



universität
wien

MASTERARBEIT

Titel der Masterarbeit

“Studying the Siderocalin NGAL and Selective
Isotope Labelling of GB1 by NMR”

verfasst von

Mag.pharm. Katharina Weinhäupl, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2013

Studienkennzahl lt. Studienblatt: A 066 834

Studienrichtung lt. Studienblatt: Masterstudium Molekulare Biologie UG2002

Betreuer: Univ. Prof. Dr. Robert Konrat

0.1 Abstract

The 25 kDa protein NGAL is a siderocalin, a subgroup of the diverse lipocalin protein family, that chelates iron via siderophores. This ability is a prerequisite for the multitude of physiological functions of NGAL. These range from antibacterial defence and regulation of the iron metabolism to differentiation, association with inflammation and cancer progression. An important partner in its function is its cellular receptor 24p3R that is responsible for the endocytosis of the NGAL-receptor complex.⁵⁰

The main part of this work deals with characterizing the interaction of NGAL with its receptor 24p3R. Two side projects are concerned with the finding of new ligands for NGAL and the development of new selective labelling techniques for the amino acids valine, tyrosine and phenylalanine.

Binding of NGAL to the N-terminal domain (NTD) of 24p3R results in severe peak broadening of the affected residues in the NMR spectrum, thus requiring the application of CPMG experiments, a NMR method to study low populated or excited states. Using ^1H - ^{15}N HSQC we detected three distinct regions in the NTD, that are differently affected by binding to NGAL. This behaviour is also reflected in the respective relaxation dispersion curves. To obtain more information about the bound state we introduced MTSL and a lanthanide binding peptide as paramagnetic tags into NGAL to enable PRE, PCS and RDC measurements.

Additionally we identified four new fluorinated ligands for NGAL, that can be used as starting points for a FBDD approach. Considering the physiological and pathophysiological importance of NGAL, the search for new ligands poses an interesting research topic.

Moreover, we successfully employed three new selective labelling strategies on the model protein GB1. Selective sidechain labelling of valines was achieved using the precursors 2-ketoisocaproate and 2-ketoisovalerate. While *p*-hydroxy- α -ketophenylpyruvate and α -ketophenylpyruvate served as precursors for the selective labelling of tyrosine and phenylalanine backbone carbonyls respectively.

In summary, we investigated the interaction of NGAL and its cellular receptor 24p3R, an interesting system for antibacterial host defence and iron metabolism. Furthermore we were able to discover four new ligands for NGAL and applied three new selective labelling strategies for valines, phenylalanines and tyrosines.

0.2 Zusammenfassung

Das 25 kDa Protein NGAL gehört zu den Siderocalinen, einer Untergruppe der vielfältigen Lipocalin-Proteinfamilie, die Eisen über Siderophore bindet. Diese Fähigkeit ist eine Voraussetzung für die Vielzahl an physiologischen Funktionen von NGAL. Diese reichen von antibakterieller Abwehr und Regulierung des Eisenstoffwechsels bis zu Differenzierung, Assoziierung mit Entzündungsprozessen und Tumorprogression. Ein wichtiger Partner in seiner Funktion ist sein zellulärer Rezeptor 24p3R, der für die Endozytose des NGAL-Rezeptor Komplexes verantwortlich ist.⁵⁰

Der Hauptteil dieser Arbeit beschäftigt sich mit der Charakterisierung der Wechselwirkung von NGAL mit seinem Rezeptor 24p3R. Zwei Nebenprojekte befassen sich mit der Suche nach neuen Liganden für NGAL und der Entwicklung neuer selektiver Markierungs Techniken für die Aminosäuren Valin, Tyrosin und Phenylalanin.

Die Bindung von NGAL an die N-terminale Domäne (NTD) von 24p3R resultiert in starker Peakverbreiterung der betroffenen Aminosäurereste im NMR-Spektrum, daher war die Anwendung von CPMG Experimenten, einer NMR-Methode, um niedrig besetzte oder angeregte Zuständen zu untersuchen, erforderlich. Mittels ^1H - ^{15}N HSQC entdeckten wir drei verschiedene Regionen in der NTD, die unterschiedlich durch die Bindung an NGAL beeinflusst werden. Dieses Verhalten wird auch in den jeweiligen Relaxations-Dispersions Kurven widergespiegelt. Um mehr Informationen über den gebundenen Zustand zu erhalten, haben wir MTSL und LBP als paramagnetische Tags in NGAL eingeführt um PRE-, PCS- und RDC-Messungen zu ermöglichen.

Zusätzlich ermittelten wir vier neue fluorierte Liganden von NGAL, die als Ausgangspunkte für einen FBDD Ansatz verwendet werden könnten. Unter Berücksichtigung der physiologischen und pathophysiologischen Bedeutung von NGAL, stellt die Suche nach neuen Liganden ein interessantes Forschungsgebiet dar.

Darüber hinaus haben wir erfolgreich drei neue selektive Markierungs Strategien für das Modellprotein GB1 entwickelt. Selektive Markierung der Seitenkette von Valin wurde mit den Ausgangsstoffen 2-Ketoisocaproat und 2-Ketoisovalerat erreicht. Während p-Hydroxy- α -ketophenylpyruvat und α -ketophenylpyruvat als Ausgangsstoffe für die selektive Markierung der Carbonyle von Tyrosinen und Phenylalaninen des Proteinrückgrats dienten.

Zusammenfassend haben wir die Wechselwirkung von NGAL und seinem zellulären Rezeptor 24p3R, einem interessanten System der antibakteriellen Immunabwehr und des Eisenstoffwechsels, untersucht. Außerdem konnten wir vier neue Liganden für NGAL entdecken und drei neue selektive Markier-

ungs Strategien für Valine, Tyrosine und Phenylalanine anwenden.

Contents

0.1	Abstract	3
0.2	Zusammenfassung	5
1	Introduction	11
1.1	NGAL	12
1.1.1	The Lipocalin Family	12
1.1.2	Structure of NGAL	13
1.1.3	Physiological Function of NGAL	14
1.1.4	NGALs Iron Binding Cofactor Enterobactin	15
1.1.5	Pathophysiology of NGAL	18
1.2	²⁴ p3R	20
1.2.1	Structure and Function	20
1.2.2	Intrinsically Disordered Proteins	22
1.3	Paramagnetic Protein Tags	25
1.3.1	Lanthanide Binding Peptides	27
1.4	Fluorinated Compounds in Drug Design	27
1.4.1	Fluorine Libraries	28
1.4.2	¹⁹ F NMR	29
1.5	Selective Labelling of Amino Acids	29
1.5.1	GB1	29
1.5.2	Purpose of Selectively Labelled Amino Acids	30
1.5.3	Methods to Selectively Label Amino Acids	30
1.6	Biomolecular NMR	32
1.6.1	HSQC	32
1.6.2	CPMG	33
1.6.3	HNCO	34
2	Aim of the Work	35
2.1	Aim of the Work	37

3	Materials and Methods	39
3.1	Materials	41
3.1.1	Cells	41
3.1.2	Buffers and Media	41
3.1.3	Cloning	46
3.1.4	Equipment	47
3.1.5	Software	47
3.2	Methods	48
3.2.1	Protein Expression and Purification	48
3.2.2	ITC	51
3.2.3	NMR	51
3.2.4	Cloning	51
4	Results	55
4.1	NGAL and its Receptor 24p3R	57
4.1.1	HMQC Titration Series of NGAL to NTD	58
4.1.2	Observing Chemical Exchange in NGAL bound NTD	60
4.1.3	Investigating the Bound State with Paramagnetic Tags	63
4.2	^{19}F Ligand Binding	69
4.2.1	Analysing Ligand Binding via ^1H - ^{15}N HSQC Chemical Shifts	69
4.2.2	Determination of Binding Affinities via ITC	69
4.3	Selective Labelling of Val, Phe and Tyr in Proteins	72
4.3.1	Independent Labelling of Valine Residues in GB1	72
4.3.2	Selective Labelling of Aromatic Residues in GB1	74
5	Discussion	77
5.1	Discussion	79
5.1.1	Investigating the Interaction of NGAL and its Cellular Receptor 24p3R	79
5.1.2	Finding New Ligands for NGAL	81
5.1.3	Selective Labelling of Val, Tyr and Phe in the Model Protein GB1	82
5.2	Conclusion	83
6	Bibliography	85
7	Glossary	93
8	Acknowledgements	99
9	Curriculum Vitae	103

List of Figures

1.1	The lipocalin fold	14
1.2	NGAL with its cofactor enterobactin	16
1.3	Enterobactin	18
1.4	24p3R	21
1.5	IDPs and folded proteins	24
4.1	^1H - ^{15}N SOFAST HMQC of the NTD	57
4.2	Titration series of ^{14}N -NGAL to ^{15}N -NTD	58
4.3	K_d value of NTD and NGAL	59
4.4	CPMG of NTD bound to NGAL - region 3	61
4.5	CPMG of NTD bound to NGAL - region 2	62
4.6	Titration series of MTSL-NGAL to ^{15}N -labelled NTD	63
4.7	LBP-NGAL constructs	65
4.8	^1H - ^{15}N HSQCs of lanthanide loaded LBP-NGAL	66
4.9	Titration of lanthanide loaded NGALstop to NTD	67
4.10	ITC - NGAL and E2	70
4.11	^{19}F ligands for NGAL	71
4.12	Precursors for selective valine labelling	72
4.13	Selective valine labelling in GB1	73
4.14	Precursors for labelling of aromatic amino acids	74
4.15	Selective labelling of phenylalanine and tyrosine in GB1	75

Chapter 1

Introduction

1.1 NGAL

Neutrophil gelatinase-associated lipocalin (NGAL) is a 25 kDa protein that was first discovered in the granules of neutrophils. It was copurified with neutrophil gelatinase (also called MMP-9). MMP-9 is a proteinase, that is responsible for the degradation of membrane and extracellular matrix components.⁸ NGAL can be found in three different forms: as a monomer (25 kDa), as a homodimer (45 kDa) and as a heterodimer (135 kDa) associated with neutrophil gelatinase.²⁶ While the monomeric form is primarily synthesized in tubular epithelial cells, the dimeric form is produced by neutrophils.⁷ NGAL homologues can also be found in mouse and rat. They show very low sequence identity (mouse 62 %, rat 63 %), but high structural identity.⁸ Secretion of NGAL is stimulated by inflammatory stimuli. NGAL is expressed in bone marrow, lung, stomach, salivary glands, appendix, colon, prostate, kidney and uterus.²⁶ NGAL has various physiological and pathophysiological functions. The most prominent being its role in host defence through an iron dependent bacteriostatic effect. This effect can be observed in NGAL knockout mice that are more susceptible for infections with gram-negative bacteria.⁴⁸ It can also modulate the intracellular iron concentration, and therefore iron dependent gene regulation and iron delivery. Furthermore, it can be used as a biomarker for acute kidney injury, psoriasis or various forms of cancer.⁵⁰ NGAL is a member of the lipocalin family, a protein family with low sequence identity, but high structural identity. Lipocalins are specialized in, among other tasks, transport of hydrophobic ligands. The biological activity of the specific lipocalin is strongly determined by the ligand it binds.⁴⁸ NGAL belongs to a specialized group of lipocalins, that are called siderocalins and that bind iron via cofactors. These cofactors are termed siderophores.

NGAL has 178 residues and, since it is a predominantly extracellular protein, comprises a 20 amino acid signal peptide, that is cleaved upon secretion. Furthermore, NGAL can be N-glycosylated at Asn65. Deglycosylated NGAL has a molecular weight of only 21 kDa.²⁶

1.1.1 The Lipocalin Family

The lipocalin family is a group of proteins that comprise very diverse sequences, where pairwise comparison of the sequences is as low as 20 %, but have a rather homogeneous tertiary structure, termed the lipocalin fold.¹⁸ The lipocalin fold comprises eight anti-parallel β -strands, that form a continuously hydrogen bonded flattened β -barrel. Strands $\beta 2$ - $\beta 4$ and $\beta 6$ - $\beta 8$ each form a β -sheet, that then fold into the barrel structure (see figure 1.1). This β -barrel constitutes the binding site for ligands and is also termed calyx.

The open end of the barrel is also believed to bind to the cellular receptor. The diversity of this calyx allows binding of various ligands and explains thereby the multifaceted functions of the different lipocalins. Additionally the lipocalin fold includes an α -helix between strands H and I and a 3_{10} -helix following strand A. In total lipocalins have seven loop regions, whereas loop 1 can fold back on the barrel and partially close it. Cys76 and Cys176 are highly conserved in all lipocalins and can be used to form intramolecular disulfide bridges. Furthermore, NGAL comprises an additional cysteine residue, Cys87, that can be used for dimerization.²⁶ These features are also termed “structurally conserved regions” or SCR. Using this classification, lipocalins can be divided into two groups: the kernel lipocalins, that possess all three SCRs and the outlier lipocalins, that possess only one or two of them.⁸ Lipocalins are comparatively small (160-180 residues) extracellular proteins, that bind small hydrophobic molecules. Although they possess a variety of different functions, the most common is transport of ligands. For this purpose they bind to specific cell surface receptors. Prominent members of the lipocalin family are the retinol-binding protein, β -lactoglobulin, odorant-binding protein, NGAL and tear lipocalin. Lipocalins are part of an even bigger protein superfamily, the calycins, with members like FABPs, avidins and metalloproteinase inhibitors. Though lipocalins are predominately found in eukaryotes, they are also present in a few prokaryotes. This promoted the hypothesis, that lipocalins first evolved in prokaryotes and were then transferred to eukaryotes via endosymbiotic uptake of mitochondria.¹⁸

Siderocalins are a subset of the lipocalin family, that bind siderophores. Siderocalins are part of innate immunity, which primarily manifests itself in sequestration of siderophores, thereby limiting the source of iron for bacteria. Siderocalins are therefore specialized in binding siderophores of pathogens, that are most dangerous to the host.⁵⁰

1.1.2 Structure of NGAL

NGAL possesses a typical lipocalin fold with an additional second 3_{10} -helix at the N-terminus. The structure of NGAL has been solved by NMR and X-ray crystallography. The calyx in the crystal structure is larger and more open compared to the NMR structure. NGAL was crystallized in its dimeric form, where the two molecules are connected via a disulfide bridge from Cys87. The crystal structure also shows an additional binding groove between the α -helix and the sixth and seventh β -strand.²⁰

Binding of NGALs ligand enterobactin is mediated via cation- π interactions with Lys125, Lys134 and Arg81.⁵⁰ This was also shown for NGALs putative eukaryotic cofactors catechol and 4-methylcatechol by X-ray crys-

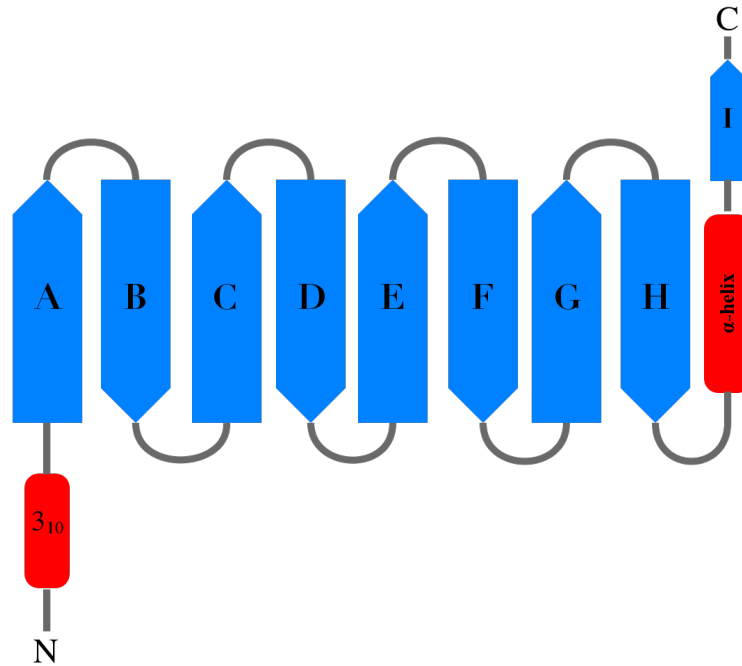


Figure 1.1: **The lipocalin fold**

The lipocalin fold comprised of eight antiparallel β -strands, one α -helix and one 3_{10} -helix (adapted from Flower et al.¹⁸)

tallography.²¹

1.1.3 Physiological Function of NGAL

In a healthy individual NGAL expression throughout the body is rather low.²¹ Only upon stimulation with certain impulses like bacterial infection, cancer, ischemia or oxidative stress is NGAL expression upregulated.⁷ Exposure to proinflammatory stimuli causes upregulation of NGAL. In human adults NGAL expression is upregulated during early stages of neutrophil maturation.⁸ Furthermore, NGAL can serve as a chemoattractant for neutrophils, attracting them to the site of inflammation. NGAL can also counteract oxidative stress, since expression of superoxide dismutase and heme oxygenase is upregulated at higher NGAL concentrations. Additionally NGAL also limits the concentration of hydroxyl free radicals by binding to Fe(III) chelated by catechols. Free catechol-Fe(III) undergoes a Fenton reaction reducing Fe(III) to Fe(II) and setting hydroxyl radicals free.⁸ In contrast to these findings *Lcn2* knockout mice are viable and reproduce normally.³⁰

One signalling pathway often associated with NGAL is the Nf κ B pathway.

NGAL regulates the $\text{Nf}\kappa\text{B}$ pathway by various cytokines, among them IL-1, IL-17, IL-22, IGF-1, TGF- α and TNF- α . NGAL expression itself can be stimulated by GM-CSF, dexamethason, vitamin A or insuline. Another, although controversial, regulation mechanism is regulation by hormones, namely estrogen.⁸

In eukaryotes NGAL is associated with processes like apoptosis, differentiation and proliferation. To fulfil these tasks there must also exist a eukaryotic cofactor for NGAL apart from enterobactin. Some possible candidates will be presented in the next section.⁴⁸

An *in vitro* study showed, that NGAL targets kidney progenitor cells in the developing kidney. Abolishing NGAL induction resulted in misassembly of tubules. NGAL also exhibits a tissue protective effect. Upon kidney injury NGAL concentration rises 25-100 fold in humans and even more in mice or rats. NGAL expression in the kidney is most prominent in the loop of Henle and the collecting ducts, although the primary site affected by renal ischemia is the proximal tubule. Interestingly post-ischemic NGAL concentration is highest in the proximal tubule. This can be explained by a transport mechanism that involves secretion of NGAL by cells of the loop of Henle and the collecting ducts into the circulation and subsequent glomerular filtration. NGAL is bound by 24p3R or megalin and gathered by cells of the proximal tubule. It is suspected, that local NGAL accounts for a bacteriostatic effect and systemic NGAL is used to quickly limit acute renal damage.⁴⁸ NGAL also induces enhanced proliferation of renal tubular cells as a physiological response to kidney damage.⁵²

1.1.4 NGALs Iron Binding Cofactor Enterobactin

Iron is an essential nutrient for humans as well as for bacteria. A 70 kg adult human carries approximately 5 g iron in his body. This corresponds to an approximate iron concentration in the body of about 10^{-3} M. Although to a lesser extent, bacteria need to maintain an intracellular iron concentration of 10^{-6} M to achieve optimal growth conditions. This can be challenging for the bacterium, since the human body tries to keep the free iron concentration low due to its toxicity. Generally free ferric iron has a low solubility at the physiological pH of 7.4. This would hypothetically limit the free ferric iron concentration to 10^{-18} M. In reality the free ferric iron concentration is even lower.⁴¹

Transport and storage of iron is accomplished either directly by specialized proteins or indirectly via a cofactor. Examples for direct binding are the iron transport protein transferrin and the iron storage protein ferritin. A prominent iron binding cofactor is for example heme in hemoglobin. In the

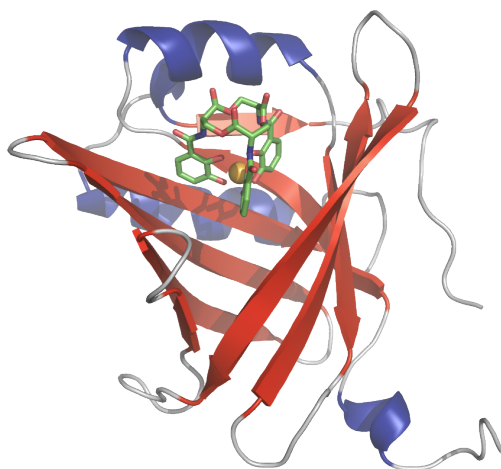


Figure 1.2: **NGAL with its cofactor enterobactin**

Typical lipocalin structure of NGAL associated with its bacterial cofactor enterobactin in the calyx [Coudeville et al.¹²].

last ten years many new proteins involved in iron metabolism have been discovered, among them hemochromatosis gene product, transferrin receptor-2, hephaestin, ferroportin-1, hepcidin or hemojuvelin. Great interest was also taken in NGAL and its receptor, that are important players in innate immunity and iron metabolism.⁴³

Iron acquisition is an important process for bacteria. This can be easily seen, since high iron concentrations in the host cause enhanced virulence of the pathogen. Examples are various strains like *Klebsiella*, *Listeria*, *Neisseria*, *Pasteurella*, *Shigella*, *Salmonella*, *Vibrio* and *Yersinia*. Furthermore, supplementation of iron, treating iron deficiencies, promotes bacterial colonization.⁴¹ For this reason humans and other higher organisms have developed a system to protect the bodies iron reserves from bacteria and keep iron concentrations in bacteria low to limit their growth.⁴⁸

