharose 6 F F.

D) Ca^{2+} charged DO3P-Sepharose 6 F F washed with 300 mM NaCl, 10 mM MES, pH 5.5.

E) Ca^{2+} charged DO3P-Sepharose 6 F F washed with 100 mM EDTA, pH 8.0.

F) Ca^{2+} charged DO3P-Sepharose 6 F F after IMAC experiments using cell lysates of *E. Coli* BL-21 in 300 mM NaCl, 10 mM MES, pH 5.5.

As expected Ca^{2+} was removed by EDTA, a well known metal chelating compound[28] with log $\beta_{Ca:L}$ = 10.65 - 10.81 at an ionic strength of 0.1 M. 300 mM NaCl at pH 5.5 also displaced virtually all Ca^{2+} (Figure 5.2, sample D).

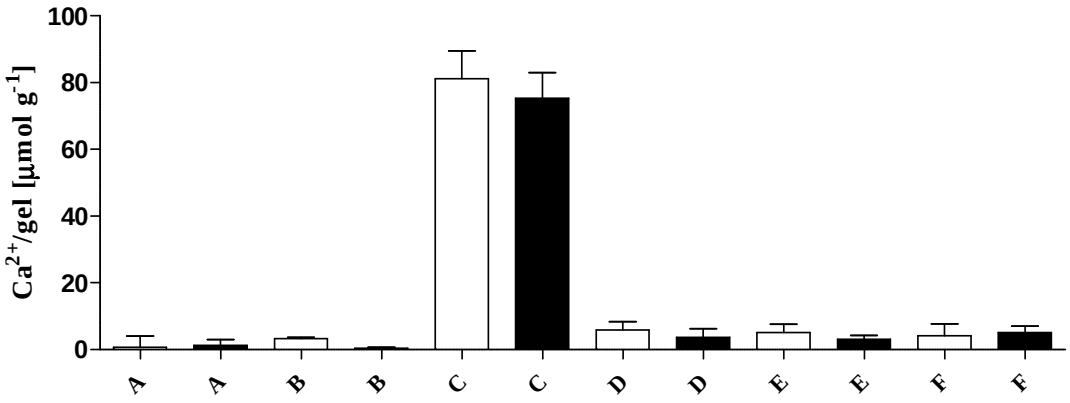


Figure 5.2. Comparison of samples analysed on AAS and ICP. All samples were in triplicate. Error bars represent standard deviations. Empty bars represent AAS samples and filled bars represent ICP samples. Samples A: Sepharose (unmodified), B: DO3P-Sepharose, C: Ca^{2+} charged DO3P-Sepharose washed with 20 mL water, D: Ca^{2+} charged DO3P-Sepharose washed with 20 mL 300 mM NaCl, 10 mM MES, pH 5.5, E: Ca^{2+} charged DO3P-Sepharose washed with 100 mM EDTA pH 8.0, F: Ca^{2+} charged DO3P-Sepharose after IMAC experiment using bacterial cell lysate of *E.coli* BL-21 as sample and 300 mM NaCl, 10 mM MES, pH 5.5 as mobile phase.

5.3.2.2 Standard curves and error analysis

Standard curves generated for ICP and AAS analyses of Ca^{2+} and Na^{+} are shown in Figure 5.3. In both cases the response was linearly dependent on the Ca^{2+} concentration over the concentration range employed with R^2 values above 0.99. Two sources of error were associated with obtaining accurate weights of the gels to be digested. Firstly, the tubes containing wet gels became static after freeze-drying which influence the behaviour of the balance, causing fluctuations in the readings. Secondly it was observed that the weight of salts remaining in the wet gels after each chromatography experiment contributed significantly to the net weight of the dried gel. The weight of remaining salts was in some cases equal to or even greater than the weight of the gel itself. The mean weight of ten samples of lyophilized gel corresponding to 1 mL wet gel containing no salt was 82.5 ± 5.4 mg. This value was used as a constant for all samples throughout.

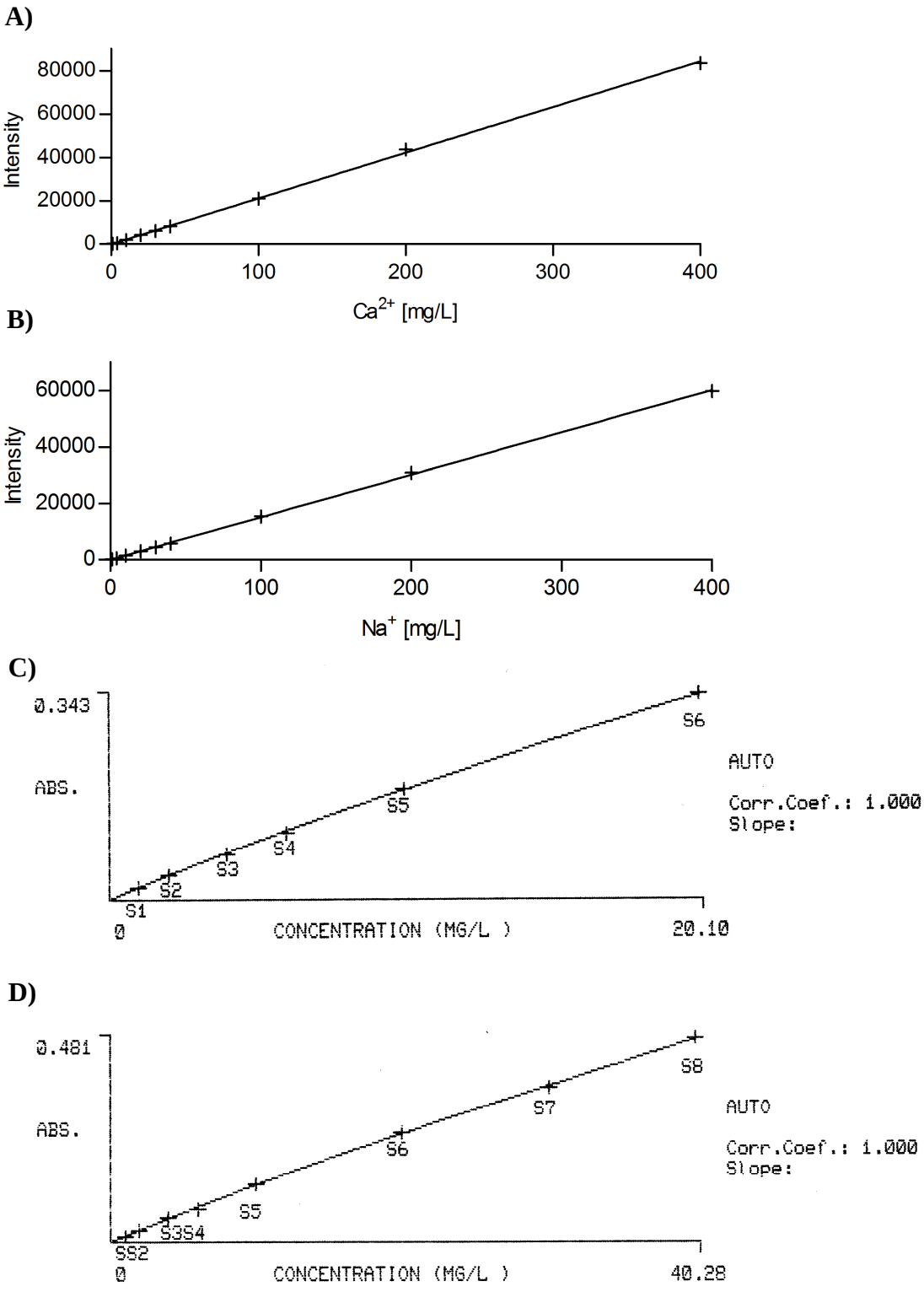


Figure 5.3. Standard curves for a) Ca^{2+} on ICP, b) Na^{+} on ICP, c) Ca^{2+} on AAS and d) Na^{+} on AAS.

5.3.3 Ca^{2+} charging of immobilized ligands

A suite of putative Ca^{2+} binding ligands were immobilized to Sepharose 6 F F and studied (Figure 5.1). The macrocyclic ligands DO2A, DO3A, DO2P and DO3P were chosen based on their high log $\beta_{Ca(II):L}$ values[28; 184; 189] of 7 – 12.5. The parent compound cyclen was included as a reference. The non-macrocyclic carboxylic acid containing ligand IDA was included as a comparison with DO2A and DO3A. IDA was also of interest as it previously had been employed as a ligand in Ca^{2+}-based IMAC[85; 86]. Alendronate was included since the ligand is a phosphonic acid containing non-macrocyclic molecule with[186] $\beta_{Ca(II):L}$ = 7.58. Ca^{2+} charged gels washed with water only were used as positive controls and gels not loaded with Ca^{2+} were used as negative controls. The calcium content of the gels was calculated using Equation 5.1 where Ca^{2+} refers to the signal obtained from the AAS or ICP experiment, V is the volume of sample, m_{gel} is the mass of freeze dried gel and Mw refers to the molar mass of the element being measured.

Equation 5.1
$$Ca^{2+}[mol \cdot g^{-1}] = \frac{Ca^{2+}[g \cdot L^{-1}] \cdot V[L]}{m_{gel}[g] \cdot Mw_{Ca^{2+}}[g \cdot mol^{-1}]}$$

The Ca^{2+} content of uncharged gels was between 0 and 4 µmol g^{-1} gel for all ligands investigated. The quote θ was defined as mol of bound Ca^{2+} per mol of immobilized ligand. θ was calculated according to Equation 5.2.

Equation 5.2
$$\theta = \frac{Ca^{2+} / gel\left[mol \cdot g^{-1}\right]}{Ligand / gel[mol \cdot g^{-1}]}$$

Ligand densities were independently determined by elemental nitrogen analysis (Dairy Tech. Services, Melbourne, Australia). The uncertainty in the values of the ligand densities was thought to originate mainly from weighing the gels, as discussed in Section 5.3.2.2. To account for errors associated with weighing the gels, and other unknown error sources, the RSD of the ligand densities was estimated to 0.15. The error of the quote θ was calculated by error propagation using Equation 5.3 where y_E, a_E and b_E are the errors associated with *Y*, *A* and *B*, respectively.

Equation 5.3
$$\frac{y_E}{Y} = \sqrt{\left(\frac{a_E}{A}\right)^2 + \left(\frac{b_E}{B}\right)^2}$$

For two batches of Ca^{2+}-DO3P-Sepharose 6 F F, θ was 130 ± 20 % and 95 ± 15 % (Table 5.4). These results indicate that one Ca^{2+} was immobilized per ligand for DO3P-Sepharose 6 F F. Immobilized alendronate, IDA, DO3A and DO2P had lower values of θ ranging between 35 ± 5 % and 65 ± 10 % (Table 5.4). Hence, it is evident that Ca^{2+} is chelated by these immobilized ligands. For DO2A-Sepharose 6 F F the obtained θ value was only 5 ± 2 % and the Ca^{2+} content was 7 ± 2 µmol g^{-1} gel (Table 5.4). The Ca^{2+} content of activated Sepahrose 6 F F was 9 ± 0.2 µmol g^{-1} gel (Table 5.4).

Table 5.4. Ca^{2+} content of gels charged with 50 mM $Ca(NO_3)_2$ and washed with water as described in detail in the Materials and Methods Section 2.4.4. The ligand density (µmol ligand g^{-1} gel) was determined by elemental analysis of nitrogen content by Dairy Tech. Services, Melbourne, Australia. The Ca^{2+} content of the gels was determined by AAS. The values are the mean of two measurements with standard deviation. θ was calculated as the quote of mol Ca^{2+} per g gel/mol ligand per g gel. The error of θ was calculated by error propagation using Equation 5.3.

Immobilized ligand or gel with batch code in brackets	**Ligand/gel [µmol g^{-1}]**	**Ca^{2+}/gel [µmol g^{-1}]**	**θ [%]**
Alendronate-Sepahrose (EC2-131)	194 ± 29	126 ± 1	65 ± 10
Cyclen-Sepharose 6 F F (EC5-63, EC5-64, EC5-71, EC5-72)	190* ± 28	30 ± 10	16 ± 6
IDA-Sepharose 6 F F (EC4-64)	298 ± 45	131 ± 9	44 ± 7
DO2A-Sepahrose (EC8-115)	139 ± 21	7 ± 2	5 ± 2
DO3A-Sepahrose (EC3-95)	178 ± 27	63 ± 2	35 ± 5
DO2P-Sepahrose (EC4-4, EC4-47)	184* ± 28	76 ± 7	41 ± 7
DO3P-Sepharose 6 F F (EC3-79)	130 ± 20	166 ± 1	130 ± 20
DO3P-Sepharose 6 F F (EC3-96)	58 ± 9	55 ± 2	95 ± 15
Sepharose 6 F F, activated (EC5-30)	0	9 ± 0.2	-
Sepharose 6 F F	0	0	-

* Batches were combined and the ligand density was calculated as the mean value.

Hence it is questionable if immobilized DO2A bound any Ca^{2+}. The only Ca^{2+} complex determined for DO2A is that of the fully deprotonated ligand[180] (log β = 7.8). According to the species distribution diagram of DO2A[196], the fully deprotonated form of the ligand only exists at pH > 9. The neutral form of the ligand with the two pendant arms carrying negative charges and two ring nitrogens carrying positive charges is predominant[196] at pH between 4 and 10. All loading was carried out by passing non pH adjusted (pH = 5.4), non buffered 50 mM $Ca(NO_3)_2$ solution through the packed column. Hence it is possible that no complex could be formed between Ca^{2+} and DO2A under these conditions. The θ value of

Cyclen-Sepharose 6 F F was 16 ± 6 % (Table 5.4) indicating that some Ca^{2+} bound to this ligand. The values of θ for DO3P-Sepharose 6 F F and DO3A-Sepharose 6 F F obtained in this thesis chapter were similar to results previously obtained in my supervisor's laboratory[191] with these ligands: 124 % for DO3P-Sepharose 6 F F and 72 % for DO3A-Sepharose 6 F F. Interestingly activated Sepharose 6 F F bound 9 ± 0.2 µmol Ca^{2+} g^{-1} while no binding was observed for nonactivated Sepharose 6 F F (Table 5.4). The θ values for the immobilized ligands could be arranged in the following order.

DO3P > alendronate > IDA ~ DO2P > DO3A > cyclen > DO2A

5.3.4 The effect of NaCl containing buffers at pH 7.5 on immobilized ligands to retain Ca^{2+}

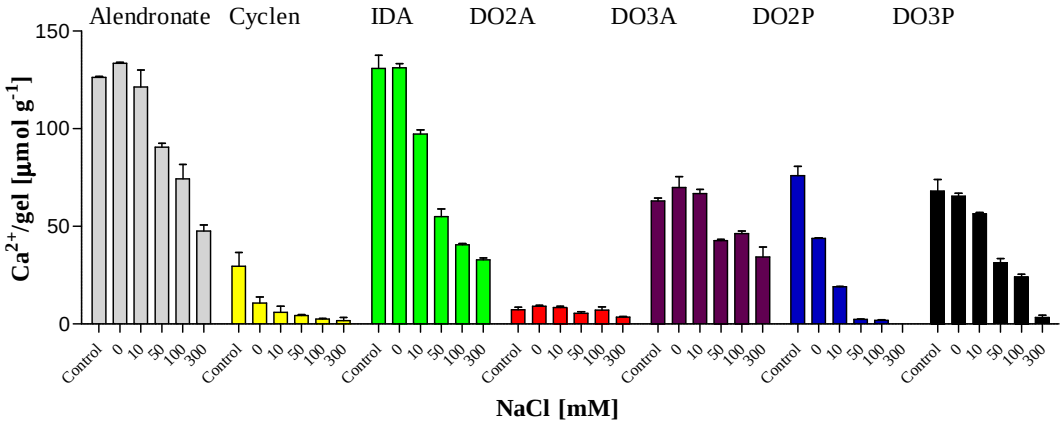


Figure 5.4. Remaining Ca^{2+} content on alendronate-Sepharose 6 F F (light grey), cyclen-Sepharose 6 F F (yellow), IDA-Sepharose 6 F F (green), DO2A-Sepharose 6 F F (red), DO3A-Sepharose 6 F F (purple), DO2P-Sepharose 6 F F (blue) and DO3P-Sepharose 6 F F (black). The gels were packed in 1 mL columns, charged with Ca^{2+} and washed with 20 mL 10 mM HEPES, pH 7.5 at 1 mL min^{-1} on an FPLC instrument. The NaCl concentration was 0, 10, 50, 100 and 300 mM. All measurements were performed in duplicates and mean values are shown. Error bars represent the standard deviation. The control experiments represent gels charged with Ca^{2+} but not washed with buffer.

To investigate the effect of NaCl on the ability of the immobilized ligands to retain Ca^{2+}, gels were charged with Ca^{2+} and washed with 10 mM HEPES, pH 7.5. The NaCl concentration of the buffers was varied between 0 and 300 mM. The results are shown in Figure 5.4. The Ca^{2+} remaining bound to the immobilized metal chelates after the excess Ca^{2+} had been washed off with water (Table 5.4) was included for comparison. Increased NaCl concentrations caused displacement of bound Ca^{2+} from all immobilized metal chelates, but to

different extents. The smallest extent of displacement was observed for Ca^{2+}-DO3A-Sepharose 6 F F (Figure 5.4). For this chelate no significant displacement was observed with NaCl-free 10 mM HEPES, pH 7.5 when 111 ± 13 % of the Ca^{2+} remained bound compared to the positive control (water wash). Buffer containing 10 mM NaCl left 106 ± 6 % Ca^{2+} bound to the gel. At 50 mM NaCl, 68 ± 3 % Ca^{2+} remained bound. Further increasing the NaCl concentration to 300 mM decreased the amount of bound Ca^{2+} to 55 ± 12 %.

A similar trend was observed for Ca^{2+}-alendronate-Sepharose 6 F F (Figure 5.4). 106 ± 1 % Ca^{2+} remained bound after a wash with NaCl free buffer. At 10 mM NaCl 96 ± 10 % Ca^{2+} remained bound. At 50 mM and 100 mM NaCl 72 ± 2 and 59 ± 8 % Ca^{2+} still remained bound. At 300 mM NaCl the corresponding number was 38 ± 3 %.

Also for Ca^{2+}-IDA-Sepharose 6 F F (Figure 5.4) no displacement was observed when the gel was washed with NaCl free buffer compared to the positive control (water wash). At 10 mM NaCl, however, only 74 ± 6 % Ca^{2+} remained bound. At 50 mM NaCl the corresponding number was 42 ± 5 %. At 300 mM NaCl only 25 ± 2 % Ca^{2+} remained bound to IDA-Sepharose 6 F F. Hence Ca^{2+}-alendronate-Sepharose 6 F F appeared to be slightly more resilient to NaCl as compared to IDA-sepahrose.

Ca^{2+}-DO3P-Sepharose 6 F F and Ca^{2+}-DO2P-Sepharose 6 F F showed significantly different behaviour (Figure 5.4). For Ca^{2+}-DO3P-Sepharose 6 F F 82 ± 13 % Ca^{2+} remained bound at 10 mM NaCl compared to the water wash. 49 ± 7 % Ca^{2+} remained bound at 50 mM NaCl. At 300 mM NaCl only 6 ± 4 % of the Ca^{2+} was still cbound (Figure 5.4). For Ca^{2+}-DO2P-Sepharose 6 F F, on the other hand, only 25 ± 2 % of the Ca^{2+} remained bound already at 10 mM NaCl. At 50 mM NaCl the corresponding number was 3 ± 1 % (Figure 5.4).

For Ca^{2+}-cyclen-Sepharose 6 F F (Figure 5.4) only 36 ± 19 % of the Ca^{2+} remained bound after the gel had been subjected to a wash with NaCl-free buffer. The majority of Ca^{2+} was displaced by 10 mM NaCl, when only 20 ± 16 % Ca^{2+} remained bound. No trend could be observed for DO2A-Sepharose 6 F F as no significant Ca^{2+} binding was evident for this ligand at any condition investigated (Figure 5.4). Based on the results in Figure 5.4, the ability of the Ca^{2+} chelates immobilized to Sepharose 6 F F to withstand NaCl in 10 mM HEPES, pH 7.5 could be ranked as follows.

Ca^{2+}-DO3A > Ca^{2+}-alendronate > Ca^{2+}-IDA > Ca^{2+}-DO3P > Ca^{2+}-DO2P > Ca^{2+}-cyclen > Ca^{2+}-DO2A.

5.3.5 The effect of NaCl and pH on Ca^{2+}-DO3P-Sepharose 6 F F and Ca^{2+}-DO3A-Sepharose 6 F F

Based on the low degree of Ca^{2+} displacement observed for Ca^{2+}-DO3A-Sepharose 6 F F (Figure 5.4) and the high θ value for Ca^{2+}-DO3P-Sepharose 6 F F (Table 5.4), these two immobilized metal chelates were further investigated. Experiments were carried out as

described above. Resins were packed in 1 mL columns and charged with Ca^{2+}, and excess Ca^{2+} was removed by washing with water. The resins were then washed with 20 mL buffer at 1 mL min^{-1} on an FPLC system. Both the pH and the NaCl concentration in the buffer compositions were varied. The buffers employed were 10 mM CHES pH 9.5, 10 mM Tris-HCl pH 8.5, 10 mM HEPES pH 7.5 and 10 mM MES pH 6.5 and pH 5.5. For Ca^{2+}-DO3A-Sepharose 6 F F the NaCl concentration was varied between 0 and 300 mM. For Ca^{2+}-DO3P-Sepharose 6 F F the NaCl concentration was varied between 0 and 600 mM. The Ca^{2+} content of the gels were subsequently analysed by AAS as described in the Materials and Methods Section 2.13. The results of the duplicate measurements are shown in Table 5.5 as μmol Ca^{2+} per g dry gel. The mean values of the two measurements are graphically represented in Figure 5.5. The ligand density was 178 ± 27 μmol g^{-1} for DO3A-Sepharose 6 F F and 58 ± 9 μmol g^{-1} for DO3P-Sepharose 6 F F (Table 5.4). The higher θ value of 95 ± 15 % for Ca^{2+}-DO3P-Sepharose 6 F F compared to 35 ± 5 % for Ca^{2+}-DO3A-Sepharose 6 F F (Table 5.4) resulted in the Ca^{2+} content of the two immobilized metal chelates being similar as seen in Figure 5.5.

For Ca^{2+}-DO3P-Sepharose 6 F F displacement of bound Ca^{2+} was observed as a function of both NaCl and pH. Relative to the positive control (Ca^{2+} charging of the gel and wash with only water), 101 ± 13 % of the Ca^{2+} remained bound when the gel was washed with 10 mM CHES, pH 9.5. After washing the gels with 10 mM Tris-HCl, pH 8.5 and 10 mM HEPES, pH 7.5, 91 ± 16 and 95 ± 16 % Ca^{2+} remained bound respectively. Hence no significant displacement of bound Ca^{2+} occurred from Ca^{2+}-DO3P-Sepharose 6 F F when the resin was washed with NaCl free buffer at pH ≥ 7.5. After a wash with 10 mM MES, pH of 6.5, 63 ± 8 % of the Ca^{2+} remained bound (Figure 5.5A). The effect of further decreasing pH was even more pronounced. After a wash with 10 mM MES pH 5.5 only 11 ± 2 % of the Ca^{2+} remained bound. Hence bound Ca^{2+} was displaced from Ca^{2+}-DO3P-Sepharose 6 F F by buffers of pH below 7.5.

Bound Ca^{2+} was displaced from Ca^{2+}-DO3P-Sepharose 6 F F by NaCl at all pH values investigated. After washes with buffers of pH 9.5, 8.5 and 7.5 containing 10 mM NaCl 72 ± 10 %, 77 ± 10 % and 82 ± 13 Ca^{2+} remained bound to Ca^{2+}-DO3P-Sepharose 6 F F. At 10 mM NaCl, pH 6.5 only 24 ± 3 % Ca^{2+} remained bound. As the degree of displacement was high at pH 5.5 no experiments with NaCl supplemented buffers were performed at this pH.

At 50 mM NaCl significant displacement of bound Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F occurred (Figure 5.5A). At 50 mM NaCl pH 9.5, 8.5 and 7.5, 41 ± 15 %, 53 ± 8 % and 49 ± 7 % Ca^{2+} respectively remained bound. At 50 mM NaCl, pH 6.5 no Ca^{2+} was detected, indicating that all or nearly all Ca^{2+} was displaced.

As the NaCl concentration was increased to 100 mM, the degree of displacement of bound Ca^{2+} increased further (Figure 5.5A). At 100 mM NaCl, pH 9.5, 8.5 and 7.5, 26 ± 4 %, 35 ± 5 % and 38 ± 6 % of the Ca^{2+} remained bound to DO3P-Sepharose 6 F F. Hence the degree of Ca^{2+} displacement appeared to be independent of pH above pH 7.5 at 100 mM NaCl.

Table 5.5. Remaining Ca^{2+} content of 1 mL Ca^{2+} charged DO3P-Sepharose 6 F F (A) and DO3A-Sepharose 6 F F (B) packed in a column and washed with 20 mL buffered solutions at 1 mL min^{-1} on an FPLC instrument. Buffers were 10 mM CHES pH 9.5, 10 mM Tris-HCl pH 8.5, 10 mM HEPES pH 7.5, 10 mM MES pH 6.5 and 10 mM MES pH 5.5. NaCl concentrations were 0, 10, 50, 100, 300 and 600 mM for DO3P-Sepharose 6 F F and 0, 10, 50, 100 and 300 for DO3A-Sepharose 6 F F. The ligand density was 58 ± 9 µmol g^{-1} dry gel for DO3P-Sepharose 6 F F and 178 ± 27 µmol g^{-1} dry gel for DO3A-Sepharose 6 F F.

	DO3P-Sepharose-6 F F						DO3A-Sepharose 6 F F				
Positive control	61.3 74.0						61.4 64.6				
NaCl / pH	0	10	50	100	300	600	0	10	50	100	300
9.5	69.5	50.9	20.9	17.9	23.3	11.2	79.7	64.6	35.7	35.0	36.7
	67.0	46.9	34.5	17.9	16.4	18.2	71.9	58.5	39.7	41.1	46.0
8.5	67.1	51.0	37.4	24.8	15.7	15.4	70.4	69.2	46.5	51.1	44.5
	56.5	53.4	34.4	22.4	25.0	13.6	72.7	66.4	39.9	49.0	40.6
7.5	69.5	52.5	31.8	24.7	6.1	5.0	75.4	68.9	43.4	44.8	29.3
	59.7	58.7	34.4	27.3	2.6	6.7	64.6	64.7	42.0	47.7	39.5
6.5	42.2	15.6	-0.9	2.9	2.0	3.3	77.8	67.1	55.7	52.4	46.3
	42.7	17.0	0.8	6.4	4.5	3.8	73.9	67.7	55.3	48.8	37.2
5.5	7.7	n.d.	n.d.	n.d.	n.d.	n.d.	71.5	53.9	45.0	42.1	33.3
	7.4						80.4	58.7	37.0	44.0	33.0

At 100 mM NaCl, pH 6.5 only 7 ± 4 % Ca^{2+} remained bound to DO3P-Sepharose 6 F F. Additional displacement was observed when the NaCl concentration was increased to 300 and 600 mM. The degree of displacement was however not linearly dependent on the NaCl concentration. The remaining bound Ca^{2+} rather appeared to reach a plateau (Figure 5.5A). At 300 mM NaCl, pH 8.5 and 9.5, 30 ± 10 % and 29 ± 8 % Ca^{2+} respectively remained bound. At 600 mM NaCl, pH 8.5 and 9.5, 21 ± 3 % and 21 ± 8 % Ca^{2+} still remained bound. At 300 and 600 mM NaCl, pH 7.5 only 6 ± 4 % and 8 ± 2 % Ca^{2+}

respectively remained bound to DO3P-Sepharose 6 F F. Hence the pH effect was more pronounced at NaCl concentrations at 300 mM and above. At 300 and 600 mM NaCl pH 6.5 only 5 ± 3 and 5 ± 1 % Ca^{2+} respectively remained bound to DO3P-Sepharose 6 F F.

For Ca^{2+}-DO3A-Sepharose 6 F F no displacement of bound Ca^{2+} was observed as a function of pH in the investigated range pH 5.5 – 9.5 (Figure 5.5B). When the concentrations of NaCl in the buffers were increased, the same trend of Ca^{2+} displacement previously obtained at pH 7.5 (Section 5.3.4) was observed for each pH investigated. For NaCl free buffers between 106 ± 6 and 121 ± 10 % Ca^{2+} remained bound compared to the positive control (water wash). Thus, no displacement of bound Ca^{2+} occurred from Ca^{2+}-DO3A-Sepharose 6 F F by pH in the range of pH 5.5 to 9.5. At 10 mM NaCl between 89 ± 6 and 107 ± 5 % Ca^{2+} remained bound to DO3A-Sepharose 6 F F. When the NaCl concentration was increased to 50 mM between 60 ± 5 % and 88 ± 3 % Ca^{2+} remained bound. At 100 mM NaCl no significant additional displacement occurred. Between 60 ± 7 % and 80 ± 5 % Ca^{2+} remained bound to DO3A-Sepharose 6 F F.

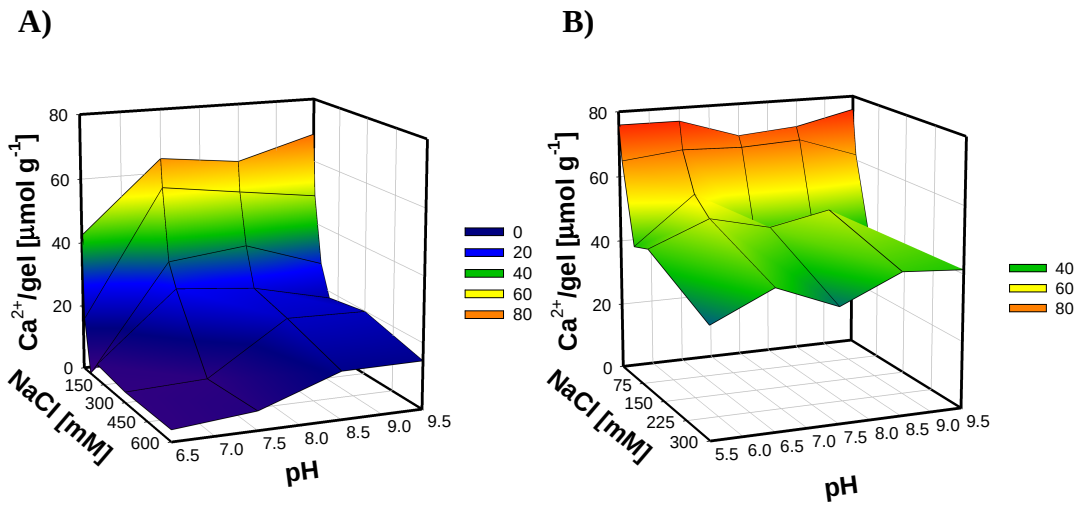


Figure 5.5. Remaining Ca^{2+} content of 1 mL Ca^{2+} charged **A)** DO3P-Sepharose 6 F F and **B)** DO3A-Sepharose 6 F F packed in a column and washed with 20 mL buffered solutions at 1 mL min^{-1} on an FPLC instrument. Buffers were 10 mM CHES pH 9.5, 10 mM Tris-HCl pH 8.5, 10 mM HEPES pH 7.5, 10 mM MES pH 6.5 and 10 mM MES pH 5.5. NaCl concentrations were 0, 10, 50, 100, 300 and 600 mM for DO3P-Sepharose 6 F F and 0, 10, 50, 100 and 300 for DO3A-Sepharose 6 F F. All measurements were performed in duplicates and mean values are shown. The ligand density was 58 ± 9 µmol g^{-1} dry gel for DO3P-Sepharose 6 F F and 178 ± 27 µmol g^{-1} dry gel for DO3A-Sepharose 6 F F.

