Development of Zinc-Finger-Based Artificial Restriction Endonucleases and Fluorescent Peptidyl Metal Sensors

DISSERTATION

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsprogramm: IRTG 1422

der Georg-August-Universität School of Science (GAUSS)

vorgelegt von

Florian Czerny

geboren in Straubing

Göttingen 2016

Thesis Committee Members:

Prof. Dr. Ulf Diederichsen   Institut für Organische und Biomolekulare Chemie,
(Referent)   Georg-August-Universität Göttingen

Prof. Dr. Franc Meyer   Institut für Anorganische und Biomolekulare Chemie,
(Co-Referent)   Georg-August-Universität Göttingen

Prof. Dr. Sofi Elmroth   Center for Molecular Protein Science,
(Co-Referent)   Lund Universitet, Lund, Sweden

Members of the Examination Board:

Prof. Dr. Ulf Diederichsen   Institut für Organische und Biomolekulare Chemie,
                                    Georg-August-Universität Göttingen

Prof. Dr. Franc Meyer   Institut für Anorganische Chemie,
                                    Georg-August-Universität Göttingen

Prof. Dr. Philipp Vana   Institut für Physikalische Chemie,
                                    Georg-August-Universität Göttingen

Prof. Dr. Ivo Feußner   Department of Plant Biochemistry,
                                    Georg-August-Universität Göttingen

Prof. Dr. Kai Tittmann   Department of Molecular Enzymology,
                                    Georg-August-Universität Göttingen

Dr. Franziska Thomas   Institut für Organische und Biomolekulare Chemie,
                                    Georg-August-Universität Göttingen

Date for thesis disputation: 8th August 2016

Development of Zinc-Finger-Based
Artificial Restriction Endonucleases
and Fluorescent Peptidyl Metal Sensors

Table of Contents

Table of Contents

1. Introduction and research objectives

Zinc finger proteins (ZFPs) are ubiquitous in the cellular environment having an excessive diversity and performing variable tasks. Their outstanding capability to develop a sequence-specific bond to deoxyribonucleic acid (DNA) makes them one of the most abundant DNA-binding domains (DBD) in many eukaryotic and prokaryotic transcription factors.[1] The ability of ZFPs to address numerous DNA sequences by an alteration of the amino acids that are involved in nucleobase binding brings them more and more into the focus of research as these characteristics seem to play an important role for the development of novel gene-editing tools.[2] Besides the recently established gene-editing systems CRISPR/Cas9 and TALENs, zinc finger nucleases (ZFN) are well examined and most frequently used in gene therapeutic applications as well as in manipulation processes of the genome of many plants, animals and microorganisms.[3] This is made possible by the connection of the sequence-specific, DNA-binding zinc finger domain with the non-specific restriction endonuclease FokI that cleaves the phosphate backbone of the DNA in the immediate vicinity of the DNA binder. Artificial restriction enzymes, which are able to establish a sequence-specific bond to DNA and to perform hydrolysis at precisely predictable positions, are nowadays essential for the development of tailor-made organisms. Especially the cultivation of crop plants has experienced an enormous upturn due to the implementation of gene-editing systems.[4] In the past decades, genetic modifications with regard to yield improvements and increased resistance against parasites and diseases mainly relied on the crossing process of different types of organisms in order to combine their best attributes in one newly cultivated plant.[5] This laborious and time-consuming process was drastically simplified by the evolvement of the latter systems, holding an unprecedented precision and effectiveness. However, the enormous molecular dimensions of the nuclease enzyme (587 aa, 65.4 kDa) in comparison to the much smaller zinc finger domain (156 aa, 17.8 kDa) and the non-specific cleaving capacity of FokI, which only hydrolyzes in an area restricted by the attached zinc finger, leave room for improvements.

For these reasons, the present study deals with the synthesis and modification of the zinc finger protein Zif268 in order to establish novel and substantially downsized artificial restriction endonucleases. This was achieved by the development of two approaches that both make use of the specific DNA-binding ability of the zinc finger but differ in the phosphodiester cleavage strategy applied. The first approach is based on the concept of small molecule model complexes, mimicking the active sites of nuclease enzymes. The latter are often composed of dinuclear complexes of transition metal ions, such as Zn(II) or Cu(II), which are able to bind phosphate groups in order to activate them and to initiate the DNA cleavage process.[6] The achievements in the field of operating organometallic model nuclease synthesis were transferred to organic peptide synthesis in order to prepare peptide

incorporable dinuclear building blocks. These building blocks were integrated in the native peptide sequence at strategic positions that come close to the phosphate backbone of the DNA. Thus, upon binding of the zinc finger to its DNA target sequence, the artificial amino acids are able to reach distinct phosphate groups to perform site-specific hydrolysis.

The second approach takes advantage of the overwhelming hydrolysis capacity of the lanthanide ion cerium(IV) which had successfully been used in the development of artificial restriction DNA cutters (ARCUTs).[7] ARCUTs rely on the cooperation of a DNA-invading polypeptide and a hydrolytically active external complex composed of Ce(IV) and EDTA. The latter is recruited to a specific cleavage site by a phosphoserine moiety, which is attached to the invading polypeptide.[8] This concept was transferred to the development of phosphoserine-modified zinc fingers, which specifically bind DNA and thereby accumulate the external Ce(IV) complex in the surrounding of the artificial amino acid. Consequently, the close proximity of the phosphoserine residue to the phosphate backbone facilitates the relocation of the lanthanide complex between both moieties in order to perform hydrolysis within a regulatory framework provided by the zinc finger.

Both approaches are based on a semi-synthetic methodology, which included the modification of the third zinc finger of Zif268 by either incorporating an internal dinuclear building block or a phosphoserine residue by means of solid phase peptide synthesis (SPPS). In addition, the tandem-peptide containing the first and the second zinc finger was expressed in *E. coli* and was used in an expressed protein ligation (EPL) approach with the modified zinc finger 3 to generate the full-lengths peptide with pronounced DNA-recognizing abilities.

1.1 The zinc finger protein

With more than thousand different members that have been discovered yet, zinc finger proteins are considered to be one of the most comprehensive classes of DNA-binding proteins found in nature.[9] This is due to their participation as DNA-binding domains in the modular structure of many eukaryotic and prokaryotic transcription factors. In this function, they are involved in the early steps of DNA transcription, an ubiquitous process that uses the genetic information by translating DNA sequences into corresponding messenger RNAs (mRNAs). Thereby, ZFPs recognize a specific promoter sequence, which is located upstream of the genetic code that should be transcribed by RNA-polymerases (Figure 1.1). Under retention of an uniform globular $\beta\beta\alpha$-structure, zinc fingers differ in their individual amino acid sequences to bind a large variety of promoter regions in a sequence-specific manner. This explains the high diversity of this protein family in the cellular environment and finally allows them to trigger transcriptional processes for a myriad of genes.
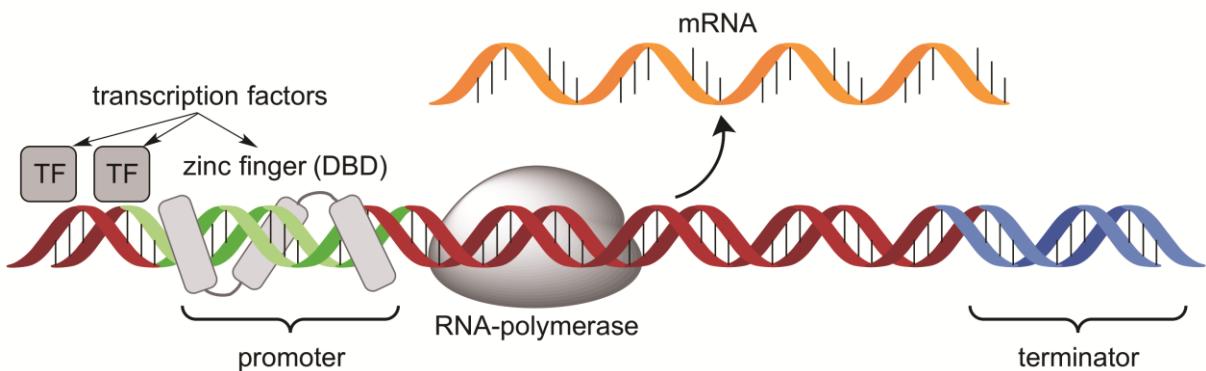
Figure 1.1 Schematic representation of transcriptional processes triggered by several transcription factors (TF) involving a zinc-finger-based DNA-binding domain (DBD).

Zinc finger proteins are composed of tandem repeated small zinc finger sub-domains containing approximately 30 amino acids, respectively.[10] Each sub-domain can be considered as individual zinc finger that folds into a simple but very stable $\beta\beta\alpha$-structure. Therefore, zinc fingers are composed of a highly conserved amino acid sequence of the general form: (F/Y)-X-Cys-X_{2-5}-Cys-X_3-(F/Y)-X_5-Φ-X_2-His-X_{3-5}-His whereby Φ stands for an hydrophobic amino acid and X indicates a freely selectable amino acid. ZFs are mainly unfolded under metal-free conditions but they receive their unique secondary structure upon Zn(II) complexation. Although ZFs differ in their amino acid sequences, the common characteristic of these proteins is the presence of an α-helix and an antiparallel β-hairpin, which are held together by a single tetrahedrally coordinated Zn(II) ion (Figure 1.2a). Individual zinc fingers are connected to each other by the short and conserved TGXKP (X represents any amino acid) linker sequence. The thus obtained protein is able to wrap around the major grove of the DNA, whereby three amino acids of each recognition helix are able to recognize exactly three nucleobases (Figure 1.2b).
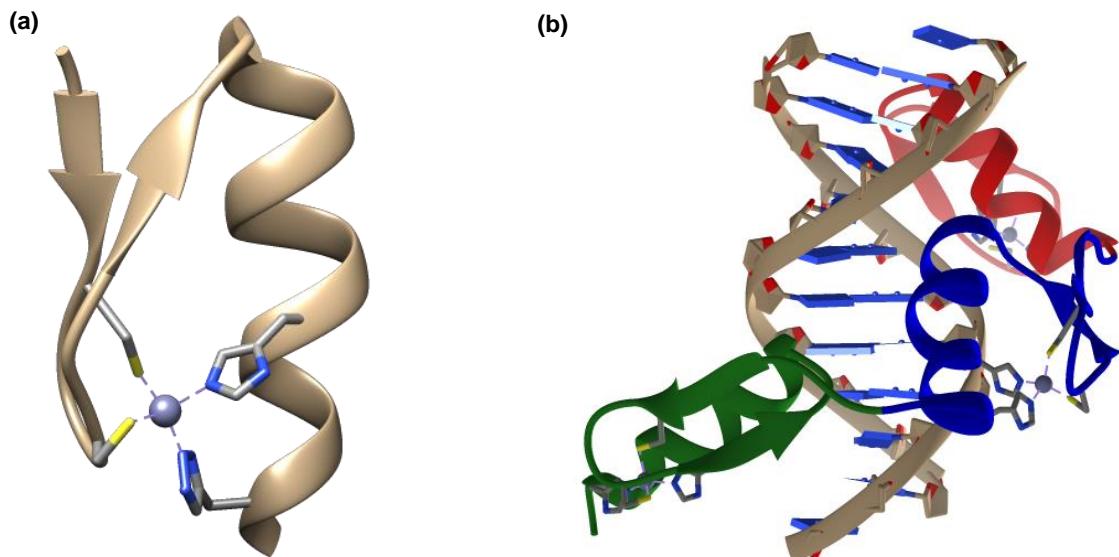
Figure 1.2 (a) βBα-Structure of a single zinc finger upon Zn(II) complexation. **(b)** The three tandem zinc finger domains (green, blue and red) of Zif268 binding to the major groove of the consensus dsDNA-binding site (PDB code 1AAY). Images were generated with UCSF Chimera.

Depending on the quantity of tandem repeated zinc fingers, the number of recognized DNA sequences can be enormous. In this meaning, gene-specific zinc fingers are usually named after the transcription factor they derive from. A prominent example is the transcription factor TFIIIA from *Xenopus laevis*.[11] It has a DNA-binding domain, which contains twelve tandem repeated zinc fingers and was the first reported example of this protein family by BRAUN and MILLER in 1985. The number of individual zinc finger motifs per protein differs significantly and ranges from 2 (ADR1) to 12 (TFIIIA) to 13 (ZFY) up to 37 (Xfin).[12-14] However, one of the most comprehensively studied ZFP, which was also subject of this study, is the mammalian transcription factor Zif268 with its three tandem zinc finger domain.[15]

1.1.1 The zinc finger motif Zif268

The DNA-binding domain of transcription factor Zif268 consists of three zinc finger motifs of the "classical" Cys_2His_2-type with a total number of 90 amino acids. Proteins of this type use two histidine and two cysteine residues to coordinate Zn(II) in a tetrahedral fashion. Cys_2His_2 coordination is most commonly found for ZFs but also Cys_3His and Cys_4 zinc fingers are thoroughly reported.[16] The consensus sequence of all three tandem zinc fingers (Zf1, Zf2 and Zf3) of Zif268 is displayed in Figure 1.3.[10]
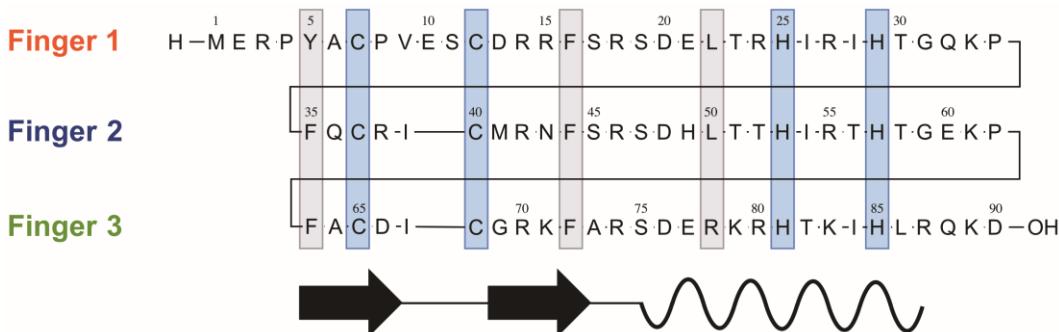
Figure 1.3 Amino acid sequence of the three tandem zinc finger domain of Zif268. Metal coordinating residues are highlighted in blue and amino acids, which form the hydrophobic core are highlighted in gray. Residue positions in the β-sheets (arrow) and the α-helix (wave) are as indicated.[10]

The sequence is highly conserved for certain amino acids to ensure the correct formation of the ββα-structure, which is required for sequence-specific DNA binding. This includes in particular the metal-coordinating cysteine and histidine residues initiating the folding process upon Zn(II) complexation (Figure 1.3, blue). The histidine residues reside in the α-helical part, three amino acids apart from each other, whereas the cysteine residues are located in the β-hairpin, one in each β-sheet. The structure is held together by the intermediate Zn(II) ion further causing the formation of a stabilizing hydrophobic core composed of three conserved hydrophobic amino acids per individual zinc finger (Figure 1.3, gray).

The amino acids, which are responsible for nucleobase recognition and binding, are of particular importance. Each zinc finger motif contains three amino acids at the helical positions –1, 3 and 6 that are responsible for direct nucleobase contact to the DNA operator region 5′-GCGTGGGCGT-3′ (Figure 1.4).[10] One additional nucleobase contact per motif, which are known to be not imperatively required, is performed from position 2 to the complementary DNA strand 5′-ACGCCCACGC-3′. All nucleobase binding residues are located in the zinc fingers helical domain, which wraps around the DNAs major groove and thereby establishes the connection through hydrogen bonding of the amino acid side chains.[17] As it can be seen from Figure 1.4, all three zinc fingers of Zif268 have an arginine residue at position –1, which respectively addresses the guanine bases of the promoter strand and an aspartic acid at position 2. The latter makes contact to either the exocyclic N^6-amine of adenine or the exocyclic N^4-amine of cytosine located in the complementary strand. Positions 3 and 6 are more variably occupied. Whereas Zf1 and Zf3 have a glutamate residue at position 3 to interact with cytosine, Zf2 has a histidine residue at this position forming a hydrogen bond to a guanine base and maintaining additional van der Waals interactions with the neighboring thymine. Position 6 is occupied by an arginine residue in Zf1 and Zf3 that binds guanine, whereas Zf2 forms a hydrogen bond to thymine through out a threonine residue at this position.[18]
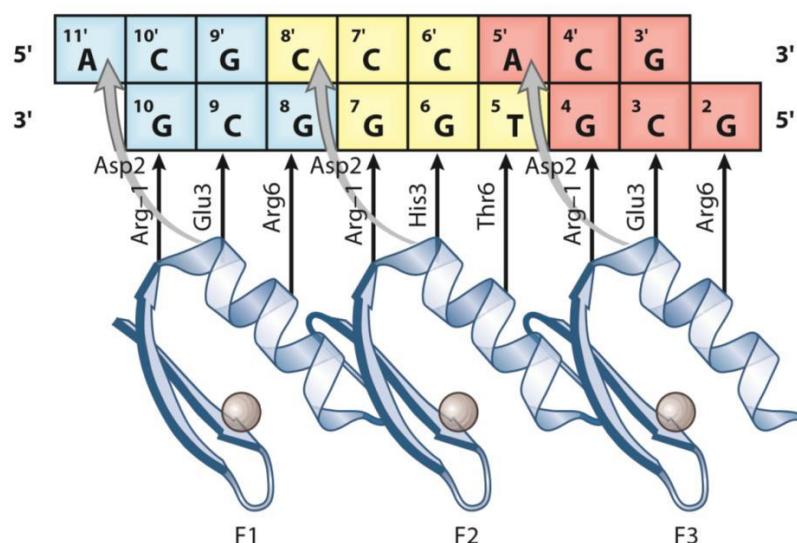
Figure 1.4 Schematic representation of the DNA recognizing amino acid residues of Zif268 establishing direct (positions -1, 3 and 6) or indirect (position 2) nucleobase contacts.[18]

The sequence-specific characteristics of the zinc finger proteins that contribute to structural properties, such as metal coordination or hydrophobic core formation, are highly conserved, and therefore, almost not alterable. In contrast, the determination of the amino acids, which are involved in nucleobase binding and their well-defined positions in the protein sequence, paved the way for the development of engineered zinc finger domains. With this knowledge, it is possible to modulate the recognition helix of each zinc finger to address any desired DNA sequence.

1.1.2 Engineered zinc finger proteins

Engineered zinc finger proteins emerged with the identification of the key positions −1, 2, 3 and 6 as well as their underlying concept of nucleobase specificity of those amino acids occupying the latter positions (Figure 1.4). Thus, the generation of tailor-made tandem repeated zinc fingers, which recognize any specific target DNA sequence, is easily accessible by following the patterns displayed in Figure 1.5.[19]
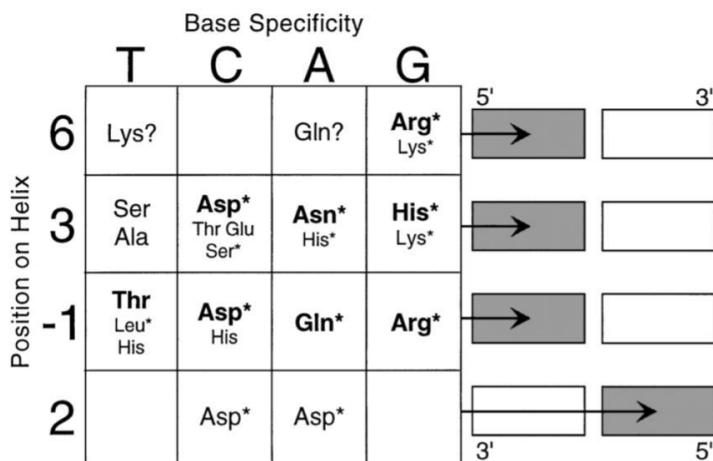
Figure 1.5 Pattern of the nucleobase specificity of the amino acids occupying the DNA-binding positions (6, 3 –1 and 2) in the zinc fingers α-helix. The residues (which were most frequently found in a phage assay) to bind a specific nucleobase are shown in bold, whereas the asterisk indicates the findings from structural analyses. Blank positions show undefined interactions and the question mark stands for uncertain relations between any amino acid at these positions and a particular nucleobase.[19]

The full potential of these achievements was used by coupling a modulated zinc finger protein to different additional (bio-)molecules, which have functional or catalytic properties themselves. Despite the therefore generated gene-activating-/repressing- or silencing-proteins, the most prominent examples created in this way belong to the DNA-hydrolyzing zinc-finger nucleases (ZFNs). ZFNs raised as important gene editing tool due to the combination of the tunable DNA recognition ability of the zinc finger with the overwhelming hydrolysis activity of the otherwise non-specific restriction endonuclease FokI.[4] The enzyme is thereby fused to the C-terminus of a zinc finger with programmed or known specificity towards a DNA target sequence. A spacer between both moieties produces a distance of approximately 7 bp to ensure efficient hydrolysis without molecular interferences. Moreover, the cleavage domain must be able to dimerize in order to perform DNA double-strand hydrolysis with the formation of blunt ends (Figure 1.6).[20]

Another consequence of the latter prerequisite is the requirement of a second FokI domain at the same binding site but facing the first one from the opposite DNA strand. This is achieved by a second zinc-finger nuclease moiety, which binds the opposite strand and leaves enough space between both ZF C-termini to allow the cleavage domain to dimerize.[20] The thus obtained double-strand cleavage opens the possibility to edit the genome in different ways. Most commonly found is the process of homology directed repair (HDR), which derives from the intrinsic cellular repair mechanism of DNA double-strand lesions by homologous recombination (Figure 1.7). The cell is only able to perform HDR if there is a homologous DNA template remaining in the nucleus, which can be used for gene correction.[21] Using the ZFN technology, it is furthermore possible to insert deviating pieces of DNA termed

transgenes. This method is excessively used in the cultivation of highly resistant crop plants or in order to investigate the influence of certain genes on organisms, which are therefore brought into the host genome of suitable organisms, such as *Drosophila melanogaster*.

The second cellular process to react on double-strand DNA breaks is non-homologous end joining (NHEJ).[21] In contrast to the aforementioned insertion of a homologous template of DNA, NHEJ relies on direct ligation of the cleaved DNA ends. ZFNs make use of the NHEJ mechanism by causing a large deletion mutation, whereby a specific deletion area is resected from the DNA by two pairs of zinc fingers arranged at both cutting edges.[22] The thus created ends are rejoined by imperfect ligation, while the deletion sequence is lost and prone to enzymatic digestion.
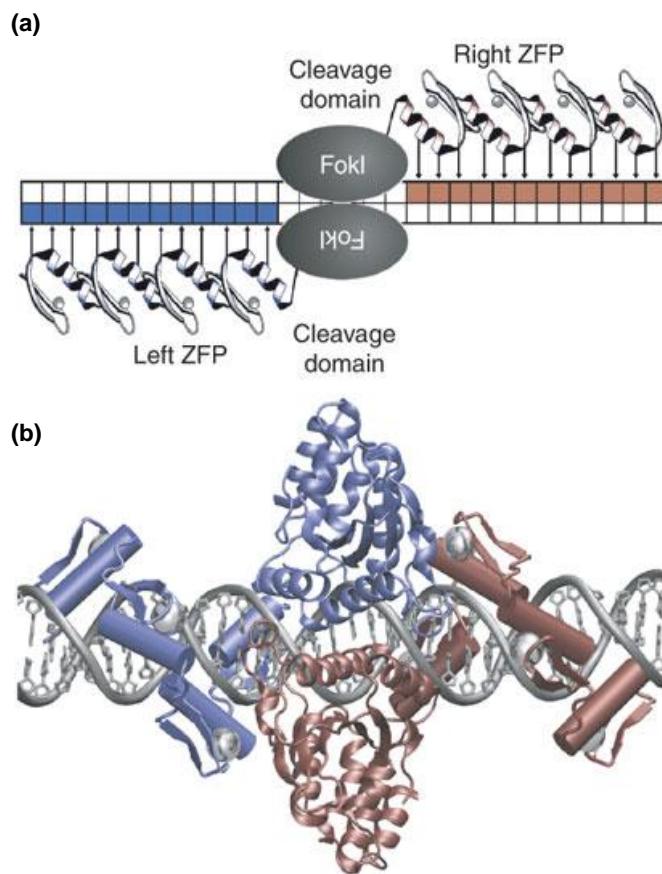
Figure 1.6 Functional principle of the zinc-finger nuclease technology. **(a)** Schematic representation of the zinc-finger domains bound to DNA and intermediate formation of the dimeric FokI cleavage domain. **(b)** Three-dimensional model showing the major-groove binding of the zinc-finger proteins and heterodimeric FokI association.[20]

The aforementioned applications are vital examples for the use of the ZFN technology in gene therapy to site specifically delete mutated genes or to induce a gene knock-out to make the mutated sequence inoperable.[22] Moreover, editing the genome was simplified in this way and became omnipresent in the development of customized agricultural products.[4]
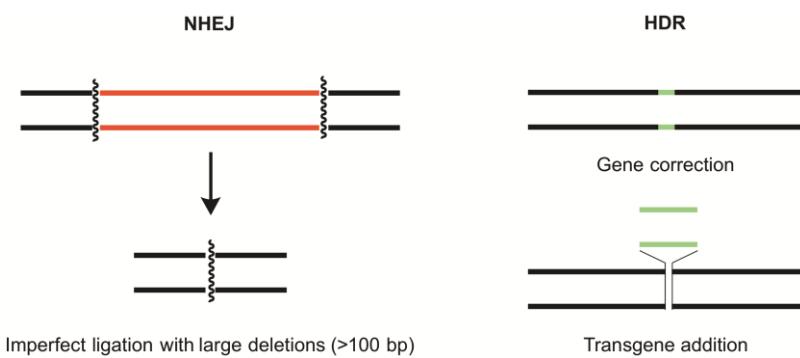
Figure 1.7 Mechanisms to repair and utilize ZFN induced DNA double-strand lesions.

1.2 Hydrolytically active metalloenzymes

Due to their catalytic involvement in almost every metabolic process in living cells, protein enzymes are of greatest importance in living organisms. To date, there are more than 5.000 reported reactions that are catalyzed by these macromolecular catalysts.[23] Enzymes are highly specialized in terms of substrate recognition and binding as well as in performing specific reactions, which are the basis to categorize these proteins in different sub-groups. For instance, oxidoreductases[24] catalyze oxidation and reduction reactions and transferases[25] are involved in the transfer of a functional group from one substrate to another. Moreover, degradation processes in living organisms are permanently required for the activation and deactivation of membrane channels, such as in aquaporins, or to digest incorrect sequences or macromolecules that are not required any longer.[26] A very important class of enzymes classified as hydrolases fulfill these tasks. The latter mainly belong to some regulatory classes of enzymes, which catalyze the hydrolysis of a variety of single bonds under physiological conditions.[27] Depending on their target substrate, they can be sub-divided into proteases/peptidases[28], which cleave amide bonds between amino acids, lipases[29], which scissor ester bonds of lipids to generate fatty acids and glycerol and nucleases[30], which are able to cleave phosphodiester bonds between the nucleotides of the DNA. Especially the last-named nucleases are quite good examples for the vast efficiency of enzymes. Under ideal conditions, the half-life of DNA can be approximately 1.7 million years implying the huge stability of phosphodiester bonds that connect individual deoxyribonucleosides.[31] Moreover, the genetic code is well protected and stays intact for long time even at elevated temperatures or under alkaline conditions. Therefore, it is remarkable that nuclease enzymes are able to cleave the latter bonds in just a fraction of time and with a notable degree of specificity.

The hydrolysis ability of nuclease enzymes is mainly based on the occurrence of a multinuclear complex of transition metals in the active site of these proteins. The suitability of

a metal cation to successfully perform hydrolysis is determined by three major criteria: (1) The Lewis acidity of the metal, (2) its affinity towards oxygen atoms and (3) its ability to rapidly substitute one substrate for another.[32] For this reasons it is not surprising that most active sites of numerous nucleases are composed of Zn(II), Mg(II) or Fe(II/III). Some prominent representatives of this class of enzymes are for example the alkaline phosphatase and the P1 nuclease (Figure 1.8).[33,34]

Whereas the first enzyme is able to cleave mono-phosphate groups (dephosphorylation) from nucleotides, alkaloids and proteins, the latter enzyme exclusively cleaves phosphodiester bonds as it forms part of single-stranded DNA. Nevertheless, both enzymes hold several zinc ions in their active sites that are responsible for the cleavage reaction. Alkaline phosphatase has a catalytic core containing a dinuclear zinc center with an additional magnesium ion (Figure 1.8a), while the P1 nuclease comprises a trinuclear zinc center without any additional metals (Figure 1.8b). The metal ions are tightly complexed by different amino acid side chains, such as the *N*-donating imidazole of histidine or the *O*-donating carboxylate originating from aspartic acid. These ligands create an ideal distance between the individual metal ions in order to enable substrate binding and water activation for the generation of the hydrolytically active species in a concerted mechanism.[35]
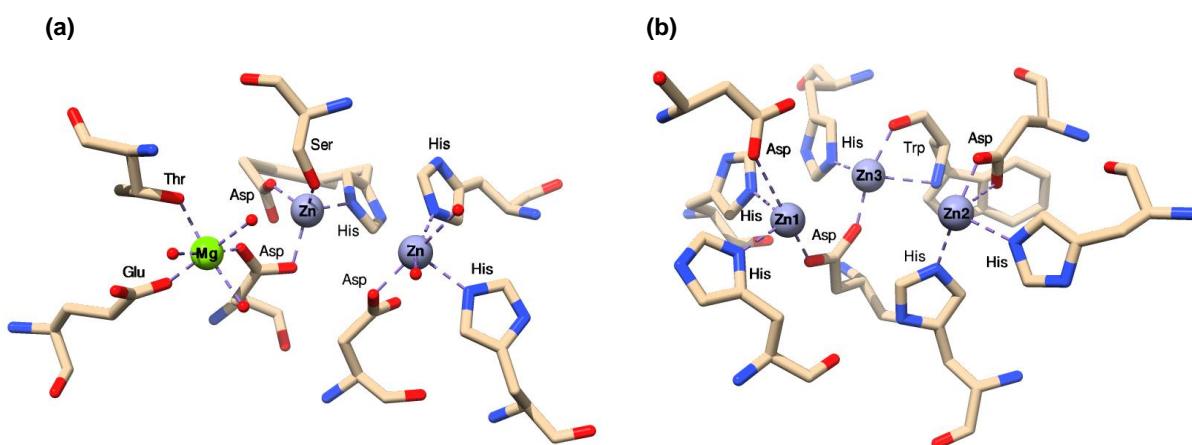
Figure 1.8 Multinuclear active centers of the alkaline phosphatase (*left*, PDB code 3WBH) and the P1 nuclease (*right*, PDB code 1AK0). Images were generated with UCSF Chimera.[33 34]

1.2.2 The mechanism of phosphodiester hydrolysis

Studies on the mechanism of phosphodiester hydrolysis revealed the importance of a hydrolytic core that is composed of a multinuclear metal center.[32] Besides the aforementioned demands on the metal ions for efficient hydrolysis, the distances between the individual metal ions as well as their coordinating amino acid ligands are also vital for an efficient cleavage.

In this sense, one of the best examined examples for single-stranded DNA hydrolysis is the nuclease P1, which cleaves the P-O (3′) bond between nucleotides.[36] The importance of having several zinc ions in close proximity to each other is illustrated by the two major tasks they have to perform in order to realize DNA hydrolysis. Figure 1.9 shows a schematic representation of the hydrolytic mechanism suggested. Step one involves binding of the substrate by the enzyme, which is attributed to the Zn2 ion. Upon coordination of the metal ion by the phosphate oxygen atom, the latter is activated and becomes susceptible for a nucleophilic attack. The attack comes from an activated water molecule located at the boundary between Zn1 and Zn3. This leads to a reduction of its pK_a resulting in the formation of a metal-bridging hydroxide species. This step is promoted by an aspartic acid residue (Asp45), which properly orients both metal ions as well as the hydroxide species in order to generate the correct distance for the activation to take place. Subsequently, the hydroxide attacks the DNA phosphate group that is activated by Zn2, whereby a pentacoordinated intermediate is generated. This reaction is generally reversible but due to the catalytic involvement of the substrate-activating metal ion, it is forced in the direction of phosphodiester cleavage. This was further confirmed by the documented stabilizing effect of the Zn2 ion on the O(3′)-oxyanion leaving group.[36]
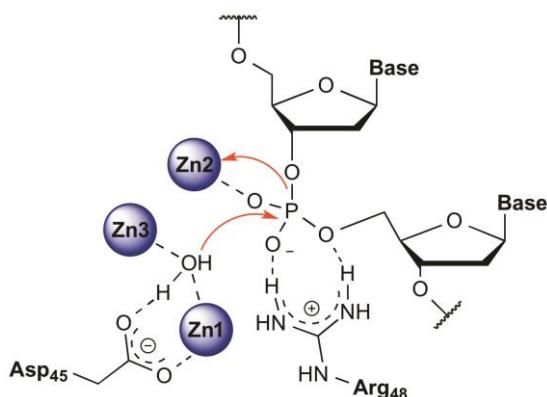
Figure 1.9 Single-stranded DNA hydrolysis by P1 nuclease.[36]

This general mechanism is applicable to various types of nucleases even though their active sites contain different metal ions, such as Mg(II) found in the alkaline phosphatase. The latter ion is known to be less involved in the hydrolysis process itself but instead it plays a crucial role in the enhancement of the overall activity of the enzyme by participating as an auxiliary.[37]

These examples demonstrate that nucleases have highly conserved active sites composed of multinuclear complexes of suitable transition metals, such as Zn(II), which are mainly coordinated by histidine and aspartic acid residues. The latter cause distinctive metal-to-metal distances enabling substrate binding and activation as well as water activation to promote phosphodiester hydrolysis. For a more profound understanding of the characteristics that allow enzymes to perform such an efficient hydrolysis process as well as to spread light on the interplay of the involved metal ions, functional small model complexes were synthesized, which mimic the active sites of their natural counterparts.

1.2.3 Model complexes for phosphodiester hydrolysis

Certain experimental techniques, such as X-ray spectroscopy or mutation analyses of amino acids, which were supposed to play crucial roles in the active sites of nuclease enzymes, contributed to the understanding of the hydrolysis mechanism described above. This paved the way for the synthesis of model complexes, which reduce the enzyme to its essential part required for hydrolysis; the catalytically active core.[38] These models spread further light on the understanding of the metal and ligand patterns while detailed kinetic studies are easily accessible due to the reduced size of the active center with a manageable number of inorganic scaffolds and ligands.

Numerous artificial metallonuclease model-complexes are found in literature.[6,39,40] Most of them transpose the parameters that specify naturally occurring enzymes, such as the attendance of a multinuclear metal center with distinct distances between the latter ions in order to bind and activate the substrate and to promote water activation.

Several studies on dinuclear model complexes involving transition metals, such as Zn(II), Fe(II/III) or Cu(II) state, so that their catalytic activities are many times greater compared to their mononuclear analogues.[41] This feature is based on the fact that two connected metals ions can much better lower the pK_a of a water molecule in order to generate the hydrolytically active hydroxide nucleophile under physiological conditions. Moreover, dinuclear metal complexes show an increased activation ability of the bound phosphate esters due to multielectron-transfer processes taking part between the metal sites and the substrate.[42]

Another vital aspect for the design of these complexes is the utilization of homo- or heterodinuclear metal combinations. Heterodinuclear metal complexes are found for example in purple acid phosphatases (PAPs), a class of enzymes which perform mono-phosphate ester cleavage at low pH values between 5 – 6.[43] An Fe(III) ion is found in all types of PAPs and is also responsible for the deep purple color caused by a ligand-to-metal charge transfer from the tyrosinate residue to the Fe(III). The occupation of the second metal ion at the active site is dependent on the organism in which the enzyme can be found. In mammalians, a redox active homodinuclear but mixed-valent Fe(III)/Fe(II) complex is found as catalytically active species (Figure 1.10a). In contrast, plant PAPs have an active site composed of either an heterodinuclear Fe(III)/Zn(II) couple or an Fe(III)/Mn(II) couple (Figure 1.10b).[44]
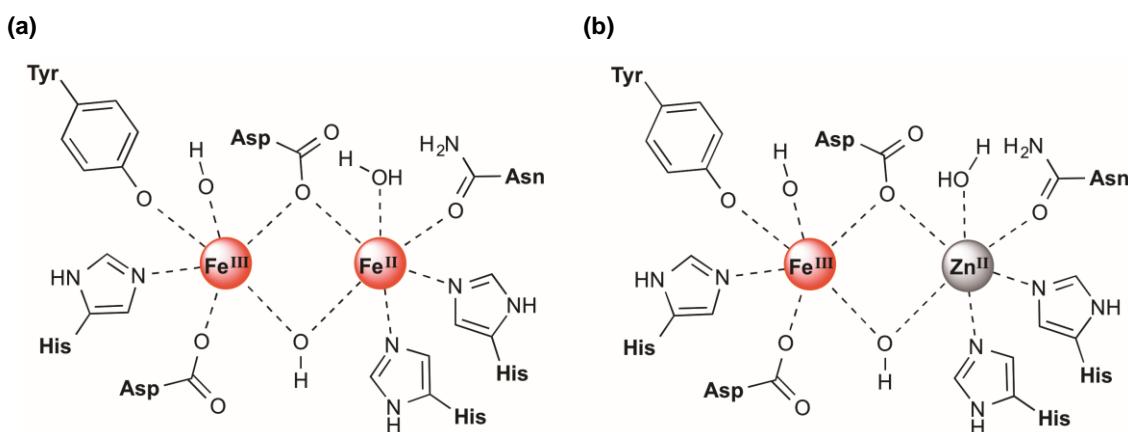
Figure 1.10 Active sites of mammalian PAPs **(a)** having a homo-dinuclear Fe(III)/Fe(II) couple (from uteroferrin) and plant PAPs **(b)** with heterodinuclear Fe(III)/Zn(II) couple (red kidney bean).

Model complexes of mammalian PAPs with an active mixed-valent Fe(III)/Fe(II) site are rare due to difficulties in avoiding the formation of the oxidized, and therefore, inactive Fe(III)/Fe(III) couple.[45] Much better results have been achieved in the synthesis of heteronuclear plant PAP-models by using an unsymmetrical coordination sphere tailored to the individual characteristics of each metal ion in order to regioselectively coordinate the latter.[46] It had been found that Fe(III) prefers rather "hard" donor ligands, and therefore, resides in an oxygen-rich environment composed of carboxylate side chains of aspartic acid residues and a phenolate side chain of tyrosine. In contrast, Zn(II) or Mn(II) prefer a "softer" environment with an additional *N*-donating ligand, such as histidine, as well as with an asparagine residue instead of an anionic aspartate ligand. Some promising PAPs model complexes were successfully synthesized transposing the aforementioned characteristics and being used to evaluate the role of each individual metal in the hydrolysis reaction (Figure 1.11).[47,48] Furthermore, ligand **B** is one of the most comprehensively studied mimic of the mixed valent mammalian PAPs. It provides a soft coordination site (N_3O_3-coordination)

suitable for Fe(II) and a hard coordination site (N_2O_4-coordination) for Fe(III). The metal-metal distance of the acetate bridged complex was determined to be 3.48 Å that is very close to the 3.31 Å reported for natural uteroferrine derived PAPs. However, the mechanism of substrate hydrolysis is still under debate due to opposite opinions regarding the origin of the attacking hydroxide species. On the one hand, the latter species is supposed to be terminally coordinated to the Fe(III) ion, from where it attacks the phosphorus atom. On the other hand, some approaches gave rise to the assumption that it might be bridged between the two metal ions.[47,48]
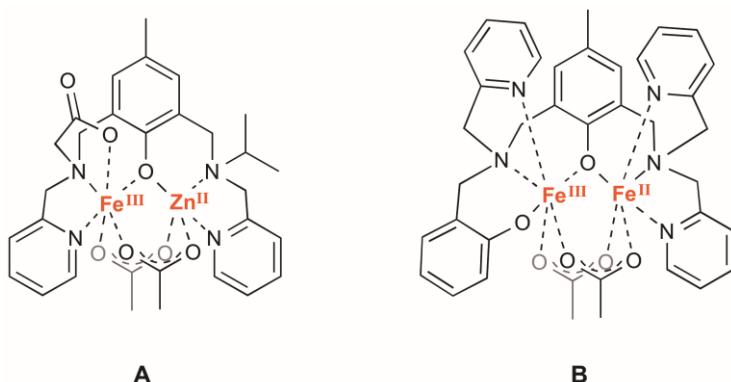
Figure 1.11 Examples of model complexes mimicking the active site of PAPs.[47,48]

However, model complexes for both, heterodinuclear and homodinuclear as well as homodinuclear but mixed valent metals were successfully synthesized. In many cases, the metals are coordinated by a scaffold, which is based on di-ortho-substituted phenols (Figure 1.12, **C+D**).[49] The phenolate acts as bridge between the two metal ions, which are held together by two tridentate ligands attached to both ortho positions of the aromatic ring (Figure 1.12). A different scaffold is based on substituted pyrazolates (Figure 1.12, **E**), which comprise the advantage of an adjustable metal-metal distance by the modulation of the attached tridentate ligands.[50] These can either be adjusted by different spacer lengths between both moieties or by the generation of an asymmetric coordination sphere.
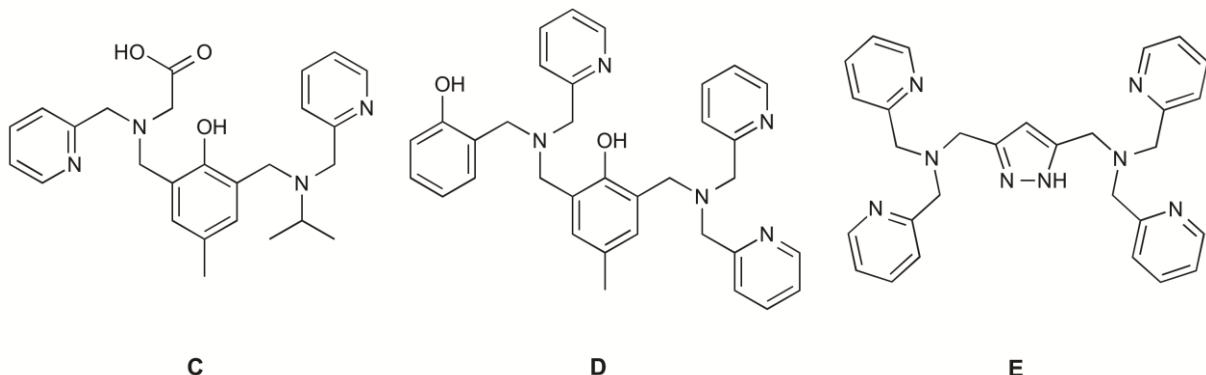
Figure 1.12 Multidentate ligands for the generation of dinuclear metal complexes. Phenol-based ligands (**C** and **D**) with two modular compartments bridged by the phenolate oxygen atom. Pyrazolate-based ligands (**E**) providing a tunable metal-metal distance depending on the topology of the attached ligands.[50,51]

The development of biomimetic model systems for naturally occurring metallonucleases over the past decades has made enormous efforts in uncovering the structural and functional patterns of this class of bio-catalysts. The use of different scaffolds as well as the attachment of a variety of ligands contributed to the understanding of the structural compositions of divers active sites of enzymes. Moreover, the individual roles of each metal ion with regard to substrate binding and activation as well as their functions in the hydrolysis mechanism were accessible. However, the models predominantly lack of a comparable catalytic activity, which is still many orders of magnitudes lower with regard to their natural paragons.[51] The mimic of heterometallic complexes is particularly challenging due to the consideration of metal dependent characteristics, which are important for a site-specific coordination. In addition, the influence of small molecules, such as transient bridging or non-bridging hydroxide nucleophiles or hydrogen-bonding substituents, is still under evaluation. Nonetheless, the reduction of a large peptide to its catalytically active site enables new areas of applications.

1.3 Summary and research focus

The enumerated examples demonstrate how different types of highly specialized biomolecules are adapted in order to generate tailor-made molecules, which comprise all key characteristics. The establishment of artificial restriction enzymes, such as zinc-finger nucleases was an important milestone in the combination of the sequence-specific DNA binding ability of a zinc finger motif with the overwhelming phosphodiester cleavage ability of an otherwise non-specific restriction endonuclease. Despite their patent advantages, the attached restriction enzymes are composed of hundreds of amino acids with enormous molecular dimensions. On this account, it would be beneficial to reduce the whole enzyme to its catalytic core unit, which could subsequently be incorporated in the DNA-recognizing zinc finger protein. The work on bioinorganic model nucleases described above can help to realize this purpose by offering information on the needs for successful hydrolysis and

furthermore, by providing functional scaffolds of multinuclear metal complexes. The following sections show the transfer of bioinorganic model complexes to peptide incorporable artificial amino acids. These were incorporation into the native peptide sequence of Zif268 at distinct positions that were chosen on the basis of molecular models using already existing crystallographic data. This allowed the evaluation of the most suitable incorporation site with regard to the accessibility of the phosphodiester backbone of the DNA in order to bind and activate the substrate as well as to enable a targeted nucleophilic attack by the hydrolytically active hydroxide species. The modified zinc finger motifs were synthesized using a semi-synthetic approach combining standard solid phase peptide synthesis (SPPS) and recombinant protein expression to circumvent length restrictions by generating the full-length peptide with enhanced sequence specificity. The influence of the incorporated building blocks on the secondary structure formation of the peptide was analyzed by means of circular dichroism spectroscopy. In addition, initial determinations of the hydrolysis abilities of the building blocks were performed using activated DNA model substrate. Gel-electrophoresis experiments were used to study the binding ability as well as to test the hydrolysis capacity of the peptides towards natural DNA.

2. Modification of Zif268 with artificial dinuclear amino acids

2.1 Dinuclear building blocks suitable for peptide incorporation

As described above, dinuclear complexes of transition metals show increased hydrolysis rates in comparison to their mononuclear analogues. In addition, the ligand system has to fulfill the requirement of providing a defined metal-metal distance in order to accommodate the two metals and enable their hydrolysis ability. The ligands must define a metal dependent coordination geometry that fits to the redox properties of the metals used. Despite all preliminary considerations in terms of the aforementioned characteristics, mimicking the active site of enzymes is challenging due to the complex interplay of numerous factors defining the enzymes reactivity.[36] Nevertheless, there are several small molecule mimics of nucleases found in literature, matching the aforementioned criteria.[51] The requirement of transferability to a system that is incorporable into a peptide sequence further reduces the number of suitable systems. Hence, the most suitable model complexes for the modification of zinc fingers were found in 2,6-substituted phenols. Models of this type have been developed in a variety of ways and their hydrolysis abilities are well documented.[52] The natural counterpart of individual phenol moieties is found in the amino acid L-tyrosine which bears a phenolic side chain. The latter provides a bridging phenolate moiety between both metal ions and can be modified in ortho-position by a MANNICH reaction.[53]

2.1.1 Ligand precursors for the modification of L-tyrosine

The development of dinuclear building blocks based on tyrosine leads to the question of which ligand precursors are suitable for the synthesis of the artificial amino acids. The most valuable nuclease mimic would involve an unsymmetrically substituted phenol with different hard and soft donor ligands to provide a metal-specific coordination sphere for two different metal ions. Chelates of this type are shown in section 1.2.3 (Figure 1.12). They have different coordinating moieties, such as nitrogen-based secondary or tertiary amines, such as pyridine or methylimidazole rings, or oxygen-based moieties, such as alcohols or carboxylic acids.[51] The use of carboxylate ligands is very challenging due to the fact that they require a suitable protecting group strategy in order to synthesize a product that is also appropriate for SPPS. The presence of a primary amine and a second carboxyl group in the amino acid scaffold further complicates the synthetic procedure. Accordingly, a different orthogonal protecting group strategy had to be applied for the synthesis of the unsymmetrically substituted tyrosine residue **11** (Figure 1.13). The unsymmetrically substituted tyrosine was successfully synthesized but provided tremendously low overall yields. Due to the enormous synthetic effort, the benefit of **11** was very low what lead to the development of symmetrically substituted tyrosine derivatives. Thus, the synthesis was facilitated by the use of

symmetrically arranged *N*-coordinating methylimidazole- (**4**) and pyridine- (**8**) based ligands, which can be used in an unprotected state during the building block synthesis as well as in the SPPS approach (Figure 2.1).[54] This concession was found to be most valuable to balance the synthetic effort with marginally lowered hydrolysis rates.

The herein described building blocks are based on the tridentate precursor ligands bis(1-methylimidazole-2-yl-methyl)amine (BMIA, **4**), bis-(2-picolyl)amine (BPA, **8**) and *N*-pyridin-2-ylmethyl-glycine (**11**), which were used in a MANNICH reaction with Boc-Tyr-OMe. The reaction was adjusted to produce the di-substituted building blocks as well as their mono-substituted analogues that were needed for comparison purposes. After changing to the Fmoc-protecting group and liberating the C-terminus, the building blocks were suitable for SPPS.
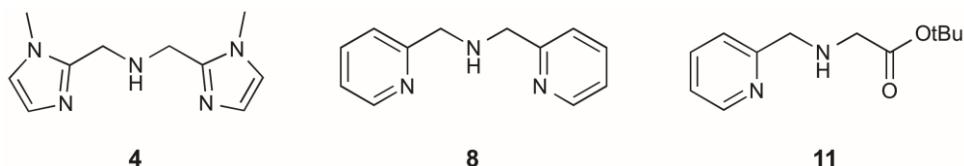
Figure 2.1 Tridentate precursor ligands synthesized for the MANNICH reaction with L-tyrosine.

2.2 Synthesis of artificial dinuclear amino acids based on L-tyrosine

The tridentate ligand precursors BMIA (**4**) was synthesized according to a customized method published by OBERHAUSEN *et al.* (Scheme 2.1) and starting with the commercially available 1-methylimidazole (**1**) which was lithiated at the C2 position using *n*-butyllithium (1.6 M) in diethyl ether at –78 °C.[55] The intermediate product was subsequently formylated by the addition of DMF and the reaction mixture was stirred over night at –60 °C resulting in the formation of 1-methylimidazole-2-carbaldehyde (**2**). The aldehyde group was transformed into the corresponding oxime derivative **3** using an ethanolic solution of hydroxylammonium chloride. Product **3** was dissolved in methanol and transferred into a high pressure vessel and catalytic amounts of palladium on active charcoal were added carefully. Hence, hydrogen gas was bubbled through the mixture for 20 min to ensure complete saturation and pre-activation of the catalyst. The vessel was clamped into a high pressure apparatus and agitated at 3 bar hydrogen pressure for 24 h. The ligand precursor BMIA (**4**) was obtained as a yellow oil after extraction from water and purification by column chromatography with an overall yield of 39%.
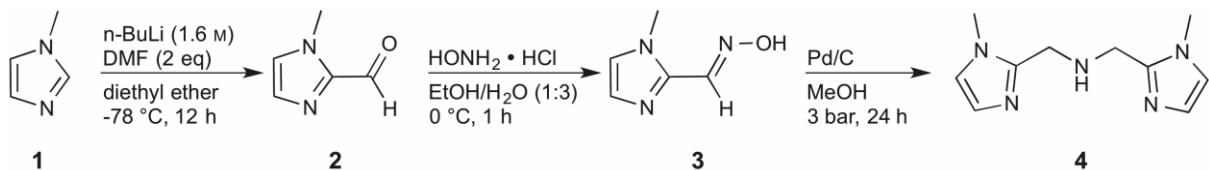
Scheme 2.1 Synthesis of the tridentate ligand precursor BMIA (**4**).[55]

The synthesis of the structurally related BPA precursor ligand **8** was performed in two different ways. The first shown in Scheme 1.2 was based on a reductive amination reaction between picolinaldehyde (**5**) and 2-picolyamine (**6**) as reported by the LIU group.[56] Equimolar amounts of both compounds were dissolved in methanol, whereby a nucleophilic attack of the amine on the aldehyde species occurred that generated the imine compound. After the validation of complete consumption of the reactants by TLC, the imine was reduced by either 5 equivalents of $NaBH_4$ or 5 equivalents of $NaCNBH_3$. Both attempts successfully generated the secondary amine as main product (98%) but, nevertheless, minor amounts of the non-reduced imine species (2%) remained after the reactions as confirmed by NMR-spectroscopy and high-resolution mass spectrometry. Repeated additions of the reducing agents could not affect the overall presence of the byproduct. The similar retardation factors of both compounds led to separation problems during column-chromatographic workup.
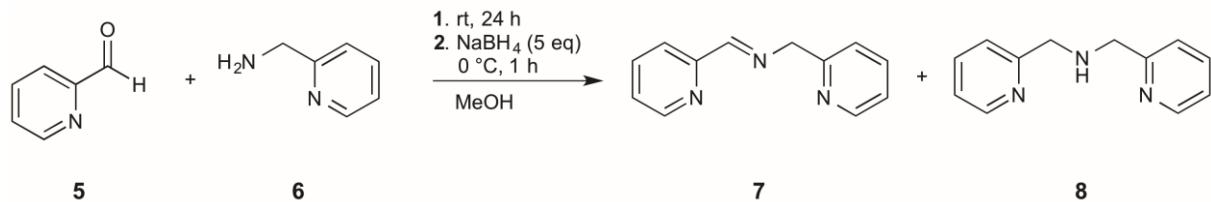
Scheme 2.2 Literature synthesis of the tridentate ligand precursor BPA (**8**).[56]

Hence, the applicability of the aforementioned method used for the preparation of the BMIA ligand **4** was evaluated for the present system (Scheme 2.3). Therefore, picolinaldehyde (**5**) was dissolved in ethanol/water (1:4) and hydroxylamine hydrochloride was added to generate the oxime compound picolinaldehyde oxime (**9**). The latter was exposed to 3 bar hydrogen pressure under the attendance of a palladium catalyst for 24 h. Excess water was added to the reaction mixture, which was filtrated through Celite to remove the catalyst on activated charcoal. The crude product was extracted with ethyl acetate and purified by flush column chromatography to obtain the BPA ligand **8** in a final yield of 78%. In contrast to the literature preparation described above, the hydrogenation of the oxime derivative at elevated pressures significantly facilitated the reaction workup due to the almost quantitative conversion of the educt. Moreover, the use of high hydrogen pressures tremendously

reduced the reaction time as confirmed by an analogues experiment at atmospheric pressure (24 h in comparison to 7 d).
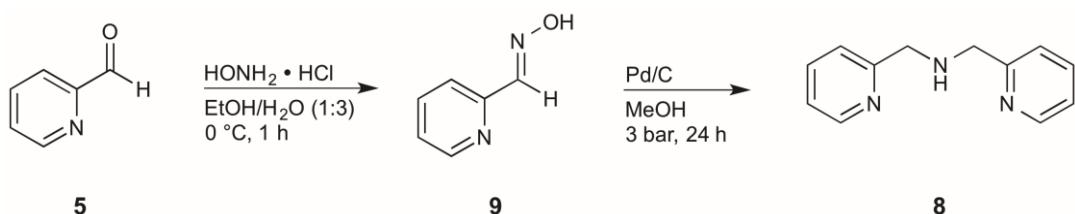
Scheme 2.3 Modified synthesis of the tridentate ligand precursor **8**.