To obtain enough iron, bacteria have to compete with transferrin. As a consequence bacteria have developed aggressive iron acquisition methods, using siderophores to sequester iron from their host. Siderophores are iron chelators and are secreted by bacteria upon iron deficiency. One example is enterobactin, that is secreted by gram-negative bacteria and uses catechols to chelate iron. In particular, enterobactin is made up of three catechol moieties that are linked via a triserine macrocycle (see figure 1.3). The iron binding ability is achieved by the catechol moieties. Catechol is biosynthesized in a side branch of the aromatic amino acid pathway, using chorismic acid as a

starting point. To obtain enterobactin, adenylated serine is added to form the triserine macrocycle.⁴¹

Iron loaded enterobactin is bound by FepA, a receptor in the outer membrane of bacteria. This receptor is only expressed upon iron deficiency and binds ferric enterobactin via a β -barrel motif. Inside of the barrel the binding is mediated by positively charged and aromatic amino acids. This mechanism is very close to the ligand binding mechanism described for NGAL. Enterobactin is transported into the periplasm, where it is bound to FepB and transported into the cytoplasm via an ATP-binding cassette transporter. For NGAL to compete with FepA its K_d must be below 50 nM, a binding constant determined by fluorescence spectroscopy. Since NGAL is binding to ferric enterobactin with a K_d of about 0.4 nM, this is the case.⁴¹

NGAL can bind to a wide array of bacterial cofactors, depending on the pathogen. Gram-negative bacteria primarily secrete enterobactin, while gram-positive bacteria secrete bacillibactin and mycobacteria carboxymycobactins.²¹ This capability of NGAL to bind to various cofactors is also termed “ligand plasticity”.⁸ To evade host defence bacteria have developed strategies to inhibit binding of the cofactor to NGAL. Mycobacteria can produce carboxymycobactins with differing length of the fatty acid chain, while salmochelin from *E.coli* strains show less aromatic amino acids to avoid cation- π interactions and petrobactin from *Bacillus anthracis* is too big for the ligand binding site and clashes with the wall of NGAL.⁵⁰

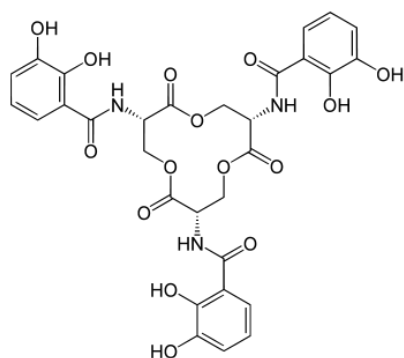
NGAL is able to effectively bind ferric enterobactin and transport it into host cells, thereby withdrawing iron from the pathogen. Thus NGAL plays an important role in the innate immune system. NGAL can bind to its cellular receptor in holo or apo form. NGAL loaded with ferric enterobactin binds 24p3R and gets endocytosed. In the endosome iron is released pH-dependently and the intracellular iron concentration increases. This can lead to the activation of iron responsive genes, assigning NGAL also a role in gene regulation. On the other hand apo-NGAL is able to scavenge intracellular iron and therefore depletes a cells iron reserves. This process was also associated with induction of apoptosis, but is rather controversial (see section 1.1.5 “Pathophysiology of NGAL”).⁴⁸

It is speculated that NGAL has its own mammalian cofactor for fulfilling these physiological tasks. The desired cofactor should have a low molecular weight, be present in the urine and be able to confer its iron binding activity to NGAL.⁴⁸ Possible candidates have been found in protein free filtrates of urine. These filtrates were able to chelate iron and, if isolated, bind to NGAL. Among these compounds were catechol, 3-methylcatechol, 4-methylcatechol and pyrogallol. All compounds lose their chelating activity upon O-methylation or O-sulfonation. These findings support the assump-

tion, that catechol must be a functional group in the putative mammalian cofactor. X-ray crystallography of NGAL bound to catechol confirmed the interaction of cationic groups in the calyx (Lys125 and Lys134) with these ligands. The observed interactions can be abrogated if the involved lysine residues are mutated to alanines. While the binding affinity of NGAL to free catechol is rather low ($K_d=0.2 \mu$), the affinity for Fe(III) bound catechol is high ($K_{d1} = 2.1 \text{ nM}$ and $K_{d2} = 0.4 \text{ nM}$). Three catechol molecules bind one Fe(III) ion in a hexadentate coordination. The binding process is performed in a stepwise manner. Release of iron in the cell is mediated by a decrease of pH in endosomes, leading to a protonation of catechol hydroxyl moieties. While iron release in enterobactin requires a pH below 4, catechol Fe(III) is already released upon pH 6. This mechanism seems to be favourable for iron trafficking. A possible source of catechols in mammals are metabolites of polyphenols like quinic or shikimic acid, aromatic amino acids like tyrosine or plant hydroxybenzenes like caffeic or chlorogenic acid, that are excreted into the urine.²¹

In short, the purpose of a mammalian cofactor for NGAL would be: mobilization of Fe(III), inhibition of the Fenton reaction and reduction of free Fe(III) in the body.²¹

A



B

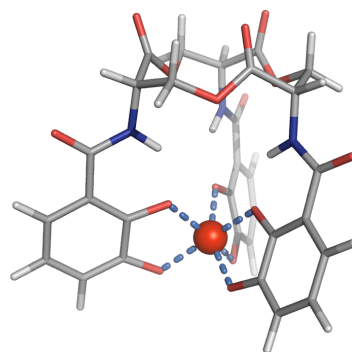


Figure 1.3: **Enterobactin**

A: Structure of enterobactin. **B:** Binding of Fe(III) by three catechol moieties [Coudevylle et al.¹²].

1.1.5 Pathophysiology of NGAL

NGAL is associated with several benign and malignant diseases.

First we will discuss benign diseases with an underlying inflammatory cause, like psoriasis, atopic dermatitis and eczema. Here only keratinocytes are affected and NGAL concentration is decreased after treatment. Another example for an inflammatory disease is periodontitis. In healthy individuals NGAL is only expressed in resident neutrophils. Patients suffering from periodontitis also express NGAL in the alveolar epithelium. Furthermore, NGAL expression is upregulated in myocarditis, ulcerative colitis, Crohn's disease and inflammatory bowel disease.⁸

An additional factor for NGAL expression is ischemia. This is mostly observed in patients with cerebrovascular or myocardial infarction. High NGAL concentrations are also associated with a higher rate of plaque instability in atherosclerosis patients. This might be due to the observed association of NGAL with MMP-9, which is an important factor in plaque instability.⁸

A third group of NGAL affected diseases are metabolic diseases. NGAL is upregulated in diabetes mellitus II and gestation diabetes. Moreover, NGAL expression is increased in response to insulin and in the fatty tissue of obese mice.⁸

The best known and best described diseases associated with NGAL are renal diseases. These range from acute to chronic kidney damage.⁸ The underlying cause is already described in section 1.1.3 "Physiology of NGAL". For this reason NGAL can also be used as biomarker for kidney damage. Normally the serum creatinine level and the urinary output are used as measures for kidney function. Disadvantages of these methods are delayed responsiveness and that they only contain information about the hemodynamic function of the kidney. Since acute kidney injuries are generally diagnosed late a new biomarker for this disease is needed. Recently NGAL was found to be a good candidate. Measuring NGAL concentration is advantageous, because it is non-invasive, since it can be directly quantified in the urine, and is already detectable at an early stage of the disease. A downside is that NGAL is not only upregulated in kidney damage, but also during sepsis or malignancies.⁵² The concentration of circulating NGAL could also be used to differentiate between bacterial and viral infections.²⁶

Upregulation of NGAL can also be found in context with cytotoxic agents and drugs. Examples are alcohol, methamphetamine, phencyclidine and other hepatotoxic agents. This effect is most probably caused by an increased expression upon liver damage, which is promoted by all of the aforementioned substances.⁸

NGAL is also involved in the development of malignancies. Key characteristic of tumor cells are: cellular independence of growth factors, insensitivity towards inhibitory growth signals, evasion of apoptosis, unlimited replication, promotion of angiogenesis, invasiveness and metastasis.¹ NGAL

expression is upregulated in, among others, malignancies of the skin, thyroid, breast, ovary, endometrium, colon, lung, liver, stomach and pancreas. Interestingly, in the endometrium a higher glycosylated 30 kDa isoform of NGAL is found. It is not surprising that NGAL is only upregulated in HER-2 driven breast cancer, since a synonym for HER-2 is “neu” and a different name for NGAL is neu-related lipocalin. Furthermore, an overexpression of 24p3R could be detected in leukemic cells and overexpression of NGAL causes myelosuppression.⁸ The role of NGAL in cancer is highly controversial since reports of both, pro- and anti-cancer effects, exist. This is maybe due to signalling pathways that can either have an activating or inhibiting effect, depending on the posttranslational modification. The most contrary results have been found for the apoptotic effect of NGAL. There are reports for a proapoptotic effect, that NGAL has no effect on apoptosis or that NGAL even causes a proliferative effect.³⁰ One reason for these controversial results could be, that the apoptotic effect was tested on a system where NGAL was binding its receptor 24p3R, carrying iron via the putative eukaryotic siderophore 2,5-dihydroxybenzoic acid. This siderophore is not a catechol, but a catechol anisomer and binds Fe(III) via the salicylate binding mode. This binding mode is thought to be too weak to cause the apoptotic effect and therefore renders the experiments invalid.⁵⁰ Another important step in cancer progression, that is often associated with NGAL, is metastasis. In several tumour cell lines, like human colon, human pancreas or mouse mammary, overexpression of NGAL leads to an inhibition of adhesiveness and invasiveness of the tumour. Although entirely contrary results have been found for other tumour cell lines.³⁰ Additionally it is not clear if the upregulation of NGAL in cancer is not only due to inflammatory processes, which are well known effects of NGAL and cancer progression.²⁷

1.2 24p3R

1.2.1 Structure and Function

The cellular receptor of NGAL, 24p3R, is a 60 kDa membrane protein with predicted twelve transmembrane helices and a 100 residue disordered N-terminal domain. It is comprised of 538 residues and is highly conserved in mice and humans. The murine form of 24p3R was first cloned by Devireddy et al. and elucidated main features of the NGAL iron cycle. It could be shown, that holo-NGAL delivers iron to the cell via 24p3R and vice versa, that apo-NGAL depletes a cell's iron reserves. Furthermore, that iron transport is mediated by recycling endosomes in the cell and that apo-

NGAL associates with its siderophore inside the cell. Additionally that iron depletion could trigger apoptosis via the proapoptotic protein Bim. Surprisingly Devireddy et al. were also able to show, that NGAL and 24p3R are linked to the oncoprotein BCR-ABL and therefore the development of chronic myeloid leukemia.¹⁷ BCR-ABL positive cells exhibit an increase in NGAL expression and a decrease in 24p3R, leading to cells that are unresponsive to possible proapoptotic signalling via NGAL. This effect could be abolished administering the BCR-ABL inhibitor imatinib.⁴³ Upregulation of NGAL expression is mediated by the JAK-Stat pathway, in particular Stat5, while downregulation of 24p3R is mediated by the Ras/MAPK-pathway.⁴⁹ 24p3R exists in different splicing variants. A long 60 kDa version and a short 30 kDa Version, that is lacking 154 N-terminal amino acids.¹⁷ Additional splicing variants were found, depending on the activation of the Wnt signalling pathway.⁶²

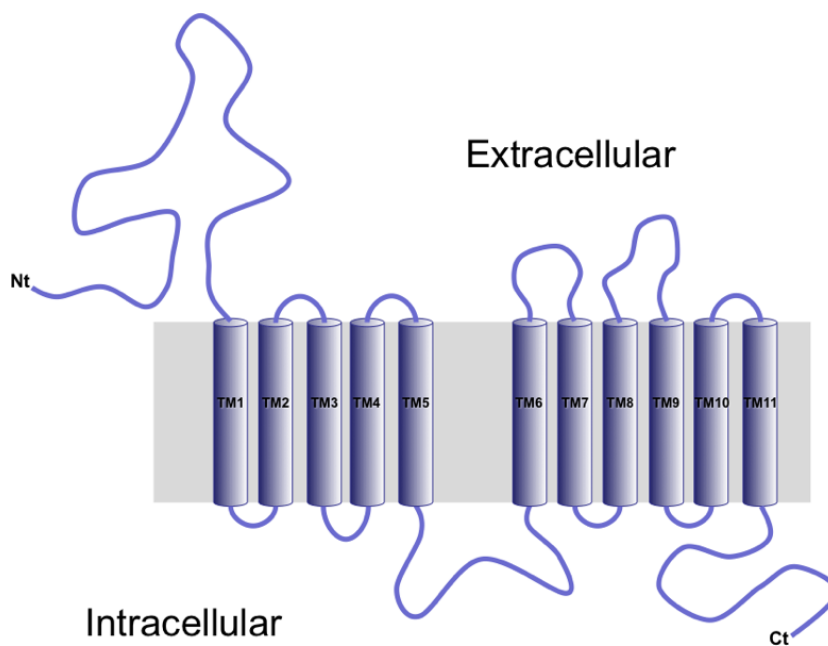


Figure 1.4: **24p3R**

The cellular receptor of NGAL 24p3R comprises twelve transmembrane helices and an approximately 100 residue N-terminal domain¹⁷

24p3R expression could be detected through antibody staining in renal epithelial cells, which is in good agreement with the physiological function of NGAL.¹⁷ NGAL expression in the kidney is found in the ascending loop of Henle, in the distal tubule and in the collecting ducts. The expression profile for 24p3R is rather similar, in particular in the distal tubule and the

collecting ducts. Renal mediated endocytosis, and hence bulk protein reabsorption, was widely believed to be exclusively taking place in the proximal tubule via the megalin-cubilin-amnionless receptor complex.²⁹ Now NGAL and 24p3R are also considered as important players in this process. Megalin is a member of the low-density lipoprotein receptor family and is an important multi-ligand endocytosis receptor. It is widely expressed throughout the body in epithelial cells and is able to bind various members of the lipocalin family, like retinol binding protein, α 1-microglobulin, odorant binding protein, lactoferrin and NGAL. Thus a major disadvantage of this receptor for NGAL is low selectivity.²² Consequently a different renal receptor for NGAL was needed and found in 24p3R, which shows much higher selectivity and affinity for NGAL ($K_d = 92$ pM) in comparison to megalin ($K_d = 60$ nM). One reason for protein reabsorption in the distal tubule and collecting ducts might be retaining valuable proteins, that have not bound to the megalin-cubilin-amnionless receptor complex due to overload or low affinity, for the body. Another advantage is defence against ascending urinary tract infections. 24p3R is also able to induce renal mediated endocytosis of metal binding proteins like metallothionein, transferrin and albumine.²⁹

1.2.2 Intrinsically Disordered Proteins

Intrinsically disordered proteins (IDPs) possess no well defined tertiary structure and exist in a highly dynamic conformational ensemble.⁴² Up until the 1990s a well defined structure was believed to be a prerequisite for protein function and a specific geometry in the binding site, as assumed in the “lock and key” model, crucial for protein interaction. In the last ten years intrinsically disordered proteins have gained a lot of attention and their importance in all kinds of physiological functions is evident. When IDPs were first encountered the lack of structure was believed to be an *in vitro* artefact and would not be found *in vivo*. Actually IDPs are not entirely unstructured and can have a certain preference for secondary structure. IDPs are present in all species, but tend to be more abundant in eukaryotes (15-45 %) than prokaryotes (10-35%). These statistics derive from a comparison of intrinsically disordered protein sequences with a length of at least 30 residues. The prevalence in eukaryotes is not surprising considering the fact, that IDPs are often associated with tasks like cell signalling, transcriptional, translational or cell-cycle regulation, which are more complex in eukaryotes. On the other hand a lack of disorder is more often found in enzymes. Common features of IDPs are adaptability, functional variety, weak but specific binding and high regulation through posttranslational modification.⁵⁵ This binding promiscuity makes them perfect hubs in large protein-protein networks. The existence

of more ubiquitinylation sites than in folded proteins makes IDPs prone to proteolytic degradation and therefore ensures a short lifetime, which can be advantageous in tasks like cell signalling.⁴²

Prominent IDPs are the proto-oncogene p53, the transcription factor CREB, Sic1 a regulator of Cdc4 involved in cell-cycle progression, the chaperone Hsp-33 and the neurodegenerative diseases associated proteins α -synuclein and τ -protein.^{35,55} Thus IDPs play an important role in health and disease. The high adaptability of IDPs can be either exploited to recognize a wide array of misfolded proteins, for example in the chaperone Hsp-33, or in viruses that mimic host proteins in virus entry, budding and replication. The high regulation of most IDPs through posttranslational modification and the quick degradation are crucial to fulfil sensitive tasks like cell-cycle control. Misregulation though, can lead to the development of cancer, neurodegenerative diseases, cardiovascular diseases or diabetes. Sequence analysis shows, that hydrophobic residues like C, W, Y, I, F, V and L are order promoting and found in ordered regions, while polar residues like E, P, Q, S, R, K and M are promoting disorder. There are different strategies to predict disorder in proteins, either by scoring charge and hydrophobicity, or by taking into account the occurrence of a certain amino acid in regular secondary structure elements like α -helices or β -sheets.⁴²

Since X-ray crystallography is an unsuitable technique to study IDPs, various other techniques must be employed. One of the most important is NMR, which has only been made possible through the development of multidimensional NMR experiments, higher field magnets and selective labelling techniques. Since IDPs generally have a poor dispersion of chemical shifts in the proton dimension (see figure 1.5 for a comparison of a HSQC spectrum of a folded protein and an IDP) and therefore often suffer from severe signal overlap. Information on IDPs can be gained by exploiting scalar coupling, secondary chemical shifts, RDCs, PREs, or relaxation measurements.⁵¹ Secondary chemical shifts are calculated via the deviation of chemical shifts from experimental random coil values. Chemical shifts are sensitive to secondary structure and hence are a good measure for secondary structure propensity. Information on transient long distance interaction is contained in RDC or PRE effects. Moreover, relaxation measurements, like T_1 and T_2 , are used to investigate backbone and sidechain dynamics⁴² and CPMG measurements to learn more about transition states or exchange processes in the micro- to millisecond timescale.⁹ Other techniques suitable for studying IDPs are SAXs, to determine the overall size and shape of the molecule or single molecule FRET to investigate interactions.⁴²

Binding Mode of IDPs

There are two major modes for an IDP to interact with its binding partner. First it can undergo folding upon binding. Here either the unfolded IDP can bind to its partner resulting in subsequent folding, this is also termed induced fit, or a low populated folded state can selectively bind to the binding partner, which is also called conformational selection.²⁵

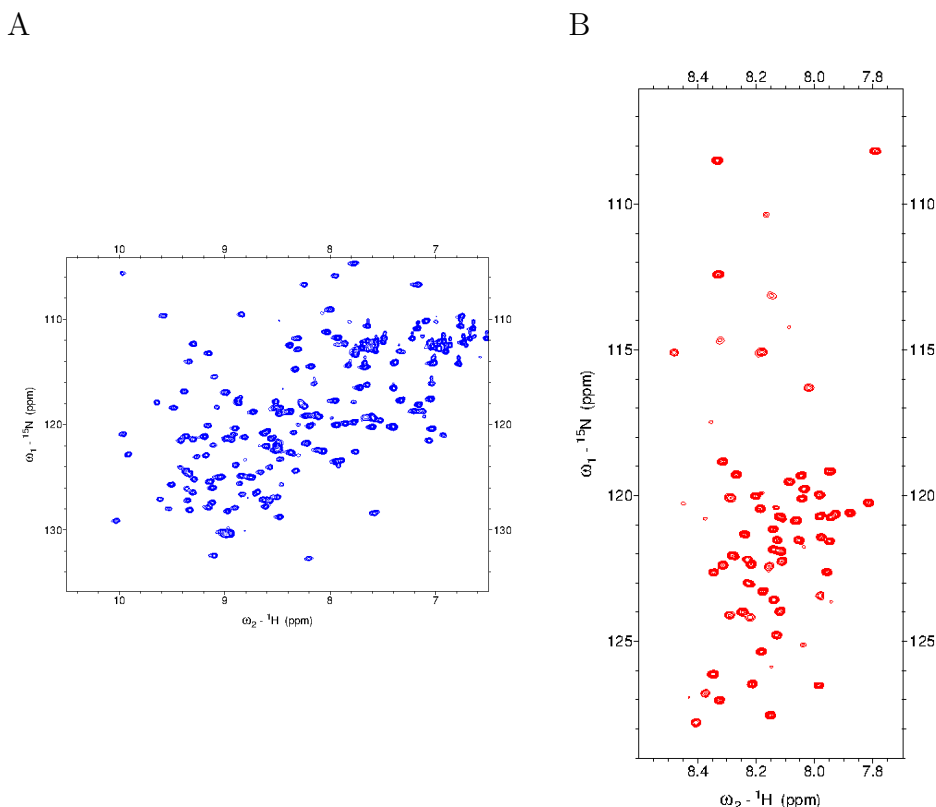


Figure 1.5: **IDPs and folded proteins**

A: ^1H - ^{15}N HSQC of NGAL. **B:** ^1H - ^{15}N SOFAST HMQC of the NTD. IDPs (NTD) exhibit a much smaller resonance dispersion in the proton dimension compared to folded proteins (NGAL).