At 300 mM NaCl between 52 ± 2 % and 68 ± 8 % Ca^{2+} remained bound. Hence the displacement of Ca^{2+} as a function of the NaCl concentration in the buffers was lower for DO3A-Sepharose 6 F F as compared to DO3P-Sepharose 6 F F even at pH 8.5 and 9.5 where the pH did not contribute to the displacement of bound Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F (Figure 5.5).

The θ value of 95 ± 15 % for the batch of Ca^{2+}-DO3P-Sepharose 6 F F used in these experiments was significantly higher compared to the 35 ± 5 % for Ca^{2+}-DO3A-Sepharose 6 F F (Table 5.4). This difference in θ values could contribute to the observed difference in displacement of Ca^{2+} from the two immobilized ligands. Relative to the ligand densities of the two gels, 35 ± 13 % Ca^{2+} remained bound to DO3P-Sepharose 6 F F after the gel had been washed with 300 mM NaCl, pH 8.5. The corresponding value for DO3A-Sepharose 6 F F was 24 ± 4 %. Hence the remaining Ca^{2+} on the ligands was similar relative to their ligand densities after a wash with buffer at pH 8.5 containing 300 mM NaCl.

5.3.6 The effect of different salts on Ca^{2+}-DO3P-Sepharose 6 F F

Aiming to find conditions under which the degree of displacement of bound Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F would be lower then observed in Section 5.3.4 and 5.3.5, a number of different salts were investigated as alternatives to NaCl. Salts were chosen so that the effect of both the anions and the cations could be studied. Toxic salts were excluded. All salts were at a concentration of 300 mM in 10 mM HEPES, pH 7.5. This condition was chosen since the majority of bound Ca^{2+} was displaced from Ca^{2+}-DO3P-Sepharose 6 F F when NaCl was used (Figure 5.5A). The pH 7.5 is also close to a physiological value, which is consistent with the overall goal of designing an IMAC system for the purification of proteins. The results are shown in Figure 5.6. As expected the salt free buffer (10 mM HEPES, pH 7.5) did not displace Ca^{2+} as compared to the positive control (Ca^{2+} charged gel washed with water).

All of the salts investigated displaced Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F, but to different extents. All salts containing divalent ions, $Na_2S_2O_3$, Na_2SO_4, $MgSO_4$, $(NH_4)_2SO_4$, $MgCl_2$ and K_2SO_4, displaced nearly all bound Ca^{2+} (Figure 5.6). The degree of displacement was lower for solutions of salts composed of monovalent ions (Figure 5.6). After a wash with buffers containing the salts NaCl, KCl, CsCl, $NaNO_3$, KNO_3, NH_4Cl, ~ 20 % of the Ca^{2+} remained bound to Ca^{2+}-DO3P-Sepharose 6 F F. After a wash with NaI, 37 ± 1 % of the Ca^{2+} remained bound to DO3P-Sepharose 6 F F. This observation could be explained by the difference in ionic strength *I* between the salts. *I* can be calculated using Equation 5.3 where c_i is the molar concentration of any ion *i* and z_i is the formal charge of the same ion.

Equation 5.3 $$I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2$$

Disregarding any contributions from the 10 mM HEPES buffer, I = 0.3 M for all 300 mM monovalent salt solutions. For the 300 mM salt solutions containing one divalent and two monovalent ions ($MgCl_2$, Na_2SO_4, K_2SO_4, $Na_2S_2O_3$), I = 0.9 M. For 300 mM $MgSO_4$ containing two divalent ions, I = 1.2 M. Hence, the displacement of bound Ca^{2+} from DO3P-Sepharose 6 F F by different salt solutons at pH 7.5 could be correlated to the ionic strength of the buffer.

The displacement of bound Ca^{2+} from Ca^{2+}-Sepharose 6 F F by different salt solutions can also be correlated to the charge of the cation. Salts composed of monovalent cations displace ~ 80 % of the bound Ca^{2+} while salts composed of divalent cations displace virtually all bound Ca^{2+} (Figure 5.6). For the salts composed of monovalent anions and cations, varying the cations did not significantly alter the displacement of Ca^{2+}.

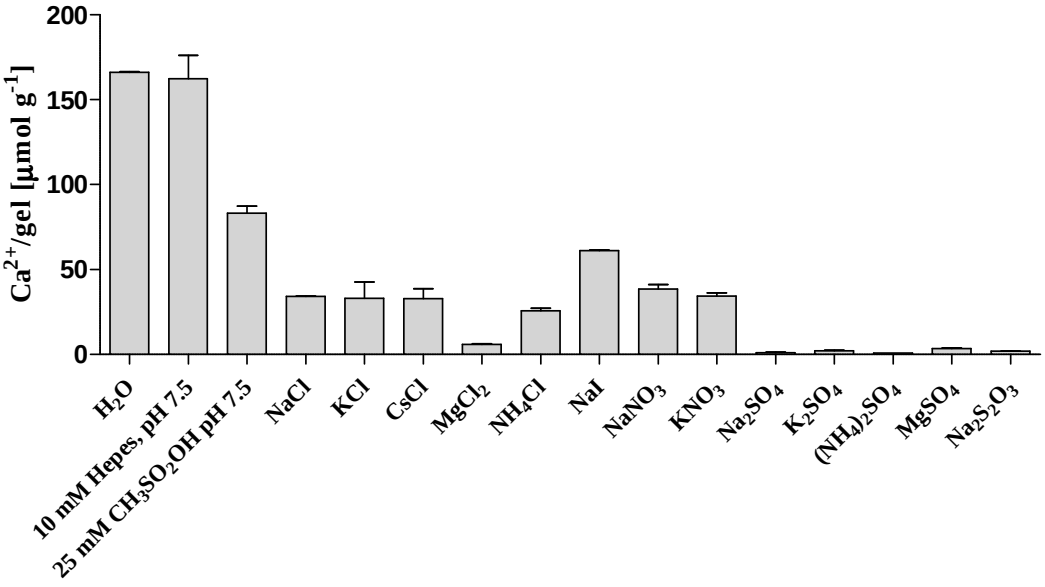


Figure 5.6. Remaining Ca^{2+} content of 1 mL Ca^{2+} charged DO3P-Sepharose 6 F F washed with 20 mL solution at 1 mL min^{-1}. The concentrations of all salts investigated were 300 mM and all solutions except methanesulfonate were buffered using 10 mM HEPES, pH 7.5. All measurements were performed in duplicates. Mean values are shown and the error bars represent the standard deviation calculated from two values. The ligand density of the DO3P-Sepharose gel was 130 µmol g^{-1} dry gel.

$MgCl_2$, composed of a divalent cation and monovalent anions, displaced slightly less Ca^{2+} as compared to the salts composed of both a divalent cation and anion (Figure 5.6). This

could suggest that the valency of the anion also play a role in the displacement. The log $\beta_{Mg:L}$ = 8.72 was determined by Kreher et al[184].

This value is similar to log $\beta_{Ca:L}$ = 10.97 determined by the same authors[184] and could explain the displacement of Ca^{2+} from DO3P-Sepharose 6 F F by Mg^{2+}-containing salts. Critical log β values of other metal complexes with DO3P have not been determined.

NaCl, KCl and CsCl displaced similar amounts of bound Ca^{2+} from DO3P-Sepharose 6 F F (Figure 5.6). Hence no trends in displacement could be observed between different monovalent cations. The displacement of bound Ca^{2+} was however lower for NaI compared to NaCl and $NaNO_3$, as mentioned above. As no other salts containing iodide anions were investigated no conclusions can be drawn regarding the role of the anion in the displacement of bound Ca^{2+} from DO3P-Sepharose 6 F F.

5.3.7 Ca^{2+} displacement by biological molecules

The work in this thesis aimed at contributing to the development of a Ca^{2+}-based IMAC system for purification of tagged proteins from complex protein mixtures such as bacterial cell lysates. Determining if Ca^{2+} is displaced from DO3P-Sepharose 6 F F by the putative Ca^{2+} binding tag HIT2 or by other components in the cell lysates was therefore important. A series of experiments were therefore performed where different protein mixtures were passed through packed columns and the Ca^{2+} remaining on the gel was measured (Figure 5.7).

Lysates of *E. coli* BL-21 bacterial cells and lysates of *E. coli* BL-21 expressing HIT2-NT1A-GFP were used as samples. The samples were used either after removal of cell debris by centrifugation or after centrifugation followed by dialysis with a molecular weight cutoff of 6-8000 Da to remove low molecular weight components. IMAC experiments were then performed employing either isocratic conditions or a linear gradient. In the isocratic experiments the mobile phase was 10 mM CHES, pH 9.5 or 10 mM CHES, 300 mM NaCl, pH 9.5. In the linear gradient experiments, samples were loaded in 10 mM CHES, pH 9.5 and elution was achieved with 10 mM MES, 300 mM NaCl, pH 5.5.

At pH 9.5 and under salt free conditions, no displacement of Ca^{2+} occurred when cell lysate of *E. coli* BL-21 were used as sample (Figure 5.7). The same behaviour was observed for *E. coli* BL-21 cell lysate containing HIT2-NT1A-GFP (Figure 5.7). When the IMAC experiments were performed using 300 mM NaCl in the buffers, significant displacement was however observed. It can therefore be concluded that the putative Ca^{2+} binding tag HIT2 and other components in cell lysate of *E. coli* BL-21 do not displace bound Ca^{2+} from DO3P-Sepharose 6 F F. The only observed displacement of bound Ca^{2+} from the gels can be attributed to the mobile phase conditions with displacement caused by low pH and the presence of salt, as observed in Section 5.3.5.

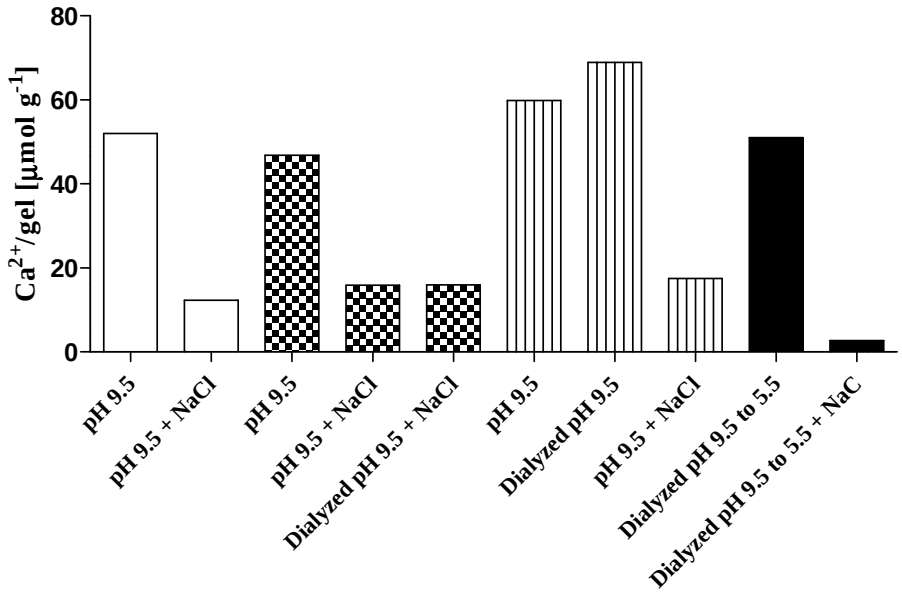


Figure 5.7. Remaining Ca^{2+} content of 1 mL Ca^{2+} charged DO3P-Sepharose subjected to FPLC experiments with and without biological samples. White bars: Wash experiments without protein. Dotted bars: FPLC experiments with *E. coli* BL-21 cell lysate not expressing tagged protein. Striped bars: FPLC experiments with *E. coli* BL-21 cell lysate expressing HIT2-NT1A-GFP. Black bars: FPLC experiments with *E. coli* BL-21 cell lysate expressing HIT2-NT1A-GFP where the elution buffer was adjusted to pH 5.5. Solutions at pH 9.5 were buffered with 10 mM CHES. Solutions at pH 5.5 were buffered with 10 mM MES. NaCl refers to a content of 300 mM NaCl in the buffer. All values are single measurements.

5.4 Discussion

5.4.1 Correlating stability constants with Ca^{2+} binding to immobilized ligands

The trends of log $\beta_{Ca:L}$ for the free ligands (Table 5.2), the θ values of the corresponding immobilized metal chelates (Table 5.4) and the extent of displacement of bound Ca^{2+} from the immobilized metal chelates by NaCl at pH 7.5 (Figure 5.4) are summarized in Figure 5.8. For alendronate a relatively high[186] log $\beta_{Ca:L}$ of 7.58 ± 0.11 corresponded to a θ value of 65 ± 10 % (Table 5.4) and a relatively low degree of Ca^{2+} displacement by NaCl at pH 7.5 (Figure 5.4). Cyclen with a low[28] log $\beta_{Ca:L}$ of 3.1 also had a low θ value of 16 ± 6 % (Table 5.4) and the bound Ca^{2+} was displaced even by 10 mM

HEPES, pH 7.5 containing no NaCl (Figure 5.4). For these two ligands their log $\beta_{Ca:L}$ values appear to correlate with the extent of metal loading and the amount of Ca^{2+} remaining bound after washes with NaCl in 10 mM HEPES, pH 7.5.

DO3A with a high[28] log $\beta_{Ca:L}$ of 12.5 had a low θ value of 35 ± 5 % but the bound Ca^{2+} was not displaced to a significant extent by NaCl (Figure 5.5B). For DO3P the high[191] log $\beta_{Ca:L}$ of 10.96 correlate to the highest θ values (95 ± 15 % and 130 ± 20 %) observed with any of the immobilized ligands (Table 5.4). The displacement of bound Ca^{2+} can be classified as intermediate with low degree of displacement at 10 mM NaCl, 10 mM HEPES, pH 7.5 but complete displacement at 300 mM NaCl, 10 mM HEPES, pH 7.5 (Figure 5.4). Interestingly IDA with a low[28] log $\beta_{Ca:L}$ of only 2.60 had a θ value of 44 ± 7 % (Table 5.4) but was relatively resilient to NaCl in 10 mM HEPES, pH 7.5 (Figure 5.4).

For DO2P the relatively high[189] log $\beta_{Ca:L}$ of 9.0 and θ of 41 ± 7 % (Tables 5.2 and 5.4) did not correlate to resilance against NaCl. Even 10 mM HEPES, pH 7.5 without NaCl displaced about half of the bound Ca^{2+} from Ca^{2+}-DO2P-Sepharose 6 F F (Figure 5.4).

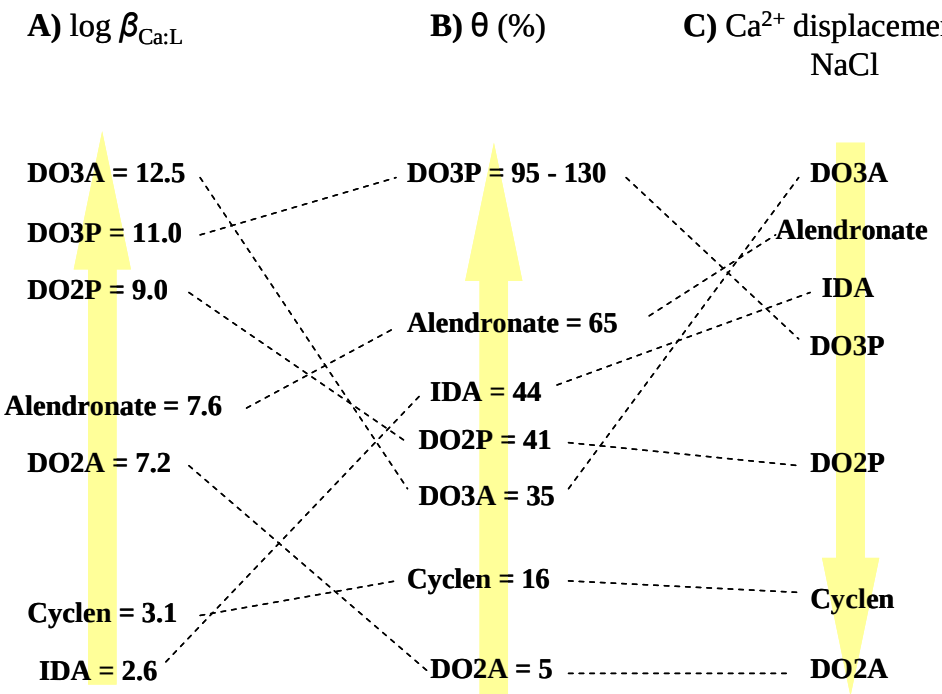


Figure 5.8. Trends of **A)** log $\beta_{Ca:L}$ of free ligands (data from Table 5.2); **B)** θ value (mol bound Ca^{2+} per mol ligand expressed as percentage, Table 5.3); **C)** the extent of displacement of bound Ca^{2+} from immobilized Ca^{2+} chelates when packed in a 1 mL column and washed with NaCl containing buffers (pH 7.5).

DO2A with[28] a log $\beta_{Ca:L}$ of 7.2 bound no Ca^{2+} (θ = 5 ± 2 %, Table 5.4). Alternatively the Ca^{2+} bound so weakly that it was displaced when the excess Ca^{2+} was washed off with water.

Hence DO2A and DO2P can be considered unsuitable as potential ligands in a Ca^{2+} based IMAC system. Hence no clear trend appear to exist between the log $\beta_{Ca:L}$ value for the free ligand, the θ value and the resilience of chelated Ca^{2+} to NaCl for the ligands investigated here.

5.4.2 Displacement of bound Ca^{2+} by NaCl and other salts

Displacement of Ca^{2+} by NaCl was observed to different extents for all the different immobilized metal chelates investigated (Figure 5.4). The displacement could be caused by formation of $CaCl_2$. However, since $CaCl_2$ is unstable[28] (log β = 0.2) compared to Ca^{2+}-alendronate[186] (log $\beta_{Ca:L}$ = 7.58), Ca^{2+}-DO3A[28] (log $\beta_{Ca:L}$ = 12.5) and Ca^{2+}-DO3P[184] (log $\beta_{Ca:L}$ = 10.97) this does not appear likely.

Na^+ could also compete with Ca^{2+} for the binging sites of the ligands. Ca^{2+} and Na^+ have effective ionic radii[197] of 1.00 Å and 1.02 Å respectively. Steric constraints can therefore not be assumed to hinder Na^+ binding at Ca^{2+} binding sites on a ligand. Moreover, alkali earth metals are known to bind to macrocyclic ligands[187; 193]. Kumar *et al*. observed binding of Na^+ and K^+ to DOTA, the latter metal ion binding more weakly[192].

Ca^{2+} would however be expected to have a higher affinity to all ligands as compared to Na^+ due to the higher charge density. This is indeed true for DO3A with[28] log $\beta_{Ca:L}$ = 12.5 >> log $\beta_{Na:L}$ = 2.2 and for IDA with log $\beta_{Ca:L}$ = 2.6 > log $\beta_{Na:L}$ = 0.36. Critically selected values of log $\beta_{Na:L}$ are not available for alendronate, DO3P, DO2A, DO2P and cyclen. Based on log $\beta_{Ca:L}$ and log $\beta_{Na:L}$ values of a selection of linear and macrocyclic ligands with phosphonic and/or carboxylic acid pendant arms (Table 5.6) it appear to be a trend that log $\beta_{Ca:L}$ is greater than log $\beta_{Na:L}$ for all ligands investigated. It is therefore reasonable to assume that $\beta_{Ca:L}$ > log $\beta_{Na:L}$ for alendronate, DO3P, DO2P and cyclen. The displacement of bound Ca^{2+} observed herein for ligands with high log $\beta_{Ca:L}$ values is therefore probably not due to the Ca^{2+}-complex being less stable compared to the corresponding Na^+ complex ($\beta_{Ca:L}$ < log $\beta_{Na:L}$).

A wide range of chelating ligands have been developed for use as ion exchange resins[198]. IDA remains to be the most commonly used[198]. Ion exchange is a reversible reaction and the equilibria can therefore be treated by the law of mass action[199]. Indeed cation exchange systems can be regenerated with highly concentrated NaCl solutions[200] even if the ions bound to the ion exchange resin have a much higher affinity to the immobilized ligand than Na^+. As the degree of displacement of bound Ca^{2+} increased with the NaCl concentration, all immobilized ligands investigated appear to behave as cation exchangers, but to various extents. Some Ca^{2+} still remained bound even after washes with solutions containing 300 mM NaCl.

Table 5.6. Comparison of stability constants for metal chelates formed with Ca^{2+} and Na^{+}. All values were from the NIST Critically Selected Stability Constants of Metal Complexes Database[28] unless otherwise stated.

Ligand	$\log \beta_{Ca:L}$	$\log \beta_{Na:L}$
N'''-(2-Hydroxypropyl)-1,4,7,10-tetraazacyclododecane-N,N',N''-triacetic acid (HP-DOTA)	14.4	3.19
N'''-(2-Hydroxypropyl)-1,4,7,10-tetraazacyclododecane-N,N',N''-tris(hydroxymethylacetic acid) (HP-DO3A-3HM)	11.09	2.5
1,4,7,10-Tetrakis(2-methoxyethyl)- 1,4,7,10-tetraazacyclododecane	5.47	2.2
EDTA	10.65	1.86
DO3A	12.5	2.2
1,4,7,10-Tetraazacyclododecane-N,N',N''-tris(hydroxymethylacetic acid) DO3A-3HM	8.84	2.58
IDA	2.6	0.36
Iminobis(methylenephosphonic acid) IDP	3.84	3.2
Phosphonoacetic acid	3.18	1.0
Methylphosphonic acid	1.64	0.5
Hydroxymethylphosphonic acid	1.68	0.61
Methylenediphosphonic acid	6.00	1.1
Ethane-1,1-diphosphonic acid	5.2	1.5
Propane-2,2-diphosphonic acid	6.3	2.08
Difluoromethylenediphosphonic acid	4.23	1.3
Dichloromethylenediphosphonic acid	4.49	1.4
1-Hydroxyethane-1,1-diphosphonic acid	6.1	2.07

As mentioned in the Introduction, metal ion leakage also occurs in borderline metal IMAC systems with ligands like IDA and NTA[37; 38; 39]. The extent of displacement of bound Ni^{2+} by NaCl is however low. The buffers employed in Chapter 3 with Ni^{2+}-NTA-Agarose, Ni^{2+}-tacn-Sepharose 6 F F, Cu^{2+}-tacn-Sepharose 6 F F, Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F and Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F contained 650 mM NaCl. Since the above resins are white when no metal is bound and blue when metal is bound, the occurance of significant displacement of Ni^{2+} or Cu^{2+} can be visually observed. No change in colour was however observed even after > 20 mL buffer was pumped through 1 mL resin packed in a column at 1 mL min^{-1}. Furthermore, the recommended NaCl content of buffers used with pre-packed HisTrap™ columns is 500 mM according to the manufacturer's instructions (available at

http://www.gelifesciences.co.jp/tech_support/manual/pdf/71502769.pdf on 2010-06-04). The log β values of borderline metal complexes[28] (log $\beta_{Ni:NTA}$ = 11.51 and log $\beta_{Ni:IDA}$ = 8.3) on the one hand and the Ca^{2+} chelates investigated here[28; 191] (log $\beta_{Ca:DO3P}$ = 11 and log $\beta_{Ca:DO3A}$ = 12.5) on the other hand are in the same range. The observed difference in displacement of bound Ca^{2+} observed in this Chapter and the displacement of bound Ni^{2+} from immobilized metal chelates can therefore not be explained only in terms of the log β values of the metal chelates. It is therefore reasonable to assume that inherent properties of the metal ion will determine the stability of the metal chelate.

5.4.3 The effect of pH on immobilized Ca^{2+}-chelates

Similar experiments as those described here were performed by Zachariou *et al*. with IDA and OPS immobilized onto Sepharose™ CL-4B. Ca^{2+} was found to be almost completely displaced from both IDA-Sepharose™ CL-4B and OPS-Sepharose™ CL-4B when the gels were incubated with 50 mM acetic acid, 0.1 M KNO_3, pH 4.0. Displacement of Ca^{2+} by low pH solutions therefore appears to be a phenomenon observed for several potential Ca^{2+} IMAC ligands. In this Chapter Ca^{2+}-IDA-Sepharose 6 F F with a ligand density of 298 μmol g^{-1} was packed in a column, charged with Ca^{2+} and washed with 20 mL 300 mM NaCl, 10 mM MES pH 5.5 at 1 mL min^{-1}. After this treatment only 2.5 ± 0.3 μmol Ca^{2+}/g dry gel remained bound. Regeneration of cation exchangers can be achieved by acidic washes[200]. For instance, the instruction manual for Chelex® (IDA-Polystyrenedivinylbenzene), available from BioRad (online at http://www.bio-rad.com/webmaster/pdfs/9184_Chelex.PDF on 2010-06-04) states that bound divalent cations can be eluted from IDA by dilute acidic solutions.

As discussed in Section 5.1.2 DO3P has several protonation states with different affinities to Ca^{2+}. At pH 6 - 9 the predominant complexes[184] are log $\beta_{Ca(II):DO3PH}$ = 9.7, log $\beta_{Ca(II):DO3PH2}$ = 7.9 and log $\beta_{Ca(II):DO3PH3}$ = 7.5. This can explain that no displacement of Ca^{2+} occur from Ca^{2+}-DO3P-Sepharose 6 F F at pH 7.5 and above. The log $\beta_{Ca(II):DO3PH4}$ and $\beta_{Ca(II):DO3PH5}$ were however not determined. Hence these complexes might not form in solution. An increasing concentration of these non Ca^{2+} chelating DO3P species at lower pH could explain the observed displacement of Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F at pH < 7.5. These results suggest that the properties of the ligand DO3P are not significantly altered when immobilized onto Sepharose 6 F F.

It is interesting to note that no pH dependent Ca^{2+} displacement occurred with Ca^{2+}-DO3A-Sepharose 6 F F at 5.5 < pH < 9.5. This can be attributed to the protonation state of the free ligand. The two first protonation constants[28] log K_1 = 10.5 - 12, log K_2 = 9.08 - 9.5 correspond to the protonation of two ring nitrogens[192]. The third and fourth protonation constants[28] (log K_3 = 4.35 - 4.5, log K_4 = 3.5 correspond to the protonation of the carboxylic acid side chains[192]. The species distribution diagram previously calculated in my supervisor's

laboratory[191] shows that the species H_2DO3A^- is the dominating species at pH between 5 and 9. The species CaH_2DO3A^+ was however not shown to form but rather complexes[191] CaHDO3A and $CaDO3A^-$. If the species CaHDO3A exist at pH between 5.5 and 9.5 then no Ca^{2+} would be displaced as an effect of pH.

5.4.4 Immobilized metal chelates as IMAC ligands

5.4.4.1 Net charge of ligands

Based on Pearson's classification of hard acids and bases[5], coordination bonds between hard acids and bases have a more ionic characteristics. The net charge of the different immobilized metal chelates under various conditions can therefore be considered an important property in a Ca^{2+} based IMAC system.

Burai *et al*.[189] found that DO2P only forms the negatively charged complexes $CaDO2P^{2-}$ and $CaDO2PH^-$. As previously discussed, the Ca^{2+}-DO3P chelates formed[184] are also all negatively charged. Thus Ca^{2+}-DO3P-Sepharose 6 F F could potentially attract positively charged groups.

Alendronate has a net charge of -2 at physiological pH[186]. Ca^{2+}-alendroante can therefore be assumed to be neutral under these conditions. As mentioned above, the complexes[191] CaDO3AH and $CaDO3A^-$ have been shown to form. It is not known which of these complexes correspond to the conditional stability constant $\log \beta'_{Ca(II):DO3A} = 5.71$ at pH 7.4 calculated by Kumar *et al*.[185]. Thus Ca^{2+}-DO3A-Sepharose 6 F F and Ca^{2+}-alendronate 6 F F could potentially form neutral complexes at pH 7.5. The coordination bounds formed with groups could then potentially be less charge dependent.

5.4.4.2 Ca^{2+} displacement

As discussed by Chaga *et al*. MIT can occur in Ca^{2+} based IMAC systems[90]. The authors argue[90] that displacement of bound Ca^{2+} from immobilized Ca^{2+} chelates can occur in IMAC experiments when samples contain Ca^{2+} binding proteins since $\log \beta_{Ca:Protein} > \log \beta_{Ca:L}$. The role of buffer components has however not been systematically investigated. From the results obtained in this Chapter it is clear that buffer components such as salts, and pH contribute to the displacement of bound Ca^{2+} from immobilized Ca^{2+} chelates. This is true even for Ca^{2+}-DO3P and Ca^{2+}-DO3A with $\log \beta_{Ca:L} > 10$ when immobilized to Sepharose 6 F F. Proteins and components in *E. coli* cell lysates did not displace Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F and neither did the putative Ca^{2+} binding tag HIT2. Hence all MIT observed here can be attributed to buffer components.

Salts such as NaCl are added to eluents to quench secondary ionic interactions in IMAC experiment[6]. The results obtained here indicate that the buffer compositions used with Ca^{2+} based IMAC system can not be chosen based on the same assumptions as with

borderline metal IMAC systems. A salt free or low salt containing buffer could instead serve as the binding condition for all immobilized metal chelates studied. The pH could be between 5.5 and 9.5 for Ca^{2+}-DO3A-Sepharose 6 F F but would need to be kept above 7.5 for Ca^{2+}-DO3P-Sepharose 6 F F to avoid displacement of bound Ca^{2+}. Elution of the bound protein(s) could be obtained by increasing the salt concentration or adding a chelator such as EDTA. The absence of salts in the load and wash steps of a chromatography experiment could potentially pose a problem if secondary ionic interactions would lead to non specific binding of proteins and / or other molecules.

5.5 Conclusions

In this Chapter the binding of Ca^{2+} to a range of different immobilized ligands and its displacement from such ligands were studied systematically under chromatographic conditions. MIT was observed for all ligands. The displacement of bound Ca^{2+} was however not caused by the sample as previously suggested[90] but rather by buffer components such as salts and by pH. Thus all immobilized ligands behaved as cation exchangers to some extent. Based on these results low-salt buffers of pH ≥ 7.5 would be suitable binding conditions for Ca^{2+}-IMAC experiments with all immobilized Ca^{2+} chelates investigated.