As described earlier, unsymmetrically 2,6-substituted phenols were found to exhibit an enhanced hydrolysis ability towards phosphodiesters when compared to their symmetrically substituted analogues. For this purpose, the synthesis of a tridentate precursor ligand was evaluated, which has a metal coordination site with different hard and soft donor ligands (cf. Figure 1.11). The ligands should predetermine the coordination preferences of different metal ions in order to result a heterodinuclear complex with distinct metal sites. Thus, *tert*-butyl-(pyridin-2-ylmethyl)glycine (**11**) was used as a precursor ligand with a metal-coordinating carboxylate group.[57] The presence of an additional carboxylic acid placed enormous demands on the orthogonal protecting group strategy since tyrosine must also be protected in the subsequent reaction with the ligand precursor. In addition, the protecting group must tolerate the conditions applied during the SPPS, without being cleaved off that would otherwise result in an uncontrolled peptide elongation at the unprotected carboxyl group. For this reason, the carboxylic acid was protected as *tert*-butyl ester, which is fully orthogonal to the base labile Fmoc protecting group. The synthesis of ligand **11** was taken from literature and performed under slightly modified conditions (Scheme 2.4).[57] Picolinaldehyde (**5**) and glycine *tert*-butyl ester hydrochloride (**10**) were dissolved in methanol, whereby the Schiff base was formed. The imine was reduced by the stepwise addition of 3 equivalents of NaBH$_4$ at 0 °C over a period of 2 h. After purification of the crude product by column chromatography on silica gel, the precursor ligand **11** was obtained in a total yield of 44%.
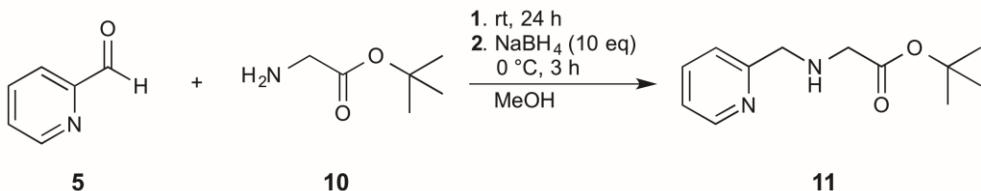
Scheme 2.4 Synthesis of the asymmetric ligand precursor *tert*-butyl-(pyridin-2-ylmethyl)glycine (**11**).[57]

2.2.1 Synthesis of dinuclear metal complexes based on tyrosine

The general approach for the preparation of the symmetrically substituted dinuclear building blocks is shown in Scheme 2.5.
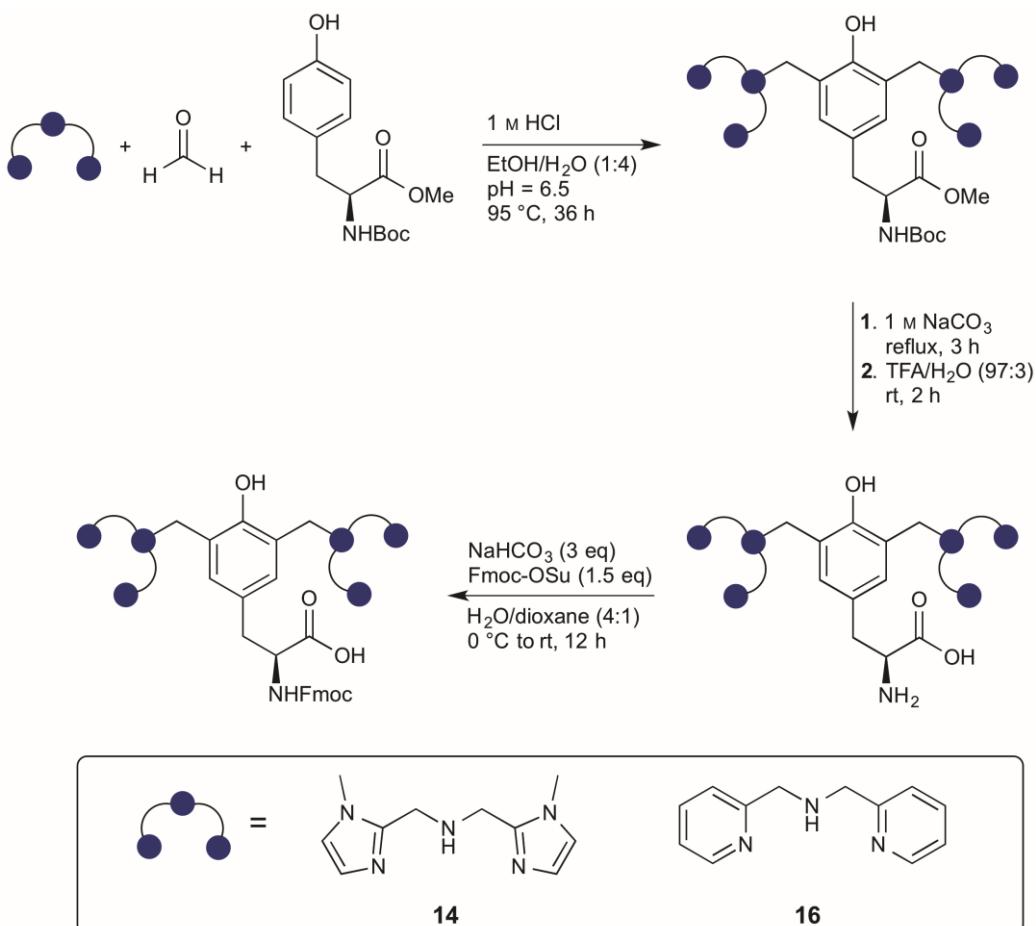
Scheme 2.5 General scheme for the preparation of the dinuclear building blocks **14** and **16** by a MANNICH reaction of Boc-Tyr-OMe and the synthesized tridentate ligand precursors **4** and **8**.

Accordingly, the aforementioned ligand precursors, paraformaldehyde and Boc-Tyr-OMe were used in a MANNICH-reaction. The highest reaction yields were obtained using a slightly modified literature protocol.[58] In contrast to the published procedure, in which all components were simultaneously added, it turned out to be beneficial to first generate the Schiff base before adding the amino acid. Hence, paraformaldehyde and the ligand precursor were suspended in ethanol/water (1:4) and stirred at 60 °C for 90 min. MANNICH reactions are known to have individual pH optima, which are dependent on the amines used and the CH-acidic compounds.[59] Due to resonance stabilization, the phenolic side chain of tyrosine almost exclusively exists in the enol form.[60] As reported by MINAKAWA, the optimum conditions for the electrophilic substitution request slightly acidic conditions (pH 5.5 – 6.5) to efficiently perform the MANNICH reaction with tyrosine derivatives.[59] However, this was restricted by the Boc protecting group, which might be cleaved by applying acidic conditions in combination with elevated temperatures for 36 Hence, Boc-Tyr-OMe was dissolved in

ethanol/water (1:4) and 2.5 equivalents of the preformed Schiff-base cocktail were added. The pH was adjusted to approximately 6.5 – 7.0 by the addition of 1 M HCl (aq.) and the reaction mixture was stirred at 95 °C for 36 h.

It has been observed that minor amounts of the undesired mono-substituted byproduct were present after the reaction workup. Due to similarly low retardation factors of both compounds, an effective chromatographic separation and purification method could not be applied. Therefore, aliquots were taken from the reaction mixture during the synthesis and analyzed on the presence of mono-substituted products by means of ESI mass spectrometry. In case of detecting the byproduct, further addition of the ligand precursor or an extension of the reaction time was applied. After the reaction has gone to completion, the crude product was extracted with chloroform and purified by RP-flash-column chromatography with water/ethanol (4:1) as eluting system.

In order to use the artificial amino acid in SPPS, a change of the N-terminal Boc protecting groups to Fmoc was necessary and, in addition, the C-terminal methyl ester had to be cleaved off. In order to generate the free carboxylic acid, the compounds were dissolved in methanol and an aqueous solution of sodium carbonate (1 M) was added in excess and refluxed until complete deprotection was observed by TLC. For the deprotection of the Boc group, trifluoroacetic acid was added to the residues and the mixture was agitated for 2 h. The volatile components were removed in a nitrogen stream and the crude products were precipitated upon the addition of ice-cold diethyl ether. The completely unprotected amino acid was dissolved in water and sodium bicarbonate (3 eq) was added. The N-terminus was Fmoc protected by the addition of Fmoc-succinimide (2.2 eq) as a solution of *para*-dioxane and the reaction mixture was subsequently stirred at room temperature overnight.[61] After the extraction of the crude product from the aqueous solution with ethyl acetate, the final products were purified by the previously described RP-flash-column chromatographic method to obtain the building blocks suitable for SPPS.

In a similar approach, the mono-substituted analogues (Figure 2.2) were synthesized in order to compare their hydrolysis rates with the di-substituted building blocks as described in section 2.3. The MANNICH conditions were slightly changed to exclude the formation of the di-substituted products. This was achieved by decreasing the amount of the ligand precursor to 0.95 equivalents with regard to the tyrosine concentration. This proceeding prevented on the one hand the formation of the di-substituted product, which was nearly inseparable from the mono-substituted compound and, on the other hand, it facilitated the purification due to different retardation factors with respect to the remaining Boc-Tyr-OMe. An exchange of the Boc protecting group for Fmoc and the removal of the C-terminal methyl ester (OMe) was not

performed due to the fact that these building blocks should not be used in the SPPS approach.
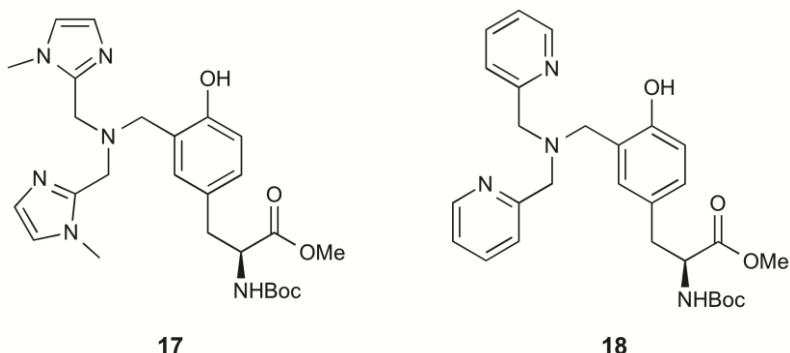
Figure 2.2 Synthesized mononuclear BMIA (**17**) and BPA (**18**) building blocks for comparative studies in kinetic experiments with their dinuclear analogues and DNA-model substrate.

As mentioned at the beginning, the synthesis of the unsymmetrically substituted building block was a big challenge due to the additionally required protecting group at the *tert*-butyl-(pyridin-2-ylmethyl)glycine (**11**) ligand precursor. The protecting group had to be fully orthogonal to the Fmoc group exclusively allowing for the use of acid labile protecting groups. Furthermore, it had to tolerate the applied conditions of the MANNICH reaction with regard to the acidic pH and longtime reflux. Thus, the Boc/OMe protected tyrosine used in the previous reactions was unsuitable due to the subsequent acidic deprotection of the Boc group. This would also cause a loss of the *tert*-butyl group of the ligand which could not be reintroduced in a region-selective manner at this position. Hence, Fmoc-Tyr-Bn (**19**) was used because this residue circumvents the need of Boc deprotection and, in addition, the benzyl group can be selectively deprotected by hydrogenation using palladium on charcoal. In order to generate the unsymmetrically modified building block, the BPA ligand **8** was first attached to tyrosine in the manner described above using 0.95 equivalents of the BPA compound (Scheme 2.6). The Schiff base was generated by the addition paraformaldehyde in EtOH/water (1:4) at 65 °C for 2 h. The pH of the reaction mixture was adjusted to approximately 6.5 because the ligand shows enhanced basicity and tends to deprotect the Fmoc group of tyrosine upon addition. Fmoc-Tyr-Bn (**19**) was added and the reaction mixture was refluxed for 36 h, whereby the mono-substituted building block **20** was generated in yields of 16%. Afterwards, the Schiff base cocktail of the reaction between ligand **11** and paraformaldehyde was added to the intermediate product in order to be attached to the remaining ortho-position of tyrosine. The reaction process was monitored by ESI-MS that revealed the formation of several byproducts. In addition, the second substitution is known to be less effective due to the deactivation of the phenol ring by the first substituent.[62] The additional aromaticity as well as the increased steric demand of the Fmoc and benzyl

protecting groups compared to the previously used Boc and OMe groups might explain the high content of side products. The unsymmetrically substituted building block **21** was obtained in 2% overall yield as confirmed by RP-HPLC. Besides, HPLC purification was the only way to isolate the target product because the high content of byproducts as well as their low retardation factors made it impossible to apply column chromatography. This fact as well as the generally difficult synthesis of **21** enormously decreased its suitability even though the hydrolysis ability might be potentially higher. Thus, future experiments were limited to the use of the successfully synthesized symmetrically modified building blocks **14** and **16**.
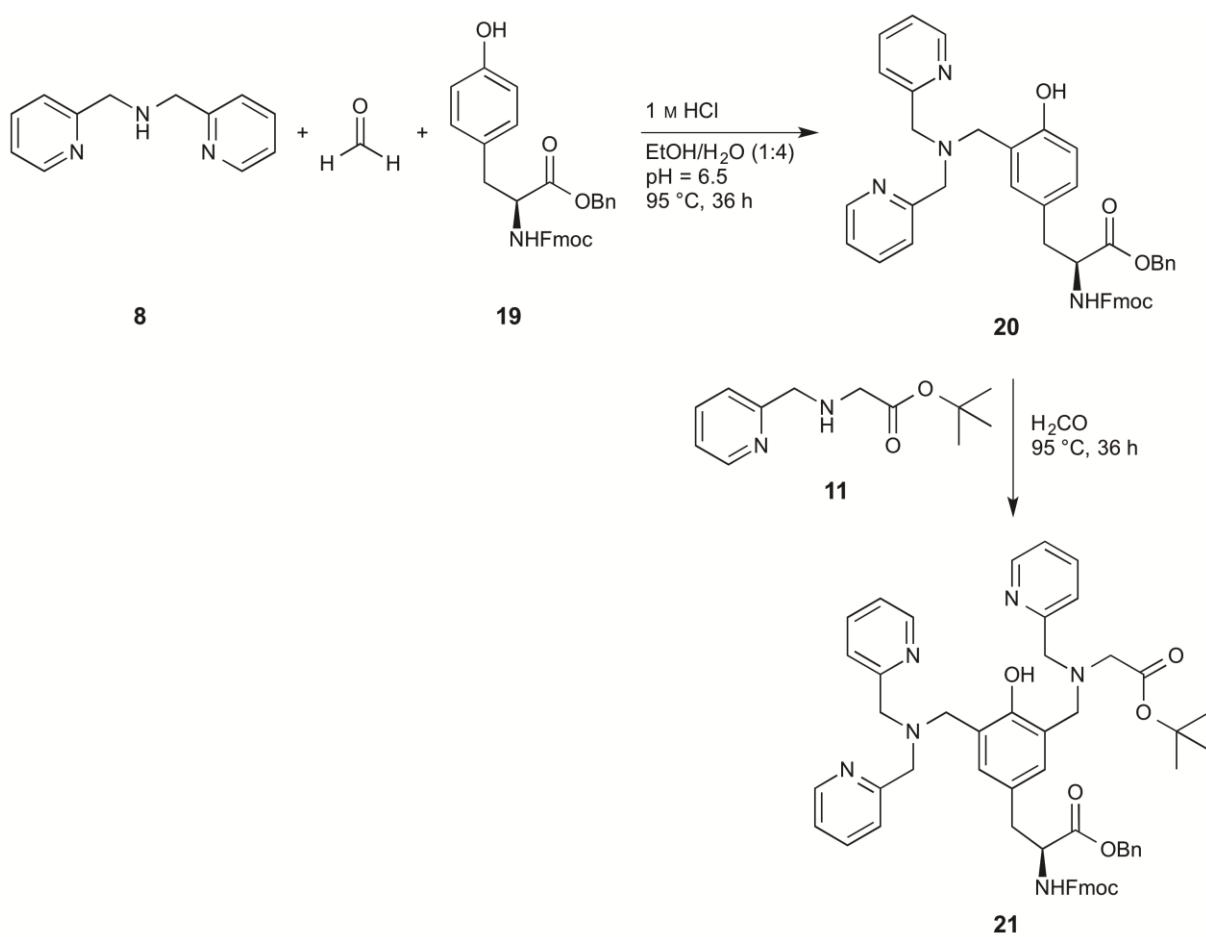
Scheme 2.6 Synthetic route for the preparation of the asymmetrically substituted building block **21** by a MANNICH reaction of Fmoc-Tyr-OBn (**19**) with the ligand precursors **8** and **11**.

2.3 Evaluation of the hydrolysis ability of the building blocks towards DNA model substrate

For the examination of the potential ability of a molecule or a complex to hydrolyze phosphodiesters, DNA-model substrates are often used for initial experiments.[63] The latter are inexpensive and can be used in higher amounts compared to natural DNA, which facilitates the handling and makes different kinetic analyses accessible. One of the most prominent examples of DNA-mimicking substrates is bis(4-nitrophenyl) phosphate (BNPP).[64] BNPP is an activated phosphodiester with two excellent *p*-nitrophenol leaving groups resulting in an increased reactivity compared to natural DNA. It offers similar prerequisites in terms of substrate binding but, in addition, it indicates the cleavage event by releasing fragments, which have a high UV-absorbance (Figure 2.3).[65] After hydrolysis, the BNPP molecule is fragmenting into a *p*-nitrophenylphosphate molecule and a *p*-nitrophenolat species which can be detected easily by means of UV-vis absorption spectroscopy due to a strong absorbance at 412 nm ($\varepsilon_{412\,nm} = 18.700\ M^{-1}\ cm^{-1}$). In aqueous solutions, the *p*-nitrophenolat cleavage product is in a pH dependent equilibrium with *p*-nitrophenol, which absorbs at 320 nm ($\varepsilon_{320} = 10.000\ M^{-1}\ cm^{-1}$).[66] With increasing pH, the equilibrium is shifted to the *p*-nitrophenolate species that must be taken into account when calculating the total concentration of the hydrolysis product for the pH-dependent determination of rate constants.
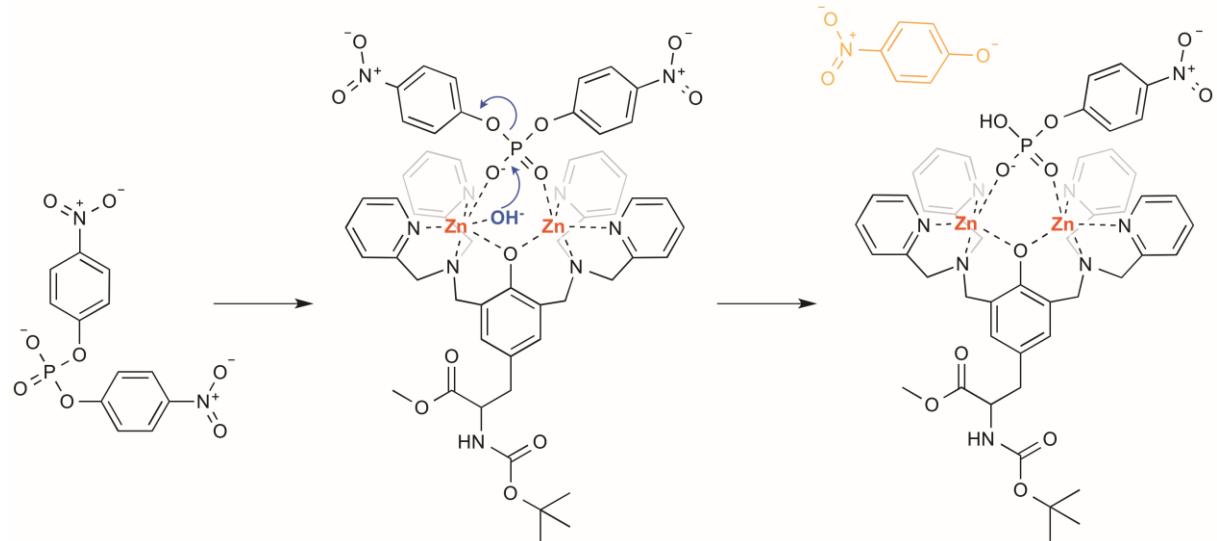
Figure 2.3 Suggested mechanism of BNPP hydrolysis by the dinuclear Zn(II) complex of building block **15**. After the coordination of the phosphate group of the substrate by the metal site, a hydrolytically active hydroxide ion attacks the phosphorous atom, which releases a strong UV-absorbing *p*-nitrophenolate species.

The hydrolytic activity for certain metal complexes is pH dependent and follows a bell-shaped curve with a maximum at the optimum pH.[67] For instance, the ideal pH for the hydrolysis reaction of the purple acid phosphatases with an active site either composed of [Fe(III)/Fe(II)] (mammalians) or [Fe(III)/Zn(II)] (plants) is in the range of 4.9 – 6.0.[68]

Consequently, an alteration of the pH away from the optimum likewise decreases the rate constant of the BNPP hydrolysis. Moreover, the zinc-coordinating ability, and therefore, the DNA-binding ability of Cys_2His_2-type zinc fingers is also pH dependent. It is stated that the pH range for obtaining correctly folded zinc fingers with intact DNA-binding ability lies between 6 – 8.[69] This mainly results from the protonation and deprotonation state of the histidine and cysteine residues being involved in metal binding. Whereas thiols are known to coordinate zinc even in the protonated state, protonated histidines progressively lose their binding capacity and, subsequently, release Zn(II) from the zinc finger complex. Moreover, the pK_a of the thiol group of cysteine is 8.4, whereas the pK_a of the imidazole side chain of histidine is 6.1.[70] Hence, under neutral conditions, histidine is in the deprotonated state and its free electron pair is available for zinc coordination. In contrast, for efficient metal binding, the thiol moiety should also be deprotonated that is supposed to be additionally achieved by the Zn(II) ion upon the formation of the zinc-finger metal complex.[69] A too strong increase of the pH to alkaline conditions is also disadvantageous for the secondary structure of the zinc finger, which is stabilized by intramolecular hydrogen bonds that would disappear otherwise. As a result of the enumerated characteristics, the applicable pH range to ensure protein functionality is rather narrow. With respect to these facts, the hydrolytic capacities of the building blocks were evaluated at pH = 7.8. This value was found to be most suitable for ensuring proper peptide folding as well as DNA binding of the zinc fingers. It was used in different other approaches allowing for the production of transferable results.[71]

In addition, the same considerations are valid for the use of an appropriate solvent system. For accurate peptide folding, a buffered aqueous solution with distinct ionic strength needs to be applied. Therefore, the hydrolysis experiments with BNPP and the individual building blocks were conducted under pseudo-aqueous conditions. A 1:1 ratio (v/v) of the buffer solution and acetonitrile has been chosen in order to increase the solubility of the organic compounds but also to have the opportunity to adjust the ionic strength of the mixture. This experimental set-up provides transferable results with regard to the aqueous-only environment in which the peptides and DNA are soluble.

For the buffer preparation, mainly sulfonate-based buffering agents, such as HEPES (useful pH range: 6.8 – 8.2), were considered because they cover a wide pH range and do not show any tendency to chelate metal ions in contrast to acetate or phosphate-based buffers.[72] This is a vital aspect with regard to a possible decrease of the catalytic activity due to the occupation of the substrate-binding site by a buffering compound at the active site of the metal complex. The ionic strength of the HEPES buffer was adjusted to $I = 0.1$ with $NaClO_4$ and perchloric acid and sodium hydroxide was used to adjust the pH.

The temperature dependence of the hydrolysis rate was another important factor to consider. Hydrolysis rates rely on the kinetic molecular theory including the Arrhenius equation suggesting that rate constants exponentially vary with the reciprocal of the absolute temperature.[73] This strong dependence gives rise to the assumption that it would be beneficial to elevate the temperature in order to increase cleavage rates. However, increased temperatures would also cause peptide and DNA degradation and would inhibit the complex formation of the latter. Thus, the temperature was kept at 37 °C for all experiments to apply physiological conditions and to generate reliable benchmarks for the experiments with natural DNA that will be described in section 2.7.

The BNPP hydrolysis studies were performed with the fully protected (Boc/OMe) dinuclear building blocks **13** and **15** and were compared with their mononuclear analogues **17** and **18**. This was motivated by the concern to determine their respective hydrolysis rates in order to acquire information about the incorporation benefit of the different building blocks into the zinc finger sequence. However, the increased hydrolytic activity of dinuclear complexes over mononuclear complexes was excessively documented and had been attributed to the cooperative effect of two metal ions, which are in close proximity to each other.[36] This provides the benefits of highly improved substrate activation and stabilization of the transition state due to geometry-optimized electronic interactions between the metal ions and the phosphate oxygen atoms. According to a study of IRANZO *et al*, who examined the hydrolysis capacities of mononuclear and dinuclear triazacyclononan (tacn)-based Zn(II) complexes towards model substrates, the dinuclear complex showed a 10-fold increased rate acceleration.[74] This was led back to the better transition state stabilization, which was found to be 9.3 kcal mol^{-1} and furthermore, dinuclear building blocks showed an increased ability to provide the hydrolytically active hydroxide species.

2.3.1 Kinetics of BNPP hydrolysis

To avoid the aforementioned negative effects on peptide folding and DNA binding, the BNPP hydrolysis studies were performed at the most suitable pH of 7.8. On the one hand, this value is most frequently used for zinc-finger DNA binding studies found in the literature and, on the other hand, it allows the comparison of the obtained production data with already reported findings of similar phenolate-based building blocks. Prior to the experiments, the building blocks (2.5 mM) and ZnCl$_2$ (5.1 mM for **13** and **15**, 2.6 mM for **17** and **18**) were dissolved in a buffer solution (600 μL, 25 mM HEPES, I = 0.1) and the pH was adjusted to 7.74. All mixtures were incubated at 37 °C over night. Subsequently, the mixtures were transferred into a stirrable cuvette and 600 μL of a solution of BNPP (2.5 mM) in acetonitrile was added. Upon the addition of BNPP, the hydrolysis reaction was initiated and its progress was monitored by means of UV-vis spectroscopy in a time-dependent manner. The

concentrations of the obtained cleavage products were calculated using LAMBERT-BEER law (Equation 2.1) with the molar extinction coefficient of the *p*-nitrophenolate ($\varepsilon_{412\ nm}$ = 14.340 M^{-1} cm^{-1}) species.[75] The possibility of non-hydrolytic BNPP decay under the applied conditions was examined beforehand by measuring a sample without the active metal complexes. It was found that over the contemplated time, the auto decay of BNPP was negligible. This made a correction of the recorded spectra redundant.

$$A = \varepsilon \cdot c \cdot l \tag{2.1}$$

(A = absorbance, ε = molar extinction coefficient, c = concentration and l = path length)

The data obtained from the spectroscopic measurements were processed in Equation 2.1 and the calculated concentrations were mathematically treated under the assumption of a pseudo-first-order rate law. This was appropriate because the hydrolytically active metal complex was used in excess and can therefore be regarded as constant over the contemplated reaction time what leads to Equation 2.2.

$$-\frac{dc(BNPP)}{dt} = \frac{dc(hydrolysis\ product)}{dt} = k[complex][substrate] = k_{obs}[substrate] \tag{2.2}$$

The equation can be further facilitated under the assumption that only values for initial rates smaller than 6% were used (initial slope method). This proceeding was also appropriate with respect to the formation of exactly one *p*-nitrophenol species and one *p*-nitrophenylphosphate species per hydrolyzed BNPP molecule. The latter can subsequently release a second *p*-nitrophenol moiety, which would negatively affect the concentration determination to provide false rate constants. By only using the initial values, the second release is assumed to be rather low and can therefore be neglected. Consequently, the substrate concentration can be seen as constant (Equation 2.3).

$$k_{obs} = \frac{dc(hydrolysis\ product)}{[substrate]_0\ dt} \tag{2.3}$$

The observed rate constants k_{obs} for each of the examined complexes were obtained by plotting the natural logarithm of the hydrolysis product concentration against the time. The slope of the thus obtained straight line equals the corresponding k_{obs} (Equation 2.4).

$$\ln(hydrolysis\ product) = \ln[substrate]_0 - k_{obs} \cdot t \tag{2.4}$$

The determined k_{obs} values for the examined building blocks are summarized in Table 2.1.

Building Block	k_{obs} [s^{-1}] (pH 7.74)
13	1.05×10^{-5}
17	1.84×10^{-6}
15	1.18×10^{-5}
18	2.21×10^{-6}

Table 2.1 Determined k_{obs} values for the BNPP cleavage reaction of the mononuclear (**17** and **18**) and dinuclear (**13** and **15**) building blocks.

It becomes apparent that the overall hydrolysis abilities of the mononuclear complexes were lower compared to their dinuclear analogues. In fact, under the applied conditions the determined k_{obs} values were approximately 5.5-fold decreased when compared to the dinuclear complexes. Considering that the aforementioned studies already revealed poor hydrolysis capacities for mononuclear complexes due to the absence of a cooperative effect with a second metal ion, this was not surprising. Thus, substrate activation and stabilization of the transition state were much less pronounced making an effective hydrolysis difficult.

However, the dinuclear building blocks **13** and **15** showed an increased BNPP cleavage ability under the applied conditions with comparable k_{obs} values for the both structurally related compounds (Table 2.1). The experiments clearly demonstrated, that the building blocks fulfill the two major criteria for phosphodiester hydrolysis. Firstly, they were able to activate the BNPP substrate upon binding of the dinuclear metal complex to the phosphate group. Secondly, the hydrolytically active hydroxide species was provided by the metal site that finally triggered the cleavage and released the *p*-nitrophenol species. This was further demonstrated by studying the turnover numbers (TON) for the dinuclear complexes (50 nM) with BNPP (25 μM) under similar conditions as mentioned above (Figure 2.4). A plot of the time against the absorbance at 412 nm revealed almost comparable TONs for both building blocks. This indicates a similar behavior in terms of substrate cleavage and product release over the contemplated reaction time.
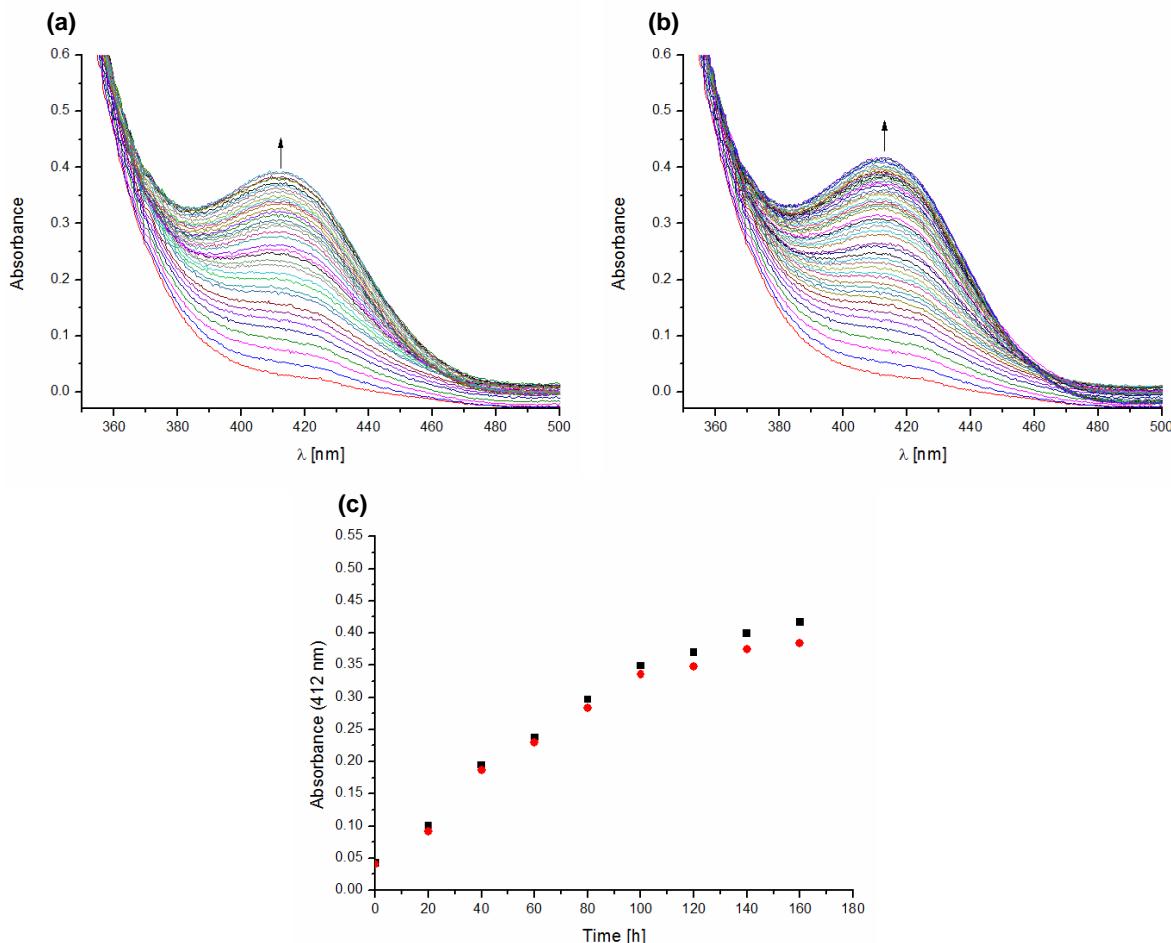
Figure 2.4 UV-vis spectra of the BNPP hydrolysis with **13** (**a**) and **15** (**b**) taken at different time points after complex addition. (**c**) Plot of the time against the *p*-nitrophenolate band (412 nm) for **15** (black squares) and **13** (red dots).

It is to note, that the here presented values were produced with an activated DNA model-substrate and the applied pH was limited to 7.74. For a more profound analysis of the individual building blocks, the experiments also need to include a pH and concentration dependent characterization for the determination of a second-order rate constant for the BNPP hydrolysis. This was neglected with respect to the restrictions during the work with peptide incorporated building blocks and natural DNA, which narrow the range of applicability. Thus, the results with natural DNA might differ due to various other factors that come into play. These include, for instance, the accessibility of the phosphodiester backbone of the DNA by the zinc-finger-incorporated building block when bound to the latter. Whereas the individual building blocks and the BNPP substrate can freely move in solution guaranteeing enhanced substrate binding and subsequent activation, the DNA-bound zinc finger is much more restricted in its movement. Moreover, the substrate release, as demonstrated by the TONs, is limited due to the fixation of the zinc finger to exactly one dsDNA moiety. Consequently, the obtained findings provide a good initial suggestion for a possible cleavage of natural DNA by the building blocks. However, the actual hydrolysis ability towards natural DNA is discussed in detail in section 2.7.

2.4 Incorporation of the dinuclear building blocks into the sequence of Zf3

The semi-synthetic methodology for the development of modified zinc finger proteins included the bacterial expression of the first two zinc-finger sub-units (Zf12) of Zif268 and the conventional peptide synthesis of the third sub-unit (Zf3). Zf3 was thereby modified by the incorporation of the previously described artificial amino acids for site-specific DNA hydrolysis. Zf12 increases the range of DNA recognition and allows sequence-specific binding of the peptide to its DNA operator sequence.[69]

The incorporation of artificial amino acids into a highly specialized molecule, such as the zinc finger, bears several risks in terms of preservation of function or loss of the DNA-binding ability by the alteration of the secondary structure. Furthermore, the incorporation site must be appropriate for the building blocks with regard to their distances to the phosphate backbone of the DNA. Hence, the metal complex must be able to reach the substrate in order to activate it and to transfer the catalytically active hydroxide species from the metal ion to the phosphorus atom. These requirements must be taken carefully into account for the selection of suitable incorporation sites for the artificial amino acids.

2.4.1 Characteristics of the peptide sequence of Zf3

The third zinc finger of Zif268 is composed of 26 amino acids (Figure 2.5).[10] As discussed earlier, most of these residues are highly conserved, and therefore, cannot be exchanged or altered. This is particularly true for the two histidine and cysteine residues (Figure 2.5, red), which are part of the zinc-coordinating site of the peptide, and therefore, are essential for the generation of the ββα-structure. This also includes the amino acids that compose the hydrophobic core (Figure 2.5, blue), whose integrity is also vital for structural reasons.[9] Moreover, each zinc finger sub-unit has three conserved residues being directly involved in nucleobase binding (Figure 2.5, green). An exchange of those amino acids would lead at best to a decrease in DNA recognition. However, it is most likely that the complete binding is inhibited.[76]
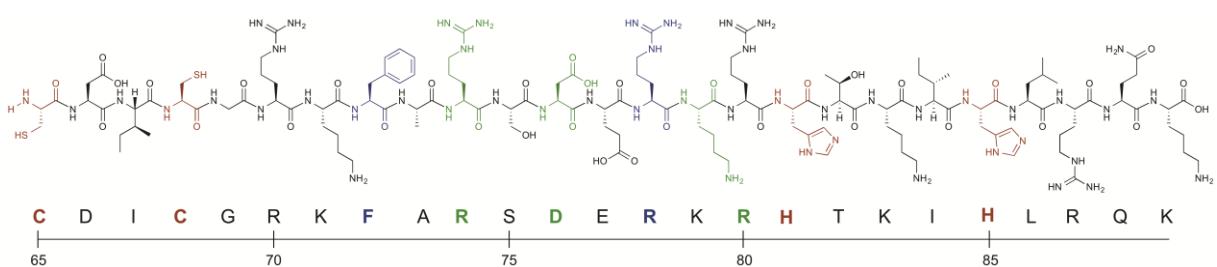
Figure 2.5 Amino acid sequence of Zf3 with metal-coordinating residues highlighted in red, hydrophobic residues being part of the hydrophobic core are highlighted in blue and residues, which make direct nucleobase contact are colored in green.[10]

2. Modification of Zif268 with artificial dinuclear amino acids

Besides the aforementioned amino acids, which mainly contribute to structural or functional properties in terms of Zn(II) and DNA binding, another vital aspect is the close proximity of the building blocks to the phosphate backbone. Only if this condition is met, an activation of the phosphodiester bond is feasible and hydrolysis can occur. This circumstance further lowers the amount of possible incorporation sites. By having a closer look into the crystal structure of Zif268 published by PABO, two positions in the sequence of Zf3 fulfill this requirement.[10] On the one hand, the arginine residue at position 70, which is known to have only an additional phosphate contact without being further involved in sequence-specific nucleobase binding (Figure 2.5 and 2.6a). And, on the other hand, the serine residue at position 75, which is not conserved but which is situated in close proximity to the phosphodiester backbone (Figure 2.5 and 26b). The serine residue is located at the beginning of the loop region connecting the α-helix and the β-hairpin. Therefore, the exchange of this amino acid for the building blocks would enable the possibility to reach and cleave the operator strand of the DNA. In contrast to that, the arginine at position 70 is part of the subsequent β-strand and the additional phosphate contact is found to occur with the complementary, non-binding strand of the DNA. This offers the possibility to reach the two opposite DNA strands and induce hydrolysis at different positions. In summary, the most promising residues of Zf3, which are suitable for the exchange for the binuclear building blocks in terms of the aforementioned reasons, are the arginine residue at position 70 and the serine residue at position 75.
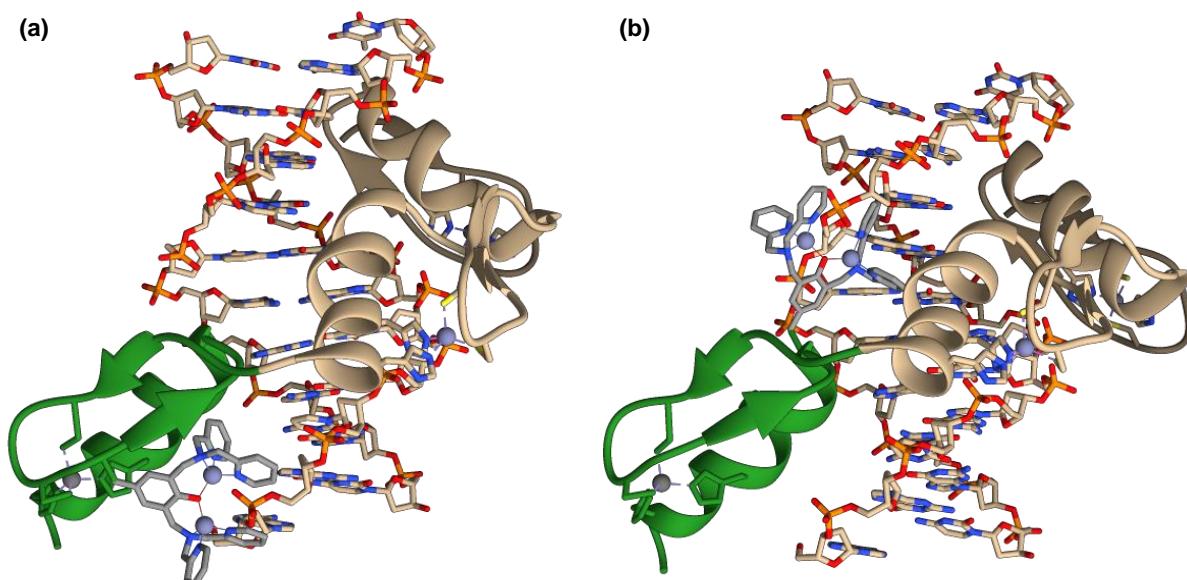
Figure 2.6 Molecular models of the zinc finger domain (green) with building block **16** (gray) incorporated at the arginine position 75 **(a)** and at the serine position 70 **(b)**. The models are based on the crystal structure of a Zif268/DNA complex (PDB code 1AAY) and were modified with UCSF Chimera.

2.4.2 Solid phase peptide synthesis of modified Zf3 constructs

With the determination of two incorporation sites, modified Zf3 peptides were synthesized by means of solid phase peptide synthesis (SPPS), including the artificial amino acids. The SPPS approach dates back to the work of MERRIFIELD in 1963.[77] It relies on the use of an insoluble solid support (resin) that is modified with an acid labile linker to which the C-terminus of the first amino acid, bearing a base labile N-terminal Fmoc protecting group, is attached. In contrast to protein biosynthesis, in which ribosomes elongate the sequence from the N-terminus to the C-terminus, the elongation in the SPPS approach is performed in opposite direction. The procedure starts with the removal of the resins N-terminal Fmoc protecting group by the addition of an organic base, such as piperidine or piperazine. The first amino acid is coupled to the resin by means of a suitable coupling cocktail containing reagents for the activation of the terminal carboxyl group by forming an active ester (Oxyma Pure®/DIC, HATU/HOAt or HBTU/HOBt). The process of Fmoc deprotection and amino acid coupling is repeated until the desired sequence is generated on the solid support (Figure 2.7). The major advantage of SPPS is the facilitated washing procedure after each reaction step.[78] The soluble side products can be easily separated, whereas the growing peptide attached to the resin remains in the filter tube. This also allows the use of all coupling reagents in excess that further increases the crude purity of the peptide and avoids the formation of deletion or truncated sequences. Finally, the peptide is cleaved from the resin by the addition of a strong acid, such as trifluoroacetic acid (TFA), which simultaneously removes all acid-labile side chain protecting groups.
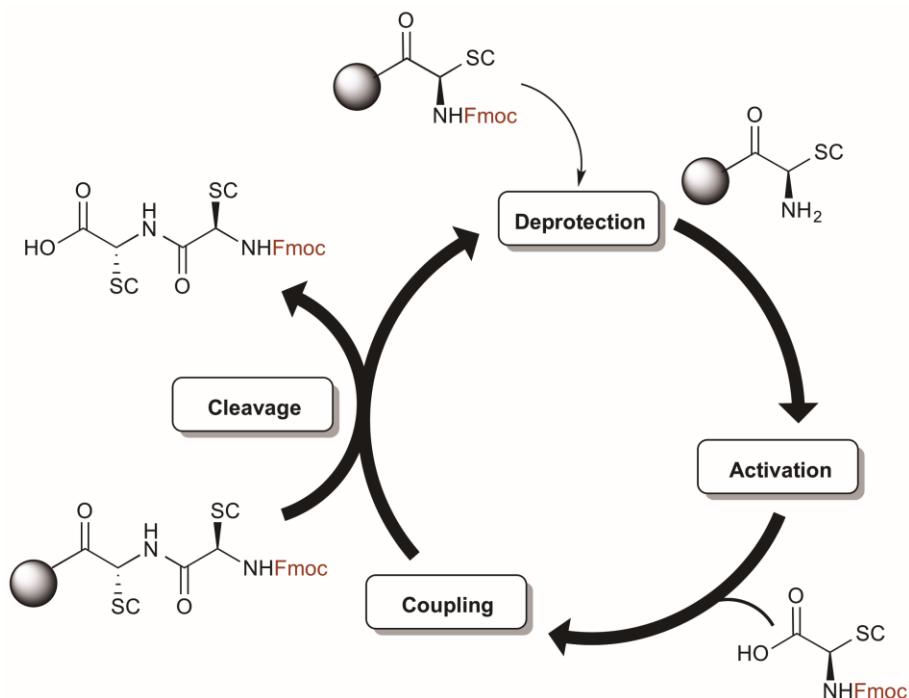
Figure 2.7 Schematic representation of the solid-phase peptide synthesis (SPPS) cycle (SC = amino acid side chain).

The zinc finger constructs in the present study were synthesized by means of the Fmoc/*tert*-butyl strategy described above using a Fmoc-L-lysine(Boc)-OH pre-loaded Wang-resin. The peptides were either synthesized using a microwave-assisted peptide synthesizer (Liberty Blue or Liberty 12 from CEM), or microwave-assisted manual peptide synthesis when particular emphasis was put on the incorporation of the bulky building blocks was applied. Standard amino acids were coupled using a coupling cocktail containing HOBt/HBTU or HOAt/HATU with DIPEA as activator base or applying the system Oxyma Pure®/DIC as indicated in the experimental section. Special attention was payed on the incorporation of the unnatural amino acids. The increased steric demand of the building blocks made it necessary to use a slightly modified coupling protocol, which was applied by manual microwave peptide synthesis. In this sense, the coupling times for the building blocks as well as for the subsequent amino acid were extended from 4 min (90 °C, 25 W) to 2 x 20 min (65 °C, 12 W), respectively. In addition, the coupling mixture for the concerned amino acids was adapted from the literature, which recommended the use of a solution containing DIPEA as activator base and PyBOP/HOBt as coupling reagents.[79] PyBOP is a very potent peptide coupling reagent that reduces racemization during elongated coupling times for critical amino acids. It is commonly used for challenging cyclization reactions and for the coupling of hindered amino acids, such as 2-aminoisobutyric acid (Aib) or bulky amino acids, as for instance phosphorylated threonine (Fmoc-Thr(PO$_3$BzlH)-OH).[80] Peptide synthesis was completed by manual coupling of the remaining amino acids under conventional condition with microwave support. The zinc finger mutants were cleaved from the resin by a mixture of TFA/EDT/H$_2$O/TIS (94:2.5:2.5:1) and purified by RP-HPLC. The target peptides could be isolated in yields ranging between 15 – 38%. However, the HPLC chromatograms revealed the formation of several deletion and truncation sequences derived from the incorporation site of the building blocks as confirmed by mass spectrometry of the isolated compounds. It can therefore be assumed that the bulky building blocks negatively affect the subsequent peptide couplings. This was slightly improved by applying a double coupling procedure at decreased temperatures (50 °C) and extended reaction time (2 x 15 min) for the remaining amino acids.

Nonetheless, the dinuclear building blocks **14** and **16** were successfully incorporated into the native sequence of Zf3 at two different positions (Arg70 and Ser75, Figure 2.8).
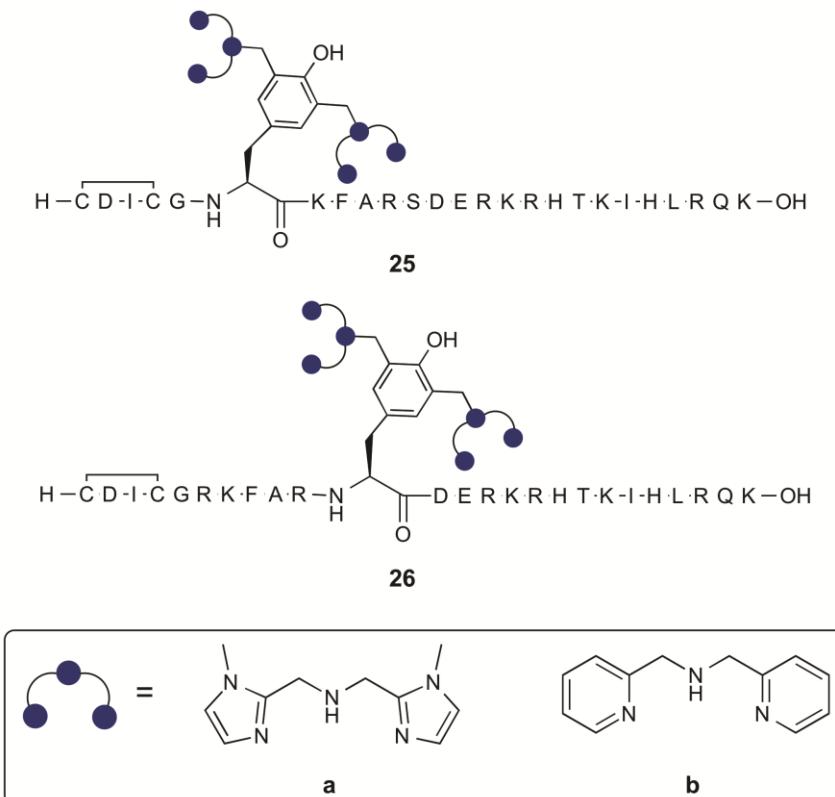
Figure 2.8 Zinc finger mutants synthesized in this work. The BMIA modified building block (**a**) and the BPA modified building block (**b**) were incorporated at the arginine position 70 and the serine position 75.

It is to note that all Zf3 mutants were obtained in the oxidized form having a disulfide bond between the two N-terminal cysteine residues. This was disadvantageous for two reasons. First, the subsequent native chemical ligation of Zf3 with the expressed Zf12 peptide requires a free N-terminal thiol moiety of the cysteine residue in Zf3, which attacks the C-terminal thioester of Zf12. Second, the peptides are only able to bind Zn(II) with both cysteine residues in their reduced state in order to receive their correct secondary structure. Hence, in all experiments presently described, the zinc fingers were reduced by the addition of the disulfide reducing agent tris(2-carboxyethyl)phosphine (TCEP) for 90 min prior to the experiment. TCEP belongs to the group of non-metal-coordinating reducing agents.[81] In contrast, several studies revealed that other commonly used reducing agents, such as dithiothreitol (DTT) or 2-mercaptoethanol (2-ME), have a pronounced tendency to coordinate divalent metal ions due to their multidentate binding sites composed of sulfur and oxygen ligands at neutral pH.[82] For this reason, the use of the latter substance had been deliberately abandoned in order to avoid interactions with the dinuclear metal complex of the building blocks. These interactions would be undesirable since the dinuclear metal complexes could occupy the allocated phosphate binding site with the reducing agent that may prevent proper substrate activation.

2.4.3 Secondary structure determination by circular dichroism spectroscopy

Circular dichroism (CD) studies were performed in order to examine the ability of the zinc finger mutants to fold into their characteristic $\beta\beta\alpha$-structure upon Zn(II) complexation, which is required for sequence-specific major groove DNA binding.

Zinc finger proteins are known to be mainly unfolded in absence of metal ions, especially of Zn(II).[83] Consequently, CD spectra of the latter differ significantly from those recorded in presence of Zn(II). This results from the higher content of random coil structure in comparison to structural motifs, such as the α-helix and β-sheets, that are found in folded zinc fingers. In general, CD spectroscopy of peptides and proteins mainly relies on the presence of chromophores in the molecule, such as the peptide bond itself (absorption below 250 nm), aromatic amino acid side chains (absorption from 260 – 320 nm) and disulfide bonds (absorption around 260 nm).[84] Nevertheless, when it comes to the determination of the secondary structure, the highest priority refers to the peptide bond, which has a characteristic absorption in the range of 190 to 240 nm. The shape of a CD spectrum in this spectral range is dependent on the structural motifs present in the protein examined, such as random coil structure, α-helices, β-sheets and loop regions. In addition, the amount of each enumerated motif in the protein is also vital for the outcome of the CD measurements since the entirety of all these factors compose the final shape of the spectrum.

CD spectroscopy is one of the most frequently used methods to determine the secondary structures of zinc fingers and their mutants due to the pronounced formation of characteristic structural motifs upon metal complexation. Under metal-free conditions, zinc fingers display a negative peak around 200 nm and a negative shoulder around 222 nm due to the high random-coil content and minor amounts of helical structure.[85] The addition of Zn(II) leads to peptide folding, and thus, generates the $\beta\beta\alpha$-structure. The changed conditions affect the CD spectrum and result in a bathochromic shift of the band at 200 nm to approximately 206 nm and a pronounced increase of the shoulder around 222 nm.[83] These characteristics are well documented for all zinc fingers of the Cys_2His_2-type and are essential for statements concerning correct secondary structure formation.

In this approach, the influence of the building blocks on the secondary structure of Zf3 was evaluated with special attention to the formation of the required $\beta\beta\alpha$-structure. Subject of evaluation were the zinc finger mutants **25a** and **25b**, which were incorporated into the sequence of Zf3 at position 70 and **26a** and **26b** incorporated at position 75 (Figure 2.8). In addition, the unmodified Zf3 peptide **24** served as reference to compare the results of the mutants with those obtained for the native sequence. All experiments were performed in absence (dashed lines) as well as in presence (solid lines) of 6 equivalents $ZnCl_2$ (120 μM)

with regard to the peptide concentration (20 μM) in HEPES buffer (20 mM, pH 7.4). Prior to the experiments, TCEP (500 μM) was added to the buffer solution containing the zinc finger and the samples were incubated for 90 min to reduce the disulfide bond. Because the reducing agent is known to cause disturbances in the far UV region of the CD spectra, its concentration was kept below 1 mM as recommended in literature.[86] To prevent the reformation of the disulfide bond, all solvents were degased and flushed with a gas mixture of 5% hydrogen in nitrogen to apply anaerobic conditions. The CD spectra were recorded in the spectral range of 190 to 250 nm.

2.4.4 Secondary structure assessment of the Zf3 mutants

The modified zinc finger constructs described here were compared to the unmodified analogue in order to evaluate similarities or oppositions in terms of secondary structure formation. The CD spectrum of the native reference peptide showed all aforementioned characteristics, which indicate the correct formation of the $\beta\beta\alpha$-structure upon metal complexation. The measurement under metal-free conditions (dashed line) revealed a strong negative band at 200 nm and a shallow shoulder around 222 nm as expected for the high random-coil content present in unstructured zinc fingers (Figure 2.9). After metal addition (solid line), the band at 200 nm decreased and showed a pronounced red-shift to 205.5 nm. In addition, the shoulder around 222 nm increased significantly indicating the increase of α-helical content in the peptide. The results are very well in line with the data published for the native Zf3 peptide found in literature.[83]
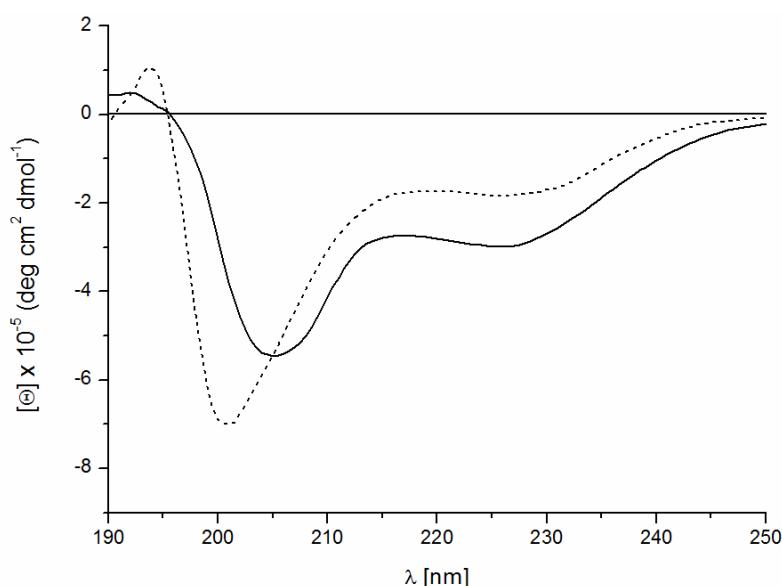
Figure 2.9 CD spectrum of the native Zf3 peptide **24** in absence (dashed line) and in presence (solid line) of Zn(II).