The second possibility for an IDP to bind is maintaining a certain disorder in the bound state, this mode of interaction is summarized as “fuzzy” binding. If there is static disorder, the protein can adopt multiple stable conformations. In dynamic disorder there is a constant fluctuation between states. Possible scenarios are either two folded regions that are connected by a disordered linker, folded short recognition elements that are embedded in large disordered parts or, in the extreme case, a protein that stays fully disor-

dered in the bound state. Reasons for maintaining disorder in the bound state could be the higher conformational freedom, thus showing higher adaptability and providing accessible sites for posttranslational modification.⁵⁶ Binding interfaces of IDPs are generally larger, form more intermolecular interactions and are more hydrophobic than in folded proteins. The hydrophobicity may counteract an unfavourable decrease in entropy upon binding and the intermolecular contacts provide a better fit with the binding partner, increasing the overall speed of the interaction. Additionally many of these binding interfaces are evolutionary conserved.³⁷

1.3 Paramagnetic Protein Tags

Paramagnetic tags are often used in NMR to obtain distance or orientation information. There are two major classes of paramagnetic tags that can be applied. The first class are nitroxide spin labels, like MTSL or TEMPO, that have unpaired electrons. Nitroxide spin labels are best attached as iodo- or bromoacetamide derivatives to improve stability. The second class are metal chelators. These tags can be attached to the protein genetically or via functional groups like Cys residues or carboxyl groups.¹¹ Paramagnetic effects were first exploited in metal binding proteins, this method was of course only applicable to a small group of proteins and soon the development of paramagnetic tags, that could be used on various proteins, began. At first whole metal-binding domains, like zinc-finger, EF-hands or calmodulin, were attached to proteins, resulting in increased molecular weight and hence reduced signal intensity, impaired function and increased mobility. These unwanted effects promoted the development of new improved paramagnetic tags and nowadays a wide range of metal chelators, used as paramagnetic tags, are available.³ Their properties vary greatly depending on the type and on the attachment mode of the probe. Therefore a suitable tag should be chosen depending on the application of the label.¹¹ Possible options are small organic metal chelators like EDTA, DTPA or DOTA, that are linked to the protein via Cys residues. A major disadvantage of these probes is extensive signal overlap due to the existence of diastereomers. The smallest metal chelator is DPA, which must be attached via a carboxyl and a thiol group at a certain distance. Absence of these moieties at certain positions requires genetic modification.³ There are also metal binding motifs, that specifically bind to one metal, like the ATCUN motif to Cu^{2+} or the HHP motif to Ni^{2+} . Both motifs must be attached N-terminally. The ATCUN motif sequence is $\text{NH}_2\text{-X1-X2-His}$ and the HHP sequence is $\text{NH}_2\text{-His-His-Pro}$. A widely used subclass of metal chelators are lanthanide binding peptides,

which will be discussed in the next section, or synthetic lanthanide binding tags like cysteinyl-phenyl-triaminohexaacetate.³⁹

There are three observables in NMR that can be measured using paramagnetic tags to obtain long-range structural information: PREs, PCSs and RDCs. While PCSs and RDCs can only be exploited in systems with an anisotropic electron g-factor, the PRE effect can be observed in any paramagnetic system. PRE effects have a distance dependence of about r^{-6} , PCSs in contrast a distance dependence of r^{-3} , making even longer distance informations available. PRE and PCS effects provide distance information, whereas RDC effects also provide orientational information.³⁸ To measure a PRE effect the difference in transverse relaxation rates, enhanced by the paramagnetic tag, of dia- and paramagnetic states is calculated. This can disclose information on dynamic processes between the paramagnetic center and the nucleus.¹¹ The PCS provides long-range distance and angular information caused by the presence of a paramagnetic center. In order to measure the effect spectra with and without a paramagnetic probe must be obtained.²⁸ A paramagnetic center with an anisotropic g-tensor causes partial alignment in the magnetic field leading to RDC effects. A combination of this weak alignment and relaxation dispersion can give information on the bond vector orientation of the minor species. An important issue for PCS and RDC effects is low flexibility of the paramagnetic tag. Hence fixed metal ions in metal binding proteins cause a very large effect and flexible linkers a small effect. Attachment of the tag at two sites instead of one is also preferable.¹¹

Additionally, paramagnetic effects can be exploited to study encounter-complexes or transition states.¹¹ Formation of a complex can be seen as a at least two-step process, where the intermediate state is comprised of an ensemble of different orientations of the molecule.⁵⁷ This gives the molecule the advantage of sampling a major part of the surface of the binding partner. The formation of the encounter-complex is dominated by long-range electrostatic forces, while interactions in the active complex are mediated by specific short-range forces.⁴⁵

Information on long-range interactions can be obtained using paramagnetic tags. PRE effects are detectable up to 35 Å, in comparison NOE effects are only active over 6 Å. Paramagnetic tags are therefore best suitable to monitor these processes, as well as conformational rearrangements in protein domains and hence domain dynamics, detecting the presence of minor interconverting conformations in the domain.¹⁶ Additional information on low-populated states can be exploited using relaxation dispersion spectroscopy, which will be discussed later.¹¹

1.3.1 Lanthanide Binding Peptides

Lanthanide binding peptides or tags (LBP) have gained popularity due to their favourable properties. They can be applied in NMR to obtain information on structure, protein-protein or protein-ligand interactions or conformational dynamics, in X-ray crystallography to solve the phase problem or in luminescence measurements to investigate trafficking or localization.² LBPs were designed after naturally occurring Ca^{2+} binding proteins, and have a length of up to 23 amino acids. As a starting point EF-hand motifs in calmodulin or zinc-finger motifs were used and optimized to selectively bind lanthanides.³ Specific mutations in the four EF-hand calcium binding sites of calmodulin were sufficient to increase the binding affinity for lanthanides compared to calcium. The most effective were the mutations N60D in site II and N97D in site III.⁵ A combination of information from calcium binding motifs and combinatorial synthesis resulted in the design of these new lanthanide binding peptides. Binding affinities of nowadays used LBPs for lanthanides are in the nM K_d range.⁶¹ A common sequence in LBPs is GYIDDTNNDGWIEGDELY. The LBP can be either anchored via Cys residues, that can be inserted through site-directed mutagenesis or by genetically engineering the peptide sequence into the protein. Anchoring of the LBP is a crucial step in the design of the paramagnetic tag. Generally high rigidity is favourable for measurements and can be achieved by linking the peptide to two Cys residues or by cloning the peptide into a loop region of the protein of interest. This has been successfully applied in IL1 β , where insertion of the LBP into a loop resulted in an effective paramagnetic tag and no disturbance of the structure. However, addition of a paramagnetic tag can always influence the protein structure or function and must therefore be carefully planned.³ Additional information can be gained by positioning the LBP at different sites, using linkers with varying length, loading the LBP with different lanthanides or using different lanthanide binding tags. Another advantage of using LBPs is the easy generation of a reference molecule, as the LBP can be loaded either with a para- or diamagnetic lanthanide, with very similar ionic radii.⁵⁴

1.4 Fluorinated Compounds in Drug Design

Fluorination of compounds is a popular method in the lead optimization phase of drug design, due to its favourable impact on pharmacokinetics and pharmacodynamics. 20 - 25 % of all drugs on the market contain at least one fluorine atom, among them blockbusters like the antidepressant fluoxetine,

the reverse-transcriptase inhibitor efavirenz, the cox-2 inhibitors celecoxib and etoricoxib, the antidiabetic sitagliptin or the statin atorvastatin. In the pharmaceutical industry fluorinated drugs are so popular because fluorination has a positive effect on absorption, distribution, metabolism and excretion (ADME). ADME are the four criteria that describe the disposition of a pharmaceutical compound in the body and therefore its pharmaceutical activity.⁵⁸

These advantages depend on some chemical properties of the fluorine atom. Fluorination increases metabolic stability of a drug, due to the higher dissociation rate of the C-F in comparison to a C-H bond. Fluorines increase the acidity of acids and decrease the basicity of bases, making it possible to fine-tune pK_a values. Permeability of the compound is increased by enhancing lipophilicity, either through placement of the fluorine atom in aromatic rings or nearby basic nitrogens or by the ability of fluorine atoms to form intramolecular H-bonds and thereby reducing the polar surface. The binding affinity can be enhanced directly by formation of H-bonds or lipophilic interactions with the binding partner or indirectly by influencing protein mediated H_2O . Nevertheless some considerations need to be taken into account when adding a fluorine to a compound. Fluorine addition can have an effect on the ligand conformation, it influences the polarization of covalent bonds and thereby changes the electron density, it has an inductive effect on vicinal functional group acidity and it can change desolvation.⁵⁸

1.4.1 Fluorine Libraries

Indeed fluorine atoms cannot only be added to preexisting structures, but are also useful in fragment based drug design (FBDD). In FBDD a library is screened for small weakly binding fragments of a biological target. Subsequently this weakly binding fragments can be combined or grown to obtain a high affinity ligand. One library that can be used for such a screening process is the fluorinated Novartis library LEF (local environment of fluorine). This library is composed of compounds from the Novartis Compound Collection and of compounds procured from commercial vendors. To limit complexity of the ^{19}F NMR spectra, compounds with more than one nonequivalent fluorine atom are not included. Another possibility for obtaining fluorinated fragments is the defragmentation of known binders and subsequent fluorination or the use of fluorinated peptides with different amino acid composition and length.⁵⁸ The fragments are of small size and are limited in functionality, therefore their binding affinities are rather weak, in the range of $10^{-4} - 10^{-2}$ M.⁵⁹ The premise of this library is that fluorines with a similar local fluorine environment will give a similar chemical shift and thus that the chemical

shift can be predicted for most compounds. If it is not possible to predict the chemical shift, it is estimated from values in the literature using typical chemical shifts for similar compounds.⁵⁸ If prediction of the correct chemical shift is not possible due to aggregation, degradation or low solubility, because of different pH or ionic strength or because the composition of the mixture is not known, the resonances can be assigned properly using a combination of 2D ^{19}F DOSY and heteronuclear 2D ^{19}F - ^1H COSY experiments.¹⁵

1.4.2 ^{19}F NMR

Knowledge of the fluorine environment can have an impact on the protein-ligand binding properties and hence be exploited for lead optimization to enhance potency or selectivity of the lead structure.¹⁴ Information on interaction properties of fluorine atoms were retrieved from X-ray crystallography structures submitted in the PDB. Shielded fluorine atoms are found in close contact to H-bond donors, while deshielded fluorine atoms are found in the neighbourhood of hydrophobic residues.¹³ Also the position and bond distance of oxygen, nitrogen or halogen atoms with respect to the fluorine atom has a large impact on the fluorine chemical shift. Increased electron density increases the shieldedness and leads to a shift to lower ppm ranges and also the other way around.⁵⁸ Fluorophilic sites in the target protein can be detected using a combination of NMR, X-ray crystallography and computational methods.⁶⁰ The binding of the target protein to the fluorinated fragments can be tested in large fragment mixtures (≥ 30) at very low protein concentration (10 - 50 μM). Thus only small amounts of target protein are needed and the screening process can be completed in a short time.⁵⁸ Moreover, multiple binders can be identified in the mixture since ^{19}F has a large chemical shift dispersion from approximately -272 ppm to +80 ppm. Binders are identified using a ^{19}F R_2 filter CPMG experiment that is recorded in presence and absence of the target protein. Binding results in a reduction of the signal intensity, hence only bound fragments can be seen in the difference spectrum and the ^{19}F resonance can be assigned to the correct molecule.¹³

1.5 Selective Labelling of Amino Acids

1.5.1 GB1

The immunoglobulin binding domain B1 of streptococcal protein G (also termed GB1) is a 56 residue single domain protein. It comprises one α -helix and one four-stranded β -sheet. Due to its small size and well resolved NMR

spectra it can be used as a model protein for new NMR experiments¹⁰ or in protein expression as a fusion protein, that enhances solubility, stability and facilitates refolding of the fusion partner.⁶

1.5.2 Purpose of Selectively Labelled Amino Acids

Isotope labelling of proteins with ^{13}C and ^{15}N is a common strategy in biomolecular NMR and enables the execution of heteronuclear experiments. Addition of ^{13}C carbon and ^{15}N nitrogen sources to the medium results in homogeneously labelled samples. Further information can be gained by selective labelling of certain amino acids.⁵³ This method can be used to facilitate assignment, avoid signal overlap in larger proteins (up to 1 MDa),¹⁹ conduct functional experiments or learn more about the active or binding site.⁵³ Of special interest are the methyl groups of isoleucines, leucines and valines, that make up 50 % of all methyl groups in proteins¹⁹ and are frequently found in the hydrophobic core of proteins and in protein-ligand interfaces.³⁶ Another important group of amino acids for selective labelling are the aromatic amino acids. They are also predominantly found in the hydrophobic core of proteins and can interact with binding partners via hydrophobic, cationic- π and polar- π forces. Unfortunately aromatic amino acids show a very narrow chemical shift dispersion and are therefore prone to signal overlap.⁴⁰

1.5.3 Methods to Selectively Label Amino Acids

Selective labelling can be either applied in cell-free or cell-based systems. Advantages of the cell-free system are a higher yield for certain proteins and the feasibility of conditions that are normally not possible in cell-based systems, like expression of proteins that are otherwise toxic for the cell. This method also causes less scrambling or dilution due to the branching of metabolic pathways and allows correct folding and modification of eukaryotic proteins, if eukaryotic cell-lysates are used. Labelled amino acids can be obtained either through chemical synthesis or hydrolysis of fully labelled proteins. Chemical synthesis has the advantage of being able to specifically position the label. However, often only racemic mixtures are produced, where the D-enantiomeres might exert an inhibitory effect. To obtain only the L-form an additional expensive purification step needs to be included. In contrast amino acids obtained through hydrolysis of fully labelled proteins, contain only the L-form of amino acids, but are not specifically labelled (e.g all carbons in the amino acid are ^{13}C labelled, not only a specific one).⁵³ Assignment of selectively labelled amino acids can be achieved using a combinatorial approach. To this end a certain number of samples with different combinations

of labelled amino acids is used.²⁴

Selective labelling of methyl groups in Ile, Leu and Val is often achieved by using α -ketoacids as precursors. α -ketobutyrate is used for the selective labelling of isoleucines,³⁶ while α -ketoisovalerate can be used for the simultaneous labelling of leucines and valines.³² A major disadvantage of using α -ketoacids is the racemic nature of the precursor, which leads to a non-stereospecific labelling of the methyl groups. This problem is negligible in small or medium sized proteins, but striking in larger proteins comprising significant signal overlap. Hence a different stereoselective approach is needed for larger proteins. This can be achieved by using methyl labelled acetolactate as a precursor for Leu and Val labelling. Acetolactate is an intermediate, that is further metabolized into α -ketoisovalerate in the biogenesis of leucines and valines. This method exploits the stereospecific rearrangement of a methyl group in acetolactate resulting in the generation of stereospecific labelled leucines and valines. Application of this method leads to a simplification of NMR spectra, enhanced sensitivity and does not interfere with other metabolic pathways.¹⁹ Another possible option to obtain stereospecifically labelled valines and leucines is to directly incorporate stereospecific labelled valines and leucines. Of course this method demands the expensive synthesis of stereospecific labelled valines and leucines. Recently a few new methods have been developed to individually label valines and leucines. Discrimination between valines and leucines can be either solved in data processing by spectral editing³⁶ or by applying one of the following expression schemes. Administering α -ketoisocaproate as a labelled precursor results in exclusive labelling of leucines.³² Instead selective labelling of valines can be achieved by adding labelled α -ketoisovalerate and unlabelled α -ketoisocaproate.³¹ A similar approach leading to selective labelling of valines, is the addition of labelled acetolactate and unlabelled L-leucine. Addition of L-leucine causes dilution of labelled leucine, derived from acetolactate, and inhibits 2-isopropylmalate synthase, resulting in exclusive labelling of valines. The use of acetolactate also has the advantage of again producing stereospecific labelled valines.³³

Selective labelling of aromatic amino acids is less well described in the literature. One approach is to administer the desired labelled aromatic amino acid and inhibit scrambling by also adding the unlabelled remaining aromatic amino acids and one g/L glyphosate, to repress the biosynthesis of aromatic amino acids in the cell.⁴ Another option is to selectively unlabel amino acids. This can be applied to all amino acids and has already been conducted with tyrosine and phenylalanine. The idea is to uniformly label the expression medium with ^{13}C and ^{15}N and to unlabel the desired amino acid using unlabelled precursors. 4-hydroxyphenylpyruvic acid was used to

unlabel tyrosines and phenylpyruvate to unlabel phenylalanines. One advantage of this method is, that less scrambling can be observed. Nevertheless relatively high amounts of unlabelled precursors need to be used, 800 mg/L for Tyr and 400 mg/L for Phe, in comparison to 100 mg/L in the standard method.⁴⁰

1.6 Biomolecular NMR

Biomolecular NMR is an important method for the determination of structure and dynamics in macromolecules like proteins or nucleic acids. A major advantage of NMR over X-ray crystallography is its applicability in dynamic and kinetic research questions, since molecular mobility is an integral part of protein function. Furthermore, NMR can be used to study interaction, either by investigating the binding site or by determining rate constants of the reaction. Due to significant signal overlap, experiments with a higher dimensionality ($\geq 2D$) are necessary for the investigation of proteins. Although NMR is a versatile tool for studying proteins, it is usually limited to proteins of a size under 30 kDa. This problem is currently addressed by the development of new NMR and selective labelling techniques, which allow the assignment of amino acids in a 1 MDa protein.⁴⁴

1.6.1 HSQC

One of the most used NMR experiments is probably the HSQC experiment, a heteronuclear NMR experiment that correlates carbon or nitrogen bound protons and carbon or nitrogen chemical shifts. The outcome of this experiment is a chemical shift pattern that is specific for the respective protein and can be seen as a sort of “fingerprint” for the protein. Typically the heteroatom is displayed in the ω_1 - and the proton in the ω_2 -dimension. In order to observe these resonances the sample needs to be fully ^{13}C or ^{15}N labelled. One drawback is the lower sensitivity of ^{13}C and ^{15}N due to their lower gyromagnetic ratios. The loss in sensitivity can be counteracted by returning the magnetization to the more sensitive proton spins for detection or by decoupling the proton-heteronuclear interaction.⁴⁴ A typical HSQC experiment starts with equilibrium magnetization on the I spin (proton), subsequent transfer of the magnetization to the S spin (heteronucleus) via an INEPT sequence, evolution of the S spin magnetization and refocusing of the coupling during time t_1 and finally transfer of the magnetization back to the I spin for detection. A very similar experiment is the HMQC experiment. The only difference is the evolution of a double-quantum state ($I_x S_x$) during

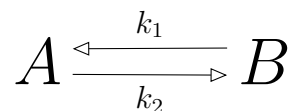
the t_1 period in the HMQC experiment compared to a single-quantum state ($I_z S_x$) in a HSQC experiment.²³

SOFAST HMQC

The SOFAST HMQC is an optimized version of the HMQC experiment and allows fast data acquisition due to very short interscan delays. Two main features are a short acquisition time and increased sensitivity. This method is ideally suited for high throughput screenings and real-time experiments studying kinetic processes. Moreover, band selective ^1H pulses, and hence only manipulation of a specific subset of protons, avoids excitement of unwanted proton spins, shortens the relaxation time T_1 and makes it a suitable experiment for IDPs.^{46,47}

1.6.2 CPMG

The CPMG experiment is used to measure the exchange of a protein between two or more environments. It was named after their inventors Carr, Purcell, Meiboom and Gill. The rate of the exchange process influences the appearance of the resulting spectra and can be used to extract certain parameters, like the apparent exchange rate k_{ex} , the populations of the different states or the frequency separation between the states $\Delta\omega$.⁴⁴



$$k_{ex} = k_1 + k_2 \quad (1.1)$$

$$\Delta\omega = \omega_A - \omega_B \quad (1.2)$$

$$p_A = \frac{k_2}{k_1 + k_2} \quad (1.3)$$

$$p_B = \frac{k_1}{k_1 + k_2} \quad (1.4)$$

In the case of slow exchange, $k_{ex} > \Delta\omega$, two lines can be discriminated in a two-state system, in the intermediate exchange regime, $k_{ex} \approx \Delta\omega$ the lines are extensively broadened and can even disappear from the spectrum and in fast exchange, $k_{ex} < \Delta\omega$, only one line can be observed. Relaxation

dispersion experiments exploit the dependence of the relaxation rate on the magnetic field strength and are therefore measured at different magnetic field strength. Subsequently the relaxation dispersion curves can be fitted to theoretical models.⁴⁴

The CPMG pulse sequence is derived from a Hahn echo. The first part was developed by Carr and Purcell and is executed as follows: A first 90° pulse flips the magnetization into the x-y plane, then a series of 180° pulses follows. Whereupon the time between the individual 180° pulses must be twice the time between the first 90° pulse and the first 180° pulse. The echoes between the subsequent 180° pulses are observed. The measured amplitude of the successive echoes must decay exponentially with a time constant equal to T_2 . A critical step in this experiment is the accurate application of the 180° pulses, since any deviation from 180° leads to a cumulative error. For this reason Meiboom and Gill developed an advanced experiment that corrects for this error by shifting the first 90° pulse with respect to the first 180° by exactly 90° .³⁴

1.6.3 HNCO

The HNCO is one of the simplest triple-resonance experiments and correlates the amide group with the carbonyl group of the preceding residue. Triple resonance experiments are in general used for the assignment of chemical shifts using the resonances of ^1H , ^{13}C and ^{15}N . Thus homogeneous labelling of the sample with ^{13}C and ^{15}N is necessary. Typical triple-resonance experiments for assigning the protein backbone are HNCA, HNCO, HN(CO)CA or HN(CA)CO.⁴⁴

Chapter 2

Aim of the Work

2.1 Aim of the Work

The 25 kDa lipocalin NGAL is an important player in many physiological and pathophysiological functions. It is associated with iron transport and metabolism, host defence against bacteria and iron dependent gene regulation, but also with pathological processes like inflammation or cancer.⁵⁰

The work described in this thesis can be divided into three distinct parts. The study of the interaction of NGAL with its cellular receptor 24p3R, finding of new ligands for NGAL and the development of new selective labelling techniques demonstrated on the model protein GB1.