Between the different immobilized metal chelates investigated Ca^{2+}-DO3P-Sepharose 6 F F had the highest metal loading of 1 mol Ca^{2+} per mol immobilized ligand. Some bound Ca^{2+} was however displaced by NaCl and a range of other salts and by pH < 7.5. Ca^{2+}-DO3A-Sepharose 6 F F had lower Ca^{2+} loading but the displacement of bound Ca^{2+} was not pH dependent in the range 5.5 < pH < 9.5. These observations could be correlated to the protonation state of the two ligands. The displacement of bound Ca^{2+} from Ca^{2+}-alendronate-Sepharose 6 F F by NaCl was also comparably low at pH 7.5.

Some Ca^{2+} remained bound to Ca^{2+}-DO3P-Sepharose 6 F F, Ca^{2+}-DO3A-Sepharose 6 F F and Ca^{2+}-alendronate-Sepharose 6 F F even after the gels had been washed with 300 mM NaCl, pH 7.5. These immobilized metal chelates therefore appear as interesting candidates in a Ca^{2+} based IMAC system.

Chapter 6. Development of new putative Ca^{2+} binding tags

6.1 Introduction

Histidine (His) has, as previously discussed, a relatively high affinity to borderline metal ions such as Ni^{2+} and Cu^{2+} with[28] log $\beta_{Ni(II):His}$ = 8.66 and log $\beta_{Cu(II):His}$ = 10.16. This can partially be explained by Pearson's classification[5] of hard and soft acids and bases whereby borderline metal ions are known to preferably form coordination bonds with nitrogen[7]. Many affinity tags intended for borderline metal IMAC systems therefore comprise linear amino acid sequences containing His residues in various numbers and arrangements. Examples include the HW-tag[53], the tag[54] $(AHGHRP)_n$ where n = 4 or 8, the His_6 tag[9], the HAT™-tag[56], the $heli_{M14}$-tag[57], the tags VYIHPFHL[59] and HQHHH[60] and the NT tags[67] including NT1A[68], developed at the Monash University Centre for Green Chemistry.

Hard metal ions preferably form coordination bonds with oxygen[7]. A logical design of an affinity tag for a Ca^{2+} based IMAC system would therefore be a sequence rich in amino acids containing oxygen rich donor groups such as aspartic acid. The affinity of Ca^{2+} to free amino acids is however very low. The highest available[28] critically determined stability constant of a Ca^{2+} complex with any free amino acid is that of cysteine (log $\beta_{Ca(II):Cys}$ = 2.5). Despite the low Ca^{2+} affinity of free amino acids, many biological molecules bind Ca^{2+} with high affinity (see Section 6.1.1 below). Hence, it is reasonable to assume that designing a putative Ca^{2+} binding tag by creating an analogue to a His containing tag, but with other amino acid residues such as aspartic acid, would yield a tag with a low Ca^{2+} affinity. An affinity tag with a high Ca^{2+} affinity could however be designed by mimicking natural motifs with high Ca^{2+} affinity.

6.1.1 The Ca^{2+} binding sites in Ca^{2+} binding proteins

Ca^{2+} binds to a wide range of proteins. These proteins can be grouped into different families based on their Ca^{2+} binding motifs. Families of Ca^{2+} binding proteins include annexins, C2 domain proteins, pentraxins, vitamin K-dependent proteins and EF-hand proteins[201].

Annexins have been found in fungi, mammals and plants. The proteins contain a core domain of four repeats (eight for annexin A6) of five α-helices and a variable domain. The Ca^{2+} binding sites are located in the loop regions connecting the α-helices in the core

domain[202]. Annexins bind phospholipids in a Ca^{2+} dependent manner. The affinity towards Ca^{2+} in the absence of negatively charged phospholipids is however low for most annexins[202]. The exception is annexin A6. The K_d values of this protein in the absence of phospholipids were determined to be 0.4 μM[203] and 1.2 μM[204] with one Ca^{2+} bound and 20 μM with four Ca^{2+} bound[205].

The C2 domain proteins contain between one and three C2 domains. This domain is present in a wide range of proteins[206]. Two four-stranded β-sheets of the C2 domain form an eight-stranded anti parallel β-sandwich. Several Ca^{2+} ions are bound by the loops connecting the β-strands. As for annexins, the Ca^{2+} affinity of some C2 domain proteins is increased in the presence of phospholipids[201; 206].

Pentraxins are a family of membrane binding proteins containing five identical subunits, non covalently bound to each other[207]. Ten Ca^{2+} ions can bind to each subunit[201]. As for annexins, the protein forms complexes with Ca^{2+} and phospholipids[201].

The vitamin K-dependent proteins contain an N-terminal Gla domain. The Gla domain contains 9 – 13 gamma-carboxyglutamate (γ-Gla), a much stronger Ca^{2+} chelator compared to glutamate[201]. Nine of the ten γ-Gla residues of the Gla domain of human factor VIIa bind seven Ca^{2+} ions[208]. The blood coagulation factors VII, IX, and X and protein C all contain an N-terminal Gla domain[209] followed by Ca^{2+} binding EGF modules. In the presence of the Gla-domain, the Ca^{2+} affinity of the N-terminal EGF module of factor X is $\log \beta = 4$. The corresponding number in the absence of the Gla-domain is[209] $\log \beta = 2.7$.

Several families of Ca^{2+} binding proteins exist in the lumen of organelles. As the Ca^{2+} concentration in these compartments is in the mM range, these proteins contain domains binding Ca^{2+} with low affinity in the mM range, but with high capacity[201].

The EF-hand is the most common Ca^{2+} binding motif in proteins[210]. The EF-hand family is one of the largest protein families, and the motif has been found in more than 1000 different proteins[211]. Proteins containing EF-hands include calmodulin, parvalbumin, troponin C, sarcoplasmic calcium-binding protein, the light chains of myosin, S100 and VIS[212]. The motifs are comprised of a Ca^{2+} binding loop, flanked by two α-helices. EF-hand motifs are divided into two major groups, the canonical and the pseudo EF-hands[213]. The Ca^{2+} binding loop of canonical EF-hands contain 12 amino acid residues. Ca^{2+} is chelated mainly via carbonyls and carboxylates of the amino acid sidechains in positions 1, 3, 5, and 12 of the loop[213]. The loop of the pseudo EF-hand contains 14 amino acid residues. Ca^{2+} is primarily chelated via backbone carbonyl residues[213] in positions 1, 4, 6 and 9 of the loop. Depending on their biological role, the Ca^{2+}-affinity of different EF-hand proteins can vary by as much[214] as $\log \beta = 6$. Some EF-hands do not bind Ca^{2+} at all[210]. The highest observed Ca^{2+} affinity is that of carp parvalbumin[212] and was determined to be $\log \beta = 9$. EF-hand proteins typically contain two to eight EF-hands in pairs which bind Ca^{2+} in a positive cooperative manner[215].

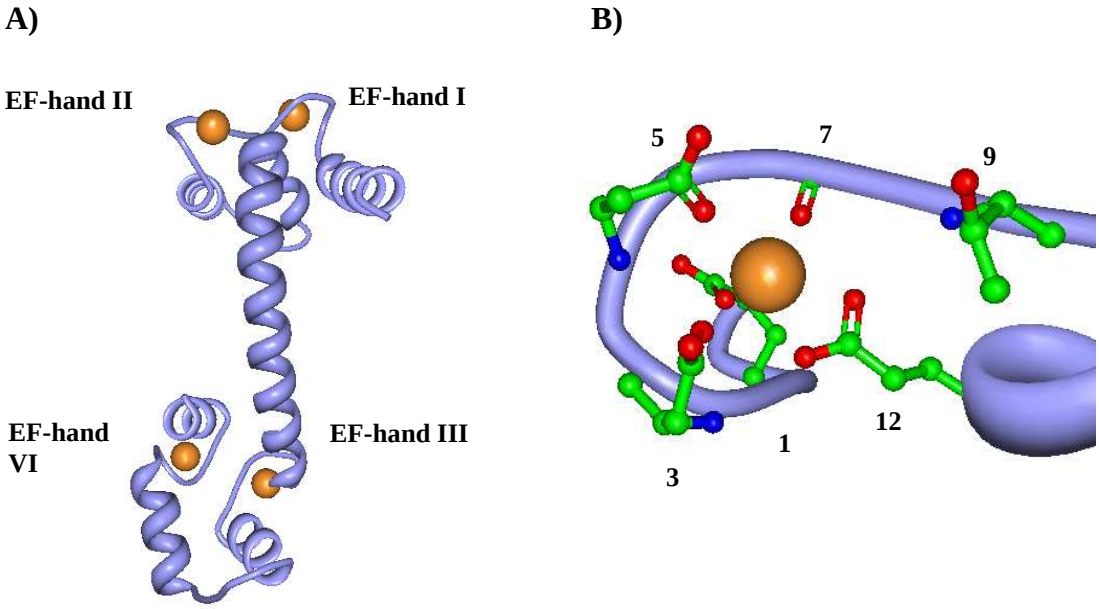


Figure 6.1. Cartoon representation of human calmodulin, pdb code: 1cll[216]. Ca^{2+} ions are coloured orange. **A)** View of the entire protein with the Ca^{2+} ions coordinated in loops I-IV. **B)** Enlargement of EF-hand I and coordinating amino acid residues 1, 3, 5, 7, 9 and 12 displayed as "balls and sticks".

Proteins with five EF-hands have however also been identified. These proteins form homo- or heterodimers through association of the fifth unpaired EF-hand.[201] Domains resembling the EF-hand have also been identified in bacteria. These domains chelate Ca^{2+} in a similar fashion to the EF-hand. Both the domain Excalibur[217] identified in *Bacillus subtilis* and the alginate-binding protein of *Sphingomonas sp.*[218] contain a 10-residue Ca^{2+} binding loop. The galactose-binding protein of *Salmonella typhimurium* contains a nine residue Ca^{2+} binding loop, flanked by a β-turn and a β- strand[219].

Calmodulin is a 148 amino acid residue EF-hand protein present in all eukaryotic cells, where it serves several functions[220; 221]. The amino acid sequence of calmodulin is identical in several organisms such as human[222], rat[223] and mouse[224]. The protein regulates several classes of proteins and enzymes in a Ca^{2+} dependent manner, thereby partiticipating in pathways regulating events such as movement, growth and proliferation[225]. Upon Ca^{2+} binding, the protein changes conformation from a closed to an open conformation. This leads to the exposure of a hydrophobic patch.

Calmodulin contains two pairs of canonical EF-hands connected with an α-helix[226] (Figure 6.1). The EF-hands are labeled I-IV in order from the N to the C-terminus of the protein. The four EF-hands bind Ca^{2+} with different affinities. The two pairs of EF-hands bind Ca^{2+} in a positive cooperative manner. The Ca^{2+} affinities were found to decrease quite significantly with increasing KCl concentration. At low KCl concentrations the Ca^{2+} affinities

of the sites were between[227] log β = 6 - 7 and at 150 mM KCl they decreased to log β = 4.0 - 6.2. As seen in Figure 6.1B the metal ion is coordinated by seven oxygen donor atoms from the 12-residue loop; 1, 3, 5 and 12, the carbonyl backbone oxygen of position 7 and a bridged water molecule (position 9). The Ca^{2+} ion is coordinated by the oxygens in a pentagonal bipyramidal or distorted octahedral configuration[228]. In a Cartesian coordinate system, the Ca^{2+} coordinating amino acids 1, 3, 5, 7, 9 and 12 in the loop have the position X * Y * Z * -Y * -X * * -Z where * denotes intervening amino acid residues[210].

6.1.2 Linear versus cyclic Ca^{2+} binding peptides

Linear peptides based on the EF-hand have been shown to bind Ca^{2+}. The binding is however rather weak, and in some cases it could not be determined at all. The synthetic peptide AcDKDGDGYISAAAQ-NH_2 (where Ac stands for acetyl) analogue to the Ca^{2+}-binding loop of rat testis calmodulin EF-hand III, bound Ca^{2+} with[229] log β = 2. The similar synthetic peptide AcDKDGDGYISAAE-NH_2 was found[230] to bind Ca^{2+} with log β = 2.6. Ca^{2+} was found to bind to a 34 residue peptides corresponding to the helix-loop-helix motif of the EF-hand of rabbit skeletal muscle troponin C. Little or no Ca^{2+}-binding was however evident for the shorter 13-residue synthetic peptide (AcDRDADGYIDAEEL-NH_2) corresponding to only the Ca^{2+}-binding site of the same EF-hand[231].

Higher Ca^{2+} affinities for EF-hand peptides have however been obtained when the secondary structure of the sequence has been constrained so that it resembles the loop shape present in the native EF-hands. Ye *et al*.[232] constrained the 12 amino acid Ca^{2+} binding loop of calmodulin EF-hands by grafting the sequences into the scaffold protein CD-2. EF-hand loop I was shown to have the highest affinity[228] for Ca^{2+} with log β = 4.5 followed by loop III with log β = 3.7, loop II with log β = 3.6 and loop IV with log β = 3.1. Glycine 'spacer regions' were required at either end of the loop in order for the Ca^{2+} binding to occur with high affinity. Furthermore, Ca^{2+} binding did not significantly alter the conformation[233] of CD2.

Research groups have also constrained the EF-hand peptide by introducing disulfide bonds. The aim was to develop lanthanide binding peptides. Le Clainche *et al*. introduced a disulfide bridge in the EF-hand I of *Paramecium tetraurelia* creating a peptide (Cam-M1c) with Ca^{2+} affinity of log β = 4.5. No measurable Ca^{2+} affinity could be determined for corresponding linear peptide sequences[234; 235]. Nitz *et al.*[236] and Franz *et al*.[237] created a series of similar lanthanide binding tags (LBT's). The tags were based on the Ca^{2+} binding EF-hand sequences since lanthanides have similar size and coordination properties to Ca^{2+} and often are employed to study Ca^{2+} binding in biological systems[238]. Nitz *et al.* also obtained higher metal affinities when the peptides were constrained by a disulfide bridge as compared to the linear sequence[236; 237]. The Ca^{2+} affinities of several Ca^{2+} binding proteins and peptides are summarized in Table 6.1.

Table 6.1. Ca^{2+} affinities of Ca^{2+} binding proteins and peptides described in the literature. n.a: not available.

Ca^{2+} binding site	**Secondary structure**	**log $\beta_{Ca(II):tag}$**	**log $\beta_{Tb(III):tag}$**	**Ref.**
Parvalbumin (carp)	Native	9	n.a	[212]
Calmodulin EF-hand I (human)	Native	7	n.a	[227]
Calmodulin EF-hand I grafted into the scaffold protein CD-2	Artificially constrained by scaffold protein	4.5	n.a	[228]
LBTC2	Artificially constrained by disulfide bridge	n.a	6.7	[237]
CaM-M1c		4.5	5.5	[234]
Linear peptide of rat testis	Unconstrained	2.6	n.a	[230]
calmodulin EF-hand III	Unconstrained	2	n.a	[229]
Linear peptide of calcium binding site III of troponin C	Unconstrained	No binding	n.a	[231]

A clear trend can be observed with the Ca^{2+} affinity being highest for the native Ca^{2+} binding proteins, in an intermediate range (log β ~ 5) for artificially constrained loops of the Ca^{2+} binding site and very low for unconstrained, linear peptides. The reason for secondary structure being a requirement for obtaining medium to high Ca^{2+} affinity to a peptide could be the chelating effect. As mentioned in the Introduction (Section 1.3) affinity tags for Ca^{2+} based IMAC have been developed at the Monash University Centre for Green Chemistry[191]. These tags were based on the 12 amino acid Ca^{2+} binding loop of the EF-hand. The tags are referred to as HIT tags and are linear sequences of 6 to 14 amino acids. The amino acid sequences of the HIT2 tag is shown in Table 6.2 together with the 12 amino acid Ca^{2+} binding loops of human calmodulin EF-hands I-IV.

Table 6.2. Amino acid sequences of the HIT2 tag used in this work and the 12-amino acid Ca^{2+}-binding loops of human calmodulin EF-hand I-IV.

Name	**Amino acid sequence**
HIT2	DIDGDGHINYEE
Calmodulin EF-hand I	DKDGDGTITTKE
Calmodulin EF-hand II	DADGNGTIDFPE
Calmodulin EF-hand III	DKDGNGYISAAE
Calmodulin EF-hand IV	DIDGDGQVNYEE

We hypothesized that a tag with a Ca^{2+} affinity in the range of log β = 4 – 5 would be suitable in a Ca^{2+}-based IMAC system. The stability constant of the tag-Ca^{2+} complex would still be at least two orders of magnitude lower than that of the immobilized metal chelates Ca^{2+}-DO3P-Sepharose 6 F F, Ca^{2+}-DO3A-Sepharose 6 F F and Ca^{2+}-alendronate-Sepharose 6 F F (log β = 7 - 10, Chapter 4). Many cyclic peptides have higher Ca^{2+} and Tb^{3+} affinities compared to linear peptides (Table 6.1). Based on the log $\beta_{Ca(II):L}$ values[228; 234] of ~ 4.5 of such cyclic peptides their Ca^{2+} affinity would be in a range suitable for an affinity tag in a Ca^{2+} based IMAC system. Designing cyclic putative affinity tags for Ca^{2+} based IMAC therefore appears as an attractive strategy.

6.2 Aim

The affinity of putative Ca^{2+} binding tags to free Ca^{2+} can serve as an estimate of the affinity between the tag and an immobilized Ca^{2+}-chelate. The aim of this Chapter was therefore to study the Ca^{2+} affinity of a range of putative Ca^{2+} binding tags. Based on the relatively high Ca^{2+} affinities of cyclic peptides in the literature[228; 234], both linear and cyclic putative Ca^{2+} binding tags were investigated.

6.3 Results

In principle, a cyclic tag can be engineered by constraining each side of the peptide sequence by a number of strategies. Amino acid residues with side chains of opposite charges could be introduced at either side of the peptide sequence. These sidechains could then bind to each other through electrostatic interactions. These attractive electrostatic interactions would be dependent on the pH and the ionic strength of the environment. Alternatively hydrophobic residues could be introduced at either side of the peptide and the ends of the peptide could bind through hydrophobic interactions. A risk of introducing hydrophobic residues would be that they might be incorporated in the core of the protein to which the tag is fused thereby causing misfolding of the protein. The tag could also be grafted into a protein sequence as described[228; 232; 233; 239] by Ye *et al*. The drawback with this approach is that the tag can not be removed. Thus the sequence of the target protein is altered, which is undesirable if the protein is intended for e.g. theraupeutic uses and structural determinations.

The cyclic putative Ca^{2+} binding tags studied in this Chapter were designed by introducing Cys residues at either side of various 12 amino acid sequences (Figure 6.2). The rationale was that the Cys residues would form a disulfide bond thus bringing the tag sequence into the shape of a loop. Introducing Cys residues also involves several potential

drawbacks. Unintended intermolecular disulfide bonds can potentially form between Cys residues of the tag and other proteins containing surface exposed Cys residues. Intramolecular disulfide bounds could also form, potentially resulting in misfolding of the protein to which the tag is fused. Moreover disulfide bonds might not form if the environment is reducing. The strategy was nevertheless chosen since it has been described previously by two independent research groups[234; 235; 236; 237]. A His_6 tag was introduced in sequence with the tag to facilitate rapid purification of the constructs by borderline metal IMAC.

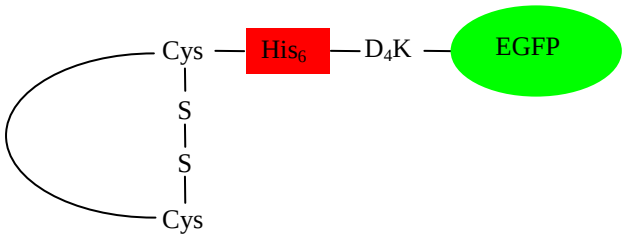


Figure 6.2. Schematic illustration of the design of cyclic, putative Ca^{2+} binding tags. The stretch between the two cystein residues contain the putative Ca^{2+} binding site.

Table 6.3. Amino acid sequences of constrained putative Ca^{2+} binding tags. Underlined sequences represent the 12-amino acid Ca^{2+} binding sequences. Cys residues intended to bond through a disulfide bond are shown in bold.

Code	Amino acid sequence	Description
A	-	Control sequence
B	**C**DKDGDGYITTKELGT**C**	Control sequence Cam-M1c[234]
C	**C**DKDGDGTITTKELGT**C**	Cam-M1c containing Calmodulin EF-I sequence
D	**C**DIDGDGHINYEELGT**C**	Cam-M1c containing HIT2
E	**C**DQDKSGFIEEDELGT**C**	Cam-M1c containing Parvalbumin EF-I sequence
F	**C**ADYNKDGWYEELE**C**	Control sequence[237] LBTC2
G	**C**ADKDGDGTITTKE**C**	LBTC2 containing Calmodulin EF-I sequence
H	**C**ADIDGDGHINYEE**C**	LBTC2 containing HIT2
I	**C**ADQDKSGFIEEDE**C**	LBTC2 containing Parvalbumin EF-I sequence
J	**C**DKDGDGTITTKE**C**	Calmodulin EF-I constrained without extra spacer
K	**C**DQDKSGFIEEDE**C**	Parvalbumin EF-I constrained without extra spacer

A D_4K enterokinase cleavage site was also introduced at the N-terminus of EGFP to facilitate the removal the tags by enterokinase cleavage. A control fusion containing no putative Ca^{2+} binding loop was also designed, thus containing His_6-D_4K-EGFP. The putative Ca^{2+} binding tags were designed by constraining the 12 amino acid Ca^{2+} binding sequences of human calmodulin EF-hand I, carp parvalbumin EF-hand I and the HIT2 sequence by

introducing Cys residues on each side of the sequences. The Cys residues were introduced in three different arrangements in respect to the 12 amino acid sequence. In the first arrangement the Cys residues were located at position -1 and 16 (relative to amino acid residues 1-12 of the Ca^{2+} binding loop) as proposed by Le Clainche *et al*[234]. In the second arrangement the Cys residues were placed at position -2 and 13 as proposed by Nitz *et al*[237].

The third arrangement was at positions -1 and 13, without introducing any spacing amino acids between the Cys residues and the Ca^{2+} binding loop. A total of ten different cyclic tags were thus created. The amino acid sequences of these cyclic peptides are listed in Table 6.3. The cyclic tags are hereafter referred to as B to K as coded in Table 6.3.

6.3.1 Gene cloning of cyclic tagged EGFP

The cloning strategy is schematically described in Figure 6.3. In the first step the gene coding for His_6-D_4K-EGFP was amplified using the polymerase chain reaction (PCR). The gene sequence coding for EGFP in the plasmid pACYC184 (available from previous cloning work in our laboratory) was used as template. The primers employed were A and HindIII-R.

The PCR product was separated on a 1 % agarose gel and purified. The gene was ligated into the plasmid pGEM®-T Easy and transformed into *E. coli* DH5-α. Colonies containing the successful clone were screened by blue/white screening by the following procedure. Transformed bacteria were plated on LB-Ampicillin plates containing 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside (X-Gal). The pGEM®-T Easy plasmid contains the α-peptide coding region of the enzyme β-galactosidase (lacZ). β-galactosidase cleave X-Gal yielding galactose and 5-bromo-4-chloro-3-hydroxyindole, the latter being oxidized to the blue coloured 5,5'-dibromo-4,4'-dichloro-indigo. The lacZ gene is disrupted if a gene is ligated into the plasmid. White colonies are hence potentially successful clones since their lacZ gene has been disrupted.

Plasmid DNA was prepared from potentially successful clones and digested with BspHI and HindIII to confirm the presence of the insert. Plasmid DNA giving the correct digestion pattern was sequenced to further confirm the presence of the insert and to verify the correct sequence.

The genes coding for the ten fusion proteins [B-K]-His_6-D_4K-EGFP were created by PCR using the previously created His_6-D_4K-EGFP sequence as template. The primers B-K were used as forward primers and the primer HindIII-R as the reverse primer. The PCR products were separated on a 1 % agarose gel, purified, ligated into the plasmid pGEM®-T Easy and transformed into *E. coli* DH5-α. Colonies were selected by blue/white screening as described above. Plasmid DNA was prepared from selected white colonies and digested with PciI and HindIII. Plasmids containing the inserts were sequenced.

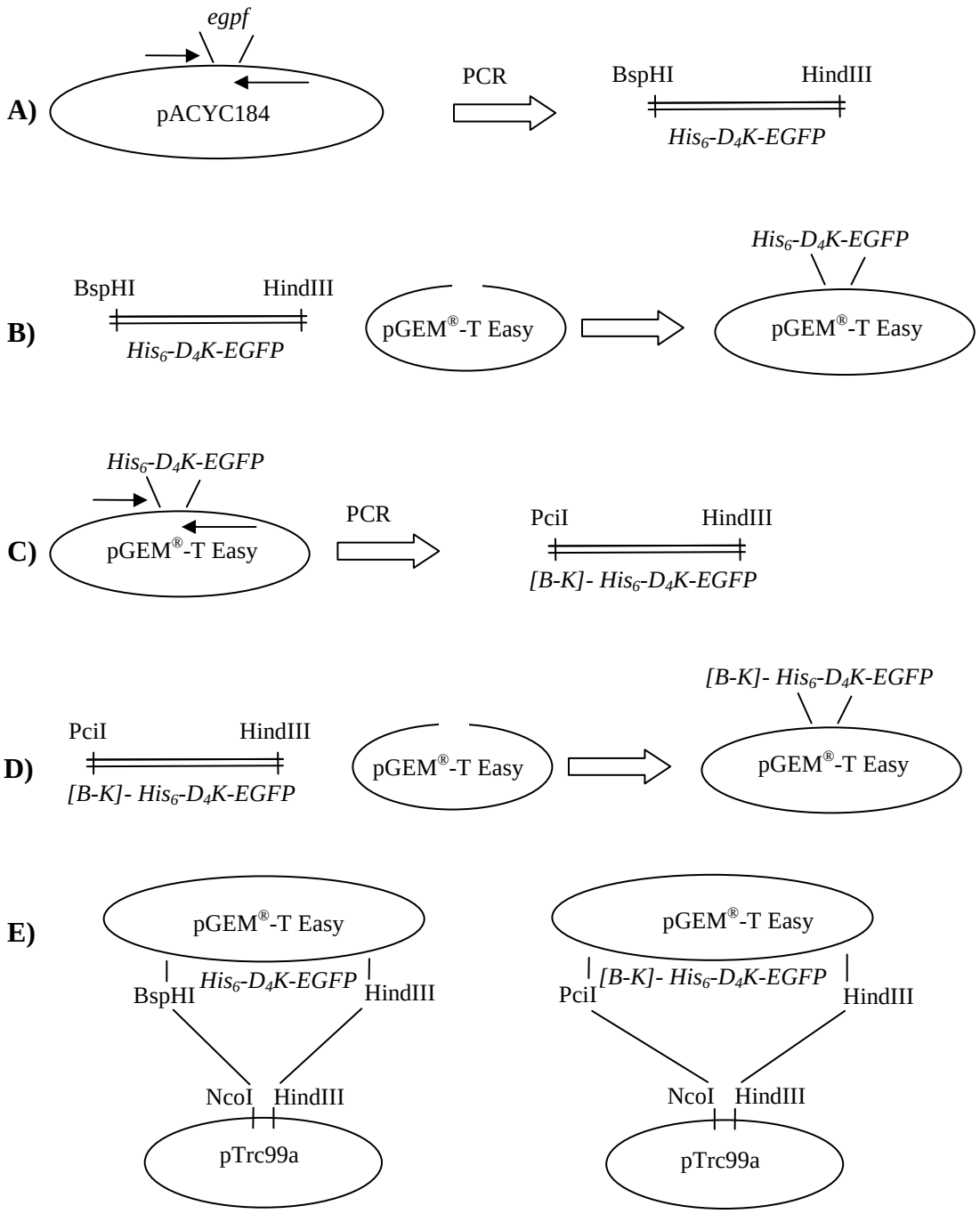


Figure 6.3. Schematic representation of the cloning strategy employed for the creation of *His_6-D_4K-EGFP* and *[B-K]- His_6-D_4K-EGFP*. a) Polymerase Chain Reaction (PCR) amplification of the gene encoding *His_6-D_4K-EGFP*. The *egfp* gene contained in the plasmid pACYC184 was employed as template. The primers were primer A and HindIII-R (section 2.1.3), b) ligation of the gene encoding *His_6-D_4K-EGFP* into the pGEM®-T Easy cloning plasmid, c) PCR amplification of the genes encoding *[B-K]-His_6-D_4K-EGFP*. The gene encoding *His_6-D_4K-EGFP* in pGEM®-T Easy was used as template. The primers were primers B-K and HindIII-R (section 2.1.3), d) ligation of the genes encoding *[B-K]-His_6-D_4K-EGFP* into the pGEM®-T Easy cloning plasmid, e) cutting of the genes encoding *His_6-D_4K-EGFP* using BspH1 and HindIII, and *[B-K]-His_6-D_4K-EGFP* using PciI and HindIII. Ligation into the expression plasmid pTrc99a digested with HindIII and NcoI. The cohesive ends generated by BspH1 and the Pci1 are compatible to those generated by NcoI.

The plasmid containing the gene coding for His_6-D_4K-EGFP was digested with BspHI and HindIII. The gene was ligated into the expression plasmid pTrc99a which had been digested with NcoI and HindIII. The genes encoding [B-K]-His_6-D_4K-EGFP were digested with PciI and HindIII and ligated into pTrc99a. The pTrc99a plasmid had previously been digested with NcoI and HindIII.

The ligation mixture was transformed into *E. coli* DH5-α. Colonies were screened by PCR using primers located on the plasmid flanking the gene insert. The NcoI site in pTrc99a is destroyed when ligated with a DNA segment cut with PciI. The egfp gene however contains an NcoI site. Plasmid DNA prepared from positive colonies was therefore digested with NcoI and HindIII to confirm the presence of the egfp gene in the plasmid. The sequence was further confirmed by sequencing. Plasmids containing the genes encoding His_6-D_4K-EGFP and [B-K]-His_6-D_4K-EGFP were transformed into *E. coli* BL-21 for expression.

6.3.2 Expression and initial characterization of [B-K]-His_6-D_4K-EGFP

The 11 fusion proteins His_6-D_4K-EGFP and [B-K]-His_6-D_4K-EGFP were expressed in *E.coli* BL-21 in 500 mL LB medium containing 100 µg mL^{-1} ampicillin at 37 °C, 220 rpm as described in detail in the Materials and Methods Section 2.3. Expression was induced by 1 mM IPTG. Upon visual inspection of the colour of the cell pellets, the level of expression was found to be His_6-D_4K-EGFP >> F- His_6-D_4K-EGFP > H-His_6-D_4K-EGFP ~ I-His_6-D_4K-EGFP > [B, C, D, E, G, J, K]-His_6-D_4K-EGFP. The control His_6-D_4K-EGFP gave brightly green cell pellets. F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP were only vaguely green/yellowish. Cell pellets containing [B, C, D, E, G, J and K]-His_6-D_4K-EGFP were not green at all.

Cell lysates of [B, C, D, E, G, J and K]-His_6-D_4K-EGFP and the control His_6-D_4K-EGFP were analysed on an anti-GFP western blot (Figure 6.4). The pattern of manually observed fluorescence of the cell pellets was in agreement with the signal intensities observed in the western blot (Figure 6.4). An intense signal is observed for His_6-D_4K-EGFP.