Afterwards, the secondary structures of the peptides **25a** (blue) and **25b** (red) with the building blocks incorporated at position 70 (Figure 2.10) were examined in a similar experiment and were compared to the reference peptide (black). The CD spectra were recorded in absence (dashed lines) and in presence (solid lines) of Zn(II). The results clearly indicate that both mutants showed similar negative maxima around 200 nm under metal-free conditions. Nevertheless, the latter are slightly increased in comparison to the reference spectrum, whereas the local negative maxima around 222 nm are almost identical. Upon Zn(II) addition, the negative bands at 200 nm decreased similarly when compared to the reference but showing a slightly less pronounced bathochromic shift. However, the negative maxima were found at 204.7 nm for peptide **25a** and 204.4 nm for peptide **25b** being well in line with the value obtained for the native peptide (205.5 nm). In addition, the spectral overlap of the bands at 222 nm indicated a comparable formation of α-helical content in all three samples. In general, the CD spectra of both mutants with the building blocks incorporated at position 70 showed all important characteristics, which indicate correct secondary structure formation and are in line with the reference spectra. Thus, it can be assumed that the exchange of the arginine residue was well tolerated by the peptides and resulted in the essential $\beta\beta\alpha$-structure required for DNA binding.
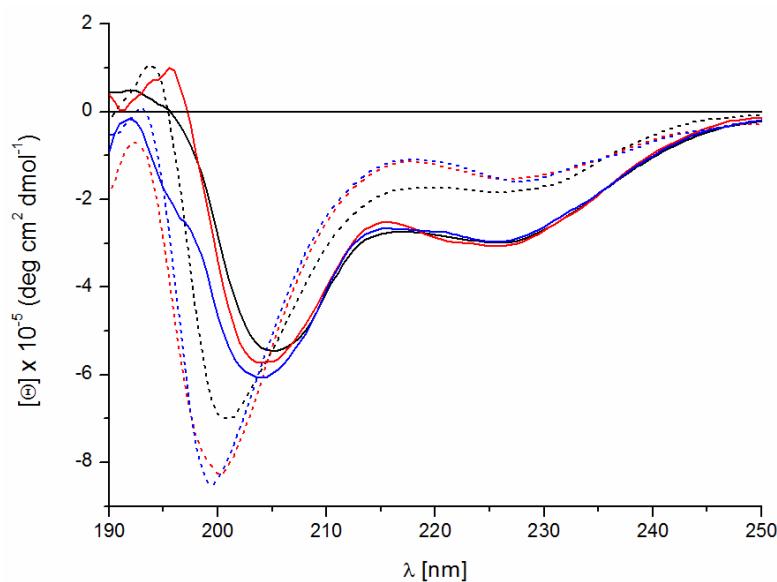
Figure 2.10 CD spectra of mutant **25a** (blue), mutant **25b** (red) and native Zf3 (black) in absence (dashed lines) and in presence (solid lines) of Zn(II).

In contrast, the CD spectra recorded for the peptides **26a** (blue) and **26b** (red) with the building blocks incorporated at position 75 (Figure 2.11) showed significant differences when compared to the reference spectra (black). In the unfolded state, both mutants showed increased bands around 200 nm, which were accompanied by slight bathochromic shifts to 199.1 nm for peptide **26a** and 199.3 nm for peptide **26b**. The shoulders around 222 nm were

considerably less pronounced especially for peptide **26a**. Consequently, a negative influence on the peptide structure induced by the building blocks incorporated at position 75 can be generally stated even under metal-free conditions. This situation was further confirmed by the corresponding spectra recorded after the addition of Zn(II). It has been found that the negative maxima just slightly decreased after metal complexation but more importantly, the bands were much less bathochromically shifted in comparison to the reference sample. Whereas the band of the native peptide was eminently red-shifted to 205.5 nm, the band of peptide **26a** shifted just to 203.1 nm and, moreover, the band of peptide **26b** even only shifted to 201.9 nm. In addition, the shoulder around 222 nm observed for peptide **26a** remained at a similar level when compared to its unfolded state, and therefore, clearly indicated a lack of α-helical content upon Zn(II) addition. In summary, the relatively high deviations observed in the CD spectra of the peptides with the building blocks incorporated at position 75 concluded a disturbed secondary structure formation. It is to note that similar differences were also observed under metal-free conditions implying a negative impact of the building blocks on the overall peptide sequence even in the unfolded state. This might be seen as a consequence of the incorporation site of the building block at position 75. It is conceivable that the bulky residues might interact with the neighboring amino acids, Arg74 and Asp76, which negatively affected the peptide folding.
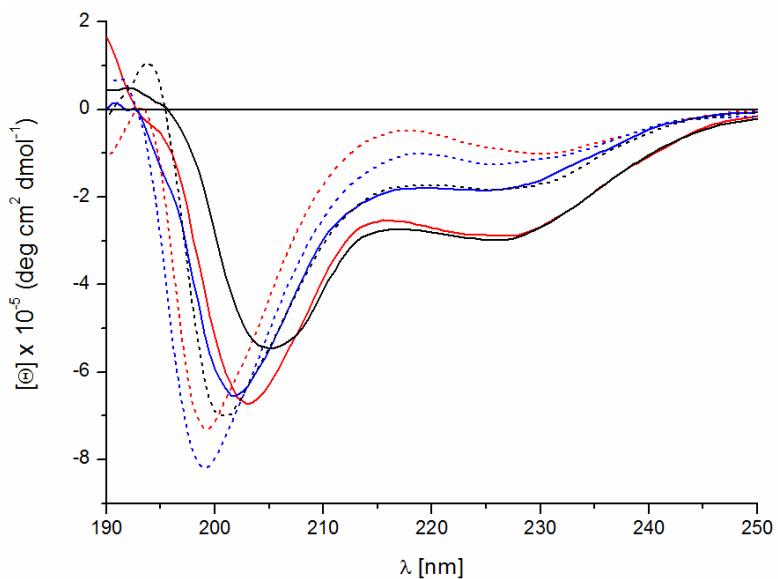
Figure 2.11 CD spectra of mutant **26a** (blue), mutant **26b** (red) and native Zf3 (black) in absence (dashed lines) and in presence (solid lines) of Zn(II).

In conclusion, the CD spectra of the zinc finger mutants **26a** and **26b** with the building blocks incorporated at position 75 were recorded and compared to the unmodified reference peptide. It was clearly demonstrated that both spectra significantly differ from the reference implying an incorrect peptide folding. In contrast, the findings for the mutants **25a** and **25b**

having the building block incorporated at position 70 were in line with the reference sample. This allows the assumption that the structural discrepancies of **26a** and **26b** derived from unsuitable intrinsic conditions for the building blocks when incorporated at position 75, which prevent correct peptide folding. These might originate from the neighboring amino acids being influenced by the bulky dinuclear complex in close proximity.[10] The arginine residue at position 74 could orientate its relatively long and nitrogen-rich side chain in the direction of the metal complex that may induce a bending of the peptide with reduced helix formation (Figure 2.12a). The same considerations are valid for the aspartic acid residue at position 76. Furthermore, both neighboring amino acids make direct nucleobase contacts, whereby a deviated orientation of their side chains would have a substantial adverse effect on DNA recognition and binding.

This hypothesis was further supported with respect to the good results obtained for the incorporated building blocks at position 70. The building blocks themselves are located in the junction region between the β-sheets just as the neighboring amino acid residues Gly69 and Lys71 (Figure 2.12b).[10] This region is known to provide more space and flexibility in comparison to the rigid helical region.[69] Moreover, the lysine residue and in particular the glycine residue are much less sterically demanding in comparison to the aforementioned residues, and therefore, allow the evasion of steric repulsions.
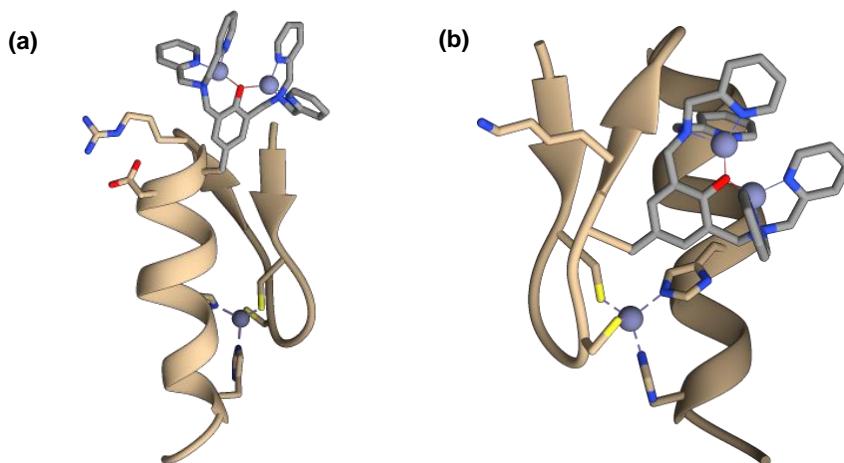
Figure 2.12 Molecular models of **26b (a)** and **25b (b)** showing the incorporated dinuclear building block as well as their neighboring amino acids. The second neighboring amino acid residue in **(b)** is a glycine residue without a side chain, which could be highlighted in the model. The models were generated with UCSF Chimera (PDB code 1AAY).

Due to the results of the CD spectroscopic measurements for both building blocks that have been incorporated at the two most promising positions of the zinc finger, it can be concluded that position 70 turned out to be most beneficial. This assumption is based on the fact that the modifications done at this position just marginally influence the secondary structure that

is vital for maintaining the DNA-binding ability of the zinc finger. However, this does not seem to be the case for the serine 75 position, which revealed significant differences in the recorded CD spectra indicating a problematic secondary structure formation. The question to what extent this fact may have an impact on DNA binding and hydrolysis is covered in section 2.7.

2.5 Expressed protein ligation

The expressed protein ligation (EPL) emerged as an approach for the development of tailor-made proteins that are much less restricted in length in comparison to proteins prepared by conventional peptide synthesis (SPPS).[87] Efficient SPPS is de facto restricted to the synthesis of small proteins of about 60 amino acids in length. This is due to the decreasing coupling efficiency with increasing peptide length that predominantly results from the limited accessibility of the N-terminus because of the progressive winding of the growing peptide chain. In addition, with increasing sequence length during SPPS, solubility problems arise for the fully protected peptide on solid support, which also lower the coupling efficiency.[88] Thus, the 90 amino acids containing Zif268 zinc finger domain with an additional artificial amino acid, which already reduced the synthetic yield, could not be synthesized in a full-synthetic manner using SPPS.

Hence, a semi-synthetic pathway was chosen in order to generate the full-length peptide by means of expressed protein ligation.[89] EPL starts with a complete gene or a gene fragment containing the encoded peptide sequence. This gene is cloned into an expression vector, which is transferred into suitable host organisms, such as yeast or *E. coli*. The recombinant protein is expressed and afterwards purified by affinity chromatography. Another characteristic of this technique is the *in situ* generation of a highly reactive thioester at the C-terminus of the protein, which can subsequently be used in a native chemical ligation (NCL) approach.[90] This semi-synthetic methodology was used in the present study to recombinantly express the Zf12 domain bearing a C-terminal thioester followed by the connection of the conventionally synthesized but modified Zf3 domains by means of NCL as described above.

2.5.1 Protein expression of the Zf12 domain in *E. coli*

The preparation of the Zf12 domain was performed using the commercially available IMPACT™ kit (Intein Mediated Purification with an Affinity Chitin-binding Tag) from *New England Biolabs*. The kit includes the pTXB1 expression vector for *E. coli* containing a modified intein, which is able to undergo a thiol-induced cleavage to release the protein with a C-terminal thioester. In addition, the expression vector includes a chitin-binding domain (CBD), which is suitable for purifying the crude product by means of affinity chromatography

using chitin beads (CB) as docking site. Due to the fact that all additionally expressed proteins and by-products can easily be separated from the target protein, this facilitates the work-up.

The expression starts with the inframe insertion of the gene coding for the Zf12 sequence into the polylinker upstream of the CBD sequence of the pTXB1 expression vector in order to generate the plasmid pTXB1-Zf12. The latter was transformed into the competent *E. coli* host strain ER2566. During incubation, the desired Zf12 protein was overexpressed in assembly with the aforementioned chitin-binding domain. After cell lysis, the lysate was repeatedly passed through a column loaded with chitin beads that have a high affinity for the chitin-binding domain. In this way, by-products were washed from the column before the Zf12 domain was released from the chitin-binding domain in a self-cleaving manner (Figure 2.13a) by the addition of thiol compounds, such as sodium 2-sulfanylethanesulfonate (MesNa). The crude protein with a liberated C-terminal thioester moiety was collected and purified by RP-HPLC (Figure 2.13b).
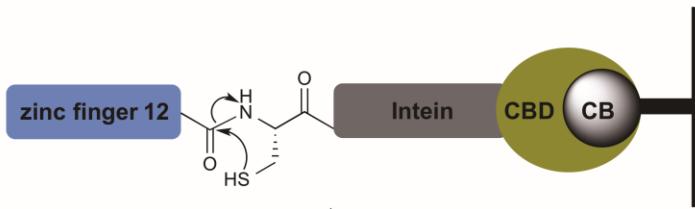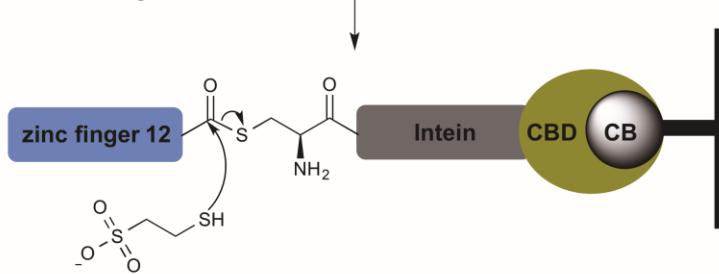
Figure 2.13 (a) Schematic representation of the C-terminal thioester formation of Zf12 during affinity chromatography. **(b)** Amino acid sequence of Zf12 bearing a C-terminal thioester.

2.5.2 Native chemical ligation of Zf12 and modified Zf3 zinc finger domains

The native chemical ligation (NCL) approach is based on the work of KENT *et al.* from 1994 and overcomes the limitations of synthesizing longer peptides and proteins ($>$60 aa) by means of conventional SPPS.[91] This was made possible due to the tethering of two completely deprotected and individually synthesized peptide chains. One of which is bearing a N-terminal cysteine residue with a freely accessible thiol moiety, while the other one bears a C-terminal thioester moiety. In a two-step reaction, the thiol first attacks the thioester leading to a thioesterification. In the second step, this thioester undergoes a fast and irreversible S- to N-acyl shift which finally produces a native amide bond between both compounds. The absence of any side chain protecting groups at both peptides illustrates the high specificity and tolerability of this reaction towards multiple other functional groups.[92] Furthermore, longer proteins are easily accessible by the consecutive application of this reaction with additional peptides bearing the required functional groups.

In this study, NCL was used to couple the aforementioned ribosomal expressed Zf12 domain, bearing a C-terminal thioester, with the conventionally synthesized but modified Zf3 domains, bearing a N-terminal cysteine residue, to generate the full length peptide with pronounced DNA-binding ability (Figure 2.14).
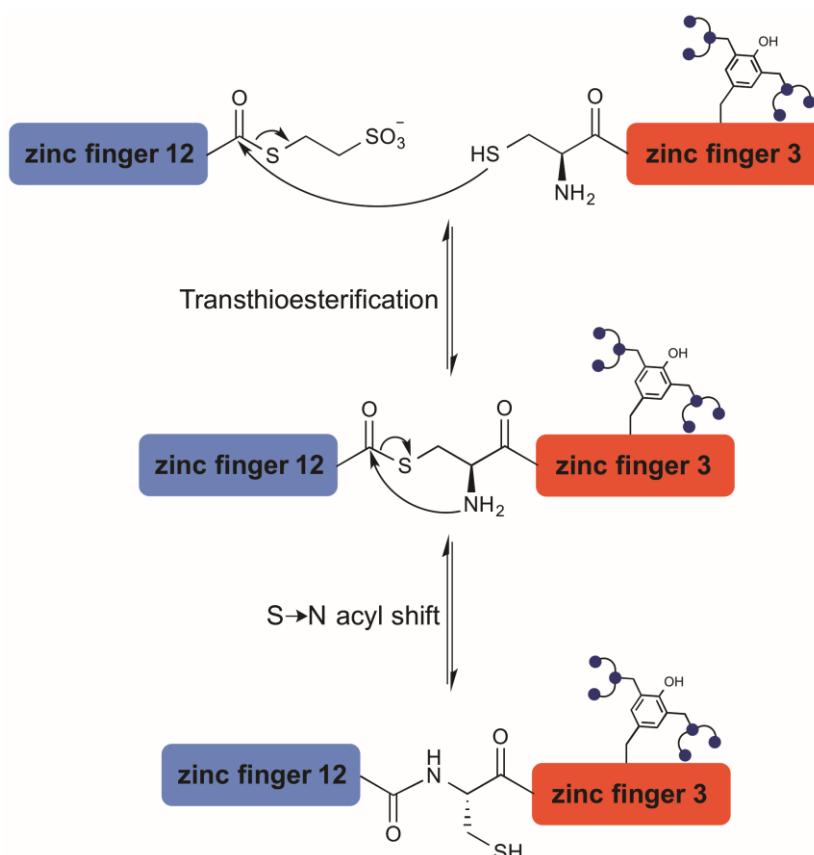
Figure 2.14 Native chemical ligation between the expressed Zf12 thioester and the building block containing Zf3 mutants.

The presence of two cysteine residues in close proximity to each other in the sequence of Zf3 led to the formation of a disulfide bond between the latter residues. This situation made it necessary to reduce the disulfide bond prior to the NCL to receive a free N-terminal thiol group. Therefore, TCEP (5 mM) was used as a reducing agent and the reaction was performed under denaturing conditions using guanidinium chloride (GdmCl, 6 M) to eliminate peptide folding and aggregation, which might have adverse effects on the reaction. NCL was performed under literature known conditions using the Zf3 peptides in a concentration of 1 mM and the Zf12 peptide in 500 μM concentration.[93] First, Zf3 and TCEP were dissolved in a buffered solution of Na_2HPO_4/NaH_2PO_4 (10 mM) adjusted to pH 4 to obtain optimal conditions for the reduction of the disulfide bond. After 2 h of incubation time, Zf12 and GdmCl (6 M) were dissolved in the same buffer system and added to the solution containing the Zf3 peptide. The pH was adjusted to 8 representing the most beneficial pH value for this system to perform NCL. This is due to the fact that thiols become increasingly deprotonated under basic conditions and the resulting thiolates react highly nucleophilic towards the electrophilic thioester. Unfortunately, thiolates are also prone to oxidation in this state facilitating the reformation of the disulfide bond. This was prevented on the one hand due to the presence of the denaturing agents mentioned above and, on the other hand, by applying anaerobic conditions with 5% hydrogen in nitrogen. All solvents were excessively degased and flushed with this gas mixture prior to use. The reaction mixture was agitated at room temperature for 24 h, while the reaction progress was monitored by RP-HPLC of reaction solution aliquots after distinct time periods. The retardation times of Zf12 and Zf3 evaluated beforehand were used to identify the decreasing educt peaks, while a new peak appeared for the Zf13 product. Finally, the full-length peptides (Zf13) **29**, **30**, **31** and **32** were purified by means of RP-HPLC (Figure 2.15).
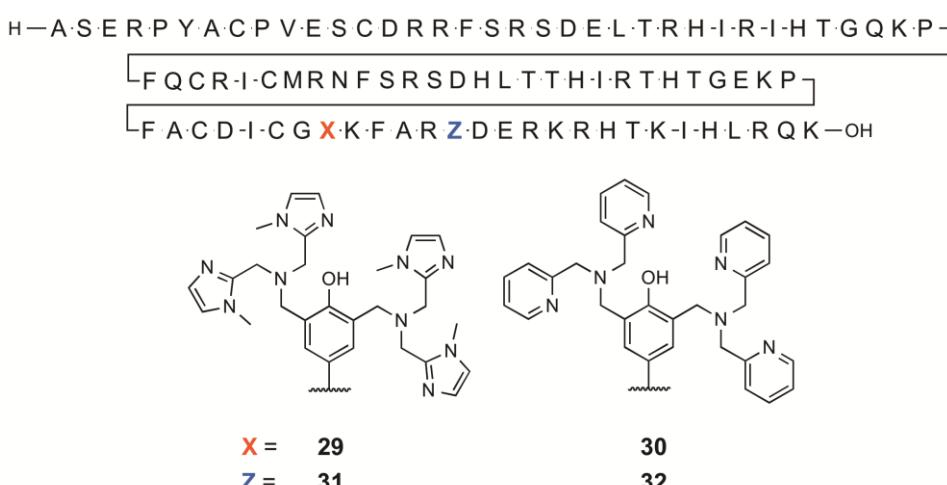
Figure 2.15 Peptide sequence of the modified Zf13 domains with the dinuclear building blocks incorporated at position 70 (**X**, Zf13BMIA70: **29** and Zf13BPA70: **30**) and 75 (**Z**, Zf13BMIA75: **31** and Zf13BPA75: **32**).

2.6 DNA binding studies with engineered Zf13 domains

Each of the three individual Zif268 zinc fingers obtain a specific ββα-structure upon binding of a Zn(II) ion.[83] Subsequently, the protein is able to wrap around the DNA, whereby the recognition helices lie in the major groove and make contacts to three nucleobases per individual zinc finger domain. More detailed information regarding DNA binding and nucleobase recognition is provided in section 1.1.1. In the following, the focus lies on the effects of the introduced dinuclear building blocks on the overall ability of the peptide mutants to accurately bind double-stranded DNA. For this purpose, polyacrylamide gel-electrophoresis was applied under native conditions in order to identify the larger, and therefore, more slowly migrating DNA/zinc finger complexes in contrast to the faster migrating individual double-stranded DNA reference sample.[94]

The here described experiments were performed with a 30-mer DNA duplex containing the Zif268 operator sequence 5'-**GCGTGGGCG**-3'.[10] The double-stranded DNA was generated by annealing the single-strand oligomers 5'-GTGTGTGTGTGATCTTA**GCGTGGGCG**T-AAG-3' and 3'-CACACACACACTAGAATCGCACCCGCATTC-5'. In order to visualize the DNA and/or protein complexes with the latter after gel-electrophoresis, a fluorescent 5'-FAM label was attached to one of the oligomers. The disulfide bonds being present in the Zf13 peptides (1 mM) were reduced with an aqueous solution of TCEP (5 mM) for 90 min and subsequently dialyzed against a zinc-containing buffer solution (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 10 μM $ZnCl_2$, pH 7.8) for 4 h. Prior to use, all solvents were degased and flushed with an anaerobic gas mixture containing 5% hydrogen in nitrogen in order to prevent reformation of the disulfide bonds. After occupation of all metal-binding sites of the peptide and the building blocks with Zn(II), excess amounts of the metal ion were removed by dialysis against a buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 1 μM $ZnCl_2$, pH 7.8) having a particularly low Zn(II) concentration.

In contrast to non-native or denaturing conditions, which are often used in gel-electrophoresis to separate and analyze macromolecules depending on their size and migration speed in an electric field, the here described experiments were performed under native conditions.[95] This means that no denaturing agents, such as sodium dodecyl sulfate (SDS) or urea, were added to preserve the native structure of the zinc finger protein as well as to keep the DNA double-helix intact. In addition, the use of additives, such as EDTA, was renounced due to its strong ability to withdraw Zn(II) from the zinc finger complex. Thus, the protein/DNA complex remains associated and can be detected as a slower migrating band in comparison to the faster migrating bands of the individual components.

In a typical experiment, distinct peptide to dsDNA ratios ($r_f = C_{Zf13}/C_{DNA}$ = 0, 1, 5, 10, 30) were pipetted into a sample buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 80 μM ZnCl$_2$, pH 7.8) and incubated at room temperature for 2 h. Subsequently, 2 μL glycerol (30%) was added to each sample to ensure the formation of a layer at the bottom of the well. The samples were pipetted into the wells of the polyacrylamide gel and electrophoresis was performed using a non-denaturing running buffer (25 mm Tris, 192 mm glycine, pH 8.2).

The binding of the peptides to DNA was analyzed using a fluorescence imager in order to detect the signal of the 5'-FAM label. The results for the Zf13 mutants were compared to the migration speed of the reference sample just containing dsDNA ($r_f = C_{Zf13}/C_{DNA}$ = 0). The use of a molecular-weight size marker was renounced because the applied native conditions prevent a reliable size separation during electrophoresis. This is a consequence of the absence of charged denaturing agents, whereby the analytes do not only differ in molecular mass but also in structure, size and intrinsic charge, which contribute to the migration speed.[96] The obtained results are presented in Figure 2.16.
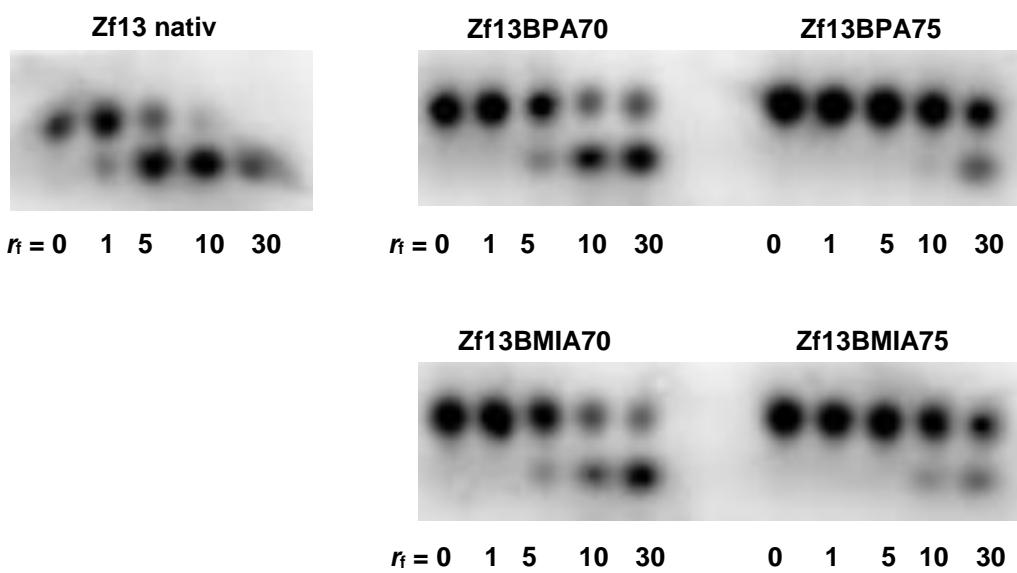
Figure 2.16 Native polyacrylamide gel electrophoresis for the evaluation of zinc-finger/DNA binding. The consensus dsDNA (0.5 μM) was incubated with the zinc finger peptides Zf13 native (**28**), Zf13BMIA70 (**29**), Zf13BPA70 (**30**), Zf13BMIA75 (**31**) and Zf13BPA75 (**32**) in different concentrations ($r_f = C_{Zf13}/C_{DNA}$ = 0, 1, 5, 10, 30) at room temperature for 2 h.

Due to the absence of peptide (r_f = 0), the reference sample showed just one migrating band for the duplex DNA, which acted as reference value for the formation of larger, and therefore, slower migrating DNA/zinc finger complexes. The native Zf13 peptide (Figure 2.16) displayed high affinity towards its operator sequence that can be concluded from the formation of a strong band, which corresponds to the associated complex and a much weaker band for the

unbound dsDNA even at low peptide to DNA ratios ($r_f = 5$). For peptide concentrations higher than $r_f = 10$, the unbound dsDNA band started to disappear implying the occupation of DNA-binding sites with the native zinc finger peptide. The Zf13BMIA70 (**29**) and Zf13BPA70 (**30**) mutants incorporated at position 70 showed a comparable behavior but with a slightly less pronounced DNA affinity (Figure 2.16). This was concluded from the overall higher protein concentrations, which were necessary to observe high occupation rates of the dsDNA with protein ($r_f = 10$) as indicated by the disappearance of the unbound dsDNA band. In addition, even at ratios of $r_f = 30$, the unoccupied dsDNA was detected. Consequently, the engineered zinc fingers are able to bind their consensus DNA sequence but with a slightly lower affinity with regard to the native peptide.

In contrast, mutants **31** and **32**, which have been incorporated at position 75, did not show any specific DNA-binding ability that can be assumed from the high concentrations of peptides, which were required for the formation of the DNA/peptide complex. Moreover, the unbound DNA band was still dominant even at the highest ratios tested indicating a poor binding affinity. This result was not surprising since the CD spectra for the latter mutants, which are described in section 2.4.4, revealed a disturbed secondary structure formation. It was concluded, that the incorporation of the building blocks at position 75 has a negative influence on the folding behavior of the zinc fingers. Hence, it can be assumed that the still observed DNA complex formation ($r_f = 30$) was a consequence of the lowered recognition capacity and predominantly resulted from unspecific electrostatic interactions and hydrogen bonding between the analytes.

In summary, the DNA-binding ability of the zinc finger mutants **29** and **30** with the building blocks incorporated at position 70 showed excellent results in terms of DNA/protein complex formation. However, the peptide to DNA ratio for complete occupation of all binding sites was higher compared to the unmodified reference peptide being indicative of minor negative influences induced by the incorporated building blocks. Those effects were clearly documented for both mutants containing the allosteric metal-binding site at position 75. A specific DNA recognition or, respectively, DNA binding cannot be assumed for these mutants with regard to the gel electrophoretic findings. This further supports the assumptions drawn from the CD spectroscopic results.

2.7 DNA cleavage studies with engineered Zf13 domains

The aforementioned DNA-binding experiments for the synthesized zinc finger mutants **29** and **30** were promising with regard to a possible phosphodiester hydrolysis. An already proceeded DNA cleavage during the previously described experiment could not be excluded due to the use of fully operational peptides. However, the applied native reaction conditions

prevented the detection of such an event. As mentioned at the beginning, the implementation of the building blocks at only one distinct position in the peptide just enabled the possibility for the latter to reach the phosphate groups of exactly one DNA strand. For this reason, a double-strand cleavage was virtually impossible. Hence, a single-strand cleavage could not be detected due to still existing WATSON-CRICK base pairing holding the duplex together over the full length of the DNA 30-mer. In addition, the incubation time of 2 h, which was applied in the previous binding studies, was rather short to scissor major amounts of the DNA substrate. This assumption was confirmed by the results obtained from the hydrolysis studies with the activated, and therefore, easier to cleave BNPP model substrate. Consequently, the PAGE conditions had to be adjusted to non-native conditions in order to visualize the formation of cleavage products.[97] The peptides were prepared according to the method described in section 2.6. After protein dialysis, incubation was carried out in a sample buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 1 μM $ZnCl_2$, pH 7.8) including the zinc finger mutants and duplex DNA (0.5 μM) under native conditions and for 72 h instead of 2 h. Moreover, the temperature was increased from room temperature to 37 °C in order to further increase hydrolysis rates. Subsequently, denaturing conditions were applied by preparing a 1:1 dilution of the sample probes with a double-concentrated solution containing SDS (2%), glycerol (25%, v/v) EDTA (2 mM) and DTT (180 mM) in Tris buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, pH 7.8). The mixtures were heated to 95 °C for 5 min and were subsequently loaded on the gel after cooling on ice. Under these conditions, the peptides as well as the duplex DNA were completely denatured that goes in hand with the release of smaller DNA fragments resulting from single-strand cleavages. Different peptide to DNA ratios ($r_f = c_{Zf13}/c_{DNA}$ = 20, 30 and 60) were examined, and therefore, prepared in parallel. After loading of the samples on the gels, electrophoresis was performed using a denaturing running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.2).
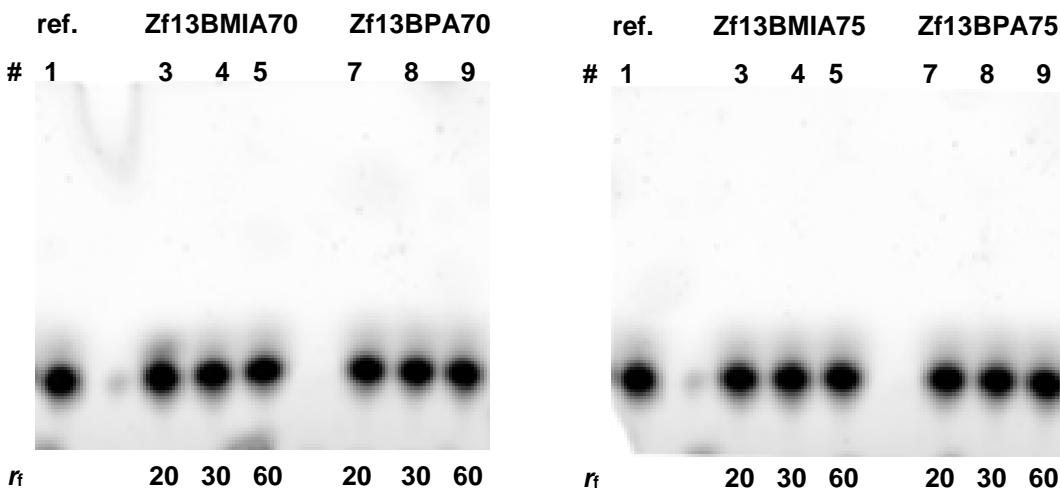
Figure 2.17 Denaturing PAGE studies to examine the cleavage ability of the engineered zinc fingers towards their dsDNA target-site. Different peptide to DNA ratios ($r_f = C_{Zf13}/C_{DNA} = 20$, 30 and 60) were prepared and incubated at 37 ° for 72 h. The single-stranded DNA oligomer containing the operator sequence for the zinc finger was 5′-labeled with a FAM fluorophore in order to visualize fragmented DNA.

The identification of the most beneficial incorporation sites for the building blocks was based on molecular models, which were generated using published crystallographic data. As mentioned in section 2.4, the building blocks incorporated at position 70 should cleave the operator strand, whereas the building blocks at position 75 should cleave the opposite strand of the DNA. As a consequence, fluorescent labeling of just the individual single-strand, which takes part in the hydrolysis reaction, appears appropriate but bears the risk of causing a false negative result in case the other strand had been cleaved instead. In order to avoid misleading interpretations, all experiments were performed twice with either the binding-strand labeled with 5'-FAM (Figure 2.17) or with the opposite non-binding strand (Figure 2.18) labeled with the same fluorescent dye.

The control lane contains solely the duplex DNA without any addition of unmodified zinc finger due to its documented lack of hydrolysis capacity and because of peptide wastage by virtue of the denaturing conditions. Hence, the only band observed after electrophoresis displayed the fully intact and labeled DNA single strand.

The results obtained for the cleavage experiment with the labeled operator strand and the zinc finger mutants **29** and **30** incorporated at position 70 showed only bands on the same level with respect to the control lane. Moreover, also the peptides **31** and **32,** which were incorporated at position 75, did not show any significant differences in terms of their migration speed indicating the integrity of the labeled DNA single strand. Also, the experiments that labeled the opposite DNA strand, revealed the same results (Figure 2.18). Consequently, all experiments performed with the four zinc finger mutants showed an unverifiable hydrolysis ability under the applied reaction conditions. This was further

supported by control experiments in which the incubation time was prolonged to 5 days and the temperature was increased to 55 °C.
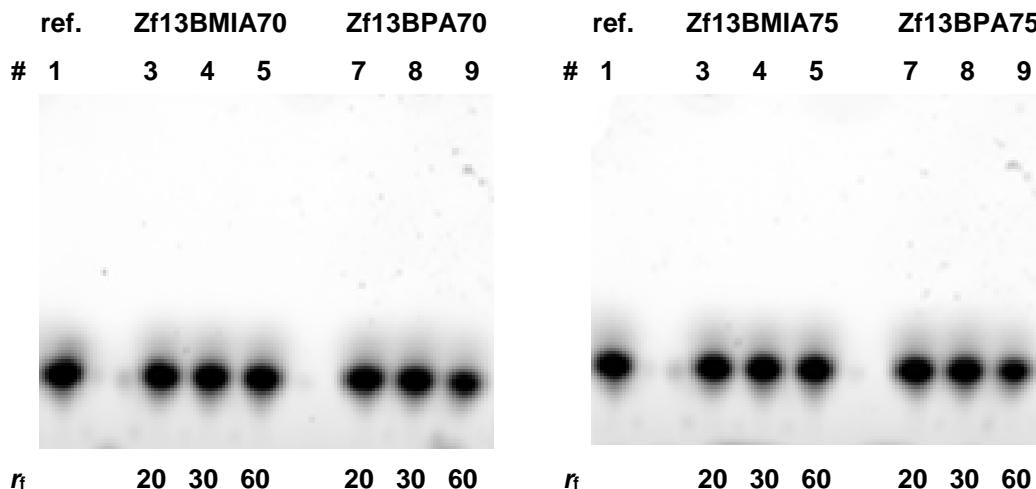
Figure 2.18 Denaturing PAGE studies to examine the cleavage ability of the engineered zinc fingers towards their dsDNA target-site. Different peptide to DNA ratios (n = C_{Zf13}/C_{DNA} = 20, 30 and 60) were prepared and incubated at 37 ° for 72 h. The single-stranded DNA oligomer opposite to the operator sequence containing strand was 5′-labeled with a FAM fluorophore in order to visualize fragmented DNA.

Finally, the question why the examined zinc finger mutants were not able to perform hydrolysis on the consensus DNA sequence, even though their binding abilities had been proven for DNA model substrate, is to be clarify.

One aspect may be a possible insufficient substrate binding, and therefore, an inadequate activation of the latter. The phosphodiester backbone of the DNA is particularly stable even at high concentrations of hydroxide ions.[98] Without an appropriate activation by a dinuclear metal complex, which forces hydrolysis in the direction of phosphodiester cleavage, the cleavage mechanism generally is a reversible process. It is even conceivable that due to the relative rigidity of the peptide structure bound to DNA, the building blocks are not perfectly able to bind the substrate. X-ray crystallography of the present DNA/peptide complex could provide useful information on the behavior of the building blocks when bound to DNA. This would allow for a more profound interpretation of the actual factors leading to an absence of DNA hydrolysis.

2.8 Summary

The ability of zinc fingers to bind to DNA in a sequence-specific manner plays a key role in living cells in terms of controlling gene transcription and to trigger protein biosynthesis.[69] The remarkable precision of DNA recognition provided by this protein family was used in this study to expand their range of applications. Thus, artificial amino acids based on tyrosine, which feature additional dinuclear metal-binding sites mimicking the active site of a

metallohydrolase, were successfully synthesized. The latter were incorporated into the native peptide sequence of zinc finger 3 at strategic positions, which should allow the building blocks to come in close proximity of the phosphate backbone of the DNA in order to perform hydrolysis on a distinct phosphodiester moiety. The major advantage of this purpose was the tremendous size reduction of an otherwise large hydrolase enzyme to just its metal derived catalytic site. The hydrolysis abilities of the synthesized building blocks were tested with the DNA model substrate BNPP and clearly revealed high affinity towards the phosphodiester group with an elevated cleavage capacity under pseudo-physiological conditions. Moreover, determination of the secondary structures of the zinc finger mutants clearly demonstrated native-like peptide folding for the building blocks, which were exchanged for the arginine residue at position 70. In contrast, the serine 75 position was proven to be inadequate for the incorporation of both building blocks due to deviated spectra with regard to the native reference. It was assumed, that this was a result of the unsuitable intrinsic conditions of this rigid helical position, which might not be able to tolerate the interference of the building blocks and the neighboring amino acids. This finding is very beneficial for the future selection of an appropriate incorporation site of relatively bulky artificial amino acids in the sequence of zinc finger proteins.

However, the modified Zf3 domains were successfully used in an EPL approach with the recombinantly expressed Zf12 domain in order to generate the engineered three tandem zinc-finger peptides of Zif268. Only this situation allows for site-specific DNA recognition and binding. Polyacrylamide gel electrophoresis under native conditions was used to examine the binding capacities of the four mutants towards their consensus dsDNA target. The peptides bearing the building blocks at the arginine 70 position showed a pronounced ability to form complexes with their consensus DNA sequence. In contrast, the mutants with an exchanged serine 75 residue showed comparably low binding affinities, which were assumed to be mainly of electrostatic nature. The findings obtained from the PAGE experiments therefore confirm the conclusions drawn from the CD experiments saying that the building blocks at this position are unfavorable for the peptides secondary structure formation.

In a similar experiment, the engineered zinc fingers were tested for their ability to cleave the phosphodiester backbone of their bound dsDNA template. Incubation of the peptides and the DNA for 72 – 120 h at elevated temperatures ranging from 37 – 55 °C and subsequent gel electrophoresis could not validate a positive hydrolysis event. This was true for all of the synthesized mutants without any preference for the incorporation site. Analogous experiments with the fluorescent label either attached to the binding or to the non-binding strand of the duplex DNA showed intact single strands even at increased peptide to DNA ratios. It was assumed, that despite the documented binding ability of **29** and **30**, there is an

activation mismatch between the dinuclear residue and the accessible phosphodiester moiety. Thus, substrate cleavage was prevented and the duplex DNA stayed intact. This clearly accounted for the evaluation of a different incorporation site, which allows for accurate substrate activation. Due to a lack of comparative studies, future approaches might have to clarify the actual conditions of the engineered zinc fingers in a complex with a DNA substrate by the means of X-ray crystallography.

However, the goal of modifying the highly sequence-specific zinc finger protein with additional dinuclear building blocks mimicking the active site of a metallohydrolase was successfully achieved even though the incorporation site needs to be improved in order to obtain efficient DNA hydrolysis. This approach represents a vital example for the combination of two different systems in order to create a tailor-made artificial restriction endonuclease. The synthesis of such a tool that combines DNA recognition, binding and cleavage could contribute to the development of novel gene-editing systems without the need for huge nuclease enzymes.

3. Phosphoserine modified zinc fingers for site-specific DNA hydrolysis by Ce(IV)/EDTA

Non-enzymatic cleavage of DNA is a major target in drug design and in the development of artificial restriction enzymes. Hence, the evaluation of different metal ions and their complexes is still under evaluation in order to enhance cleavage rates with respect to phosphodiester hydrolysis.[99]

Since the discovery of the overwhelming hydrolysis abilities of lanthanides, a lot of research has been done in this field in order to cleave site-specifically double-stranded DNA under mild and physiological conditions.[100]

First approaches were performed using Ce(III) salts in nitrogen-saturated solvents under physiological conditions. However, only low catalytic activity towards DNA was observed. In contrast, the same salts treated with molecular oxygen showed a tremendously increased cleavage ability as a direct result of the oxidation of Ce(III) to Ce(IV).[101] After the detection of tetravalent cerium ions to be the active species by YASHIRO and CHIN, independently, great effort was expended on the investigation of possible hydrolysis mechanisms and numerous kinetic and spectroscopic studies were conducted.[101,102]

The studies disclosed documentary evidence that the catalytically active species is a dinuclear hydroxo-clusters of Ce(IV), which activates the phosphodiester of the DNA and boosts the formation of a pentacoordinated phosphorus intermediate. In a fully reversible back-reaction, the initial state is normally preferred, and therefore, recovered immediately. Nevertheless, the influence of the lanthanides forces the reaction to shift to the irreversible phosphodiester cleavage with rate accelerations of $10^8 - 10^{12}$ fold.[99] In comparison to other metal ions used for DNA hydrolysis, such as Zn(II), Fe(III) or Mg(II), these outstanding values have been so far constrained to Ce(IV) ions. Moreover, Ce(IV) is even superior to all other lanthanides, such as La(III), Tm(III), Lu(III) and Yb(III).[103]

This raises the question of which feature does these ions make special in terms of phosphodiester activation and cleavage. The hydrolysis of DNA takes place in a two-step process (Figure 3.1). The first step starts with the attack of a reactive nucleophile (e.g. hydroxide ion) to the phosphorus atom, leading to a pentacoordinated intermediate. The second step is initiated by the scission of the 5′-OH, which belongs to the 2′-deoxyribonucleotide, from the phosphorus atom which results in the cleavage of the phosphodiester bond.[99]
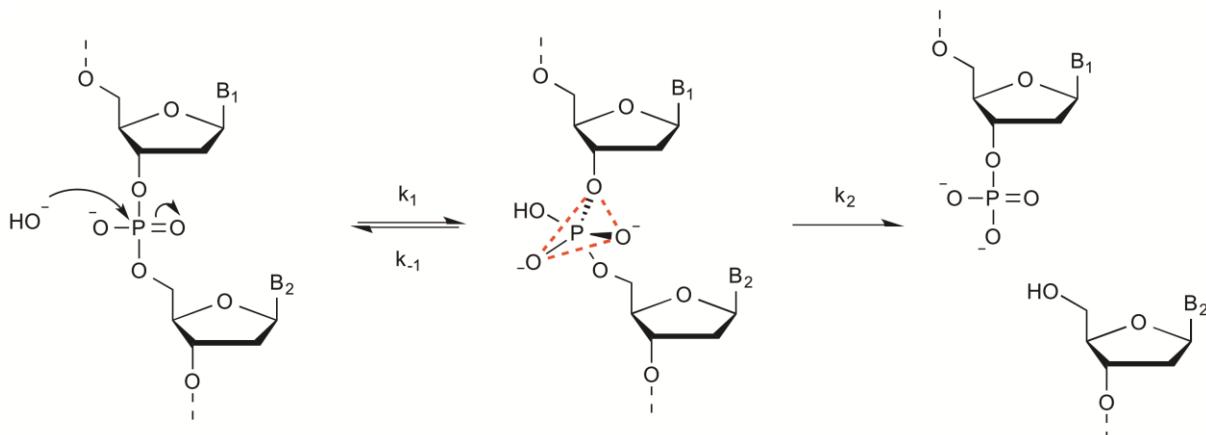
Figure 3.1 Mechanism of DNA hydrolysis via a two-step procedure.

In order to obtain increased hydrolysis-rates, the Ce(IV) species must promote the rate limiting step of the two-step mechanism described above. It is not clearly proven which of the two steps is the rate limiting step. Different approaches tried to concretize this question by the examination of 2′-deoxyribonucleotide-related leaving groups and their impacts on hydrolysis rates. After the evaluation of different energy diagrams obtained from experimental data, a rate acceleration by 10^{11} fold was observed for Ce(IV)-promoted hydrolysis with regard to the uncatalyzed alkaline hydrolysis (Figure 3.2).[104] Consequently, the aforementioned second step must be enhanced by the same value. Due to the fact that the second step continues to limit the rate even in the catalyzed reaction, the first step must also be promoted by the Ce(IV) ion to the same or higher proportions. If this was not the case, the first step would limit the rate. Consequently, the participation of the Ce(IV) species promotes both steps by decreasing the Gibbs free energy of the reaction, and it therefore, lowers the height of the energy barrier.
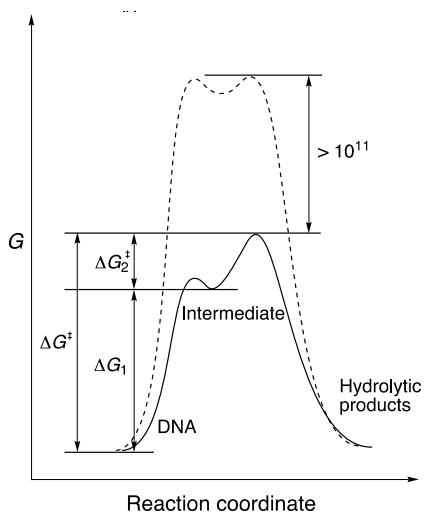
Figure 3.2 Energy diagram for the Ce(IV)-promoted DNA hydrolysis (solid line) in comparison to non-catalyzed alkaline hydrolysis (doted line).[104]

3.1 Determination of the catalytically active species

The cleavage of DNA by Ce(IV) ions in an aqueous environment works well over a broad pH range. However, at a pH value above 4, Ce(IV) forms hydroxide-gels of various compositions, which are unsuitable for kinetic studies.[105] Below this value, homogeneous hydroxide complexes of Ce(IV) are formed, which show just a minor decrease in activity towards DNA hydrolysis compared to the hydroxide-gels. Under these conditions, cerium is expected to form a bimetallic complex of the composition $[Ce^{IV}_2(OH)_4]^{4+}$, which represents 20–30% of the conceivable species in the mixture.[106] This complex showed the highest catalytic activity of all possible homonuclear Ce(IV) hydroxide complexes (e.g. $[Ce^{IV}(OH)]^{3+}$, $[Ce^{IV}(OH)_2]^{2+}$, $[Ce^{IV}_2(OH)_2]^{6+}$, $[Ce^{IV}_2(OH)_3]^{5+}$, $[Ce^{IV}_2(OH)_4]^{4+}$, and $[Ce^{IV}_6(OH)_{12}]^{12+}$). Therefore, different studies focused on the electronic nature of the lanthanide ions to draw conclusions on the patterns behind the activation and cleavage of phosphodiesters. For instance, core-level spectroscopy was performed in order to get insights in the electron distributions and the involved orbitals.[107] This method uses X-ray energy to eject a core electron from a specific atomic site. Different sub-types of this technique (XPS, XAS, AES and XES) were used to obtain information on the electronic structure, the local geometry and the chemical state at the absorption site. Thus, core-level spectroscopy performed for the 2p orbitals of a phosphorous substrate within a complex of various lanthanide ions revealed an enormous electron-withdrawing effect induced by Ce(IV). Accordingly, its electrophilicity was strongly increased promoting the following attack of a hydroxide nucleophile that results in the pentacoordinated intermediate. In addition to this effect, L_3-XANES spectroscopy disclosed that 0.67 electrons were present in the 4f orbitals of the Ce(IV) ions when these were in the corresponding phosphorus complex.[108] As a matter of fact, the latter orbitals should be

55

empty for tetravalent cerium. Consequently, the detected electron density had to originate from the coordinated phosphate groups. It was assumed that the electron-withdrawing effect resulted from the formation of hybrid-orbitals between the lanthanide and either the phosphorus-atom or the bound oxygen-atoms. Some studies consider the simultaneous participation of both species possible. This was demonstrated by X-ray photoelectronic measurements, which documented the occurrence of a new energetic state near the Fermi-level.[109,110] The energetic state neither belonged to the individual cerium species nor to the phosphodiester. As a result of this, the spatial proximity of the Ce(IV) species to the phosphodiester is vital for an efficient substrate activation in order to enhance its electrophilicity and susceptibility for a nucleophilic attack by the hydroxide ion.

The increased acceleration of phosphodiester cleavage by Ce(IV) is a result of the formation of mainly dinuclear complexes, which were identified to be the active species. In addition, the electronic structure of Ce(IV) shows a vast potential in order to withdraw electrons from the phosphate group to the metal and hence increases its electrophilicity. Due to this fact, the nucleophilic attack of a nearby formed hydroxide is promoted, the formation of the pentacoordinated phosphorus atom is stabilized and the actual cleavage is further boosted. In comparison to the hydrolysis mechanism of other metal complexes described above, the fundamental steps remain the same but the rate acceleration results from the unique electronic features of the lanthanide ion cerium(IV).

3.1.1 Recruitment of Ce(IV)/EDTA complexes by phosphorylated amino acids

The hydrolytic activity of Ce(IV) complexes towards double-strand DNA was repeatedly demonstrated and profoundly studied. However, the site specificity of such complexes is limited to address mainly unnatural modifications of the DNA, such as gap or bulge structures (Figure 3.3).[111]
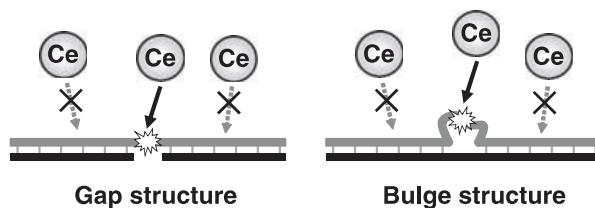
Figure 3.3 The concept of site-specific DNA cleavage by Ce(IV) ions at bulge- or gap-structures.[111]

The latter results from DNA-invading molecules or induced kink points caused by DNA-bending molecules, such as cisplatin derivatives. These modifications function as predetermined breaking points due to local changes in geometry and altered WATSON-CRICK base pairing. In addition, even nature makes use of those sites to initiate the programed cell death called apoptosis, which leads to a chromosomal DNA fragmentation as a part of the DNA repair machinery.[112] The altered geometric conditions, and therefore, the reduced stability at these positions makes them an ideal attachment point for either natural restriction enzymes or artificial DNA cutters, namely ARCUTS (artificial restriction DNA cutters, Figure 3.4).[113]
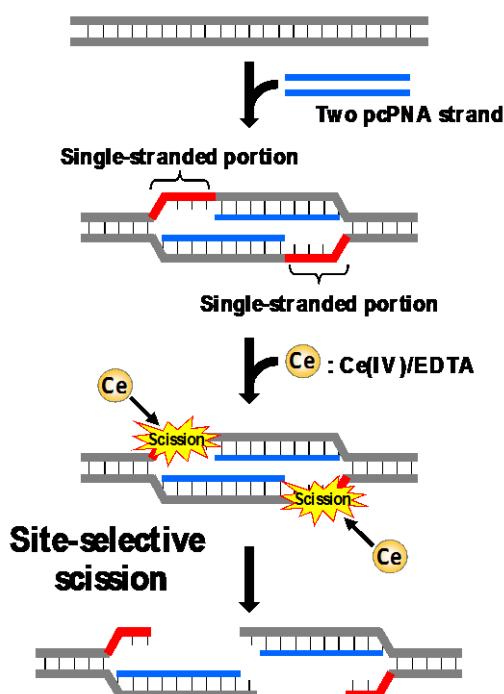
Figure 3.4 Site selective hydrolysis of dsDNA by artificial restriction DNA cutters (ARCUT). Two pcPNA strands (blue) invade double-stranded DNA to induce the formation of single-stranded proportions (red). Subsequently, a Ce(IV)/EDTA complex selectively hydrolyses on the modified sites.[113]

Nevertheless, Ce(IV) complexes also show high catalytic activity for double-stranded DNA hydrolysis without having any site specificity.[114] Every available phosphodiester group in the backbone is cleaved with equally high probability, which was observed in experiments where dsDNA was incubated with an excess of Ce(IV)/EDTA at 50 °C. This resulted in random fragmentation of the DNA as confirmed by gel mobility shift assays.

In order to enhance the site specificity of the latter complexes, it was found that lanthanide ions show high affinities towards phosphate groups. This effect is responsible for the recruitment of the catalytically active species to the desired target site. With an accumulation

at this site, the complex is able to relocate between the phosphate ligand and the phosphodiester groups of the DNA.[115]

It was recently reported, that appropriate phosphate containing ligands are e.g. phosphorylated amino acids, such as threonine, tyrosine or serine. These unnatural amino acids can easily be attached to DNA recognizing and/or invading molecules in order to obtain sequence specificity. In this context, it was possible to recruit external lanthanide complexes to a desired cleavage site as reported by KAMESHIMA *et al.* (Figure 3.5).[116] Due to these facts, it is greatly favorable to combine a highly DNA-recognizing protein such as the naturally-occurring zinc finger with a recruiting phosphate moiety. Thus, due to the accumulation of the external Ce(IV) complex at the position of the phosphorylated amino acid, the site-specific hydrolysis could be enabled after the protein has bound to the DNA.
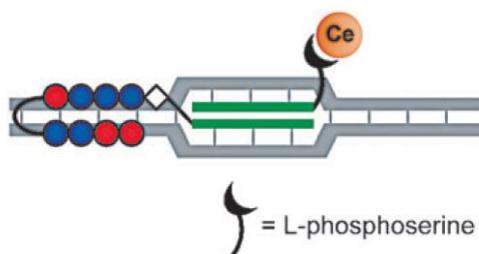
Figure 3.5 ARCUT approach using phosphoserine-modified pcPNA (green lines) to invade dsDNA combined with a sequence-recognizing polyamide (colored circles). The phosphoserine residue site-selectively recruits an external Ce(IV) complex to the bulge-sites where it is located itself.[116]

3.2 Preparation of phosphoserine suitable for solid phase peptide synthesis

The phosphorylation and dephosphorylation of certain amino acids, such as serine, threonine, tyrosine, and to a lesser extend of histidine and aspartate, is a commonly found process in nature.[117] It is associated with specific tasks that are switched on or off. For instance, the permeability of the transmembrane water channel protein aquaporine-2 (AQP-2) is regulated by a phosphorylation and dephosphorylation process, which leads to the opening and closing of the channel.[118] In the cell, phosphorylation is typically performed by protein kinases, whereas the dephosphorylation is achieved by phosphatases.[117] For the preparation of phosphorylated products by means of organic synthesis, a distinction is made between direct phosphorylation of residues within already existing peptides or proteins (global phosphorylation strategy)[119] and the synthesis of single phosphorylated amino acids suitable for SPPS (prephosphorylation strategy)[120]. The latter strategy also involves the elaboration of a sufficient side-chain protecting group strategy in order to avoid side reactions during the coupling and cleavage steps in Fmoc-based SPPS. Thus, the prephosphorylation strategy appeared to be most suitable for the present approach since it guarantees the specific incorporation of a phosphorylated amino acid at the desired target site. In contrast,

enzymatically performed modifications on the entire peptide bear the risk of multi-phosphorylation events, which would occur on all amino acid residues of the same kind. Moreover, the global phosphorylation approach is restricted to cases the residue which should be phosphorylated is already in the right position in the peptide sequence. This would prohibit an exchange of any other amino acid in the sequence for a phosphorylated residue.