The objective of the first part was to shed more light on the interaction of NGAL with its cellular receptor 24p3R and thus to learn more about the physiological function of NGAL. In this study only the soluble N-terminal part of 24p3R was used. It has been already shown, that NGAL is binding to this N-terminal domain and that this system can be used to study their interaction via solution state NMR [Nicolas Coudeville, unpublished data]. To this end we performed various titration series of NGAL to 24p3R, recording SOFAST HMQC experiments at every titration point and extracting information about the binding site and the binding affinity. To obtain further insight into the binding event, we designed NGAL mutants carrying paramagnetic tags, like MTSL or lanthanide binding peptides, to employ NMR experiments that exploit the presence of paramagnetic tags to gain long-range distance information.¹¹

The second part was dedicated to the finding of new ligands for NGAL. It was conducted in cooperation with Marina Veronesi, who screened a library of fluorinated-fragments for putative NGAL ligands. These ligands were then tested for their binding properties to NGAL by HSQC and ITC. Fluorines are a widespread moiety in drugs, because of their positive influence on metabolic stability, lipophilicity, permeability and binding affinity. Furthermore, fluorinated ligands can be easily investigated by ¹⁹F NMR and weakly binding fragments can be used as a starting point for a FBDD approach.⁵⁸

Selective labelling of amino acids is a useful tool in NMR to facilitate assignments, investigate binding sites, avoid signal overlap or to make larger proteins accessible for NMR experiments. In the third part of this work we aimed to develop new selective labelling techniques for valines, tyrosines and phenylalanines, which are all frequently found in binding sites or in interaction surfaces.³³ As a model protein we used the 56 residue protein GB1, that can be easily expressed and offers well resolved NMR spectra. To test correct incorporation of the label we used standard ¹H-¹³C HSQC and HNCOC experiments.

In summary we aimed to elucidate binding properties of NGAL to its cel-

lular receptor and to new ligands. Additionally we tried to develop selective labelling techniques that could be further used to examine this interaction.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 Cells

Cells	
Cell Strain	Resistance
<i>BL21pLysS</i>	Chloramphenicol
<i>DH5a</i>	none
<i>W3110 Δ<i>fur::cat</i></i>	Chloramphenicol

Table 3.1: Cell Strains

3.1.2 Buffers and Media

Buffers

10x Tris pH 7.4	
Compound	Concentration
Tris	200 mM
NaCl	500 mM
dissolve in 1 L ddH ₂ O, adjust pH to 7.4, filtrate and degas before use dilute 1:10 with ddH ₂ O	

Table 3.2: 10x Tris pH 7.4

10x PBS	
Compound	Concentration
NaCl	1.37 M
KCl	27 mM
Na ₂ HPO ₄	100 mM
KH ₂ PO ₄	18 mM
dissolve in 1 L of ddH ₂ O, adjust pH to 7.4, filtrate and degas before use dilute 1:10 with ddH ₂ O	

Table 3.3: 10x PBS

HS PBS	
Compound	Concentration
NaCl	1.5 M
Imidazole	20 mM
dissolve in desired amount of 1x PBS, adjust pH to 7.4, filtrate and degas	

Table 3.4: HS PBS

HiTrap Elution Buffer NGAL	
Compound	Concentration
Imidazole	500 mM
dissolve in desired amount of 1x PBS, adjust pH to 7.4, filtrate and degas	

Table 3.5: HiTrap Elution Buffer NGAL

HiTrap Elution Buffer 24p3R	
Compound	Concentration
Imidazole	500 mM
dissolve in desired amount of 7 M urea, adjust pH to 7.4, filtrate and degas	

Table 3.6: HiTrap Elution Buffer 24p3R

TEV-Buffer	
Compound	Concentration
DTT	1 mM
EDTA	0.5 mM
add to 100 mL of Tris-buffer and adjust pH to 7.4	

Table 3.7: TEV-Buffer

Media

LB Broth	
Compound	Concentration
Tryptone	10 g/L
Yeast Extract	5 g/L
NaCl	5 g/L
dissolve 20 g in 1 L ddH ₂ O and autoclave	

Table 3.8: LB Broth

M9	
Compound	Concentration
Na ₂ HPO ₄ ·2H ₂ O	7.5 g/L
KH ₂ PO ₄	3.0 g/L
NaCl	0.5 g/L
¹⁴ N/ ¹⁵ N-NH ₄ Cl	1.0 g/L
dissolve in 970 mL ddH ₂ O and autoclave	
add after autoclaving and before expression:	
Glucose	4 g/L (¹² C) or 3 g/L (¹³ C)
Trace Elements	10 mL
Antibiotic	1 mL (each)
1 M MgSO ₄	2 mL
1 M CaCl ₂	300 µL

Table 3.9: M9

Chemicals

M9 Additives 1	
Compound	Concentration
CaCl ₂	1 M
MgSO ₄	1 M
autoclaved 121°C, 30 min	

Table 3.10: M9 Additives 1

M9 Additives 2		
Name	Compound	Concentration
20 % Glucose	D-Glucose	200 g/L
sterile filtered, 20 mL aliquots, stored at room temperature		
Trace Elements		
sterile filtered, 50 mL aliquots, stored at room temperature		

Table 3.11: M9 Additives 2

Antibiotics		
Compound	Concentration	Solvent
Kanamycine	25 mg/mL	ddH ₂ O
Chloramphenicol	25 mg/mL	Ethanol
sterile filtered, stored in 1 mL aliquots at -20°C		

Table 3.12: Antibiotics

Chemicals	
Compound	Concentration
NiCl ₂	0.1 M
EDTA	50 mM
Guanidinium HCl	6 M
IPTG	0.4 M
Ketoisocaproate	10 mg/mL
Ketoisovalerate	5 mg/mL
Hydroxy- α -ketophenylpyruvate	10 mg/mL
α -ketophenylpyruvate	10 mg/mL
TEV	1.8 mg/mL
Urea	7 M

Table 3.13: Chemicals

3.1.3 Cloning

Cloning		
Compound	Concentration	Provider
Pfu DNA Polymerase	2.5 U/ μ L	Thermo Scientific
Pfu Buffer with MgSO ₄	10x	Thermo Scientific
MassRuler Low Range DNA Ladder		Thermo Scientific
QIAquick PCR Purification Kit		Quiagen
NcoI	10 U/ μ L	Thermo Scientific
NdeI	10 U/ μ L	Thermo Scientific
Buffer O	10x	Thermo Scientific
petM11	200 ng/ μ L	Thermo Scientific
T4 DNA Ligase		Thermo Scientific
T4 DNA Ligase Buffer	10x	Thermo Scientific

Table 3.14: Cloning

Primers		
Primer	Sequence	Provider
NGAL LBP 3 Ct for	CAACGATGGCTGGATTGAAGGCGATGAA CTGTATGACCTGATCCAGCCCCACCTCTG	Sigma- Aldrich
NGAL LBP 3 Nt rev	CAATCCAGCCATCGTTGTTGGTATCAATAT AGCCTGAGGTGGAGTCCTGACCCATGGC	Sigma- Aldrich
NGAL LBP Stop Ct for	CAACGATGGCTGGATTGAAGGCGATGAA CTGTATTGAGCGGCCGCACTCGAGCACC	Sigma- Aldrich
NGAL LBP Stop Nt rev	CAATCCAGCCATCGTTGTTGGTATCAATAT AGCCGCCGTCGATACACTGGTCGATTGGG	Sigma- Aldrich
T7 Promotor	TAATACGACTCACTATAGGG	Invitrogen
T7 Terminator	GCTAGTTATTGCTCAGCGG	Invitrogen

Table 3.15: Primers

3.1.4 Equipment

Equipment	
Name	Provider
Äkta Explorer	GE Healthcare
HiLoad 26/60 Superdex 75 prepgrade	GE Healthcare
HiTrap Chelating HP	GE Healthcare
Varian Inova 500	Agilent
Varian Direct-Drive 600	Agilent
Varian Inova 800	Agilent
MicroCal ITC200	GE Healthcare
Mastercycler Gradient	Eppendorf
NanoDrop 2000C	Thermo Scientific
Homogenizer EmulsiFlex-C5	Avestin
Amicon Ultra-15 Centrifugal Filter Units	Merck Millipore
PD-10 Desalting Columns	GE Healthcare

Table 3.16: Equipment

3.1.5 Software

Software	
Name	Provider
NMRDraw Version 3.0	NIH Laboratory of Chemical Physics
Sparky 3.110	T.D. Goddard and D.G. Kneller, University of California, San Francisco
NMRPipe Conversion Utility Version 2007.030.16.06	NIH Laboratory of Chemical Physics
OpenOffice 3.2 Calc	Apache Software Foundation
VNMRJ 2.1	Agilent
MATLAB 7.8.0	MathWorks

Table 3.17: Software

3.2 Methods

3.2.1 Protein Expression and Purification

NGAL

Expression: For the transformation 10 ng plasmid (petM11) are added to 100 μ L competent *BL21pLysS* cells, that were thawed on ice. The mixture is kept on ice for another 15 minutes and then heat shocked for 90 seconds at 42°C. The cells are immediately put on ice for 2 minutes. After addition of 500 μ L of LB the cells are incubated at 37°C for 1 hour. 100 μ L of these cells are plated on a kanamycine/chloramphenicol LB agar plate and incubated over night at 37°C. On the next day the plate is sealed with parafilm and stored at 4°C until further use. In the evening before expression a preculture is started. 5-10 colonies of the transformed cells are incubated in 10 mL LB containing kanamycine and chloramphenicol per 1 L culture and grown over night at 37°C or over the weekend at room temperature. 10 mL preculture are added to 1 L M9 containing the corresponding antibiotics and supplemented with either $^{14}\text{N-NH}_4\text{Cl}$ or $^{15}\text{N-NH}_4\text{Cl}$. For detailed description of the constituents see the Materials section. Cells are grown at 37°C until they reach an OD_{600} of 0.8, then expression is induced with 0.8 mM IPTG at 30°C over night.

In the morning cells are harvested at 6000 rpm, 15 minutes and 4°C. The cell pellet is resuspended in ca. 30 mL PBS and passed through a homogenizer for 30 minutes with a pressure of 1 bar. To remove cell debris the suspension is centrifuged at 18 000 rpm for 30 minutes at 4°C. The supernatant is set aside for purification and the pellet discarded.

Purification: A HiTrap Chelating HP 5 mL is loaded with NiCl_2 and equilibrated as recommended by the provider. Then the supernatant is loaded on the column with a flow of 2 mL/min. The column is washed with PBS until the absorption at A_{280} is at the baseline again. Next the column is washed with three column volumes of HS PBS, followed again by a wash with PBS until the absorption at A_{280} has reached the baseline again. Finally the protein is eluted with HiTrap elution Buffer and collected in 1 mL fractions in a 96 well plate. The protein containing fractions are pooled and dialysed over night in 500 mL Tris pH 7.4. On the next day the dialysed protein is concentrated to ca. 500 μ L with an Amicon Centrifugal Filter and washed two times with 10 mL TEV-Buffer. The HisTag is cleaved off with 1 mg TEV per 50 mg protein over night rotating at 4 °C.

On the next day the sample is loaded on a pre-equilibrated HiLoad 26/60

Superdex 75 prep grade column and eluted with one column volume of Tris pH 7.4 with a flow of 3 mL/min. 1.5 mL fractions are collected in 96 well plates. Peak fractions are chosen according to the chromatogram and concentrated to about 500 μ L with an Amicon Centrifugal Filter. To unfold NGAL, and thereby remove its cofactor enterobactin, Guanidinium HCl is added to the sample to reach a final concentration of 6 M. This is followed by a 20 minute incubation at 70°C. Afterwards the sample volume is adjusted to 2.5 mL with Guanidinium HCl and applied on a PD-10 Desalting Column already pre-equilibrated with 25 mL 6 M Guanidinium HCl. The first flow-through is discarded and another 3.5 mL Guanidinium HCl applied to the column. The second flow-through contains NGAL without the cofactor, while enterobactin is retained in the desalting column. The flow-through is collected in a 50 mL BD Falcon Tube, immediately mixed with 15 mL Tris pH 7.4 and stepwise dialysed with 1.5 L of Tris pH 7.4. Finally the dialysed protein is concentrated to a volume of about 1 mL, the concentration measured with a NanoDrop 2000C and the sample stored at -20°C.

This purification procedure was applied to all versions of NGAL: NGAL, NGAL60, NGAL85, NGAL3 and NGALstop.

LBP Loading: NGAL mutants containing a lanthanide binding peptide (LBP) have to be loaded with the corresponding lanthanide. For NMR experiments we used TbCl_3 as a paramagnetic probe and LuCl_3 as a non-paramagnetic probe. The loading was achieved by diluting the protein sample to 25 μ M with Tris pH 7.4 and stepwise addition of 10 times 1.1 equivalents of lanthanide. The sample is then reconcentrated to 500 μ L, resulting in a final concentration of 0.5 mM NGAL and 0.55 mM lanthanide.

Enterobactin

Enterobactin is derived from *E.coli* W3110 $\Delta fur::cat$ cells, a strain that is deregulated in enterobactin biosynthesis. 1 mL overnight LB preculture are added to 500 mL M9 medium and grown for approximately 24 h at 37°C. Cells are separated by centrifugation at 6000 rpm for 15 min and the supernatant retained. Extraction is performed by solvent extraction in a separatory funnel with two times 100 mL ethylacetate. The resulting ethylacetate fractions are washed with 100 mL citric buffer pH 4.5 and ddH₂O and subsequently evaporated in a rotavapor. The residue is dissolved in a few mL of ethylacetate and crystallized in 40 mL hexane on ice. The crystals are harvested by centrifugation at 4000 rpm 15 min, dried at room temperature for 30 min and dissolved in 500 μ L methanol. The concentration is determined by measuring the absorption at A_{495} upon addition of 1.5 equivalents of Fe^{3+} and

six equivalents of KOH. The obtained sample is stored at -20 °C.

24p3R

Expression: Expression is conducted as described for NGAL (see section NGAL Expression).

Purification: After breaking the cells in a homogenisator and centrifugation at 18 000 rpm, 30 minutes, 4°C the supernatant is discarded and the pellet resuspended in ca. 12 mL 7 M urea. To increase the solubilized part the sample is rotated at 4 °C over night. On the next day the sample is sonicated two times three minutes at 40 % amplitude with a 10 minute break. Then the sample is filtered through a 0.45 µm filter and loaded on a pre-equilibrated Hitrap Chelating HP with a flow of 2 mL/min. As a binding buffer 7 M urea is used. The column is washed with 7 M urea until the absorption at A_{280} has reached the baseline. The column is washed with 3 column volumes of HS 7 M urea and 7 M urea until A_{280} is back at the baseline. Finally the protein is eluted with 0.5 M imidazole and collected in 1 mL fractions. Protein fractions are pooled, concentrated, diluted 1:10 with Tris pH 7.4 with 1 mM DTT and dialysed over night in 500 mL Tris pH 7.4 with 1 mM DTT. The dialysed protein is concentrated and washed with two times 10 mL TEV-Buffer. TEV cleavage of the His-Tag is performed at 4°C over night at a ratio of 1 mg TEV per 50 mg protein. The protein is then further purified on a HiLoad 26/60 Superdex 75 prep grade. Fractions that contain, according to the chromatogram, 24p3R are pooled, concentrated and stored at -20°C. Purity of the protein sample is checked by SDS PAGE or via NMR.

GB1

GB1 expression and purification was conducted as described for NGAL and 24p3R, except that no TEV-cleavage and no refolding step were performed.

Selective Labelling: The labelled and unlabelled precursors are dissolved in ddH₂O and sterile filtered with a 0.22 µm filter. Sterile addition of the desired amount of precursor is performed 1 h prior to induction of the cells. Induction, expression and purification are performed as described above. All samples are fully ¹⁵N labelled.

3.2.2 ITC

ITC measurements were performed on a MicroCal ITC200 with 0.1 mM NGAL in the reservoir and 3 mM ligand E2 in the syringe. All experiments were conducted in Tris pH 7.4 at 25°C. The first injection has a volume of 0.5 μ L, the following six a volume of 1 μ L and the last 15 a volume of 2 μ L, resulting in a total injection volume of 36.5 μ L. A delay of 180 s is set between injections and the rotating speed of the syringe is 1000 rpm.

3.2.3 NMR

NMR experiments were conducted on Varian spectrometers operating at 500, 600, or 800 MHz at 25 °C. All samples were measured in Tris pH 7.4 (details see Materials section) containing 10 % D₂O. Enterobactin containing NMR samples are loaded with gallium instead of iron to avoid paramagnetic effects. NMR data was processed using NMRPipe, Sparky, MATLAB and OpenOffice Calc. CPMG data was analysed using the MATLAB program rdNMR. Fitting of K_d-Values was performed with QtiPlot.

3.2.4 Cloning

PCR

The PCR to generate NGAL3 and NGALstop inserts is divided into 2 steps. The first PCR is conducted to obtain the C-terminal and N-terminal fragments with overlaps in the middle. The second PCR is performed to anneal these two parts. For the edges of the construct T7 promotor and terminator primers are used. To insert the LBP sequence at the N(3)- or C(stop)-terminus special primers termed Nt rev (3 or stop) and Ct for (3 or stop) are designed. The primers T7 promotor and Nt rev generate the N-terminal fragment and the primers T7 terminator and Ct for the C-terminal fragment. For the first PCR 100 ng of the respective primers, 10x Pfu Buffer + MgSO₄, 10x dNTPs, 50 ng NGAL template, 1 μ L Pfu DNA Polymerase and ddH₂O are mixed to culminate in a total volume of 50 μ L. The applied PCR programme is shown in figure 3.18. The PCR product is purified using the Quiagen DNA Purification Kit according to manufacturers instruction and the purified DNA eluted with 40 μ L ddH₂O. For the second PCR the whole purified PCR product is used and only 10x Pfu Buffer + MgSO₄, 10x dNTPs and 1 μ L Pfu DNA Polymerase are added. The same PCR program as for the first PCR is used and the obtained PCR product analysed on a 1 % agarose gel.

PCR		
Step	Temperature	Time
1 Initialization	98°C	2' 30"
2 Denaturation	98°C	30"
3 Annealing	55°C	1'
4 Elongation	68°C	2'
repeat step 2-4 24x		
5 Final Elongation	68°C	10'
6 Hold	4°C	

Table 3.18: PCR

Restriction

The PCR product is purified using the Quiagen DNA Purification Kit according to the manufacturers instructions and the purified DNA eluted with 40 μ L ddH₂O. 1 μ L NcoI and NdeI each and 5 μ L Buffer O are added to the purified PCR product and incubated over night at 37°C.

Ligation

Concentration of the inserts is determined by measuring the absorption at A_{260} . Three equivalents of insert (600 bp) are mixed with one equivalent (50 ng) of vector petM11 (5000 bp), 10x T4 DNA Ligase Buffer, 1 μ L T4 DNA Ligase and ddH₂O to a final volume of 20 μ L. The mixture is incubated for one hour at room temperature.

Transformation in DH5 α

The transformation of the NGAL3 and NGALstop plasmids into DH5 α cells is carried out as described in the expression section of NGAL. The cells are plated on kanamycine containing agar plates and incubated over night at 37°C.

Miniprep

The plasmid is extracted from DH5 α cells using the QIAprep Spin Miniprep Kit according to the manufacturers instructions.

Transformation in BL21pLysS

For expression NGAL3 and NGALstop plasmids are transformed into BL21pLysS cells as described in the NGAL expression section.