The signals for the cyclically tagged F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP are stronger as compared to the rest of the cyclically tagged fusion proteins (Figure 6.4). Some degradation of several of these tagged EGFP proteins is evident from the immunoreactive bands at ~ 20 kDa (Figure 6.4).

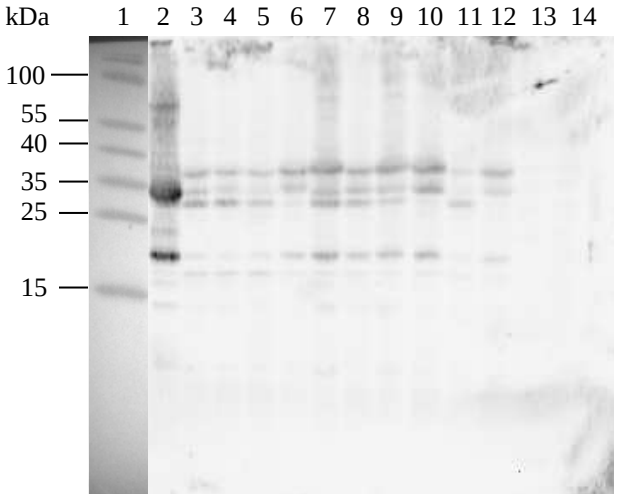


Figure 6.4. Anti-GFP western blot (15% SDS-PAGE) of lysed *E. coli* BL-21 expressing the cyclic tags. Lane 1 represents the size marker. Lane 2 represents His_6-D_4K-EGFP. Lanes 3-12 represents [B-K]-His_6-D_4K-EGFP. Lane 14 represents Xylanase (negative control). 15 represents *E. Coli* BL-21 lacking expression vector (negative control). All samples were diluted 1/10.

In addition dimer formation can be observed for His_6-D_4K-EGFP giving rise to a band at ~ 60 kDa (Figure 6.4). The lack of dimer formation for the cyclic tags most likely reflects the concentration dependence of dimer formation observed for these GFP's[164].

6.3.3 Purification of F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP

Due to the relatively low expression levels, F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP were expressed in batches of 10 L liquid culture. During large scale purification the proteins were found to bind irreversibly to prepacked IMAC resins (HisTrap™) and to Q-Sepharose 6 F F columns. The organic extraction procedure employed by Yakhnin *et al.*[171] to purify EGFP and by Samarkina *et al.*[172] to purify the fluorescent proteins TagCFP, TagGFP, TagYFP, TagRFP, TurboGFP, Dendra2, TurboFP602 and KillerRed, was therefore employed to purify F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP. The method was modified by introducing a SEC step to increase the final purity of the fusion proteins.

The experimental procedures are described in detail in the Materials and Methods Sections 2.3.2 and 2.4.12. Cells were lysed in 20 mM Tris-HCl, 150 mM NaCl, 5 mM EDTA pH 8.0. The lysates was cleared by centrifugation. To 10 mL supernatant, 3 mL 5 M NaCl, 23.3 mL 4 M $(NH_4)_2SO_4$ and 12 mL 96 % ethanol were added. The mixture was shaken vigorously for one minute and centrifuged for 7 minutes at 3000 x g. Two distinct phases could be observed, separated by a layer of cell debris - a top ethanolic layer containing the EGFP as judged by the green colour, and a bottom colourless aqueous layer. The EGFP containing ethanolic phase was transferred to a new tube and 3 mL n-butanol was added.

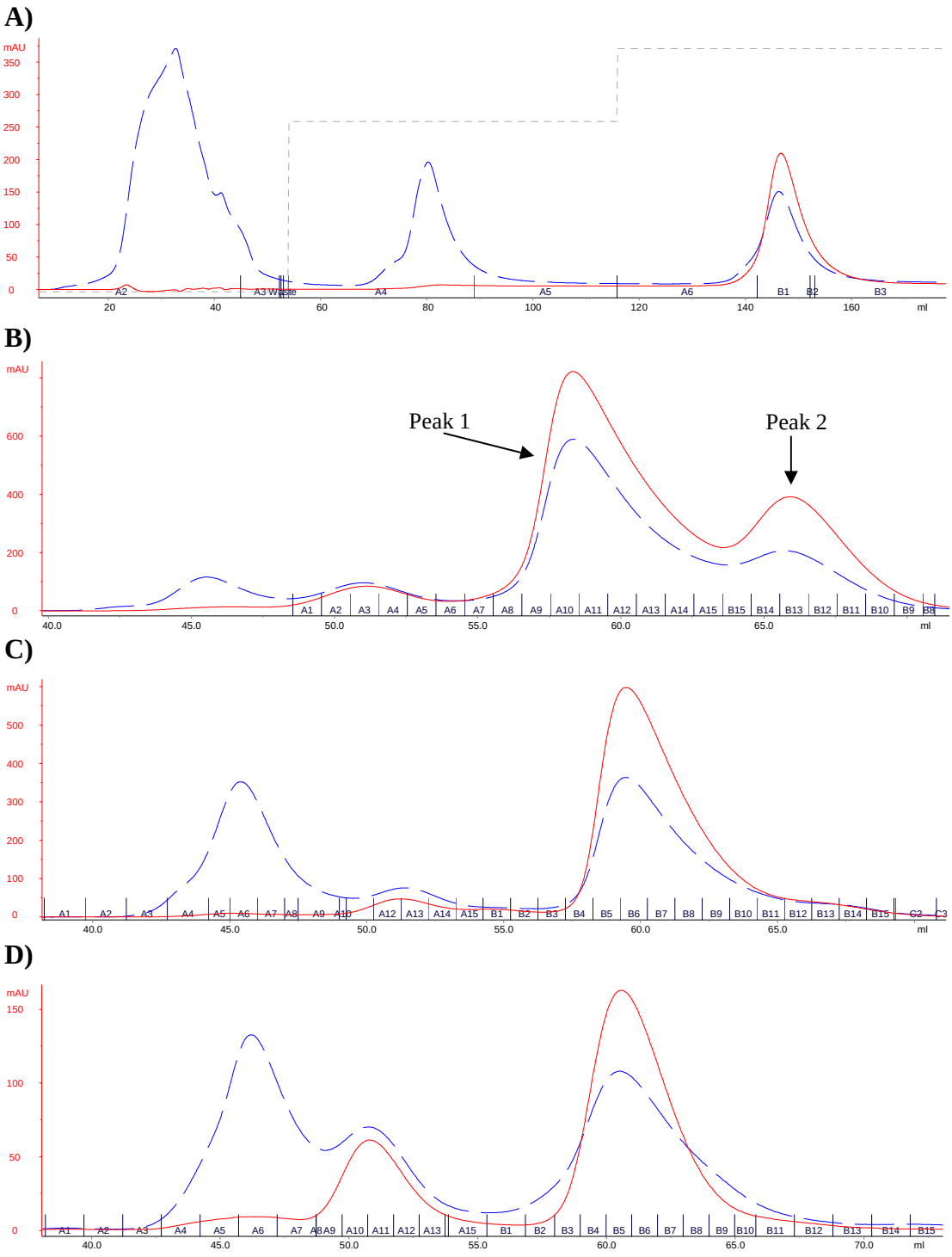
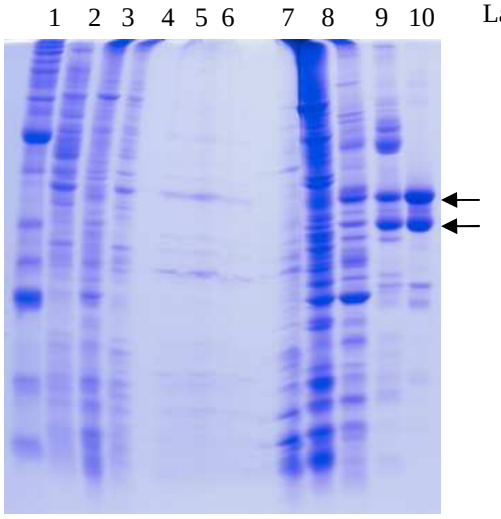


Figure 6.5. A) Hydrophobic interaction chromatography purification of F-his$_6$-D$_4$K-EGFP on an XK 45 glass column packed with 20 mL Phenyl Sepharose Low Sub. Mobile phase A: 1.2 M Ammonium Sulphate, 20 mM Tris-HCl pH 8.0. Mobile phase B: 20 mM Tris-HCl pH 8.0. **B)- D)** Size exclusion chromatography purifications of F-his$_6$-D$_4$K-EGFP, H-his$_6$-D$_4$K-EGFP and I-his$_6$-D$_4$K-EGFP respectively on a Superdex^{TM} 75 HiLoad^{TM} 16/60 column. Mobile phase: 25 mM Tris-HCl, 150 mM NaCl, pH 7.5. Solid trace: absorbance at 485 nm. Dashed trace: absorbance at 280 nm. Dotted trace: concentration of eluent B in eluent A.

After repeating the shaking and centrifugation steps two new phases could be observed, a bright green bottom aqueous layer containing EGFP and a top layer containing butanol/ethanol. The level of purity obtained in the different steps of the purification of H-His_6-D_4K-EGFP is shown in Figure 6.6. The low expression level of the target protein is evident by the lack of an intense band at ~ 35 kDa in the lane representing the cell lysate. No significant enrichment was observed until the second step of the extraction where two bands can be identified at ~ 30 and ~ 35 kDa respectively (Figure 6.6, lane 8).

The organic extraction method was successfully employed for F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP. The bottom layer was further purified by HIC on a 20 mL Phenyl Sepharose 6 F F (Low Sub) column employing a step gradient.



Lane	Sample
1	Standard, 5 µL
2	Cleared cell lysate, 40 µg
3	Ethanolic top layer, extraction 1, 40 µg
4	Precipitated protein, extraction 1, 40 µg
5	Aqueous layer, extraction 1, 12 µL
6	Butanolic layer, extraction 2, 12 µL
7	Precipitated protein, extraction 2, 40 µg
8	Aqueous layer, extraction 2, 40
9	Hydrophobic interaction chromatography elution fractions, 4 µg
10	Size Exclusion Elution Fractions, 4 µg

Figure 6.6. 15% SDS-PAGE gel of fractions taken at different stages of purification of H-His_6-D_4K-EGFP. The band corresponding to the reduced and oxidized form of the construct is shown with arrows.

The HIC purification gave identical elution profiles for all proteins. An example is shown in Figure 6.5A. The protein was loaded and washed in 1.2 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0. A second wash step was employed using 0.36 M $(NH_4)_2SO_4$, 20 mM Tris-HCl, pH 8.0. Both wash steps removed unbound protein as can be seen by the peaks absorbing at 280 nm. The protein was eluted in 20 mM Tris-HCl pH 8.0. The elution peak had higher absorbance at 485 nm indicating relatively high protein purity. When analysed on SDS-PAGE the elution fraction from the HIC step can be considered as semi-pure with the majority of contaminating proteins being removed (Figure 6.6, lane 9). The SEC steps gave different elution profiles for the three cyclic tagged proteins. F-His_6-D_4K-EGFP gave two

unresolved elution peaks (Figure 6.5B). Both peaks absorbed at 485 nm (Figure 6.5B) and it can thus be concluded that they contained EGFP.

Anti-GFP western blot (Figure 6.7C and D, lanes 4 and 6) of fractions taken from the first and second peak (Figure 6.5B) revealed that the second peak contained a protein with a size corresponding to untagged EGFP. The genes encoding His_6-D_4K-EGFP and [B-K]-His_6-D_4K-EGFP contain the start codon *atg* at the beginning of the *egfp* gene. It is therefore possible that untagged EGFP is expressed in parallel to F-His_6-D_4K-EGFP.

6.3.4 Investigation of disulfide bond formation in the cyclic tags

To confirm the correct formation of a disulfide bond in the cyclic tags, the fusion proteins were cut with enterokinase. The experimental conditions are described in the Materials and Methods Section 2.7.

The cut proteins were analysed on SDS-PAGE under non-reducing and reducing conditions and anti-GFP western blot using untagged EGFP as a control. Enterokinase cuts the D_4K recognition site between the tag and the protein (Figure 6.2). Enterokinase treatment of fusion protein containing correctly formed tags will therefore release untagged EGFP. If an intramolecular disulfide bond was formed between Cys 1 or Cys 15 in the tag and the surface accessible Cys 48 of EGFP, digestion with enterokinase would not reduce the size of the protein in a non-reducing SDS-PAGE. As can be seen in Figure 6.7A, the enterokinase treated fractions of F-His_6-D_4K-EGFP (lane 5), H-His_6-D_4K-EGFP (lane 9) and I-His_6-D_4K-EGFP (lane 11) migrated the same distance as untagged EGFP (lane 1).

It can therefore be concluded that no disulfide bond was formed between Cys 1 or Cys 15 in the tags and Cys 48 in EGFP in any of the fusion proteins F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP. Enterokinase treated untagged EGFP gave the same pattern as untreated untagged EGFP proving that the enzyme did not cut unspecifically in the EGFP sequence. (Figure 6.7A, lanes 1 and 2).

Figure 6.7A show non-reduced samples of F-His_6-D_4K-EGFP (lane 4), H-His_6-D_4K-EGFP (lane 8) and I-His_6-D_4K-EGFP (lane 10). Each of these lanes contains two bands, with an estimated molecular weight of ~ 32 and ~ 38 kDa. The bands contain EGFP since they all give a signal in the corresponding western blot (Figure 6.7C). The band at ~ 35 kDa in Figure 6.7A, lane 10, was a contaminant as it did not give a signal in the western blot (Figure 6.7C, lane 10). Figure 6.7B show reduced samples of F-His_6-D_4K-EGFP (lane 4), H-His_6-D_4K-EGFP (lane 8) and I-His_6-D_4K-EGFP (lane 10). In these lanes only one band is observed at ~ 38 kDa. Since Cys residues will be reduced by β-mercaptoethanol, it can be concluded that the band at ~ 38 kDa correspond to the fusion proteins with a reduced, linear tag. The bands at ~ 32 kDa correspond to fusion proteins with an oxidized, cyclic tag. The intensities of the

bands corresponding to the cyclic and linear form of the tag give an estimate of the relative abundance of the two forms.

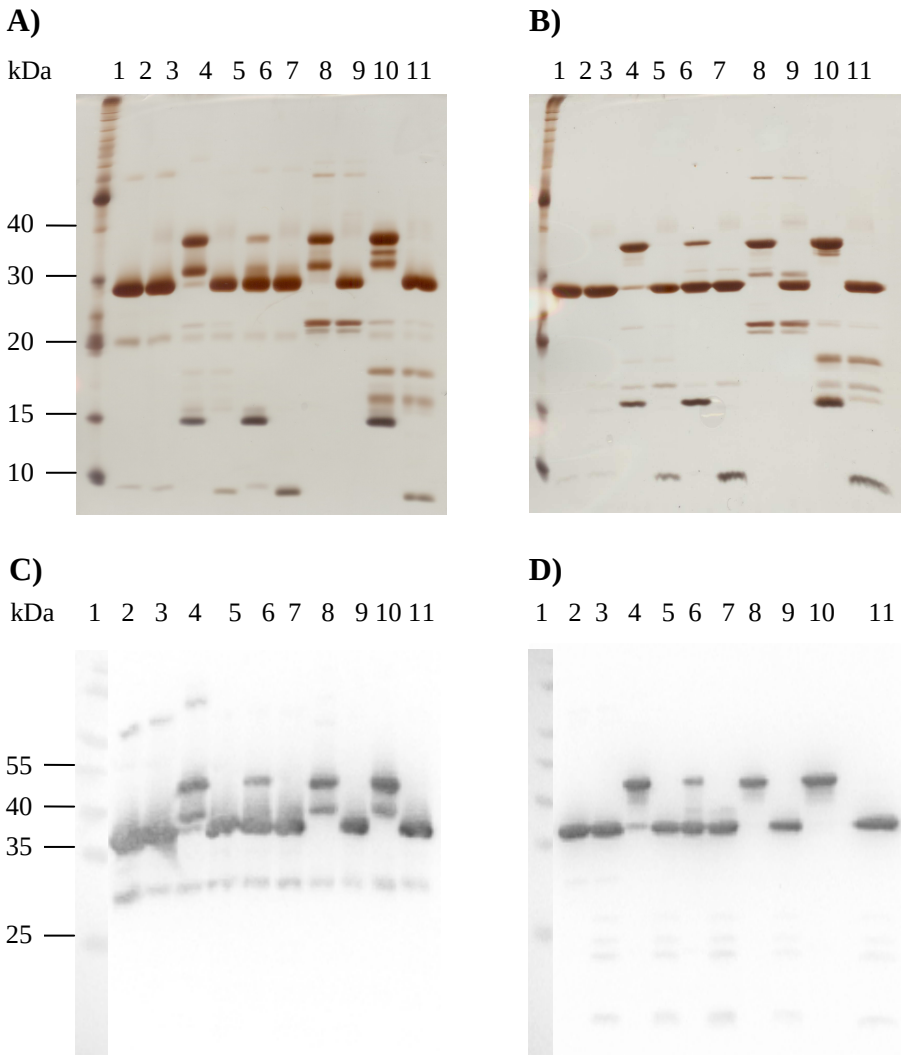


Figure 6.7A) Non reducing 15 % SDS-PAGE. **B)** 15 % SDS-PAGE gel of samples reduced with β-mercaptoethanol. **C)** Non reducing anti-GFP western blot. **D)** Anti-GFP western blot of samples reduced with β-mercaptoethanol. Lane 1: size marker, lane 2: EGFP, lane 3: EGFP cut with enterokinase, lane 4: F-His_6-D_4K-EGFP peak 1, lane 5: F-His_6-D_4K-EGFP peak 1 cut with enterokinase, lane 6: F-His_6-D_4K-EGFP peak 2, lane 7: F-His_6-D_4K-EGFP peak 2 cut with enterokinase, lane 8: H-His_6-D_4K-EGFP, lane 9: H-His_6-D_4K-EGFP cut with enterokinase, lane 10: I-His_6-D_4K-EGFP, lane 11: I-His_6-D_4K-EGFP cut with enterokinase. Peak 1 and 2 refer to the marked peaks for F-His_6-D_4K-EGFP in Figure 6.5.

Hence it can be concluded that a disulfide bond in fact is formed between Cys 1 and Cys 15 in the cyclic tags F, H and I. The fusion proteins containing the oxidized cyclic form of the tag appear to be slightly less abundant as compared to the linear form for F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP.

6.3.5 Purification of Calmodulin

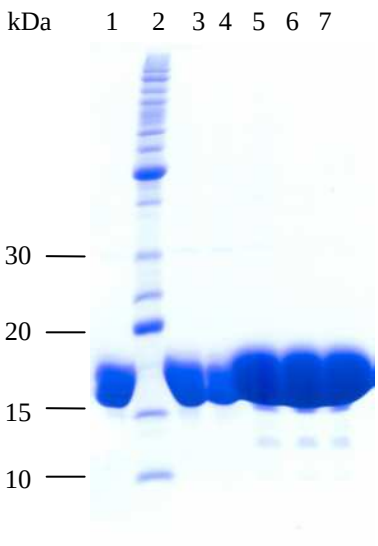


Figure 6.8. 15% SDS-PAGE of Calmodulin purified on Phenyl-Sepharose 6 F F (low sub) on an Econocolumn. Lanes 1 and 3-7 represent calmodulin elution fractions. Lane 2 represents the size marker.

Calmodulin was used as a positive control for all assays of Ca^{2+} binding since the Ca^{2+} association and dissociation of this protein has been well studied. Furthermore the linear tag HIT2 and the cyclic tags B-K were engineered based on the EF-hand sequence of calmodulin (and other Ca^{2+} binding proteins). Another advantage of using calmodulin is that it is easily purified from cell extracts or tissues. Here we employed a protocol for purification of calmodulin from *S. cerevisiae*[102]. Briefly we expressed the gene of human calmodulin cloned into the expression plasmid pTrc99a in *E. coli* BL-21. 50 mL cell culture worth of pellet was lysed as described in the Materials and Methods Section 2.3.2 and purified on Phenyl Sepharose 6 F F (Low Sub) using a bench column as described in detail in the Materials and Methods Section 2.4.3. Figure 6.8 shows a 15 % SDS-PAGE gel of elution fractions from several columns run in parallel. The band corresponding to calmodulin at ~ 17 kDa is intense in all lanes. The purity can be assessed as being > 90 % as minor contaminating bands can be identified at ~ 12 kDa. The protein was dialyzed into 2 mM Tris-HCl, pH 7.5 to remove any residual EGTA from the purification procedure which would otherwise interfere with the downstream Ca^{2+} binding assays.

6.3.6 Studies of Ca^{2+} binding using ESI-MS

Metal binding to proteins can be measured stoichiometrically by ESI-MS by the observed difference in mass of the metal free and metal bound form of the protein. This approach has been successfully employed to study Ca^{2+} binding to several calcium binding proteins. Calbindin D_{28K} was found to bind four Ca^{2+} among its six putative binding sites[240]. Bovine calmodulin, bovine α-lactalbumin and rabbit parvalbumin were found to bind four, one and two Ca^{2+} respectively[241]. For calmodulin the relative Ca^{2+} affinity of the different EF-hands could also be distinguished[242]. Ca^{2+} binding has also been demonstrated in artificial

Ca^{2+} binding sites such as the EF-hands grafted into the scaffold protein CD-2 by Ye *et al.*[228; 232; 233; 239]. Binding of Ca^{2+} to Ca^{2+} binding proteins has been detected both in the positive[228; 242]and negative[240; 241] ion mode. The experiments described herein were performed in the positive ion mode based on instrumental considerations. Ca^{2+}-free protein samples and samples containing Ca^{2+} in molar ratios protein : Ca^{2+} 1 : 2, 1 : 10, 1 : 100 and 1 : 500 were studied. Calmodulin was used as a positive control and untagged EGFP as a negative control. EGFP-HIT2 and untagged EGFP were purified as described in Chapter 4.

As metal ions and buffer salts suppress the ionization in ESI-MS all samples were in 2 mM Tris-HCl, pH 7.5 to maintain buffer capacity while operating at a low salt concentration. When Ca^{2+} binds to Calmodulin two protons are lost to maintain the charge balance and the observed mass shift is therefore + 40 – 2 = + 38 Da[242]. This mass difference upon Ca^{2+} binding has also been observed for the grafted[228] EF-hand in CD-2. Considering that the precision of the deconvoluted masses have a potential error of ± 1 Da and the possibility of Ca^{2+} binding giving rise to unexpected protonation or deprotonation events, mass differences of 38 ± 2 Da were considered to be an indication of Ca^{2+} binding herein. The atomic weight of K is 39 Da. Hence binding of K^+ and loss of one proton could potentially also result in a mass increase of 38 Da.

As all samples were free of K^+, mass increases of 38 ± 2 Da were attributed to the binding of Ca^{2+}. Examples of two deconvoluted mass spectra demonstrating Ca^{2+} binding to calmodulin and to EGFP-HIT2 are shown in Figure 6.9. In both cases both the Ca^{2+}-free and Ca^{2+}-bound forms of the protein were present in the same spectrum. A minimum of three spectra were acquired for each protein at each Ca^{2+} concentration. A summary of the mass shifts observed for all proteins is shown in Table 6.4.

The deconvoluted mass of commercial calmodulin was 16793 Da in the absence of Ca^{2+} (Table 6.4). This mass is in agreement with the manufacturer's instructions (available at http://www.sigmaaldrich.com/catalog/ProductDetail.do?D7=0&N5=Product+No.|BRAND_KEY&N4=P2277|SIGMA&N25=0&QS=ON&F=SPEC on 2010-06-16) stating that the molecular weight of the protein is 16790 Da. Three deconvoluted masses of 16706, 16709 and 16730 Da were obtained for calmodulin expressed in *E. coli* BL-21 (Table 6.4). Using the online tool at http://ca.expasy.org/tools/pi_tool.html the theoretical monoisotopic molecular weight of calmodulin was determined to be 16706 Da. The deconvoluted mass of 16709 Da was present at several experimental conditions (Table 6.4). Hence these masses of 16706 Da and 16709 Da were considered to correspond to calmodulin. The atomic weight of Na is 22.99 g mol^{-1}. The deconvoluted mass of 16730 Da could hence reflect binding of one Na^+ to calmodulin.

Table 6.4. Summary of ESI-MS results of proteins with and without Ca^{2+} in different molar ratios. All samples were in 2 mM Tris-HCl, pH 7.5. Numbers in brackets: theoretical monoisotopic molecular weights of proteins calculated from their amino acid sequence using the online tool http://ca.expasy.org/tools/pi_tool.html. Δ: Difference observed between deconvoluted masses in the same spectrum. n: number of spectra where a mass shift was observed/total number of spectra acquired. Commercial bovine brain calmodulin was from Sigma Aldrich (P2277). Expressed calmodulin was expressed in *E. coli* BL-21. *The different masses were obtained from two different mass spectra of the same sample.

Sample	**No Ca^{2+}**	**EGTA 1:10**	**Ca^{2+} 1:2**	**Ca^{2+} 1:10**	**Ca^{2+} 1:100**	**Ca^{2+} 1:500**
Calmodulin, commercial (16790 Da)	16793	16793	16793	16793 16830 Δ = 37, n = 4/7	16793 16830 16866 16905 Δ = 37, 36, 39, n = 3/9	n.a
Calmodulin, expressed (16706 Da)	16706 16709 16730	n.a	16709 16747 Δ = 38, n = 1/6	16706 16745 Δ = 39, n = 3/9	16706 16745 16785 Δ = 39, 40, n = 5/6	n.a
F-His_6-D_4K-EGFP (30237 Da)	30310 30312 30318	30311	30306 30310 30318	30310 30312 30318 30327	30310 30341	30310 30346 30348 Δ = 36, 38, n = 1/10
H-His_6-D_4K-EGFP (30054 Da)	30164 30169 30172	30165 30190	30169 30171	30164 30168 30170	30166	n.a
I-His_6-D_4K-EGFP (30089 Da)	30161 30167 30187	30160	30162	30161 30200 Δ = 39, n = 1/4*	30161 30163 30312	n.a
EGFP-HIT2 (28382 Da)	28251	28251	28251	28251	28251	28251 28287 Δ = 36, n = 3/4
EGFP (26877 Da)	26746 26843	n.a	26746 26843 26866	26745 26842 26867	26746 26784 Δ = 38, n = 3/6	26747 26784 Δ = 37, n = 1/3

A mass difference of 87 Da can be observed between commercial calmodulin and calmodulin expressed in *E. coli* BL-21 (Table 6.4). This difference can be attributed to the post-translational modifications occurring in the eukaryotic system; a trimethylation of lysine at position 115 and acetylation of the terminal alanine residue according to the manufacturer's instructions. These post translational modifications result in mass increases of + 42 Da and + 43 Da respectively.

For calmodulin expressed in *E. coli* BL-21 two deconvoluted masses of 16709 and 16830 Da were obtained at a protein : Ca^{2+} ratio of 1 : 2 (Table 6.4). The difference between these masses is 38 Da, corresponding to the binding of one Ca^{2+}. At a protein : Ca^{2+} ratio of 1 : 100 three deconvoluted masses of 16706, 16747 and 16785 Da were detected (Table 6.4). The differences between these masses were 39 and 40 Da. Hence the obtained masses corresponded to Ca^{2+}-free calmodulin and calmodulin to which one and two Ca^{2+} were bound.

For commercial calmodulin two deconvoluted masses of 16793 and 16830 Da were detected at a protein : Ca^{2+} ratio of 1 : 10 (Table 6.4). The mass difference was 37 Da, which corresponded to the binding of one Ca^{2+}. At a protein : Ca^{2+} ratio of 1 : 100 the deconvoluted masses 16793, 16830, 16866 and 16905 Da were detected (Table 6.4). This corresponded to Ca^{2+}-free calmodulin and calmodulin with one, two and three Ca^{2+} bound.

For F-His_6-D_4K-EGFP, three deconvoluted masses of 30310, 30312 and 30318 Da were detected in the absence of Ca^{2+} (Table 6.4). The masses of 30310 and 30312 Da could correspond to the reduced and oxidized form of the cyclic tag. Formation of the disulfide bridge would result in the loss of two hydrogens corresponding to a mass difference of 2 Da. The theoretical monoisotopic mass of F-His_6-D_4K-EGFP was however 30237 Da. Hence the deconvoluted masses obtained were 73 - 84 Da heavier than the theoretical monoisotopic mass. One or several post translational modifications can therefore potentially have occurred. Post translational modifications resulting in mass increases in the range 70 - 74 Da include 3,5-dichlorination of Tyr with one Cl^{35} and one Cl^{37} (+ 70 Da), 3,5-dichlorination of Tyr with two Cl^{37} (+ 72 Da), S-(sn-1-glyceryl) on Cys (+ 74 Da) according to the ABRF delta mass database (http://www.abrf.org/index.cfm/dm.home on 2010-06-17). Since no mass increase of ~ 74 Da occurred for proteins not containing a cyclic tag, addition of S-(sn-1-glyceryl) on one of the Cys residues of the cyclic tags appear as the most likely post translational modification. At a protein : Ca^{2+} ratio of 1 : 500 deconvoluted masses of 30310, 30346 and 30348 Da were detected. Hence Ca^{2+} binding might have occurred at this Ca^{2+} concentration.

Three deconvoluted masses of 30164, 30169 and 30172 Da were detected for H-His_6-D_4K-EGFP in the absence of Ca^{2+} (Table 6.4). As the theoretical monoisotopic mass of the recombinant protein was determined to 30054 Da, the obtained masses were 110 - 118 Da heavier than the theoretical monoisotopic mass. No sign of Ca^{2+} binding was observed for H-His_6-D_4K-EGFP at any protein : Ca^{2+} ratio investigated (Table 6.4).

For I-His_6-D_4K-EGFP three deconvoluted masses were detected in the absence of Ca^{2+}; 30161, 30167 and 30187 Da (Table 6.4). Again the theoretical monoisotopic mass of 30089 Da was 72 - 98 Da lower compared to the obtained deconvoluted masses. It is hence reasonable to assume that S-(sn-1-glyceryl) was added on a Cys residue of the cyclic tag I. At a protein : Ca^{2+} ratio of 1 : 10 two deconvoluted masses of 30161 and 30200 Da were detected. This mass difference of 39 Da could correspond to binding of one Ca^{2+}. No sign of Ca^{2+} binding was however observed when the protein : Ca^{2+} ratio was increased to 1 : 100 and 1 : 500 (Table 6.4).

One deconvoluted mass of 28251 was detected for EGFP-HIT2 at Ca^{2+}-free conditions and up to a protein : Ca^{2+} ratio of 1 : 100. At a protein : Ca^{2+} ratio of 1 : 500 two deconvoluted masses of 28251 and 28287 Da were however detected (Table 6.4 and Figure 6.9B). This corresponds to a mass difference of 36 Da, potentially representing the binding of one Ca^{2+} to this fusion protein.