A well-established phosphorylated amino acid that is suitable for the aforementioned purposes is Fmoc-O-benzyl-L-phosphoserine, which was introduced by WAKAMIYA in 1994.[121] The use of a single-benzyl side-chain protecting group is completely sufficient with regard to the also available dibenzylated residue. Moreover, this protecting group is entirely orthogonal to the Fmoc-based coupling protocol and efficiently prevents the elimination of the phosphate group to form dehydroalanine under basic conditions.[122] Due to its facile removement under acidic conditions simultaneously with most other side-chain protection groups, the monobenzylated product is also preferably used in this context.

The original preparation of Fmoc-O-benzyl-L-phosphoserine by WAKAMIYA was slightly modified by the approach of PETRILLO (Scheme 3.1).[120] Equimolar amounts of the inexpensive PCl$_3$ (33), which readily provides the phosphorus source, and benzyl alcohol, which acts as side-chain protection group, were mixed and a nucleophilic substitution of a chlorine for a benzyl group occurred. Subsequently, Fmoc-L-serine was added to the mixture in order to form the cyclic phosphite intermediate 34, which was subsequently oxidized to the benzyl protected phosphate. After recrystallization, product 35 was obtained in a yield of 83%.
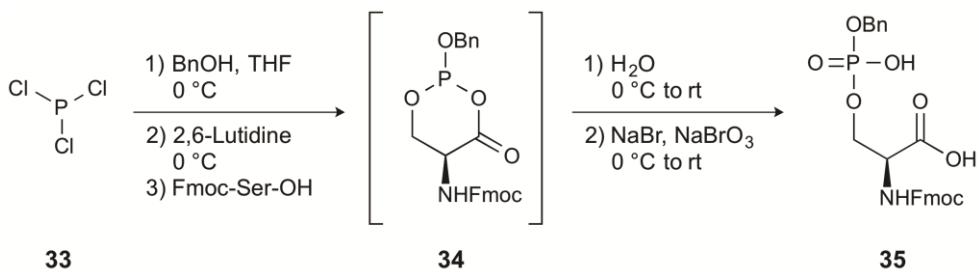
Scheme 3.1 Synthetic route for the preparation of Fmoc-O-benzyl-L-phosphoserine.[120]

3.2.1 Synthesis of Zf3 mutants with incorporated phosphoserine

The aforementioned Fmoc-O-benzyl-L-phosphoserine residue was used in a SPPS approach to synthesize zinc finger mutants containing the latter residue. Considerations concerning the choice of the most valuable incorporation site within the peptide sequence were already subject to section 2.4. In summary, molecular models revealed two positions in the Zf3 peptide that come in close proximity to the phosphodiester backbone of the DNA. This

circumstance was not equally important when compared to the previously described internal dinuclear building blocks but even though, it might also facilitate DNA hydrolysis in the present study. Consequently, phosphoserine was incorporated at the arginine position 70 and the serine position 75.

The use of phosphoserine in SPPS is accompanied by a common side reaction. It was reported that a piperidine mediated β-elimination can occur during the Fmoc deprotection cycle. It eliminates the phosphate group to form 3-(1-piperidinyl)alanyl (Scheme 3.2).[123]
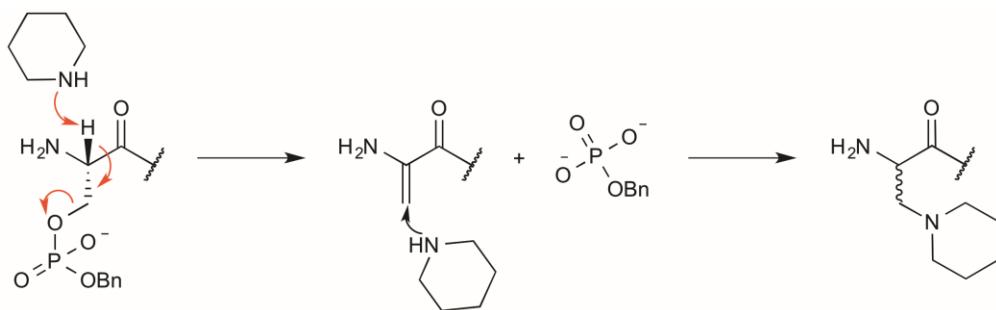
Scheme 3.2 Mechanism of the 3-(1-piperidinyl)alanyl formation from the β-elimination of the phosphoserine residue during Fmoc deprotection with piperidine.[123]

The mechanism is supposed to proceed due to ionization of the unprotected phosphoryl oxygen atom. In doing so, it might decrease the acidity of the proton at the α-carbon making it vulnerable towards the piperidine base. This reaction was successfully suppressed by the use of monobenzyl protected phosphoserine instead of using the fully di-protected derivative. However, a certain proportion of elimination product was observed even with this species. It was confirmed that this side reaction only occurs during Fmoc deprotection when the phosphoserine residue is in the N-terminal position. All subsequent deprotections for the following amino acids are not prone to β-elimination. In addition, the use of microwave energy during Fmoc deprotection to rapidly increase the temperature to 75 °C also contributed to the formation of the side product. Hence, the Fmoc deprotection for the phosphoserine residue was performed at room temperature for 2 x 15 min.

The steric demand of the fully protected phosphoserine residue places particular demands on the coupling cycles for the residue itself but also on the coupling of the subsequent amino acid. Therefore, a literature known procedure was applied for both amino acids, which involved the use of microwave energy at 72 °C and a double coupling procedure with 15 min reaction time for each coupling step.[124] These adjustments to the SPPS protocol successfully suppressed the formation of 3-(1-piperidinyl)alanyl as confirmed by HR-ESI mass spectrometry. Both modified Zf3 mutants (Scheme 3.3) were successfully synthesized and obtained in yields of 41% (**36**) and 27% (**37**).
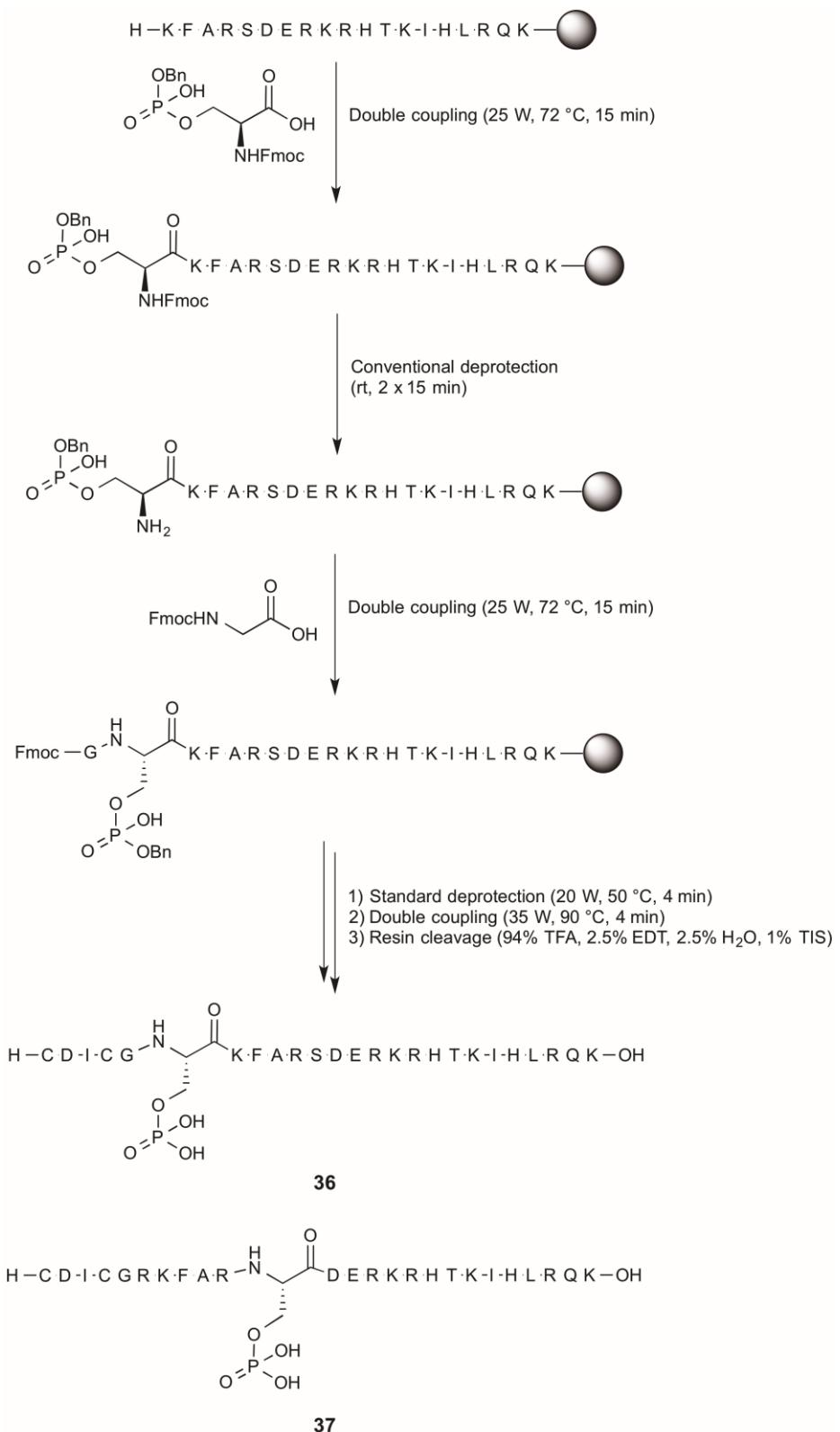
Scheme3.3 Applied SPPS approach for the synthesis of phosphoserine modified Zf3 domains. The complete procedure is shown for peptide **36** with the modification at position 70. Peptide **37** was synthesized in the same manner but with phosphoserine residue incorporated at position 75.

3.2.2 Secondary structure evaluation of phosphoserine modified Zf3 domains

As described in detail in section 2.4.3, prior to all experiments, which focus on DNA binding of the modified zinc-finger domains, it is essential to test for correct secondary structure formation of the peptides by the means of CD spectroscopy. Accurate peptide folding is a basic precondition that ensures proper nucleobase recognition and facilitates DNA-binding by the zinc-fingers. For the fundamental CD-spectroscopic characteristics, which indicate the successful formation of the required $\beta\beta\alpha$-structure upon Zn(II) addition, the reader is referred to section 2.4.3.[83]

To evaluate the secondary structure by means of CD spectroscopy, the phosphoserine modified Zf3 domains **36** and **37** (20 μM) where dissolved in HEPES buffer (20 mM, pH 7.4) and TCEP (500 μM) was added in order to reduce the disulfide bond between the N-terminal cysteine residues. To prevent the reformation of the latter, all solvents were degased and flushed with a gas mixture of 5% hydrogen in nitrogen to apply anaerobic conditions. The CD spectra were recorded in the spectral range of 190 to 250 nm either in absence (dashed lines) or in presence (bold lines) of ZnCl$_2$ (120 μM). Peptide folding was evaluated by comparing the spectra of the engineered zinc-fingers with the spectra obtained for the native reference sample **24**.

The CD spectra of the zinc-finger mutants, which have been modified with phosphoserine at position 70 (red) or at position 75 (blue), are summarized in Figure 3.5 and were compared to the native sequence (black). It can be seen immediately, that the findings for Zf3Pser70 (**36**) much better fit to the reference spectra when compared to peptide Zf3Pser75 (**37**). This concerns both measurements, in absence and in presence of Zn(II). Zf3Pser70 showed a native-like spectrum under metal-free conditions with respect to the location of the negative maximum at 200 nm and of the shallow shoulder around 222 nm. Moreover, upon metal addition, similar bathochromic shifts of the bands at 200 nm to 205.2 nm (native Zf3: 205.5 nm) and equally increased shoulders at 222 nm were observed. In summary, the good correlation with respect to the reference spectra indicated a proper formation of the $\beta\beta\alpha$-structure for peptide **36**.
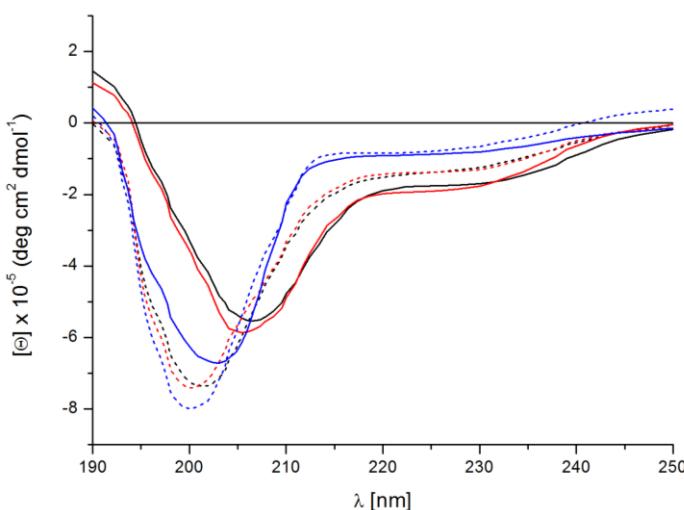
Figure 3.5 CD spectra of Zf3Pser70 (**36**, red), Zf3Pser75 (**37**, blue) and native Zf3 (**24,** black) in absence (dashed lines) and in presence (solid lines) of Zn(II).

In contrast to that, Zf3Pser75 showed deviated spectra under both applied conditions and with regard to the reference peptide. Without the addition of Zn(II), the negative maximum was located at 199.5 nm, and was therefore, slightly more blue-shifted when compared to the reference. In addition, the shoulder at 222 nm was less pronounced what accounts for a disturbed preformation of the zinc-fingers secondary structure. This had become apparent upon Zn(II) addition what caused a just marginally red-shifted band from 199.5 nm to 202.8 nm and a much more decreased shoulder at 222 nm.

Thus, the CD spectroscopic measurements for the Zf3 peptides with phosphoserine incorporated at position 70 and 75 confirmed the findings obtained for the dinuclear buildings blocks, which were incorporated at the same positions (section 2.4.4). It was observed, that position 75 seems not to be suitable for the incorporation of the here described residues. This fact might originate from disturbances triggered by the bulky and charged residues on the neighboring amino acids what prevents correct peptide folding. This situation was not observed in both approaches for the building blocks being incorporated at position 70. These spectra matched well to those of the native reference what clearly speaks for the suitability of this position in order to incorporate the presented building blocks.

3.3 Binding studies between phosphoserine and Ce(IV)/EDTA complexes by microscale thermophoresis

In order to realize a site-specific cleavage of the DNA, it is of great importance to have concentrated amounts of the catalytically-active Ce(IV)/EDTA complex in the surrounding of the phosphoserine moiety of the zinc finger. From this state, a dynamic exchange of the neighboring phosphoserine bound Ce(IV)/EDTA complex to the DNAs phosphodiester backbone is possible and allows for a regionally-limited scissoring of the latter. To do so, the complex recruitment must be caused by attractive forces between the negatively-charged phosphate group of the phosphoserine residue and the positively-charged Ce(IV) ion. In literature, the overwhelming catalytic ability of lanthanides is often attributed to their high binding affinities towards organophosphates.[104] However, reliable values in the form of association or dissociation constants are lacking or differ over a wide range. In order to quantify binding affinities for the recruitment of the Ce(IV)/EDTA complex by the phosphoserine building block, microscale thermophoresis measurements were performed.

Microscale thermophoresis (MST) emerged as a new approach in determining binding constants for a large variety of applications, such as DNA/protein-, protein/protein- and protein/small-molecule interactions.[125] Thus, the method represents an enrichment to the established gold-standard methods, such as isothermal titration calorimetry (ITC) or surface plasmon resonance spectroscopy (SPR) due to the high sensitivity towards molecular interactions under native conditions.[126]

3.3.1 Phosphoserine/(Ce(IV)/EDTA) binding studies

Microscale thermophoresis is based on the motion of molecules along a microscopic temperature gradient referred to as the thermophoretic effect and induced by an IR-laser.[125] By applying the temperature gradient in a time-resolved manner, the movement speed differs with regard to the charge and conformation of the molecule and, more importantly, with its size. Hence, it is possible to differentiate between monomeric or unbound states and oligomeric or bound states due to different molecular dimensions of the examined species. In combining a MST experiment with a titration approach, it is possible to determine affinity constants due to the different migration speeds of an associated complex and dissociated molecules.

The temperature gradient is induced by an IR-laser, which is focused on a glass capillary with small diameter containing the analytes. Simultaneously, a fluorescence excitation beam is coupled in the IR-beam by a dichromatic beam splitter. Fluorescence excitation is necessary in order to detect the movement of exactly one species, unbound or bound, which therefore needs to have an attached fluorophore. Hence, statements concerning molecular

interactions of the latter with other molecules are accessible via the determination of the half maximal effective concentration (EC_{50}), whereby 50% of the fluorescence-labeled molecule is bound, or by calculating the dissociation constant (K_d).[127] A titration experiment is applied in which a serial dilution of the unlabeled molecule is prepared and, subsequently, the fluorescence-labeled molecule is added with constant concentration, respectively. After centrifugation of the probes with 10.000 rpm for 15 min, each probe is soaked into a capillary. One after another, the temperature gradient is applied to each capillary accompanied by simultaneous recording of the molecular movement by means of fluorescence excitation and detection of the corresponding emission. The obtained thermophoretic data for each measured probe includes four major stages (Figure 3.6).[125] The initial stage involves the undirected movement of all molecules in the sample. This is displayed by a plateau in the MST plot with consistent fluorescence. By applying the IR-induced temperature gradient, the molecules start to diffuse out of the heat spot area, which can be observed by a decrease in fluorescence. This is a relatively fast process resulting in a steady state, where out-diffusion and in-diffusion on the edge of the heat-spot area are in an equilibrium state. Deactivation of the IR-beam finally favors the back-diffusion process resulting a jump in fluorescence back to a homogeneous species distribution.
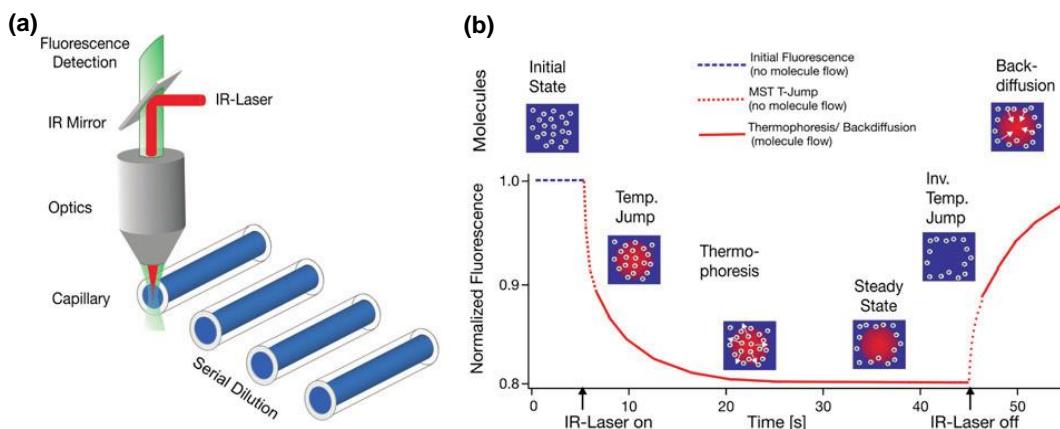
Figure 3.6 Schematic representation of the measurement setup of a dilution series and production data obtained from a typical MST experiment. **(a)** Construction of the measuring cell and **(b)** experimental thermophoresis curve passing through different stages during the measurement.[125]

The major benefit of this set up is the low amount of analytes required for thermophoretic measurements. Typically, low volumes of $10-20\ \mu L$ per individual dilution sample are necessary with a concentration of the fluorescence-labeled species in the nanomolar range. Evaluation of the experimental data is performed by comparing the recorded thermophoretic curves obtained for all diluted samples. Accordingly, at high ligand concentrations, where all labeled molecules are in the bound state, the out-diffusion proceeds slower with respect to low ligand concentrations, whereby all molecules are in the unbound state and therefore

migrate faster. Thus, a plot of fluorescence intensity versus ligand concentration generates a sigmoidal curve, which is subsequently subject of calculating the binding constant K_d according to the following equitation with [L] = ligand, [S] = substrate and [LS] = complex.

$$LS \rightleftharpoons L + S$$

$$K_d = \frac{[L] \cdot [S]}{[LS]} \qquad\qquad (3.1)$$

3.3.2 Evaluation of the binding affinity of the Ce(IV)/EDTA complex towards phosphoserine

In this study, the advantages of MST were used in order to experimentally visualize the binding affinity of the metal complex towards the phosphoserine residue. In various publications, the recruitment ability of phosphoserine for the complex was documented.[116,128] However, there is a lack of substantial values, such as the binding constant, in the literature. Therefore, the here described zinc finger mutants were subject to the experimental evaluation of binding affinities by means of MST.

3.3.3 Preparation of labeled phosphoserine suitable for MST experiments

Initial experiments were performed using only the individual phosphoserine building block instead of the full-length peptide in order to reduce the system to the least number of required components. This was necessary because no reference K_d values were available for the present systems and, moreover, the experimental work with zinc finger proteins requires additional caution on disulfide bond cleavage and Zn(II) addition for a proper formation of the secondary structure. Consequently, the phosphoserine building block **35** was attached to a Rink Amide MBHA resin in order to equip it with a short PEG spacer and a fluorophore (fluorescein isothiocyanate, FITC) by means of solid phase organic chemistry. C-terminal amidation eliminated the negative charge of the carboxylate, which might also contribute to complex binding. The PEG spacer should increase the distance between the recruiting phosphate group and the fluorophore in order to avoid intramolecular or intermolecular interactions between the fluorophore and the molecules taking part in the bonding. This could either be a narrowing of the coordination sphere of the phosphoserine by the spatial demand of the fluorophore in close proximity or a distraction of the Ce(IV)/EDTA complex during the binding event.

Different fluorophores can be utilized in MST experiments as long as they are detectable in the red or blue channel of the device.[129] At first, the phosphoserine residue was modified using 4-chloro-7-nitrobenzofurazan (NBD) because this fluorophore is relatively small in size

and easy to attach to N-terminal amines. During initial MST measurements, a major disadvantage appeared, which finally impeded the use of NBD. The fluorophore was strongly quenched in aqueous media, such as buffer solutions.[130] To compensate this drawback, the concentration of the labeled phosphoserine had to be increased. Consequently, the titrant concentration also had to be increased restricting the setup of a proper dilution series to determine only high dissociation constants. More importantly, the quenching effect only allowed inaccurate concentration determinations of the labeled molecule by means of UV-vis spectroscopy. However, this is crucial for MST experiments, in which concentrations in the low nanomolar range are required. Hence, the NBD fluorophore was exchanged for FITC, which offers high fluorescence intensity without being quenched in aqueous media. The literature reported a dependence of the molar extinction coefficient of FITC from the pH that was well characterized and that was applied to the UV spectroscopic determination of the phosphoserine concentration.[131]

The synthesis was performed on solid support because this method allows for the use of all reagents in excess and, moreover, it allows for the removal of unreacted materials and byproducts by applying a straightforward washing step based on filtration. Thus, the N-terminally Fmoc protected phosphoserine residue **35** was immobilized on the resin using a coupling cocktail containing HATU (4.9 eq) and HOAt (5.0 eq) as activators and DIPEA (9.8 eq) as activator base, under microwave irradiation (35 W, 75 °C) for 20 min. The protecting group was cleaved off by the addition of piperidine (20% in DMF) and by applying microwave irradiation (25 W, 55 °C) for 5 minutes. Subsequently, the Fmoc protected PEG spacer, 8-(Fmoc-amino)-3,6-dioxaoctanoic acid, was coupled to the amino acid in the same fashion. After a further deprotection step, the FITC fluorophore was coupled to the N-terminus of the PEG spacer. The final product was cleaved from the resin under conventional conditions. A cleavage cocktail containing trifluoroacetic acid (TFA), 1,2-ethanedithiol (EDT), triisopropylsilane (TIS) and water (94:2.5:2.5:1, v/v/v/v) was added to the resin, which was placed in a syringe equipped with a polyethylene frit. In this way, the benzyl side-chain protecting group of the phosphoserine residue was simultaneously removed. After concentrating the solution in a nitrogen stream, the crude product was precipitated by the addition of ice-cold diethyl ether. Finally, RP-HPLC purification was performed and the desired product was maintained as C-terminal amide suitable for MST measurements (Figure 3.7, PSer$_{PEG}$FITC: **39**). In addition, a reference probe containing native serine instead of phosphoserine was synthesized in the same manner (Figure 3.7, Ser$_{PEG}$FITC: **38**).
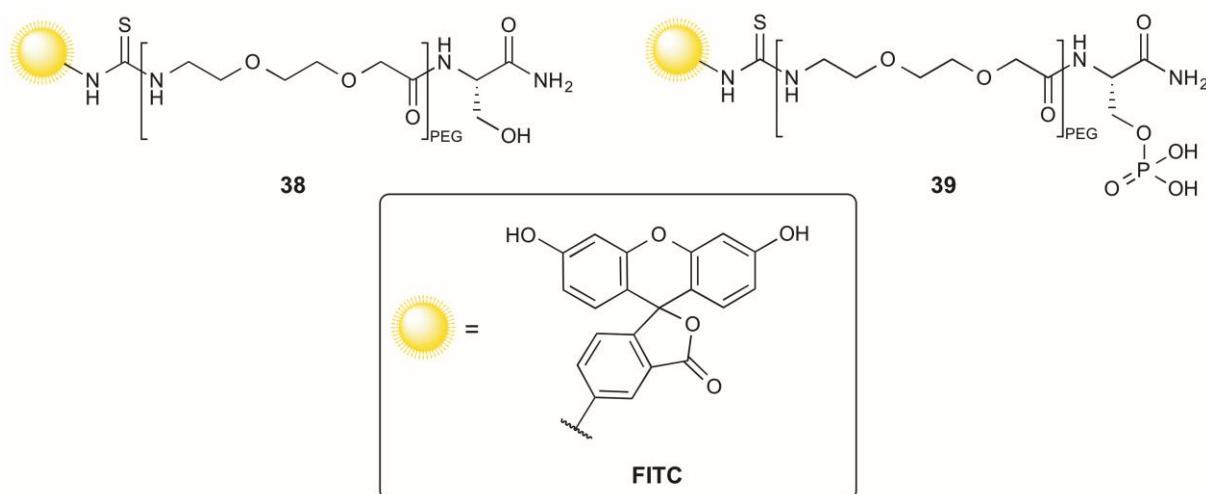
Figure 3.7 Simplified systems used for MST experiments. Serine with C-terminal amide and N-terminal PEG spacer as well as an attached FITC fluorophore (**38**) and equally modified phosphoserine residue (**39**).

3.3.4 Synthesis of the phosphoserine modified zinc finger 3 mutant suitable for MST experiments

The zinc finger peptides used for MST experiments were synthesized by means of microwave-assisted solid phase peptide synthesis (SPPS) on pre-loaded Wang resin. Hence, amino acid coupling including the incorporation of the phosphoserine residue at the arginine position 70 was performed automatically on a Liberty Blue peptide synthesizer (CEM). Coupling of the PEG spacer and the subsequent FITC label was performed manually by applying the same conditions, which were used for the modification of the single-phosphoserine residue described in section 3.3.3. After cleavage of the peptide from the resin with simultaneous removal of all side-chain protecting groups, the peptide was purified by RP-HPLC and lyophilized to obtain Zf3Pser70FITC (**41**) in 15.2% overall yield. In addition, the native peptide without a phosphoserine residue but with the N-terminal PEG and FITC modifications was synthesized in the same manner. The reference peptide Zf3FITC (**40**) was maintained in a final yield of 22.2% (Figure 3.8).
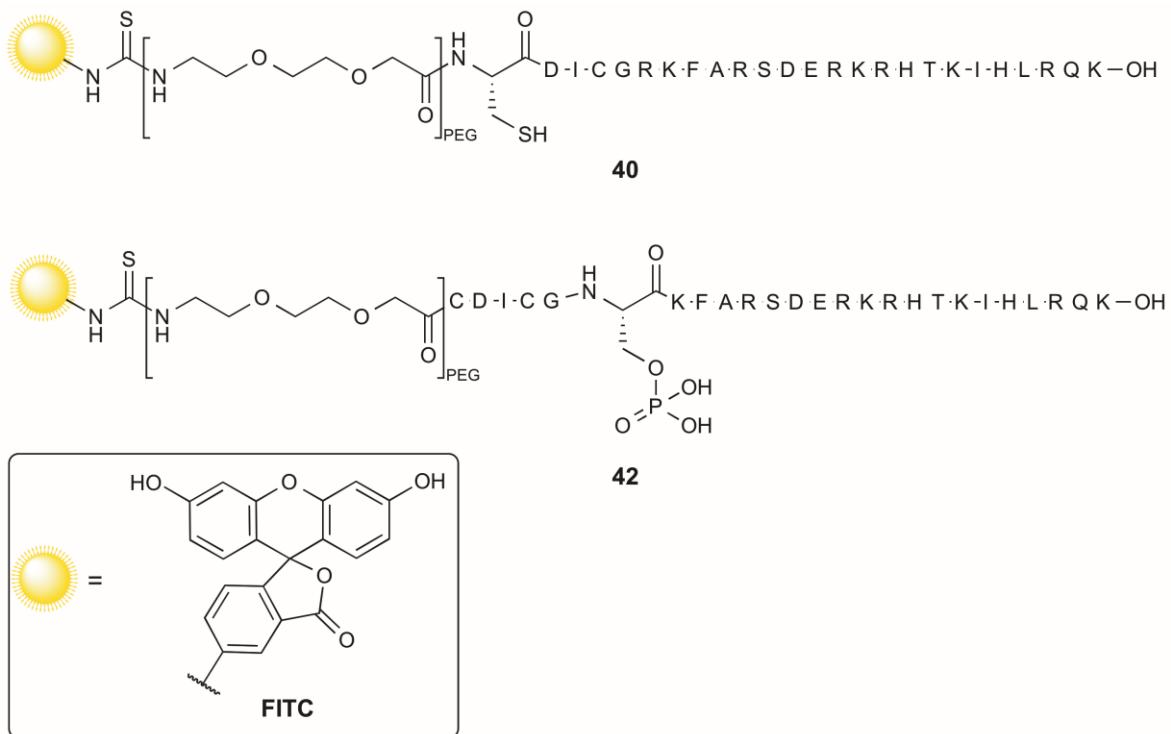
Figure 3.8 Peptides used in microscale thermophoresis experiments bearing a FITC fluorophore, which was connected to the N-terminal cysteine residue by a short PEG spacer. **(a)** Native reference peptide **40**. **(b)** Phosphoserine modified zinc finger 3 with the residue incorporated at position 70 (**42**).

3.3.5 Microscale thermophoresis experiments with labelled phosphoserine and Ce(IV)/EDTA

In order to perform MST measurements, some considerations need to be taken into account. As mentioned in the beginning, the choice of buffers and solvents is crucial for the detection of molecular movement and substrate binding. The use of HEPES buffers is preferred over other buffer systems (i.e. phosphate buffers) due to their low ability to interact with metal ions or even to act as additional chelator besides EDTA.[132] To obtain comparable experimental data for the individual building block and the later discussed experiments with the building block incorporated in the peptide, buffer concentration, pH and ionic strength must be comparable. Therefore, the solubility of the individual building block must be ensured even after the interaction with the Ce(IV)/EDTA complex. Precipitation would lead to a decrease in concentration, and therefore, would result in a decrease in fluorescence intensity of the MST signal. This might lead to distorted titration curves, which consequently give incorrect binding affinities. In the following section, all measurements were performed in HEPES buffer (10 mM, 150 mM KCl, pH 7.5, I = 0.1 M) as it was successfully used in the previous sections. Likewise, the use of Tris buffer and the addition of 3% DMSO or DMF and/or 0.05% Tween20 as detergent was evaluated. But neither the change of the buffer system nor the

addition of organic solvents or detergents influenced the results notably and were therefore neglected.

The quality of the prepared Ce(IV)/EDTA complex is crucial for the generation of reliable data. Thus, the titrant stock solution was prepared by mixing almost equimolar amounts of $(NH_4)_2[Ce(NO_3)_6]$ and EDTA, whereas the latter was used in slight excess to avoid interactions with free Ce(IV) ions.[133] As already mentioned, the addition of ammonium cerium(IV) nitrate in aqueous media yields a highly acidic solution. Increasing the pH of this solution to 7.5 leads to the precipitation of cerium hydroxide. This was circumvented by dissolving the cerium salt in water to prepare a concentrated stock solution. In a second graduated flask, the EDTA disodium dihydrate salt was dissolved in the aforementioned HEPES buffer and the pH was adjusted to 8.5. Addition of the cerium containing solution directly generated the desired complex and the solution remained homogeneous even after the adjustment of the pH to 7.5.

For MST experiments on the simplified system, the previously described phosphoserine residue with N-terminally attached PEG spacer and FITC label (PSer$_{PEG}$FITC, **39**) was used in a final concentration of 60 nM in HEPES buffer (10 mM, 150 mM KCl, pH 7.5). A dilution series of Ce(IV)/EDTA starting at 40 mM in HEPES buffer (10 mM, 150 mM KCl, pH 7.5) was prepared. Therefore, 20 μL of the double concentrated complex containing solution was added to a microcentrifuge tube. A 1:1 dilution was obtained by the addition of 10 μL from the latter probe to a second microcentrifuge tube, which contained 10 μL of the buffer solution. The probe was excessively mixed before another 10 μL were removed and again pipetted into a tube containing 10 μL of the buffer solution. This process was repeated until 16 probes with decreasing Ce(IV)/EDTA concentrations were prepared. Finally, 10 μL of the solution containing the labeled molecule (120 nM) was pipetted to the 16 samples, respectively, to obtain each sample in a final volume of 20 μL with a consistent concentration of 60 nM of the labeled compound.

In theory, the initial fluorescence of each diluted sample must be the same within a 10% range in order to plot binding curves accurately and to minimize the error of the K_d determination (Figure 3.9c).[125] In contrast, the measurements performed for Pser$_{PEG}$FITC showed a different behavior (Figure 3.9a). A linear decrease in fluorescence was observed for titrant concentrations ranging from 40 mM to 2.5 mM. This was followed by a plateau of minimal fluorescence (2.5 mM to 0.6 mM). At higher dilutions, a renewed increase in fluorescence intensity was observed in a sigmoidal manner (0.6 mM to 9.7×10^{-3} mM). Finally, another plateau was found with constant fluorescence intensity for any of the remaining Ce(IV)/EDTA concentrations.

On the one hand, these experimental results are not satisfactory because a major requirement for the application of microscale thermophoresis is a constant fluorescence signal for each dilution step in a range of ±10%. On the other hand, a change of fluorescence in a sigmoidal manner is another indication for the occurrence of a binding event (Figure 3.9a). In order to validate the theory of a binding event, the reference sample $Ser_{PEG}FITC$ was examined similarly. From the constant fluorescence signals for each dilution step within the required ±10% range (Figure 3.9b), it can be concluded that the data did not reveal any fluorescence dependency on the titrant concentration.
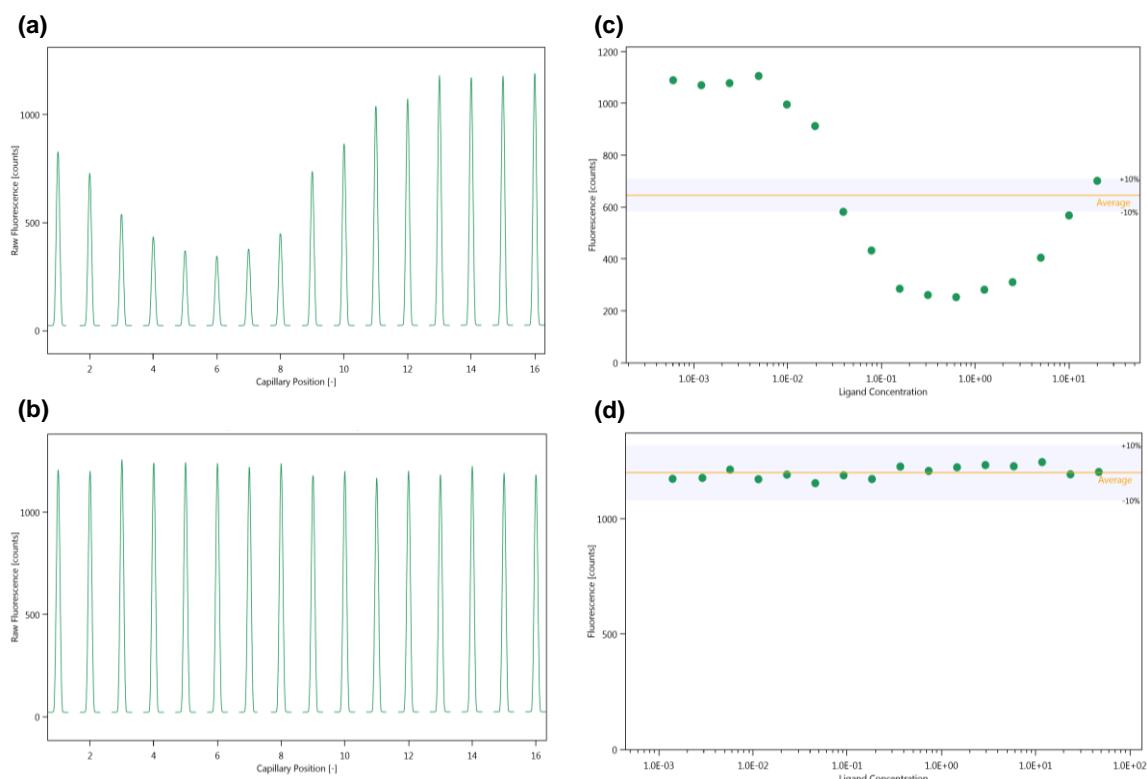
Figure 3.9 Fluorescence charts obtained from capillary scans for the dilutions series of $Pser_{PEG}FITC$ **(a)** and $Ser_{PEG}FITC$ **(b)** with the Ce(IV)/EDTA titrant solution. The initial fluorescence intensities of $Pser_{PEG}FITC$ **(c)** and $Ser_{PEG}FITC$ **(d)** were plotted as a function of the titrant concentration with the highlighted ±10% range.

It can therefore be concluded that the change in fluorescence intensity resulted from the binding event of the Ce(IV)/EDTA complex to the phosphoserine moiety. It was reported that high valent metals, such as cerium (IV), can influence the electronic environment of the fluorophore.[134] This event is referred to as "cold fluorescence" being a frequently occurring event during MST experiments.[135,136] For this reason, the determination of the dissociation constant is not accessible by MST. However, the latter can be determined by a devious route using the change in fluorescence intensity instead of using the molecular movement of the species. The observed increase in fluorescence represented the dissociation of the [Ce(IV)/EDTA]/phosphoserine complex due to the continuing separation of the fluorescence quenching Ce(IV)/EDTA complex from the phosphoserine moiety. The determination of the

affinity constants via this route was accessible using the original MST approach since the data acquisition and processing survey stayed the same. Therefore, the MST measurements were performed at two different MST power settings, 40% MST power and 95% excitation power and 80% MST power and 95% excitation power. In addition, the starting concentration of the titrant solution was varied to cover more data points to enhance a reliable curve fitting procedure. The starting concentrations for the titrant solutions were 40 mM, 35 mM, 30 mM and 25 mM.

According to the method described above, each solution was used in a 1:1 dilution series with the FITC labeled phosphoserine (PserPEGFITC) and the native serine (SerPEGFITC). All measurements were performed in triplicates and plotted with standard deviations.
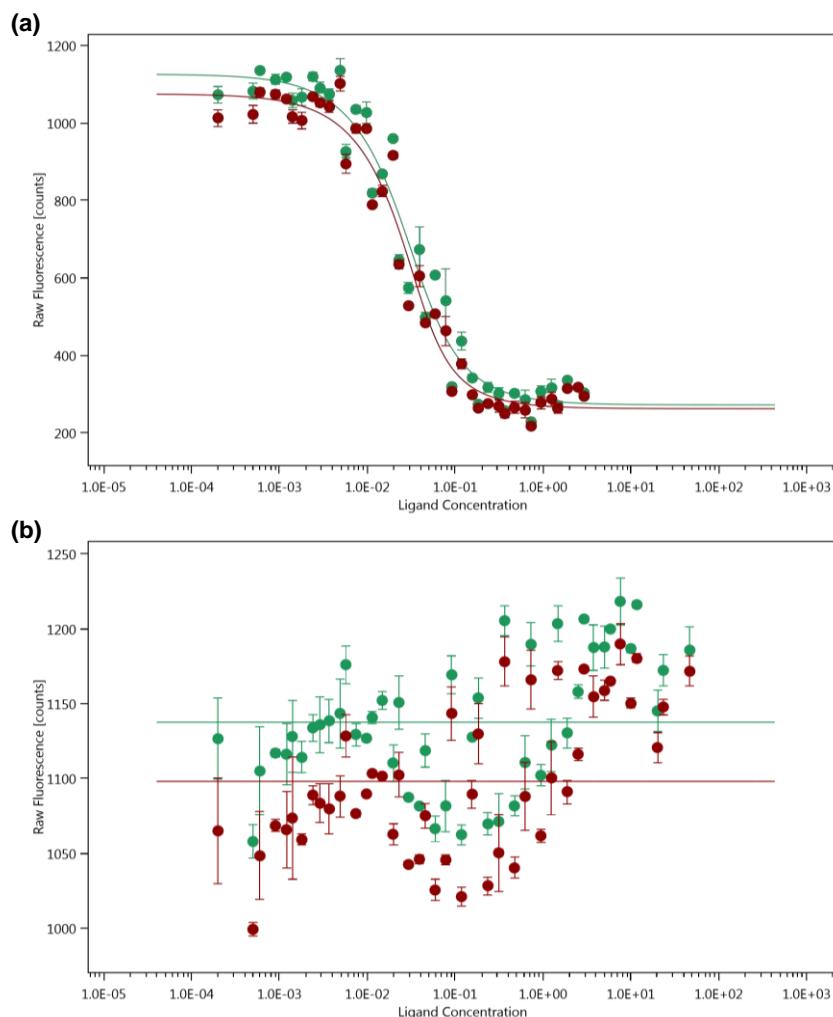
Figure 3.10 Measurement of the FITC fluorescence signal as a function of the Ce(IV)/EDTA addition. **(a)** Experimental data-points obtained for PserPEGFITC at 40% MST power and 95% excitation power (green) and 80% MST power and 95% excitation power (red) involving non-linear curve fits. **(b)** Experimental data-points for the reference probe SerPEGFITC at 40% MST power and 95% excitation power (green) and 80% MST power and 95% excitation power (red).

The experimental data points obtained for Pser$_{PEG}$FITC (Figure 3.10a) and Ser$_{PEG}$FITC (Figure 3.10b) were plotted as a function of the Ce(IV)/EDTA concentration. In doing so, the reference probe showed an inhomogeneous distribution of the fluorescence intensity at both applied MST power settings. The data points were fluctuating within a very narrow range of about 150 fluorescence counts. The applied fits did not reveal any significant trends in complex association or dissociation as it had been expected for a molecule without the recruiting phosphate group. In contrast, Pser$_{PEG}$FITC showed the aforementioned binding curves with low fluorescence intensities at high titrant concentrations, representing the bound-state, and high fluorescence intensities at low titrant concentrations, representing the unbound-state. At high initial titrant concentrations, the fluorescence intensity first decreased (Figure 3.11, gray dots). This was interpreted as an artifact of oligomeric Ce(IV)/EDTA species and was therefore neglected from the actual dissociation event. Due to this effect, the curves, however, were not ideally sigmoidal. In order to apply reasonable non-linear fits, the affected data-points were excluded from the plots. Hence, the resulting sigmoidal binding curves obtained at the two different MST power settings were fitted in order to determine the dissociation constants, respectively. Thus, the K_d at 40% MST (Figure 3.11, green) was determined to be $13.7 \pm 0.7 \mu M$ and the K_d at 80% MST power (Figure 3.11, red) was determined to be $8.7 \pm 0.6 \mu M$.
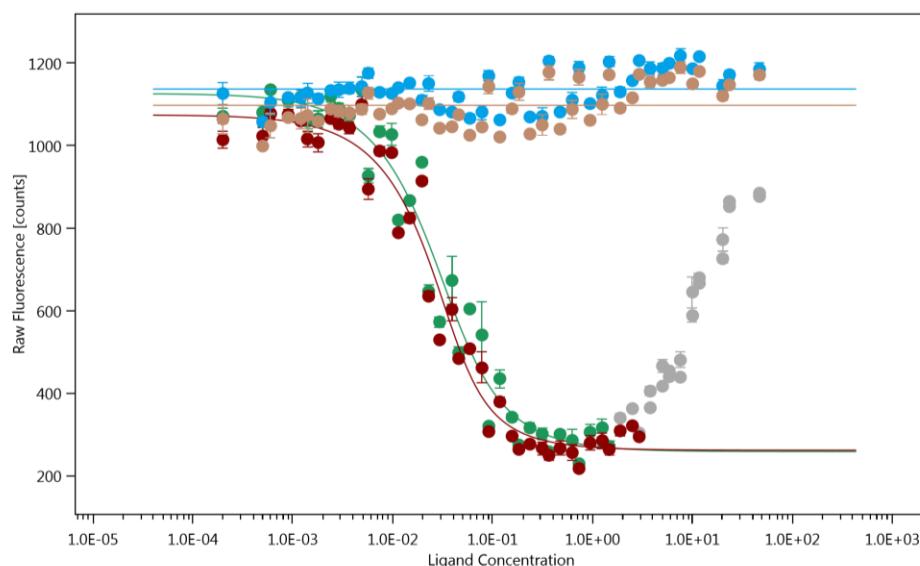
Figure 3.11 Overlaying spectra of Pser$_{PEG}$FITC at 40% MST power (green) and 80% MST power (red) and Ser$_{PEG}$FITC at 40% MST power (blue) and 80% MST power (ochre). Data-points at higher titrant concentrations were neglected as indicated (gray dots) in order to apply a non-linear curve fit.

3.3.6 MST measurements of phosphoserine modified Zf3 mutants

The determination of the binding constant of the individual building block revealed a significant binding of the phosphoserine moiety to the Ce(IV)/EDTA complex. Hence, it was investigated whether the building block incorporated in the zinc finger motif would be able to do the same under native conditions. These included the addition of Zn(II) for appropriate peptide folding. It was known that binding constants for individual building blocks and building blocks incorporated into peptides might show differences because of several conditions. For instance, these conditions result from differences in the hydrophobic/hydrophilic environment, the accessibility of the binding group due to secondary structure formation and electron donating or withdrawing effects of neighboring amino acids.

The peptide Zf3Pser70FITC (**42**) was modified by exchanging the arginine residue at position 70 for a phosphoserine moiety (Figure 3.8). Additionally, the N-terminus was equipped with a PEG spacer in order to provide some distance between the zinc-coordinating N-terminal cysteine residue and the FITC fluorophore. For comparison reasons, the native peptide Zf3FITC (**40**) was synthesized containing the same N-terminal modifications but without the phosphoserine residue incorporated (Figure 3.8).

Unfortunately, zinc coordination is a major problem when it comes to MST measurements. The addition of zinc salts to buffered solutions with neutral or slightly alkaline pH leads to a continuing precipitation of insoluble zinc hydroxide.[137] This process is rather slow but the solid material disturbs the homogeneity of the dilution series, and therefore, negatively affects the fluorescence signal in each capillary. To circumvent this problem, buffered solutions with additional weak metal chelators, such as tricine, were prepared.[138] The advantage of these additives over strong metal chelators, such as EDTA, are their rather low binding constants for divalent zinc ions (e.g. tricine: 10^{-6} M contra EDTA: 10^{-16} M).[139] Hence, metal ions are just loosely bound in order to keep them in solution and make them available for the much stronger binding zinc finger (Zf: 10^{-10} M)[69]. Thus, the aforementioned HEPES buffer was modified by the addition of 10 mM tricine. Moreover, $ZnCl_2$ was just added in slight excess (360 nM) with regard to the used amounts of labeled zinc finger peptides (120 nM) to ensure complete occupation of all metal binding sites in the molecule. Furthermore, this proceeding also excluded the presence of unfolded peptide which might contribute to a false determination of the K_d.

Consequently, the solutions remained homogeneous and no significant influence of precipitated zinc hydroxide on the fluorescence intensity was observed in time-depended fluorescence examinations. Pursuing the same objective, the addition of the macromolecular crowding agent PEG 8000 was tested. This is a polyethylene glycol moiety with an average molecular mass of 8 kDa, which loosely binds Zn(II) ions to keep them in solution.[140] In

principle, it acts as an elongated and opened crown ether. Moreover, it mimics the macromolecular crowding in a cell, which is supposed to contribute to the overall structural stability of proteins in a liquid environment.[141] In comparative studies, PEG 8000 was added to the aforementioned measurement buffer in a final concentration of 5% (m/v). It was found that the addition of the additive has no significant influence on the experimental quality. The results were consistent with those obtained by using only the tricine modified buffer. Consequently, the use of PEG 8000 was abandoned and all experiments were performed without the addition of the latter.

Prior to the experiments, the zinc finger peptides (1.2 μM) were dissolved in the sample buffer containing TCEP (2.5 μM) and ZnCl$_2$ (2.5 μM) and agitated for 90 min at room temperature in order to reduce the disulfide bond. Afterwards, the stock solutions were diluted to the final peptide concentrations of 240 nM, respectively. The preparation of the dilution series with the Ce(IV)/EDTA solution was performed according to the method described in section 3.3.5.

Initial capillary scans for both peptides revealed a similar behavior with regard to the concentration dependent changes in fluorescence intensity as observed for the simplified system (Figure 3.12).
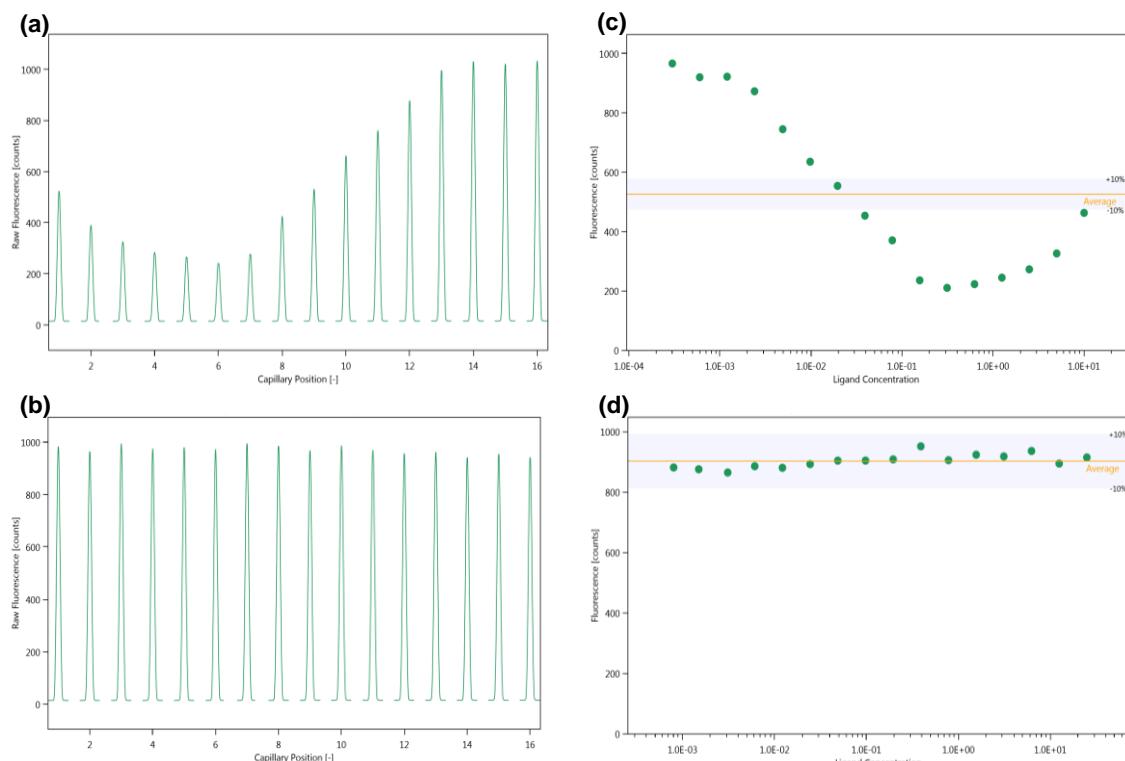
Figure 3.12 Fluorescence charts obtained from capillary scans for the dilutons series of Zf3Pser70FITC **(a)** and Zf3FITC **(b)** at different Ce(IV)/EDTA titrant concentrations. The initial fluorescence intensities of Zf3Pser70FITC **(c)** and Zf3FITC **(d)** were plotted as a function of the titrant concentration with highlighted ±10% range.

While the fluorescence of Zf3Pser70FITC (Figure 3.12, a and c) increased during the titration, the fluorescence of the reference sample Zf3FITC remained almost constant within the required ±10% range (Figure 3,12, b and d).

Consequently, the formation of the Zf3Pser70FITC/[Ce(IV)/EDTA] complex was accompanied by the same fluorescence quenching effect as observed before. The occurrence of the cold-fluorescence effect was again observed for high titrant concentrations, which belonged to the bound-state. With decreasing amounts of the Ce(IV)/EDTA species and, subsequently, with the dissociation of the complex, the intrinsic fluorescence of the fluorophore was recovered. This was assumed to be due to the spatial separation of the analytes at low titrant concentrations. Thus, the determination of the dissociation constants for the latter complex was performed in the previously described manner using the changes in fluorescence intensity. Due to the fact, that the general MST approach includes fluorescence measurements by default, the method could be maintained. Solely the data processing and interpretation had to be adjusted.

To cover a wide range of data points in order to map the curve progression in detail, Ce(IV)/EDTA solutions with starting concentrations of 40, 35, 30 and 25 mM were prepared. The latter were used in a 1:1 dilution series, respectively, whereby the concentration of the labeled peptides (120 nM) was kept constant. MST experiments were performed in duplicates at 40% MST power and 60% excitation power and at 80% MST power and 60% excitation power. The experimental data points were plotted as a function of the titrant concentration and a non-linear fit was applied.

The native reference sample (Zf3FITC) showed no dependence of its fluorescence on the titrant concentration for any of the applied power settings (Figure 3.13). The fluorescence signals detected for each diluted capillary remained almost the same. These findings demonstrated that the native reference zinc finger was not able to bind the Ce(IV)/EDTA complex.
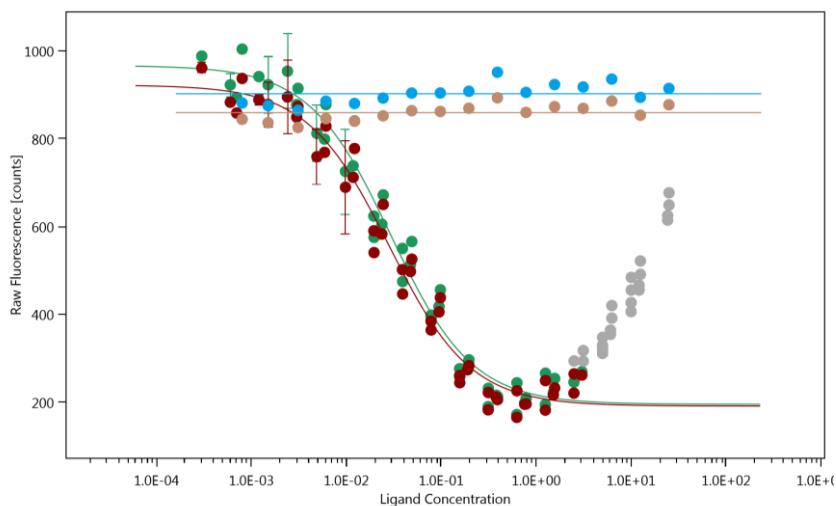
Figure 3.13 Overlaying spectra of Zf3Pser70FITC at 40% MST power (green) and 80% MST power (red) and Zf3FITC at 40% MST power (blue) and 80% MST power (ochre). Data-points at higher titrant concentrations were neglected as indicated (gray dots) in order to apply a non-linear curve fit.

On the contrary, the phosphoserine-modified zinc finger **(42)** showed a similar curve progression as observed for the simplified system. In order to determine the K_d values for the different power settings, the data points at higher titrant concentrations were neglected (Figure 3.13, gray dots). Thus, a sigmoidal binding curve was obtained and a non-linear fit could be applied. The determination of the dissociation constants resulted in a K_d of 28.2 $\pm$ 0.7 μM at 40% MST power (Figure 3.13, green) and 27.4 $\pm$ 0.6 μM at 80% MST power (Figure 3.13, red).