Chapter 4

Results

4.1 NGAL and its Receptor 24p3R

The membrane protein 24p3R is the cellular receptor of NGAL and responsible for the recognition and subsequent endocytosis of NGAL, a crucial step in the NGAL cycle.¹⁷

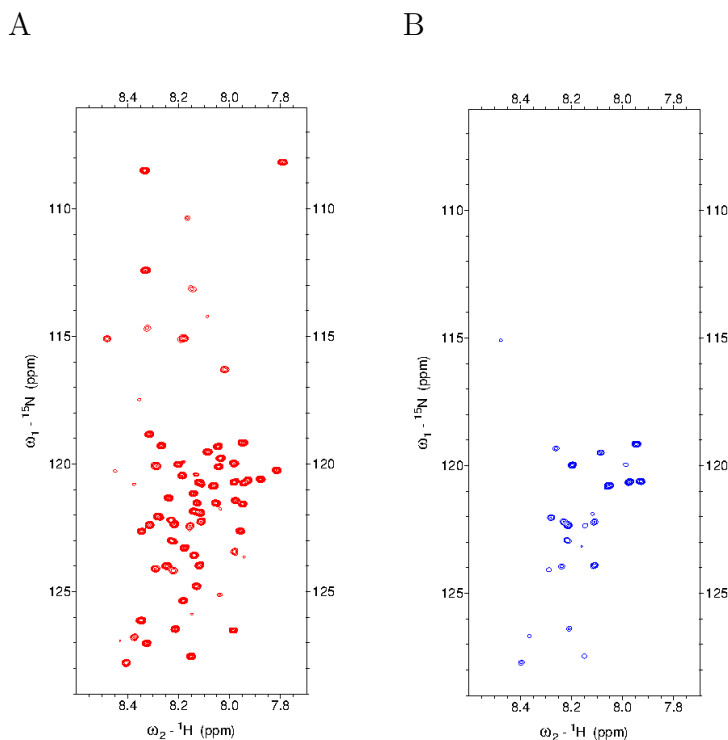


Figure 4.1: ^1H - ^{15}N SOFAST HMQC of the NTD without and with NGAL

A: ^1H - ^{15}N SOFAST HMQC of 0.5 mM NTD without NGAL. **B:** ^1H - ^{15}N SOFAST HMQC of 0.5 mM NTD with one equivalent of unlabelled NGAL.

In this study we used the intrinsically disordered soluble N-terminal domain of 24p3R, which was previously found to bind to NGAL [Nicolas Coudeville, unpublished data]. Additionally we used a cysteine mutant (C15S, C59S, C70S and C94S) of the N-terminal domain that is providing nicer NMR spectra than the wild type (further referred to as NTD). A ^1H - ^{15}N SOFAST HMQC spectrum of the NTD is shown in figure 4.1A. Upon addition of unlabelled NGAL a local reduction in signal intensity can be observed (figure 4.1B). Reduction in signal intensity can be a sign for chemical exchange between different environments. Hence this behaviour lead us to the

hypothesis, that NTD and NGAL form a disordered complex, where NTD and NGAL adopt more than one conformation upon binding.

4.1.1 HMQC Titration Series of NGAL to NTD

To investigate which parts of the NTD are affected most by binding to NGAL, we performed a titration series of unlabelled NGAL to ^{15}N labelled NTD. The result of this experiment can be seen in figure 4.2A. The experiment was conducted at a concentration 0.5 mM NTD with increasing amounts of NGAL, ranging from 0 to 2 equivalents.

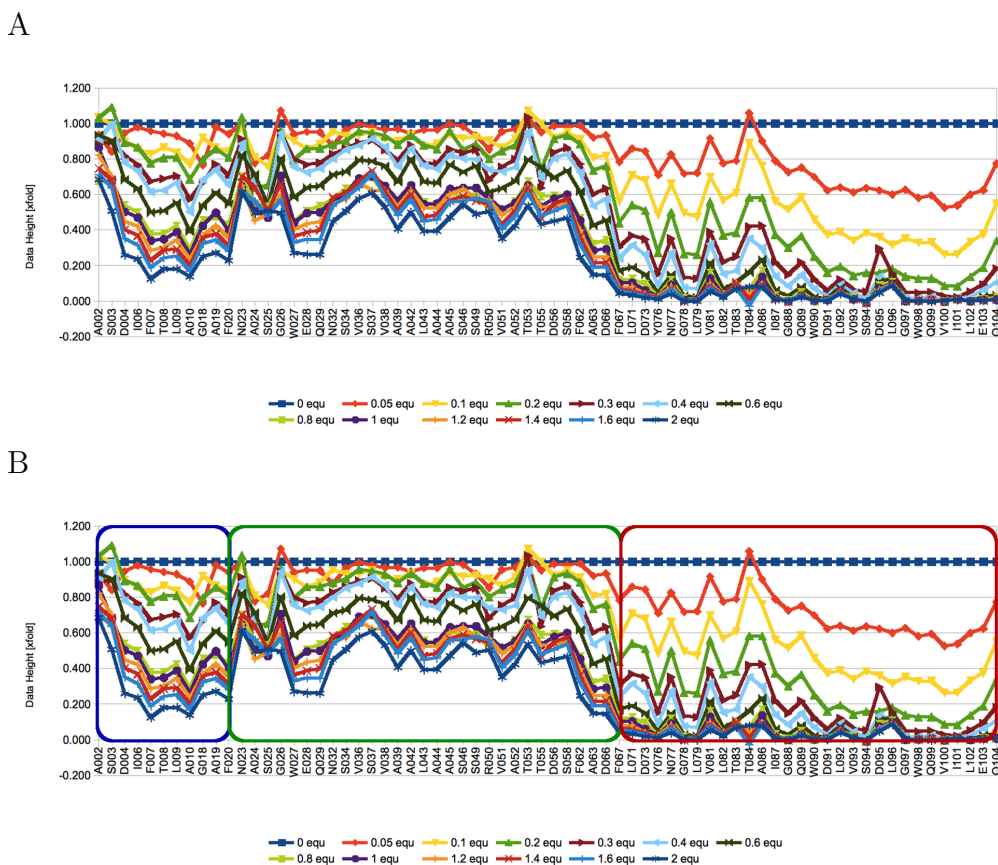


Figure 4.2: Titration series of ^{14}N -NGAL to ^{15}N -NTD

A: ^1H - ^{15}N SOFAST HMQCs were recorded for ^{15}N -NTD with increasing amounts of unlabelled NGAL (concentration see figure legend).

B: Differently affected regions of the NTD upon binding of NGAL. Region 1 (highlighted in blue) is affected by the binding, region 2 (highlighted in green) is not involved in the binding and region 3 (highlighted in red) seems to be the main binding site.

On closer inspection three distinctly affected parts can be detected in the NTD. The least affected is the middle part from residue F020 to F067 (highlighted in green in figure 4.2B). The part that is most affected is the C-terminal part from residue L071 to Q104 (highlighted in red in figure 4.2B). Whereas the N-terminal part from residue A002 to A019 (highlighted in blue in figure 4.2B) seems to be involved in the binding, but not as strongly as the C-terminal part. Thus these three areas will from now on be referred to as region 1 for the N-terminal part, region 2 for the middle part and region 3 for the C-terminal binding part.

$$I = I_{free} - I_{bound} * \frac{P + x + K_d - \sqrt{(P + x + K_d)^2 - 4 * P * x}}{2 * P} \quad (4.1)$$

I_{free} = intensity of the free state

I_{bound} = intensity of the bound state

P = protein concentration

x = ligand concentration

K_d = dissociation constant

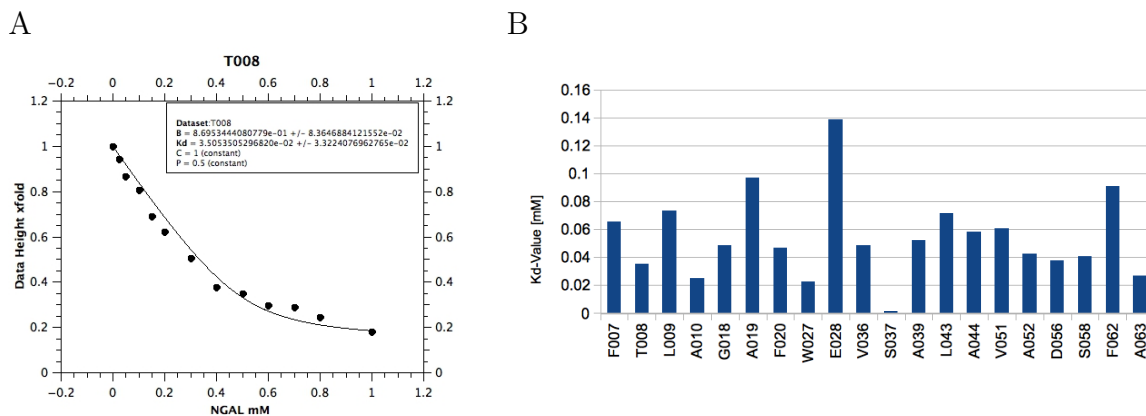


Figure 4.3: **K_d value of NTD and NGAL**

A: Fitting curve for residue T008 in QtiPlot using equation 4.1. The resulting K_d value for this residue is 35 μM. **B:** K_d values for all residues that show non-overlapping peaks in the ¹H-¹⁵N SOFAST HMQC spectrum and were thus used for fitting. The average K_d value amounts to 54 μM.

Based on this titration curve we aimed to determine the binding affinity of the NTD to NGAL. Although the binding affinity of this interaction

has already been measured by ITC [Nicolas Coudevylle, unpublished data] yielding a K_d value of 15 μM . This result was inconsistent with fluorescence measurements conducted by our cooperation partner. Hence we tried to fit the intensities of our titration experiment, using QtiPlot with equation 4.1. Thus we obtained an average K_d value of 54 μM (see figure 4.3), which is consistent with our previous ITC measurement.

4.1.2 Observing Chemical Exchange in NGAL bound NTD

CPMG experiments can be used to elucidate exchange between two or more chemical environments in the micro- to millisecond time scale. Thus it can be used to investigate low populated or excited states that are invisible to standard NMR methods.⁴⁴

Based on our previous observation, that NTD ^1H - ^{15}N SOFAST HMQC signal intensity is reduced upon addition of NGAL, we assumed that this behaviour was due to chemical exchange and exploited this method to study the invisible bound state of NTD and NGAL. Analysis of the relaxation dispersion curves of residues in region 3 of the NTD with 0.05, 0.1 and 0.15 equivalents of NGAL confirmed chemical exchange in the intermediate exchange regime. In figure 4.4 relaxation dispersion curves for two residues in region 3 are shown. The experiments were conducted at two fields (500 and 800 MHz) at a protein concentration of 0.5 mM. As expected $\Delta R_{2,eff}$ increases with higher NGAL concentration. No chemical exchange was found in region 2 of NTD, supporting the hypothesis that this part is not affected by binding to NGAL. Fitting of the relaxation dispersion curves of residues in region 3 of the NTD with a two-state model yielded constant values for the parameters $\Delta\omega$ and k_{ex} , but contradictory results for the populations p_A and p_B . Assuming a K_d of 15 μM addition of 0.1 equivalents of NGAL should render 10 % bound form. In contrast CPMG measurements predicted 1 % in the bound state [Nicolas Coudevylle, unpublished data].

Consequently a two-state model is not sufficient to describe the binding of NTD and NGAL. The absence of chemical exchange in region 2 and the discrepancy of the theoretical and observed populations lead us to the conclusion, that the measured chemical exchange is not between a free- and a bound-state, but between two bound-states, and thus must be described by a higher order model. For now we assume a three-state model, though a higher N-state model would also be plausible.

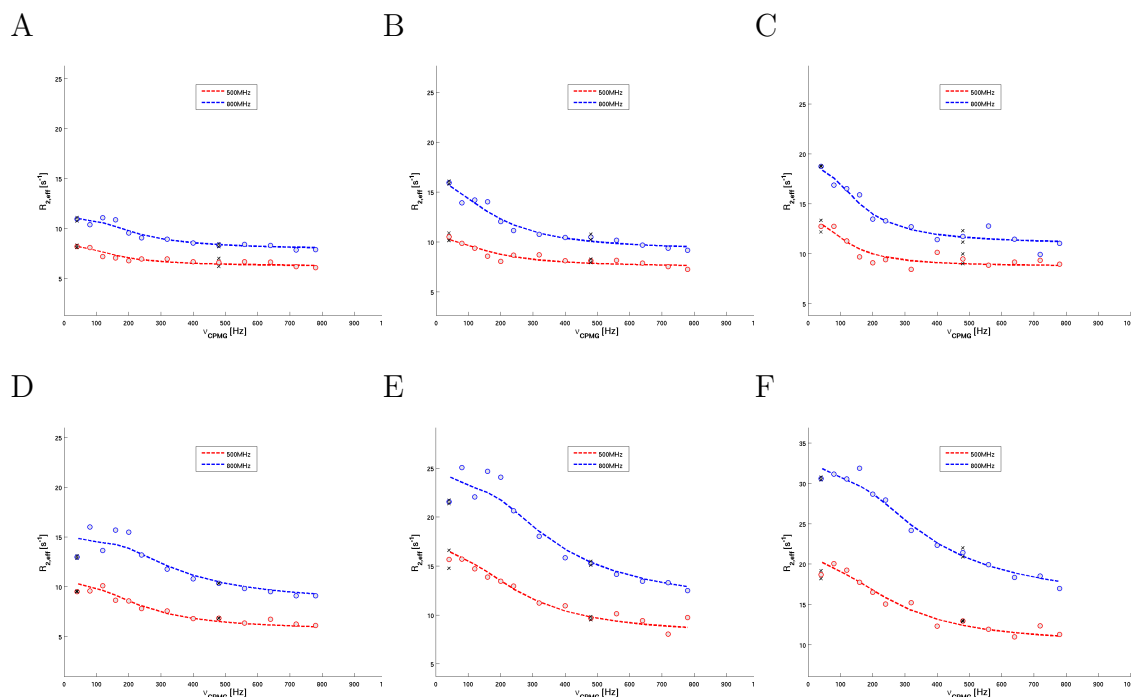


Figure 4.4: **CPMG of NTD bound to NGAL - region 3**

Relaxation dispersion curves of two residues situated in the binding region of NTD (81,100). CPMG measurements were conducted at 500 and 800 MHz. CPMG data was evaluated using rdNMR.

A: Residue 81 with 0.05 equ NGAL. **B:** Residue 81 with 0.1 equ NGAL. **C:** Residue 81 with 0.15 equ NGAL. **D:** Residue 100 with 0.05 equ NGAL. **E:** Residue 100 with 0.1 equ NGAL. **F:** Residue 100 with 0.15 equ NGAL.

Moreover, we observed a reduction in signal intensity in region 2 of the NTD upon addition of higher NGAL amounts (> 0.4 equ). To exclude a mere dilution effect, we performed a reference titration series using the same volume of buffer instead of NGAL. In this experiment no signal reduction could be observed. Thus we asked ourselves if the observed signal reduction is due to an involvement in the binding process of region 2 at higher NGAL concentrations. To address this question we performed another CPMG experiment at 600 and 800 MHz and higher NGAL concentrations. Figure 4.5 shows the relaxation dispersion curves of two residues in region 2 with 0.6 and 1.2 equivalents of NGAL. Here no exchange can be detected, resulting in flat relaxation dispersion curves.

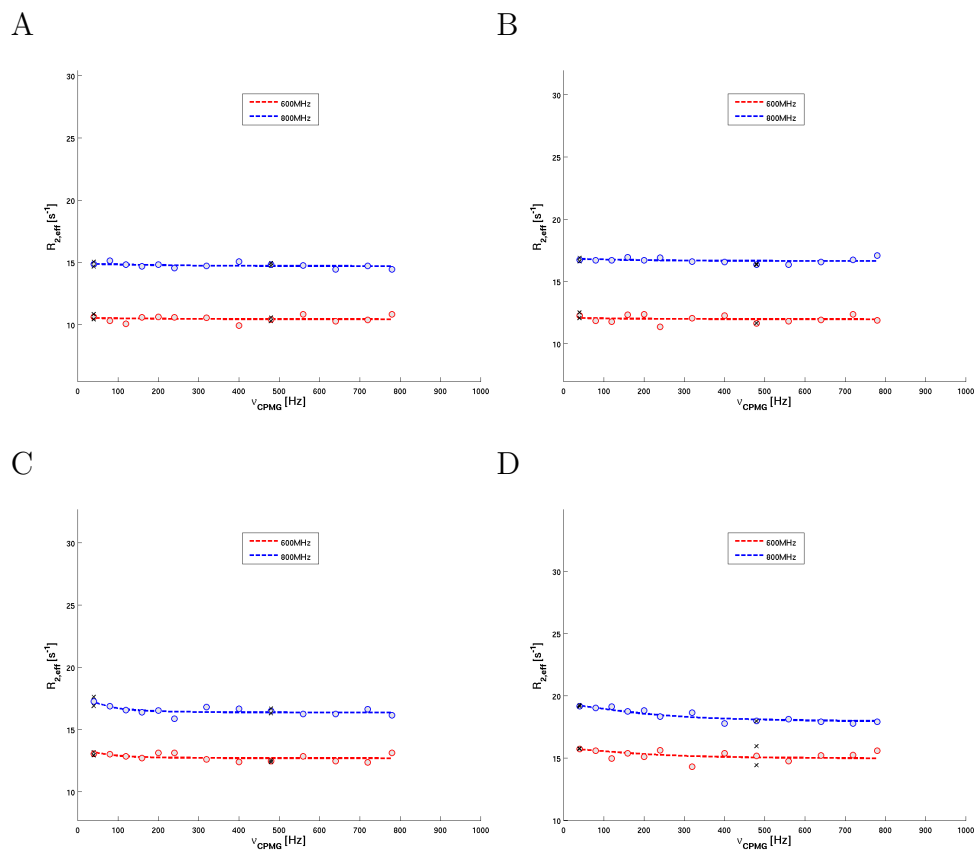


Figure 4.5: **CPMG of NTD bound to NGAL - region 2**

Relaxation dispersion curves of two residues situated in the non-binding region of NTD (36,51). CPMG measurements were conducted at 600 and 800 MHz. CPMG data was evaluated using rdNMR.

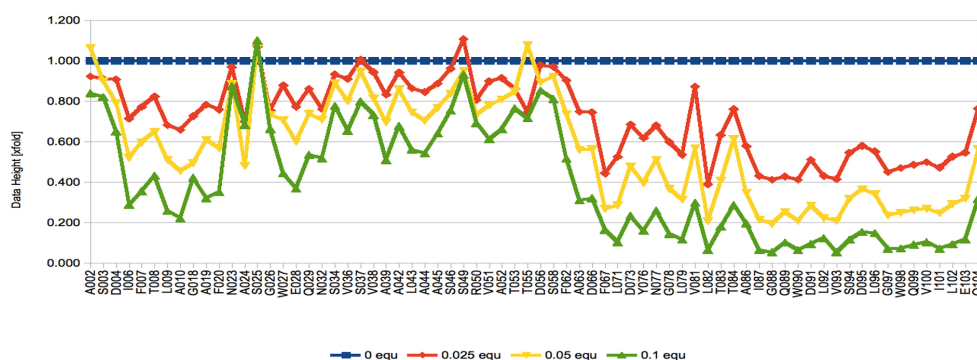
A: Residue 36 with 0.6 equ NGAL. **B:** Residue 36 with 1.2 equ NGAL. **C:** Residue 51 with 0.6 equ NGAL. **D:** Residue 51 with 1.2 equ NGAL.

4.1.3 Investigating the Bound State with Paramagnetic Tags

Paramagnetic tags are widely used in macromolecular NMR to obtain informations about dynamics, long-range distance interactions or bound states. A broad selection of paramagnetic tags is available and can be chosen according to application. There are two major groups of paramagnetic tags: compounds that carry a nitroxide spin radical, like MTSL, or lanthanide binding tags, like the LBP.¹¹

Our aim was to investigate the bound state of NTD and NGAL more closely with the introduction of paramagnetic tags. This would enable the exploitation of PRE, PCS and RDC effects.

A



B

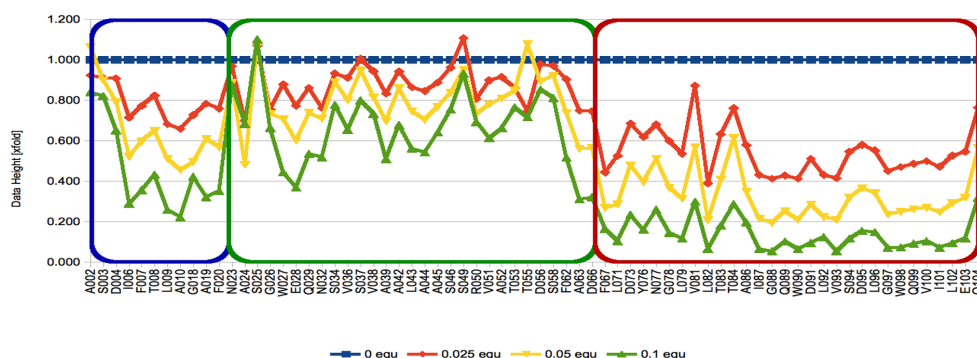


Figure 4.6: **Titration series of MTSL-NGAL to ^{15}N -labelled NTD**
A: ^1H - ^{15}N SOFAST HMQCs were recorded for NTD with increasing amounts of unlabelled MTSL-NGAL (concentration see figure legend). **B:** Differently affected regions of NTD upon binding of MTSL-NGAL (see also figure 4.2).

MTSL

MTSL is a nitroxide spin label and can be used to perform PRE measurements.¹¹ In this experiment we attached MTSL to NGAL and conducted a titration series to NTD as described in the previous section. The result of this experiment is depicted in figure 4.6. In agreement with the first titration series, three differently affected regions can be discriminated. The only difference being, that the signal reduction is much faster in the MTSL-NGAL titration due to the relaxation properties of MTSL. If we compare the signal intensities upon addition of 0.1 equivalents of NGAL or MTSL-NGAL, we can see that the signal intensity in the MTSL-NGAL titration is already down to 10 %, while the signal intensity in the normal NGAL titration is only down to about 50 %. Furthermore, if we compare the three regions inside the NTD we can see that the signal reduction is more pronounced in region 1 and 3 compared to region 2. This allows the assumption, that regions 1 and 3 are closer to the MTSL label and hence closer to the bound NGAL.