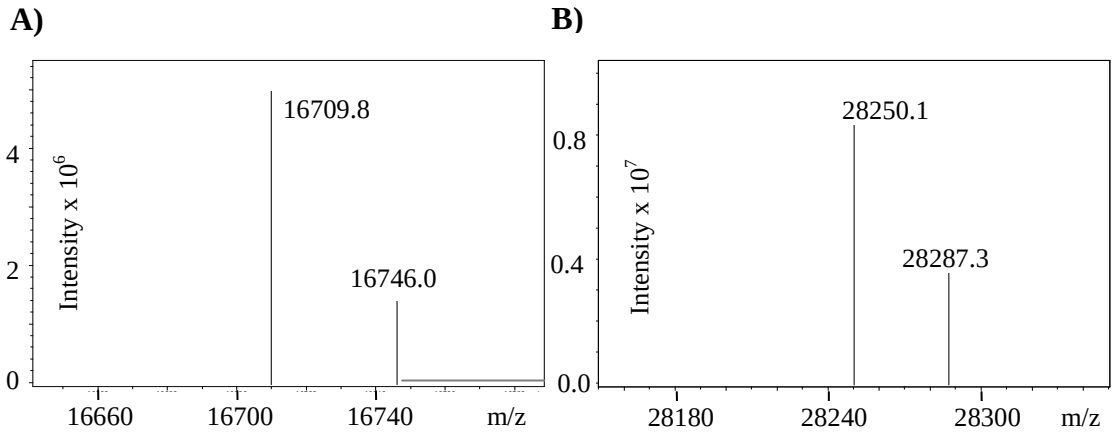


Figure 6.9. Deconvoluted electrospray mass spectra of **A)** 100 µM purified calmodulin supplemented with $CaCl_2$ to a molar ratio Ca^{2+} : protein 2 : 1. **B)** 25 µM EGFP-HIT2 supplemented with $CaCl_2$ to a molar ratio Ca^{2+} : protein 500 : 1. Samples were in 2 mM Tris-HCl, pH 7.5.

For untagged EGFP two deconvoluted masses of 26746 and 26843 were detected under Ca^{2+}-free conditions. At a protein : Ca^{2+} ratio of 1 : 100 EGFP two deconvoluted masses of 26746 and 26784 Da were detected (Table 6.4). This corresponds to a mass difference of 38 Da and hence binding of one Ca^{2+}. The binding of Ca^{2+} to the negative control (untagged EGFP) indicates that unspecific binding of Ca^{2+} can occur at protein : Ca^{2+} ratios of 1 : 100. The observed binding of Ca^{2+} to EGFP-HIT2 might therefore not be related

to the HIT2 tag but could reflect non-specific binding of Ca^{2+} to the EGFP part of the fusion protein.

As previously discussed (Chapter 3) the difference between theoretical monoisotopic masses and the observed deconvoluted masses of EGFP and EGFP-HIT2 (131 Da for both proteins) is likely due to removal of the starting methionines in these two proteins. Based on the results obtained herein it is hard to draw the conclusion that any of the tags bind Ca^{2+} with any measurable affinity.

6.3.7 Determining the Ca^{2+} affinity of putative Ca^{2+} binding tags

6.3.7.1 Methodological considerations

Many different methods for determining stability constants for Ca^{2+} and various ligands such as proteins and peptides have been reported. If the Ca^{2+} bound and Ca^{2+} free state of the protein show different spectroscopic properties such as absorbance or NMR shift, direct spectroscopic measurements are the most straightforward as the concentration of the Ca^{2+}-protein complex [Ca^{2+}_{bound}] is measured directly.

If the Ca^{2+} bound and Ca^{2+} free states of the protein do not show different spectroscopic properties the free Ca^{2+} in solution [Ca^{2+}_{free}] can be measured using a Ca^{2+} selective electrode. [Ca^{2+}_{bound}] can then be calculated by subtracting [Ca^{2+}_{free}] the total Ca^{2+} [Ca^{2+}_{tot}]. Ideally, the protein concentration should be in the same range as the dissociation constant K_d so that bound and free protein is present in reasonable concentrations at several Ca^{2+} concentrations. If K_d is in the range 1 to 100 nM (log β = 7 - 9) the amount of [Ca^{2+}_{free}] will be too low until the point where the protein is saturated. K_d can then be more accurately determined by introducing a competitive chelator with a similar affinity to Ca^{2+} as the protein. The chelator changes its optical properties upon Ca^{2+} binding. The concentration of the chelator is measured either in its free form or its Ca^{2+}-bound form. The concentration of the Ca^{2+}-protein complex [Ca^{2+}_{bound}] can then be calculated by subtracting the Ca^{2+} bound by the chelator from the total Ca^{2+}, [Ca^{2+}_{tot}]. Other techniques available are equilibrium and flow dialysis experiments. These techniques are useful for determining stability constants in the same range as dye competition experiments[243]. Ye *et al*. determined the stability constant of a grafted peptide by dye competition studies using the fluorescent dye Rhodamine-5N[228; 232; 233; 239]. Stability constants for synthetic Ca^{2+} binding peptides have been measured by several techniques such as time-resolved Laser-Induced Fluorescence (TRLIF), CD spectroscopy[234; 235] and NMR[229].

6.3.7.2 Determining stability constants with 5,5'-Br_2-BAPTA

Competitive chelator experiments were carried out to determine log $\beta_{Ca(II):L}$ of EGFP-HIT2. Untagged EGFP was used as a negative control and calmodulin as a positive control.

HIT2-GST and untagged GST (purified as described in Chapter 2, Section 2.4.2) was also tested. The competitive chelator 5,5'-Br_2-BAPTA employed has a log $\beta_{Ca(II):L}$ = 7. Stability constants were determined using the program CaLigator[107] which employs a Levenberg-Marquardt nonlinear fitting routine. The experimental procedures are detailed in the Materials and Methods Section 2.14. The log $\beta_{Ca(II):L}$ of the four Ca^{2+} binding sites in calmodulin were determined to 7.2, 6.8, 6.6 and 6.4 respectively in 2 mM Tris-HCl, 25 mM KCl, pH 7.5. These values are in reasonable agreement with the log β values of 7.0, 5.8, 5.7 and 5.4 determined by Linse *et al.*[227] using 5,5'-Br_2-BAPTA in 2 mM Tris-HCl, 25 mM KCl, pH 7.5.

No reliable data could however be obtained for EGFP-HIT2, untagged EGFP, HIT2-GST and untagged GST (data not shown). This can reflect that the log $\beta_{Ca(II):L}$ value of the chelator is too high compared to the log $\beta_{Ca(II):L}$ value of EGFP-HIT2[107]. Hence the competitive chelator method using 5,5'-Br_2-BAPTA was unsuitable for determining $\beta_{Ca(II):L}$ of EGFP-HIT2.

6.3.7.3 Determining Ca^{2+} affinity with a Ca^{2+} selective electrode

A Ca^{2+}-selective electrode method was employed as it allows to measure lower $\beta_{Ca(II):L}$ values compared to the competitive chelator method. The experimental procedures are described in detain in the Materials and Methods Section 2.15. Five mL 8 µM protein solutions in 50 mM KCl, 2 mM Tris-HCl, pH 7.5 were used for each titration experiment. Standard curves were obtained by titrating a buffer blank (50 mM KCl, 2 mM Tris-HCl, pH 7.5). The buffer was chosen to be as similar to the ESI-MS experiments as possible. Addition of 50 mM KCl was required due to the operational requirements of the Ca^{2+}-selective electrode, which include using KCl as electrolyte in the reference element. Due to sample limitations, only duplicate titrations were performed on F-His_6-D_4K-EGFP and H-His_6-D_4K-EGFP. All other titrations were triplicates. From the raw data of the titrations (Figure 6.10) it is evident that there is a linear relationship between the electrode output in mV and log[Ca^{2+}_{added}] in the blank experiments at Ca^{2+} concentrations above 3 µM (corresponding to log[Ca^{2+}_{added}] = -5.5). For calmodulin a significant deviation from the standard curve can be observed in Figure 6.10 showing binding of Ca^{2+} to the protein. Slight deviations can also be observed for EGFP-HIT2 and H-His_6-D_4K-EGFP.

The Ca^{2+} bound per protein (r) was plotted versus the concentration of free Ca^{2+} as described in the Materials and Methods Section 2.15.4 (Figure 6.11).

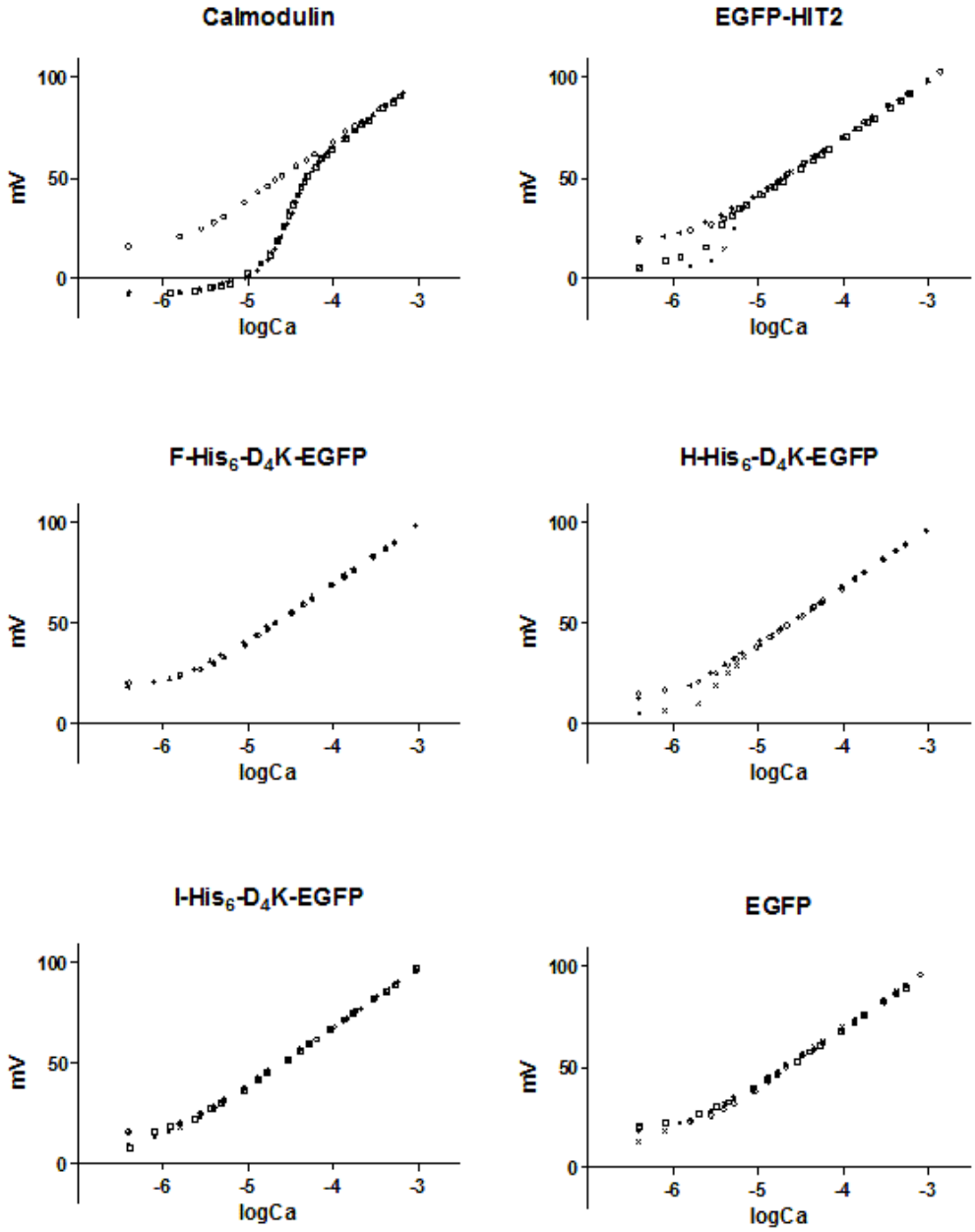


Figure 6.10. Raw data from Ca^{2+}-selective electrode titrations with the electrode output in mV as a function of added $CaCl_2$. All samples were 8 µM protein in 50 mM KCl, 2 mM Tris-HCl, pH 7.5. Protein samples and buffers were decalcified on IDA-Sepharose. Blank experiments were protein-free buffers. Experiment number 1 (x), 2 (+), 3 (□), buffer blanks (○).

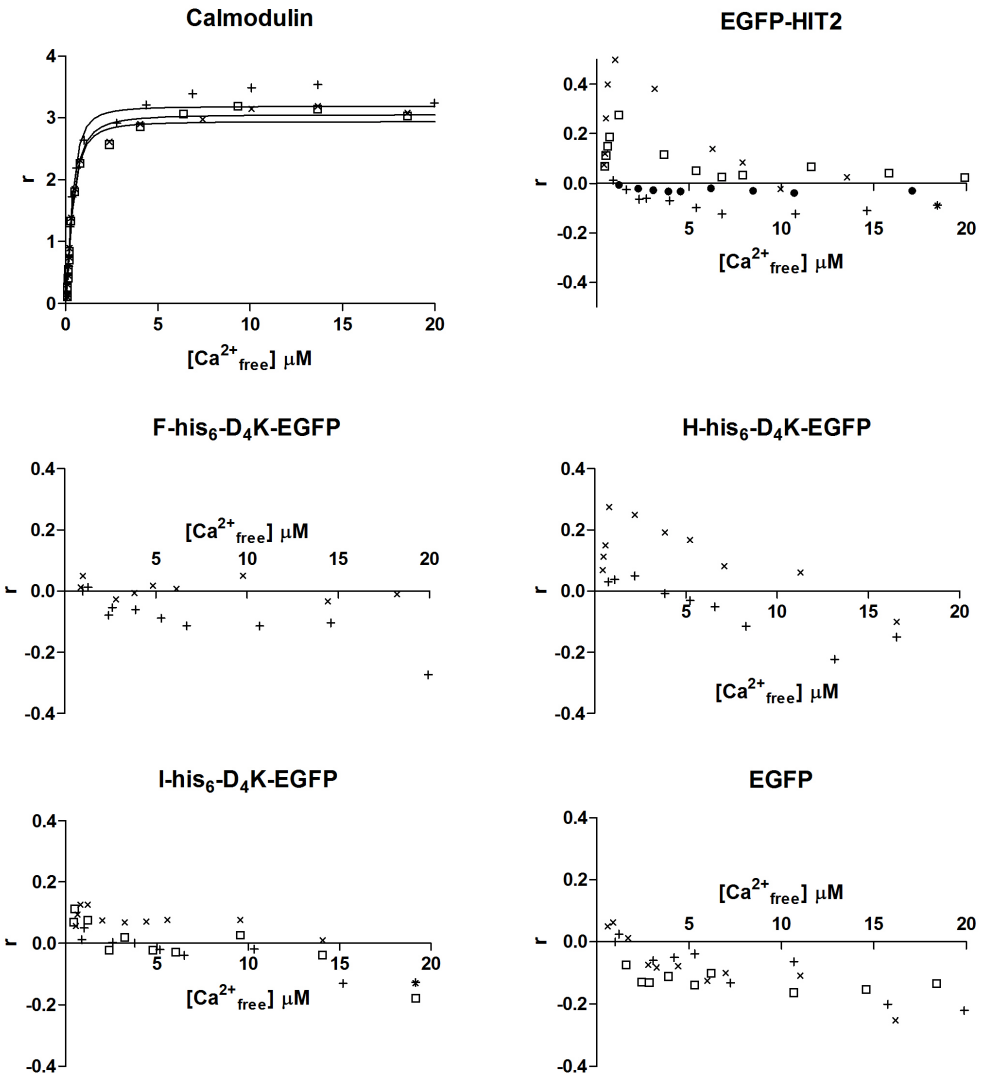


Figure 6.11. Data from Ca^{2+} electrode titrations plotted as bound *versus* free Ca^{2+}. r represents [Ca^{2+} bound]/protein. All samples were 8 µM protein in 50 mM KCl, 2 mM Tris-HCl, pH 7.5. Protein samples and buffers were decalcified using IDA-Sepharose 6 F F. Experiment number 1 (x), 2 (+), 3 (□), experiment with 27 µM EGFP-HIT2 (■). The lines in the calmodulin experiment represent a line of best fit to a binding Equation assuming one binding site with a Hill slope.

For calmodulin the binding of Ca^{2+} can be clearly observed with saturation of the protein at r ~ 3 at a concentration of ~ 10 µM free Ca^{2+} (Figure 6.11). The data were fitted to a binding Equation for one-site binding with a Hill slope (Equation 6.1) where Kd is the dissociation constant and h is the Hill slope.

Equation 6.1
$$r = \frac{r_{\max} \cdot [Ca^{2+}]^h}{K_d{}^h + [Ca^{2+}]^h}$$

A Kd of 0.38 ± 0.1 μM (mean ± sd), corresponding to log β = 6.42 ± 0.02 was obtained for the three measurements. The Hill slope (h) was 1.7 ± 0.1 indicating positive cooperativity between the binding sites in the protein. This is in agreement with the literature as the two EF-hands in Calmodulin are known to bind Ca^{2+} in a positive cooperative manner[227].

The model for proteins containing four Ca^{2+} binding sites with a Ca^{2+} selective electrode in the program CaLigator[107] was also employed, returning binding constants of log β = 5 - 7 for the binding sites (data not shown). The stability constants obtained here are in agreement with literature values of log β = 4.6 – 6.7 for the four binding sites of calmodulin in 50 mM KCl[227]. It can thus be concluded that log β for the binding of Ca^{2+} to calmodulin could be accurately determined with a Ca^{2+} selective electrode.

F-His6-D4K-EGFP, H-His6-D4K-EGFP, I-His6-D4K-EGFP and EGFP-HIT2 contain one Ca^{2+} binding site while calmodulin contain four. An increase of r towards 1 at increasing concentrations of free Ca^{2+} would therefore indicate binding of Ca^{2+} to these fusion proteins. This was however not observed in the plots of r versus free Ca^{2+} for of F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP and EGFP-HIT2 (Figure 6.11). Instead the value of r varied between + 0.3 and – 0.3 (Figure 6.11). From these results, there is therefore no evidence of Ca^{2+}-binding to any of these tagged proteins.

The lack of binding of Ca^{2+} could be due to a too low concentration of fusion proteins containing cyclic Ca^{2+} binding tags in the sample of F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP. Compared to calmodulin which contain four Ca^{2+} binding sites, the concentration of proteins containing one Ca^{2+} binding site would need to be four times higher to achieve the same concentration of Ca^{2+} binding sites. The low expression levels of these fusion proteins (see Section 6.3.2) led to sample limitations. The amount of purified F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP did not permit experiments at higher protein concentrations than 8 μM in 5 mL. For EGFP-HIT2, one experiment was performed using a higher protein concentration of 27 μM. No binding of Ca^{2+} could however be observed for this condition either (Figure 6.11).

6.3.8 Studies of Tb^{3+} binding with LRET

Lanthanides have been extensively used as surrogates to study Ca^{2+} binding. The lanthanides Eu^{3+} and Tb^{3+} are both luminescent in aqueous solution and retain this property when chelated. Furthermore their coordination chemistry[238] closely resembles that of Ca^{2+}. Tb^{3+} mediated luminescence has therefore been employed as a probe for Ca^{2+} binding in a large number of Ca^{2+} binding proteins[244]. The principle behind this technique is that the luminescent amino acids tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe) can

sensitize chelated Tb^{3+}. These amino acids absorb light in the UV range of the spectrum and emit broad-band fluorescence (280-500 nm). Tb^{3+} exhibits several absorption bands in this region (238-385 nm and 476-500 nm). Energy can thus be transferred to chelated Tb^{3+} through non-radiative energy transfer processes.

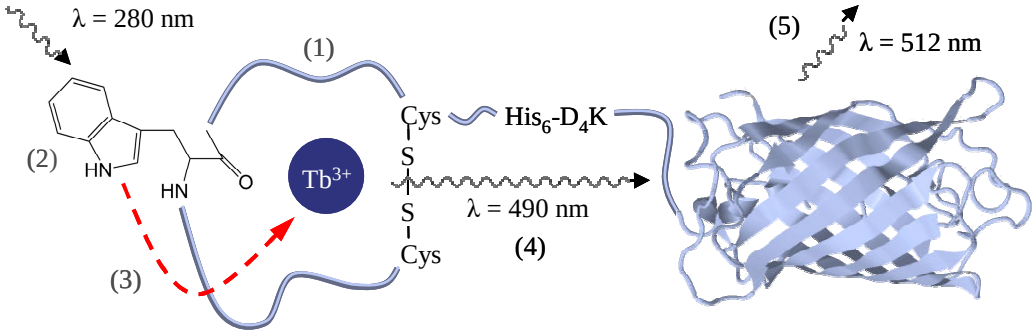


Figure 6.12. Schematic representation of Tb^{3+} mediated LRET in EGFP proteins fused to Tb^{3+} binding tags. **(1)** Disulfide bond constrained Tb^{3+} binding tag. **(2)** Sensitizing Trp residue excited at 280 nm. **(3)** Non radiative energy transfer between Trp and Tb^{3+}. **(4)** Tb^{3+} emission at λ = 490 nm excites EGFP. **(5)** EGFP emission at 512 nm being detected.

This ligand to Tb^{3+} energy transfer nearly always leads to enhanced Tb^{3+} luminescence. The Tb^{3+} emission spectra contain several bands; 485-500 nm (medium-strong), 540-555 nm (strongest), 580-595 nm (medium), 615-625 nm (medium-weak) and 645-655 nm (weak)[238]. Using Tb^{3+} sensitized luminescence methods the order of Ca^{2+} binding to different sites has been determined[245]. In our system, we could employ Tb^{3+} as a probe to investigate Ca^{2+} binding to the putative Ca^{2+} binding tags through luminescence resonance energy transfer (LRET) between Tb^{3+} and EGFP. The principle is schematically described in Figure 6.12. Trp and Tyr residues were excited by irradiating at 280 nm thus potentially exciting chelated Tb^{3+} through nonradiative energy transfer as described above. The Tb^{3+} emission band at 485-500 nm could then excite EGFP with an absorption maxima at ~ 488 nm[154]. The EGFP fluorescence emission could then be detected at 512 nm.

The emission and excitation wavelengths of EGFP do not overlap with the harmonic doubling observed at 560 nm from the excitation of Trp and Tyr at 280 nm (Figure 6.13). The irradiation at 280 nm however lead to emission of EGFP fluorophore fluorescence when a delaytime of 20 μs between excitation and detection was employed (data not shown). This was also observed by Stepanenko *et al.*[166] These authors attributed this observation to possible occurrence of fluorescence resonance energy transfer (FRET) between the Trp residue in EGFP and the EGFP fluorophore[166].

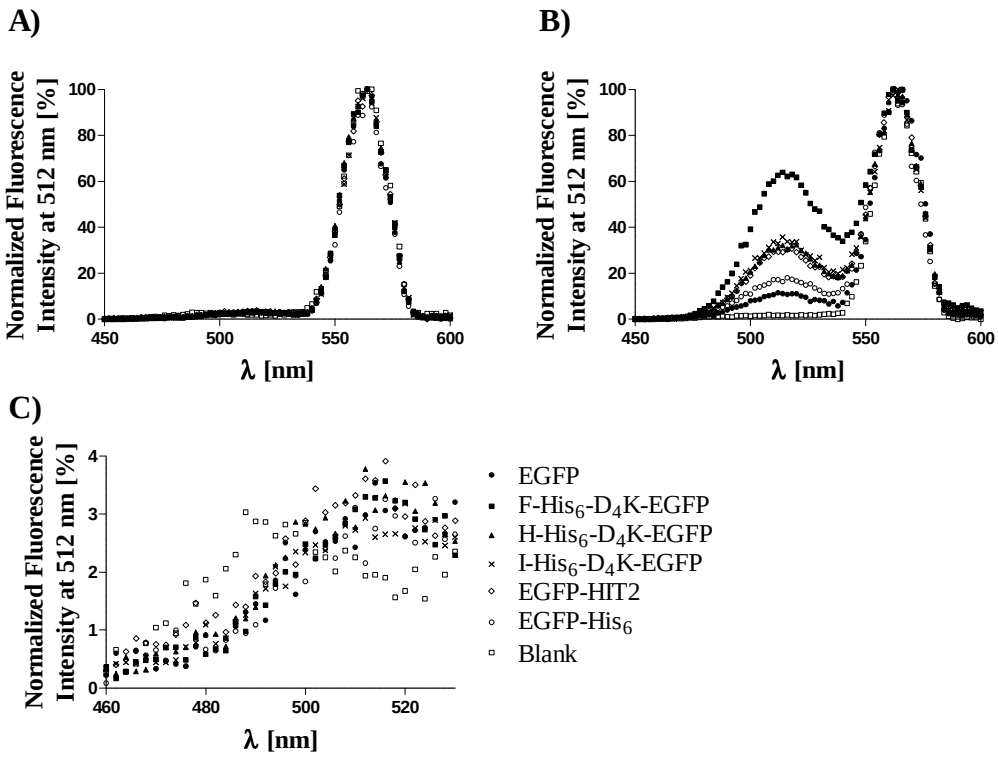


Figure 6.13. Normalized fluorescence emission spectra of **A)** Tb^{3+} free samples, **B)** 40 µM Tb^{3+} and **C)** signal between 0 and 4% of spectrum **A)** between 460 and 530 nm. λ_{exc} = 280 nm, delay time: 50 µs. All proteins were in 2 mM Tris-HCl, pH 7.5. EGFP (●), F-His_6-D_4K-EGFP (■), H-His_6-D_4K-EGFP (▲), I-His_6-D_4K-EGFP (x), EGFP-HIT2 (◊), EGFP-His_6 (○), blank (buffer) (□).

Time gated experiments can however be employed to avoid contribution of non Tb^{3+} related fluorescence, due to the long lifetime of the Tb^{3+} luminescence[246]. A delay time of 50 µs between excitation and measurement was therefore employed in the LRET experiments described in this Chapter. Over this time interval virtually all nonspecific EGFP fluorescence at 512 nm decayed.

This can be seen in Figure 6.13 where the emission spectra of F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP, EGFP-HIT2 and EGFP-His_6 in the absence (Figure 6.13A) and in the presence (Figure 6.13B) of Tb^{3+} are compared. For the Tb^{3+}-free samples the EGFP fluorescence signal at 512 nm is only marginally higher as compared to the buffer blank samples as shown in the enlargement of the spectrum of the Tb^{3+} free samples (Figure 6.13C). Hence the observed fluorescence at 512 nm was due to LRET between chelated Tb^{3+} and EGFP. Solutions of 2 µM of the different fusion proteins in 2 mM Tris-HCl, pH 7.5 were titrated with $TbCl_3$ in the concentration range 0-30 µM as described in detail in the Materials and Methods Section 2.17. The samples were measured in dark 96 well plates in a platereader.

Tb^{3+} binding was measured by exciting at 280 nm. The LRET between chelated Tb^{3+} and EGFP was measured by measuring the EGFP fluorescence intensity at 512 nm after a delay time of 50 μs. The EGFP fluorescence (LRET signal) was plotted as a function of the Tb^{3+} concentration (Figure 6.14A).

Binding was clearly evident for F-His_6-D_4K-EGFP with a steep increase in the LRET signal as a function of the Tb^{3+} concentration reaching a plateau at 20 μM and with half saturation at ~ 10 μM (Figure 6.14A). H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP and EGFP-HIT2 also showed an increase in LRET signal as a function of Tb^{3+} concentration.

This increase was however linear and less pronounced compared to F-His_6-D_4K-EGFP (Figure 6.14A). The increase in LRET signal as a function of Tb^{3+} concentration of EGFP-His_6 and EGFP was low (Figure 6.14A) but still higher than the buffer blank. It can hence be concluded that some non-specific Tb^{3+} binding occurred. The binding data obtained for F-His_6-D_4K-EGFP were fitted to the model "one-site total and nonspecific binding" contained in the software package Graphpad Prism 5.0. The calculation employs Equations 6.2 where $B_{Specific}$ represents the specific binding, B represents the nonspecific binding, K_d is the dissociation constant, BGR represents the background signal, NS is the slope of the nonspecific binding, Y_{max} represents the signal at saturation and x is the concentration of Tb^{3+}. Two curves were then simultaneously fitted to ($B_{specific}$ + B) and to B. The data set obtained for untagged EGFP was employed for representing the nonspecific binding (B). A K_d of 6.9 μM or $\log\beta$ = 5.1 was obtained for F-His_6-D_4K-EGFP.

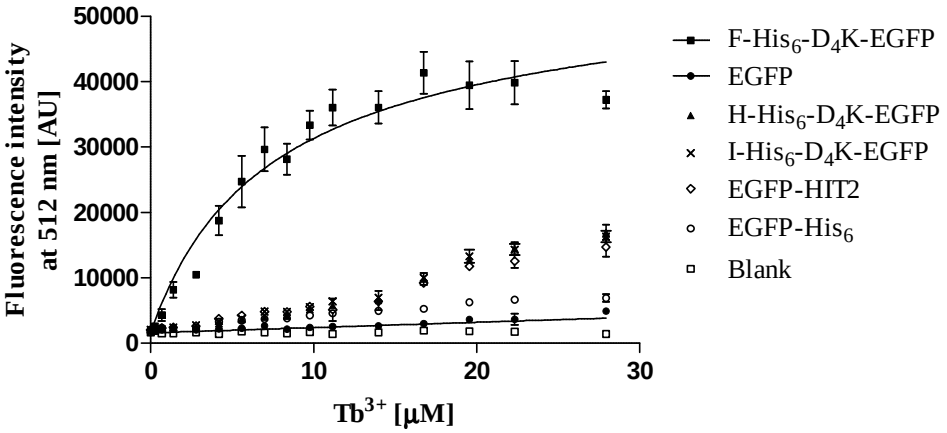
Equations 6.2

$$\begin{cases} B_{Specific} = \dfrac{Y_{max} \cdot x}{x + K_d} \\ B = NS \cdot x + BGR \end{cases}$$

The K_d of 6.9 μM obtained here was considerably higher compared to the K_d of 0.22 μM reported by Franz *et al.*[237] for the cyclic tag F as a peptide, and the slightly lower K_d obtained by the same authors of the fusion protein His_6-F-ubiquitin[237].

To investigate if the high Tb^{3+} affinity of F-His_6-D_4K-EGFP was related to the tag being cyclic, the experiment was repeated under reducing conditions. The increase in LRET signal as a function of Tb^{3+} concentration of non-reduced samples and samples reduced in 100 mM DTT is shown in Figure 6.14B. For the reduced samples the increase in the LRET signal was linear as a function of Tb^{3+} concentration (Figure 6.14B). Hence the observed binding of Tb^{3+} to F-His_6-D_4K-EGFP can be attributed to the tag F being cyclic. The lower LRET signal of the reduced samples could potentially be due to degradation of the EGFP fluorescence by the DTT in the samples.