In summary, the binding affinity of the phosphoserine residue towards the Ce(IV)/EDTA complex was evaluated in two different approaches. In the first approach, the residue was treated in an individual manner, without being incorporated into the zinc finger. This was carried out in order to reduce the system to the lowest number of components required. Initial capillary scans for MST measurements revealed that the thermophoretic effect was superimposed by cold-fluorescence. This represented an electronic influence exerted by the charged Ce(IV)/EDTA complex on the fluorophore. Thus, the dissociation constants could not be calculated by using standard molecular movement considerations due to disturbing titrant-concentration dependent changes in fluorescence intensity. However, the fluorescence increase observed for progressive dilution of the titrant, corresponds to the dissociation of the [Ce(IV)/EDTA]/phosphoserine complex. This was confirmed by control experiments with the native serine analogue, which did not show a similar effect. Consequently, it was possible to determine the dissociation constants by an approach, which uses these changes in fluorescence intensity.

In the second approach, the phosphoserine residue was incorporated at the arginine position 70 to generate the zinc finger mutant Zf3Pser70FITC (**42**). Hereby, the suitability of this system under native conditions was examined. This involved the accurate formation of

the zinc fingers secondary structure upon Zn(II) addition and, moreover, the tolerance of the latter during the addition of the Ce(IV)/EDTA complex. The effect of cold-fluorescence was still observed and, therefore, the changes in fluorescence intensity were used again for the determination of the dissociation constant.

Thus, the experiments revealed an existing attractive force exerted by the phosphoserine residue towards the hydrolytically active Ce(IV)/EDTA complex. It was not possible to draw conclusions from the determined K_d values whether they are high or low due to the lack of published dissociation constants for similar systems. However, it is to note that the determined constants are acceptable for an artificial system that relies on attractive forces between two oppositely charged species. Reversely, a higher K_d can actually be of use with respect to the required relocation of the metal complex from the phosphoserine to the phosphodiester backbone in order to promote hydrolysis. Hence, the tendency of phosphoserine to accumulate the hydrolytically active species in a limited area was confirmed. The discrepancy between the dissociation constants obtained from both approaches can be explained by the different conditions found in the peptide. Neighboring amino acids might shield the binding moiety or could electronically influence the binding event.

Nevertheless, this approach was originally designed to use microscale thermophoresis to determine the dissociation constants. Due to the occurrence of cold-fluorescence, the thermophoretic effect had been superimposed by changes in fluorescence intensity. In order to create a system, which actually is suitable for MST, the fluorophore needs to be attached to a site in the sequence in far distance to the phosphoserine residue. The PEG spacer used in this study was inserted between the N-terminal residue and the fluorophore to avoid undesired interactions of the latter. However, it turned out that this could not be prevented by the flexibility of the PEG spacer. Moreover, it can be assumed that the PEG spacer further contributed to the occurrence of the cold-fluorescence effect. The long and flexible PEG chain enabled the fluorophore to come in close proximity to the phosphoserine residue. An attachment of the fluorophore to the side chain of the C-terminal lysine might be more beneficial in this context.

The findings clearly account for the use of phosphoserine to recruit external Ce(IV)/EDTA complexes. The often suggested ability to bind the latter complex was documented and substantiated with solid dissociation constants. The determined K_d values, especially those for the individual phosphoserine residue, can contribute to the design of artificial restriction enzymes in different future approaches. For this trial, the verification of the binding ability of the zinc finger mutant towards Ce(IV)/EDTA was very promising with regard to a possible DNA cleavage. In conjunction with the sequence-specific DNA binding of the zinc finger, the

scission might be constrained to distinct cleavage sites within the surrounding of the phosphoserine moiety.

3.4 Preparation of phosphoserine modified Zf13 peptides

The two modified zinc finger 3 peptides bearing a phosphoserine residue either at the arginine position 70 (**36**) or at the serine position 75 (**37**) were used in a native chemical ligation approach with the recombinant expressed Zf12 thioester **27**.[89] According to the method described in section 2.5.2, the Zf3 mutants were dissolved in a denaturing phosphate buffer (Na$_2$HPO$_4$/NaH$_2$PO$_4$, 10 mM), which contained guanidinium chloride (6 M) as denaturing agent. The disulfide bond between the N-terminal cysteine residues was reduced by the addition of TCEP (5 mM) at pH 4 for 2 h. Subsequently, the Zf12 thioester **27** was added as a solution in phosphate buffer (10 mM, 6 M GdmCl) and the pH of the reaction mixture was adjusted to pH 8. The reaction progress was followed by analyzing small aliquots by RP-HPLC. When the reaction has gone to completion, the phosphoserine modified full-length peptides Zf13Pser70 (**43**) and Zf13Pser75 (**44**) were purified by semi-preparative RP-HPLC. The peptides were lyophilized and stored at –22 °C under an argon atmosphere until they were needed (Figure 3.14).
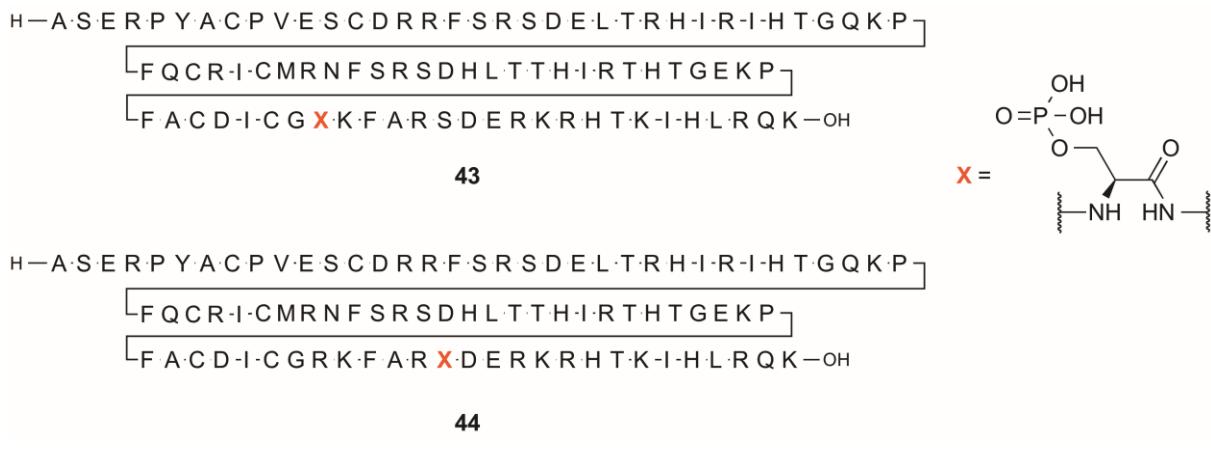
Figure 3.14 Phosphoserine modified Zf13 mutants Zf13Pser70 (**43**) and Zf13Pser75 (**44**) obtained after NCL of the corresponding Zf3 peptide with the Zf12 thioester.

3.5 DNA binding studies with phosphoserine modified Zf13 peptides

According to the method described in section 2.6, the DNA binding ability of the engineered zinc fingers was tested by means of polyacrylamide gel electrophoresis under native conditions.[95] Thus, the duplex DNA was prepared by annealing the two single-stranded oligomers 5'-GTGTGTGTGTGATCTTAGCGTGGGCGTAAG-3' and 3'-CACACACACACT-AGAATCGCACCCGCATTC-5'. The detection of a peptide/DNA complex was ensured by the attachment of a 5'-FAM label to one of the oligomers. The peptides (1 mM) were reduced with an aqueous solution of TCEP (5 mM) for 90 min and subsequently dialyzed against a

zinc containing buffer solution (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 10 μM ZnCl$_2$, pH 7.8) for 4 h. Prior to use, all solvents were degased and flushed with an anaerobic gas mixture containing 5% hydrogen in nitrogen in order to prevent reformation of the disulfide bonds. Excess amounts of Zn(II) were removed by dialysis against a buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 1 μM ZnCl$_2$, pH 7.8) having a particular low Zn(II) concentration.

Distinct peptide to DNA ratios ($r_f = C_{Zf13}/C_{DNA} = 0, 1, 5, 10, 30$) were prepared in order to quantify the binding abilities of the engineered zinc fingers towards their target DNA sequence. The latter were prepared in a sample buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 1 μM ZnCl$_2$, pH 7.8) and incubated at room temperature for 3 h. The further addition of the Ce(IV)/EDTA complex had been renounced to exclude misleading results obtained from possible hydrolysis events or other interferences. Thus, the samples were pipetted into the wells of the polyacrylamide gel and electrophoresis was performed using a non-denaturing running buffer (25 mM Tris, 192 mM glycine, pH 8.2). Figure 3.15 shows the PAGE findings for the examined peptides.
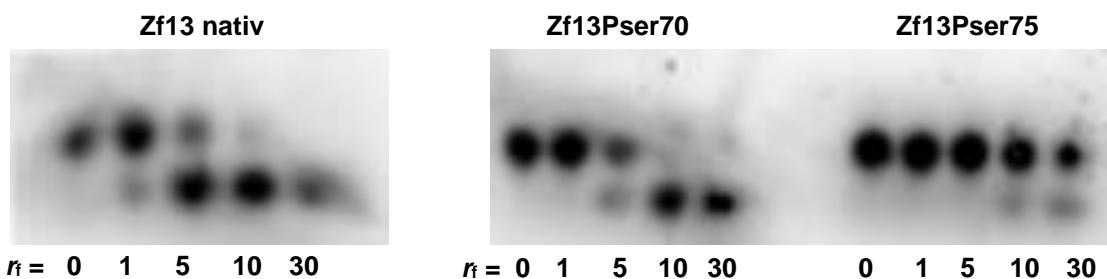
Figure 3.15 Native polyacrylamide gel electrophoresis for the evaluation of zinc-finger/DNA binding. The consensus dsDNA (0.5 μM) was incubated with the zinc finger peptides Zf13 native (**28**), ZF13Pser70 (**43**), and Zf13Pser75 (**44**) in different concentrations ($r_f = C_{Zf13}/C_{DNA} = 0, 1, 5, 10, 30$) at room temperature for 2 h.

As already demonstrated in section 2.6, the vast majority of DNA binding sites were occupied with the native peptide at approximately $r_f = 5$. This was concluded from the disappearance of the individual DNA band. The gel mobility shift assays for the zinc finger peptides with the phosphoserine residue incorporated either at position 70 or at position 75 showed similar complex formation abilities as observe before for the dinuclear complexes (section 2.6). Zf13Pser70 (**43**) showed pronounced DNA binding, especially at ratios of $r_f = 10$ and 30. The disappearance of the faster migrating individual dsDNA band indicated the complete occupation of all DNA binding sites by the peptide. It is to note, that nearly complete DNA occupation with **43** occurred at a slightly higher ratio ($r_f = 10$) with regard to the native zinc finger ($r_f = 5$). In contrast to that, Zf13Pser75, which was modified at position 75 did not show any indications for sequence-specific DNA binding or DNA binding in general. For ratios up

to $r_f = 10$, there was no complex formation detected what corresponds to the observations for the dinuclear building blocks incorporated at the same position. It was assumed, that the still appearing weak complex bands at $r_f = 10$ and 30 mainly resulted from electrostatic interactions but the unbound DNA was still the predominant species. This was not surprising considering the deviated CD spectroscopic findings for peptide **37**. In conjunction with the comparable results obtained from the PAGE studies with the dinuclear building blocks incorporated at position 75, it can be concluded, that this position seems to be unsuitable for any of the herein described modifications. In contrast to that, correct secondary structures as well as acceptable DNA binding affinities were documented for all modifications done at position 70.

3.6 DNA cleavage studies with phosphoserine modified Zf13 peptides and Ce(IV)/EDTA

In a similar approach, the site-specific hydrolysis ability of the Ce(IV)/EDTA complex towards double-stranded DNA was under evaluation using denaturing PAGE.[97] In general, the latter relies on the increased accumulation of the hydrolytically active metal complex at the phosphoserine species, which occupies a distinct position in the peptide. Thus, upon binding of the zinc finger mutants to the dsDNA, the phosphate side-chain of phosphoserine comes in close proximity to the phosphodiester backbone of the DNA. This situation promotes the relocation of the phosphoserine bound Ce(IV)/EDTA complex to the neighboring phosphodiester backbone. Consequently, DNA fragments of predictable molecular size should originate from the cleavage that have a different migration speed in the PAGE assay in comparison to the intact DNA single-strands.[142]

Prior to the experiments, peptides **43** and **44** were prepared according to the method described in section 2.6. After protein dialysis, incubation was carried out in a sample buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, 1 μM $ZnCl_2$, pH 7.8) including the zinc finger mutants and the consensus duplex DNA at room temperature for 3 h. Based on the findings obtained from the previous DNA-binding studies (section 3.5), the most suitable peptide to DNA ration ($r_f = 20$) was applied and kept constant in order to examine the influence of different Ce(IV)/EDTA concentrations on the DNA cleavage. Hence, a buffered stock-solution of Ce(IV)/EDTA (20 mM, 20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, pH 7.8) was added to the Zf13/DNA samples to obtain the metal complex in final concentrations of [M] = 10, 20, 50, 100 or 200 μM. The probes were incubated under the exclusion of molecular oxygen at 37 °C for 72 h before denaturing conditions were applied by preparing a 1:1 dilution of the samples with a double-concentrated solution containing SDS (2%), glycerol (25%, v/v), EDTA (2 mM) and DTT (180 mM) in Tris buffer (20 mM Tris, 150 mM NaCl, 0.5 mM TCEP, pH 7.8). The mixtures were heated to 95 °C for 5 min and immediately loaded on the gel after cooling on ice. Electrophoresis was performed using a denaturing running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.2) and the migrated DNA was visualized by a fluorescence imager, which detected the attached 5'-FAM label. The latter was either attached to the binding strand (Figure 3.16), which included the operator sequence for the zinc finger peptides, or to the opposite strand (Figure 3.17). Consequently, the aforementioned procedure was performed in parallel for the two differently labeled dsDNA samples under the same condition.
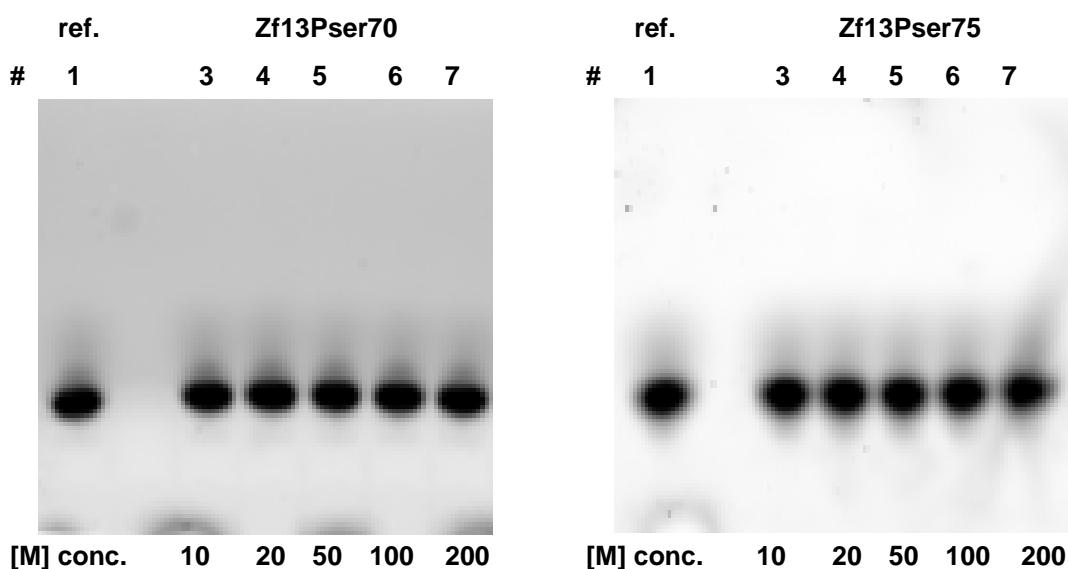
Figure 3.16 Denaturing PAGE studies to examine the cleavage ability of the phosphoserine modified zinc fingers towards their dsDNA (0.5 μM) target-site. The peptide to DNA ratio was fixed to $n = C_{Zf13}/C_{DNA} = 20$ for all samples. The Ce(IV)/EDTA concentration was varied ([M] = 10, 20, 50, 100 or 200 μM). Incubation was performed at 37 °C for 72 h. The single-stranded DNA oligomer containing the operator sequence was 5′-labeled with a FAM fluorophore in order to visualize fragmented DNA.

Figure 3.16 shows the findings for the experiment with dsDNA having the label attached to the binding strand. Well 1 in each gel represents the reference, which contained only the dsDNA without further addition of peptide or Ce(IV)/EDTA. Wells 3, 4, 5, 6 and 7 contained the engineered zinc fingers, either with the phosphoserine residue incorporated at position 70 or 75 as well as with different Ce(IV)/EDTA concentrations ([M] = 10, 20, 50, 100 and 200 μM). It is conspicuous that all migrating bands were on comparable levels, which also included the reference. This strongly accounts for consistently intact single-stranded DNA samples that were not prone to hydrolysis over the contemplated reaction time. This result was also observed for the similar experiment including the dsDNA bearing the fluorescent label at the opposite strand (Figure 3.17). The experiments were repeated with a prolonged incubation time of 120 h instead 72 h but with the same experimental result. Also a mass spectrometric examination of the samples after the incubation showed predominantly the intact oligomers.
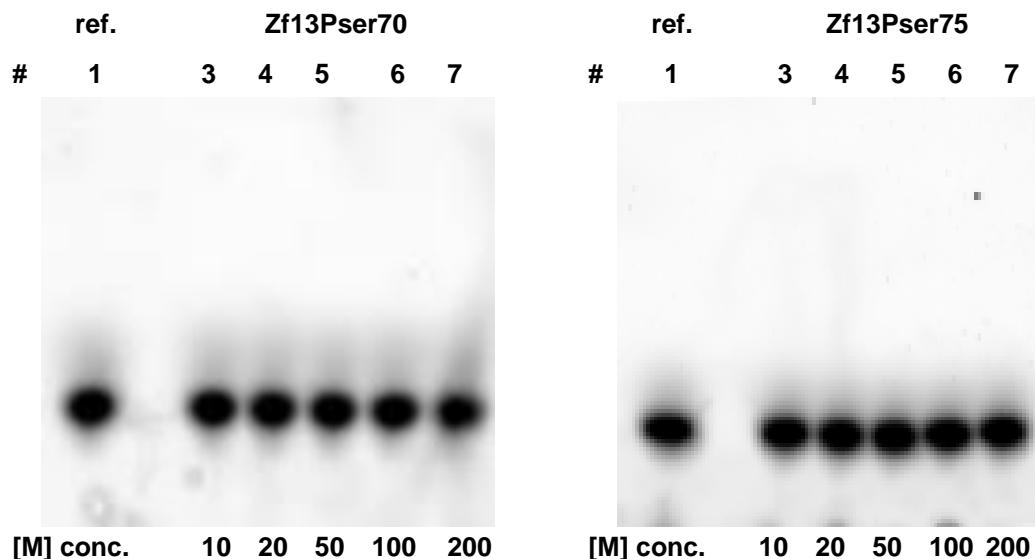
Figure 3.17 Denaturing PAGE studies to examine the cleavage ability of the phosphoserine modified zinc fingers towards their dsDNA target-site (0.5 μM). The peptide to DNA ratio was fixed to $r_f = C_{Zf13}/C_{DNA} = 20$ for all samples. The Ce(IV)/EDTA concentration was varied ([M] = 10, 20, 50, 100 or 200 μM). Incubation was performed at 37 °C for 72 h. The single-stranded DNA oligomer opposite to the operator-sequence-containing strand was 5'-labeled with a FAM fluorophore in order to visualize fragmented DNA.

The limited cleavage ability for the zinc finger bearing the phosphoserine at position 75 can be at least partially attributed to its distorted secondary structure. The latter was presumably responsible for the limited DNA-binding ability as confirmed by native PAGE. For this reason, it is not surprising that Zf13Pser75 had no tendency to cleave DNA because the zinc finger was not able to recruit the catalytically active species to the desired target site. The absence of cleavage products in the PAGE studies observed for Zf13Pser70 with the phosphoserine residue at position 70 is not so general to answer. Possible reasons are summarized in section 3.7 and a discussion is subject to section 4.

3.7 Summary

This section focused on the concept of modifying the DNA-binding zinc finger domain by the incorporation of a phosphoserine residue in order to recruit external Ce(IV)/EDTA complexes to a desired target site. The documented high hydrolysis capacity of lanthanide complexes should be thus utilized to trigger DNA cleavage.[104] The lack of reliable values that describe the molecular attraction of the phosphoserine moiety towards the Ce(IV)/EDTA complex was successfully overcome by the establishment of a microscale thermophoresis approach. In doing so, the phosphoserine residue was coupled to a PEG spacer, which was equipped with a FITC fluorophore. A similar system with unmodified serine served as a reference sample and both moieties were used to determine the dissociation constants for the corresponding complexes with Ce(IV)/EDTA by means of fluorescence quenching. It was clearly demonstrated that only the phosphorylated serine residue was able to form a complex with the hydrolytically active species, whereby it was possible to calculate a profound dissociation constant. This simplified system served as benchmark for the evaluation of the phosphoserine modified zinc finger 3 domain. It was found that the dissociation constant obtained for the Zf3/[Ce(IV)/EDTA] complex was in a comparable range implying an increased accumulation of the lanthanide complex in close proximity of the phosphoserine residue. These findings were beneficial in conjunction with the binding of the zinc finger to its consensus DNA sequence, which predetermines possible hydrolysis sites with regard to the location of the phosphoserine in the peptide sequence. Thus, CD spectroscopic studies were performed in order to examine the influence of the unnatural amino acid on the secondary structure formation of the zinc finger peptide. As previously observed for the incorporation of the internal dinuclear building blocks (section 2.4.4), the incorporation at the arginine 70 position provided best results with respect to the formation of the required $\beta\beta\alpha$-structure. In contrast, the incorporation at the serine 75 position resulted in a deviated CD spectrum, which indicated a negative influence of the phosphoserine on the peptide folding.

Moreover, the phosphoserine modified Zf3 mutants were successfully used in an EPL approach to generate the full-length three tandem zinc-finger domain of Zif268 with enhanced DNA binding ability. With regard to the native zinc finger reference sample, PAGE studies under non-denaturing conditions successfully revealed a similar peptide/DNA complex formation ability for the engineered zinc finger with the phosphoserine residue incorporated at position 70. The disturbed peptide folding ability, which was found for the mutant modified at position 75 was further proven due to its apparently poor ability to interact with DNA.

A similar experiment should shed light on the hydrolysis activity of the engineered zinc finger peptides or to be more precise, on the site-specific cleavage by the recruited Ce(IV)/EDTA

complex. Hence, the dsDNA samples were incubated with the ZF-mutants and different amounts of the hydrolytically active Ce(IV)/EDTA complex were added. After the incubation at 37 °C and for different reaction times varying from 72 – 120 h, there was no indication for phosphodiester hydrolysis observed after gel electrophoresis. A possible explanation might be a limited substrate accessibility of the recruited lanthanide species. The preformed zinc-finger/DNA complex could be to compact to additionally allow the binding of the bulky Ce(IV)/EDTA complex to the partially shielded phosphoserine residue. During the MST experiments, solely the individual zinc fingers were examined for their ability to bind the metal complex. The additional use of DNA was neglected due to the possible occurrence of a peptide/DNA-derived binding event, which might superimpose the actual target. The changed conditions after DNA binding might have caused a lowered accessibility of the phosphoserine residue leading to an unspecific movement or binding of the active species. Consequently, the utilization of its full hydrolytic potential, even at elevated concentrations, was inhibited. The actual conditions after the peptide/DNA complex formation need to be evaluated by means of X-ray crystallography in order to develop a detailed understanding. Due to the fact that the hydrolysis capacity of the examined cerium complex was multiply demonstrated in literature, this gives reason to review the structural conditions, such as the incorporation site of the phosphoserine residue in the zinc finger in order to establish a functional system. However, the results obtained from this approach can help to refine the current system to develop a powerful artificial restriction endonuclease with enhanced site-specificity.

4. Discussion and conclusions

The different approaches described in the previous two sections shared the overall goal of modifying the DNA-binding zinc-finger motif of Zif268 with an additional DNA hydrolyzing moiety. The development of such artificial restriction endonucleases, which have a predictable and site-specific DNA cleavage ability is of great importance for the establishment of novel gene-editing systems and for gene-therapeutic applications.[144,145] This was aimed to be achieved by applying two different methodologies.

The first method relied on the synthesis of tyrosine-based artificial amino acids, which mimic the dinuclear active-sites of natural nucleases.[146] Two building blocks suitable for SPPS were successfully synthesized and incorporated into the peptide sequence of the third zinc-finger subdomain of Zif268 at two distinct positions. The second method relied on a different concept. Instead of directly introducing the hydrolytically active species into the peptide sequence, an external Ce(IV)/EDTA complex should be attracted by a peptide-incorporated phosphoserine residue.[116] Both approaches should allow for site-specific hydrolysis of the phosphodiester backbone of the DNA either by direct phosphate binding and subsequent hydrolysis by the internal building blocks or by accumulating the external lanthanide complex in a narrow area in-between the phosphoserine residue and the phosphate backbone. The different zinc-finger mutants were successfully synthesized following a semi-synthetic approach, which included the direct modification of Zf3 by means of SPPS followed by the generation of the full-length peptide by means of NCL with the expressed Zf12 domain. A native PAGE approach was used in order to test the ability of the engineered zinc fingers on binding their consensus DNA target sequence. The obtained results were in excellent agreement with the native peptide except for modifications, which were done at position 75. This was referred to unsuitable intrinsic conditions for the bulky building blocks, such as the neighboring arginine and aspartate residues, which have long and charged side chains and are furthermore located in the rigid α-helical part of the zinc finger.[69] These findings are very valuable for the future selection of appropriate incorporation sites for artificial amino acids in zinc-finger proteins. However, both approaches showed poor hydrolysis abilities towards natural dsDNA. This result was quite surprising due to the fact that especially the dinuclear building blocks were beforehand examined with DNA model substrate, which was clearly hydrolyzed. One particular reason for the deviant behavior of the building blocks when incorporated into the zinc finger might be a limited substrate accessibility. Despite the fact that the positions were chosen on the basis of molecular models, the actual realities might differ. Moreover, the documented unsuitability of position 75 for the incorporation of the building blocks further reduced the number of potential hydrolysis candidates. Thus, the remaining position has an ideal secondary structure formation but suffers from insufficient substrate activation that might consequently prevent hydrolysis although the system might be

generally able to do so. In order to evaluate this hypothesis, crystal structures of the present peptides in complexes with DNA can help to shed light on the actual locations and orientations of the building blocks. This would furthermore contribute to the development of a functioning zinc-finger-based nuclease with an incorporated second metal-binding site.

The aforementioned considerations are also true for the second approach, where nevertheless even another possibility has to be taken into account. This concerns the phosphoserine residue used, which might be susceptible to enzymatic digestion under physiologic conditions.[147] The presence of phosphatases could cause a loss of the phosphate group that ultimately would reconstitute native serine. Thus, the latter would not have any tendency to attract the hydrolytically active species as confirmed by the MST experiments described here. In order to circumvent this problem, the phosphate group of phosphoserine needs to be modified in a way to resist enzymatic digestion. Phosphorylated amino acids have a phosphate monoester moiety (C-O-P) and this readily undergoes hydrolysis in a biological environment that contains suitable enzymes. Potent phosphate mimics were found in phosphonates, which have tremendously more stable C-C-P bonds that resist enzymatic cleavage (Figure 4.1).[148]

Figure 4.1 Phosphoserine (*left*) is prone to enzymatic digestion, which results in the loss of the phosphate group. Phosphoserine mimics, such as (α,α-difluoroalkyl)phosphonates (*right*), provide stable C-C-P bonds, which resist dephosphorylation. Difluorination at the α-carbon generates a phosphate-like polarity and a similar steric demand.

Moreover, fluorination of such alkylphosphonates at the α-carbon compensates for the lost oxygen atoms and perfectly surrogates the polar character and steric demand of the original phosphate group.[149] Phosphoserine mimics, which are furthermore suitable for SPPS, are well established and are used in a broad variety of applications, such as in phosphatase inhibition studies.[150] The use of such compounds in the previously described approach could eliminate the possibility of enzymatic digestion and would provide a stable recruiting moiety over the complete course of DNA hydrolysis.

One important limitation of both approaches has to be emphasized. This concerns the general interaction mechanism of natural restriction enzymes towards a DNA substrate in comparison to the here presented artificial nucleases. Natural enzymes use an induced-fit mechanism for substrate binding and interaction. This concept dates back to the work of KOSHLAND in 1958 and describes the cooperative effect when enzymes are binding to their

substrates.[151] Thus, initial interactions are assumed to be rather weak but after target-site recognition, both moieties pass through a series of conformational changes, which strengthen the binding.[152] A distinction can be made in two different substrate-binding modes. On the one hand, there is a uniform binding mechanism that implies strong substrate binding and, on the other hand, there is a differential binding mechanism that exclusively has a strong effect on transition-state binding.[153] Consequently, the activation energy for the reaction is reduced. The induced-fit mechanism is often attributed to large enzymes, which are able to undergo conformational changes because of their relatively flexible protein scaffold. Moreover, they are also capable to induce conformational changes to their bound substrate by establishing additional contacts to the latter. It is important to note that the induced-fit mechanism only describes enzyme-substrate interactions without allowing conclusions regarding the actual hydrolysis process.[154] The latter is dependent on different energetic considerations, such as the Gibbs free energy of phosphodiester activation, as mentioned at the beginning (cf. Figure 3.2).[104] Nevertheless, the stated conformational changes assist in the preorganization of the substrate and the active-site of the enzyme and thus, achieves the correct alignment of both moieties. The modified zinc finger mutants described here have a lack of prerequisites, which would allow to mimic such a process. However, this does not generally argue against DNA hydrolysis by zinc-finger-based nucleases as, for example, confirmed by the SUGIURA group.[155] In their approach, they placed a peptide-based Ce(IV) binding loop between two distinct zinc finger proteins (Figure 4.2) and incubated the artificial nuclease with its dsDNA target. Their findings revealed an increased cleavage preference in a narrow range within the DNA sequence, which directly faced the aforementioned lanthanide binding loop. Although the cleavage rates were found to be rather low, the study confirmed the general applicability of such approaches. Thus, future ventures on the development of artificial restriction endonucleases might have to focus on a more native-like substrate/active-site interaction in order to encourage hydrolysis and to increase cleavage rates.

Figure 4.2 Zinc finger-type artificial nuclease with an inserted cerium-binding peptide for the site-specific cleavage of double-stranded DNA.[155]

However, the combination of different approaches for the development of artificial restriction endonucleases is a vigorous field, which comprises multiple natural science disciplines. The work on functional small-molecule model complexes, which mimic the active sites of

hydrolases were used in this study in order to combine them with the overwhelming sequence specificity of the zinc finger proteins. Future improvements can contribute to the enrichment of the range of well-established technologies, such as zinc-finger nucleases or ARCUTs.

5. Zinc-finger-based peptidyl metal sensors

Since their discovery in 1982, zinc finger proteins (ZFPs) have frequently been subject to multiple tailor-made modifications.[156] First attempts tried to use the overwhelming specificity of ZFs towards multiple DNA sequences. After the identification of the pattern behind this recognizing ability, the way was paved for the development of artificial restriction enzymes.[4] These zinc-finger nucleases (ZFN) were developed by combining the DNA-binding domain of a zinc-finger motif with a DNA-cleaving endonuclease. Thus, ZFNs emerged as fully programmable genome editing tools, whereby almost every desired DNA sequence can be addressed.

Furthermore, the unique folding behavior and secondary structure formation of zinc fingers upon metal binding were recognized as very useful characteristics for the development of peptidyl chemosensors.[157] The sensing mechanisms of most chemosensors rely on the detection of a specific interaction between a receptor molecule and a target analyte. For this reason, zinc fingers offer an ideal scaffold for the development of peptidyl metal sensors due to their remarkable metal binding ability and particular secondary structure formation.[69] Based on the approach of NADLER, fluorophore-labeled zinc fingers were synthesized to detect and distinguish between different divalent transition metal ions.[158] The sensor assembly described here differs from other published zinc-finger-based sensors. In this approach, a metal-coordinating histidine residue of zinc finger 3 (Zf3) of the Zif268 domain was exchanged for a propargylglycine residue. This allows for the *in-situ* generation of a 1,2,3-triazole by the means of click-chemistry. The triazole moiety mimics the metal coordination properties of the imidazole side chain of histidine and moreover, bears a fluorophore derived from an azide-modified precursor. The conjugated fluorophore undergoes metal-specific changes in its electronic environment due to the bound metal ion in close proximity. These changes may be observed as an increase or decrease in fluorescence intensities and can be used for the distinction between different transition metals.

5.1 Structural properties and metal coordination abilities of zinc fingers

DNA-binding proteins, such as Zif268, occur as tandem repeats of several individual zinc fingers. Each zinc finger sub-unit consists of 20 to 30 amino acids and features enormously high affinity towards divalent zinc ions. Zn(II) complexation by four distinct residues leads to the formation of a simple but highly stable $\beta\beta\alpha$-structure, which is composed of an antiparallel β-sheet hairpin followed by an α-helix.[10] Both motifs are connected by a short loop region and held together by an intermediate Zn(II) ion. The stability of this globular structure is further enhanced by hydrophobic residues, which form a hydrophobic core in the

interspace between both structural motifs. The folding process is made possible by a highly conserved peptide sequence in which the coordinating residues and the hydrophobic residues hold significant positions. Tetrahedral metal coordination is predominantly attained through cysteine and histidine ligands. Depending on their composition and frequency, ZFs are subdivided into different types, such as the Cys_4-, Cys_3His- or the most commonly found Cys_2His_2-type.[69]

The dissociation constants for Cys_2His_2-type zinc finger complexes with Zn(II) are extremely low and lie in the range of 10^{-11} to 10^{-9} M.[159] However, several studies revealed that ZFs are also able to bind many other transition metal ions, such as Co(II), Ni(II), Cu(II) or Fe(II) with comparable secondary structure formation but with lower binding affinities with regard to Zn(II).[160] This can be attributed to the redox inactive d^{10} configuration of Zn(II), which is ideal for structural purposes as the coordinating ligands remain unaffected. Moreover, Zn(II) is a Lewis acid and accepts electron pairs from donor ligands, such as histidine and cysteine.[161] Its vacant $4p4s^3$ orbitals only allow for occupation with four individual donor electron pairs prescribing tetrahedral coordination. In combination with the entropic benefit of exchanging six water ligands for one tetradentate ZF-chelator, the listed characteristics explain why especially Zn(II) is the most suitable metal within these proteins. In contrast, zinc fingers interact poorly with cations, such as Na^+, K^+, Mg^{2+} and Ca^{2+}, which are found in high concentrations in the cell lysate.[69] This is explained by a coordination mismatch between these ions and the zinc finger due to the "hard and soft acids and bases" (HSAB) theory. The ions are considered to be "hard" due to their small ionic radii, and therefore, prefer "hard" ligands, such as carboxylates, over "soft" ligands, such as the provided thiol groups of the cysteine residues. In this context, Zn(II) ions are intermediating between hard and soft with a particular preference for histidines over cysteines.[162] However, Zn(II) has the best overall prerequisites in terms of bond lengths and coordination geometry to perfectly fit into complexes with zinc fingers.

5.2 From small molecule metal sensors to peptidyl metal sensors

The detection of a specific interaction between a receptor molecule and a target analyte requires an explicit and prompt visualization method. A well-established sensing method involves the use of fluorophores, which respond to an interaction by changes in their fluorescence intensities or by showing spectral shifts of their emission wavelength.[163] Diverse zinc-finger-based fluorescence metal sensors can be found in literature.[157,164] However, most approaches rely on the metal-induced peptide folding and the generation of the ββα-structure. An extensively studied system is based on the use of a FRET pair, which shows a distance-dependent energy transfer from a donor molecule to an acceptor fluorophore (Figure 5.1a).[165] Secondary structure formation leads to an increased or

decreased Förster distance between both moieties that is accompanied by a change in fluorescence from which the actual distance can be derived. Another approach established by the IMPERIALI group used an environmental sensitive fluorophore, which was attached to a residue that is part of the hydrophobic core of the peptide (Figure 5.1b).[166] Under metal-free conditions, the zinc finger is unfolded and the fluorophore exposed to the polar environment of the solvent. Upon metal-induced peptide folding, the latter moves to the less polar environment of the hydrophobic core and shows a different fluorescence response.
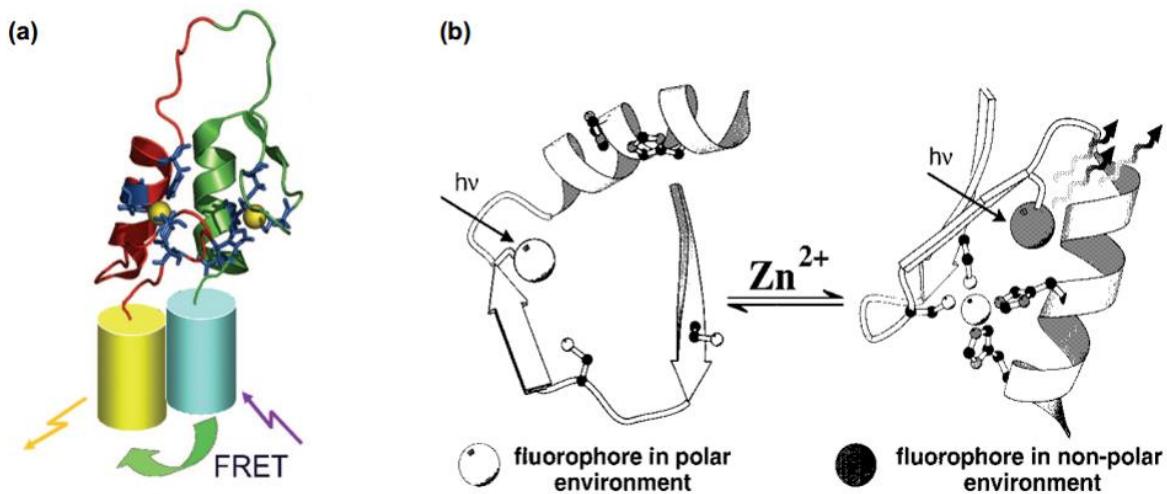
Figure 5.1 Folding and structure dependent zinc-finger fluorescent metal sensors. **(a)** FRET-based sensor. **(b)** ZF-sensor with incorporated environmental sensitive fluorophore.[167,157]

As can be seen, the above-mentioned examples have a lack of distinguishability between different metal ions due to the same $\beta\beta\alpha$-structure formation upon complexation of various divalent metal ions. For this reason, the fluorophores used in this study were directly conjugated to the metal-coordination site of the peptide. This concept was successfully used in the development of multiple small molecule metal sensors. Some prominent examples, which also include the herein described dansyl fluorophore, are summarized in Figure 5.2.[167,168]

It is to note that the BANDYOPADHYAY group synthesized a histidine-based fluorescent sensor for the selective detection of Hg(II) (Figure 5.2, right).[168] They coupled a dansyl fluorophore to the imidazole side chain of histidine and attached two of these moieties to a bipodal thiocarbamate scaffold. The complexation of Hg(II) caused a 19-fold enhancement of the fluorescence intensity, which was significantly greater in comparison to other tested metal ions. However, this type of imidazole modification is challenging due to the formation of a mixture containing the N(ε) and N(δ)-labeled products, which are difficult to separate. More importantly, the sulfonamide bond between the fluorophore and the pyrrole-like nitrogen of the imidazole ring is very labile and the use of strong acids, such as TFA, would cause the

elimination of the fluorophore. However, this reagent is required in SPPS for the cleavage of the peptide from standard Wang or Rink-Amide resins and for the removal of all side-chain protecting groups. A change to the very acid labile 2-chlorotrityl chloride resin, from which peptides can be cleaved under extremely mild acidic conditions with trifluoroethanol (TFE), is in conflict with the protecting group strategy. 2-Chlorotrityl chloride resins would give fully side-chain protected peptides after cleavage, which are unsuitable for metal-binding assays. These facts finally restrict the applicability of direct dansylation of histidines rather to the synthesis of small molecule metal sensors than to peptidyl sensors.
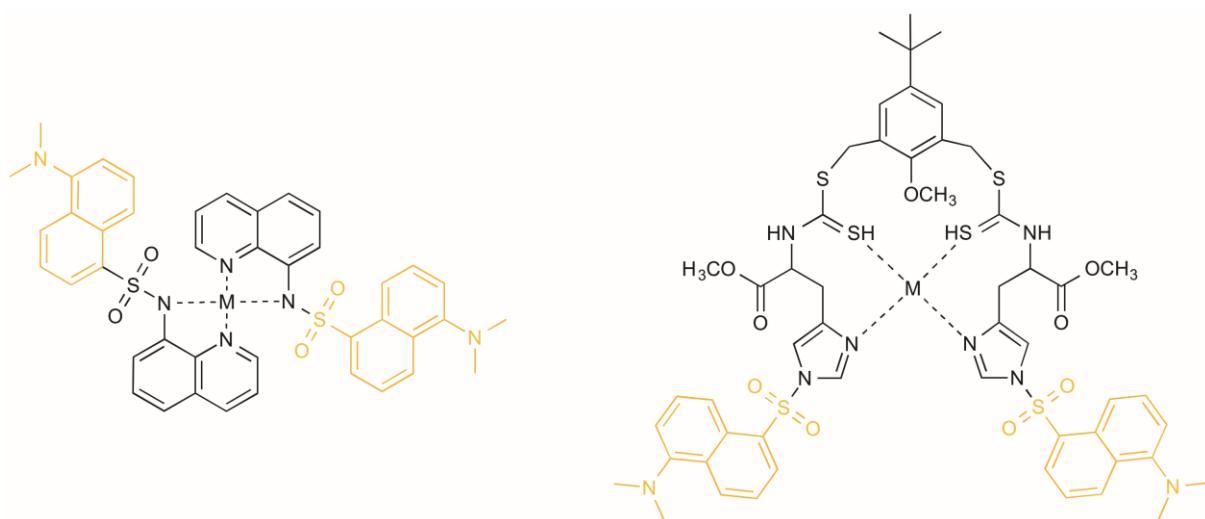
Figure 5.2 Small molecule fluorescent metal sensors: Asymmetric half-*salen* ligand with attached dansyl fluorophore (*left*) and bipodal chemosensor with dansyl modified histidines (*right*). M = metal ion.[167,168]

5.2.1 1,2,3-Triazoles as mimics for histidine

A method that elegantly compensates for this disadvantage emerged with the discovery of click-chemistry by SHARPLESS in 1998.[169] In this approach, 1,2,3-triazoles can be generated from a concerted Huisgen [3+2] cycloaddition between an alkyne moiety and an azide group. The use of copper(I) catalysts (CuAAC) exclusively generate 1,4-regioisomers whereas the use of ruthenium(II) catalysts (RuAAC) only provides 1,5-substituded triazoles.[170] The outstanding feature of triazoles generated in this way manifests in the similar metal-coordination capacity with regard to the imidazole side chain of histidine.[171] Furthermore, it can be easily modulated through the choice of the used azide moiety, which is attached in closed proximity to the metal coordination site with high stability.[172]

Among all proteinogenic amino acids, L-histidine has specific significance for protein structure formation and the activity of biomolecules.[173] Its imidazole side chain is an aromatic motif with an acidic ionization constant around $pK_a = 6.5$. Hence, it acts as a

hydrogen bond donor or acceptor and is used as ligand in biomolecules for a broad variety of metal ions, such as Zn(II), Cu(II) or Ca(II).

The imidazole of histidine is a five-membered heterocycle composed of a trigonal pyrrole-like nitrogen (Nε) with two electrons in an unhybridized p-orbital and a trigonal pyridine-like nitrogen (Nδ) with a lone pair in a hybrid orbital and a single electron in the p-orbital. Metal coordination is predominantly performed by the N(ε) nitrogen because it offers the only completely unshared electron pair (Figure 5.3, *left*). In contrast, metal coordination by the N(δ) atom would negatively affect the aromaticity of the system because the π-electrons at this position are part of the aromatic sextet. Binding through the N(ε) site is additionally favored due to the electron-withdrawing effect induced by the pyridine-like nitrogen, which is lowering the basicity at this position.[174]
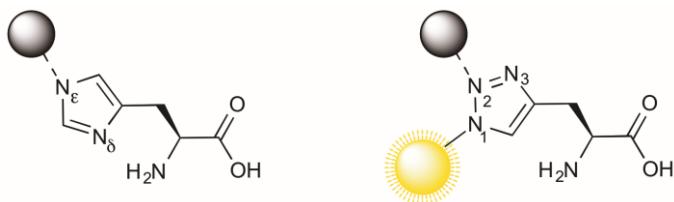
Figure 5.3 Comparison of the metal binding properties of naturally occurring L-histidine (*left*) and CuAAC-derived triazoles generated from L-propargylglycine and azide-modified fluorophores (*right*).

Histidine mimicking triazoles can be generated from L-propargylglycine and an azide-bearing moiety in a copper(I)-catalyzed azide-alkyne cycloaddition (Figure 5.3, *right*). In this approach, azide-functionalized fluorophores were used establishing a connection to the N1 nitrogen atom of the triazole moiety after the click-reaction. The resulted artificial amino acid offers two possible nitrogen atoms, which are suitable for metal binding. Monodentate metal coordination can be achieved by either the N2 nitrogen atom or the N3 nitrogen atom of the heterocycle.[172] N2 coordination is similar to the N(ε) binding mode of histidine, while N3 coordination would rather correspond to the unusual N(δ) binding mode. Especially for zinc fingers, it is known that tetrahedral metal coordination is just observed for N(ε) binding that would correspond to the involvement of the N2 nitrogen of the triazole moiety.[69] Metal binding via the N2 atom generates "reverse" ligands due to the lower electron density at this position, while "regular" ligands are observed for N3 coordination.[175] Although there are more examples published for regular ligands, it has been shown that the peptide scaffolds, such as the zinc finger, constrain metal binding to the energetically unfavorable N3 atom.[176]

An approach of BERG *et al*. gave rise to the assumption that the exchange of histidine for a single triazole is unlikely to have significant negative impact on zinc finger folding.[177] This assumption is based on experiments with a C-terminally shortened zinc finger peptide in

which the last metal-coordinating histidine residue was truncated (Figure 5.4). Metal binding studies revealed that the vacant coordination site was automatically reoccupied by either a single water molecule or by other external ligands, such as *N*-methylimidazole, to recover the tetrahedral coordination sphere. More importantly, the dissociation constant of the mutated zinc finger complex was just within one order of magnitude higher compared to the consensus sequence.
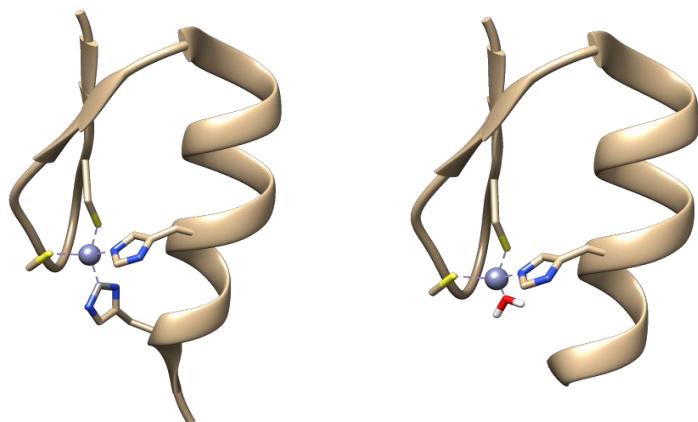
Figure 5.4 Native Cys$_2$His$_2$ zinc finger with tetrahedral metal coordination sphere (*left*) and modified zinc finger with truncated C-terminal histidine (*right*). The vacant position was automatically substituted by an external ligand (e.g. H$_2$O). The models were generated with UCSF Chimera (PDB code 1AAY).

In addition, there is a large body of literature found for different small molecule metal sensors with fluorophores coupled to triazoles.[178,179] These models demonstrate the suitability of both nitrogens of the triazole, the N2 atom and the N3 atom, for metal coordination depending on the provided ligand scaffold and the coordination geometry of the metal ion. For instance, ZHU *et al.* synthesized a nitrogen-rich ligand scaffold based on a tridentate bis-(2-picolyl)amine (BPA) ligand.[180] The latter was equipped with either a propyne group (Figure 5.5, *left*) or an azidomethane moiety (Figure 5.5, *right*) and used in a click reaction with the corresponding counterpart, respectively. Thus, the resulted triazoles were compelled to coordinate the metal ion with either the N3 or the N2 nitrogen atom in compliance to the local coordination geometry. The determined stability constants for both complexes revealed only minor differences that further supports the assumption drawn from the BERG approach saying that zinc-finger incorporated triazoles would adapt to the locally found conditions.

Moreover, it has been observed that metal-induced changes in the electronic environment of the ligands are effectively transferred to the fluorophore even across several bonds. The WANG group synthesized an anthracene-based fluorescence sensor for the specific detection of Cu(II).[181] In this approach, the click-derived triazole moiety simply acts as a connector between the fluorophore and two metal-coordinating iminodiacetate groups (Figure 5.5).

However, complexation of Cu(II) led to a 5-fold increased fluorescence quenching compared to other metal ions despite the fact that the anthracene dye was several bonds apart from the coordination site.
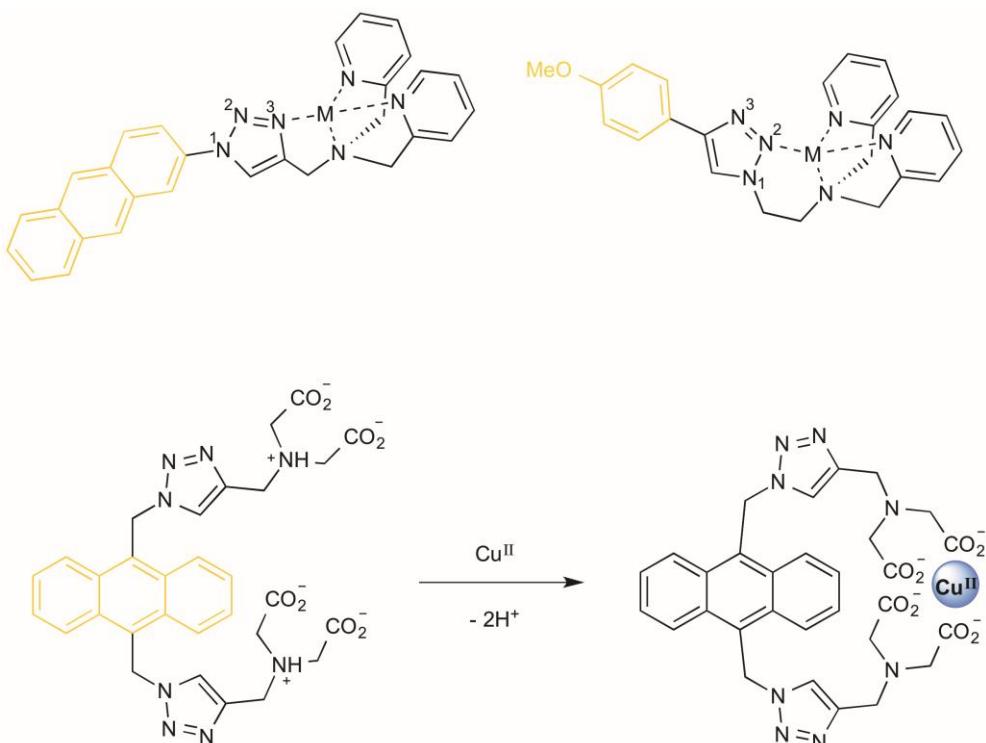
Figure 5.5 Fluorescence sensors using either the N3 nitrogen atom or the N2 nitrogen atom (*above*) of triazole for metal coordination. Fluorescence quenching of anthracene upon Cu(II) complexation across several bonds (*below*).[180,172,181]

5.3 Theory and mechanisms of fluorescence metal sensing

The energy transfer mechanism for the discussed sensors is predominantly attributed to metal-induced stimuli in the coordination environment upon complex formation. In contrast to the well-established FRET effect, which relies on near-field transmission of a virtual photon from an emitting chromophore to a receiving chromophore, the majority of click-based metal sensors rely on concepts, such as photoinduced electron transfer (PET) or photoinduced charge transfer (PCT).[163] Both sensing mechanisms produce either a "turn-on" or a "turn-off" fluorescence response upon metal binding. The PET mechanism is most strongly represented and has also particular importance for the present study.[182] In general, PET sensors are composed of a fluorophore, which is connected to a metal-binding receptor ligand by a spacer. The electron transfer mechanism between the receptor and the fluorophore requires a separation of charge, which is achieved by photoexcitation of the donor fluorophore. The receptor ligand has a relatively high-energy, non-bonding electron pair, which is able to transfer an electron to the fluorophore and, thereby, quenches its fluorescence. Upon metal complexation, the electron pair is used for metal binding increasing

the redox potential of the receptor, and therefore, lowers its HOMO energy level below that of the fluorophore (Figure 5.6). Consequently, the PET process is inhibited and the fluorescence is turned on.[183] Triazoles have been found to be suitable receptor ligands due to their delocalized π-system and their ability to bind metals via two nitrogen atoms within the heterocycle.[184] Hence, the PET mechanism for triazoles coupled to spacer-separated fluorophores was excessively demonstrated for small molecule metal sensors. Depending on the electronic nature of the coordinating ligand, the choice of the intermediate spacer and the used fluorophore, the PET process can be modulated in terms of fluorescence brightness and sensitivity. It was further stated that the spacer region should not exceed three carbon atoms to assure maximum PET efficiency.[183]
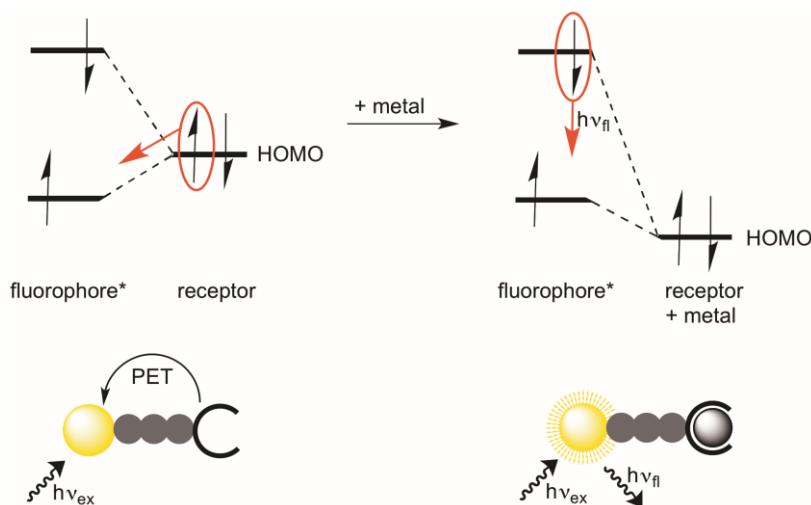
Figure 5.6 Schematic representation of a "turn-on" PET mechanism upon metal complexation.

A different type of fluorescence metal sensors is characterized by a similar structure as described above but without a clear spacer in-between the fluorophore and the receptor unit. Moreover, the fluorophore is either supposed to actively take part in metal coordination by a suitable heteroatom within its delocalized π-system or is directly conjugated to a receptor through an extended intermediate π-system. Thus, an internal charge transfer (ITC) from the receptor-bound metal to the fluorophore is achieved, which often leads to a pronounced hypsochromic or bathochromic shift of the fluorescence emission.[185] These ratiometric sensors are dependent on several factors, such as the electronic nature of the metal ion, the provided coordination sphere and the fluorophore used. Therefore, most of the reported sensors based on this concept are small molecule sensors because they offer better conditions to control the above-mentioned criteria in comparison to their peptidyl analogues. Nevertheless, research efforts concentrate on the development of ratiometric sensors due to the enhanced metal distinguishability because of metal-dependent spectral shifts in contrast to the solely changing fluorescence intensity observed for PET derived sensors.[186]

5.4 Preparation of azido-functionalized fluorophores

Special attention must be paid on the selection of the fluorophores, that should be incorporated into the peptide sequence of the zinc finger. The direct conjugation to the metal-binding site requires fluorophores that are relatively small in size to exclude interferences upon metal coordination and peptide folding. Fluorophores that match these criteria and were therefore used in this study are dansyl, coumarin or anthracene based dyes, which are widely used in peptide and protein labeling and provided good results in the development of other peptidyl chemosensors as described above.[187] Thus, the aforementioned fluorophores were modified with an azide moiety and were subsequently used in the click reaction with the propargylglycine-modified zinc finger.