Lanthanide Binding Peptide

The lanthanide binding peptide we used in our experiments is a 17 amino acid sequence (GYIDTNNDGWIEGDELY), that selectively binds to lanthanides and can be engineered into known loop-regions of the protein structure.³ Due to their anisotropic g-tensor LBPs can be used as labels for PRE, PCS and RDC measurements.¹¹

The first two constructs we used had the LBP sequence genetically integrated into a loop at position 60 or 85 in the NGAL structure (figure 4.7 A and B). Expression and purification of the constructs NGAL60 and NGAL85 was conducted, as described for the wild type version, without any difficulties and a good yield. Loading of lanthanide ions Tb^{3+} and Lu^{3+} resulted in severe precipitation of the protein. After removing the precipitate by centrifugation and concentration of the remaining protein solution, a sufficient concentration could be reached for NMR measurements. To test the quality of the constructs we recorded ^1H - ^{15}N HSQCs for the free, Lu^{3+} - and Tb^{3+} loaded NGAL versions. However the quality of the spectra was unsatisfying and could not even be improved by addition of enterobactin, which would provide higher stability. Thus insertion of the LBP sequence into the loop regions disrupts the tertiary structure of NGAL too severely and is hence not suitable for NTD-NGAL titration experiments.

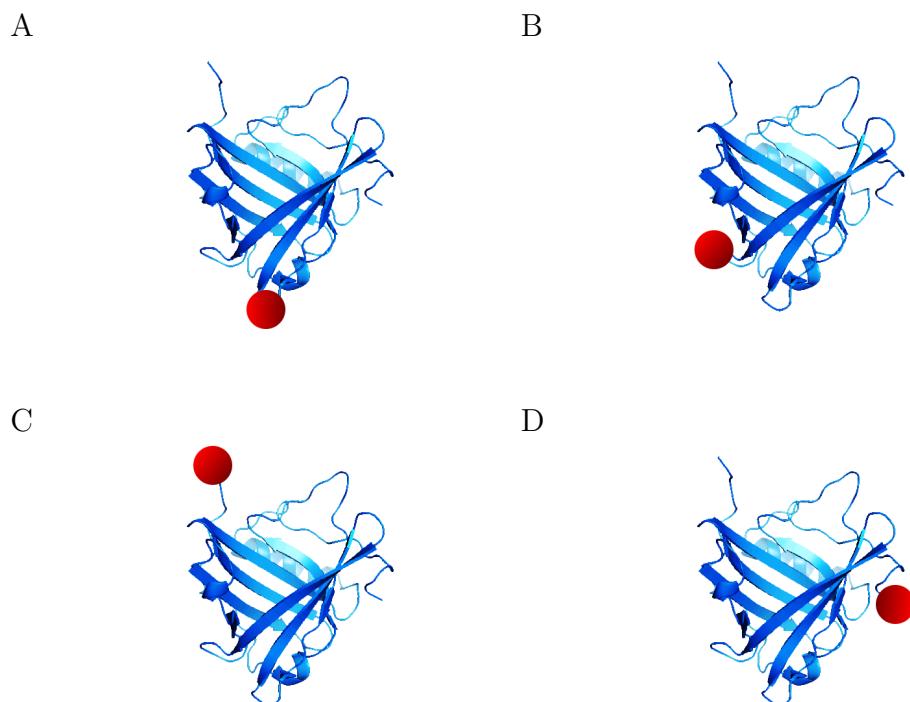


Figure 4.7: **LBP-NGAL constructs**

Four NGAL mutants each carrying a LBP at different positions. A and B contain the LBP in a loop, C and D at the N- and C-terminus of the protein.

A: LBP inserted at residue 60. **B:** LBP inserted at residue 85. **C:** LBP positioned at the N-terminus. **D:** LBP at the C-terminus.

The next step was to design NGAL constructs carrying the LBP sequence, without interfering with the tertiary structure of the protein. To be on the safe side, we decided for two constructs that would contain the LBP tag at the N- and C-terminus respectively. In the C-terminal mutant (NGAL3) the LBP sequence was positioned at residue 3, in the N-terminal mutant (NGALstop) before the stop codon (figure 4.7 C and D). Expression and purification of the two constructs was again unproblematic. Moreover, the loading of the proteins with lanthanide ions resulted in negligible precipitation and the recorded test ^1H - ^{15}N HSQC spectra were nicely resolved (see figure 4.8, displayed spectra show enterobactin bound NGAL). Consequently we were able to perform a titration series with these two NGAL constructs.

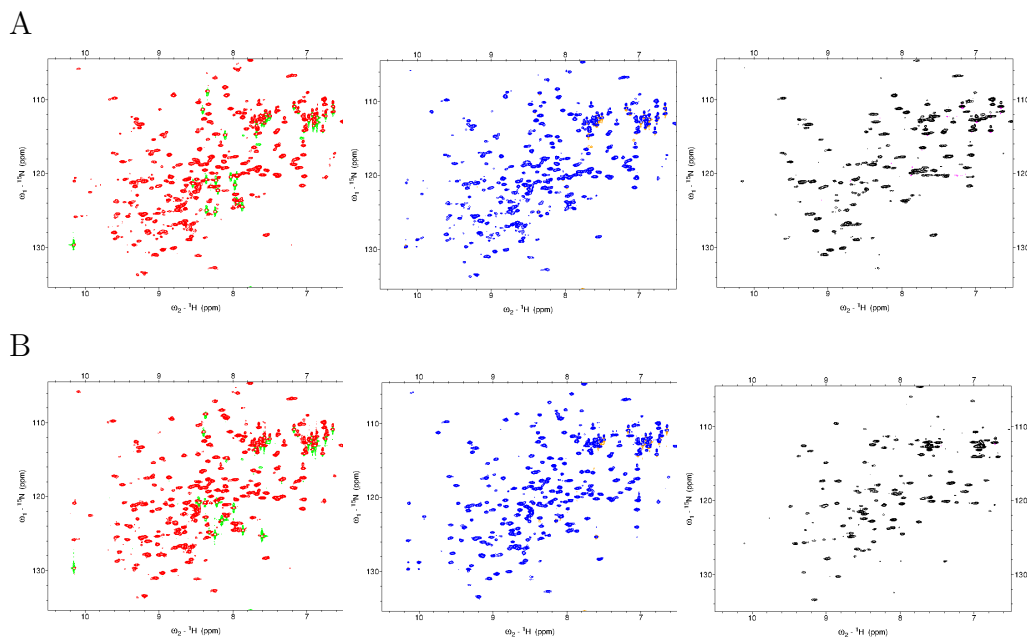


Figure 4.8: ^1H - ^{15}N HSQCs of lanthanide loaded LBP-NGAL

^1H - ^{15}N HSQC spectra of NGAL3 and NGALstop without lanthanide (red), with 1.1 equ of LuCl_3 (blue) and 1.1 equ of TbCl_3 (black). **A:** NGAL3. **B:** NGALstop.

The titration of lanthanide loaded NGALstop to NTD was conducted as described for the wild type. Two titration series, one with Lu^{3+} and one with Tb^{3+} loaded NGAL, were performed (see figure 4.9 for the results). Each diagram shows residues from one of the three NTD regions for the Lu^{3+} and the Tb^{3+} titration. Due to the diamagnetic nature of Lu^{3+} , the Lu^{3+} titration series was used as a reference. Indeed the Lu^{3+} titration series is in good agreement with the wild type titration series, concerning the affected regions as well as the rate of signal intensity decrease, implying that lanthanide loaded NGALstop actually binds to NTD. The decrease in signal intensity is again most pronounced in region 3 (figure 4.9C), while region 2 and 1 are barely affected (figure 4.9A and B). Nevertheless titration of Tb^{3+} loaded NGALstop to NTD resulted in a rapid overall decrease in intensity, making a distinction of different binding regions and any further exploitation of information impossible (figure 4.9D, E and F).

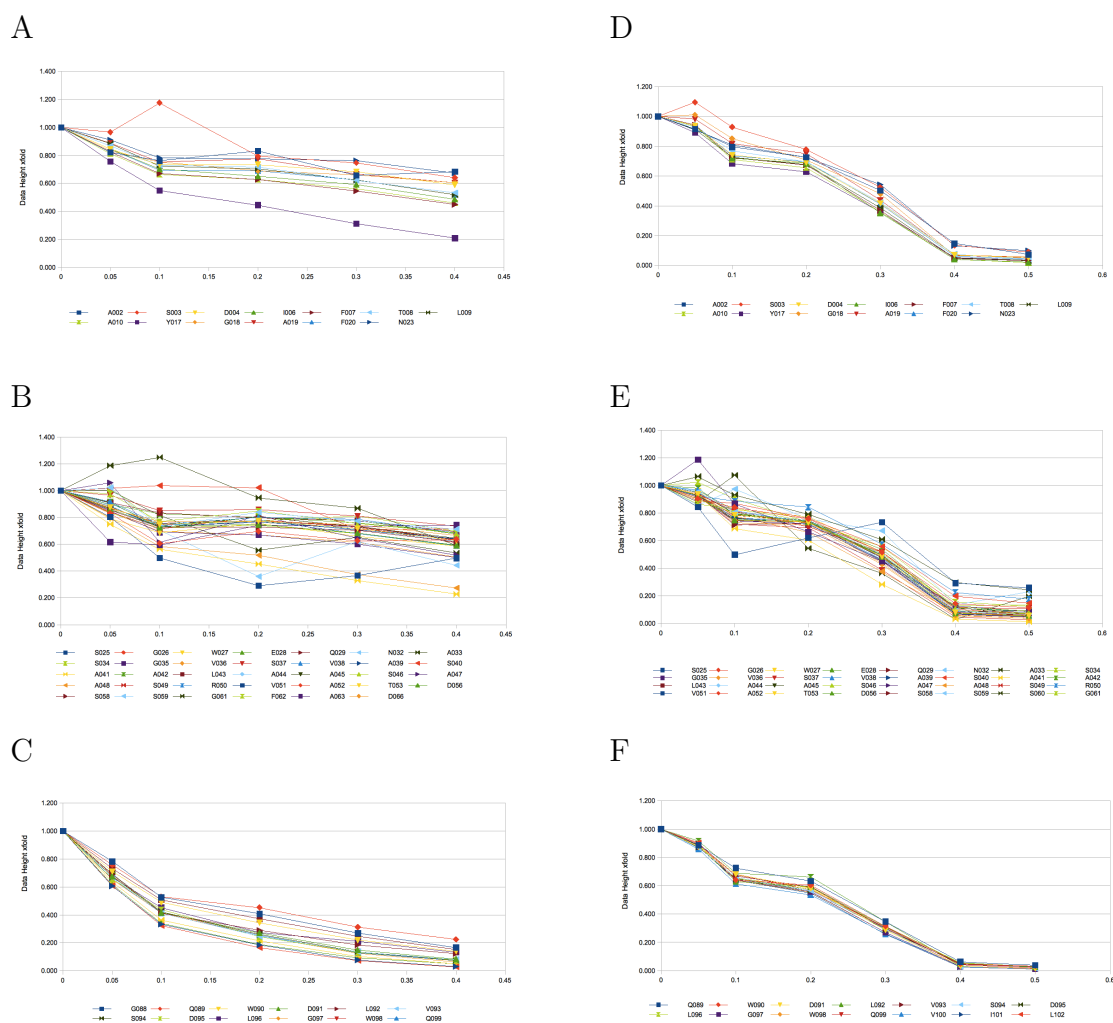


Figure 4.9: **Titration of lanthanide loaded NGALstop to NTD**

Lu^{3+} and Tb^{3+} loaded NGALstop was titrated to ^{15}N labelled NTD and the decrease in signal intensities, for the three NTD regions, plotted in these graphs. Y-axis: Relative signal intensity. X-axis: NGALstop concentration in equivalents. **A/D**: Residues from NTD region 1 with 0-1 equivalents of Lu^{3+} (**A**) or Tb^{3+} (**D**) loaded NGALstop. **B/E**: Residues from NTD region 2 with 0-1 equivalents of Lu^{3+} (**B**) or Tb^{3+} (**E**) loaded NGALstop. **C/F**: Residues from NTD region 3 with 0-1 equivalents of Lu^{3+} (**C**) or Tb^{3+} (**F**) loaded NGALstop.

In summary we show here that NGAL is binding to the N-terminal domain of 24p3R and that certain regions of the NTD are differently affected, possibly because of their varying closeness to NGAL. This could be corroborated by NMR measurements exploiting the presence of paramagnetic tags, namely MTSL and LBP. Additionally we employed the CPMG method to learn more about the invisible bound state of NGAL and the receptor. Furthermore, we demonstrated that the binding process cannot be described by a two-state model, since exchange is most probably not going on between a free- and a bound-state, but rather between two bound-states. Therefore the process must be described by at least a three- or N-state model. Moreover, we determined the binding affinity of NGAL for NTD in two different ways, on one hand by ITC, on the other hand by fitting the titration data with a non-cooperative binding curve, obtaining consistent values for both.

4.2 ^{19}F Ligand Binding

NGAL plays a crucial role in various physiological and pathological processes and thus is an interesting target for FBDD. Fluorination of ligands offers several advantages for a drugs pharmacodynamic and pharmacokinetic properties. Among them ameliorated permeability, metabolic stability, increased lipophilicity and improved binding affinities.⁵⁸

Our aim was to confirm the binding of putative ligands, that were found in a fluorine library screen. To this end we performed a ^1H - ^{15}N HSQC titration series for all putative binders and an ITC measurement for one selected compound to quantify the binding affinity.

4.2.1 Analysing Ligand Binding via ^1H - ^{15}N HSQC Chemical Shifts

We recorded ^1H - ^{15}N HSQC spectra of ^{15}N labelled NGAL with increasing ligand concentration. As a control experiment we conducted a titration series with DMSO, the solvent of the ligand samples. Indeed all putative ligands caused extensive chemical shift changes in the NGAL ^1H - ^{15}N HSQC spectrum, indicating interaction between NGAL and the ligand. No chemical shift changes were observed for the control sample. The ligand code, structure of the ligand, an overlay of the ^1H - ^{15}N HSQC spectra with increasing amounts of ligand and the affected area in the NGAL structure are shown in figure 4.11. Interestingly all ligands caused chemical shift changes in the β -barrel, suggesting binding of the ligands inside the calyx. Apart from E2 all ligands caused chemical shift changes throughout the barrel. One reason for the deviating binding behaviour of E2 might be the small size of the ligand compared to the remaining ligands.

4.2.2 Determination of Binding Affinities via ITC

We determined the binding affinity of the NGAL ligand interaction on one selected example ligand. For this purpose we chose E2, because it could be a good starting point for FBDD, it was binding to NGAL in a more specific manner than the other ligands and it was commercially available. The experiment was conducted as described in the “Materials and Methods” section and the obtained result can be seen in figure 4.10. Although the binding affinity of 6.4 mM is rather low, E2 could be the basis for designing a ligand with higher binding affinity and selectivity.

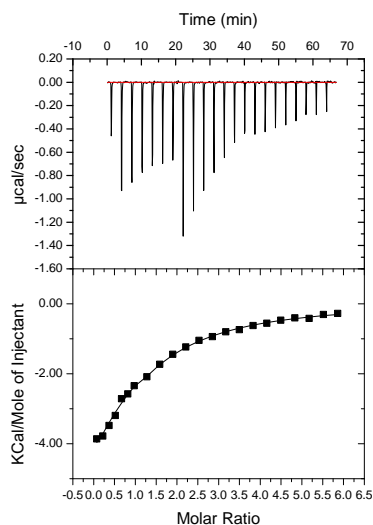


Figure 4.10: **ITC - NGAL and E2**

ITC titration of 3 mM E2 in 0.1 mM NGAL. The volume of the first injection was 0.5 μL , the preceding six injections 1 μL and the last 15 injections were 2 μL each.

In short we could demonstrate that all putative fluorine ligands are actually binding to NGAL and that the affected area is deviating depending on the ligand. Furthermore, we determined the binding affinity for the ligand E2, providing a starting point for optimization of the ligand.

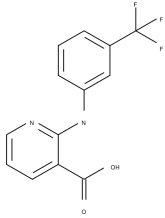
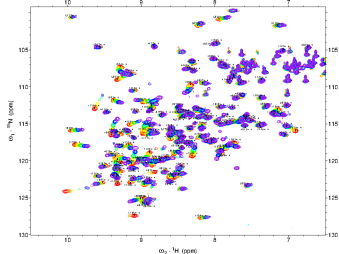

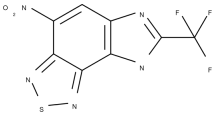
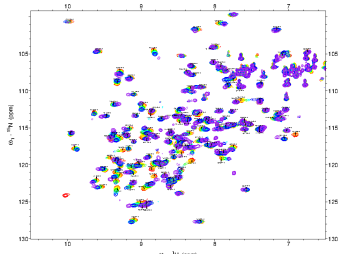
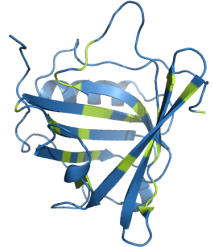
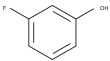
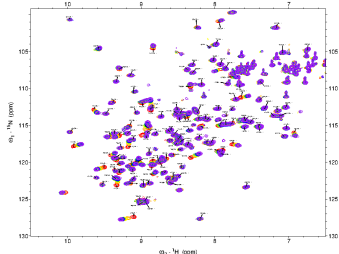

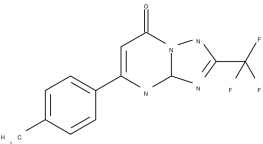
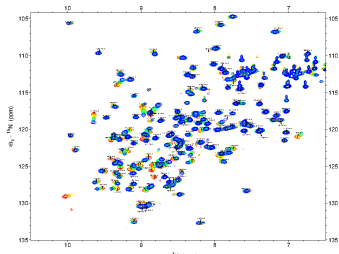
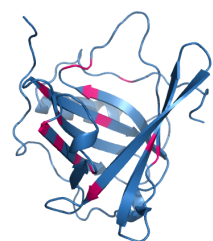
compounds	Structure	HSQC	Affected Area
A2			
A10			
E2			
E5			

Figure 4.11: ^{19}F ligands for NGAL

From left to right: ligand code, ligand structure, ^1H - ^{15}N HSQC titration of NGAL with the corresponding ligand and affected area in the structure of NGAL. For the HSQC titration 0, 0.1, 0.25, 0.5, 0.75 and 1 mM ligand were added to 0.5 mM NGAL. The picture shown is an overlay of all HSQC spectra. The affected peaks were identified via the HSQC chemical shift change and their position marked on the NGAL structure.

4.3 Selective Labelling of Val, Phe and Tyr in Proteins

Selective isotope labelling of amino acids is a useful tool for the study of large proteins, to avoid signal overlap or to obtain specific information on interactions in the binding site. Of particular interest are amino acids that are frequently found in binding sites or interaction surfaces, which is the case for isoleucines, leucines, valines and aromatic amino acids.³²



Figure 4.12: **Precursors for selective valine labelling**

A: ^{13}C 2-ketoisovalerate, **B:** 2-ketoisocaproate. Dissolved and sterile filtered precursors were added one hour prior to induction to the growth medium.

4.3.1 Independent Labelling of Valine Residues in GB1

Selective labelling of valine residues, without labelling of leucine residues, was up until now not feasible. Recently a few labelling strategies, addressing this problem, have been developed.^{31,33,36} The following is one of them. As a model protein we used GB1 instead of NGAL, avoiding elaborate purification procedures. Selective labelling of valines was achieved by adding ^{13}C labelled 2-ketoisovalerate (see figure 4.12A) and unlabelled 2-ketoisocaproate (see figure 4.12B) to the ^{15}N supplemented growth medium (for a detailed description refer to chapter 3). 2-ketoisovalerate is a precursor for valines, that can be metabolised into 2-ketoisocaproate, which subsequently is metabolised into leucines. The rational behind this labelling technique is a feedback inhibition of the valine/leucine biosynthetic pathway by 2-ketoisocaproate and hence no metabolism of 2-ketoisovalerate into leucines.