A)



B)

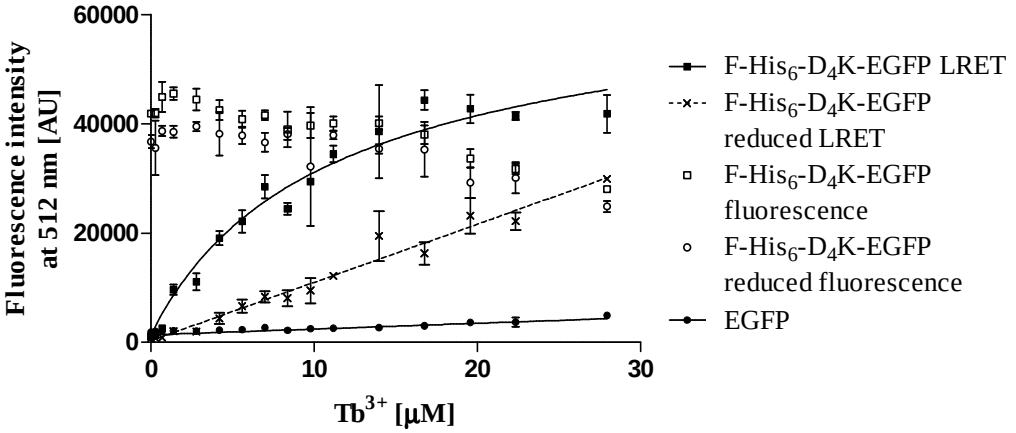


Figure 6.14. Terbium titrations of EGFP constructs. All samples were 2 µM protein solutions in 2 mM Tris-HCl, pH 7.5. LRET experiments were performed with λ_{exc} = 280 nm and a 50 µs delay time. EGFP fluorescence measurements were performed with λ_{exc} = 480 nm and a 20 µs delay time. All points represent the mean of three samples. The error bars represent the standard deviation. **A)** LRET experiments. EGFP (●), F-His_6-D_4K-EGFP (■), H-His_6-D_4K-EGFP (▲), I-His_6-D_4K-EGFP (x), EGFP-HIT2 (◊), EGFP-His_6 (○), blank (buffer) (□). Curves represent lines of best fit to a one-site binding model of total binding accounting for non-specific binding. Binding to untagged EGFP was assumed to represent non-specific binding (lower line). **B)** Samples non-reduced and reduced with 100 mM DTT. LRET of non-reduced F-His_6-D_4K-EGFP (■). LRET of reduced F-His_6-D_4K-EGFP (x). Fluorescence of non-reduced F-His_6-D_4K-EGFP (□). Fluorescence of reduced F-His_6-D_4K-EGFP (○). LRET of EGFP (●), representing non-specific Tb^{3+} binding. Curves represent lines of best fit to a one-site binding model of total binding. Binding to untagged EGFP was assumed to represent non-specific binding (lower line). The dashed line represents a linear regression curve fitted to the dataset obtained with reduced F-His_6-D_4K-EGFP.

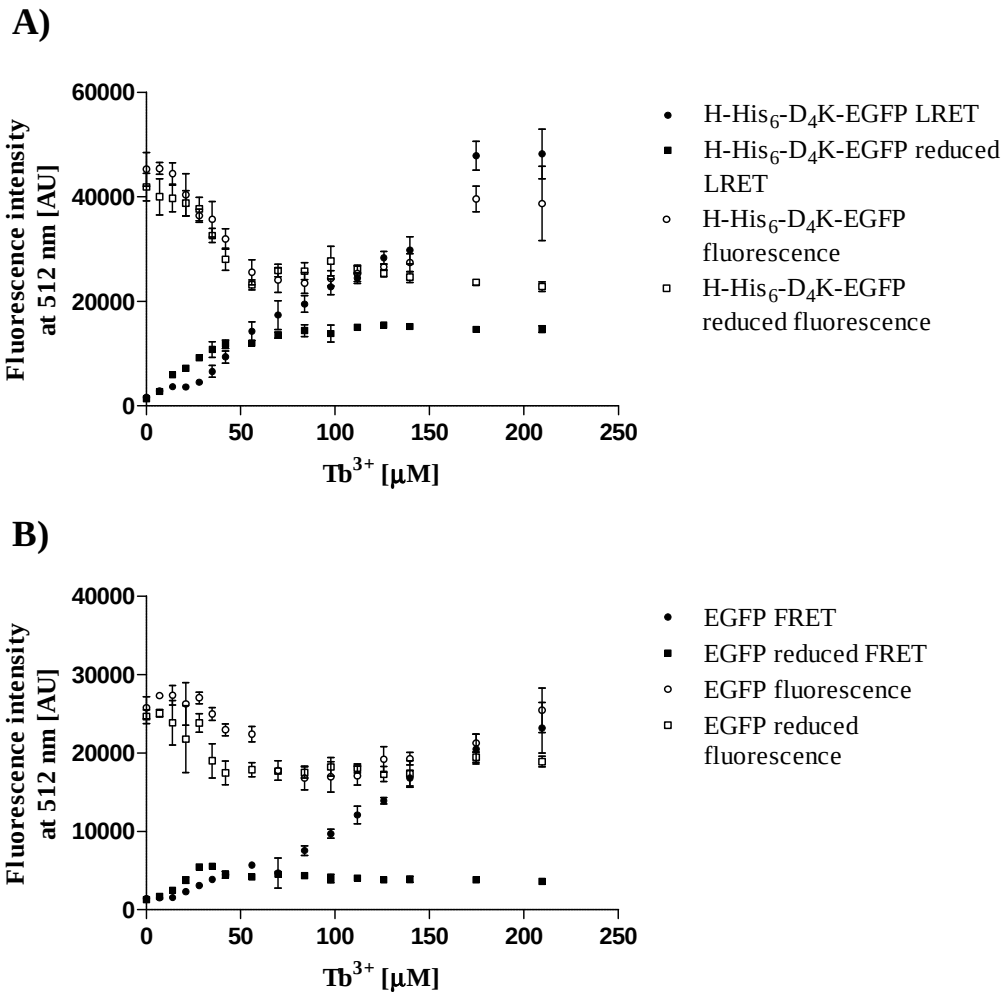


Figure 6.15. Terbium titration of 2 µM protein 2 mM Tris-HCl, pH 7.5. LRET experiments were performed with λ_{exc} = 280 nm and a 50 µs delay time. EGFP fluorescence was performed with λ_{exc} = 480 nm and a 20 µs delay time. All points represent the mean of three samples. The error bars represent the standard deviation. LRET (●). LRET of protein in 100 mM DTT (■). Fluorescence of protein (○). Fluorescence of protein in 100 mM DTT (□). **A)** H-His_6-D_4K-EGFP, **B)** EGFP (untagged).

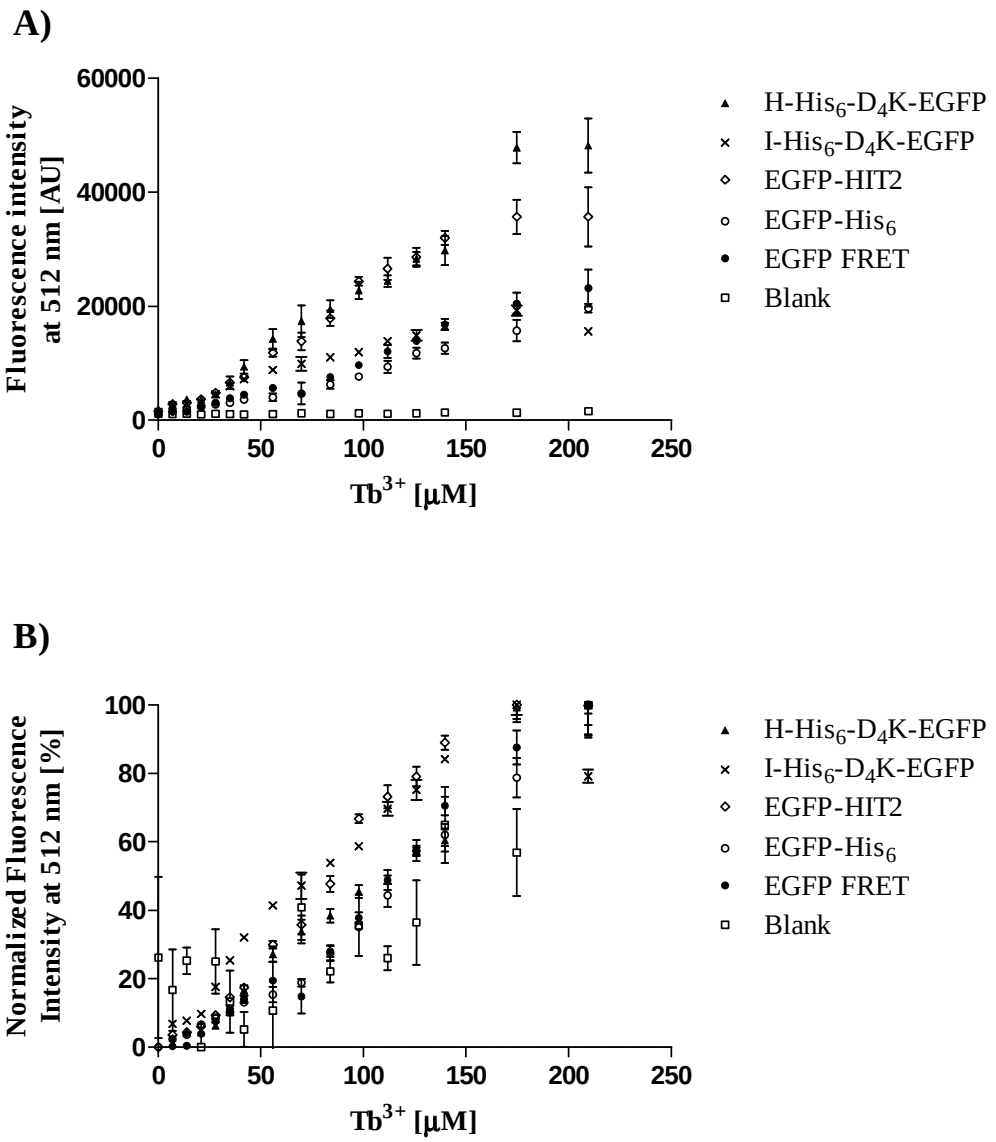


Figure 6.16. Terbium titrations of EGFP fusion proteins. All samples were 2 µM protein in 2 mM Tris-HCl, pH 7.5. LRET experiments were performed with λ_{exc} = 280 nm and a 50 µs delay time. All points represent the mean of three samples. The error bars represent the standard deviation. H-His_6-D_4K-EGFP (▲), I-His_6-D_4K-EGFP (x), EGFP-HIT2 (◊), EGFP-His_6 (○), EGFP (●), buffer blank (□). **A)** Fluorescence intensity. **B)** Normalized fluorescence intensity.

The EGFP fluorescence was therefore measured for all samples by exciting the samples at 480 nm and measuring the EGFP fluorescence at 512 nm after a delay time of 20 µs. The intensities were not significantly different in the presence and absence of DTT (Figure 6.14B). It can hence be concluded that the lower LRET signal in the presence of DTT was not caused by degradation of the EGFP by DTT. The results in Figure 6.14B clearly show that the Tb^{3+}-affinity of the cyclic tag F is higher than the affinity of the same linear peptide when fused to EGFP.

Both in the presence and in the absence of DTT the EGFP fluorescence intensity decreased slightly with increasing concentrations of Tb^{3+} up to a concentration of ~ 50 µM Tb^{3+} (Figure 6.14B and Figure 6.15). This pattern was observed both for proteins tagged with cyclic putative Tb^{3+} binding tags F, H and I and for untagged EGFP. It is therefore reasonable to assume that this observation was related to the nonspecific binding of Tb^{3+} to the proteins.

Higher Tb^{3+} concentrations (0-200 µM) were employed to investigate if binding of Tb^{3+} could be observed for H-His_6-D_4K-EGFP, I -His_6-D_4K-EGFP or EGFP-HIT2. The increase in LRET signal remained linear as a function of Tb^{3+} concentration (Figure 6.16). Hence no evidence of binding could be obtained for any of these fusion proteins.

Cyclic tag F contains a Trp residue at position 7. A Trp residue at this position in a Ca^{2+} binding loop was previously found to enhance Tb^{3+} luminescence[247]. In addition the cyclic tag F contains two Tyr residues at position 2 and 8, further increasing the Tb^{3+} luminescence[236]. Cyclic tag H and the linear HIT2 tag contain one Tyr residue at position 9 and cyclic tag I contain one Phe residue at position 7. The higher abundance of luminescent amino acids in cyclic tag F and the optimized location of the sensitizing Trp residue could explain why binding of Tb^{3+} was observed only for this cyclic tag. The steeper slope observed for reduced F-His_6-D_4K-EGFP as compared with non-reduced H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP could be due to the sensitizing amino acid residues enhancing nonspecific Tb^{3+}-binding. Alternatively the sequence could retain some Tb^{3+} affinity even when linear.

6.3.9 Preliminary results with Ca^{2+}-based IMAC

The properties of putative immobilized Ca^{2+}-chelates and putative Ca^{2+} binding tags fused to EGFP were investigated in Chapter 5 and 6 respectively. Ca^{2+}-alendroante-Sepharose 6 F F, Ca^{2+}-DO3A-sepharose 6 F F and Ca^{2+}-DO3P-Sepharose 6 F F were identified as potentially suitable immobilized metal chelates for use in a Ca^{2+} based IMAC system (Chapter 5). When washed with buffers at pH 7.5, Ca^{2+} was displaced from all the immobilized metal chelates by NaCl (Chapter 5, Sections 5.3.4 and 5.3.5).

The Ca^{2+} affinity to the cyclic tags F, H and I and the linear tag HIT2 fused to EGFP could not be determined (Sections 6.3.6 and 6.3.7). The affinity of Tb^{3+} to the fusion protein F-His_6-D_4K-EGFP was however determined to $\log \beta = 5.1$ in 2 mM Tris-HCl, pH 7.5 (Section

6.3.8). As such, F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP and EGFP-HIT2 were considered as putative Ca^{2+} binding tags. Based on these results, Ca^{2+} IMAC experiments were performed. Several different experimental procedures were employed.

IMAC experiments of EGFP-HIT2 and of [B, C, D, E, F, G, H, I, J and K]-His_6-D_4K-EGFP were performed on Ca^{2+}-DO3P-Sepharose 6 F F. Cell lysates in 10 mM HEPES, pH 7.5 were loaded and eluted using 300 mM NaCl, 10 mM HEPES, pH 7.5. An example of a chromatographic profile obtained for F-His_6-D_4K-EGFP is shown in Figure 6.17A. A small elution peak can be observed early in the gradient, corresponding to fractions 14-17. Similar chromatograms were obtained for [B, C, D, E, G, H, I, J and K]-His_6-D_4K-EGFP and for the negative control His_6-D_4K-EGFP (data not shown). To investigate if the elution peaks contained the tagged EGFP, fractions 13-18 were analysed on SDS-PAGE gels and silver stained. Gels of elution fractions from IMAC experiments of F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP are shown in Figure 6.17B-D. The negative control His_6-D_4K-EGFP was included as a reference. The protein bands in the lanes corresponding to the elution peak have low protein content with only low intensity bands visible on the gels. Bands corresponding to the cyclic tagged proteins can not be clearly distinguished. This could either be due to lack of binding of the fusion proteins to the chromatographic resin, or that the fusion proteins remained bound to the resin during elution.

The resin was not coloured at the end of the chromatography experiment. This observation support that no protein was bound to the column during the experiment. To screen a larger number of conditions more rapidly, gels were packed in 1 mL Tricorn™ columns and used in isocratic chromatography experiments. In these experiments, binding was manually assessed by visually inspecting if the EGFP fusion proteins bound to the column. The mobile phases 10 mM HEPES, pH 7.5 and 10 mM HEPES, 50 mM $CaCl_2$, pH 7.5 were investigated. The immobilized Ca^{2+} chelates Ca^{2+}-alendroante-Sepharose 6 F F, Ca^{2+}-DO3A-sepharose 6 F F, Ca^{2+}-DO3P-Sepharose 6 F F and Ca^{2+}-IDA-Sepharose 6 F F were investigated. *E. coli* BL-21 cell lysates containing the fusion proteins EGFP-HIT2, F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP were tested. In all experiments no green fluorescence was observed in the elution fractions. The flow-through fractions were however green. Based on these observations, it was concluded that no binding could be observed for any of the immobilized metal chelate – affinity tag combinations investigated.

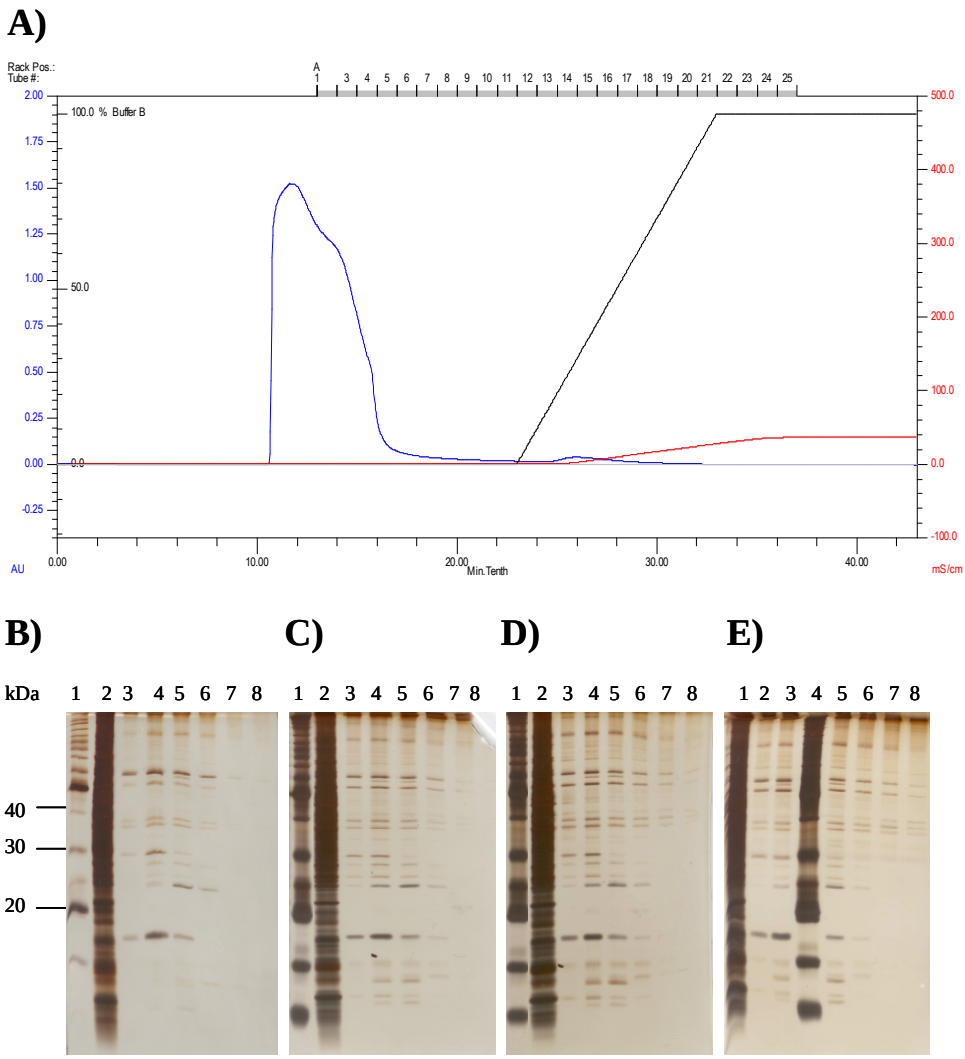


Figure 6.17A) FPLC IMAC experiment of F-His_6-D_4K-EGFP on a 1 mL column packed with Ca^{2+}-DO3P-Sepharose 6 F F. The eluents were A: 10 mM HEPES pH 7.5 B: 300 mM NaCl, 10 mM HEPES, pH 7.5. The blue trace represents the absorbance at 280 nm, the red trace represents the conductivity and the black trace represents per cent eluent B in eluent A. **B-E)** Silver stained 15% SDS-PAGE of elution fractions obtained from FPLC IMAC experiments of cyclic tags on 1 mL columns packed with Ca^{2+}-DO3P-Sepharose 6 F F. 12 µL sample was loaded in each lane. **B)** His_6-D_4K-EGFP (negative control), lane 1: size marker, lane 2: cell lysate diluted 1/100, lanes 3 to 8: elution fractions 13-18. **C)** F-His_6-D_4K-EGFP, lane 1: size marker, lane 2: cell lysate diluted 1/100, lanes 3-8: elution fractions 13-18. **D)** H-His_6-D_4K-EGFP lane 1: size marker, lane 2: cell lysate diluted 1/100, lanes 3-8: elution fractions 13-18. **E)** I His_6-D_4K-EGFP, lane 1: cell lysate diluted 1/100, lanes 2-3: elution fractions 13-14, lane 4: size marker, lanes 5-8: elution fractions 15-18.

6.4 Discussion

6.4.1 The low expression levels of fusion proteins containing cyclic tags

As discussed in Chapter 3, low expression levels can be caused by numerous factors including secondary structure formation at the 5' end of the mRNA[132]. Alternatively properties of the tags at the protein level could be responsible for the low expression levels. The expression levels of F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP were higher than for the other recombinant proteins containing an engineered disulfide bridge (Figure 6.4). F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP all had the Cys residues arranged Cys-Ala-[12 amino acid tag]-Cys (Table 6.3), as previously described by the Imperiali group[236; 237]. Hence this arrangement of the Cys residues appeared to favour the expression of recombinant EGFP with a cyclic tag fused to the N-terminus of the protein. An exception was the fusion protein G-His_6-D_4K-EGFP which expressed lower than F-His_6-D_4K-EGFP, H-His_6-D_4K-EGFP and I-His_6-D_4K-EGFP even if the arrangement of the Cys residues was the same (Figure 6.4).

Interestingly, the Imperiali group reported no negative impact on protein expression levels by introducing a disulfide constrained cyclic tag. The expressed fusion protein was His_6-[cyclic tag]-Ubiquitin[237]. It is therefore possible that expression is diminished by a cyclic tag if it is arranged at the N-terminus of the fusion protein rather than between the protein and another tag. The expression levels of the [cyclic tag]-His_6-D_4K-EGFP fusion proteins were not increased by changing the temperature of expression to 25 °C or to 42 °C (data not shown). A possible strategy to increase the expression of fusion proteins containing a disulfide constrained cyclic tag could be to arrange the tag between the protein and another tag. Alternatively the cyclic tag could be fused to the C-terminus of EGFP.

6.4.2 Purification

Small scale purification of all fusion proteins was successfully carried out by Ni^{2+}-NTA chromatography. Both the cyclic and the linear form of the tags could be observed. The level of purity was however < 90 % based on SPS-PAGE (data not shown). Furthermore the yields were too low to facilitate the Ca^{2+} and Tb^{3+} binding experiments described in Section 6.3.6, 6.3.7 and 6.3.8.

The strong binding of the cyclic tagged proteins onto HisTrap™ columns and MonoQ™ resins could be related to the Cys residues in the tags interacting with functional groups on the resins. Cys residues can potentially bind to immobilized metal ions in their reduced form[6]. The phenomenon was clearly observed here due to the fluorescence of EGFP,

but could potentially occur unnoticed with a non coloured protein. Stripping of the metal with EDTA followed by a 1 M NaOH wash was required to elute the protein from the HisTrap™ column. Similarly the MonoQ™ column required regeneration with 1 M NaOH followed by 6 M GuHCl to remove the bound protein. The strong binding of the linear reduced cyclic tags could potentially be used as a purification strategy where all protein tagged with non-cyclic tags containing free Cys residues was irreversibly bound to the IMAC column. If the bound proteins containing free Cys residues were difficult to remove during column regereration, it could however be a very material consuming strategy. The strategy was therefore not further investigated herein.

6.4.3 Ca^{2+}-binding experiments

ESI-MS proved to be a rapid assay to qualitatively study Ca^{2+} binding. Differences in deconvoluted masses of 38 ± 2 Da could be used to assess if Ca^{2+} binding occurred. As the ionization of proteins is suppressed by salts, the successive adding of Ca^{2+} leads to a decrease in the signal intensity leading to difficulties of detection at higher metal concentrations. As a consequence metal binding can potentially be difficult to observe for proteins with low Ca^{2+} affinity, when higher Ca^{2+} concentrations are required for binding events to occur. Moreover, different proteins ionize to different extent. In the results obtained herein, the spectra of calmodulin and EGFP-HIT2 were of high quality compared to spectra of the other fusion proteins. Hence these samples could generate spectra of sufficiently high quality at higher Ca^{2+} concentrations compared to other samples.

The cyclic tags exist in one reduced linear and one oxidized cyclic form (Section 6.3.4). As ESI-MS is a quantitative method, the reduced species of the fusion proteins containing a cyclic tag could act as a contaminant. Signals in the mass spectra showing binding of Ca^{2+} to the oxidized species of the fusion protein could therefore be too low to be detected. Hence Ca^{2+} binding could potentially have occurred without being detected.

6.4.4 Tb^{3+}-binding experiments

Measuring Tb^{3+} binding through LRET proved to be a quick experimental procedure consuming low amounts of recombinant protein compared to the Ca^{2+} selective electrode method. Hence Tb^{3+} binding using LRET can serve as a relatively fast and simple assay when investigating if novel affinity tags have the potential of binding Ca^{2+}. There are however two potential drawbacks of this method.

Firstly, the energy transfer to the chelated Tb^{3+} requires the presence of sensitizing amino acid residues at or near the Tb^{3+} binding site. Hence the lack of a signal might not reflect lack of binding of Tb^{3+} but could also potentially originate in the lack of energy transfer between excited amino acid residues and the chelated Tb^{3+}. In this Chapter, Tb^{3+} was

shown to bind to F-His_6-D_4K-EGFP (Figure 6.14) but not to H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP and EGFP-HIT2 (Figure 6.14A and Figure 6.16). As mentioned in Section 6.3.8, the cyclic tag F contains a Trp residue at position 7 and Tyr residues at positions 2 and 8, shown to optimize the Tb^{3+} luminescence of the tag[236]. It can therefore not be concluded if the lack of a LRET signal with H-His_6-D_4K-EGFP, I-His_6-D_4K-EGFP and EGFP-HIT2 was due to actual lack of binding or to the absence of Tb^{3+} sensitizing amino acids in optimized positions of these tags.

Secondly, mutations of single amino acids in disulfide constrained Ca^{2+} binding loops have previously been shown to alter their metal affinities[234]. Introducing Tb^{3+} sensitizing amino acid residues might therefore potentially affect the Ca^{2+} affinity of putative Ca^{2+} binding tags. Thus tags optimized for binding Tb^{3+} might not be ideal for binding Ca^{2+}, even if Ca^{2+} and lanthanides generally are considered to have very similar coordination properties[238].

6.5 Conclusions

In conclusion disulfide constrained putative Ca^{2+} binding cyclic tags were generated with the position of the Cys residues forming the disulfide bond in three different arrangements with respect to the 12 amino acid Ca^{2+} binding sequence. The arrangement Cys-Ala-[12 amino acid tag]-Cys was giving the highest expression. Three such cyclic tags were purified and their affinity towards Ca^{2+} was analysed using ESI-MS and a Ca^{2+} selective electrode. The cyclic tags were compared with the linear HIT2 tag. No binding of Ca^{2+} to the cyclic tags and HIT2 could be detected by the two assays employed. Tb^{3+} binding could however be measured through the LRET between sensitized chelated Tb^{3+} and EGFP. The binding was stronger for a cyclic tag as compared to the corresponding reduced linear sequence. Cyclic tags therefore present promising candidates as Ca^{2+} binding tags in IMAC. None of the cyclic tags bound to the resins in preliminary Ca^{2+}-IMAC experiments.

Ca^{2+} affinities can vary dramatically between different Ca^{2+} binding sites in Ca^{2+} binding proteins[210; 214]. It is therefore likely that a cyclic tag with high Ca^{2+} affinity could be identified from a larger library of disulfide bond constrained putative Ca^{2+} binding tags.

The amounts of protein required for the Tb^{3+} binding assays in 96-well format were relatively low. This method could therefore be employed for screening larger numbers of putative Ca^{2+} binding tags.

Chapter 7. Discussion and conclusion

The overall aim in this thesis was to contribute to the development of more environmentally benign IMAC systems. Current IMAC systems, based on borderline metal ions according to the classification of Pearson[5], are known to leak[37; 38; 39] metal ions. The leakage of borderline metal ions such as Ni^{2+}, Cu^{2+} and Co^{2+} can be considered problematic from a green chemistry perspective due to their toxicity[22; 23], carcinogenity[25] and allergenicity[24]. To address the overall aim, two strategies were employed. The first strategy was to use an IMAC system employing borderline metal ions based around the macrocycle tacn. The high[40; 41; 42] stability constants of Ni^{2+} and Cu^{2+} chelates of tacn and its derivatives ($16 < \log \beta < 30$) compared to non macrocyclic ligands like IDA and NTA[28] ($8 < \log \beta < 13$) would minimize metal ion leakage from such IMAC systems. The toxicity of metal ions has been correlated to their hardness/softness with hard metal ions being less toxic than soft metal ions[22; 23]. The second strategy was therefore to contribute to the development of an IMAC system based on the hard metal ion Ca^{2+}.

The first strategy was addressed by characterizing the borderline metal IMAC systems based on the tacn motif and the NT1A tag, previously developed at the Monash University Centre for Green Chemistry[67; 68]. In Chapter 3 the influence of different arrangements of the borderline NT1A tag, and the putative hard metal binding tag HIT2, was investigated. EGFP fusion proteins containing the NT1A and HIT2 tags in sequence were studied to facilitate sequential purification with a borderline IMAC system followed by a hard metal system or *vice versa*. The differences in expression levels observed with N-terminally tagged fusion proteins were not observed when the tags were instead fused to the C-terminus of EGFP (Chapter 3, Section 3.3.3). The chromatographic behaviour of the fusion proteins were however affected slightly. C-terminally tagged EGFP fusion proteins eluted at a slightly lower imidazole concentration compared to their N-terminal counterparts on Ni^{2+}-NTA-Agarose, Ni^{2+}-tacn-Sepharose 6 F F, Ni^{2+}-propyl-*bis*-tacn-Sepharose 6 F F, Cu^{2+}-tacn-Sepharose 6 F F and Cu^{2+}-propyl-*bis*-tacn-Sepharose 6 F F (Chapter 3, Table 3.4). The purity of EGFP-His_6 on Cu^{2+}-tacn-Sepharose 6 F F was > 90 % as judged by SDS-PAGE (Chapter 3, Figure 3.9B). Thus purification similar to that obtained with "classic" IMAC systems such as Ni^{2+}-IDA-Sepharose 6 F F and Ni^{2+}-NTA-Agarose could be obtained with Cu^{2+}-tacn-Sepharose 6 F F.

In Chapter 4 it was shown that the tags NT1A and HIT2 did not alter the structure and stability of their partner EGFP. These results verify the suitability of the NT1A and HIT2 tags in IMAC. Hence the characterization of the borderline metal IMAC systems based on tacn and its derivatives led to the conclusion that these systems can be used as alternatives to classic IMAC systems based on ligands such as IDA and NTA.

In chapter 5 and 6 key parameters of Ca^{2+} based IMAC systems were studied. In contrast to borderline metal IMAC, hard metal IMAC has not gained widespread acceptance and is currently not routinely used for protein purification. As discussed in Section 7.1.2, the mechanism of Ca^{2+} based IMAC is not currently well understood. Based on the limited knowledge in the field of Ca^{2+} based IMAC, the work in this thesis was focused on the fundamental principles of such a system.