N-(2-Azidoethyl)dansyl amide (**48**) was synthesized according to a modified preparation reported by INVERARITY and HULME (Scheme 5.1).[188] First, the 2-azidoethanamine linker **46** was synthesized by substituting the bromine moiety of 2-bromoethanamine (**45**) for an azide group. Second, the linker was attached to the commercially available dansyl chloride (**47**) to generate the fluorescent molecule **48**. After purification by silica gel chromatography, the product was obtained in an overall yield of 85%.
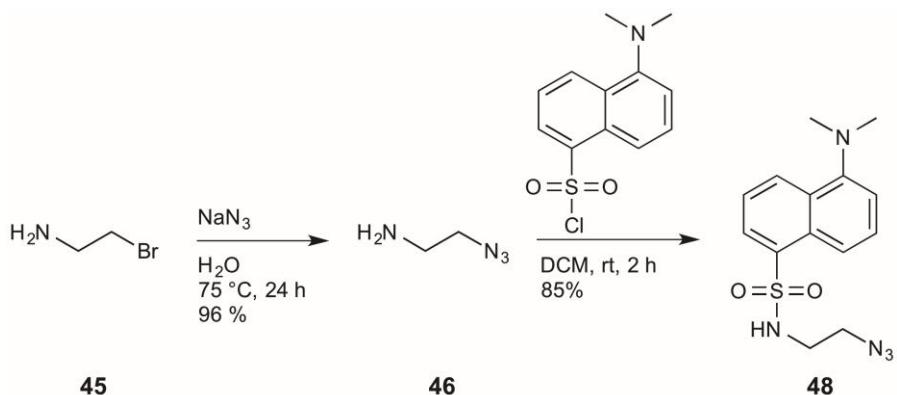
Scheme 5.1 Synthetic route for the preparation of *N*-(2-azidoethyl)dansyl amide (**48**).

The synthesis of 3-azido-7-(diethylamino)coumarin (**53**) was performed using the method of SIVAKUMAR *et al* (Scheme 5.2).[189] In this context, 4-(diethylamino)-2-hydroxybenzaldehyde (**49**) and ethyl 2-nitroacetate (**50**) were used in an aldol condensation to yield 7-diethylamino-3-nitrocoumarin (**51**), which was further reduced to the amine (**52**). Afterwards, the latter was transformed into a diazonium compound using sodium nitrite, which was finally substituted by an azide group using sodium azide.
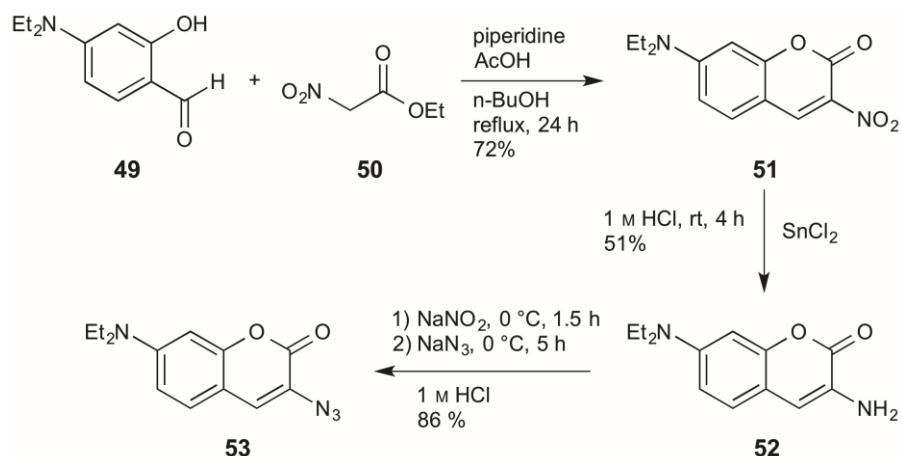
Scheme 5.2 Synthetic route for the preparation of 3-azido-7-(diethylamino)coumarin (**53**).

The third fluorescence sensor 9-azidomethylanthracene (**55**) was synthesized by activating 9-hydroxymethylanthracene (**54**) with tosyl chloride and subsequent replacement for an azide moiety derived from a reaction with sodium azide at room temperature for 24 h (Scheme 5.3).[190]
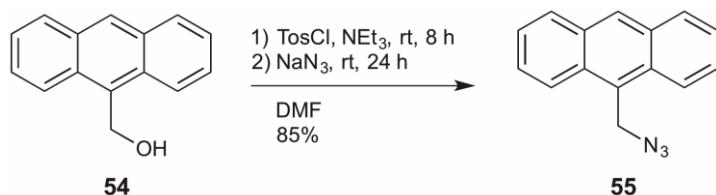
Scheme 5.3 Synthetic route for the preparation of 9-azidomethylanthracene (**55**).

5.4.1 Synthesis of fluorophore modified zinc-finger motifs

As mentioned at the beginning, the zinc finger motif includes two histidine residues that form part of the metal-coordination site and that therefore are suitable targets to be exchanged by triazole units. In this approach, the C-terminal histidine (Figure 5.7) was chosen for that purpose because it is located at the very beginning of the zinc fingers α-helix and offers enough conformational freedom for the fluorophore. According to the already mentioned BERG approach, this position has documented certainty that secondary structure formation does not solely rely on the C-terminal histidine.[177] Even a loose metal coordination by the triazole generated would be sufficient for peptide folding and fluorescence sensing. Moreover, this position prevents the fluorophore to slip into the hydrophobic core of the peptide after metal coordination that could influence the fluorescence response due to environmental changes as reported by the IMPERIALI group.[166]
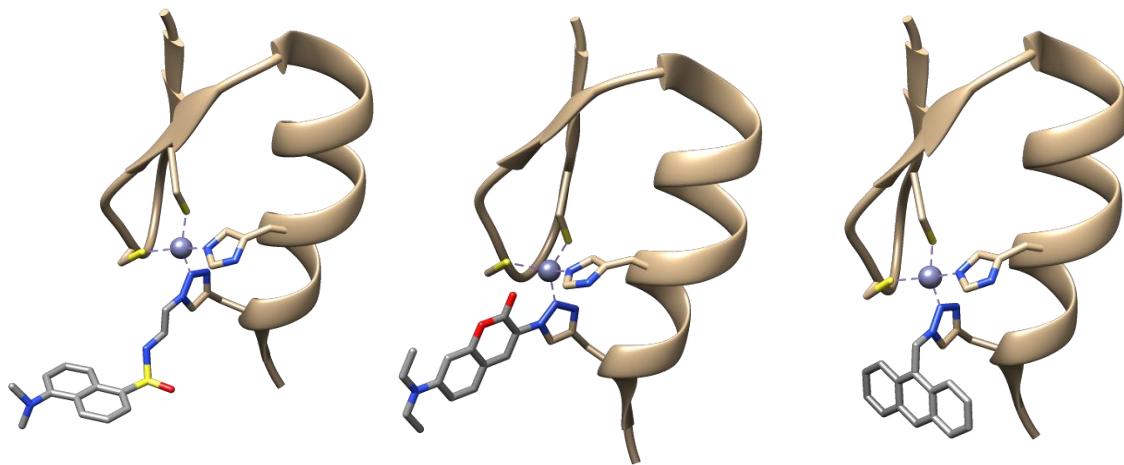
Figure 5.7 Molecular models of the synthesized fluorophore-modified zinc finger constructs Zf3DNS (*left*), Zf3COUM (*middle*) and Zf3ACEN (*right*). Fluorophores and metal coordinating residues are highlighted. The models were generated with UCSF Chimera using the crystal structure of ZIF268 published by PABO (PDB: 1AAY).

The peptide precursor was synthesized by automated solid-phase peptide synthesis (SPPS) using the Fmoc/*tert*-butyl protocol. In place of the C-terminal histidine, Fmoc-propargylglycine was incorporated into the native peptide sequence. After synthesis of the full-length peptide, the N-terminal Fmoc group was not cleaved off by treatment with a piperidine solution (20% in DMF), in order to avoid side reactions of the free amine function in the following click reaction. The copper(I)-catalyzed azide-alkyne cycloadditions (CuAAC) between the propargylglycine residue and the azide-modified fluorophores were performed on solid support by the addition of catalytic amounts of CuI and sodium ascorbate in DMF for 6 h, respectively. Unfortunately, the yield of this reactions were surprisingly low ($<$12%), even though click reactions are known to be fast and nearly quantitative reactions.[191] A closer examination revealed that almost insoluble Cu(I) could not be efficiently separated from the resin by extensive washing with common solvents, such as NMP, DMF, MeOH or DCM. Subsequent cleavage from the resin using TFA solutions and following precipitation of the crude peptide from diethyl ether at –20 °C resulted in peptide precipitations that were nearly insoluble as confirmed by dialysis against zinc rich solutions. It can be assumed that peptide precipitation was due to high concentrations of the Cu(I) species used for the click reaction. Cu(I) is known to either directly coordinate to thiol ligands, such as cysteines, or to slowly oxidize to Cu(II).[192] The latter might be finally responsible for the formation of insoluble peptide aggregates. The hypothesis was confirmed by SCHROEDER *et al.* who observed peptide precipitation as well as structural perturbations of the zinc finger after copper-mediated click reactions.[193] To overcome these problems, cysteine building blocks were used, which were side-chain protected as MeOBzl-ethers (Mob) instead of trityl-ethers and

that are thus stable to TFA-cleavage conditions.[194] In this sense, the metal coordination site of the zinc finger is inoperable after diethyl ether precipitation and the crude peptide was separated from any residual metal ions by reverse phase HPLC. Subsequent treatment of the lyophilized peptide with TFMSA led to the deprotection of the cysteine side chain and precipitation from diethyl ether yielded the fluorophore-modified zinc finger peptides (Scheme 5.4).
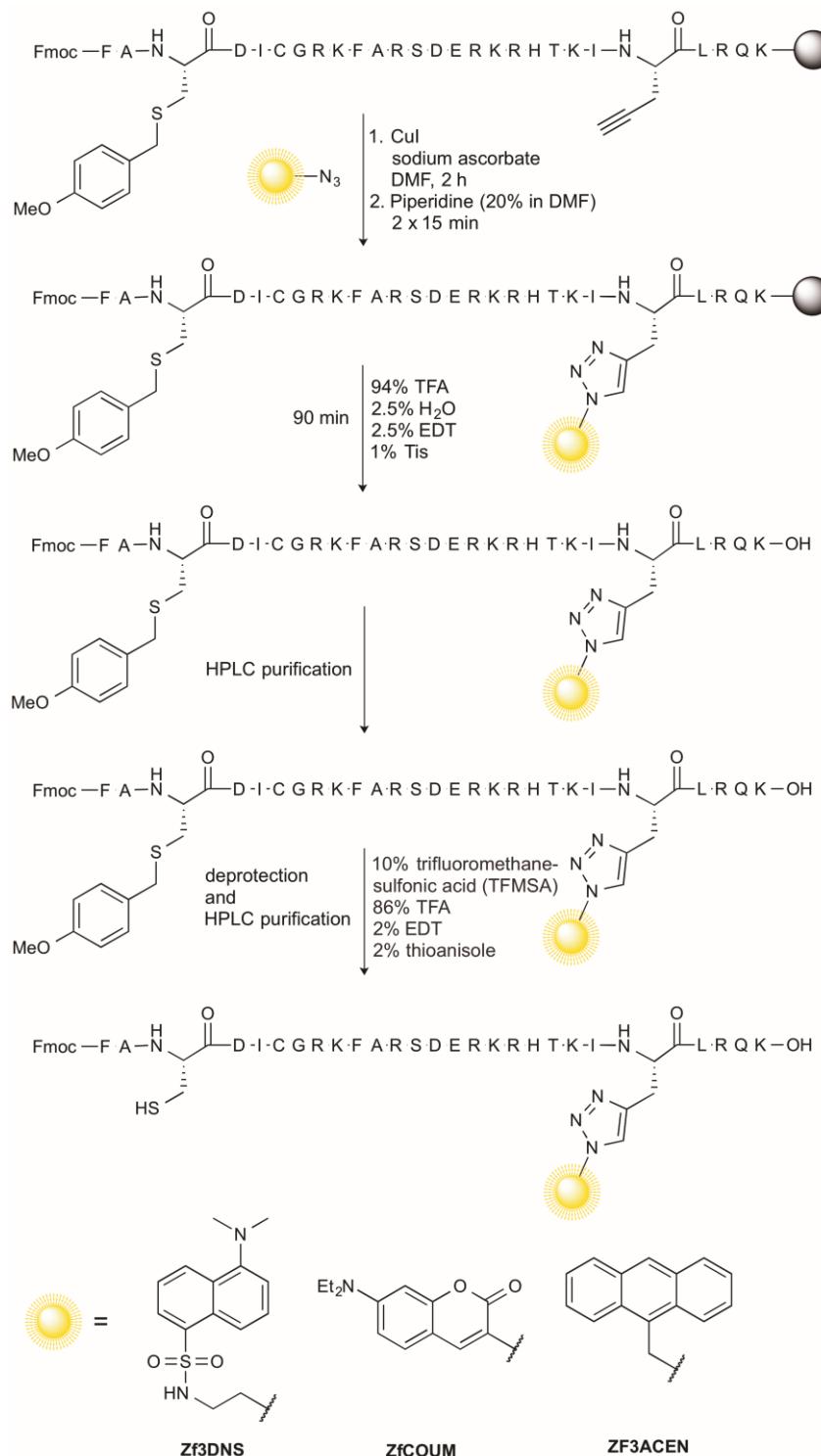
Scheme 5.4 Synthesis route for the preparation of fluorescent labelled zinc finger peptides.

5.4.2 Secondary structure formation of modified zinc finger constructs

In order to evaluate the influence of the incorporated fluorescent labels in close proximity to the metal-binding 1,2,3-triazole moiety, circular dichroism (CD) spectroscopy was applied. While zinc fingers are mainly disordered in absence of metal ions, they obtain their unique $\beta\beta\alpha$-structure upon metal complexation.[83] For this purpose, CD spectra were recorded for the modified zinc-finger motifs (red) and compared with the native peptide (black) in absence (dashed lines) and presence (solid lines) of Zn(II). The experiments were conducted in HEPES buffer (25 mM, pH 7.0), since sulfonate-based buffers are known to have almost no influence on metal complexation or even act as ligands themselves.[72] All peptides (15 μM) were incubated with TCEP (2.5 mM) for 90 min prior to the measurements in order to reduce the disulfide bond between the two cysteine residues and to prevent reformation of the latter. ZnCl$_2$ was used in excess (15 eq with regard to the peptide concentration) to ensure complete occupation of all metal-binding sites. The folding event was monitored by a characteristic bathochromic shift of the negative maximum at 200 nm to 204.5 nm. Furthermore, the shallow shoulder at 222 nm increased significantly due to the increased content of α-helical structure. In comparison to the reference peptide, the dansyl modified zinc finger (Zf3DNS: **58**) showed excellent accordance with the aforementioned characteristics (Figure 5.8). The spectra recorded for both peptides in absence and presence of zinc are almost congruent. It can therefore be considered, that the fluorophore does not negatively influence the secondary structure of the peptide.
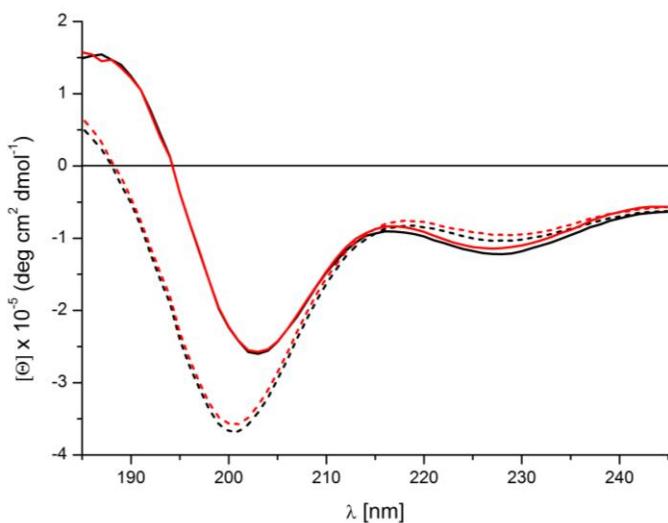
Figure 5.8 CD-spectra recorded for Zf3DNS (red) in absence (dashed lines) and presence of Zn(II) (solid lines) and in comparison to the native zinc finger (black).

In contrast, the spectrum of the coumarin modified peptide (Zf3COUM: **59**) showed a bathochromic shift from 200 nm to 199 nm in absence of Zn(II), which indicated a higher content of random coil structure. This minimum only shifted marginally to higher wavelengths

upon the addition of Zn(II). In addition, the shoulder at 222 nm was considerable less pronounced compared to the reference sample (Figure 5.9).
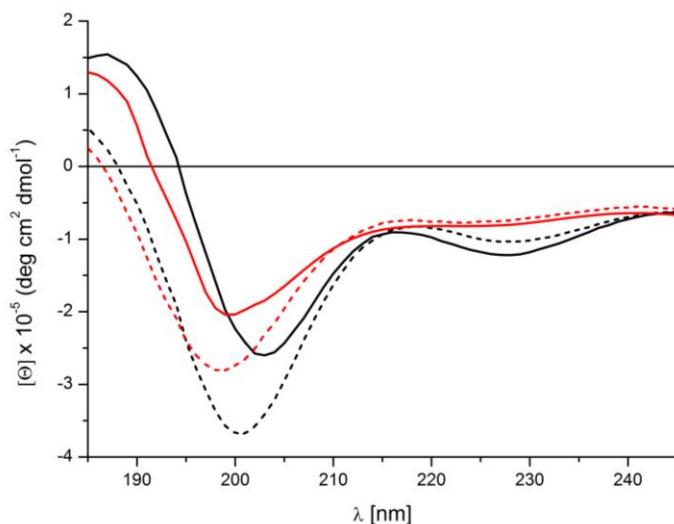
Figure 5.9 CD-spectra recorded for Zf3COUM (red) in absence (dashed lines) and presence of Zn(II) (solid lines) and in comparison to the native zinc finger (black).

An analogous investigation of the anthracene-modified zinc-finger mutant (Zf3ACEN: **60**) revealed a similar result although the spectra were not as distorted as those observed for Zf3COUM. Whereas the spectrum recorded in absence of Zn(II) was quite comparable to the reference, the associated spectrum upon Zn(II) addition failed to show any of the required characteristics, which would indicate accurate peptide folding (Figure 5.10).
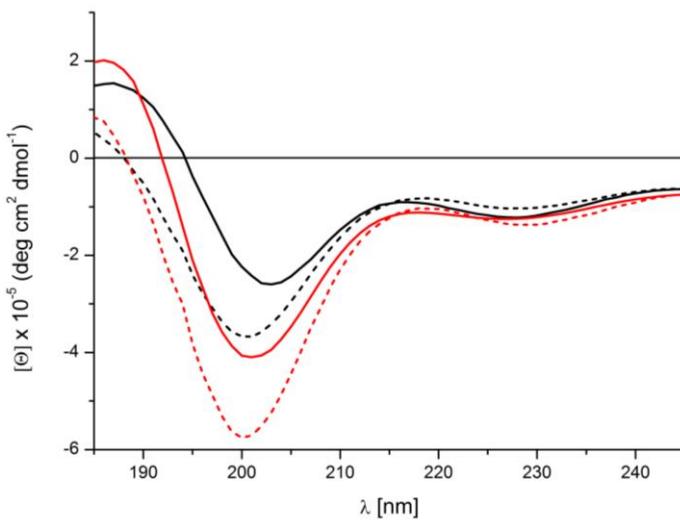
Figure 5.10 CD-spectra recorded for Zf3ACEN (red) in absence (dashed lines) and presence of Zn(II) (solid lines) and in comparison to the native zinc finger (black).

On the basis of the CD-spectra of Zf3COUM and Zf3ACEN, it is reasonable to assume that both peptides show limited capacities for accurate secondary structure formation. On the

contrary, the findings for Zf3DNS were in excellent agreement with those obtained for the native peptide showing similar shifts of the characteristic bands upon Zn(II) addition.

Thus, the question arises how these different observations are connected with each other. A possible explanation could be the use of different spacers, which separated the fluorophores from the metal-coordinating triazole moiety (Figure 5.7). Whereas the dansyl fluorophore was separated by an ethylamine unit, the linker of anthracene was one methylene unit shorter and an inhibited peptide folding was observed. This correlation was even more pronounced for the coumarin derivative, in which the azido group was directly attached to the fluorophore. This situation provided even less distance between the fluorophore and the triazole moiety, whereby the carbonyl oxygen of the benzopyran scaffold might additionally influence metal binding. Hence, the generation of a deviated binding site, which leads to an altered secondary structure, cannot be excluded from CD results. Consequently, both fluorophores, could be too close to the metal-coordinating N2 atom probably causing the inhibited metal complexation for these zinc finger mutants. Thus, the findings clearly demonstrated a vital relationship between the spacer length and the folding ability of the zinc finger.

5.5 Investigation of metal binding by UV-vis spectroscopy

To investigate the metal complexation ability of the engineered zinc fingers as well as the stoichiometry of metal binding, UV-vis spectroscopy was applied. Zn(II) has a closed d10 electron configuration and is therefore referred to as "spectroscopically silent".[195] Hence, it does not show any transition that could be detected by UV-vis spectroscopy. However, zinc competes with other divalent metal ions, such as Co(II) (d7), for the same binding site, even though the competitors show a lower binding affinity.[69] For this reason, it is possible to use cobalt instead of zinc in UV-vis experiments to study the metal to peptide binding ratio.[196] Isostructural Cys_2His_2 complexes with Co(II) (S_2N_2Co coordination) show a pronounced d-d transition in the visible region and a charge transfer in the UV region. To investigate the molar ratio of Co(II) to peptide in the formed metal complex as well as the dissociation constant, Co(II) titration experiments were performed. Therefore, dissociation constants are accessible by titrating distinct amounts of Co(II) to a known concentration of the peptide, while monitoring the increase of the d-d band at 662 nm. The spectroscopic band between 300 – 365 nm, representing the ligand to metal charge transfer (LMCT) from cysteine to cobalt (S-Co[II]), could not be used in this study due to the spectral overlap with the UV-vis signals of the attached fluorophores. The K_d for the corresponding ZF-complex with Zn(II) is subsequently accessible in a reverse Zn(II) titration experiment.[197] Hence, the peptide is saturated with excess amounts of Co(II) and Zn(II) is titrated to the Zf3-Co[II] complex. This is why the disappearance of the d-d band can be observed, indicating the displacement of Co(II) by Zn(II) due to stronger zinc binding. Subsequently, the obtained spectroscopic data

can be plotted as a function of the Zn(II) concentration and fitted to a competitive binding model to determine the K_d for the ZF-Zn^{II} complex.

First, the stoichiometry of Co(II) binding to the modified zinc finger peptides was examined. Prior to the experiments, the disulfide bond between the cysteine residues was reduced by incubation with 2.5 mM TCEP for 90 min. Despite the fact that the reducing agent belongs to the group of non-coordinating reducing agents, measurement errors originating from TCEP/metal interactions should be avoided by purifying the fully-reduced peptides on a Sep-Pak® C18 cartridge (Waters) under non-oxidizing conditions. This included the use of an anaerobic atmosphere containing 5% hydrogen in nitrogen. All solvents were degassed and flushed with this gas mixture in order to prevent reformation of the disulfide bond.

Then, 1.0 μL of a concentrated stock solution of CoCl$_2$ (1.34 mM) in HEPES buffer (20 mM, pH 7.0) was pipetted into a cuvette containing a solution of Zf3DNS (800 μL, 16.8 μM) to obtain a final Co(II) concentration of 1.68 μM. In this way, the Co(II) concentration was increased in steps of 0.1 equivalents with regard to the peptide concentration and the associated UV-vis spectra were recorded between 550 and 720 nm (Figure 5.11a).
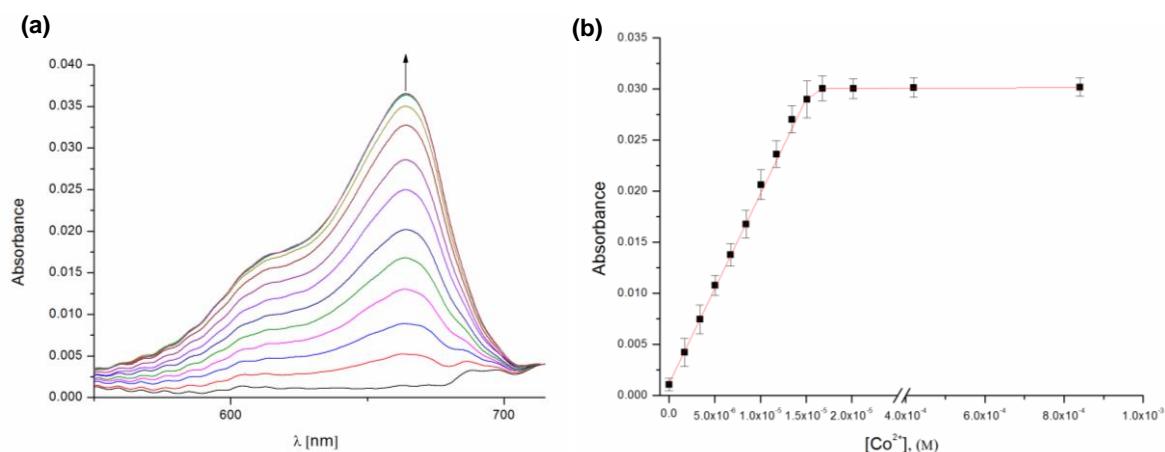
Figure 5.11 (a) Plot of the change in absorption during the titration of Co(II) (1.68 μM) to Zf3DNS (16.8 μM) in HEPES buffer (20 mM, pH 7.0). **(b)** Plot of the absorbance values at 655 nm against their Co(II) concentration.

Complete saturation of the zinc finger was indicated by a consistent absorption signal after the addition of equimolar amounts of Co(II) confirming the formation of a 1:1 complex. The titration was continued by adding excess amounts of Co(II) (1.2, 1.6 and 4.0 eq) to collect enough data points to apply a reliable least squares fit. All measurements were performed in triplicates with freshly prepared solutions. Hence, the obtained maximum absorbance values for each metal concentration were averaged and plotted against their Co(II) concentration what gave a hyperbolic binding curve, which was used to calculate the K_d (Figure 5.11b) according to the method of BERG and MERKLE.[197]

Thus, the dissociation constant is a specific type of equilibrium constant, which measures the propensity of the ZF-Co^{II} complex to loose metal binding and breaking apart. According to Equation 5.1, this process requires the knowledge of the individual concentrations of components that are in equilibrium.[198]

$$K_d = \frac{[M] \cdot [L]}{[ML]}$$

(5.1)

The UV-vis spectra provided information about the relation of the metal ion (M) bound to the ligand (L) by recording the characteristic absorbance (A) at which the formation of the corresponding complex (ML) was indicated. The experimentally obtained absorbance values range between a minimal value (A$_{min}$) under metal-free conditions and a maximal value (A$_{max}$) at saturating levels of the metal species. The equilibrium concentrations in Equation 5.2 were replaced by the total values obtained from the individual titration points. The latter values were used in Equation 5.3 where c$_M$ was the total metal concentration and c$_L$ was the ligand concentration.

$$[ML] = \frac{K_d + c_M + c_L - \sqrt{(K_d + c_M + c_L)^2 - 4 \cdot c_L \cdot c_M}}{2}$$

(5.2)

$$A = A_{min} - \left(\frac{A_{min} - A_{max}}{2 \cdot c_L}\right) \times \left(K_d + c_M + c_L - \sqrt{(K_d + c_M + c_L)^2 - 4 \cdot c_L \cdot c_M}\right)$$

(5.3)

Thus, the dissociation constant K$_d$ of the Zf3DNS-Co^{II} complex was determined in this way to be 7.16 ± 0.4 µM.

In the reverse Zn(II) titration, 1.0 µL of a concentrated stock solution of ZnCl$_2$ (3.0 mM) in HEPES buffer (20 mM, pH 7.0) was pipetted into a cuvette containing a solution of Zf3DNS (800 µL, 15.4 µM) and a 30-fold molar excess of Co(II). By increasing the Zn(II) concentrations stepwise to 0.25, 0.5, 0.75, 1.0, 1.25 and 1.5 equivalents with regard to the peptide concentration, the transition band disappeared and indicated the successful displacement of Co(II) for Zn(II) (Figure 5.12a). A plot of the averaged absorbance data at 655 nm against their Zn(II) concentration provided a hyperbolic binding curve (Figure 5.12b). The determination of the dissociation constant of the Zf3DNS-Zn^{II} complex is based on a competitive binding model between Co(II) and Zn(II) and was conducted according to the method of KREZEL and MILOCH.[199] Consequently, the individual equilibrium equations (Equation 5.4 and 5.5) for every metal (M1 and M2) must be considered and rearranged, according to the mathematical derivation descried in the appendix, into a single equation

(Equation 5.6). This includes the equilibrium concentrations of the metal complexes M1L and M2L as well as of the metals M1 and M2 and the dissociation constants K_d^{M1L} and K_d^{M2L} of the aforementioned metal complexes.

$$K_d^{M1L} = \frac{[M1] \cdot [L]}{[M1L]} \tag{5.4}$$

$$K_d^{M2L} = \frac{[M2] \cdot [L]}{[M2L]} \tag{5.5}$$

$$[M2L] = \frac{[M2] \cdot [M1L] \cdot K_d^{M1L}}{[M1] \cdot K_d^{M2L}} \tag{5.6}$$

Equation 5.6 was transformed into an expression that considered the experimentally obtained absorbance values for the reverse titration in a similar fashion as previously described with the detected absorbance values A within both extremes A_{min} and A_{max} (Equation 5.7). The latter values resulted from the vanishing d-d band upon the displacement of Co(II) (M1) for Zn(II) (M2) in the quantitatively corresponding zinc finger complex (M1L to M2L). The experimental data points for the reverse titration of Zn(II) to the Zf3DNS-Co^{II} complex were processed using the linear Equation 5.7 and provided a K_d of 150.2 ± 14.4 pM.

$$
\begin{aligned}
A = A_{min} &- \left(\frac{A_{min} - A_{max}}{2 \cdot c_L \cdot (K_d^{M1L} - K_d^{M2L})} \right) \\
&\times \left(c_{M2} \cdot K_d^{M1L} + c_L \cdot K_d^{M1L} + c_{M1} \cdot K_d^{M2L} - c_L \cdot K_d^{M2L} \right. \\
&\left. - \sqrt{(c_{M2} \cdot K_d^{M1L} + c_L \cdot K_d^{M1L} + c_{M1} \cdot K_d^{M2L} - c_L \cdot K_d^{M2L} - 4 \cdot c_{M2} \cdot c_L \cdot K_d^{M1L}(K_d^{M1L} - K_d^{M2L})} \right)
\end{aligned}
$$

$$\tag{5.7}$$
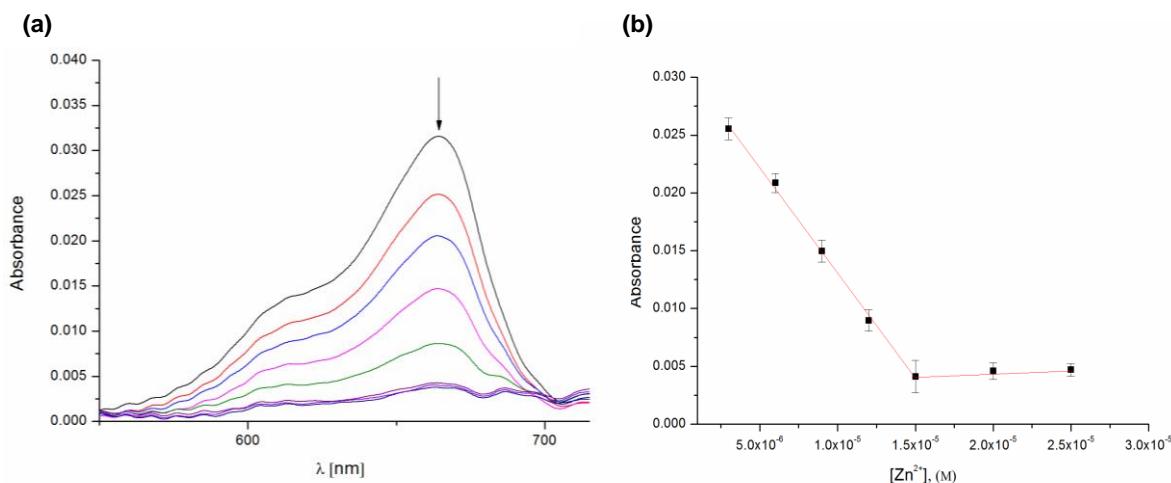
Figure 5.12 (a) Plot of the change in absorption during the reverse titration of Zn(II) (3.75 μM) into a solution containing Zf3DNS (15.4 μM) and Co(II) (500 μM) in HEPES buffer (20 mM, pH 7.0). **(b)** Plot of the absorbance values at 655 nm against their Zn(II) concentration.

The Co(II) titration was also performed for Zf3COUM (15.2 μM) and Zf3ACEN (14.8 μM). However, both peptides showed no significant dependence of the absorbance from the Co(II) concentration due to the absence of the required d-d transition band (Figure 5.13). In this regard, the previous assumption that both peptides have a coordination inhibition caused by the attached fluorophore was further confirmed by these experiments.
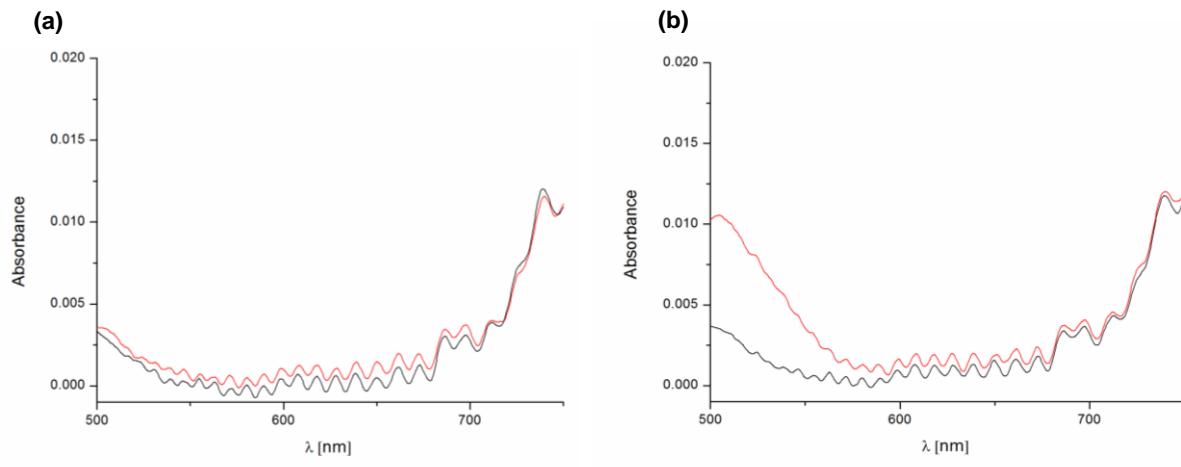
Figure 5.13 UV-vis spectra of **(a)** Zf3COUM (15.2 μM) and **(b)** Zf3ACEN (14.8 μM) before (black line) and after the addition of Co(II) (50 μM).

The calculated dissociation constant for the complexes of Zf3DNS with either Co(II) or Zn(II) are approximately one order of magnitude higher compared to published data for the native sequence.[69] However, they still are in the average range when compared to other zinc finger mutants published. The difference might be a consequence of the introduced 1,2,3-triazol moiety that slightly differs from the exchanged imidazole side chain of histidine. The reduced

electron density at the coordinating N2 atom compared to the native metal-coordinating N(ε) atom might decrease the overall strength of metal complexation of the engineered zinc finger in comparison to the native zinc finger. In addition, the experimental data revealed a slight redshift of the observed absorbance maxima from the d-d transition. In native Cys_2His_2-type zinc fingers, the maximum is predominantly found at 662 nm, whereas the maximum absorbance for Zf3DNS was located at 655 nm.[197] However, the overall shape of the observed curve with regard to the maximum accompanied by a smaller shoulder at shorter wavelengths was in good accordance to literature known ZF-UV spectra and suggested a tetrahedral N_2S_2-coordination. The shifted transition band might also result from the changed electronic structure at the triazole binding site induced by the conjugated fluorophore or due to solvatochromic effects upon peptide folding.[200]

5.6 Fluorescence emission properties of Zf3 metal sensors

The engineered zinc finger sensors were used in fluorescence titration experiments to examine their fluorescence responses upon metal addition as well as to test their ability to distinguish between different divalent metal ions.

Since they are known to be coordinated by the zinc finger with relatively high dissociation constants, the following metal salts were used in this approach: $ZnCl_2$, $CuCl_2$, $CoCl_2$, $NiCl_2$ and $FeCl_2$.[69] In a typical experiment, 1.0 μL of a concentrated stock solution of the metal salt (4 mM) in HEPES buffer (20 mM, 500 μM TCEP, pH 7.0) was added to the sensor peptides (800 μL, 25 μM) to obtain the metal in a final concentration of 5 μM. The metal concentration was increased stepwise to 0.2, 0.4, 0.6 0.8, 1.0, 1.2 and 1.6 equivalents with regard to the peptide concentration. Metal sensing was followed for each concentration by recording the fluorescence emission spectrum by simultaneous excitation of the fluorophore at its most intense excitation wavelength (Table 5.1).

peptide	λ_{ex} [nm]	λ_{em} [nm]	ε [M^{-1} cm^{-1}]
Zf3DNS	334	500 - 550	3.400
Zf3COUM	416	430 - 540	23.500
Zf3ACEN	369	390 - 460	9.700

Table 5.1 Spectral properties of the fluorophores incorporated in the zinc finger peptide by means of CuAAC. λ_{ex} is the excitation wavelength, λ_{em} is the wavelength range of fluorescence emission and ε is the molar extinction coefficient.

Prior to the experiments, the disulfide bond was reduced as described in section 5.5 with the difference that TCEP (500 μM) was also added to the buffer solutions especially for experiments involving the redox active metals Cu(II) and Fe(II). This was necessary because

these metals are known to rapidly induce disulfide bond formation between the two metal coordinating cysteine residues. In order to receive a metal-corrected baseline for each sensor, EDTA (50 μM) was added to the metal-free samples to remove traces of residual ions, which were possibly present in the solution.

As expected and with respect to the afore mentioned CD and UV-vis measurements, the fluorescence spectra of Zf3COUM and Zf3ACEN did not show any significant trend in terms of concentration and metal dependent fluorescence changes for any of the examined metal salts. This was not surprising considering a limited metal binding ability due to disadvantageous interactions between the fluorophore and the metal coordinating 1,2,3-triazol moiety. In contrast to that, Zf3DNS showed a nearly linear behavior of its fluorescence response on the addition of all examined metal ions. Only in the case of Zn(II), a strong increase in fluorescence was observed complying to a switch-on mechanism (Figure 5.14). In addition, the increasing emission band was accompanied by a slight blue shift from 539 nm to 527 nm. At nearly equimolar concentrations of Zf3DNS and Zn(II), the emission band reaches a plateau that remains even after the addition of excess Zn(II). On saturating levels, the peptide showed a total increase in fluorescence intensity by 40% with regard to its initial fluorescence. The addition of EDTA (to a final concentration of 500 μM) into the ZF-Zn^{II} solution completely restored the initial fluorescence. Interestingly, the excess use of zinc salts (>50 eq) led to a precipitation of the peptide. This resulted in the formation of fine fibrils in the cuvette, which arbitrarily increased the fluorescence signal. This fact might be indicative for the formation of additional intermolecular metal-binding-sites and must be carefully considered when utilizing the zinc finger in a sample containing high metal concentrations.
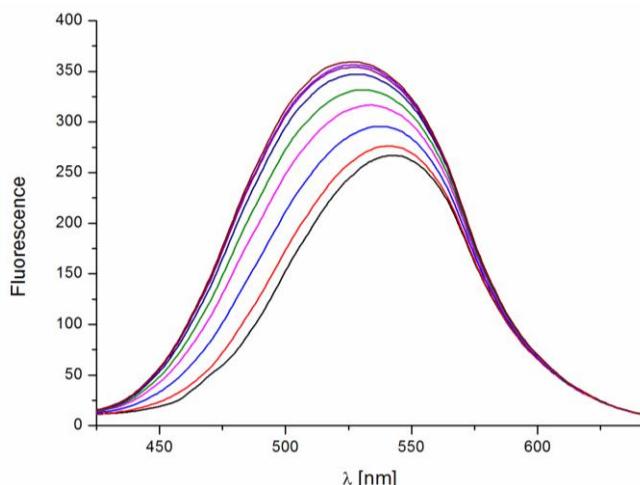
Figure 5.14 Fluorescence spectra obtained for the titration of Zn(II) in steps of 0.2 equivalents into a solution of Zf3DNS (25 μM) in HEPES buffer (20 mM, 500 μM TCEP, pH 7.0).

The fluorescence spectra recorded for the other divalent metal ions used in this study showed a different behavior in terms of fluorescence response upon metal titration. With increasing concentrations of these metals, a slight decrease in fluorescence intensity of the band at 539 nm was observed. The decrease was accompanied by either a slight red shift (Cu(II) and Fe(II)) or by a slight blue shift (Ni(II)). In case of Co(II), the fluorescence bands remained at 539 nm (Figure 5.15). All fluorescence titrations showed a nearly linear decrease of the dansyl emission band after the addition of distinct concentrations of the metal ions ending in a plateau with consistent fluorescence.
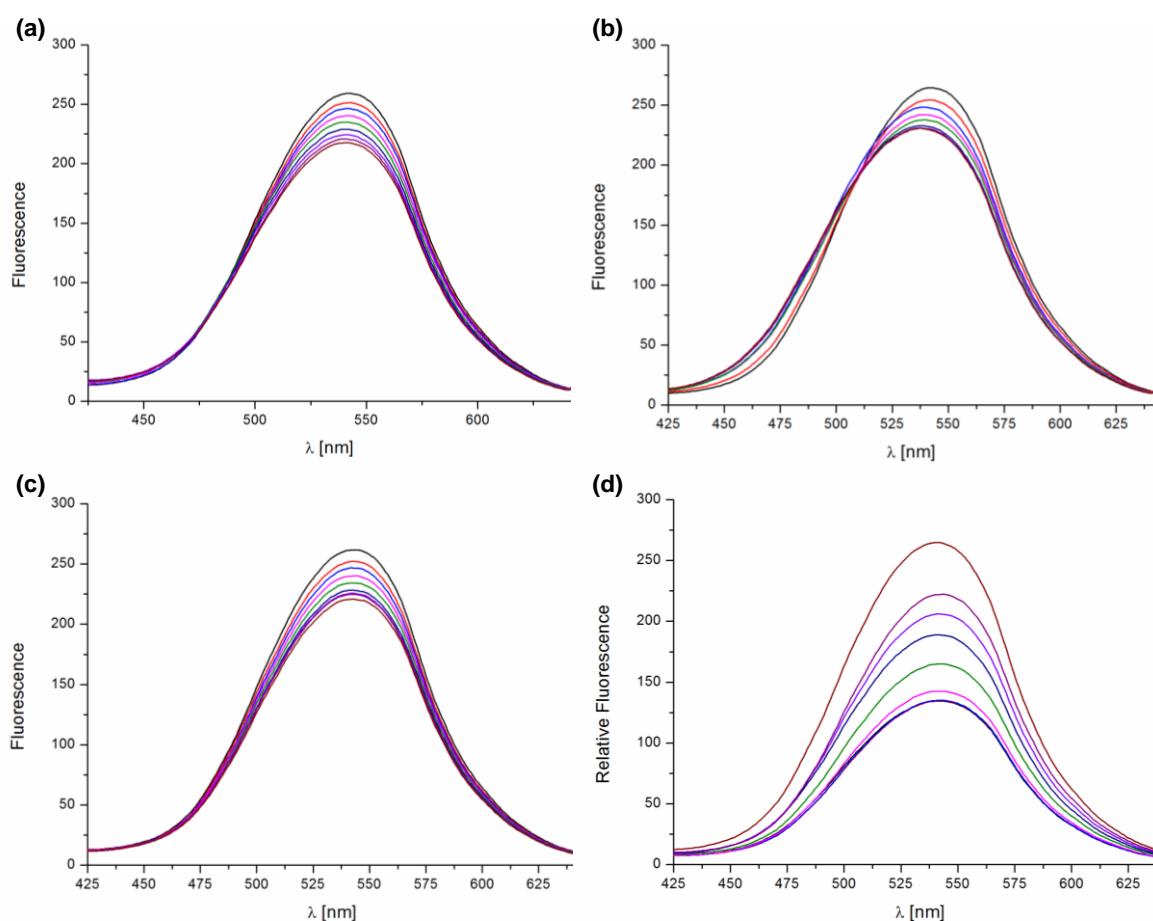
Figure 5.15 Fluorescence spectra of Zf3DNS (25 μM) for the titration with **(a)** Co(II), **(b)** Ni(II), **(c)** Fe(II) and **(d)** Cu(II). The metals were added in steps of 0.2 equivalents with respect to the peptide concentration.

Cu(II) showed the highest decrease in fluorescence intensity among all examined metals with only 55% of the initial dansyl fluorescence. In contrast, the fluorescence intensities of Co(II), Ni(II) and Fe(II) were only slightly decreased within a range of 8% with regard to their metal-free initial fluorescence. The addition of an EDTA solution (to a final concentration of 500 μM) to the saturated ZF-metal complexes completely recovered the initial fluorescence. For other fluorescent chemo sensors found in literature, a similar behavior for Cu(II) was observed.[201] It was stated that the redox-active Cu(II) ion induces a strong electron-

withdrawing effect on the coordinating ligands, which might change the electronic structure of the conjugated fluorophore. In this sense, Cu(II) has 9 d-electrons (d^9 configuration) and it therefore has a high ambition to obtain a closed-shell d^{10} configuration in comparison to the less fluorescence attenuating ions Ni(II) (d^8) and Co(II) (d^7).

5.7 Summary and conclusions

The described zinc-finger-based metal sensors derived from click-chemistry successfully demonstrated the use of CuAAC to introduce a fluorophore to a metal-coordination site of a peptide. The *in-situ* generated 1,2,3-triazole moiety was able to mimic the coordination properties of histidine and thereby maintained all characteristic structural features as confirmed by CD spectroscopy. The same method revealed the importance of the spacer-length that separates the fluorophore from the metal coordination site. With decreasing length, the metal binding ability of the peptide was inhibited and led to an inaccurate secondary structure formation as observed for Zf3COUM and Zf3ACEN. It can be assumed that the close proximity of the fluorophore to the metal coordinating N2 atom of the triazole moiety negatively influenced the metal coordination. The influence might either derive from shielding effects of the fluorophore on the metal binding site or due to additional functional groups of the fluorophores, which are also able to perform metal coordination such as the carboxyl oxygen of coumarin. This could have established a competing binding mode, which led to inaccurate peptide folding. To evaluate these assumptions, future experiments could target the precise impact of the spacer length. It remains to be emphasized that a spacer of similar length as the one used for Zf3DNS might also be beneficial for other fluorophores. The hypothesis of a competing binding site needs to be evaluated for example by nuclear magnetic resonance spectroscopy to exactly prove if metal binding proceeds via the N2 or N3 nitrogen atom of the triazole. Such experiments were performed for the present systems but due to the enormous abundance of signals it was not possible to draw solid conclusions. The use of a reduced model system could remedy this deficiency.

The aforementioned coordination mismatches for the metal complexes of Zf3COUM and Zf3ACEN were further confirmed by UV-vis titration experiments, which were performed to investigate the molar ratio of metal to peptide binding and to determine the dissociation constants for the zinc finger complexes with Co(II) and Zn(II). Whereas the spectra of Zf3COUM and Zf3ACEN showed no d-d transition upon Co(II) addition, the latter was obtained for Zf3DNS and used to determine the K_d for the ZF-Co^{II} complex to be 7.16 ± 0.4 μM. In a reverse titration with Zn(II), the dissociation constant for the ZF-Zn^{II} complex was determined to be 150.2 ± 14.4 pM. These values are in overall good agreement with published dissociation constants for other zinc finger mutants found in the literature.[69]

Fluorescence titration experiments for Zf3DNS with different divalent metal ions were performed to test its ability to function as fluorescent metal sensor. The findings clearly demonstrated that the peptide shows an increase in fluorescence intensity exclusively upon Zn(II) addition. This switch-on mechanism resulted in a 40% increased fluorescence at saturating Zn(II) levels with regard to the initial metal-free value. In contrast to that, all other examined divalent metal ions showed moderately decreased fluorescence intensities within an 8% range with regard to metal-free conditions (Figure 5.16).
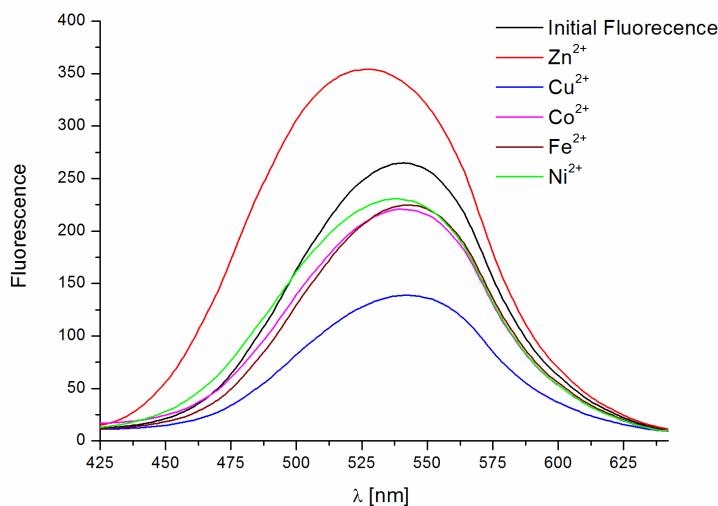
Figure 5.16 Fluorescence spectra of the dansyl modified zinc finger (Zf3DNS) saturated with different divalent metal ions (~1.6 eq), as indicated. The peptide was used in a concentration of 25 μM in HEPES buffer (20 mM, 500 μM TCEP, pH 7.0). The initial fluorescence was determined by the addition of EDTA (50 μM) to complex residual metals.

These findings imply that the metal sensing ability of Zf3DNS relies on the PET mechanism described in section 5.3. This assumption is based on two facts. Firstly, the presence of a spacer that connects the metal coordination site and the fluorophore is a major indication for a PET mechanism. This condition was met with the ethylamine moiety that connects the dansyl group and the triazole. The absence of delocalized electrons within the spacer makes an ICT mechanism rather unlikely due to an insufficient signal transmission from the metal binding site to the fluorophore. Secondly, the metal addition led to a change in fluorescence intensity under retention of the absorbance wavelength. The still observed hypsochromic or bathochromic shifts may be attributed to peptide folding and therefore, to local changes in the environment of the fluorophore upon changing from the unfolded state to the folded state where the latter is partially shielded. An ICT mechanism as mentioned in the beginning can be virtually excluded due to the observation of minor spectral shifts within 7 nm. Due to the findings for multiple other ICT-based metal sensors, the expected spectral shift should be significantly more pronounced.[185] In addition, the turn-on fluorescence sensing mechanism exclusively observed for Zn(II) is in very good agreement with other PET based metal

sensors.[202] The fluorescence decrease observed for the other examined metals is rather due to the aforementioned folding event than due to a simultaneously occurring turn-off effect. Despite the fact, that only a few examples are known, which combine a turn-on and turn-off mechanism, the energy diagrams shown in Figure 5.6 clearly prohibit such a mechanism for solely PET based metal sensors.[203] Hence, metal sensors that comprise a simultaneous on-off mechanism are usually composed of two different binding sites that also show different metal selectivity and fluorescence response upon complex formation.

The resolution of the Zf3DNS sensor to detect Zn(II) is rather limited to the relatively low brightness of the fluorophore than to the metal binding ability of the peptide, which was found to be in the upper nanomolar range. Compared to other fluorophores, the azido-*N*-dansylethylamine fluorophore has a comparably low molar extinction coefficient of 3400 M^{-1} cm^{-1}.[204] The abolition of the PET quenching upon Zn(II) addition increased its fluorescence intensity by 40%. In contrast to other fluorescence metal sensors this is a rather small increase, which might be attributed to the relatively long spacer. Thus, the efficiency of the PET could be reduced as indicated at the beginning. These characteristics leave room for improvements concerning the attachment of brighter fluorophores with variable spacer-lengths. However, the reported zinc-finger-based peptidyl metal sensor was found to be fully operable and able to distinguish between Zn(II) and other divalent metal ions the zinc finger is able to bind.

The successful synthesis of a zinc-finger-based peptidyl metal sensor suggests that histidines, which are part of a metal binding site in peptides, proteins or other molecules can be exchanged for a propargylglycine residue to enable metal coordination by a click-derived triazole moiety. The applied CuAAC allows the direct modification of the metal binding site with conjugated fluorophores in a mild, fast and orthogonal reaction. This could lead to new insights in processes, which involve metal coordination in biomolecules due to the possibility of metal sensing in an aqueous environment. A potential field of application is for example the pathogenesis of Alzheimer's disease. In current studies, the involvement of histidine residues in Zn(II) and Cu(II) complexation is under investigation because this process is expected to cause the formation of highly neurotoxic amyloid-β fibrils and plaques.[205] An exchange of the eligible histidines for the herein described system, could help to spread further light on the progression of this disease. The modification of triazoles with other functional molecules instead of fluorophores was demonstrated by RAO. In her approach, a second metal binding site derived from a 1,4,7-triazacyclononane (TACN) modified triazole was *in-situ* generated in the zinc finger sequence in close proximity to the native metal binding site.[206] The thus generated dinuclear Zn(II) complex was intended to be used as sequence specific artificial nuclease. It is important to note, that even with such a complex

system, the secondary structure of the zinc finger mutant was successfully established as confirmed by CD spectroscopy.

In summary, tailor-made post synthetic modifications of metal binding peptides or proteins, either synthesized by the means of solid phase peptide synthesis or recombinantly expressed, are easily accessible by CuAAC. 1,2,3-Triazoles are excellent surrogates for histidines, which can easily be generated from propargylglycine that is furthermore genetically encoded for the use in direct ribosomal synthesis.[207] In combination with the highly orthogonal click reaction, this would even allow the modification of larger proteins.

6. Experimental section

6.1 Materials and methods: organic synthesis

Solvents

All solvents used for synthesis were purchased in the highest quality available (puriss. p.a., absolute). Technical grade solvents were distilled prior to use and were stored over molecular sieves. Analytic and HPLC grade solvents were supplied from *Acros Organics* (Geel, Belgium), *Sigma-Aldrich* (Taufkirchen, Germany) and *Fisher Scientific* (Nidderau, Germany). Demineralized water was further purified on the water purification system *Simplicity* from *Merck Millipore* (Billerica, USA) prior to use.

Reagents

All reagents were purchased in the highest grade available and were used without further purification. Amino acids, coupling reagents and resins used in SPPS were purchased from *NovaBiochem* (Darmstadt, Germany), *GL Biochem* (Shanghai, China), *ABCR* (Karlsruhe, Germany), *Bachem* (Bubendorf, Switzerland), and *IRIS Biotech* (Marktredwitz, Germany). Other reagents were purchased from *Merck* (Darmstadt, Germany), *Carl Roth GmbH* (Karlsruhe, Germany), *Fisher Scientific GmbH* (Nidderau, Germany), *Alfa Aesar* (Karlsruhe, Germany), *Bachem* (Bubendorf, Switzerland), *TCI* (Eschborn, Germany), *VWR* (Darmstadt, Germany), *NovaBiochem* (Darmstadt, Germany), *Sigma Aldrich* (Taufkirchen, Germany) and *Acros Organics* (Geel, Belgium). DNA oligomers were purchased in a HPLC-purified state from *Biomers* (Ulm, Germany). Biochemical reagents were purchased from *Fermentas* (St. Leon-Rot, Germany), *Biorad* (Munich, Germany) and *GERBU* (Heidelberg, Germany).