Protein quality was evaluated recording a ^1H - ^{15}N HSQC, while correct incorporation of the precursor was verified in a ^1H - ^{13}C HSQC experiment. The optimum 2-ketoisocaproate concentration, for exclusive labelling of valines, was determined in a series of experiments with increasing precursor concentration (0, 50, 100 and 200 mg/L). Figure 4.13 shows the outcome of this experiment. We can see that addition of 100 mg/L 2-ketoisovalerate and 100 mg/L 2-ketoisocaproate suppresses leucine labelling and gives the

4.3. SELECTIVE LABELLING OF VAL, PHE AND TYR IN PROTEINS73

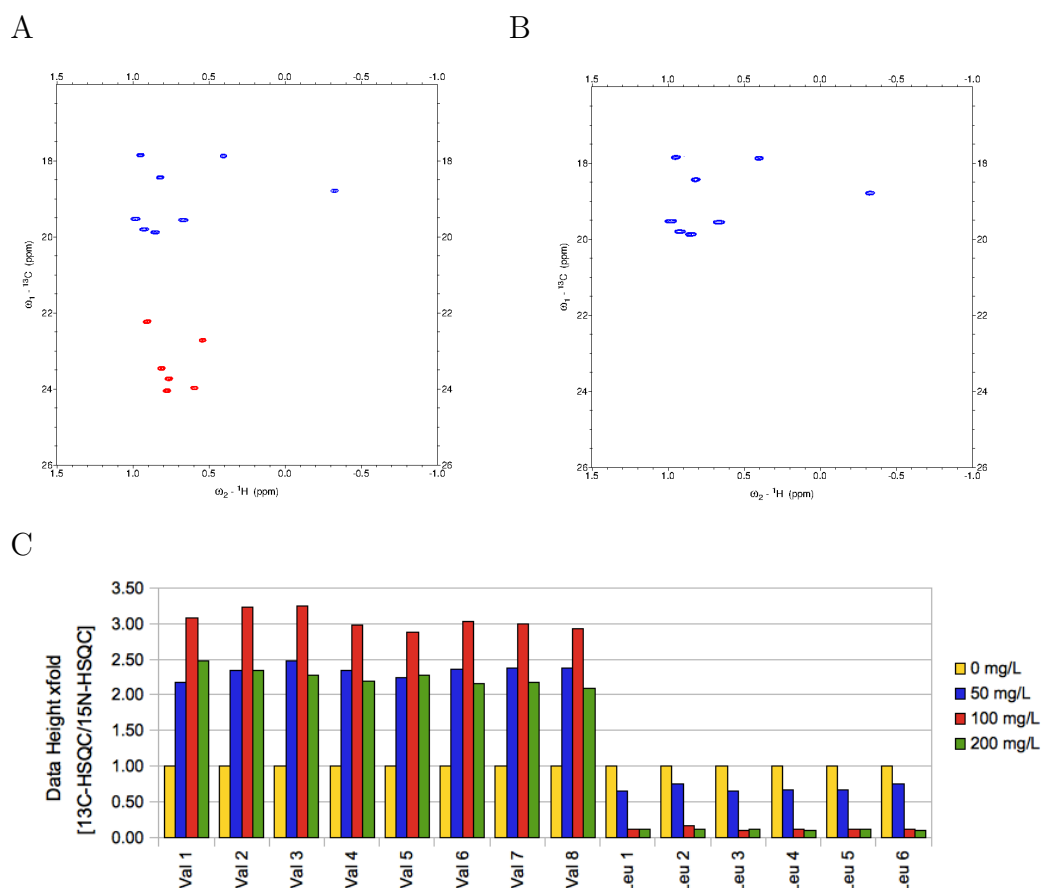


Figure 4.13: **Selective valine labelling in GB1**

A: ^1H - ^{13}C HSQC of GB1 with 100 mg/L ^{13}C ketoisovalerate. **B:** 100 mg/L ^{13}C ketoisovalerate plus 100 mg/L ketoisocaproate. Valines are shown in blue, Leucines in red. **C:** Relative ^1H - ^{13}C HSQC intensities of valine and leucine methyl groups in GB1. All samples were expressed in the presence of 100 mg/L ^{13}C 2-ketoisovalerate. In addition 0, 50, 100 and 200 mg/L 2-ketoisocaproate were added.

highest signal intensity for valines (see figure 4.13C). The same result is illustrated in the ^1H - ^{13}C HSQC spectra. Figure 4.13A shows the spectrum upon addition of 100 mg/L 2-ketoisovalerate, where valines (blue) as well as leucines (red) are ^{13}C labelled. Figure 4.13B in contrast shows the spectrum resulting in the addition of 100 mg/L ^{13}C labelled 2-ketoisovalerate and 100 mg/L 2-ketoisocaproate. In this spectrum only ^{13}C labelled valines can be observed. Surprisingly the signal intensity of valines increases upon addition of 2-ketoisocaproate until 100 mg/L, without adding more ^{13}C labelled 2-ketoisovalerate, and decreases at higher 2-ketoisocaproate concentrations. This might be explained by a feedback mechanism in the valine/leucine biosynthetic pathway.

4.3.2 Selective Labelling of Aromatic Residues in GB1

Selective labelling of aromatic amino acids is underrepresented in the literature, but offers great opportunities for NMR studies, due to their specific ability to interact with various binding partners, like proteins, nucleic acids or carbohydrates via cation- π , polar- π , stacking or hydrophobic interactions.⁴⁰

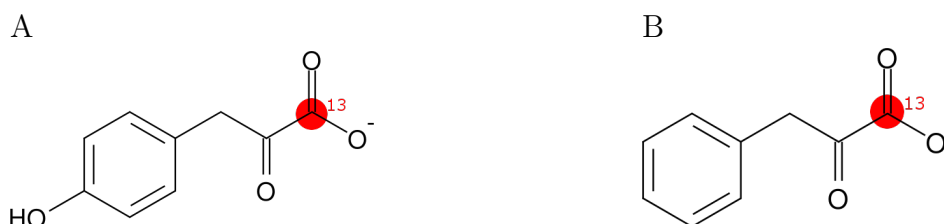


Figure 4.14: **Precursors for labelling of aromatic amino acids**

A: ^{13}C p-hydroxy- α -ketophenylpyruvate, **B:** ^{13}C α -ketophenylpyruvate. Precursors were added one hour prior to induction to the growth medium.

The selective labelling approach for tyrosines and phenylalanines is in principal the same as for valines, differing only in the position of the ^{13}C label, that is situated in the backbone and not in the side chain. The precursors are shown in figure 4.14. p-hydroxy- α -ketophenylpyruvate was used for the selective labelling of tyrosines and α -ketophenylpyruvate for phenylalanines. According to the position of the label in the protein backbone, we recorded HNCO experiments to verify the correct incorporation of the precursors. The outcome of the selective labelling experiment for tyrosines and phenylalanines can be seen in figure 4.15C/D. For the selective labelling of tyrosines 100 mg/L ^{13}C p-hydroxy- α -ketophenylpyruvate are sufficient, a further increase in precursor concentration even exceeds the signal intensity obtained from homogeneous ^{13}C labelling. The same phenomenon could be observed for

4.3. SELECTIVE LABELLING OF VAL, PHE AND TYR IN PROTEINS75

phenylalanine labelling. Figure 4.15A shows the HNC0 spectra of the three tyrosine residues of tyrosine labelled GB1, in the graph only two tyrosine residues are displayed due to signal overlap in the HNC0 spectrum of the fully ^{13}C labelled GB1 sample. Figure 4.15B shows the HNC0 spectrum of the two phenylalanine residues in GB1 and figure 4.15D the relative signal intensities of the α -ketophenylpyruvate concentration series for one of the phenylalanine residues. Again the second phenylalanine overlaps in the ^{13}C labelled sample.

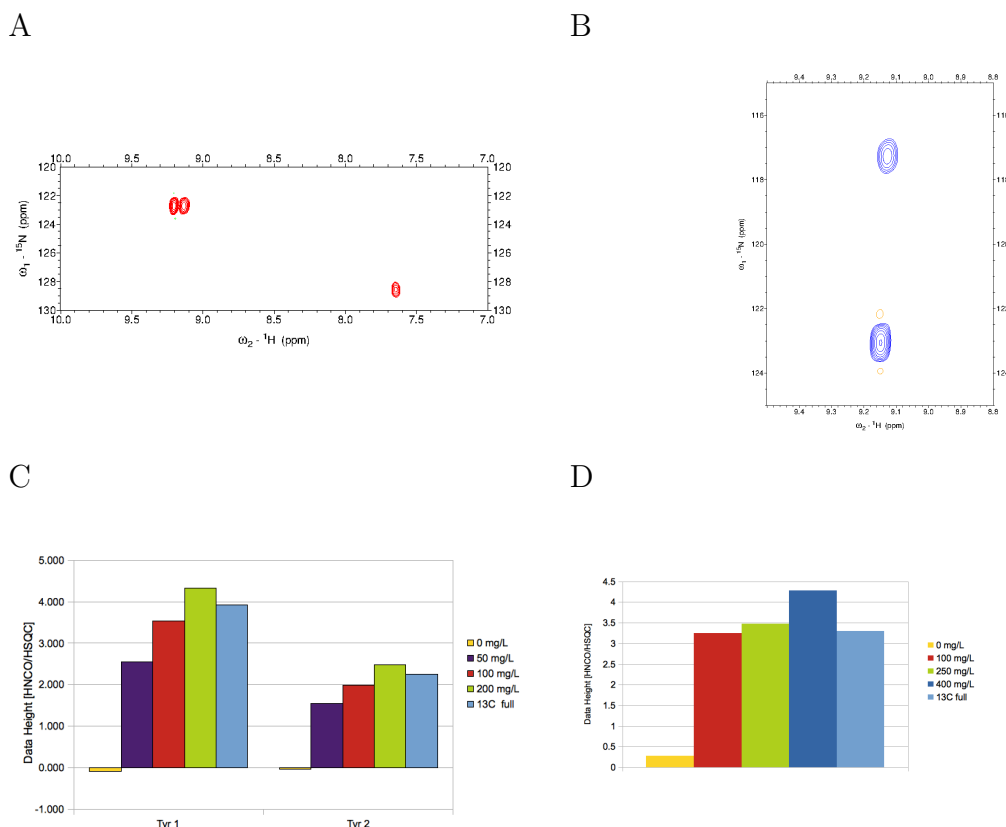


Figure 4.15: **Selective labelling of phenylalanine and tyrosine in GB1**

A: HNC0 of selectively labelled tyrosines in GB1 (100 mg/L ^{13}C p-hydroxy- α -ketophenylpyruvate). **B:** HNC0 of selectively labelled phenylalanines in GB1 (100 mg/L ^{13}C α -ketophenylpyruvate). **C:** Relative HNC0/HSQC intensities of tyrosine residues in the presence of 0, 50, 100 and 200 mg/L ^{13}C p-hydroxy- α -ketophenylpyruvate compared to tyrosine residues in full ^{13}C labelled GB1. **D:** Relative HNC0/HSQC intensities of a phenylalanine residue in the presence of 0, 100, 250 and 400 mg/L ^{13}C α -ketophenylpyruvate compared to a phenylalanine residue in full ^{13}C labelled GB1.

Thus we demonstrated a new easily applicable technique to selectively label valines, phenylalanines and tyrosines in proteins. Furthermore, we can

report here that no scrambling was observed, using these strategies. Neither between valines and leucines, nor the aromatic amino acids, even without addition of glyphosate.

Chapter 5

Discussion

5.1 Discussion

The siderocalin NGAL is an intriguing research topic on account of its diverse physiological functions, that range from iron metabolism to antibacterial defence and gene regulation. Only recently the location of 24p3R, the cellular receptor of NGAL, has been determined in the kidney, assigning the NGAL-24p3R complex a crucial role in renal filtration via receptor mediated endocytosis.²⁹ In addition to this vital physiological functions, NGAL has been associated with pathological processes like cancer progression or inflammation.¹⁷ Acknowledging these findings the investigation of the 24p3R-NGAL interaction and the search for new ligands of NGAL is of particular interest to understand the functions of NGAL.

Accordingly this thesis is divided into three parts. The first two address the aforementioned questions and the third is concerned with the development of new selective labelling techniques that simplify the study of protein structure and function by NMR.

5.1.1 Investigating the Interaction of NGAL and its Cellular Receptor 24p3R

The interaction of NGAL with its cellular receptor 24p3R and subsequent endocytosis is crucial for the physiological function of NGAL, for it allows iron metabolism, subsequent iron dependent gene regulation and inflammatory effects, renal filtration and host defence against bacteria through sequestration of siderophores.⁵⁰

In this study we used the N-terminal domain of 24p3R offering the following advantages. 24p3R is a large membrane protein and thus not applicable for conventional solution state NMR methods. However, the intrinsically disordered, approximately 100 residue, N-terminal domain of 24p3R is soluble and suitable for solution state NMR techniques. A quadruple cysteine mutant of this N-terminal domain even provides well resolved NMR spectra and thus was used for this work. Most importantly it was shown that this N-terminal domain binds to NGAL with a K_d of about 15 μ M.

Our first aim was to characterize the binding mode of the N-terminal domain to NGAL. A titration of NGAL to NTD provided a general overview of the binding site. Apparently there are three regions in the NTD that are differently affected by binding to NGAL. The middle area of the protein does not seem to be affected by the bound NGAL, while the C-terminal part is most affected. The same profile could be observed for a titration series performed with MTSL tagged NGAL except, that the N-terminal and C-terminal part, which were already more affected in the NGAL titration,

are more severely affected compared to the middle part. This suggests a proximity of these two regions to the MTSL label and thus the bound NGAL. Therefore we could further substantiate our assumption, that region 1 and 3 are the main sites involved in the binding event.

Furthermore, we extracted the binding affinity of NGAL to NTD fitting the intensities of a titration experiment with equation 4.1 using QtiPlot. Indeed the derived value is in good agreement with a previous ITC measurement.

In a subsequent experiment we attempted to study the bound state of NGAL and NTD more closely. Since binding of NGAL to NTD results in severe peak broadening of the affected residues we were in need of a method that is able to detect this invisible bound state. CPMG, normally used to analyse low populated or excited states, is a suitable method for this purpose. In a previous CPMG experiment varying behaviour of the three differently affected regions in NGAL bound NTD could be observed, showing chemical exchange for the putative binding site. Fitting of the CPMG data using a two-state model was not feasible, strengthening the assumption that a model with a higher order must be used. Thus leading to the conclusion, that NGAL and the NTD form a disordered complex. Additionally we carried out CPMG experiments at higher NGAL concentrations to prove, that the reduction of signal intensity at this concentrations is not due to chemical exchange in the micro- to millisecond range. The presence of a mere dilution effect was excluded by a control experiment. Indeed no chemical exchange could be measured for region 2 at higher NGAL concentrations. One possible explanation for the signal decrease is an increase in molecular weight upon binding of NGAL to the NTD.

A third approach to learn more about the NGAL NTD interaction was the introduction of a paramagnetic label in form of a LBP into the protein structure of NGAL. This would enable PRE, PCS and RDC experiments of NGAL binding to NTD and provide long-range distance information. The first two NGAL constructs carried the LBP inside two loop regions, as was described by Barthelmes et al.³ Unfortunately presence of these long flexible LBPs had detrimental effects on the structure of NGAL and thus this approach was dismissed. The following two constructs were designed to carry the LBP at the N- or C-terminus of NGAL. This should avoid destruction of the tertiary structure. Indeed both constructs provided well resolved ^1H - ^{15}N HSQC spectra, as well for lanthanide bound and free forms, and therefore were used for further experiments. To observe PCSs we conducted the same titration experiment for NTD and Lu^{3+} or Tb^{3+} loaded NGALstop, as we did for NGAL. The diamagnetic Lu^{3+} should serve as a control and the paramagnetic lanthanide Tb^{3+} should induce the PCSs. The NTD spectra for

the Lu^{3+} loaded NGALstop titration were in every respect comparable to the NGAL titration and even yielded the same rate of intensity decrease. Nevertheless, Tb^{3+} loaded NGALstop enhanced the relaxation too much, so that no PCSs effect could be exploited. Even a discrimination between the three regions, normally present in NGAL bound NTD, was impossible. A plausible explanation would be a rather compact structure of the NTD, such that the paramagnetic label would be close to all residues in the NTD. To solve this problem application of a different paramagnetic lanthanide ion with a shorter radius would be convenient.

5.1.2 Finding New Ligands for NGAL

Owing to its association with inflammatory diseases and cancer the discovery of new ligands for NGAL could be an important step for the treatment and diagnosis of these diseases. One way to find new lead compounds is fragment based drug design. FBDD is a popular method in the pharmaceutical industry to identify weak binders and turn them into high affinity ligands. In contrast to high throughput screenings, FBDD is not aimed at finding binders for a biological target with nM affinity, but settles for finding binders in the mM range. These weak binders can then be combined and optimized to obtain specific high affinity ligands.⁵⁹

Fluorination of compounds is another popular method in the pharmaceutical industry to improve parameters like metabolic stability, permeability and binding affinity.⁵⁸ Thus we decided to test fluorinated putative binders of NGAL, that were preselected from a fluorine fragment library via a ^{19}F NMR screening.

^1H - ^{15}N HSQC titration of these fragments verified the binding of NGAL to the ligands and even displayed the binding site via chemical shift changes. The identified area was located inside the β -barrel for all ligands. Interestingly only one ligand showed a deviating binding mode. Whereas all other ligands affected the chemical shifts throughout the β -barrel, E2 only disturbed the chemical shifts in one half of the calyx. This might be due to the smaller size of fluorophenol (E2) compared to the other ligands. Special care had to be taken in preparing the ligand stock solutions for the titration experiments, since a small deviation in the pH of the sample lead to extensive chemical shift changes in the NGAL HSQC spectrum. A reason for this behaviour might be the pH dependent release of the siderophore inside the endosome. To determine the binding affinity of the ligands for NGAL we selected one compound that was commercially available and performed an ITC experiment. The resulting K_d of 6.4 mM indicates a good suitability for a further FBDD approach. Additionally ^{19}F NMR experiments could eluci-

date the binding mode of the different ligands to NGAL and might be used in the FBDD optimization process.

5.1.3 Selective Labelling of Val, Tyr and Phe in the Model Protein GB1

Isotopic ^{13}C and ^{15}N labelling of proteins is a common strategy to enable heteronuclear NMR experiments. Similarly, selective isotope labelling of amino acids has gained a wide popularity in protein NMR, owing to its advantages for large proteins, signal overlap and the study of binding sites.⁵³ However, not all amino acids are accessible with this technique yet and we present here new methods for the selective labelling of valines, tyrosines and phenylalanines. Additionally these amino acids are of special interest due to their frequent presence in binding sites and their potential to interact via methyl groups (valine) or aromatic rings (tyrosine, phenylalanine).³⁶

Exclusive labelling of valines without labelling leucines was not feasible up until now, when several groups published strategies to selectively label valines. These approaches either exploit the possibility to stereoselectively label valines and leucines with acetolactate and inhibit leucine labelling by adding unlabelled leucine³³ or by directly adding labelled valine and unlabelled leucine to the expression medium.³⁶ Our strategy involves the well-known valine and leucine precursors 2-ketoisovalerate and 2-ketoisocaproate. While introduction of 2-ketoisovalerate to the expression medium leads to selective labelling of valines and leucines, due to subsequent metabolism into 2-ketoisocaproate and leucine, 2-ketoisocaproate solely labels leucines. To achieve exclusive labelling of valines we supplement the expression medium with ^{13}C labelled 2-ketoisovalerate and suppress metabolism of this labelled precursor by adding unlabelled 2-ketoisocaproate, resulting in pure valine labelling of the protein without scrambling. In conclusion this approach offers an elegant way to selectively label valine residues and simultaneously homogeneously ^{15}N label proteins at a moderate cost. One downside is the non-stereospecific labelling of the valine residues, which could lead to problems in very large proteins or proteins with severely overlapping NMR spectra.

The second strategy we tested was selective labelling of the aromatic amino acids tyrosine and phenylalanine. Surprisingly labelling of aromatic amino acids is not frequently found in the literature, despite the crucial role of aromatic amino acid interactions in the binding of proteins, carbohydrates and nucleic acids.⁴ Although in principle identical to the valine labelling technique the precursors differ in the location of the ^{13}C label. In ^{13}C 2-ketoisovalerate the label is positioned in the side chain methyl group while in

the tyrosine and phenylalanine precursors p-hydroxy- α -ketophenylpyruvate and α -ketophenylpyruvate the label is located in the backbone carbonyl moiety. Correct incorporation of the label was verified via a HNCO experiment and was successful for both precursors without any scrambling. Anyway, synthesis of precursors that achieve labelling of the sidechains of aromatic amino acids could prove even more useful and has already been tested with a sidechain ^{13}C labelled phenylalanine precursor in a preliminary experiment with a successful outcome.

Nevertheless, application of these labelling techniques in larger proteins than the 7.5 kDa protein GB1 and studying the interaction in a relevant biological system would be even more interesting.

5.2 Conclusion

In summary we could show that the N-terminal domain of the cellular receptor of NGAL possesses three regions that are differently affected upon binding, probably due to varying proximity to NGAL. Furthermore, we determined the binding affinity of NGAL and the N-terminal domain to be in the μM range. Additionally we found out, that the binding process must be described by at least a three- or N-state model, making the existence of a disordered complex plausible. Moreover, we were able to design two functioning LBP-NGAL constructs that bind the NTD with the same affinity as NGAL. In addition to these preliminary studies on the NTD and LBP-NGAL more experiments employing different paramagnetic lanthanides are necessary.

Likewise we demonstrated that all selected putative fluorine ligands of NGAL are actual binders and bind inside the NGAL calyx similar to all other ligands known for NGAL. Thus all of these fragments are possible starting points for creating a high affinity ligand in a FBDD approach.

Finally we employed three new selective amino acid labelling techniques for valines, tyrosines and phenylalanines, that can prove useful in facilitating assignment in large protein, avoid signal overlap and help characterize interactions in the binding site of proteins.