In Chapter 5 one component of the Ca^{2+} based IMAC system – the metal chelate – was investigated. The immobilized Ca^{2+} chelates were studied under chromatographic conditions. The immobilized metal chelates Ca^{2+}-DO3A-Sepharose 6 F F, Ca^{2+}-alendronate-Sepharose 6 F F, Ca^{2+}-DO3P-Sepharose 6 F F, Ca^{2+}-DO2P-Sepharose 6 F F, Ca^{2+}-DO2A-Sepharose 6 F F, Ca^{2+}-IDA-Sepharose 6 F F and Ca^{2+}-cyclen-Sepharose 6 F F were investigated. It was found that salts such as NaCl in 10 mM HEPES, pH 7.5 displaced Ca^{2+} from all immobilized chelates (Chapter 5, Figure 5.4). Previous studies with Ca^{2+} binding proteins concluded that the displacement of Ca^{2+} was due to the protein forming more stable Ca^{2+} complexes compared to the immobilized ligand[90]. Herein, no MIT of Ca^{2+} from Ca^{2+}-DO3P-Sepharose 6 F F was observed when the chelate was subjected to *E. coli* cell lysates containing EGFP-HIT2 (Chapter 5, Figure 5.7). The finding that displacement of Ca^{2+} is caused by salts in the buffers is therefore an important contribution to the field of Ca^{2+} based IMAC.

The macrocyclic ligands DO3A and DO3P formed more stable immobilized Ca^{2+} chelates as compared to the corresponding ligands containing only two pendant arms (DO2A and DO2P). Ca^{2+}-DO3P-Sepharose 6 F F was stable between pH 7.5 and 9.5 while Ca^{2+}-DO3A-sepharose was stable between pH 5.5 and 9.5. At pH 7.5 Ca^{2+}-alendronate-Sepharose 6 F F showed promising results with relatively low displacement of Ca^{2+} as a function of NaCl. Ca^{2+}-DO3A-Sepharose 6 F F, Ca^{2+}-DO3P-Sepharose 6 F F and Ca^{2+}-alendronate-Sepharose 6 F F were identified as potential ligands in a Ca^{2+}-based IMAC system. The results obtained herein point to substantial differences between Ca^{2+} based IMAC systems and borderline metal IMAC systems. While salt concentrations in the range 0.1 M – 0.3 M displace a significant proportion of the Ca^{2+} from immobilized Ca^{2+}-chelates, these salt concentrations (and higher) are routinely used in borderline metal IMAC systems.

In Chapter 6 the second critical component of a Ca^{2+} based IMAC system – the affinity tag - was studied. From the literature[212; 227; 228; 229; 230; 231; 234; 237] cyclic peptides based on

the 12 amino acid Ca^{2+} binding loop of the EF-hand were identified as putative Ca^{2+} binding tags. Such tags were therefore created by introducing cysteine (Cys) residues forming a disulfide bridge. These cyclic tags were shown to bind Tb^{3+}, a metal commonly employed to study Ca^{2+} binding, with log β = 5.1 (Chapter 6, Section 6.3.8). No binding was observed for the corresponding linear tag. Hence, cyclic tags are promising candidates as Ca^{2+} binding tags.

One difficulty with developing a Ca^{2+} based IMAC system was the lack of positive controls. In Ca^{2+}-IMAC experiment, if binding was not observed, this could potentially be due to the ligand, the tag or the buffer compositions. The results obtained about the ligands (Chapter 5) and the affinity tags (Chapter 6) constitute valuable knowledge regarding the fundamentals of Ca^{2+}-based IMAC systems. With these results, the potential choices of chromatographic conditions can be significantly reduced.

As mentioned in the Introduction (Section 1.3), only a handful of papers describe the use of Ca^{2+} based IMAC. The early papers all describe isolation of naturally occurring Ca^{2+} binding proteins. The separation of *Dolichos biflorus* seed lectin and a carbohydrate binding protein on Ca^{2+}-IDA-agarose reported by Borrebaeck *et al.*[85; 86] in 1981 and 1984 involved equilibration of the Ca^{2+} charged gel with Tris-acetate buffer at pH 8.2 containing 0.5 M NaCl. In Chapter 5 it was shown that NaCl in 10 mM HEPES, pH 7.5 displace significant amounts of Ca^{2+} from IDA-Sepharose 6 F F (Figure 5.4). This can be attributed to IDA acting as a cation exchanger. It is hence possible that a significant proportion of the bound Ca^{2+} was displaced from the IDA-agarose by the NaCl in the wash solution in the studies by Borrebaeck *et al*[85; 86]. The authors achieved elution of the target proteins by using the binding buffer supplemented with EDTA[85; 86]. Elution could then be achieved by (at least) two mechanisms. Firstly, the Ca^{2+} chelated by IDA-Agarose could be chelated by EDTA, thereby breaking the bonds between the protein(s) and the immobilized metal chelate. It is also possible that Ca^{2+} was bound to the proteins themselves. If displacement of bound Ca^{2+} from the proteins caused them to elute from the Ca^{2+}-IDA-agarose, then the absorption and desorption process could be dependent on the tertiary structure of the protein, since many Ca^{2+} binding proteins change conformation upon Ca^{2+} binding[210]. This would suggest that retention could occur through other modes of chromatography than IMAC, such as ion exchange chromatography or hydrophobic interaction chromatography.

Mantovaara *et al.*[87] purified human serum amyloid P component (SAP) using Ca^{2+} chelated to carboxymethylated aspartic acid (CM-Asp), Ca^{2+}-CM-Asp-agarose. The authors used 20 mM imidazole, 250 mM $CaCl_2$ at pH 7.0 as the binding buffer. Elution was achieved with 20 mM Na-acetate-acetic acid at pH 5.0. The authors proposed that lowering the pH would displace Ca^{2+} from the gel, causing the protein to elute by metal ion transfer (MIT). In their next study[88] Mantovaara *et al*. isolated factor VIII, again on Ca^{2+}-CM-Asp-agarose. The binding buffer was 20 mM Na-acetate-acetic acid, 0.5 M K_2SO_4, pH 5.5. The protein was

eluted in 20 mM Na-acetate-acetic acid pH 5.5. The Ca^{2+} content of the gel was however determined[88] to be 133 µmol g^{-1}. Hence the occurrence of MIT can not have been extensive, as the Ca^{2+} content of the gel was still significant.

In a more recent paper[90] the Porath group described purification of calmodulin using Mn^{2+}, La^{3+}, Nd^{3+} and Eu^{3+} instead of Ca^{2+}, chelated to TED-agarose. The best results were obtained with Eu^{3+}-TED-agarose. In these experiments the protein was loaded in 0.1 M Tris-HCl, 2.0 M NaCl, 0.2 M $CaCl_2$ at pH 7.5. Unbound proteins were removed by two wash steps with 0.1 M Tris-HCl, 0.6 M Na_2SO_4, 2.0 M NaCl, pH 7.5 and 0.1 M Tris-HCl, 40 mM malonate, 0.6 M Na_2SO_4, 2.0 M NaCl, pH 7.5 respectively. Elution was achieved by 0.2 M potassium phosphate, 3.0 M NaCl, 0.2 M citrate, pH 7.5. The authors[90] proposed two mechanism of retention where several calmodulin molecules bound to each other in the presence of Ca^{2+}. One calmodulin molecule would bind to Eu^{3+}-TED-agarose. Other calmodulin molecules would in turn bind to the immobilized calmodulin. The authors argued that the calmodulin molecules could be bound via Ca^{2+}, thus forming a calmodulin-Ca^{2+}-calmodulin linkage. Alternatively calmodulin molecules would bind to each other via binding sites exposed in the presence of Ca^{2+}. The proposed mechanism of elution was displacement of Ca^{2+} by Ca^{2+} chelators such as citrate, phosphate and malonate.

Indeed calmodulin and several other EF-hand proteins are known to expose a hydrophobic patch upon Ca^{2+} binding[210]. This property has been exploited to purify calmodulin with HIC as described in Section 6.3.5. The Ca^{2+} dependent hydrophobicity of calmodulin has even led to the use of the protein as an affinity tag[66]. Hence, the Ca^{2+} dependent retention of calmodulin to Eu^{3+}-TED-agarose observed by Chaga *et al.*[90] could be attributed to Ca^{2+} dependent hydrophobic interactions between calmodulin and the resin.

Based on the results obtained in this thesis (Chapter 5 and 6) and the discussion above, it is evident that further work is required to develop a Ca^{2+} based IMAC system. The development of such a system might require an entirely different approach compared to that taken when developing IMAC systems based on borderline metal ions. A number of strategies which could be employed to develop a Ca^{2+} based IMAC system are considered below.

A large library of disulfide constrained cyclic tags could be cloned. Their Tb^{3+} affinity could be screened as described in Chapter 6, Section 6.3.8.

New ligands could be designed, immobilized and screened with buffers as described in Chapter 5. The aim would be to identify an immobilized Ca^{2+} chelate where Ca^{2+} is not displaced by buffer components such as NaCl.

A Ca^{2+}-based IMAC system could be created where the ligand mimicks a half structure of the 12 amino acid sequence in the EF hand of Ca^{2+}-binding proteins. The affinity tag could constitute the other half. Ca^{2+} binding would then be obtained in the presence of Ca^{2+} and the binding could be broken by introducing EGTA or EDTA.

In conclusion it was shown in Chapters 3 and 4 that the NT1A and HIT2 tags could be employed both at the N- and C-terminus of EGFP. The chromatographic properties were not altered significantly by moving the tags from the N to the C-terminus of EGFP. The tags did not affect the structure and stability of EGFP and are thus suitable as affinity tags. Significant knowledge was gained regarding both the properties of immobilized Ca^{2+} chelates and the requirements for putative Ca^{2+} binding tags (Chapters 5 and 6). It was shown that immobilized metal chelates are sensitive to salt and in some cases pH, and that putative Ca^{2+} binding tags showed greater affinities towards Tb^{3+} when constrained.

Reference list

1. Kastner, M. (1999). *Protein Liquid Chromatography*. Journal of Chromatography Library, Elsevier Science & Technology.
2. Hage, D. S. (2006). *Handbook of Affinity Chromatography* (Cazes, J., Ed.), 92. 92 vols, CRC Press, Taylor & Francis Group, Boca Raton.
3. Everson, R. J. & Parker, H. E. (1974). Zinc binding and synthesis eight-hydroxy-quinoline-agarose. *Bioinorganic Chemistry* **4**, 15-20.
4. Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-9.
5. Pearson, R. G. (1990). Hard and soft acids and bases-the evolution of a chemical concept. *Coordin. Chem. Rev.* **100**, 403-425.
6. Gaberc-Porekar, V. & Menart, V. (2001). Perspectives of immobilized-metal affinity chromatography. *J. Biochem. Biophys. Methods* **49**, 335-360.
7. Chaga, G. S. (2001). Twenty-five years of immobilized metal ion affinity chromatography: past, present and future. *J. Biochem. Biophys. Methods* **49**, 313-34.
8. Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biot.* **60**, 523-533.
9. Hochuli, E. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology* **6**, 1321.
10. Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H. W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S. H., Rao, Z., Shi, Y., Terwilliger, T. C., Kim, C. Y., Hung, L. W., Waldo, G. S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J. L., Stevens, R. C., Lesley, S. A., Wilson, I. A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M. I., Eschenfeldt, W. H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S. K., Emtage, J. S., Sauder, J. M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S. C., Bonanno, J. B., Fiser, A., Swaminathan, S., Studier, F. W., Chance, M. R., Sali, A., Acton, T. B., Xiao, R., Zhao, L., Ma, L. C., Hunt, J. F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C. K., Wang, D., Wang, H., Jiang, M., Montelione, G. T., Stuart, D. I., Owens, R. J., Daenke, S., Schutz, A., Heinemann, U., Yokoyama, S., Bussow, K. & Gunsalus, K. C. (2008). Protein production and purification. *Nat. Methods* **5**, 135-46.

11. Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expres. Purif.* **48**, 1-13.
12. Walters, R. R. (1985). Affinity-Chromatography. *Anal. Chem.* **57**, 1099-1114.
13. Sundberg, L. & Porath, J. (1974). Preparation of Adsorbents for Biospecific Affinity Chromatography .1. Attachment of Group-Containing Ligands to Insoluble Polymers by Means of Bifunctional Oxiranes. *J. Chromatogr.* **90**, 87-98.
14. Porath, J., Låås, T. & Janson, J.-C. (1975). Agar derivatives for chromatography, electrophoresis and gel-bound enzymes : III. Rigid agarose gels cross-linked with divinyl sulphone (DVS). *J. Chromatogr. A* **103**, 49-62.
15. Axen, R. (1971). Ribonuclease chemically attached to beads of epichlorohydrin cross-linked agarose. *Enzymologia* **41**, 359-364.
16. Crawford, J., Ramakrishnan, S., Periera, P., Gardner, S., Coleman, M. & Beitle, R. (1999). Immobilized metal affinity membrane separation: Characteristics of two materials of differing preparation chemistries. *Separ. Sci. Technol.* **34**, 2793-2802.
17. Pearson, R. G. (1963). Hard and soft acids and bases. *J. Am. Chem. Soc.* **85**, 3533-3539.
18. Pearson, R. G. (1986). Absolute electronegativity and hardness correlated with molecular orbital theory. *Proc. Natl. Acad. Sci. U S A* **83**, 8440-8441.
19. Ahrland, S. (1968). Thermodynamics of complex formation between hard and soft acceptors and donors. *Structure and Bonding* **5**, 118-49.
20. Pearson, R. G. & Mawby, R. J. (1967). Nature of metal-halogen bonds. *Halogen Chem.* **3**, 55-84.
21. Pearson, R. G. (1989). Absolute electronegativity and hardness: applications to organic chemistry. *J. Org. Chem.* **54**, 1423-30.
22. Jones, M. M. & Vaughn, W. K. (1978). HSAB theory and acute metal ion toxicity and detoxification processes. *J. Inorg. Nucl. Chem.* **40**, 2081-2088.
23. McCloskey, J. T. J. (1996). Predicting the relative toxicity of metal ions using ion characteristics: Microtox bioluminescence assay. *Environ. Toxicol. Chem.* **15**, 1730.
24. Thyssen, J. P. (2010). Metal allergys-A Review on Exposures, Penetration, Genetics, Prevalence, and Clinical Implications. *Chem. Res. Toxicol.* **23**, 309-318.
25. Bal, W., Kozlowski, H. & Kasprzak, K. S. (2000). Molecular models in nickel carcinogenesis. *J. Inorg. Biochem.* **79**, 213-8.
26. Cabbiness, D. K. & Margerum, D. W. (2002). Macrocyclic effect on the stability of copper(II) tetramine complexes. *J. Am. Chem. Soc.* **91**, 6540-6541.
27. Martell, A. E. & Smith, R. M. (1974). *Critical stability constants*, Plenum Press, New York.

28. Motekaitis, R. (2004). NIST CRITICALLY SELECTED STABILITY CONSTANTS OF METAL COMPLEXES. In *NIST Standard Reference Database 46* 8.0 edit. NIST, Gaithersburg.
29. Cotton, F. A. (1999). *Advanced inorganic chemistry*. 6th edit, John Wiley & Sons, New York.
30. Porath, J. & Olin, B. (1983). Immobilized metal ion affinity adsorption and immobilized metal ion affinity chromatography of biomaterials. Serum protein affinities for gel-immobilized iron and nickel ions. *Biochemistry* **22**, 1621-30.
31. Hochuli, E., Dobeli, H. & Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.* **411**, 177-84.
32. Chaga, G., Hopp, J. & Nelson, P. (1999). Immobilized metal ion affinity chromatography on Co2+-carboxymethylaspartate-agarose Superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle. *Biotechnol. Appl. Bioc.* **29 (Pt 1)**, 19-24.
33. Chaouk, H. & Hearn, M. T. W. (1999). New ligand, N-(2-pyridylmethyl)aminoacetate, for use in the immobilised metal ion affinity chromatographic separation of proteins. *J. Chromatogr. A* **852**, 105-115.
34. Zachariou, M., Traverso, I. & Hearn, M. T. (1993). High-performance liquid chromatography of amino acids, peptides and proteins. CXXXI. O-phosphoserine as a new chelating ligand for use with hard Lewis metal ions in the immobilized-metal affinity chromatography of proteins. *J. Chromatogr.* **646**, 107-20.
35. Zachariou, M. & Hearn, M. T. V. (1992). High-performance liquid chromatography of amino acids, peptides and proteins : CXXI. 8-Hydroxyquinoline-metal chelate chromatographic support: an additional mode of selectivity in immobilized-metal affinity chromatography. *J. Chromatogr. A* **599**, 171-177.
36. Shahwan, G. J. & Jezorek, J. R. (1983). Liquid chromatography of phenols on an 8-quinolinol silica gel--iron(III) stationary phase. *J. Chromatogr. A* **256**, 39-48.
37. Oswald, T., Hornbostel, G., Rinas, U. & Anspach, F. B. (1997). Purification of (His)6EcoRV [recombinant restriction endonuclease EcoRV fused to a (His)6 affinity domain] by metal-chelate affinity chromatography. *Biotechnol Appl Biochem* **25 (Pt 2)**, 109-15.
38. Nakagawa, Y., Yip, T. T., Belew, M. & Porath, J. (1988). High-performance immobilized metal ion affinity chromatography of peptides: analytical separation of biologically active synthetic peptides. *Anal. Biochem.* **168**, 75-81.
39. Jiang, W. & Hearn, M. T. (1996). Protein interaction with immobilized metal ion affinity ligands under high ionic strength conditions. *Anal. Biochem.* **242**, 45-54.

40. Haidar, R., Ipek, M., DasGupta, B., Yousaf, M. & Zompa, L. J. (1997). Copper(II) Complexes of Bis(1,4,7-triazacyclononane) Ligands with Polymethylene Bridging Groups: An Equilibrium and Structural Study. *Inorg. Chem.* **36**, 3125-3132.
41. Tanaka, N., Kobayashi, Y. & Takamoto, S. (1977). Synthesis of a New Hexadentate Ligand and Measurement of Stability-Constants of Its Metal-Chelates. *Chem. Lett.*, 107-108.
42. Zhang, X. P., Hsieh, W. Y., Margulis, T. N. & Zompa, L. J. (1995). Binuclear Copper(II) Complexes of Bis(1,4,7-Triazacyclononane) Ligands Containing Trimethylene and Tetramethylene Bridging Groups - an Equilibrium and Structural Study. *Inorg. Chem.* **34**, 2883-2888.
43. Jiang, W., Graham, B., Spiccia, L. & Hearn, M. T. W. (1998). Protein Selectivity with Immobilized Metal Ion-tacn Sorbents: Chromatographic Studies with Human Serum Proteins and Several Other Globular Proteins. *Anal. Biochem.* **255**, 47-58.
44. Zompa, L. J. (1978). Metal complexes of cyclic triamines. 2. Stability and electronic spectra of nickel(II), copper(II), and zinc(II) complexes containing nine- through twelve-membered cyclic triamine ligands. *Inorg. Chem.* **17**, 2531-2536.
45. Graham, B., Comba, P., Hearn, M. T. & Spiccia, L. (2007). An examination of the binding behavior of histidine-containing peptides with immobilized metal complexes derived from the macrocyclic ligand, 1,4,7-triazacyclononane. *J. Biol. Inorg. Chem.* **12**, 11-21.
46. Graham, B., Spiccia, L. & Hearn, M. T. W. (2008). Examination of the binding behaviour of several proteins with the immobilized copper(II) complexes of o-, m- and p-xylylene bridged bis(1,4,7-triazacyclononane) macrocycles. *J. Chromatogr. A* **1194**, 30-37.
47. Jiang, W., Prescott, M., Devenish, R. J., Spiccia, L. & Hearn, M. T. (2009). Separation of hexahistidine fusion proteins with immobilized metal ion affinity chromatographic (IMAC) sorbents derived from M(N+)-tacn and its derivatives. *Biotechnol Bioeng* **103**, 747-56.
48. Waugh, D. S. (2005). Making the most of affinity tags. *Trends Biotechnol.* **23**, 316-320.
49. Hearn, M. T. & Acosta, D. (2001). Applications of novel affinity cassette methods: use of peptide fusion handles for the purification of recombinant proteins. *J. Mol. Recognit.* **14**, 323-69.
50. Davis, G. D., Elisee, C., Newham, D. M. & Harrison, R. G. (1999). New fusion protein systems designed to give soluble expression in Escherichia coli. *Biotechnol Bioeng* **65**, 382-8.

51. Zhang, Y. B., Howitt, J., McCorkle, S., Lawrence, P., Springer, K. & Freimuth, P. (2004). Protein aggregation during overexpression limited by peptide extensions with large net negative charge. *Protein Expres. Purif.* **36**, 207-16.
52. Kapust, R. B. & Waugh, D. S. (1999). Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8**, 1668-74.
53. Smith, M. C., Furman, T. C., Ingolia, T. D. & Pidgeon, C. (1988). Chelating Peptide-Immobilized Metal-Ion Affinity-Chromatography - a New Concept in Affinity-Chromatography for Recombinant Proteins. *J. Biol. Chem.* **263**, 7211-7215.
54. Ljungquist, C., Breitholtz, A., Brink-Nilsson, H., Moks, T., Uhlen, M. & Nilsson, B. (1989). Immobilization and affinity purification of recombinant proteins using histidine peptide fusions. *Eur. J. Biochem.* **186**, 563-9.
55. Listwan, P., Martin, J., Kobe, B., et al. (2004). Methods for high-throughput protein expression, purification and structure determination adapted for structural genomics. *Australian Biochemist* **35**, 43-46.
56. Chaga, G., Bochkariov, D. E., Jokhadze, G. G., Hopp, J. & Nelson, P. (1999). Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt(II)-carboxymethylaspartate crosslinked agarose. *J. Chromatogr. A* **864**, 247-56.
57. Enzelberger, M. M., Minning, S. & Schmid, R. D. (2000). Designing new metal affinity peptides by random mutagenesis of a natural metal-binding site. *J. Chromatogr. A* **898**, 83-94.
58. Fujii, I., Takaoka, Y., Suzuki, K. & Tanaka, T. (2001). A conformationally purified [alpha]-helical peptide library. *Tetrahedron Lett.* **42**, 3323-3325.
59. Beitle, R. R. & Ataai, M. M. (1993). One-step purification of a model periplasmic protein from inclusion bodies by its fusion to an effective metal-binding peptide. *Biotechnol Prog* **9**, 64-9.
60. Sloane, R. P., Ward, J. M., O'Brien, S. M., Thomas, O. R. T. & Dunnill, P. (1996). Expression and purification of a recombinant metal-binding T4 lysozyme fusion protein. *J. Biotechnol.* **49**, 231-238.
61. Smith, D. B. & Johnson, K. S. (1988). Single-Step Purification of Polypeptides Expressed in Escherichia-Coli as Fusions with Glutathione S-Transferase. *Gene* **67**, 31-40.
62. Ullmann, A. (1984). One-step purification of hybrid proteins which have beta-galactosidase activity. *Gene* **29**, 27-31.
63. Moks, T., Abrahmsen, L., Holmgren, E., Bilich, M., Olsson, A., Uhlen, M., Pohl, G., Sterky, C., Hultberg, H., Josephson, S. & et al. (1987). Expression of human insulin-

like growth factor I in bacteria: use of optimized gene fusion vectors to facilitate protein purification. *Biochemistry* **26**, 5239-44.

64. Nilsson, B. & Abrahmsen, L. (1990). Fusions to staphylococcal protein A. *Methods Enzymol* **185**, 144-61.
65. Uhlen, M., Nilsson, B., Guss, B., Lindberg, M., Gatenbeck, S. & Philipson, L. (1983). Gene fusion vectors based on the gene for staphylococcal protein A. *Gene* **23**, 369-78.
66. McCluskey, A. J., Poon, G. M. K. & Gariepy, J. (2007). A rapid and universal tandem-purification strategy for recombinant proteins. *Protein Sci.* **16**, 2726-2732.
67. Christensen, T., Hearn, M. T. W., Spiccia, L., Jiang, W., Mooney, J. T. & Graham, B. (2003). Peptide purification by means of metal ion affinity chromatography, pp. 193 pp, WO 2003042249.
68. Mooney, J. T., Monash University. & Monash University. Dept. of Biochemistry and Molecular Biology. (2005). Development of a novel IMAC system using the MCL tacn, pp. ix, 420 leaves.
69. Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F. & Fletterick, R. J. (1989). Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. *Virology* **172**, 302-10.
70. Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Pat Cerretti, D., Urdal, D. L. & Conlon, P. J. (1988). A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. *Nat Biotech* **6**, 1204-1210.
71. Schmidt, T. G., Koepke, J., Frank, R. & Skerra, A. (1996). Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *J. Mol. Biol.* **255**, 753-66.
72. Zwicker, N., Adelhelm, K., Thiericke, R., Grabley, S. & Hanel, F. (1999). Strep-tag II for one-step affinity purification of active bHLHzip domain of human c-Myc. *Biotechniques* **27**, 368-75.
73. Chen, Y. T., Holcomb, C. & Moore, H. P. (1993). Expression and localization of two low molecular weight GTP-binding proteins, Rab8 and Rab10, by epitope tag. *Proc. Natl. Acad. Sci. U S A* **90**, 6508-12.
74. Kim, J. S. & Raines, R. T. (1993). Ribonuclease S-peptide as a carrier in fusion proteins. *Protein Sci.* **2**, 348-56.
75. Raines, R. T., McCormick, M., Van Oosbree, T. R. & Mierendorf, R. C. (2000). The S.Tag fusion system for protein purification. *Methods Enzymol* **326**, 362-76.
76. Zheng, C. F., Simcox, T., Xu, L. & Vaillancourt, P. (1997). A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins. *Gene* **186**, 55-60.

77. Wyborski, D. L., Bauer, J. C., Zheng, C. F., Felts, K. & Vaillancourt, P. (1999). An Escherichia coli expression vector that allows recovery of proteins with native N-termini from purified calmodulin-binding peptide fusions. *Protein Expres. Purif.* **16**, 1-10.

78. Stofko-Hahn, R. E., Carr, D. W. & Scott, J. D. (1992). A single step purification for recombinant proteins. Characterization of a microtubule associated protein (MAP 2) fragment which associates with the type II cAMP-dependent protein kinase. *FEBS Lett* **302**, 274-8.

79. Neri, D., de Lalla, C., Petrul, H., Neri, P. & Winter, G. (1995). Calmodulin as a versatile tag for antibody fragments. *Biotechnology (N Y)* **13**, 373-7.

80. Smith, J. C., Derbyshire, R. B., Cook, E., Dunthorne, L., Viney, J., Brewer, S. J., Sassenfeld, H. M. & Bell, L. D. (1984). Chemical synthesis and cloning of a poly(arginine)-coding gene fragment designed to aid polypeptide purification. *Gene* **32**, 321-7.

81. Sassenfeld, H. M. & Brewer, S. J. (1984). A Polypeptide Fusion Designed for the Purification of Recombinant Proteins. *Nat Biotech* **2**, 76-81.

82. Stubenrauch, K., Bachmann, A., Rudolph, R. & Lilie, H. (2000). Purification of a viral coat protein by an engineered polyionic sequence. *J. Chromatogr. B* **737**, 77-84.

83. Pedersen, J., Lauritzen, C., Madsen, M. T. & Weis Dahl, S. (1999). Removal of N-terminal polyhistidine tags from recombinant proteins using engineered aminopeptidases. *Protein Expres. Purif.* **15**, 389-400.

84. Kenig, M., Peternel, S., Gaberc-Porekar, V. & Menart, V. (2006). Influence of the protein oligomericity on final yield after affinity tag removal in purification of recombinant proteins. *J. Chromatogr. A* **1101**, 293-306.

85. Borrebaeck, C. A. K., Lonnerdal, B. & Etzler, M. E. (1981). Metal chelate affinity chromatography of the Dolichos biflorus seed lectin and its subunits. *FEBS Lett.* **130**, 194-196.

86. Borrebaeck, C. A., Mattiasson, B. & Nordbring-Hertz, B. (1984). Isolation and partial characterization of a carbohydrate-binding protein from a nematode-trapping fungus. *J. Bacteriol.* **159**, 53-56.

87. Mantovaara, T., Pertoft, H. & Porath, J. (1989). Purification of human serum amyloid P component (SAP) by calcium affinity chromatography. *Biotechnol. Appl. Bioc.* **11**, 564-70.

88. Mantovaara, T., Pertoft, H. & Porath, J. (1991). Purification of factor VIII:c coagulant activity from rat liver nonparenchymal cell culture medium by immobilized metal ion affinity chromatography. *Biotechnol Appl Biochem* **13**, 120-6.

89. Mantovaara, T. T. (1991). Carboxymethylated aspartic acid agarose, a selective adsorbent for calcium-binding proteins, preliminary studies. *Biotechnol. Appl. Bioc.* **13**, 315.
90. Chaga, G. S., Ersson, B. & Porath, J. O. (1996). Isolation of calcium-binding proteins on selective adsorbents. Application to purification of bovine calmodulin. *J. Chromatogr. A* **732**, 261-9.
91. Baldwin, C., Cumming, J. & Timperman, A. T. (2005). Isolation of naturally occurring aluminium ligands using immobilized metal affinity chromatography for analysis by ESI-MS. *Analyst* **130**, 318-324.
92. Zachariou, M. & Hearn, M. T. W. (1995). Protein Selectivity in Immobilized Metal Affinity-Chromatography Based on the Surface Accessibility of Aspartic and Glutamic-Acid Residues. *J. Protein Chem.* **14**, 419-430.
93. Zachariou, M. (1996). Application of immobilized metal ion chelate complexes as pseudocation exchange adsorbents for protein separation. *Biochemistry* **35**, 202-211.
94. Zachariou, M. & Hearn, M. T. W. (2000). Adsorption and selectivity characteristics of several human serum proteins with immobilised hard Lewis metal ion-chelate adsorbents. *J. Chromatogr. A* **890**, 95-116.
95. Hearn, M. T. W., Spiccia, L., Daly, R. & Kreher, U. (2005). Peptide purification by means of hard metal ion affinity chromatography, pp. 50 pp. (Novo Nordisk A/S, Den.; Monash University). WO Patent 2005120700.
96. Chao, H., Bautista, D. L., Litowski, J., Irvin, R. T. & Hodges, R. S. (1998). Use of a heterodimeric coiled-coil system for biosensor application and affinity purification. *J. Chromatogr. B* **715**, 307-29.
97. Phillips, T. M. (1989). High-performance immunoaffinity chromatography. *Adv Chromatogr* **29**, 133-73.
98. Valenti, L. E., De Pauli, C. P. & Giacomelli, C. E. (2006). The binding of Ni(II) ions to hexahistidine as a model system of the interaction between nickel and His-tagged proteins. *J. Inorg. Biochem.* **100**, 192-200.
99. Nieba, L., Nieba-Axmann, S. E., Persson, A., Hamalainen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A. F. & Pluckthun, A. (1997). BIACORE Analysis of Histidine-Tagged Proteins Using a Chelating NTA Sensor Chip. *Anal. Biochem.* **252**, 217-228.
100. Lata, S., Reichel, A., Brock, R., Tampe, R. & Piehler, J. (2005). High-affinity adaptors for switchable recognition of histidine-tagged proteins. *J. Am. Chem. Soc.* **127**, 10205-10215.
101. Lindner, P., Guth, B., Wulfing, C., Krebber, C., Steipe, B., Muller, F. & Pluckthun, A. (1992). Purification of native proteins from the cytoplasm and periplasm of

Escherichia coli using IMAC and histidine tails: A comparison of proteins and protocols. *Methods* **4**, 41-56.