Reactions

Air and moisture sensitive reactions were performed under an argon or nitrogen atmosphere. Laboratory glassware was flame dried under reduced pressure and purged with dried nitrogen or argon. Solvents used in air sensitive reactions were degassed and flushed with argon prior to use. Reactants and solvents were added through a septum with the help of a syringe equipped with a cannula.

Lyophilization

Products were lyophilized from an aqueous solution, which was allowed to contain minimal amounts of acetonitrile, methanol or dioxane on a *Christ Alpha-2-4* attached to a high vacuum pump. Freeze drying of larger samples was performed in a single-neck round bottom flasks. Smaller samples (<2 mL) were lyophilized in an *Eppendorf* safe-lock microcentrifuge tube in an evacuable *Christ RCV-2-18* centrifuge, which was connected to the lyophilizer. All solvents were completely frozen in liquid nitrogen prior to lyophilization.

Thin layer chromatography (TLC)

Aluminium-backed plates coated with *silica gel 60 F_{254}* (layer thickness: 0.20 mm) from *Merck* were used for TLC. Substances were visualized by dipping the plate into a ninhydrin-staining solution (1.5 g ninhydrin, 3 mL acetic acid, 100 mL n-butanol) followed by spot development in a stream of hot air.

Flash column chromatography

Flush column chromatography was performed by filling a glass column with an appropriate diameter and equipped with a frit with *Merck Silica Gel 60* (particle size: 40 – 62 μm). The silica gel was suspended in the elution system under the exclusion of air pockets. The compounds were dissolved in the smallest possible amount of eluent. The solution was carefully added to the top layer of the silica gel. The elution was performed by applying a pressure of 0.1 – 1 bar and fractions were collected in glass test tubes.

High performance liquid chromatography (HPLC)

Reverse phase HPLC (RP-HPLC) was performed on an *Amersham Pharmacia Biotech* system (Äkta basic, Pump-type P-900, UV detector 900). Time-dependent chromatograms were recorded by the detection of the compounds by means of UV absorption at 215 nm, 254 nm and 280 nm.

The following solvent systems were used:

A [99.9% H_2O, 0.1%] to B [79.9% MeCN, 20% H_2O, 0.1% TFA],
A [99.9% H_2O, 0.1%] to C [99.9% MeOH, 0.1% TFA]

The samples were dissolved in MilliQ water and either MeCN or MeOH was added in order to increase their solubility. Prior to injection, the samples were filtered using CHROMAFIL® RC-45/15 MS (*Macherey-Nagel*) disposable syringe filters. The flow rates for analytical,

semi-preparative and preparative RP-HPLC runs were 1 mL/min, 3 mL/min and 10 mL/min. The used columns were as indicated in the analytical section of the particular substances.

Nuclear magnetic resonance spectroscopy (NMR)

NMR spectroscopy was performed on a *Varian* instrument (Mercury 300, Unity 300, INOVA 500, INOVA 600). Chemical shifts are indicated in parts per million (ppm) downfield of the internal standard ($\delta_{TMS} = 0$ ppm). $CDCl_3$, [D6]-DMSO, CD_3OD and D_2O were used as deuterated solvents. Signal multiplicities were abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. Coupling constants $^n J_{XX}$ are stated in hertz (Hz), where n is the order of coupling and X stands for the nuclei.

Mass spectrometry

The compounds were characterized by electron spray ionization mass spectrometry (ESI-MS) on a *Finnigan instruments* (LCQ or TSQ 7000) or *Bruker* spectrometer (Apex-Q IV 7T and micrOTOF API). High resolution mass spectra were recorded on a *Bruker* Apex-Q IV 7T or on a *Bruker* micrOTOF.

UV spectroscopy

UV-vis spectra were recorded either on a *Nanodrop* ND-2000c (d = 0.1 cm) or on a *JASCO* V-550 UV-vis spectrophotometer (d = 1 cm). *Hellma* absorption cells with a path-length of 10 mm and a chamber volume of 1300 μL or 700 μL were used. Prior to the measurements, a blank spectrum was recorded, which was subtracted from the measured spectra. In cases UV-vis spectroscopy was used to determine the concentration of a compound, Lambert-Beer law was used for the calculations.

CD spectroscopy

Circular dichroism spectroscopy was either performed on a *Jasco-810* or on a *Jasco-1500* spectropolarimeter, which were equipped with a *Jasco-PTC432S* or a *Julabo-F250* temperature controller. *Hellma* quartz precision cells with a thickness of either 1.0 mm or 10 mm were used for the measurements. The spectra were recorded in the range of 190–280 nm at a constant temperature of 25 °C. The following parameters were applied for secondary structure measurements with Zf3 peptides: band width: 0.1 nm, response time: 0.1 s, sensitivity: high, data pitch: 0.1 nm and scanning speed: 50 nm/min. The final spectra were background corrected and were averaged from 12 accumulations. The means movement method was used to smooth the obtained curves.

Fluorescence spectroscopy

Fluorescence measurements were performed on a *Jasco FP-6500* spectrofluorometer at 25 °C. The excitation and emission bandwidth was set to 5 nm, the data pitch was 1 nm, the response time was 0.1 s and the sensitivity was high. *Hellma* fluorescence cuvettes (semi micro Suprasil® quartz, limit 200–2.500 nm spectral range, path-length 10 x 4 mm, chamber volume 1400 µL) were used for measurements. All spectra were background subtracted and averaged from five measurements.

Microscale thermophoresis

MST measurements were performed on a Monolith NT.115 from *NanoTemper Technologies* (Munich, Germany). Fluorescence detection of the FITC labeled compounds was performed using the blue channel of the device. The LED power was set between 50–95% and the MST-power was set between 20–80%. The laser-on time was 30 s and the laser-off time was 5 s. The labeled small molecules were used in a final concentration of 60 nM whereas the labeled peptides were used in a final concentration of 120 nM, respectively. Ligand titrations were performed by preparing 1:1 dilution series starting with the highest concentration of the titrant and subsequent dilution with the buffer solution. Prior to capillary loading, all samples were centrifuged at 9.000 rpm for 10 min. All experiments were performed with *premium capillaries,* which were immediately sealed after loading.

6.2 Materials and methods: biochemistry

General operations

Sterilized equipment and laboratory glassware was used in all biochemical experiments described here. Single-use items were collected and autoclaved after usage and prior to disposal. Reusable materials were autoclaved after usage and subsequently sealed with sterile aluminum foil. All experiments, which involved the work with genetically modified organisms, were performed under a laminar flow biological safety cabinet Bio-II-A from *Prettl-Telstar*. Aqueous media and solutions were prepared with sterilized MilliQ water.

Materials

Luria-Bertani (LB)-media: Tryptone (1%, w/v), yeast (0.5%, w/v) and NaCl (1%, w/v) were dissolved in MilliQ water and autoclaved at 120 °C for 25 min.

LB plates: Tryptone (1%, w/v), yeast (0.5%, w/v), NaCl (1%, w/v) and agar (7.5%, w/v) were dissolved in MilliQ water and autoclaved at 120 °C for 25 min. After cooling to 50 °C, carbenicillin was added to the LB-media, which was subsequently casted into sterile Petri dishes.

Antibiotics: Carbenicillin (100 µg/mL) was used as additive in the prepared media to suppress undesired bacterial growth except for the resistant *E. coli* host strain ER2566.

Bacterial host strain: The commercially available *E. coli* host strain ER2566 (*New England Biolabs*) was used for protein expression, which was made competent prior to use.

Preparation of competent *E. coli* cells

The preparation of competent *E. coli* cells facilitates the uptake of plasmid DNA through the bacterial cell membrane. Thus, a single colony was picked from a LB plate and transferred into 5 mL of LB medium. The cultivation was performed at 37 °C overnight before 2 mL of the medium was inoculated into an Erlenmeyer flask, which contained another 50 mL of LB medium. The medium was cultivated at 37 °C until an OD_{600} of approximately 0.3 was obtained. The cell culture was immediately cooled on ice and was subsequently centrifuged at 3.000 rpm and 4 °C for 15 min. The supernatant liquid was removed and an aqueous solution of $CaCl_2$ (0.1 M, 10 mL, 5 °C) was added to the cell pellet. The cell suspension was incubated on ice for 20 min followed by centrifugation at 3.000 rpm and 4 °C for 15 min. After the removal of the liquid components, an aqueous solution of $CaCl_2$ (0.1 M, 5 mL), which also contained glycerol (15%) was added to the cell pallet. The solution was divided into sterile

Eppendorf safe-lock tubes (500 μL each) that were frozen in liquid nitrogen and stored at –80 °C.

Vectors: The commercially available (*New England Biolabs*) plasmid vector pTXB1 was used for the recombinant protein expression of the Zf12 domain. The vector also included a gene for carbenicillin resistance.

Transformation of competent *E. coli* cells

A 500 μL glycerol stock of competent *E. coli* cells was slowly thawed on ice. Subsequently, 5 μL of the plasmid DNA was added to 100 μL of the *E. Coli* stock and the sample was incubated on ice for 30 min. The sample was heat-shocked at 42 °C for 45 sec and placed on ice again for 5 min. LB medium (900 μL) was added and the sample was incubated at 37 °C for 1 h. The solution was centrifuged at 3.000 rpm for 2 min. 100 μL of the supernatant was placed on a LB plate, which was subsequently incubated at 37 °C overnight.

Recombinant protein expression

A single colony was picked and transferred to LB medium (5 mL), which also contained 5 μL of carbenicillin. The medium was cultivated at 37 °C for 24 h. Afterwards, the medium was used to inoculate the expression culture that was cultivated in 400 mL of LB medium containing an appropriate amount of carbenicillin. The expression culture was incubated at 37 °C for 4 h before 40 μL of IPTG (1 mM) was added and the incubation was continued at 16 °C for 24 h. The suspension was centrifuged at 4.000 rpm and 4 °C for 20 min. The supernatant was removed and the cell pallet was transferred into a centrifuge tube before the cell lysis reagent BPer (1 g/mL) was added. The suspension was placed on ice and was subject to pulsed sonification (5 x 45 sec, 50% power, 50% impulse) using a Sonoplus GM 7 (*Bandelin*). The cell material was separated by centrifugation (14.000 rpm, 4 °C, 20 min) and the supernatant that contained all expressed proteins was purified by affinity chromatography.

Purification of Zf12

The cell lysate was purified by affinity chromatography using Chitin beads (*New England Biolabs*). Hence, the expressed Zf12 domain contained a fused chitin binding domain (CBD), which showed high affinity for the corresponding chitin beads. To do so, the chitin beads (100 mL) were loaded to a XK26/20 column (*GE Healthcare*) and were equilibrated with two bed volumes of water and ten bed volumes of column buffer (20 mM HEPES, 500 mM NaCl, 0.1% Tween20, 20 mM TCEP, pH 8) using an Äkta Prime Plus system (*GE Healthcare*). The cell lysate was repeatedly loaded onto the column at a flow rate of 0.5 mL/min overnight,

whereby only the Zf12 protein stuck to the chitin beads. All other proteins were washed from the column using 20 bed volumes of column buffer at a flow rate of 2 mL/min. The Zf12 thioester protein was liberated from the chitin beads by adding 1.5 bed volumes of cleavage buffer (20 mM HEPES, 500 mM NaCl, 0.1% Tween20, 20 mM TCEP, 250 mM MeSNa, pH 8) and incubation overnight. The desired protein was eluted by adding column buffer at a flow rate of 2 mL/min until the detected UV absorbance as well as the electric conductivity reached a minimum. The aqueous solution was immediately lyophilized and the solid components were solubilized again in the minimum amount of water required to obtain a homogeneous solution. Subsequently, the solution was subject to RP-HPLC purification, the peptide was lyophilized and stored at –20 °C.

Oligonucleotide annealing protocol

The DNA oligomers were purchased as a freeze-dried solid, and were therefore, dissolved in 900 μL MilliQ water. The absorption values were determined by means of UV-vis spectroscopy and the concentrations were calculated using Lambert-Beers law with the molar absorption coefficients provided by the DNA supplier. Distinct concentrations of complementary DNA strands were pipetted into an *Eppendorf* safe-lock microcentrifuge tube, which contained the annealing buffer (10 mM Tris, 50 mM NaCl, pH 7.8). The tube was sealed and heated at 95 °C for 5 min and subsequently cooled to room temperature. Prior to use, the samples were centrifuged at 6.000 rpm using a *Roth* microcentrifuge.

Gel electrophoresis

Polyacrylamide gel electrophoresis (PAGE) was used to examine the binding ability of the engineered zinc-fingers for their DNA target sequence as well as to test for their DNA hydrolysis capacity. Gel electrophoresis is used to separate different macromolecules according to their electrophoretic mobility, which is dependent on the length, conformation and intrinsic charge of the analytes.

Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gels are cross-linked porous materials having small pore sizes, which are suitable for the separation of polynucleotides with a length between 5–500 bp. The gel matrix was prepared by co-polymerization of acrylamide with N,N′-methylenebisacrylamide in the presence of APS and TEMED. The Zf13 peptides as well as the DNA samples were pipetted into the wells of the polyacrylamide gel and electrophoresis was performed with a suitable running buffer and by applying an appropriate voltage. The electrophoretic migration of the macromolecules was evaluated by the detection of the fluorescent 6-FAM label that was attached to the 5′-end of a single DNA strand. This was conducted on a *Typhoon 9400*

Variable Mode Gel Imager from *Amersham Biosciences (GE Healthcare)* by scanning the gel with the blue laser channel and using *ImageQuant* software for visualization.

Native PAGE

Native PAGE is performed under non-denaturing conditions that allow the detection of higher-order structures, such as peptide/DNA complexes after electrophoresis. This method was used to evaluate the binding ability of the engineered zinc-fingers towards their consensus dsDNA sequence.

Native polyacrylamide gels were prepared according to the following procedure using the Mini-PROTEAN® Tetra Handcast Systems (*Bio-Rad*). The glass plates and the spacer plates were washed with double-distilled water followed by an aqueous solution of EtOH (70%) and were allowed to dry in air. The plates were assembled to generate a gel thickness of 0.75 mm. The cassette was placed on a gasket and filled with a freshly prepared acrylamide gel solution until the solution was 1 cm below the teeth of the comb. A layer of distilled water was rapidly but carefully added on top of the acrylamide gel and the gel was allowed to polymerize. Subsequently, the water was removed using filter paper. The remaining space in the cassette was filled completely with stacking gel and a comb was immediately placed in the cassette. After complete polymerization, the cassette was transferred to the electrophoresis apparatus, which was filled with running buffer. After the addition of 2 μL glycerol (30%) to each sample, the comb was removed and the samples were pipetted into the wells of the gel. Electrophoresis was performed by applying a constant voltage of 170 V.

All buffers and solvents were prepared under the exclusion of denaturing or metal-chelating substances, such as EDTA or SDS, to maintain native conditions.

10x TB Buffer (1 L): 108 g Tris
                   55 g Boric acid
                   adjusted to pH 8.0

Gel Solution (50 mL): 15 mL/25 mL Rotiphorese Gel 40 for 12%/20% acrylamide gel
                   5 mL of 10x TB buffer
                   270 μL of aqueous APS (10%)
                   27 μL of TEMED
                   Final volume adjusted to 50 mL with *MilliQ* water

SDS-denaturing PAGE

Polyacrylamide gel electrophoresis under denaturing conditions was performed to evaluate the DNA cleavage ability of the engineered zinc-finger mutants towards their consensus DNA sequence. In contrast to native PAGE, denaturing PAGE eliminates all non-covalent interactions allowing for the detection of smaller and therefore faster migrating DNA fragments, which resulted from a possible hydrolysis. The polyacrylamide gels were prepared according to the procedure described above that however differs in the fact that the gel solution as well as the running buffer contained an additional 0.1% SDS. 2 μL of TriTrack DNA loading dye (*Fermentas*) was added to the samples, which were subsequently loaded on the gel. Electrophoresis was performed by applying a constant voltage of 170 V.

6x TriTrack Loading Dye: 10 m$\textsc{m}$ Tris HCl (pH 7.6)

60 m$\textsc{m}$ EDTA

0.03% bromophenol blue

0.03% xylene cyanol FF

0.15% orange G

Glycerol (60%)

6.3 Solid phase peptide synthesis

The peptides were synthesized using the Fmoc/*tert*-butyl SPPS protocol on either a Rink amide MBHA resin (loading density 0.56 mmol/g) in order to receive a C-terminal amide or on a Wang resin (loading density 0.33 mmol/g) to receive a carboxylic acid. Coupling of the first amino acid to the solid support was performed in a syringe (BD) equipped with a polyethylene-frit. The resin was swollen in DMF for 2 h. Subsequently, a solution containing the first Fmoc protected amino acid (5.0 eq relative to the resin loading), N,N′-diisopropylcarbodiimid (DIC) (5 eq) and Oxyma Pure (5 eq) in DMF (0.4 mL) was added to the syringe and the mixture was allowed to react in a Discover microwave reaction cavity (CEM, 75 °C, 35 W, 15 min). The liquid components were drained and the resin was washed with NMP (5x), DMF (5x) and finally with DCM (5x), before the resin was dried under reduced pressure. The loading density was estimated using the UV absorbance of the dibenzofulvalen species resulting from the Fmoc deprotection. Thus, 5 mg of the dry resin was placed in a graduated flask (10 mL) and 2% DBU in DMF (2 mL) was added and the mixture was allowed to react for 1 h. The solution was diluted with acetonitrile (1/12.5, v/v) and the absorption of the Fmoc cleavage product was detected by means of UV spectroscopy at 304 nm.

6.3.1 Coupling protocols for automated SPPS

Automated SPPS was performed on a microwave-assisted peptide synthesizer, either on the *Liberty 12 system* or on the *Liberty Blue* system (CEM, Kamp-Lintfort, Germany), which both were equipped with a *Discover* microwave reaction cavity (CEM). The pre-loaded resin was swollen in DMF or NMP for 2 h prior to SPPS. Standard Fmoc-amino acids were prepared as 0.2 M solutions in NMP in cases the Liberty 12 system was used or in DMF in cases the Liberty Blue system was used.

SPPS-A1

The SPPS was performed on a Liberty 12 peptide synthesizer. Amino acid coupling was achieved by a coupling cocktail, which contained HBTU/HOBt (0.50 M/0.45 M) in DMF as carboxyl activators and DIPEA (0.2 M) in NMP as activator base. Deprotection of the Fmoc group was achieved by piperidine (20% in NMP, 0.1 M HOBt). NMP and DCM were used for resin washing after each coupling or deprotection step. Microwave-assisted Fmoc deprotection was performed in two steps after the addition of the deprotection solution (2.5 mL). The first step involved microwave irradiation for 30 s at 55 °C (25 W) and the second step involved irradiation for 180 s at 75 °C (25 W). Standard amino acids except for cysteine and arginine residues were single-coupled for 600 s at 75 °C (25 W) or double-coupled for 2 x 600 s at 75 °C (25 W). Fmoc-Cys(Trt)-OH was double coupled for 2 x 900 s

at 50 °C (15 W) and Fmoc-Arg(Pbf)-OH was double coupled in two steps. The first step was performed without microwave irradiation for 600 s at 25 °C (0 W) and the second step was performed under microwave irradiation for 300 s at 75 °C (25 W).

SPPS-A2

The SPPS was performed on a Liberty Blue peptide synthesizer. Amino acid coupling was achieved by a coupling cocktail, which contained DIC/Oxyma Pure (0.50 M/1.0 M) in DMF. Deprotection of the Fmoc group was achieved by piperidine (20% in DMF, 0.1 M HOBt). DMF was used for resin washing after each coupling or deprotection step. Microwave assisted Fmoc deprotection was performed in two steps after the addition of the deprotection solution (2.5 mL). The first step involved microwave irradiation for 30 s at 55 °C (25 W) and the second step involved irradiation for 180 s at 75 °C (25 W). Standard amino acids except for cysteine, arginine and histidine residues were either single-coupled for 180 s at 90 °C (35 W) or double-coupled for 2 x 180 s at 90 °C (35 W). Fmoc-Cys(Trt)-OH and Fmoc-His(Trt)-OH were double-coupled for 2 x 360 s at 50 °C (35 W) and Fmoc-Arg(Pbf)-OH was double-coupled in two steps. The first step was performed without microwave irradiation for 1500 s at 25 °C (0 W) and the second step was performed under microwave irradiation for 120 s at 75 °C (30 W).

SPPS-A3

This method was used to couple the Fmoc-*O*-benzyl-L-phosphoserine building block as well as the subsequent amino acid on a Liberty Blue peptide synthesizer. The coupling was achieved by a coupling cocktail, which contained DIC/Oxyma Pure (0.50 M/1.0 M) in DMF. Deprotection of the Fmoc group was achieved by piperidine (20% in DMF, 0.1 M HOBt). DMF was used for resin washing after each coupling or deprotection step. Fmoc deprotection was performed under conventional conditions. These included the initial deprotection for 300 s at 25 °C (0 W) and the actual deprotection for 600 s at 25 °C (0 W). The residues were double-coupled for 2 x 900 s at 72 °C (25 W).

Manual solid phase peptide synthesis

Manual SPPS was performed using a Discover microwave reaction cavity (CEM). The pre-loaded resin was placed in a syringe (BD) equipped with a polyethylene frit. Prior to SPPS, the resin was swollen in DMF for 2 h.

6.3.2 Coupling protocols for manual SPPS

SPPS-M1

The Fmoc protecting groups were removed by the addition of piperidine (20% in NMP, 0.1 M HOBt) under microwave irradiation for 30 s at 55 °C (35 W) followed by irradiation for 180 s at 75 °C (45 W). Amino acid coupling was achieved by the addition of a coupling cocktail, which contained HATU/HOAt (0.50 M/0.45 M) in DMF as carboxyl activators and DIPEA (0.2 M) in NMP as activator base. All amino acids except for cysteine and arginine residues were either single-coupled for 600 s at 75 °C (25 W) or double-coupled for 2 x 600 s at 75 °C (25 W). Fmoc-Cys(Trt)-OH was double-coupled for 2 x 900 s at 50 °C (15 W) and Fmoc-Arg(Pbf)-OH was double-coupled in two steps. The first step was performed without microwave irradiation for 600 s at 25 °C (0 W) and the second step was performed under microwave irradiation for 300 s at 75 °C (25 W). The resin was excessively washed with NMP after each coupling and deprotection step.

SPPS-M2

This method was used to couple the artificial dinuclear building blocks as well as the subsequent amino acid to the pre-loaded resin. The Fmoc protecting groups were removed by the addition of piperidine (20% in NMP, 0.1 M HOBt) under microwave irradiation for 30 s at 55 °C (35 W) followed by irradiation for 180 s at 65 °C (45 W). Coupling was achieved by the addition of a coupling cocktail, which contained PyBOP/HOBt (1.0 M/0.5 M) in DMF as carboxyl activators and DIPEA (0.2 M) in NMP as activator base. The residues were double-coupled under microwave irradiation for 2 x 20 min at 65 °C (12 W). The resin was excessively washed with NMP after each coupling and deprotection step.

6.4 General methods in SPPS

Kaiser test

The coupling efficiency was tested by applying the Kaiser test in order to detect free primary amino groups. This was performed by the subsequent addition of 0.2 mL of the following solutions to a small sample of the resin (3 mg) as follows:

KT1 = Ninhydrin (1 g) in ethanol (20 mL)

KT2 = Phenol (40 g) in n-butanol (20 mL)

KT3 = KCN (1 mM in water, 1 mL) in pyridine (49 mL)

The samples were incubated at 95 ° for 5 min. The presence of free primary amino groups was indicated by a deep-blue-colored solution as well as resin beads.

Cleavage

The resin was transferred into a syringe (BD) equipped with a polyethylene frit. It was excessively washed with 5 x 4 mL NMP, 5 x 4 mL DMF, 5 x 4 mL MeOH and lastly with 5 x 4 mL DCM. The liquid components were removed under reduced pressure overnight.

Cleavage of the peptide from the resin as well as simultaneous deprotection of all acid labile side-chain protecting groups was performed in a solution containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) for 2 h. The solid support was separated from the liquid components and the solution was concentrated in a nitrogen stream. The crude peptide was precipitated by the addition of ice-cold diethyl ether and isolated by centrifugation (9.000 rpm, 15 min). The peptide pellet was dissolved in water/dioxane (5:1) and lyophilized. The product was purified by RP-HPLC, lyophilized again and stored under argon at –20 °C.

Native chemical ligation (NCL)

The modified Zf3 mutants were coupled to the expressed Zf12 thioester domain by means of NCL. Thus, the Zf3 mutants (1 mM) were dissolved in a degassed and argon-saturated buffer solution of Na_2HPO_4/NaH_2PO_4 (10 mM) which contained GdmCl (6 M) and TCEP (5 mM). The pH was adjusted to 4 and the peptides were incubated for 2 h at room temperature. The Zf12 thioester (0.5 mM) was dissolved in a buffer solution of Na_2HPO_4/NaH_2PO_4 (200 mM), which contained GdmCl (6 M) and the pH of the solution was adjusted to 8. Subsequently, both solutions were combined and the pH was validated to be in the range of 7.8 – 8.0. The reaction mixture was agitated at room temperature for 24 h. The ligation products were isolated and purified by semi-preparative RP-HPLC. After lyophilization, the Zf13 peptides were stored under an argon atmosphere at –20 °C.

6.5 Synthesis of peptide incorporable artificial amino acids as metallohydrolase mimics

1-Methyl-2-imidazolecarboxaldehyde
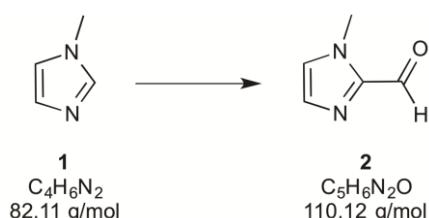
Under a nitrogen atmosphere, *n*-butyllithium (1.60 M, 30.5 cm^3, 48.7 mmol, 1 eq) was slowly added to a suspension of 1-methylimidazol (4.00 g, 48.7 mmol, 1 eq) in dry diethyl ether (140 mL) at −78 °C. The reaction mixture was allowed to stir at −50 °C for 1 h. Subsequently, a solution of DMF (5.66 mL, 73.1 mmol, 1.5 eq) in diethyl ether (10 mL) was added and the resulting mixture was stirred over night at −60 °C. The temperature was slowly allowed to rise to 0 °C before water (5 mL) was added to the reaction mixture over a period of 15 min followed by an aqueous solution of HCl (4 N, 35 mL). The layers were separated and the organic phase was washed with HCl (4 N, 5 x 10 mL). The aqueous extracts were saturated with K_2CO_3 and the crude product was extracted with chloroform (4 x 15 mL). The combined organic layers were dried over $MgSO_4$ and the solvents were evaporated under reduced pressure. The crude product was distilled under reduced pressure to give the product (2.21 g, 20.1 mmol, 41%) as colorless oil, which crystallized on standing.

^{1}H-NMR (300 MHz, CDCl$_3$): δ = 4.02 (s, 3H, CH$_3$), 7.07 (s, 1H, C4), 7.14 (s, 1H, C3), 9.8 (s, 1H, CHO) ppm.

ESI-MS *m/z*: 111.1 [M+H]$^+$.

ESI-HRMS: calculated for [C$_5$H$_6$N$_2$O]$^+$ ([M+H]$^+$) = 111.0553, found = 111.0557.

1-Methyl-2-imidazolecarboxaldehyde oxime

2
$C_5H_6N_2O$
110.12 g/mol

3
$C_5H_7N_3O$
125.13 g/mol
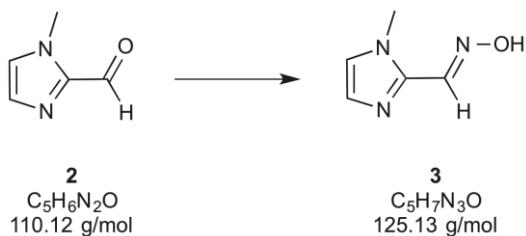
Hydroxylamine hydrochloride (1.27 g, 17.9 mmol, 1.1 eq) and Na_2CO_3 (0.87 g, 8.17 mmol, 0.5 eq) were added to a stirred solution of **2** (1.80 g, 16.4 mmol, 1 eq) in $EtOH/H_2O$ (12 mL, 1:4) at 0 °C. The reaction progress was followed by TLC (EtOAc/Pentane 12:1) until complete consumption of the reactants was observed. The resulted precipitate was filtered and subsequently washed with ice cold ethanol (50 mL). The product was obtained as a white solid in nearly quantitative yields (2.02 g, 16.2 mmol, 99%).

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.79 (s, 3H, CH_3), 6.91 (s, 1H, CH), 7.20 (s, 1H, CH), 8.09 (s, 1H, CH), 10.05 (s_{br}, 1H, OH) ppm.

Bis[(1-methylimidazol-2-yl)methyl]amine (BMIA)

3
$C_5H_7N_3O$
125.13 g/mol

4
$C_{10}H_{15}N_5$
205.26 g/mol
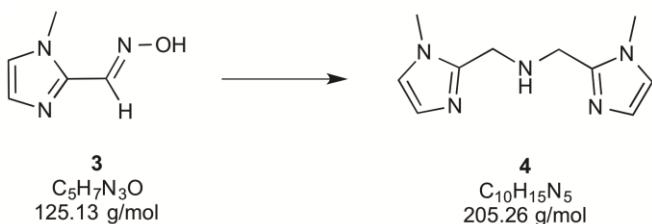
A suspension of 10% palladium on charcoal (250 mg) in methanol (5 mL) was added to a high pressure vessel containing a solution of **3** (1.50 g, 12.0 mmol) in MeOH (25 mL). The reaction mixture was hydrogenated at 3 bar hydrogen pressure for 24 h. Subsequently, the mixture was filtered over Celite to remove the catalyst and the solvent was evaporated to dryness under reduced pressure. The crude product was purified by flash column chromatography on silica gel (EtOAc/pentane, 12:1) to yield **4** (1.22 g, 5.93 mmol, 99%) as a yellow oil.

^{1}H-NMR (300 MHz, [D_6]-DMSO): δ = 3.59 (s, 6H, CH_3), 3.72 (s, 4H, CH_2), 6.76 (d, $^3J_{HH}$ = 1.2 Hz, 2H, C4), 7.04 (d, $^3J_{HH}$ = 1.2 Hz, 2H, C3) ppm.

ESI-MS: 206.1 [M+H]$^+$, 228.1 [M+Na]$^+$.

ESI-HRMS: calculated for [C$_{10}$H$_{15}$N$_5$Na]$^+$ ([M+Na]$^+$) = 228.1220, found = 228.1220.

Bis(2-pyridylmethyl)amine (BPA)

5	**6**	**7**	**8**
C$_6$H$_5$NO	C$_6$H$_8$N$_2$	C$_{12}$H$_{11}$N$_3$	C$_{12}$H$_{13}$N$_3$
107.11 g/mol	108.14 g/mol	197.24 g/mol	199.26 g/mol
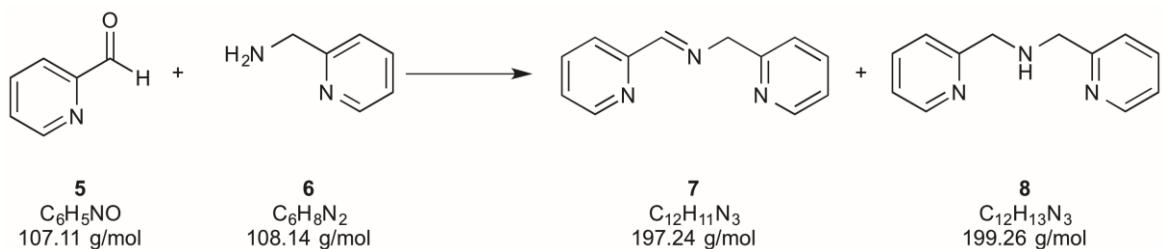
A solution of 2-picolyamine (2.24 g, 20.8 mmol, 1.1 eq) in MeOH (10 mL) was added to a solution of picolinaldehyde (2.02 g, 18.9 mmol, 1.0 eq) in MeOH (10 mL) and the reaction mixture was stirred at room temperature for 24 h. A solution of NaBH$_4$ (0.71 g, 18.9 mmol, 1.0 eq) in water (10 mL) was added at 0 °C and the mixture was allowed to stir for 1 h. The temperature was increased to room temperature and stirring was continued for another 24 h. The solution was concentrated under reduced pressure and water (10 mL) was added to the residue. The organic compounds were extracted with DCM (3 x 20 mL) and dried over MgSO$_4$. The solvent was evaporated under reduced pressure and the crude product was purified by flash column chromatography on silica gel (EtOAc/MeOH, 10:1). Product **8** was obtained as mixture with the non-reduced imine species **7** in a yield of 3.40 g (17.1 mmol, 90.5%).

^{1}H-NMR (300 MHz, CDCl$_3$): δ = 3.88 (s, 4H, CH$_2$), 7.04 (m, 2H, CH$_{ar}$), 7.52 (td, $^3J_{HH}$ = 1.8 Hz, 2H, CH$_{ar}$), 8.45 (m, 2H, CH$_{ar}$) ppm.

^{13}C-NMR (300 MHz, CDCl$_3$): δ = 54.36 (CH$_2$), 121.64 (C5), 121.98 (C3), 136.15 (C4), 148.95 (C6) ppm.

ESI-MS: 200.1 [M+H]$^+$.

ESI-HRMS: calculated for [C$_{12}$H$_{14}$N$_3$]$^+$ ([M+H]$^+$) = 200.1182, found = 200.1187.

Picolinaldehyde oxime

5
C_6H_5NO
107.11 g/mol

9
$C_6H_6N_2O$
122.13 g/mol
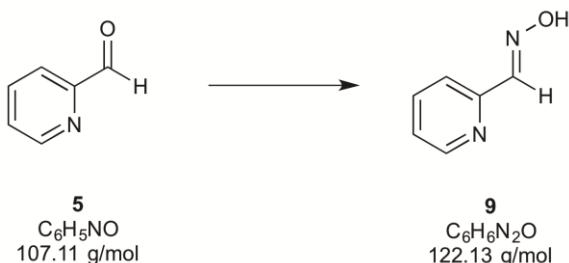
Hydroxylamine hydrochloride (1.81 g, 25.7 mmol, 1.1 eq) and Na_2CO_3 (1.24 g, 11.7 mmol, 0.5 eq) were added to a stirred solution of **5** (2.50 g, 23.3 mmol, 1 eq) in $EtOH/H_2O$ (16 mL, 1:4) at 0 °C. The reaction progress was followed by TLC (EtOAc/Pentane 12:1) until complete consumption of the reactants was observed. The resulted precipitate was filtered and subsequently washed with ice cold ethanol (65 mL). The product was obtained as a white solid in nearly quantitative yields (2.82 g, 23.1 mmol, 99%).

^{1}H-NMR (300 MHz, CDCl$_3$): δ = 7.16 (m, 1H, CH$_{ar}$), 7.62 (m, 1H, CH$_{ar}$), 7.88 (m, 1H, CH$_{ar}$), 8.29 (s, 1H, CH), 8.52 (s, 1H, CH$_{ar}$), 9.91 (s$_{br}$, 1H, OH) ppm.

Bis(2-pyridylmethyl)amine (BPA)

9
$C_6H_6N_2O$
122.13 g/mol

8
$C_{12}H_{13}N_3$
199.25 g/mol
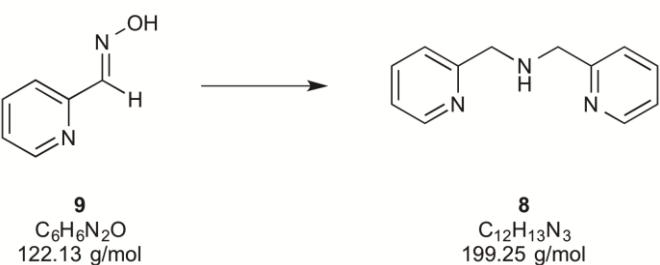
A suspension of 10% palladium on charcoal (350 mg) in methanol (5 mL) was added to a high pressure vessel containing a solution of **9** (2.25 g, 18.4 mmol) in MeOH (30 mL). The reaction mixture was hydrogenated at 3 bar hydrogen pressure for 24 h. Water (20 mL) was added and the mixture was filtered over Celite to remove the catalyst. The solution was concentrated *in vacuo* and the organic compounds were extracted with EtOAc (2 x 15 mL). The combined organic extracts were washed with brine (2 x 15 mL) and water (2 x 15 mL) before the solvent was evaporated to dryness under reduced pressure. The crude product was purified by flash column chromatography on silica gel (EtOAc/pentane, 12:1) to yield **8** (1.82 g, 9.12 mmol, 99%) as a yellow oil.

¹H-NMR (300 MHz, CDCl₃): δ = 3.88 (s, 4H, CH₂), 7.04 (m, 2H, CH$_{ar}$), 7.52 (td, $^3J_{HH}$ = 1.8 Hz, 2H, CH$_{ar}$), 8.45 (m, 2H, CH$_{ar}$) ppm.

¹³C-NMR (300 MHz, CDCl₃): δ = 54.36 (CH₂), 121.64 (C5), 121.98 (C3), 136.15 (C4), 148.95 (C6) ppm.

ESI-MS: 200.1 [M+H]⁺.

ESI-HRMS: calculated for [C₁₂H₁₄N₃]⁺ ([M+H]⁺) = 200.1182, found = 200.1187.

tert-Butyl-(pyridine-2-ylmethyl)glycine

5	10	11
C₆H₅NO	C₆H₁₃NO₂	C₁₂H₁₈N₂O₂
107.11 g/mol	131.18 g/mol	222.29 g/mol
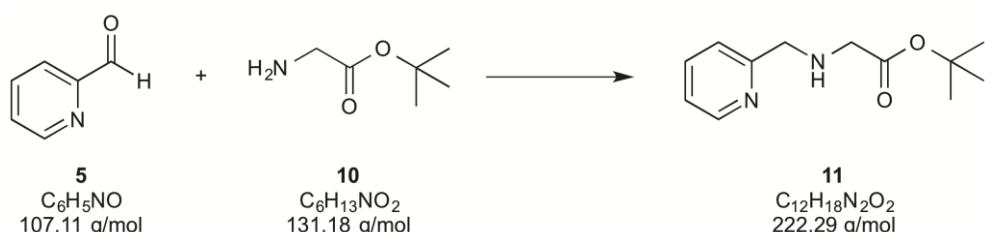
A solution of glycine *tert*-butyl ester hydrochloride (3.44 g, 20.5 mmol, 1.1 eq) in MeOH (10 mL) was added to a solution of picolinaldehyde (2.00 g, 18.7 mmol, 1.0 eq) in MeOH (10 mL) and the reaction mixture was stirred at room temperature for 24 h. The reaction mixture was cooled to 0 °C and NaBH₄ (2.12 g, 56.9 mmol, 3.0 eq) was added in small portions over a period of 2 h while stirring. The temperature was increased to room temperature and stirring was continued for another 24 h. The solution was concentrated under reduced pressure to a final volume of 5 mL before water (10 mL) was added to the residue. The organic compounds were extracted with DCM (3 x 20 mL) and dried over MgSO₄. The solvent was evaporated under reduced pressure and the crude product was purified by flash column chromatography on silica gel (EtOAc/MeOH, 10:1). Product **11** was obtained as a yellow oil (1.83 g, 8.22 mmol, 44%).

¹H-NMR (300 MHz, CDCl₃): δ = 1.40 (m, 9H, *tert*-Bu), 3.87 (s, 2H, CH₂), 4.69 (s, 2H, CH₂), 7.11 (m, 2H, CH$_{ar}$) 7.61 (m, 1H, CH$_{ar}$), 8.47 (m, 1H, CH$_{ar}$) ppm.

ESI-MS: 223.2 [M+H]⁺.

ESI-HRMS: calculated for [C₁₂H₁₉N₂O₂]⁺ ([M+H]⁺) = 223.1441, found = 223.1443.

Boc-Tyr-[2,6-(bis(2(*N*-methylimidazol)methyl)amine]-OMe

4	**12**	**13**
$C_{10}H_{15}N_5$	$C_{15}H_{21}NO_5$	$C_{36}H_{49}N_{11}O_5$
205.26 g/mol	295.33 g/mol	715.86 g/mol
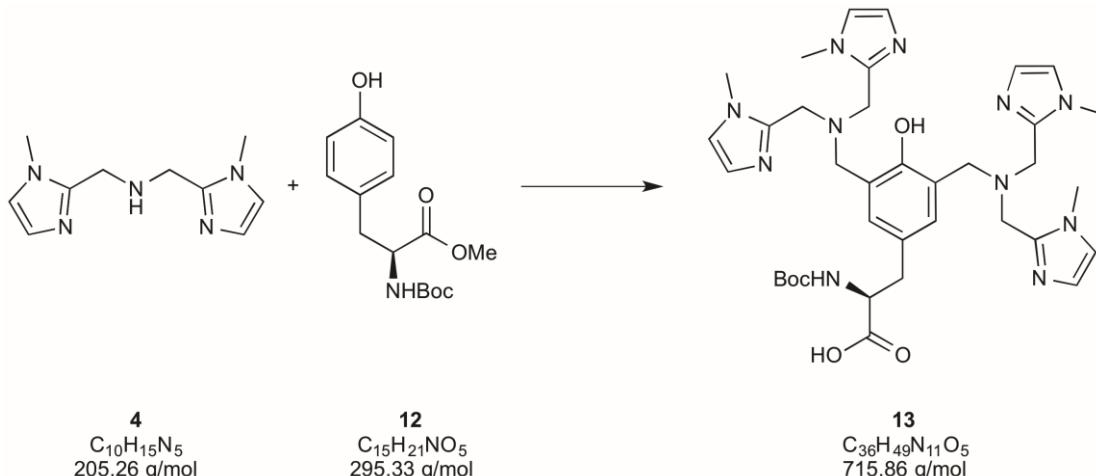
Ligand **4** (2.61 g, 12.7 mmol, 2.5 eq) was added to a suspension of paraformaldehyde (0.38 mg, 12.7 mmol, 2.5 eq) in $EtOH/H_2O$ (24 mL, 1:4). The reaction mixture was stirred at 60 °C for 90 min. Boc-Tyr-OMe (1.50 g, 5.08 mmol, 1 eq) was dissolved in $EtOH/H_2O$ (8 mL, 1:4) and added to the previously prepared Schiff base cocktail while stirring. The pH was adjusted to approximately 6.5 – 7.0 by the addition of 1 M HCl (aq.) and the reaction mixture was allowed to stir at 95 °C for 36 h. The reaction progress was monitored by taking aliquots, which were analyzed by means of ESI mass spectrometry

The organic compounds were extracted with DCM (3 x 30 mL). The combined organic extracts were washed with brine (25 mL) and water (2 x 25 mL) and dried over $MgSO_4$. The solvent was removed under reduced pressure and the crude product was purified by RP-flash column chromatography on C18 silica gel with water/ethanol (4:1) as eluting system. Product **13** was obtained as a yellow solid (872.6 mg, 1.22 mmol, 24%).

¹H-NMR (300 MHz, [D_6]-DMSO): δ = 3.57 (s, 9H, *tert*-Bu), 3.59 (s, 6H, CH_2), 3.82 (s, 12H, CH_3), 3.86 (s, 8H, CH_2), 4.31 (s$_{br}$, 1H, α-H), 5.41 (s$_{br}$, 1H, OH), 6.78 (d, $^3J_{HH}$ = 1.4 Hz, 4H, CH), 6.91 (s, 2H, CH_{ar}), 6.93 (d, $^3J_{HH}$ = 1.3 Hz, 4H, CH), 7.11 (s$_{br}$, 1H, NH) ppm.

¹³C-NMR (300 MHz, [D_6]-DMSO): δ = 28.48 (CH_3, *tert*-Bu), 32.54 (CH_3, Me), 44.60 (CH_2), 49.31 (CH_2), 54.93 (β-CH_2), 77.21 (C1, *tert*-Bu), 121.29 (CH_{im}), 123.12 (C_{ar}), 126.22 (CH_{ar}), 127.15 (CH_{im}), 145.13 (C_{im}), 146.19 ($C_{ar}OH$), 154.83 (COOH), 155.13 ($\underline{C}OOt$Bu) ppm.

ESI-MS: 716.4 [M+H]$^+$, 738.4 [M+Na]$^+$

ESI-HRMS: calculated for [$C_{36}H_{50}N_{11}O_5$]$^+$ ([M+H]$^+$) = 716.3991, found = 716.3984.

Boc-Tyr-[2,6-(bis(2(*N*-methylimidazol)methyl)amine]-OH

13
C$_{36}$H$_{49}$N$_{11}$O$_5$
715.86 g/mol

14
C$_{46}$H$_{51}$N$_{11}$O$_5$
837.98 g/mol
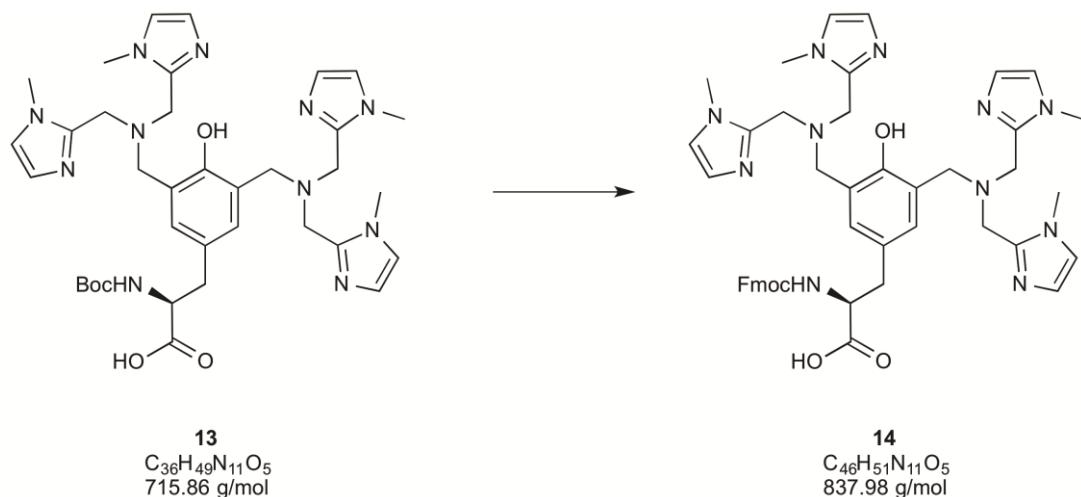
A solution of TFA/H$_2$O (5 mL, 95:5) was carefully added to a centrifuge tube containing product **13** (0.75 g, 1.05 mmol) and the reaction mixture was stirred at room temperature for 2 h. The volatile components were removed in a nitrogen stream. Ice-cold diethyl ether (35 mL) was added and the tube was centrifuged (9.000 rpm, 15 min). The supernatant was removed and the precipitate was collected to yield the fully unprotected building block. The intermediate product (0.64 g, 1.04 mmol, 1 eq) was dissolved in an aqueous solution of 10% Na$_2$CO$_3$ (5 mL) and cooled to 0 °C. Fmoc-OSu (0.77 g, 2.28 mmol, 2.2 eq) was dissolved in dioxane (5 mL), which was also cooled to 0 °C and subsequently added to the aqueous solution. The reaction mixture was stirred at 0 °C for 30 min and allowed to warm to room temperature overnight. The reaction was quenched by the addition of water (5 mL) and the organic compounds were extracted with EtOAc (3 x 15 mL). The combined aqueous layers were acidified to pH 1 with HCl (1 M) and extracted with EtOAc (3 x 15 mL). The combined organic layers were dried over MgSO$_4$ and the solvent was removed under reduced pressure. The crude product was purified by RP-flash column chromatography on C18 silica gel with water/ethanol (4:1) as eluting system. Product **14** (0.83 g, 0.99 mmol, 96%) was obtained as a brownish solid.

^{1}H-NMR (300 MHz, [D$_6$]-DMSO): δ = 2.61 (m, 1H, CH), 2.93 (m, 1H, CH), 3.52 (s, 12H, CH$_3$), 3.78 (s, 8H, CH$_2$), 3.96 (s, 6H, CH$_2$), 4.23 (m, 1H, α-CH), 7.00 (s, 2H, CH$_{tyr}$), 7.27 (d, $^3J_{HH}$ = 6.2 Hz, 2H, CH$_2$-Fmoc), 7.37 (d, $^3J_{HH}$ = 1.2 Hz, 4H, CH$_{im}$), 7.39 (d, $^3J_{HH}$ = 1.2 Hz, 4H, CH$_2$-Fmoc), 7.47 (d, $^3J_{HH}$ = 1.2 Hz, 4H, CH$_{im}$), 7.55 (m, 2H, H2-Fmoc), 7.58 (m, 2H, H3-Fmoc), 7.63 (m, 2H, H1-Fmoc), 7.84 (d, 2H, H4-Fmoc), 13.14 (s, 1H, COOH) ppm.

^{13}C-NMR (300 MHz, [D$_6$]-DMSO): δ = 52.43 (CH$_3$), 43.65 (CH$_2$), 109.53 (C-OH), 119.94 (C4-Fmoc), 121.32 (C1-Fmoc), 121.62 (CH$_{im}$), 125.49 (CH$_{im}$), 127.12 (C2-Fmoc), 128.84 (C3-Fmoc), 139.45 (C6-Fmoc), 145.42 (C5-Fmoc), 172.67 (CO-Fmoc), 185.32 (COOH) ppm.

ESI-MS: 838.4 [M+H]$^+$, 860.4 [M+Na]$^+$.

ESI-HRMS: calculated for [C$_{47}$H$_{52}$N$_{11}$O$_5$]$^+$ ([M+H]$^+$) = 838.4147, found = 838.4156.

Boc-Tyr-[2,6-bis(2,2'(pyridylmethyl)amine)]-OMe

8	**12**	**15**
C$_{12}$H$_{13}$N$_3$	C$_{15}$H$_{21}$NO$_5$	C$_{41}$H$_{47}$N$_7$O$_5$
199.26 g/mol	295.33 g/mol	717.87 g/mol
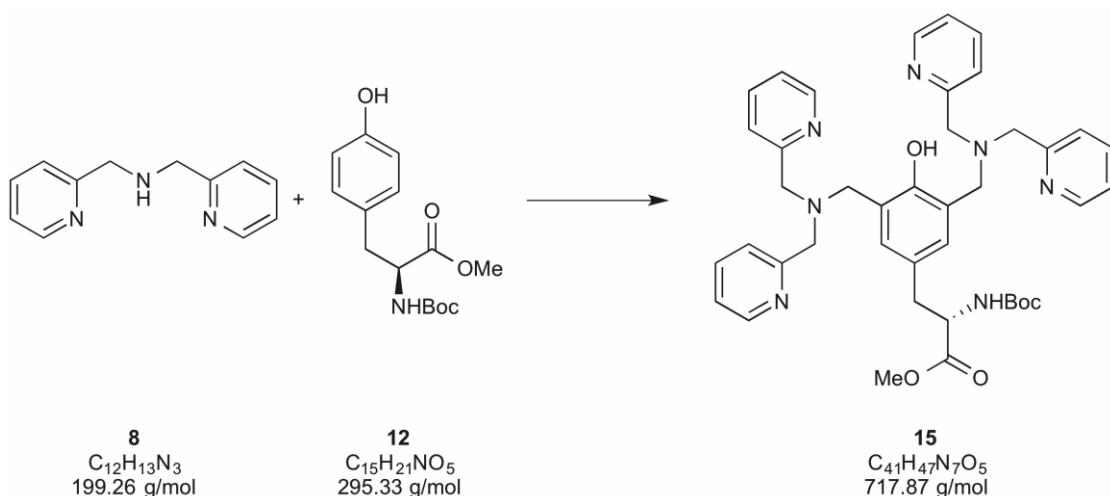
Ligand **8** (2.02 g, 10.2 mmol, 2.5 eq) was added to a suspension of paraformaldehyde (0.31 mg, 10.2 mmol, 2.5 eq) in EtOH/H$_2$O (24 mL, 1:4). The reaction mixture was stirred at 60 °C for 90 min. Boc-Tyr-OMe (1.10 g, 4.06 mmol, 1 eq) was dissolved in EtOH/H$_2$O (8 mL, 1:4) and added to the previously prepared Schiff base cocktail while stirring. The pH was adjusted to approximately 6.5 - 7.0 by the addition of 1 M HCl (aq.) and the reaction mixture was allowed to stir at 95 °C for 36 h. The reaction progress was monitored by taking aliquots, which were analyzed by means of ESI mass spectrometry

The organic compounds were extracted with DCM (3 x 30 mL). The combined organic extracts were washed with brine (25 mL) and water (2 x 25 mL) and dried over MgSO$_4$. The solvent was removed under reduced pressure and the crude product was purified by RP-flash column chromatography on C18 silica gel with water/ethanol (4:1) as eluting system. Product **13** was obtained as a brownish solid (2.76 g, 3.84 mmol, 95%).

¹H-NMR (300 MHz, [D₆]-DMSO): δ = 2.43 (s, 9H, *tert*-Bu), 3.51 (m, 6H, CH₂), 3.76 (m, 8H, CH₂), 3.88 (s, 3H, CH₃), 4.42 (s_br, 1H, α-H), 5.66 (s_br, 1H, OH), 6.91 (s, 2H, CH_ar), 7.11 (s_br, 1H, NH), 7.31 (m, 8H, CH_py), 7.75 (m, 4H, CH_py), 8.22 (m, 4H, CH_py) ppm.

¹³C-NMR (300 MHz, [D₆]-DMSO): δ = 28.41 (CH₃, *tert*-Bu), 32.37 (CH₃, Me), 45.62 (CH₂), 49.11 (CH₂), 55.08 (β-CH₂), 77.32 (C1, *tert*-Bu), 120.31 (CH_py), 123.52 (C_ar), 128.56 (CH_py), 129.98 (CH_py), 131.42 (C_ar), 145.13 (C_py), 146.19 (C_arOH),151.78 (C_py), 154.43 (<u>C</u>OOMe), 155.95 (<u>C</u>OO*t*Bu) ppm.

ESI-MS: 718.4 [M+H]⁺, 740.4 [M+Na]⁺.

ESI-HRMS: calculated for [C₄₁H₄₈N₇O₅]⁺ ([M+H]⁺) = 718.3711, found = 718.3713.

calculated for [C₄₁H₄₇O₅Na]⁺ ([M+Na]⁺) = 740.3531, found = 740.3532.

Fmoc-Tyr-[2,6-bis(2,2'(pyridylmethyl)amine)]-OH

15
C₄₁H₄₇N₇O₅
717.87 g/mol

16
C₅₀H₄₇N₇O₅
825.97 g/mol
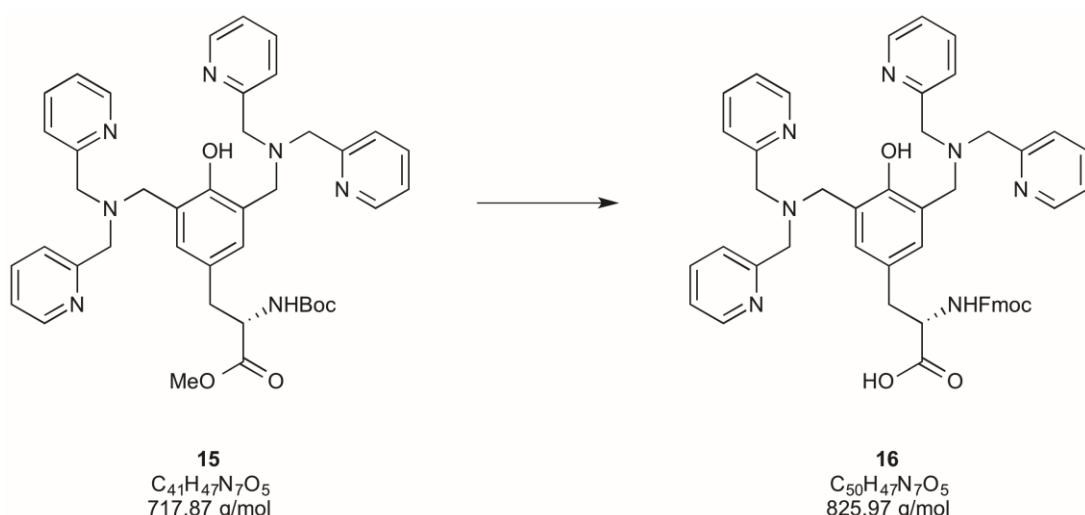
Compound **15** (2.10 g, 2.93 mmol) was dissolved in MeOH (10 mL) and an aqueous solution of NaOH (1 M) was added in excess. The reaction mixture was heated to 95 °C and stirred for 2 h. The solution was concentrated *in vacuo* before a solution of TFA/H₂O (5 mL, 95:5) was carefully added and the mixture was stirred at room temperature for 2 h. The volatile components were removed in a nitrogen stream. Ice-cold diethyl ether (35 mL) was added and the formed precipitate was centrifuged (9.000 rpm, 15 min). The supernatant was removed and the precipitate was collected to yield the fully unprotected building block. The intermediate product (2.00 g, 3.31 mmol, 1 eq) was dissolved in an aqueous solution of 10% Na₂CO₃ (10 mL) and cooled to 0 °C. Fmoc-OSu (2.46 mg, 7.29 mmol, 2.2 eq) was dissolved

in dioxane (10 mL), which was also cooled to 0 °C and subsequently added to the aqueous solution. The reaction mixture was stirred at 0 °C for 30 min and allowed to warm to room temperature overnight. The reaction was quenched by the addition of water (15 mL) and the organic compounds were extracted with EtOAc (3 x 25 mL). The combined aqueous layers were acidified to pH 1 with HCl (1 M) and extracted with EtOAc (3 x 25 mL). The combined organic layers were dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude product was purified by RP-flash column chromatography on C18 silica gel with water/ethanol (4:1) as eluting system. Product **16** (1.23 g, 1.49 mmol, 45%) was obtained as a brownish solid.