Chapter 6

Bibliography

Bibliography

- [1] AKTORIES, FÖRSTERMANN, HOFMANN, AND STARKE. Allgemeine und spezielle pharmakologie und toxikologie. *Elsevier 10.Auflage* (2009).
- [2] ALLEN, K. N., AND IMPERIALI, B. Lanthanide-tagged proteins: an illuminating partnership. *Current Opinion in Chemical Biology* 14, 2 (2010), 247 – 254.
- [3] BARTHELMES, K., REYNOLDS, A. M., PEISACH, E., JONKER, H. R. A., DENUNZIO, N. J., ALLEN, K. N., IMPERIALI, B., AND SCHWALBE, H. Engineering encodable lanthanide-binding tags into loop regions of proteins. *Journal of the American Chemical Society* 133, 4 (2011), 808–819.
- [4] BATURIN, S., OKON, M., AND MCINTOSH, L. Structure, dynamics, and ionization equilibria of the tyrosine residues in bacillus circulans xylanase. *Journal of Biomolecular NMR* 51, 3 (2011), 379–394.
- [5] BERTINI, I., GELIS, I., KATSAROS, N., LUCHINAT, C., AND PROVENZANI, A. Tuning the affinity for lanthanides of calcium binding proteins†. *Biochemistry* 42, 26 (2003), 8011–8021. PMID: 12834353.
- [6] BYEON, I.-J. L., LOUIS, J. M., AND GRONENBORN, A. M. A protein contortionist: Core mutations of {GB1} that induce dimerization and domain swapping. *Journal of Molecular Biology* 333, 1 (2003), 141 – 152.
- [7] CAI, L., RUBIN, J., HAN, W., VENGE, P., AND XU, S. The origin of multiple molecular forms in urine of hnl/ngal. *Clinical Journal of the American Society of Nephrology* 5, 12 (2010), 2229–2235.
- [8] CHAKRABORTY, S., KAUR, S., GUHA, S., AND BATRA, S. K. The multifaceted roles of neutrophil gelatinase associated lipocalin (ngal) in inflammation and cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1826, 1 (2012), 129 – 169.
- [9] CHEN, J. Towards the physical basis of how intrinsic disorder mediates protein function. *Archives of Biochemistry and Biophysics* 524, 2 (2012), 123 – 131.
- [10] CHENG, Y., AND PATEL, D. J. An efficient system for small protein expression and refolding. *Biochemical and Biophysical Research Communications* 317, 2 (2004), 401 – 405.
- [11] CLORE, G. M., AND IWAHARA, J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chemical Reviews* 109, 9 (2009), 4108–4139.

- [12] COUDEVYLLE, N., GEIST, L., HÖTZINGER, M., HARTL, M., KONTAXIS, G., BISTER, K., AND KONRAT, R. The v-myc-induced q83 lipocalin is a siderocalin. *Journal of Biological Chemistry* 285, 53 (2010), 41646–41652.
- [13] DALVIT, C., AND VULPETTI, A. Fluorine-protein interactions and 19f nmr isotropic chemical shifts: An empirical correlation with implications for drug design. *ChemMedChem* 6, 1 (2011), 104–114.
- [14] DALVIT, C., AND VULPETTI, A. Intermolecular and intramolecular hydrogen bonds involving fluorine atoms: Implications for recognition, selectivity, and chemical properties. *ChemMedChem* 7, 2 (2012), 262–272.
- [15] DALVIT, C., AND VULPETTI, A. Technical and practical aspects of 19f nmr-based screening: toward sensitive high-throughput screening with rapid deconvolution. *Magnetic Resonance in Chemistry* 50, 9 (2012), 592–597.
- [16] DASGUPTA, S., HU, X., KEIZERS, P., LIU, W.-M., LUCHINAT, C., NAGULAPALLI, M., OVERHAND, M., PARIGI, G., SGHERI, L., AND UBBINK, M. Narrowing the conformational space sampled by two-domain proteins with paramagnetic probes in both domains. *Journal of Biomolecular NMR* 51, 3 (2011), 253–263.
- [17] DEVIREDDY, L. R., GAZIN, C., ZHU, X., AND GREEN, M. R. A cell-surface receptor for lipocalin 24p3 selectively mediates apoptosis and iron uptake. *Cell* 123, 7 (2005), 1293 – 1305.
- [18] FLOWER, D. R., NORTH, A. C., AND SANSOM, C. E. The lipocalin protein family: structural and sequence overview. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1482, 1–2 (2000), 9 – 24.
- [19] GANS, P., HAMELIN, O., SOUNIER, R., AYALA, I., DURÁ, M., AMERO, C., NOIRCLERC-SAVOYE, M., FRANZETTI, B., PLEVIN, M., AND BOISBOUVIER, J. Stereospecific isotopic labeling of methyl groups for nmr spectroscopic studies of high-molecular-weight proteins. *Angewandte Chemie International Edition* 49, 11 (2010), 1958–1962.
- [20] GOETZ, D. H., WILLIE, S. T., ARMEN, R. S., BRATT, T., BORREGAARD, N., AND STRONG, R. K. Ligand preference inferred from the structure of neutrophil gelatinase associated lipocalin†. *Biochemistry* 39, 8 (2000), 1935–1941.
- [21] GUANHU, B., MATTHEW, C., M, H. T., KIYOSHI, M., ANDONG, D. S.-X., MELANIE, V., DAVID, W., NEAL, P., THOMAS, L., RITWIJ, K., XIANGPO, L., BELINDA, L., AVTANDIL, K., ADAM, R., CARLOS, J. J., M, S.-O. K., W, L. D., K, R. K. N. R., AND JONATHAN, B. Iron traffics in circulation bound to a siderocalin ngal-catechol complex. *Nat Chem Biol* 6, 8 (2010), 602–609.
- [22] HVIDBERG, V., JACOBSEN, C., STRONG, R. K., COWLAND, J. B., MOESTRUP, S. K., AND BORREGAARD, N. The endocytic receptor megalin binds the iron transporting neutrophil-gelatinase-associated lipocalin with high affinity and mediates its cellular uptake. *{FEBS} Letters* 579, 3 (2005), 773 – 777.
- [23] JAMES, K. Understanding nmr spectroscopy. *Wiley* 2 (2010).
- [24] JEREMY CRAVEN, C., AL-OWAIS, M., AND PARKER, M. A systematic analysis of backbone amide assignments achieved via combinatorial selective labelling of amino acids. *Journal of Biomolecular NMR* 38, 2 (2007), 151–159.

- [25] KIEFHABER, T., BACHMANN, A., AND JENSEN, K. S. Dynamics and mechanisms of coupled protein folding and binding reactions. *Current Opinion in Structural Biology* 22, 1 (2012), 21 – 29.
- [26] KJELDSSEN, L., COWLAND, J. B., AND BORREGAARD, N. Human neutrophil gelatinase-associated lipocalin and homologous proteins in rat and mouse. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1482, 1–2 (2000), 272 – 283.
- [27] KJELDSSEN, L., JOHNSEN, A. H., SENGELOV, H., AND BORREGAARD, N. Isolation and primary structure of ngal, a novel protein associated with human neutrophil gelatinase. *Journal of Biological Chemistry* 268, 14 (1993), 10425–10432.
- [28] KOBASHIGAWA, Y., SAIO, T., USHIO, M., SEKIGUCHI, M., YOKOCHI, M., OGURA, K., AND INAGAKI, F. Convenient method for resolving degeneracies due to symmetry of the magnetic susceptibility tensor and its application to pseudo contact shift-based protein–protein complex structure determination. *Journal of Biomolecular NMR* 53, 1 (2012), 53–63.
- [29] LANGEUEDDECKE, C., ROUSSA, E., FENTON, R. A., WOLFF, N. A., LEE, W.-K., AND THÉVENOD, F. Lipocalin-2 (24p3/neutrophil gelatinase-associated lipocalin (ngal)) receptor is expressed in distal nephron and mediates protein endocytosis. *Journal of Biological Chemistry* 287, 1 (2012), 159–169.
- [30] LI, C., AND CHAN, Y. R. Lipocalin 2 regulation and its complex role in inflammation and cancer. *Cytokine* 56, 2 (2011), 435 – 441.
- [31] LICHTENECKER, R., WEINHÄUPL, K., REUTHER, L., SCHÖRGHUBER, J., SCHMID, W., AND KONRAT, R. Independent valine and leucine isotope labeling in escherichia coli protein overexpression systems. *Journal of Biomolecular NMR* (2013), 1–5.
- [32] LICHTENECKER, R. J., COUDEVILLE, N., KONRAT, R., AND SCHMID, W. Selective isotope labelling of leucine residues by using α -ketoacid precursor compounds. *ChemBioChem* 14, 7 (2013), 818–821.
- [33] MAS, G., CRUBLET, E., HAMELIN, O., GANS, P., AND BOISBOUVIER, J. Specific labeling and assignment strategies of valine methyl groups for nmr studies of high molecular weight proteins. *Journal of Biomolecular NMR* (2013), 1–12.
- [34] MEIBOOM, S., AND GILL, D. Modified spin-echo method for measuring nuclear relaxation times. *Review of Scientific Instruments* 29, 8 (1958), 688–691.
- [35] MITTAG, T., MARSH, J., GRISHAEV, A., ORLICKY, S., LIN, H., SICHERI, F., TYERS, M., AND FORMAN-KAY, J. D. Structure/function implications in a dynamic complex of the intrinsically disordered sic1 with the cdc4 subunit of an {SCF} ubiquitin ligase. *Structure* 18, 4 (2010), 494 – 506.
- [36] MIYANOIRI, Y., TAKEDA, M., OKUMA, K., ONO, A., TERAUCHI, T., AND KAINOSHO, M. Differential isotope-labeling for leu and val residues in a protein by e. coli cellular expression using stereo-specifically methyl labeled amino acids. *Journal of Biomolecular NMR* (2013), 1–13.
- [37] MÉSZÁROS, B., TOMPA, P., SIMON, I., AND DOSZTÁNYI, Z. Molecular principles of the interactions of disordered proteins. *Journal of Molecular Biology* 372, 2 (2007), 549 – 561.

- [38] OTTING, G. Protein nmr using paramagnetic ions. *Annual Review of Biophysics* 39 (2010), 387–405.
- [39] PETERS, F., MAESTRE-MARTINEZ, M., LEONOV, A., KOVAČIČ, L., BECKER, S., BOELEN, R., AND GRIESINGER, C. Cys-ph-taha: a lanthanide binding tag for rdc and pcs enhanced protein nmr. *Journal of Biomolecular NMR* 51, 3 (2011), 329–337.
- [40] RASIA, R. M., BRUTSCHER, B., AND PLEVIN, M. J. Selective isotopic unlabeled of proteins using metabolic precursors: Application to nmr assignment of intrinsically disordered proteins. *ChemBioChem* 13, 5 (2012), 732–739.
- [41] RAYMOND, K. N., DERTZ, E. A., AND KIM, S. S. Enterobactin: An archetype for microbial iron transport. *Proceedings of the National Academy of Sciences* 100, 7 (2003), 3584–3588.
- [42] REZAEI-GHALEH, N., BLACKLEDGE, M., AND ZWECKSTETTER, M. Intrinsically disordered proteins: From sequence and conformational properties toward drug discovery. *ChemBioChem* 13 (2012), 930 – 950.
- [43] RICHARDSON. 24p3 and its receptor: Dawn of a new iron age? *Cell* 123, 7 (2005), 1175 – 1177.
- [44] RULE, G. S., AND T., H. K. Fundamentals of protein nmr spectroscopy. *Springer* 5 (2006).
- [45] SCANU, S., FOERSTER, J. M., ULLMANN, G. M., AND UBBINK, M. Role of hydrophobic interactions in the encounter complex formation of the plastocyanin and cytochrome f complex revealed by paramagnetic nmr spectroscopy. *Journal of the American Chemical Society* 135, 20 (2013), 7681–7692.
- [46] SCHANDA, P., AND BRUTSCHER, B. Very fast two-dimensional nmr spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *Journal of the American Chemical Society* 127, 22 (2005), 8014–8015.
- [47] SCHANDA, P., KUPCE, E. R., AND BRUTSCHER, B. Sofast-hmqc experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *Journal of Biomolecular NMR* 33, 4 (2005), 199–211.
- [48] SCHMIDT-OTT, K. M., MORI, K., LI, J. Y., KALANDADZE, A., COHEN, D. J., DEVARAJAN, P., AND BARASCH, J. Dual action of neutrophil gelatinase-associated lipocalin. *Journal of the American Society of Nephrology* 18, 2 (2007), 407–413.
- [49] SHENG, Z., WANG, S.-Z., AND GREEN, M. R. Transcription and signalling pathways involved in bcr-abl-mediated misregulation of 24p3 and 24p3r. *EMBO J* 28, 7 (2009), 866–876.
- [50] SIA, A. K., ALLRED, B. E., AND RAYMOND, K. N. Siderocalins: Siderophore binding proteins evolved for primary pathogen host defense. *Current Opinion in Chemical Biology* 17, 2 (2013), 150 – 157.
- [51] SIBILLE, N., AND BERNADO, P. Structural characterization of intrinsically disordered proteins by the combined use of nmr and saxs. *Biochemical Society Transactions* 40, 5 (2012), 955–962.

- [52] SINGER, E., MARKÓ, L., PARAGAS, N., BARASCH, J., DRAGUN, D., MÜLLER, D. N., BUDDE, K., AND SCHMIDT-OTT, K. M. Neutrophil gelatinase-associated lipocalin: pathophysiology and clinical applications. *Acta Physiologica* 207, 4 (2013), 663–672.
- [53] STAUNTON, D., SCHLINKERT, R., ZANETTI, G., COLEBROOK, S. A., AND CAMPBELL, I. D. Cell-free expression and selective isotope labelling in protein nmr. *Magnetic Resonance in Chemistry* 44, S1 (2006), S2–S9.
- [54] SU, X.-C., MCANDREW, K., HUBER, T., AND OTTING, G. Lanthanide-binding peptides for nmr measurements of residual dipolar couplings and paramagnetic effects from multiple angles. *Journal of the American Chemical Society* 130, 5 (2008), 1681–1687.
- [55] TOMPA, P. Intrinsically disordered proteins: a 10-year recap. *Trends in Biochemical Sciences* 37, 12 (2012), 509 – 516.
- [56] TOMPA, P., AND FUXREITER, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends in Biochemical Sciences* 33, 1 (2008), 2 – 8.
- [57] UBBINK, M. The courtship of proteins: Understanding the encounter complex. *{FEBS} Letters* 583, 7 (2009), 1060 – 1066.
- [58] VULPETTI, A., AND DALVIT, C. Fluorine local environment: from screening to drug design. *Drug Discovery Today* 17, 15–16 (2012), 890 – 897.
- [59] VULPETTI, A., HOMMEL, U., LANDRUM, G., LEWIS, R., AND DALVIT, C. Design and nmr-based screening of lef, a library of chemical fragments with different local environment of fluorine. *Journal of the American Chemical Society* 131, 36 (2009), 12949–12959.
- [60] VULPETTI, A., SCHIERING, N., AND DALVIT, C. Combined use of computational chemistry, nmr screening, and x-ray crystallography for identification and characterization of fluorophilic protein environments. *Proteins: Structure, Function, and Bioinformatics* 78, 16 (2010), 3281–3291.
- [61] WÖHNERT, J., FRANZ, K. J., NITZ, M., IMPERIALI, B., AND SCHWALBE, H. Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *Journal of the American Chemical Society* 125, 44 (2003), 13338–13339.
- [62] ZIEGLER, S., RÖHRS, S., TICKENBROCK, L., LANGERAK, A., CHU, S.-T., FELDMANN, I., JAKUBOWSKI, N., AND MÜLLER, O. Lipocalin 24p3 is regulated by the wnt pathway independent of regulation by iron. *Cancer Genetics and Cytogenetics* 174, 1 (2007), 16 – 23.

Chapter 7

Glossary

Glossary

ADME Absorption, distribution, metabolism and excretion 27, 93

ATP Adenosine triphosphate 93

BCR-Abl Philadelphia chromosome 93

Cdc4 Cell division control protein 4 23, 93

COSY Correlation spectroscopy 29, 93

Cox-2 Cyclooxygenase-2 93

CPMG Carr, Purcell, Meiboom, Gill 3, 23, 29, 33, 34, 51, 60, 61, 68, 80, 93

CREB cAMP response element-binding protein 23, 93

DNA Deoxyribonucleic acid 51, 52, 93

DOSY Diffusion-ordered spectroscopy 29, 93

DOTA 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid 25, 93

DPA D-Penicillamine 25, 93

DTPA Diethylene triamine pentaacetic acid 25, 93

EDTA Ethylenediaminetetraacetic acid 25, 93

FABP Fatty acid-binding protein 13, 93

FBDD Fragment-based drug discovery 3, 28, 37, 69, 81–83, 93

FRET Förster resonance energy transfer 23, 93

GB1 Immunoglobulin binding domain B1 of streptococcal protein G 3, 29, 37, 50, 72, 75, 83, 93

GMCSF Granulocyte macrophage colony-stimulating factor 15, 93

HER-2 Human epidermal growth factor receptor 2 20, 93

HMQC Heteronuclear multiple quantum coherence 32, 33, 37, 57, 93

Hsp Heat shock protein 23, 93

- HSQC** Heteronuclear single quantum coherence 3, 23, 32, 37, 64, 65, 69, 72, 74, 80, 81, 93
- IDP** Intrinsically disordered protein 22–24, 33, 93
- IGF** Insulin-like growth factor 15, 93
- IL** Interleukin 15, 93
- INEPT** Insensitive nuclei enhanced by polarization transfer 32, 93
- IPTG** Isopropyl β -D-1-thiogalactopyranoside 48, 93
- ITC** Isothermal titration calorimetry 37, 51, 60, 68, 69, 80, 81, 93
- JAK** Janus kinase 21, 93
- K_d** Dissociation constant 17, 18, 21, 27, 51, 60, 79, 81, 93
- LB** Lysogeny broth 48, 49, 93
- LBP** Lanthanide-binding peptide 26, 27, 51, 63–65, 68, 80, 83, 93
- LEF** Local environment of fluorine 28, 93
- MAPK** Mitogen-activated protein kinase 21, 93
- MMP** Matrix metalloproteinase 12, 19, 93
- MTSL** S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate 3, 25, 37, 63, 64, 68, 79, 80, 93
- NGAL** Neutrophil gelatinase-associated lipocalin 3, 12–21, 37, 49–53, 57–61, 63–66, 68, 69, 71, 72, 79–83, 93
- NMR** Nuclear magnetic resonance 3, 13, 23, 25, 26, 28–32, 37, 49–51, 57, 60, 63, 64, 68, 74, 79, 81, 82, 93
- NTD** N-terminal domain of 24p3R 3, 57–61, 63, 64, 66, 68, 79–81, 83, 93
- OD** Optical density 93
- PBS** Phosphate buffered saline 48, 93
- PCR** Polymerase chain reaction 51, 52, 93
- PCS** Pseudocontact chemical shift 3, 25, 26, 64, 80, 81, 93
- PDB** Protein data bank 29, 93
- PRE** Paramagnetic relaxation enhancement 3, 23, 25, 26, 64, 80, 93
- Ras** Rat sarcoma 21, 93
- RDC** Residual dipolar coupling 3, 23, 25, 26, 64, 80, 93
- SAXS** Small-angle X-ray scattering 93

SCR Structurally conserved regions 13, 93

SOFAST Band-selective optimized-flip-angle short-transient 33, 37, 57, 93

Stat Signal transducer and activator of transcription 21, 93

TEMPO Piperidine-type nitroxide 2,2,5,5-tetramethyl-4-piperidin-1-oxyl 25, 93

TEV Tobacco etch virus 48, 50, 93

TGF Transforming growth factor 15, 93

TNF Tumor necrosis factor 15, 93

Chapter 8

Acknowledgements

Acknowledgements

I thank Univ. Prof. Dr. Robert Konrat for providing me the opportunity to perform my master's thesis at the NMR group. In addition I want to thank Nicolas Coudevylle for his effort, support and patience, while supervising me, and Fabian Gruber for critical reading of my master's thesis. I would also like to thank all members of the NMR group for their scientific advice and moral support.

Chapter 9

Curriculum Vitae

Personal Data	
Name:	Katharina Weinhäupl
Address:	Wilhelm-Dachauer-Weg 3
Telephone:	+436801433876
e-mail:	kathi.weinhaeupl@gmail.com
Date of Birth:	September 29, 1987
Place of Birth:	Wr. Neustadt, Austria
Nationality:	Austrian
Education	
1994 - 1998:	Elementary school: “Volksschule II”, Ried im Innkreis
1998 - 2006:	Grammar school: “BG/BRG Ried im Innkreis”, graduated with distinction
December 2004:	Student exchange with the “Liceo Simone Weil” in Treviglio, Italy
2006 - 2013:	Diploma Programme Pharmacy at the University of Vienna - awarded title: Mag.pharm.
2007 - 2010:	Bachelor’s Degree Programme Biology at the University of Vienna - awarded title: Bsc.
March 2011:	Enrollment Master’s Degree Programme Molecular Biology at the University of Vienna
Since February 2013:	Master’s Thesis in the Group of Prof. Konrat
Language Skills	
German	mother tongue
English	CEFR: C1
Italian	CEFR: B2
Spanish	CEFR: A2
Work Experience	
July 2005	GST Global Sports Technologies GmbH
July 2007	“Apotheke zur heiligen Dreifaltigkeit”, Berndorf
July - August 2008	“Apotheke zur heiligen Dreifaltigkeit”, Berndorf
August - September 2009	“Apotheke zur heiligen Dreifaltigkeit”, Berndorf
July 2010	“Apotheke zur heiligen Dreifaltigkeit”, Berndorf
September 2011	“Apotheke zur heiligen Dreifaltigkeit”, Berndorf