102. Ohya, Y., Uno, I., Ishikawa, T. & Anraku, Y. (1987). Purification and Biochemical-Properties of Calmodulin from Saccharomyces-Cerevisiae. *Eur. J. Biochem.* **168**, 13-19.

103. McRae, S. R., Brown, C. L. & Bushell, G. R. (2005). Rapid purification of EGFP, EYFP, and ECFP with high yield and purity. *Protein Expres. Purif.* **41**, 121-127.

104. Steiner, R. F., Juminaga, D., Albaugh, S. & Washington, H. (1996). A comparison of the properties of the binary and ternary complexes formed by calmodulin and troponin C with two regulatory peptides of phosphorylase kinase. *Biophys. Chem.* **59**, 277-288.

105. Swain, M. & Ross, N. W. (1995). A silver stain protocol for proteins yielding high resolution and transparent background in sodium dodecyl sulfate-polyacrylamide gels. *Electrophoresis* **16**, 948-51.

106. Malmedal, A. (2001). Basic Instructions for Calculating Calcium-Binding Constants Using the Competitive Chelator Method.

107. André, I. & Linse, S. (2002). Measurement of Ca2+-Binding Constants of Proteins and Presentation of the CaLigator Software. *Anal. Biochem.* **305**, 195-205.

108. Pace, N. C. & Tanford, C. (1968). Thermodynamics of Unfolding of Beta-Lactoglobulin a in Aqueous Urea Solutions between 5 and 55 Degrees. *Biochemistry* **7**, 198-&.

109. Kieczykowski, G. R., Jobson, R. B., Melillo, D. G., Reinhold, D. F., Grenda, V. J. & Shinkai, I. (1995). Preparation of (4-Amino-1-Hydroxybutylidene)bisphosphonic Acid Sodium Salt, MK-217 (Alendronate Sodium). An Improved Procedure for the Preparation of 1-Hydroxy-1,1-bisphosphonic Acids. *J. Org. Chem.* **60**, 8310-8312.

110. Kovacs, Z. & Sherry, A. D. (1995). A general synthesis of 1,7-disubstituted 1,4,7,10-tetraazacyclododecanes. *J. Chem. Soc. Chem. Commun.*, 185-186.

111. Kovacs, Z. & Sherry, A. D. (1997). pH-Controlled Selective Protection of Polyaza Macrocycles. *Synthesis*, 759-763.

112. De León-Rodríguez, L. M., Kovacs, Z., Esqueda-Oliva, A. C. & Miranda-Olvera, A. D. (2006). Highly regioselective N-trans symmetrical diprotection of cyclen. *Tetrahedron Lett.* **47**, 6937-6940.

113. Dischino, D. D., Delaney, E. J., Emswiler, J. E., Gaughan, G. T., Prasad, J. S., Srivastava, S. K. & Tweedle, M. F. (1991). Synthesis of nonionic gadolinium chelates useful as contrast agents for magnetic resonance imaging: 1,4,7-tris(carboxymethyl)-10-substituted-1,4,7,10-tetraazacyclododecanes and their corresponding gadolinium chelates. *Inorg. Chem.* **30**, 1265-1269.

114. Sun, X., Wuest, M., Kovacs, Z., Sherry, A. D., Motekaitis, R., Wang, Z., Martell, A. E., Welch, M. J. & Anderson, C. J. (2003). In vivo behavior of copper-64-labeled methanephosphonate tetraaza macrocyclic ligands. *J. Biol. Inorg. Chem.* **8**, 217-25.
115. Richman, J. E. & Atkins, T. J. (1974). Nitrogen analogs of crown ethers. *J. Am. Chem. Soc.* **96**, 2268-2270.
116. Atkins, T. J., Richman, J. E. & Oettle, W. F. (1978). Macrocyclic Polyamines: 1,4,7,10,13,16-Hexaäzacycloöctadecane. *Org. Synth.* **58**, 86-97.
117. Tsunasawa, S., Stewart, J. W. & Sherman, F. (1985). Amino-terminal processing of mutant forms of yeast iso-1-cytochrome c. The specificities of methionine aminopeptidase and acetyltransferase. *J. Biol. Chem.* **260**, 5382-91.
118. Ben-Bassat, A., Bauer, K., Chang, S. Y., Myambo, K., Boosman, A. & Chang, S. (1987). Processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure. *J. Bacteriol.* **169**, 751-757.
119. Vassileva-Atanassova, A., Mironova, R., Nacheva, G. & Ivanov, I. (1999). N-terminal methionine in recombinant proteins expressed in two different Escherichia coli strains. *J. Biotechnol.* **69**, 63-67.
120. Broadwater, J. A. & Fox, B. G. (1999). Spinach Holo-Acyl Carrier Protein: Overproduction and Phosphopantetheinylation inEscherichia coliBL21(DE3),in VitroAcylation, and Enzymatic Desaturation of Histidine-Tagged Isoform I. *Protein Expres. Purif.* **15**, 314-326.
121. Levitt, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* **104**, 59-107.
122. Dalboge, H. (1990). In vivo processing of N-terminal methionine in E. coli. *FEBS Lett.* **266**, 1-3.
123. Casadaban, M. J. & Cohen, S. N. (1980). Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J. Mol. Biol.* **138**, 179-207.
124. Halliwell, C. M., Morgan, G., Ou, C.-P. & Cass, A. E. G. (2001). Introduction of a (Poly)histidine Tag in -Lactate Dehydrogenase Produces a Mixture of Active and Inactive Molecules. *Anal. Biochem.* **295**, 257-261.
125. Pietzsch, M., Wiese, A., Ragnitz, K., Wilms, B., Altenbuchner, J., Mattes, R. & Syldatk, C. (2000). Purification of recombinant hydantoinase and -N-carbamoylase from Arthrobacter aurescens expressed in Escherichia coli: comparison of wild-type and genetically modified proteins. *J. Chromatogr. B* **737**, 179-186.
126. Seidler, A. (1994). Introduction of a histidine tail at the N-terminus of a secretory protein expressed in Escherichia coli. *Protein Eng.* **7**, 1277-80.

127. Huang, C.-Y., Chao, Y.-P. & Yang, Y.-S. (2003). Purification of industrial hydantoinase in one chromatographic step without affinity tag. *Protein Expres. Purif.* **30**, 134-139.
128. Kim, G. J., Lee, D. E. & Kim, H.-S. (2001). High-Level Expression and One-Step Purification of Cyclic Amidohydrolase Family Enzymes. *Protein Expres. Purif.* **23**, 128-133.
129. Vessoni Penna, T. C. & Ishii, M. (2002). Selective permeation and organic extraction of recombinant green fluorescent protein (gfpuv) from Escherichia coli. *BMC Biotechnol* **2**, 7.
130. Tsumoto, K. (2003). Solubilization of active green fluorescent protein from insoluble particles by guanidine and arginine. *Biochem. Bioph. Res. Co.* **312**, 1383-1386.
131. Iafolla, M. A. J. (2008). Dark proteins: Effect of inclusion body formation on quantification of protein expression. *Proteins* **72**, 1233-1242.
132. Schumann, W. & Ferreira, L. C. S. (2004). Production of recombinant proteins in Escherichia coli. *Genet. Mol. Biol.* **27**, 442-453.
133. Stryer, L. (1995). *Biochemistry*. 4:th edit, W. H. Freeman and Company, New York.
134. Bolanos-Garcia, V. M. & Davies, O. R. (2006). Structural analysis and classification of native proteins from E-coli commonly co-purified by immobilised metal affinity chromatography. *BBA-Gen. Subjects* **1760**, 1304-1313.
135. Chant, A., Kraemer-Pecore, C. M., Watkin, R. & Kneale, G. G. (2005). Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein Expres. Purif.* **39**, 152-9.
136. Cadel, S., Gouzy-Darmon, C., Petres, S., Piesse, C., Pham, V. L., Beinfeld, M. C., Cohen, P. & Foulon, T. (2004). Expression and purification of rat recombinant aminopeptidase B secreted from baculovirus-infected insect cells. *Protein Expres. Purif.* **36**, 19-30.
137. Fonda, I., Kenig, M., Gaberc-Porekar, V., Pristovaek, P. & Menart, V. (2002). Attachment of histidine tags to recombinant tumor necrosis factor-alpha drastically changes its properties. *Scientific World Journal* **2**, 1312-25.
138. Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P. & Kobe, B. (2003). Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **12**, 1313-22.
139. Tsien, R. Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509-44.
140. Snapp, E. L. (2009). Fluorescent proteins: a cell biologist's user guide. *Trends Cell Biol.* **19**, 649-655.
141. Ward, T. H. & Lippincott-Schwartz, J. (2006). The uses of green fluorescent protein in mammalian cells. *Methods Biochem. Anal.* **47**, 305-37.

142. Wang, S. & Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis. *Nature* **369**, 400-03.
143. Lalonde, S., Ehrhardt, D. W. & Frommer, W. B. (2005). Shining light on signaling and metabolic networks by genetically encoded biosensors. *Curr. Opin. Plant Biol.* **8**, 574-81.
144. Magliery, T. J. & Regan, L. (2006). Reassembled GFP: detecting protein-protein interactions and protein expression patterns. *Methods Biochem. Anal.* **47**, 391-405.
145. Shimomura, O., Johnson, F. H. & Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J. Cell. Comp. Physiol.* **59**, 223-39.
146. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. (1992). Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111**, 229-233.
147. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-5.
148. Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y. & Remington, S. J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **v273**, p1392(4).
149. Yang, F., Moss, L. G. & Phillips, G. N. (1996). The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246-1251.
150. Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormo, M. & Remington, S. J. (1997). Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proc. Natl. Acad. Sci. U S A* **94**, 2306-11.
151. Enoki, S., Saeki, K., Maki, K. & Kuwajima, K. (2004). Acid denaturation and refolding of green fluorescent protein. *Biochemistry* **43**, 14238-48.
152. Heim, R., Cubitt, A. B. & Tsien, R. Y. (1995). Improved green fluorescence. *Nature* **373**, 663-4.
153. Miller, W. G. & Lindow, S. E. (1997). An improved GFP cloning cassette designed for prokaryotic transcriptional fusions. *Gene* **191**, 149-53.
154. Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-8.
155. Crameri, A., Whitehorn, E. A., Tate, E. & Stemmer, W. P. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* **14**, 315-319.
156. Fukuda, H., Arai, M. & Kuwajima, K. (2000). Folding of Green Fluorescent Protein and the Cycle3 Mutant. *Biochemistry* **39**, 12025-12032.

157. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E. & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567-72.

158. Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M., Lukyanov, S., Zaraisky, A. G., Davidson, M. W. & Chudakov, D. M. (2009). Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J.* **418**, 567-74.

159. Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K. & Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87-90.

160. Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A. & Tsien, R. Y. (2001). Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* **276**, 29188-94.

161. Subach, O. M., Gundorov, I. S., Yoshimura, M., Subach, F. V., Zhang, J., Gruenwald, D., Souslova, E. A., Chudakov, D. M. & Verkhusha, V. V. (2008). Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.* **15**, 1116-24.

162. Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. (2004). An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol.* **22**, 445-9.

163. Pédelacq, J. D. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology* **24**, 79-88.

164. William W. Ward, H. J. P. A. F. R. C. W. C. S. C. R. (1982). Spectral Perturbations of the Aequorea Green-Fluorescent Protein. *Photochemistry and Photobiology* **35**, 803-808.

165. Verkhusha, V. V., Kuznetsova, I. M., Stepanenko, O. V., Zaraisky, A. G., Shavlovsky, M. M., Turoverov, K. K. & Uversky, V. N. (2003). High stability of Discosoma DsRed as compared to Aequorea EGFP. *Biochemistry* **42**, 7879-84.

166. Stepanenko, O. V., Verkhusha, V. V., Kazakov, V. I., Shavlovsky, M. M., Kuznetsova, I. M., Uversky, V. N. & Turoverov, K. K. (2004). Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, "dimer2", and DsRed1. *Biochemistry* **43**, 14913-23.

167. Dietz, H. & Rief, M. (2004). Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc. Natl. Acad. Sci. U S A* **101**, 16192-16197.

168. Enoki, S., Maki, K., Inobe, T., Takahashi, K., Kamagata, K., Oroguchi, T., Nakatani, H., Tomoyori, K. & Kuwajima, K. (2006). The Equilibrium Unfolding Intermediate Observed at pH 4 and its Relationship with the Kinetic Folding Intermediates in Green Fluorescent Protein. *J. Mol. Biol.* **361**, 969-982.

169. Huang, J.-r., Craggs, T. D., Christodoulou, J. & Jackson, S. E. (2007). Stable Intermediate States and High Energy Barriers in the Unfolding of GFP. *J. Mol. Biol.* **370**, 356-371.

170. Xie, J. B. & Zhou, J. M. (2008). Trigger Factor Assisted Folding of Green Fluorescent Protein. *Biochemistry* **47**, 348-357.

171. Yakhnin, A. V., Vinokurov, L. M., Surin, A. K. & Alakhov, Y. B. (1998). Green Fluorescent Protein Purification by Organic Extraction. *Protein Expres. Purif.* **14**, 382-386.

172. Samarkina, O. N., Popova, A. G., Gvozdik, E. Y., Chkalina, A. V., Zvyagin, I. V., Rylova, Y. V., Rudenko, N. V., Lusta, K. A., Kelmanson, I. V., Gorokhovatsky, A. Y. & Vinokurov, L. M. (2009). Universal and rapid method for purification of GFP-like proteins by the ethanol extraction. *Protein Expres. Purif.* **65**, 108-113.

173. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-23.

174. Whitmore, L. & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392-400.

175. Stepanenko, O. V., Verkhusha, V. V., Shavlovsky, M. M., Kuznetsova, I. M., Uversky, V. N. & Turoverov, K. K. (2008). Understanding the role of Arg96 in structure and stability of green fluorescent protein. *Proteins* **73**, 539-51.

176. Huang, J. R., Hsu, S. T., Christodoulou, J. & Jackson, S. E. (2008). The extremely slow-exchanging core and acid-denatured state of green fluorescent protein. *HFSP Journal* **2**, 378-87.

177. Sinnecker, D., Voigt, P., Hellwig, N. & Schaefer, M. (2005). Reversible photobleaching of enhanced green fluorescent proteins. *Biochemistry* **44**, 7085-94.

178. Goldstein, D. A. (1990). Serum Calcium. In *Clinical Methods, The History, Physical, and Laboratory Examinations* 3:d Edition edit. (Walker, K. H., Hall, D. W. & Hurst, W. D., eds.), Boston.

179. Bousquet, J. C., Saini, S., Stark, D. D., Hahn, P. F., Nigam, M., Wittenberg, J. & Ferrucci, J. T., Jr. (1988). Gd-DOTA: characterization of a new paramagnetic complex. *Radiology* **166**, 693-8.

180. Huskens, J., Torres, D. A., Kovacs, Z., Andre, J. P., Geraldes, C. F. & Sherry, A. D. (1997). Alkaline Earth Metal and Lanthanide(III) Complexes of Ligands Based upon 1,4,7,10-Tetraazacyclododecane-1,7-bis(acetic acid). *Inorg. Chem.* **36**, 1495-1503.

181. Buster, D. C., Castro, M. M., Geraldes, C. F., Malloy, C. R., Sherry, A. D. & Siemers, T. C. (1990). Tm(DOTP)5-: a 23Na+ shift agent for perfused rat hearts. *Magn Reson Med* **15**, 25-32.

182. Malloy, C. R., Buster, D. C., Castro, M. M., Geraldes, C. F., Jeffrey, F. M. & Sherry, A. D. (1990). Influence of global ischemia on intracellular sodium in the perfused rat heart. *Magn Reson Med* **15**, 33-44.

183. Meyer, M., Dahaoui-Gindrey, V., Lecomte, C. & Guilard, R. (1998). Conformations and coordination schemes of carboxylate and carbamoyl derivatives of the tetraazamacrocycles cyclen and cyclam, and the relation to their protonation states. *Coordin. Chem. Rev.* **178-180**, 1313-1405.

184. Kreher, U., Hearn, M. T. W. & Spiccia, L. (2009). X-Ray Crystal Structure, Acid-Base Properties and Complexation Characteristics of a Methylenephosphonate Derivative of 1,4,7,10-Tetraazacyclododecane. *Aust. J. Chem.* **62**, 1583-1592.

185. Kumar, K. (1995). Synthesis, Stability, and Crystal Structure Studies of Some Ca2+, Cu2+, and Zn2+ Complexes of Macrocyclic Polyamino Carboxylates. *Inorg. Chem.* **34**, 6472.

186. Cohen, H., Alferiev, I. S., Monkkonen, J., Seibel, M. J., Pinto, T., Ezra, A., Solomon, V., Stepensky, D., Sagi, H., Ornoy, A., Patlas, N., Hagele, G., Hoffman, A., Breuer, E. & Golomb, G. (1999). Synthesis and preclinical pharmacology of 2-(2-aminopyrimidinio) ethylidene-1,1-bisphosphonic acid betaine (ISA-13-1) - A novel bisphosphonate. *Pharmaceutical Research* **16**, 1399-1406.

187. Delgado, R. & da Silva, J. J. (1982). Metal complexes of cyclic tetra-azatetra-acetic acids. *Talanta* **29**, 815-22.

188. Delgado, R., Siegfried, L., C. & Kaden, T., A. . (1990). Metal Complexes with Macrocyclic Ligands. Part XXXI. Protonation studies and complexation properties of tetraazamacrocyclic methylenephosphonates with earth-alkali ions. *Helv. Chim. Acta* **73**, 140-148.

189. Burai, L., Ren, J., Kovacs, Z., Brucher, E. & Sherry, A. D. (1998). Synthesis, Potentiometry, and NMR Studies of Two New 1,7-Disubstituted Tetraazacyclododecanes and Their Complexes Formed with Lanthanide, Alkaline Earth Metal, Mn(2+), and Zn(2+) Ions. *Inorg. Chem.* **37**, 69-75.

190. Kalman, F. K., Baranyai, Z., Toth, I., Banyai, I., Kiraly, R., Brucher, E., Aime, S., Sun, X., Sherry, A. D. & Kovacs, Z. (2008). Synthesis, potentiometric, kinetic, and NMR Studies of 1,4,7,10-tetraazacyclododecane-1,7-bis(acetic acid)-4,10-bis(methylenephosphonic acid) (DO2A2P) and its complexes with Ca(II), Cu(II), Zn(II) and lanthanide(III) ions. *Inorg. Chem.* **47**, 3851-62.

191. Kreher, U. (2006). Hard Metal ion chelating complexes for the application in immobilised metal affinity chromatography (IMAC) of biomolecules, Monash University.

192. Kumar, K., Chang, C. A., Francesconi, L. C., Dischino, D. D., Malley, M. F., Gougoutas, J. Z. & Tweedle, M. F. (1994). Synthesis, Stability, and Structure of Gadolinium(Iii) and Yttrium(Iii) Macrocyclic Poly(Amino Carboxylates). *Inorg. Chem.* **33**, 3567-3575.

193. Desreux, J. F., Merciny, E. & Loncin, M. F. (1981). Nuclear magnetic resonance and potentiometric studies of the protonation scheme of two tetraaza tetraacetic macrocycles. *Inorg. Chem.* **20**, 987-991.

194. Smith, R. V. & Nessen, M. A. (1971). Atomic absorption analysis of sodium, potassium, and calcium in Ringer's solution. *J. Pharm. Sci.* **60**, 907-8.

195. Schinkel, H. (1984). Determination of Calcium, Magnesium, Strontium, Potassium, Sodium, Lithium, Iron, Manganese, Chromium, Nickel, Copper, Cobalt, Zinc and Cadmium by Flame Aas - a Universal Method for the Analysis of Waters, Coals, Ashes, Ores, Rocks, Building-Materials, Metals and Similar Sample. *Fresenius Zeitschrift Fur Analytische Chemie* **317**, 10-26.

196. Weeks, J. M., Taylor, M. R. & Wainwright, K. P. (1997). Formation constants for complexes of 1,4,7,10-tetraazacyclododecane-1,7-diacetic acid and the crystal structure of its nickel(II) complex. *J. Chem. Soc., Dalton Trans.*, 317-322.

197. Shannon, R. D. (1976). Revised Effective Ionic-Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crys. A* **32**, 751-767.

198. Sahni, S. K. & Reedijk, J. (1984). Coordination chemistry of chelating resins and ion exchangers. *Coord. Chem. Rev.* **59**, 1-139.

199. Skoog, D. A., West, D. M. & Holler, F. J. (1996). *Fundamentals of analytical chemistry* (edit., t., Ed.), Saunders College Publishing, Orlando.

200. Alexandratos, S. D. (2009). Ion-Exchange Resins: A Retrospective from Industrial and Engineering Chemistry Research. *Ind. Eng. Chem. Res.* **48**, 388-398.

201. Schwaller, B. (2005). Calcium-binding proteins. In *Encyclopedia of Life Sciences*, pp. 1-8. John Wiley & Sons, Ltd.

202. Rosengarth, A. & Luecke, H. (2004). Annexins: calcium binding proteins with unusual binding sites. In *Handbook of Metalloproteins* (Messerschmidt, A., Bode, W. & Cygler, M., eds.), Vol. 3, pp. 1-15. John Wiley & Sons, Ltd, Chichester.

203. Moore, P. B. (1986). 67 kDa calcimedin, a new Ca2+-binding protein. *Biochem. J.* **238**, 49-54.

204. Owens, R. J. & Crumpton, M. J. (1984). Isolation and characterization of a novel 68,000-Mr Ca2+-binding protein of lymphocyte plasma membrane. *Biochem. J.* **219**, 309-16.

205. Mani, R. S. & Kay, C. M. (1989). Purification and spectral studies on the Ca2+-binding properties of 67 kDa calcimedin. *Biochem. J.* **259**, 799-804.

206. Nalefski, E. A. & Falke, J. J. (1996). The C2 domain calcium-binding motif: structural and functional diversity. *Protein Sci* **5**, 2375-90.

207. Schwalbe, R. A., Dahlback, B., Coe, J. E. & Nelsestuen, G. L. (1992). Pentraxin family of proteins interact specifically with phosphorylcholine and/or phosphorylethanolamine. *Biochemistry* **31**, 4907-15.

208. Ohkubo, Y. Z. & Tajkhorshid, E. (2008). Distinct structural and adhesive roles of Ca2+ in membrane binding of blood coagulation factors. *Structure* **16**, 72-81.

209. Stenflo, J., Stenberg, Y. & Muranyi, A. (2000). Calcium-binding EGF-like modules in coagulation proteinases: function of the calcium ion in module interactions. *BBA-Protein Struct. M.* **1477**, 51-63.

210. Lewit-Bentley, A. & Rety, S. (2000). EF-hand calcium-binding proteins. *Curr. Opin. Struc. Biol.* **10**, 637-643.

211. Henikoff, S., Greene, E. A., Pietrokovski, S., Bork, P., Attwood, T. K. & Hood, L. (1997). Gene families: the taxonomy of protein paralogs and chimeras. *Science* **278**, 609-14.

212. Allouche, D., Parello, J. & Sanejouand, Y. H. (1999). Ca2+/Mg2+ exchange in parvalbumin and other EF-hand proteins. A theoretical study. *J. Mol. Biol.* **285**, 857-73.

213. Zhou, Y., Yang, W., Kirberger, M., Lee, H. W., Ayalasomayajula, G. & Yang, J. J. (2006). Prediction of EF-hand calcium-binding proteins and analysis of bacterial EF-hand proteins. *Proteins* **65**, 643-55.

214. Kawasaki, H. & Kretsinger, R. H. (1995). Calcium-binding proteins 1: EF-hands. *Protein Profile* **2**, 297-490.

215. Linse, S. & Chazin, W. J. (1995). Quantitative measurements of the cooperativity in an EF-hand protein with sequential calcium binding. *Protein Sci.* **4**, 1038-44.

216. Chattopadhyaya, R., Meador, W. E., Means, A. R. & Quiocho, F. A. (1992). Calmodulin structure refined at 1.7 A resolution. *J. Mol. Biol.* **228**, 1177-92.

217. Rigden, D. J., Jedrzejas, M. J. & Galperin, M. Y. (2003). An extracellular calcium-binding domain in bacteria with a distant relationship to EF-hands. *FEMS Microbiol Lett* **221**, 103-10.

218. Momma, K., Mikami, B., Mishima, Y., Hashimoto, W. & Murata, K. (2002). Crystal structure of AlgQ2, a macromolecule (alginate)-binding protein of Sphingomonas sp. A1 at 2.0A resolution. *J. Mol. Biol.* **316**, 1051-9.

219. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. (1987). A novel calcium binding site in the galactose-binding protein of bacterial transport and chemotaxis. *Nature* **327**, 635-8.

220. Baba, M. L., Goodman, M., Bergercohn, J., Demaille, J. G. & Matsuda, G. (1984). The Early Adaptive Evolution of Calmodulin. *Molecular Biology and Evolution* **1**, 442-455.

221. Means, A. R., Tash, J. S. & Chafouleas, J. G. (1982). Physiological Implications of the Presence, Distribution, and Regulation of Calmodulin in Eukaryotic Cells. *Physiological Reviews* **62**, 1-39.

222. Rhyner, J. A., Ottiger, M., Wicki, R., Greenwood, T. M. & Strehler, E. E. (1994). Structure of the human CALM1 calmodulin gene and identification of two CALM1-related pseudogenes CALM1P1 and CALM1P2. *Eur. J. Biochem.* **225**, 71-82.

223. Ni, B., Rush, S., Gurd, J. W. & Brown, I. R. (1992). Molecular cloning of calmodulin mRNA species which are preferentially expressed in neurons in the rat brain. *Brain Res. Mol. Brain Res.* **13**, 7-17.

224. Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smailus, D. E., Schnerch, A., Schein, J. E., Jones, S. J. & Marra, M. A. (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. U S A* **99**, 16899-903.

225. Chin, D. & Means, A. R. (2000). Calmodulin: a prototypical calcium sensor. *Trends Cell Biol.* **10**, 322-328.

226. Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R. & Cook, W. J. (1985). 3-Dimensional Structure of Calmodulin. *Nature* **315**, 37-40.

227. Linse, S., Helmersson, A. & Forsen, S. (1991). Calcium binding to calmodulin and its globular domains. *J. Biol. Chem.* **266**, 8050-4.

228. Ye, Y., Lee, H. W., Yang, W., Shealy, S. & Yang, J. J. (2005). Probing site-specific calmodulin calcium and lanthanide affinity by grafting. *J. Am. Chem. Soc.* **127**, 3743-50.

229. Siedlecka, M., Goch, G., Ejchart, A., Sticht, H. & Bierzynski, A. (1999). alpha -Helix nucleation by a calcium-binding peptide loop. *Proc. Natl. Acad. Sci. U S A* **96**, 903-908.

230. Wojcik, J., Goral, J., Pawlowski, K. & Bierzynski, A. (1997). Isolated Calcium-Binding Loops of EF-Hand Proteins Can Dimerize To Form a Native-Like Structure. *Biochemistry* **36**, 680-687.

231. Nara, M., Morii, H., Yumoto, F., Kagi, H. & Tanokura, M. (2006). Fourier transform infrared spectroscopic study on the Ca2+ -bound coordination structures of synthetic peptide analogues of the calcium-binding site III of troponin C. *Biopolymers* **82**, 339-43.

232. Ye, Y., Shealy, S., Lee, H. W., Torshin, I., Harrison, R. & Yang, J. J. (2003). A grafting approach to obtain site-specific metal-binding properties of EF-hand proteins. *Protein Eng.* **16**, 429-34.

233. Ye, Y., Lee, H. W., Yang, W., Shealy, S. J., Wilkins, A. L., Liu, Z. R., Torshin, I., Harrison, R., Wohlhueter, R. & Yang, J. J. (2001). Metal binding affinity and structural properties of an isolated EF-loop in a scaffold protein. *Protein Eng.* **14**, 1001-13.

234. Le Clainche, L. L. (2003). Engineering new metal specificity in EF-hand peptides. *J. Biol. Inorg. Chem.* **8**, 334.

235. Clainche, L. L., Figuet, M., Montjardet-Bas, V., Blanchard, S. & Vita, C. (2006). Modulating the affinity and the selectivity of engineered calmodulin EF-Hand peptides for lanthanides. *Biotechnol Bioeng* **95**, 29-36.

236. Nitz, M., Franz, K. J., Maglathlin, R. L. & Imperiali, B. (2003). A powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides. *ChemBioChem* **4**, 272-6.

237. Franz, K. J., Nitz, M. & Imperiali, B. (2003). Lanthanide-binding tags as versatile protein coexpression probes. *ChemBioChem* **4**, 265-71.

238. Richardson, F. S. (1982). Terbium(Iii) and Europium(Iii) Ions as Luminescent Probes and Stains for Biomolecular Systems. *Chem. Rev.* **82**, 541-552.

239. Ye, Y., Lee, H.-W., Yang, W. & Yang, J. J. (2005). Calcium and lanthanide affinity of the EF-loops from the C-terminal domain of calmodulin. *J. Inorg. Biochem.* **99**, 1376-1383.

240. Veenstra, T. D., Johnson, K. L., Tomlinson, A. J., Naylor, S. & Kumar, R. (1997). Determination of calcium-binding sites in rat brain calbindin D28K by electrospray ionization mass spectrometry. *Biochemistry* **36**, 3535-42.

241. Hu, P., Ye, Q. Z. & Loo, J. A. (1994). Calcium stoichiometry determination for calcium binding proteins by electrospray ionization mass spectrometry. *Anal. Chem.* **66**, 4190-4.

242. Lafitte, D., Capony, J. P., Grassy, G., Haiech, J. & Calas, B. (1995). Analysis of the ion binding sites of calmodulin by electrospray ionization mass spectrometry. *Biochemistry* **34**, 13825-32.

243. Porumb, T. (1994). Determination of Calcium-Binding Constants by Flow Dialysis. *Anal. Biochem.* **220**, 227-237.

244. Brittain, H. G. (1976). Terbium(III) emission as a probe of calcium(II) binding sites in proteins. *J. Am. Chem. Soc.* **98**, 8255-8260.

245. Borin, G., Ruzza, P., Rossi, M., Calderan, A., Marchiori, F. & Peggion, E. (1989). Conformation and ion binding properties of peptides related to calcium binding domain III of bovine brain calmodulin. *Biopolymers* **28**, 353-69.

246. Sculimbrene, B. R. & Imperiali, B. (2006). Lanthanide-binding tags as luminescent probes for studying protein interactions. *J. Am. Chem. Soc.* **128**, 7346-52.

247. MacManus, J. P., Hogue, C. W., Marsden, B. J., Sikorska, M. & Szabo, A. G. (1990). Terbium luminescence in synthetic peptide loops from calcium-binding proteins with different energy donors. *J. Biol. Chem.* **265**, 10358-66.