^{1}H-NMR (300 MHz, [D$_6$]-DMSO): δ = 2.79 (m, 1H, CH$_{\beta,cis}$), 2.97 (m, 1H, CH$_{\beta,trans}$), 3.48 (s, 3H, CH$_3$), 3.74 (m, 8H, CH$_{py}$), 4.22 (s, 8H, CH$_{py}$), 7.35 (m, 4H, CH$_{ar}$), 7.52 (m, 7H, CH$_{ar}$), 7.68 (m, 7H, CH$_{ar}$), 8.01 (m, 4H, CH$_{ar}$), 8.67 (m, 4H, CH$_{ar}$) ppm.

ESI-MS: 840.4 [M+H]$^+$.

ESI-HRMS: calculated for [C$_{51}$H$_{50}$N$_7$O$_5$]$^+$ ([M+H]$^+$) = 840.3870, found = 840.3860.

Boc-Tyr-[2-(bis(2-(*N*-methylimidazol)methyl)amine]-OH

4	**12**	**17**
C$_{10}$H$_{15}$N$_5$	C$_{15}$H$_{21}$NO$_5$	C$_{26}$H$_{36}$N$_6$O$_5$
205.26 g/mol	295.33 g/mol	512.61 g/mol
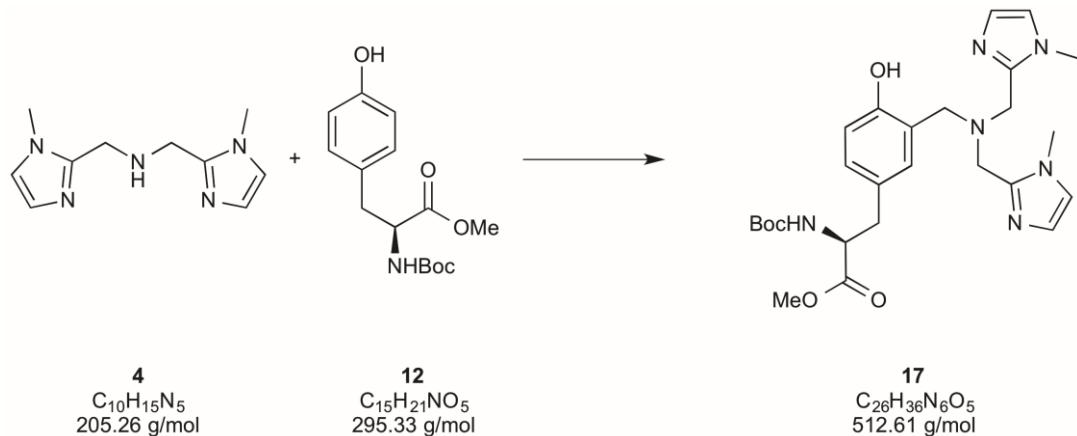
Ligand **4** (0.99 g, 4.83 mmol, 0.95 eq) was added to a suspension of paraformaldehyde (0.15 g, 5.08 mmol, 1.0 eq) in EtOH/H$_2$O (12 mL, 1:4). The reaction mixture was stirred at 60 °C for 90 min. Boc-Tyr-OMe (1.50 g, 5.08 mmol, 1.0 eq) was dissolved in EtOH/H$_2$O (4 mL, 1:4) and added to the previously prepared Schiff base cocktail while stirring. The pH was adjusted to approximately 6.5 – 7.0 by the addition of 1 M HCl (aq.) and the reaction mixture was allowed to stir at 95 °C for 36 h.

The organic compounds were extracted with DCM (3 x 30 mL). The combined organic extracts were washed with brine (25 mL) and water (2 x 25 mL) and dried over MgSO$_4$. The solvent was removed under reduced pressure and the crude product was purified by RP-flash column chromatography on C18 silica gel with water/ethanol (4:1) as eluting system. Product **17** was obtained as a yellow solid (2.40 g, 4.67 mmol, 92%).

^{1}H-NMR (300 MHz, [D$_6$]-DMSO): δ = 2.12 (s, 9H, *tert*-Bu), 3.01 (m, 1H, CH$_2$), 3.28 (m, 1H, CH$_2$), 3.72 (m, 5H, CH$_3$, CH$_2$), 3.92 (s, 6H, OCH$_3$), 4.41 (s, 4H, CH$_2$), 4.61 (s$_{br}$, 1H, α-H), 6.67 (d, $^3J_{HH}$ = 1.4, 2H, CH$_{im}$), 6.87 (d, $^3J_{HH}$ = 1.4, 2H, CH$_{im}$), 6.93 (s, 2H, CH$_{ar}$), 7.19 (s$_{br}$, 1H, NH) ppm.

^{13}C-NMR (300 MHz, [D$_6$]-DMSO): δ = 28.38 (CH$_3$, *tert*-Bu), 33.51 (CH$_3$), 37.12 (CH$_2$), 50.41 (CH$_2$), 51.87 (OCH$_3$), 56.32 (CH$_2$), 57.87 (α-CH), 79.46 (*tert*-Bu), 115.83 (CH$_{ar}$), 120.34 (CH$_{im}$), 122.78 (C$_{ar}$), 126.04 (CH$_{im}$), 128.22 (CH$_{ar}$), 130.22 (CH$_{ar}$), 131.32 (C$_{ar}$), 147.20 (C$_{im}$), 148.54 (C$_{ar}$OH), 149.52 (COOtBu), 153.49 (COOMe) ppm.

ESI-MS: 513.3 [M+H]$^+$.

ESI-HRMS: calculated for [C$_{26}$H$_{36}$N$_6$O$_5$]$^+$ ([M+H]$^+$) = 513.2820, found = 513.2819.

Boc-Tyr-[2-bis(2,2'(pyridylmethyl)amine)]-OMe
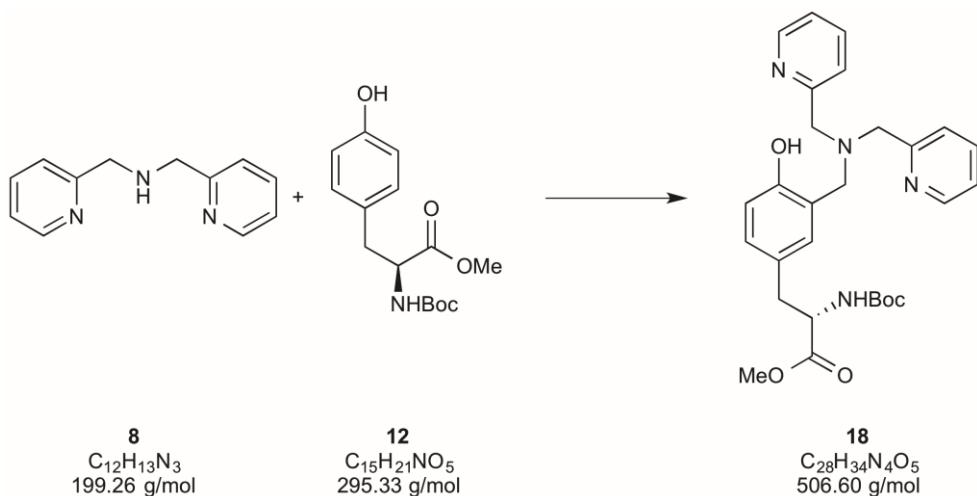
Ligand **8** (0.70 g, 3.86 mmol, 0.95 eq) was added to a suspension of paraformaldehyde (0.12 g, 4.06 mmol, 1.0 eq) in EtOH/H$_2$O (12 mL, 1:4). The reaction mixture was stirred at 60 °C for 90 min. Boc-Tyr-OMe (1.20 g, 4.06 mmol, 1.0 eq) was dissolved in EtOH/H$_2$O

140

(4 mL, 1:4) and added to the previously prepared Schiff base cocktail while stirring. The pH was adjusted to approximately 6.5 – 7.0 by the addition of 1 M HCl (aq.) and the reaction mixture was allowed to stir at 95 °C for 36 h.

The organic compounds were extracted with DCM (3 x 30 mL). The combined organic extracts were washed with brine (25 mL) and water (2 x 25 mL) and dried over MgSO$_4$. The solvent was removed under reduced pressure and the crude product was purified by RP-flash column chromatography on C18 silica gel with water/ethanol (4:1) as eluting system. Product **18** was obtained as a yellow solid (1.87 g, 3.70 mmol, 91%).

^{1}H-NMR (300 MHz, [D$_6$]-DMSO): δ = 2.03 (s, 9H, *tert*-Bu), 2.95 (m, 1H, CH$_2$), 3.18 (m, 1H, CH$_2$), 3.67 (s, 3H, CH$_3$), 3.79 (s, 2H, CH$_2$), 3.95 (s, 4H, CH$_2$), 4.62 (s$_{br}$, 1H, α-H), 6.90 (m, 3H, CH$_{ar}$), 7.17 (s$_{br}$, 1H, NH), 7.25 (m, 4H, CH$_{py}$), 7.78 (m, 2H, CH$_{py}$), 7.91 (m, 2H, CH$_{py}$) ppm.

^{13}C-NMR (300 MHz, [D$_6$]-DMSO): δ = 28.36 (CH$_3$, *tert*-Bu), 37.17 (CH$_2$), 51.37 (CH$_2$), 51.68 (OCH$_3$), 55.10 (CH$_2$), 56.27 (α-CH), 78.89 (*tert*-Bu), 115.39 (CH$_{ar}$), 120.87 (CH$_{py}$), 122.65 (C$_{ar}$), 124.07 (CH$_{py}$), 126.74 (CH$_{ar}$), 129.02 (CH$_{ar}$), 130.34 (C$_{ar}$), 139.13 (CH$_{py}$), 144.22 (CH$_{py}$), 146.89 (C$_{ar}$OH), 149.36 (C$_{py}$), 151.32 (<u>C</u>OOtBu), 153.21 (<u>C</u>OOMe) ppm.

ESI-MS: 507.3 [M+H]$^+$.

ESI-HRMS: calculated for [C$_{28}$H$_{34}$N$_4$O$_5$]$^+$ ([M+H]$^+$) = 507.2602, found = 507.2611.

Fmoc-Tyr-[2-bis(2,2'(pyridylmethyl)amine)]-OBn

8	**19**	**20**
C$_{12}$H$_{13}$N$_3$	C$_{31}$H$_{27}$NO$_5$	C$_{44}$H$_{40}$N$_4$O$_5$
199.26 g/mol	493.56 g/mol	704.83 g/mol
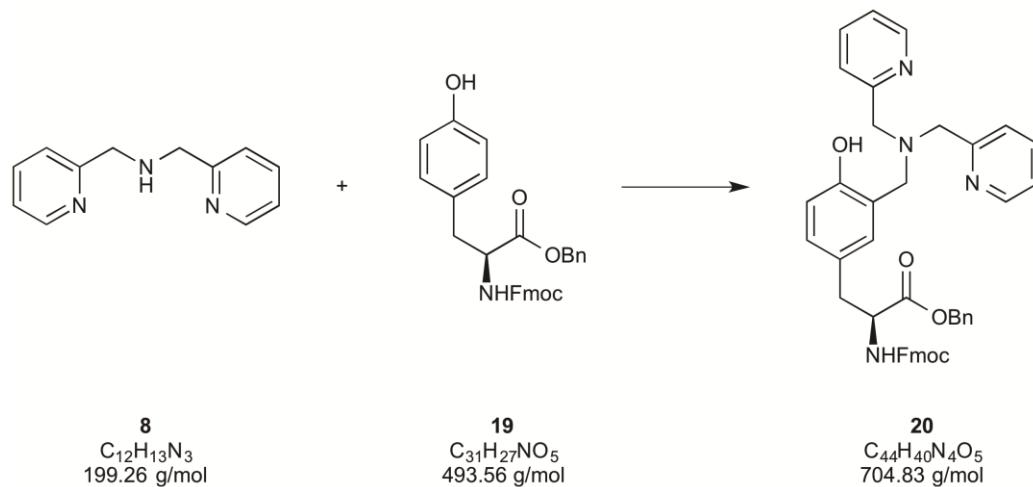
Ligand **8** (0.19 g, 0.96 mmol, 0.95 eq) was added to a suspension of paraformaldehyde (0.03 g, 1.01 mmol, 1.0 eq) in EtOH/H_2O (8 mL, 1:4). The pH was adjusted to 6.5 with 1 M HCl (aq.) and the reaction mixture was stirred at 65 °C for 2 h. Fmoc-Tyr-OBn (0.50 g, 1.01 mmol, 1.0 eq) was dissolved in EtOH/H_2O (4 mL, 1:4) and added to the previously prepared Schiff base cocktail while stirring. The reaction mixture was allowed to stir at 95 °C for 36 h.

The organic compounds were extracted with DCM (3 x 15 mL). The combined organic extracts were washed with brine (20 mL) and water (2 x 20 mL) and dried over $MgSO_4$. The solvent was removed under reduced pressure to yield product **20** as a yellow oil (114.2 mg, 0.162 mmol, 16%).

HPLC (A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-70% B, 30 min): t_R = 10.53 min.

ESI-MS: 705.3 [M+H]$^+$.

Fmoc-Tyr-[2-bis(2,2'(pyridylmethyl)amine)-6-*(tert*-Butyl-(pyridin-2-ylmethyl)glycine)]-OBn

11	20	21
$C_{12}H_{18}N_2O_2$	$C_{44}H_{40}N_4O_5$	$C_{57}H_{58}N_6O_7$
222.29 g/mol	704.83 g/mol	939.13 g/mol
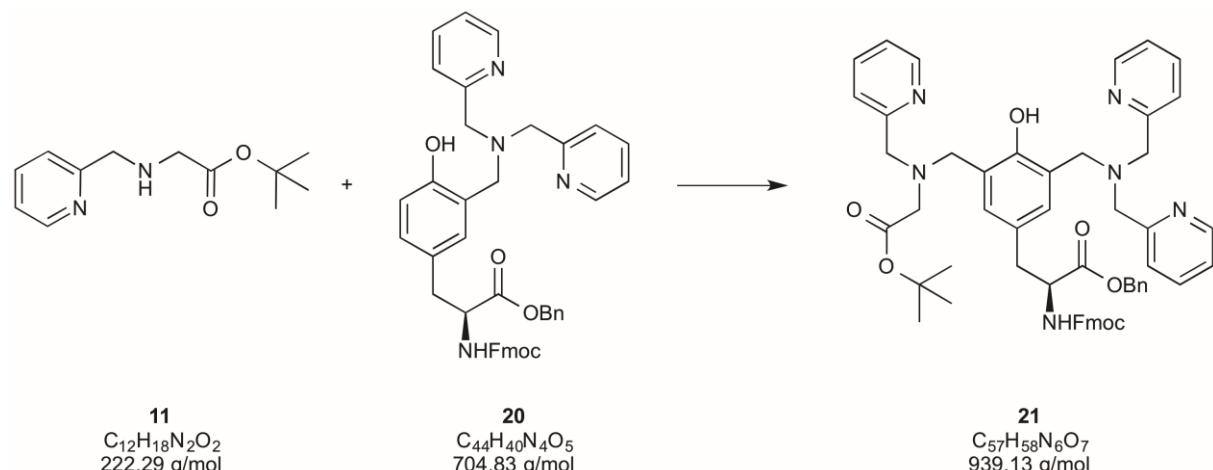
Ligand **11** (34.7 mg, 0.156 mmol, 1.1 eq) was added to a suspension of paraformaldehyde (4.69 mg, 0.156 mmol, 1.1 eq) in EtOH/H_2O (5 mL, 1:4). The pH was adjusted to 6.5 with 1 M HCl (aq.) and the reaction mixture was stirred at 65 °C for 2 h. Compound **20** (0.10 g, 0.142 mmol, 1.0 eq) was dissolved in EtOH/H_2O (3 mL, 1:4) and added to the previously prepared Schiff base cocktail while stirring. The reaction mixture was allowed to stir at 95 °C for 36 h.

The organic compounds were extracted with DCM (3 x 15 mL). The combined organic extracts were washed with brine (10 mL) and water (2 x 10 mL) and dried over MgSO$_4$. The solvent was removed under reduced pressure to yield product **21** as a yellow oil (2.66 mg, 0.028 mmol, 2%).

HPLC (A (99.9% H$_2$O, 0.10% TFA), B (79.9% ACN, 20.0% H$_2$O, 0.10% TFA), gradient 10-70% B, 30 min): t_R = 12.24 min.

ESI-MS: 939.4 [M+H]$^+$.

Solid phase peptide synthesis of Zf3 domains

Zf3 Wildtype
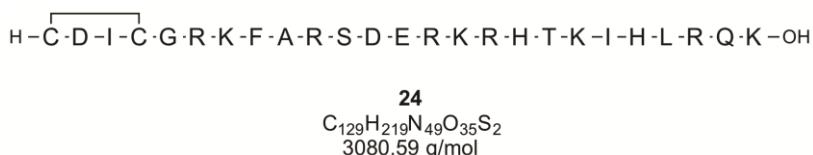
H–C·D·I·C·G·R·K·F·A·R·S·D·E·R·K·R·H·T·K–I·H·L·R·Q·K–OH

24
C$_{129}$H$_{219}$N$_{49}$O$_{35}$S$_2$
3080.59 g/mol

The native sequence of Zf3 was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol using an automated peptide synthesizer (Liberty 12, CEM). Coupling was performed according to the Fmoc/*tert*-butyl protocol SPPS-A1 described in section 6.3.1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H$_2$O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **24** (108 mg, 33.1 mmol, 38.2%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H$_2$O, 0.10% TFA), B (79.9% ACN, 20.0% H$_2$O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 16.58 min.

ESI-MS *m/z*: 514.1 [M+6H]$^{6+}$, 616.7 [M+5H]$^{5+}$, 770.9 [M+4H]$^{4+}$.

ESI-HRMS: calculated for [C$_{129}$H$_{225}$N$_{49}$O$_{35}$S$_2$]$^{6+}$ ([M+6H]$^{6+}$) = 514.1124, found = 514.1125, calculated for [C$_{129}$H$_{224}$N$_{49}$O$_{35}$S$_2$]$^{5+}$ ([M+5H]$^{5+}$) = 616.7334, found = 616.7335.

Zf3BMIA70

25a
$C_{154}H_{246}N_{56}O_{36}S_2$
3522.14 g/mol
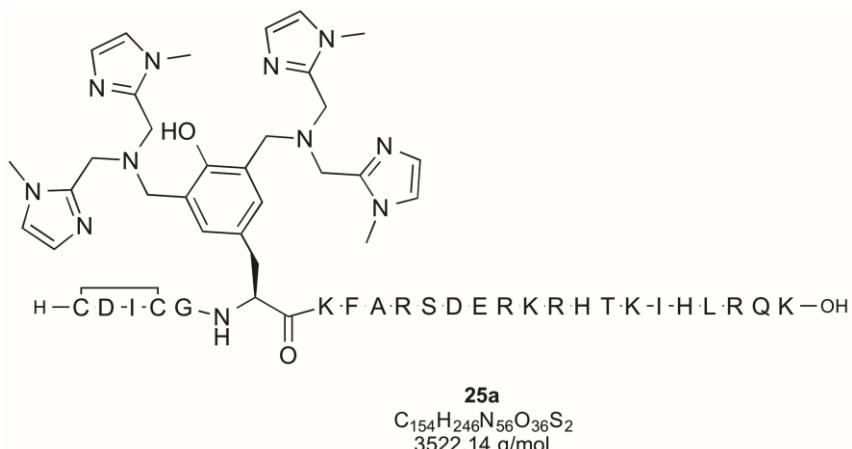
Zf3BMIA70 (**25a**) was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol. The first part of the sequence until the lysine residue at position 71 was synthesized by means of an automated peptide synthesizer (Liberty 12, CEM) using protocol SPPS-A1 (section 6.3.1). The building block **14** and the subsequent amino acid were coupled manually according to protocol SPPS-M2. The remaining amino acid residues were also coupled manually according to conventional protocol SPPS-M1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **25a** (28 mg, 0.79 mmol, 18.4%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 17.46 min.

ESI-MS *m/z*: 687.6 $[M+6H]^{6+}$, 704.9 $[M+5H]^{5+}$, 880.9 $[M+4H]^{4+}$.

ESI-HRMS: calculated for $[C_{154}H_{251}N_{56}O_{36}S_2]^{5+}$ ($[M+5H]^{5+}$) = 704.9789, found = 704.9794, calculated for $[C_{154}H_{250}N_{56}O_{36}S_2]^{4+}$ ($[M+4H]^{4+}$) = 880.9718, found = 880.9723.

Zf3BPA70

25b
$C_{158}H_{242}N_{52}O_{36}S_2$
3510.12 g/mol
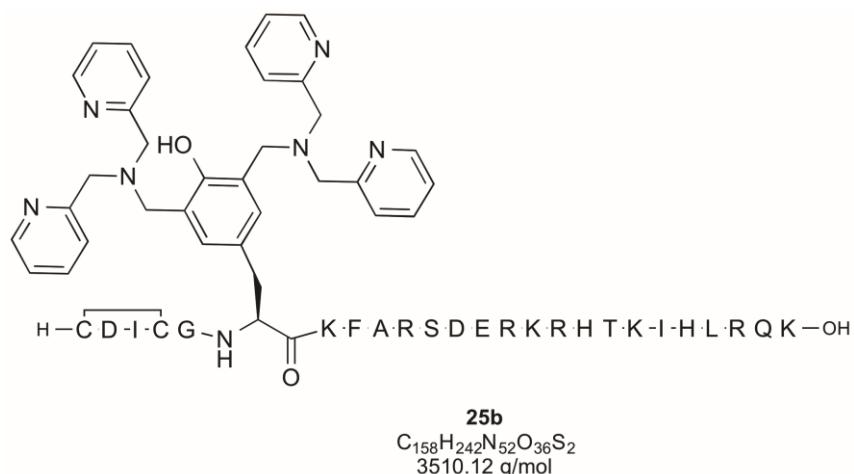
Zf3BPA70 (**25b**) was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol. The first part of the sequence until the lysine residue at position 71 was synthesized by means of an automated peptide synthesizer (Liberty 12, CEM) using protocol SPPS-A1 (section 6.3.1). The building block **16** and the subsequent amino acid were coupled manually according to protocol SPPS-M2. The remaining amino acid residues were also coupled manually according to conventional protocol SPPS-M1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **25b** (24 mg, 0.68 mmol, 16.3%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 17.12 min.

ESI-MS *m/z*: 585.6 [M+6H]$^{6+}$, 702.6 [M+5H]$^{5+}$, 877.9 [M+4H]$^{4+}$.

ESI-HRMS: calculated for $[C_{158}H_{248}N_{52}O_{36}S_2]^{6+}$ ([M+6H]$^{6+}$) = 585.6430, found = 585.6428, calculated for $[C_{158}H_{247}N_{52}O_{36}S_2]^{5+}$ ([M+5H]$^{5+}$) = 702.5702, found = 702.5702.

Zf3BMIA75

26a
$C_{157}H_{253}N_{59}O_{35}S_2$
3591.25 g/mol
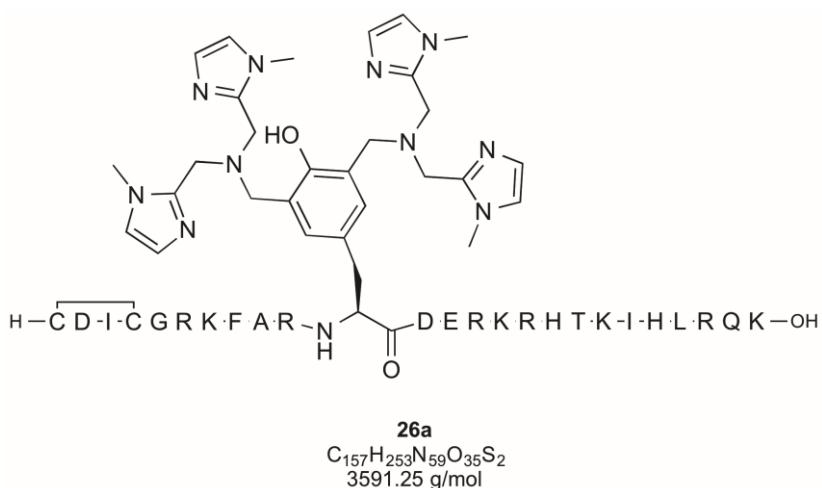
Zf3BMIA75 (**26a**) was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol. The first part of the sequence until the aspartic acid residue at position 76 was synthesized by means of an automated peptide synthesizer (Liberty 12, CEM) using protocol SPPS-A1 (section 6.3.1). The building block **14** and the subsequent amino acid were coupled manually according to protocol SPPS-M2. The remaining amino acid residues were also coupled manually according to conventional protocol SPPS-M1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **26a** (35 mg, 9.74 mmol, 23.4%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 17.53 min.

ESI-MS *m/z*: 599.2 [M+6H]$^{6+}$, 718.8 [M+5H]$^{5+}$, 898.2 [M+4H]$^{4+}$.

ESI-HRMS: calculated for $[C_{157}H_{259}N_{59}O_{35}S_2]^{6+}$ ([M+6H]$^{6+}$) = 599.1618, found = 599.1621, calculated for $[C_{157}H_{258}N_{59}O_{35}S_2]^{5+}$ ([M+5H]$^{5+}$) = 718.7927, found = 718.7929.

Zf3BPA75

26b
$C_{161}H_{249}N_{55}O_{35}S_2$
3579.23 g/mol
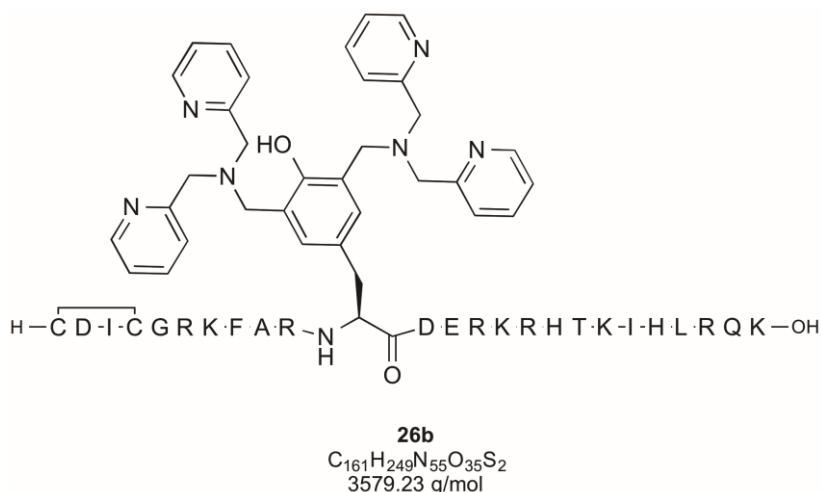
Zf3BPA75 (**26b**) was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol. The first part of the sequence until the aspartic acid residue at position 76 was synthesized by means of an automated peptide synthesizer (Liberty 12, CEM) using protocol SPPS-A1 (section 6.3.1). The building block **16** and the subsequent amino acid were coupled manually according to protocol SPPS-M2. The remaining amino acid residues were also coupled manually according to conventional protocol SPPS-M1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **26b** (29 mg, 0.81 mmol, 15.2%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 17.25 min.

ESI-MS *m/z*: 716.4 [M+5H]$^{5+}$, 895.2 [M+4H]$^{4+}$.

ESI-HRMS: calculated for [$C_{161}H_{254}N_{55}O_{35}S_2$]$^{5+}$ ([M+5H]$^{5+}$) = 716.3840, found = 716.3841, calculated for [$C_{161}H_{253}N_{55}O_{35}S_2$]$^{4+}$ ([M+4H]$^{4+}$) = 895.2282, found = 895.2281.

6.6 Protein expression of Zf12

Zf12 thioester

H—A S E R P Y A C P V E S C D R R F S R S D E R T R H·I·R·I·H T G⌐

⌐Q K P F Q C R·I·C N R M F S R S D H L T·T·H·I·R T H T T G K P F—N(H)—...—S—...—SO$_3$H

27
$C_{328}H_{520}N_{110}O_{98}S_7$
7796.86 g/mol
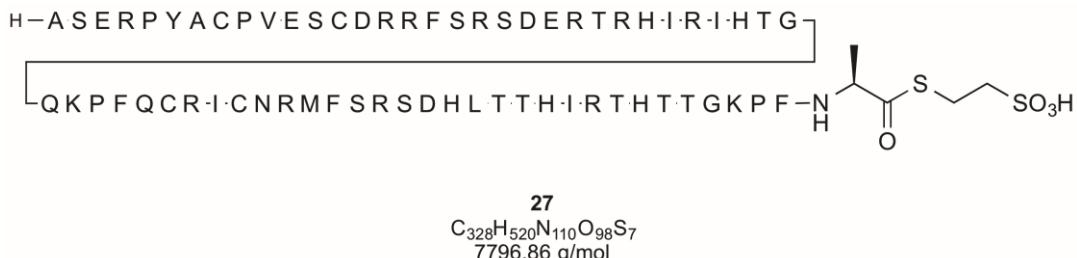
The Zf12-thioester domain **27** was expressed in *E. coli* and purified according to the method described in section 6.2.

HPLC (RP-C18, preparative, A (99.9% H$_2$O, 0.10% TFA), B (79.9% ACN, 20.0% H$_2$O, 0.10% TFA), gradient 20-50% B, 30 min, 10 mL/min): t_R = 23.24 min.

ESI-MS *m/z*: 600.4 [M+13H]$^{13+}$, 650.2 [M+12H]$^{12+}$, 709.2 [M+11H]$^{11+}$, 780.1 [M+10H]$^{10+}$, 866.5 [M+9H]$^{9+}$, 974.8 [M+8H]$^{8+}$, 1114.1 [M+7H]$^{7+}$, 1299.8 [M+6H]$^{6+}$.

ESI-HRMS: calculated for [C$_{328}$H$_{526}$N$_{110}$O$_{98}$S$_7$]$^{10+}$ ([M+10H]$^{10+}$) = 780.0762, found = 780.0769, calculated for [C$_{328}$H$_{525}$N$_{110}$O$_{98}$S$_7$]$^{9+}$ ([M+9H]$^{9+}$) = 866.5281, found = 866.5286, calculated for [C$_{328}$H$_{524}$N$_{110}$O$_{98}$S$_7$]$^{8+}$ ([M+8H]$^{8+}$) = 974.8434, found = 974.8431.

6.7 Preparation of Zf13 domains by means of NCL

Zf13 Wildtype

28
$C_{455}H_{735}N_{159}O_{130}S_7$
10737.29 g/mol
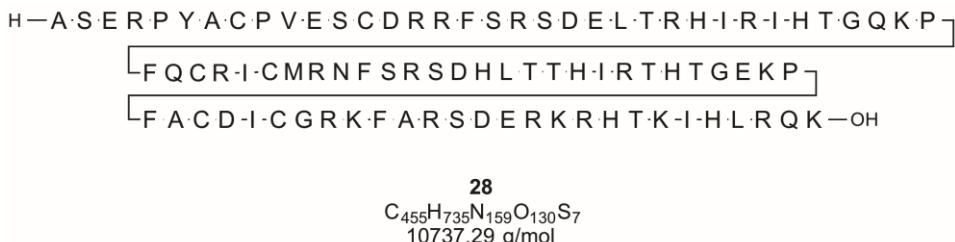
The Zf13 wildtype domain **28** was prepared by NCL between the Zf3 wildtype domain **24** and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 1.32 mg (0.13 μmol, 24%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-50% B, 30 min, 1 mL/min): t_R = 21.14 min.

ESI-MS *m/z*: 716.5 [M+15H]$^{15+}$, 767.6 [M+14H]$^{14+}$, 826.5 [M+13H]$^{13+}$.

ESI-HRMS: calculated for $[C_{455}H_{746}N_{159}O_{130}S_7]^{15+}$ ([M+15H]$^{15+}$) = 716.1643, found = 716.1641, $[C_{455}H_{745}N_{159}O_{130}S_7]^{14+}$ ([M+14H]$^{14+}$) = 767.2469, found = 767.2474.

Zf13BMIA70

29
$C_{480}H_{762}N_{166}O_{131}S_7$
11178.83
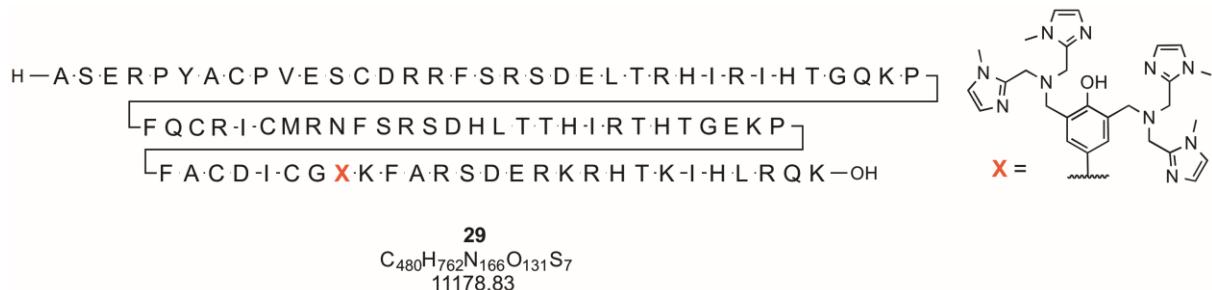
Peptide **29** was prepared by NCL between Zf3BMIA70 (**25a**) and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 0.89 mg (0.08 μmol, 11.5%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-60% B, 30 min, 1 mL/min): t_R = 22.43 min.

ESI-MS *m/z*: 745.9 $[M+15H]^{15+}$, 799.2 $[M+14H]^{14+}$, 860.4 $[M+13H]^{13+}$.

ESI-HRMS: calculated for $[C_{480}H_{762}N_{166}O_{131}S_7]^{15+}$ ($[M+15H]^{15+}$) = 745.7814, found = 745.7821, $[C_{480}H_{762}N_{166}O_{131}S_7]^{14+}$ ($[M+14H]^{14+}$) = 798.9795, found = 798.9798.

Zf13BPA70

30
$C_{484}H_{758}N_{162}O_{131}S_7$
11166.81
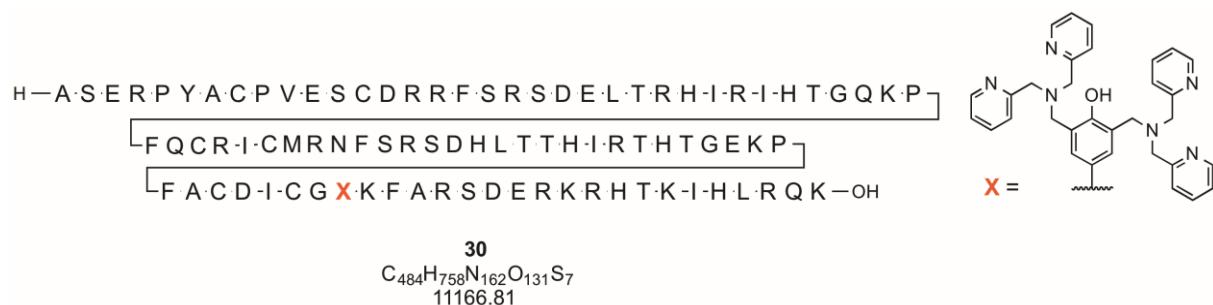
Peptide **30** was prepared by NCL between Zf3BPA70 (**25b**) and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 1.07 mg (0.09 µmol, 14.2%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-60% B, 30 min, 1 mL/min): t_R = 21.49 min.

ESI-MS *m/z*: 745.4 $[M+15H]^{15+}$, 798.3 $[M+14H]^{14+}$, 859.5 $[M+13H]^{13+}$.

ESI-HRMS: calculated for $[C_{484}H_{758}N_{162}O_{131}S_7]^{15+}$ ($[M+15H]^{15+}$) = 744.9785, found = 744.9789, $[C_{484}H_{758}N_{162}O_{131}S_7]^{14+}$ ($[M+14H]^{14+}$) = 798.1193, found = 798.1199.

Zf13BMIA75

31
$C_{483}H_{769}N_{169}O_{130}S_7$
11247.94
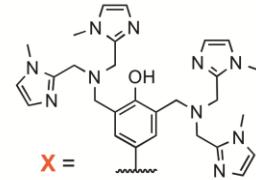
Peptide **31** was prepared by NCL between Zf3BMIA75 (**26a**) and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 1.16 mg (0.10 μmol, 17.8%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-60% B, 30 min, 1 mL/min): t_R = 21.33 min.

ESI-MS *m/z*: 750.4 $[M+15H]^{15+}$, 803.9 $[M+14H]^{14+}$, 865.7 $[M+13H]^{13+}$.

ESI-HRMS: calculated for $[C_{483}H_{769}N_{169}O_{130}S_7]^{15+}$ ($[M+15H]^{15+}$) = 750.3860, found = 750.3866, $[C_{483}H_{769}N_{169}O_{130}S_7]^{14+}$ ($[M+14H]^{14+}$) = 803.9130, found = 803.9132.

Zf13BPA75

32
$C_{487}H_{765}N_{165}O_{180}S_7$
11235.92
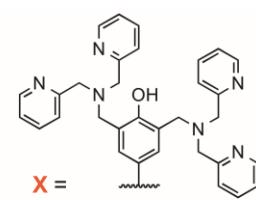
Peptide **32** was prepared by NCL between Zf3BPA75 (**26b**) and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 0.97 mg (0.09 μmol, 12.1%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-60% B, 30 min, 1 mL/min): t_R = 20.53 min.

ESI-MS *m/z*: 749.6 [M+15H]$^{15+}$, 803.1 [M+14H]$^{14+}$, 864.7 [M+13H]$^{13+}$.

ESI-HRMS: calculated for $[C_{487}H_{765}N_{165}O_{130}S_7]^{15+}$ ([M+15H]$^{15+}$) = 749.5831, found = 749.5840, $[C_{487}H_{765}N_{165}O_{130}S_7]^{14+}$ ([M+14H]$^{14+}$) = 803.0528, found = 803.0539.

6.8 Phosphoserine modified zinc finger domains

Fmoc-*O*-benzyl-L-phosphoserine
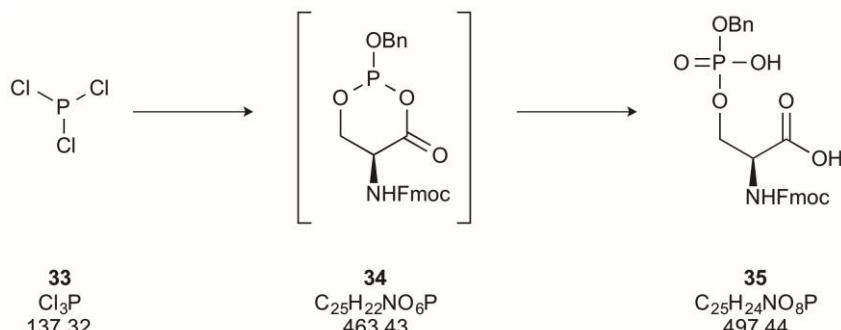
PCl$_3$ (0.70 mL, 8.00 mmol, 1.3 eq) was dissolved in THF (15.0 mL) and cooled to 0 °C. BnOH (0.96 mL, 9.20 mmol, 1.5 eq) was added slowly to keep the temperature of the solution below 5 °C. The reaction mixture was allowed to stir for 60 min at 0 °C. 2,6-Lutidine (2.12 mL, 18.3 mmol, 3.00 eq) was slowly added to keep the internal temperature between -5 to 5 °C. During rigorous stirring, a thick slurry formed.

2,6-Lutidine (0.72 mL, 6.20 mmol, 1.0 eq) was added to a solution of Fmoc-L-serine-OH (2.00 g, 6.10 mmol, 1.0 eq) in THF (10.0 mL). The solution was subsequently added to the slurry while the temperature was kept at 0 °C. The reaction mixture was allowed to stirred at 5 °C for 12 h. The reaction was quenched by the addition of H$_2$O (8.0 mL) at a rate that kept the temperature below 10 °C. A biphasic mixture was formed which was continuously stirred while NaBr (1.46 g, 14.2 mmol, 2.3 eq) was added at 0 °C. Subsequently, an aqueous solution of NaBrO$_3$ (0.44 g, 3.0 mmol, 0.48 eq) was added to the solution at 0 °C. The reaction mixture was allowed to warm to room temperature and stirring was continued for 2 h. An aqueous solution of Na$_2$S$_2$O$_5$ (10 wt% in H$_2$O, 2.0 mL) was quickly added and stirring was continued for 1 h.

2-MeTHF (20 mL) was added to the biphasic mixture and the layers were separated. The organic layer was washed with brine (2 x 25 mL) and dried over MgSO$_4$. The solvent was removed under reduced pressure to yield the crude product as a brown oil. The latter was diluted with 2-MeTHF (7 mL/g) and placed in a freezer at –22 °C over night. A white participate formed, which was filtrated and recrystallized from 2-MeTHF to give product **35** (2.52 g, 5.07 mmol, 83%) as a white solid.

¹H-NMR (300 MHz, [D$_6$]-DMSO): δ = 4.08 – 4.48 (m, 6H), 4.97 (d, J_{HH} = 7.1 Hz, 2H), 7.27 –
7.46 (m, 8H), 7.73 (d, J_{HH} = 7.4 Hz, 2H), 7.79 (d, J_{HH} = 8.2 Hz, 1H), 7.87 (d, J_{HH} = 7.5 Hz,
2H) ppm.

¹³C-NMR (300 MHz, [D$_6$]-DMSO): δ = 46.61, 54.44 (d, J = 7.9 Hz), 65.33 (d, J = 7.6 Hz),
65.98, 67.56 (d, J = 7.6 Hz), 120.04, 125.27, 127.06, 127.56, 127.62, 127.99, 128.33, 136.76
(d, J = 7.6 Hz), 140.70, 143.75, 156.00, 170.73 ppm.

³¹P-NMR (122 MHz, [D$_6$]-DMSO): δ = −1.48 ppm.

ESI-MS: 498.1 [M+H]$^+$, 520.1 [M+Na]$^+$.

ESI-HRMS: calculated for [C$_{25}$H$_{24}$NO$_8$P]$^+$ ([M+H]$^+$) = 498.1312, found = 498.1311,
calculated for [C$_{25}$H$_{24}$NO$_8$PNa]$^+$ ([M+Na]$^+$) = 520.1132, found = 520.1127.

Zf3Pser70

36
C$_{126}$H$_{213}$N$_{46}$O$_{39}$PS$_2$
3091.45 g/mol
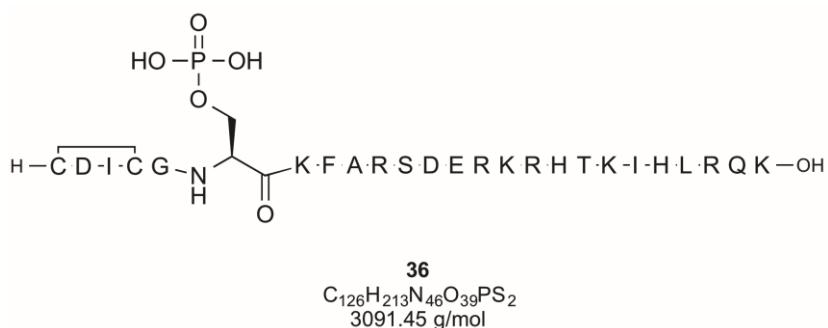
Peptide **36** was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at
a scale of 0.05 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Standard
amino acids were coupled according to protocol SPPS-A2 described in section 6.3.1.
Coupling of the Fmoc-O-benzyl-L-phosphoserine (**35**) building block as well as of the
subsequent amino acid was performed according to protocol SPPS-A3. Cleavage of the
peptide from the resin was achieved by the addition of a cleavage cocktail containing
TFA/H$_2$O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-
HPLC and subsequently lyophilized. Peptide **36** (103 mg, 33.3 mmol, 41.1%) was obtained
as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 15.38 min.

ESI-MS *m/z*: 516.4 [M+6H]$^{6+}$, 619.5 [M+5H]$^{5+}$, 774.2 [M+4H]$^{4+}$.

ESI-HRMS: calculated for [C$_{126}$H$_{215}$N$_{46}$O$_{39}$PS$_2$]$^{6+}$ ([M+6H]$^{6+}$) = 516.2645, found = 516.2650, calculated for [C$_{126}$H$_{215}$N$_{46}$O$_{39}$PS$_2$]$^{5+}$ ([M+5H]$^{5+}$) = 619.3159, found = 619.3163, calculated for [C$_{126}$H$_{215}$N$_{46}$O$_{39}$PS$_2$]$^{4+}$ ([M+4H]$^{4+}$) = 773.8931, found = 773.8933.

Zf3Pser75
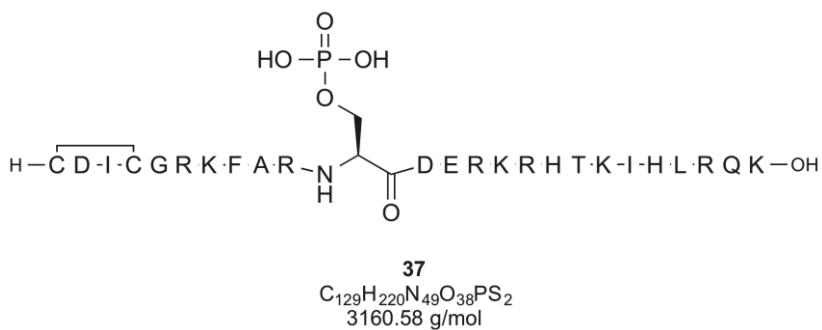
37
C$_{129}$H$_{220}$N$_{49}$O$_{38}$PS$_2$
3160.58 g/mol

Peptide **37** was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Standard amino acids were coupled according to protocol SPPS-A2 described in section 6.3.1. Coupling of the Fmoc-*O*-benzyl-L-phosphoserine (**35**) building block as well as of the subsequent amino acid was performed according to protocol SPPS-A3. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **37** (98.0 mg, 31.6 mmol, 27.2%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 14.42 min.

ESI-MS *m/z*: 527.6 [M+6H]$^{6+}$, 632.9 [M+5H]$^{5+}$, 790.9 [M+4H]$^{4+}$.

ESI-HRMS: calculated for [C$_{129}$H$_{220}$N$_{49}$O$_{38}$PS$_2$]$^{6+}$ ([M+6H]$^{6+}$) = 527.4401, found = 527.4414,

calculated for $[C_{129}H_{220}N_{49}O_{38}PS_2]^{5+}$ ($[M+5H]^{5+}$) = 632.7266, found = 632.7273, calculated for $[C_{129}H_{220}N_{49}O_{38}PS_2]^{4+}$ ($[M+4H]^{4+}$) = 790.6565, found = 790.6570.

Ser$_{PEG}$FITC

38
$C_{30}H_{30}N_4O_{10}S$
638.65
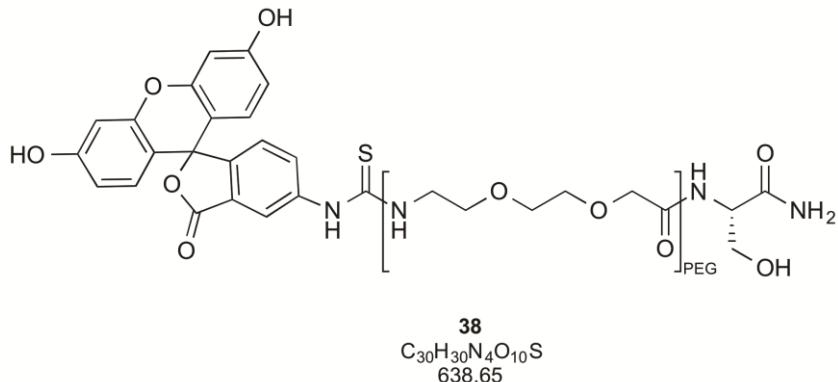
Compound **38** was synthesized by the means of solid phase organic synthesis on a Rink amide MBHA resin (0.18 mmol/g) at a scale of 0.05 mmol. Coupling was performed manually with microwave support (CEM) according to the protocol SPPS-M1 described in section 6.3.1. Cleavage of the compound from the resin was achieved by the addition of a cleavage cocktail containing TFA/H$_2$O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Compound **38** (21.0 mg, 0.32 mmol, 72%) was obtained as a white solid.

HPLC (RP-C18, semi-preparative, A (99.9% H$_2$O, 0.10% TFA), B (79.9% ACN, 20.0% H$_2$O, 0.10% TFA), gradient 10-70% B, 30 min, 3 mL/min): t_R = 24.38 min.

ESI-MS *m/z*: 639.2 [M+H]$^+$, 661.2 [M+Na]$^+$.

ESI-HRMS: calculated for $[C_{30}H_{30}N_4O_{10}S]^+$ ($[M+H]^+$) = 639.1755, found = 639.1749, calculated for $[C_{30}H_{30}N_4O_{10}SNa]^+$ ($[M+Na]^+$) = 661.1575, found = 661.1565.

Pser$_{PEG}$FITC

39
C$_{30}$H$_{31}$N$_4$O$_{13}$PS
718.63
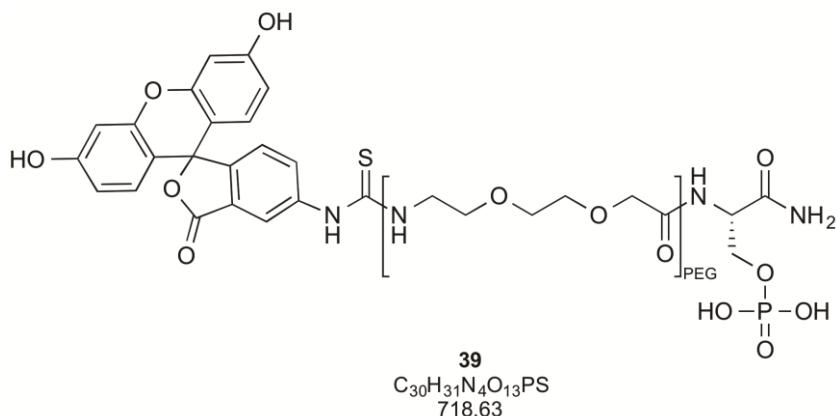
Compound **39** was synthesized by the means of solid phase organic synthesis on a Rink amide MBHA resin (0.18 mmol/g) at a scale of 0.05 mmol. Coupling was performed manually with microwave support (CEM) according to the protocol SPPS-M1 described in section 6.3.1. Cleavage of the compound from the resin was achieved by the addition of a cleavage cocktail containing TFA/H$_2$O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Compound **39** (18.3 mg, 0.25 mmol, 64%) was obtained as a white solid.

HPLC (RP-C18, semi-preparative, A (99.9% H$_2$O, 0.10% TFA), B (79.9% ACN, 20.0% H$_2$O, 0.10% TFA), gradient 10-70% B, 30 min, 3 mL/min): t_R = 25.31 min.

ESI-MS *m/z*: 717.1 [M-H]$^-$, 358.1 [M-2H]$^{-2-}$.

ESI-HRMS: calculated for [C$_{30}$H$_{31}$N$_4$O$_{13}$PS]$^-$ ([M-H]$^-$) = 717.1273, found = 717.1272, calculated for [C$_{30}$H$_{30}$N$_4$O$_{10}$S]$^{-2}$ ([M-2H]$^{-2}$) = 358.0600, found = 358.0604.

Zf3FITC wildtype

40
$C_{156}H_{245}N_{51}O_{43}S_3$
3617.15
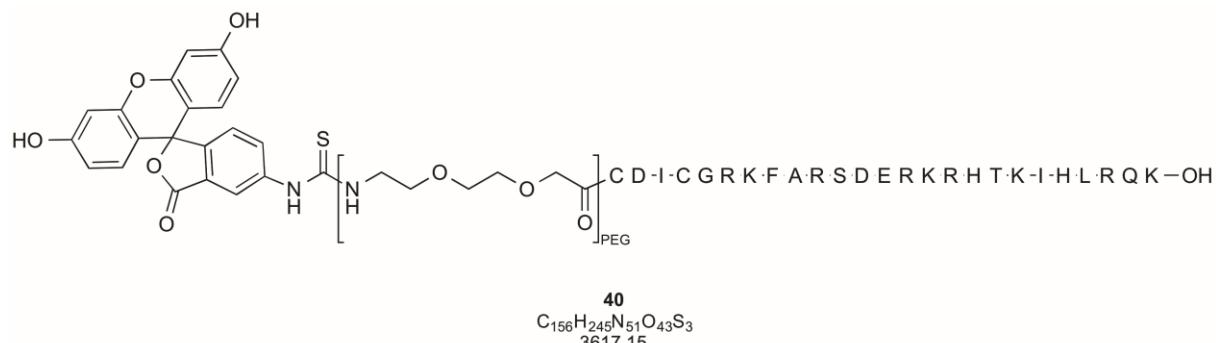
Peptide **40** was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Standard amino acids were coupled according to protocol SPPS-A2 described in section 6.3.1. Coupling of the N-terminal PEG spacer (8-(Fmoc-amino)-3,6-dioxaoctanoic acid) and of the FITC fluorophore was performed manually according to protocol SPPS-M1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **40** (42.0 mg, 11.1 mmol, 22.2%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 18.45 min.

ESI-MS *m/z*: 517.7 $[M+7H]^{7+}$, 603.8 $[M+6H]^{6+}$, 724.4 $[M+5H]^{5+}$.

ESI-HRMS: calculated for $[C_{156}H_{243}N_{51}O_{43}S_3]^{7+}$ ($[M+7H]^{7+}$) = 517.4010, found = 517.4006, calculated for $[C_{156}H_{243}N_{51}O_{43}S_3]^{6+}$ ($[M+6H]^{6+}$) = 603.4666, found = 603.4661, calculated for $[C_{156}H_{243}N_{51}O_{43}S_3]^{5+}$ ($[M+5H]^{5+}$) = 723.9584, found = 723.9582.

Zf3Pser75NBD

41
$C_{135}H_{221}N_{52}O_{41}PS_2$
3323.67
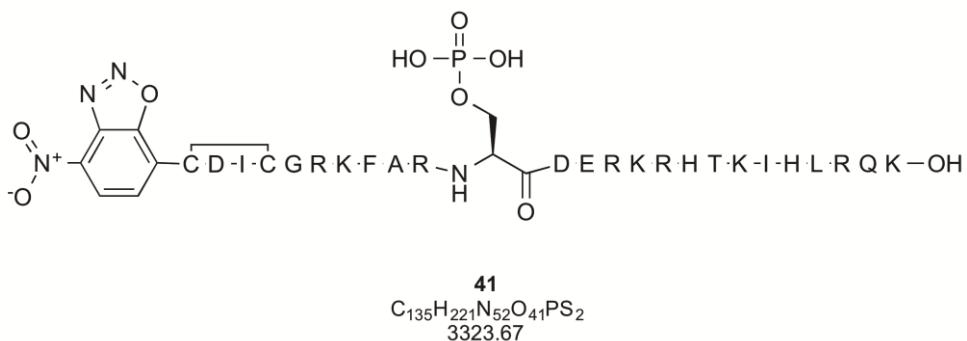
Peptide **41** was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Standard amino acids were coupled according to protocol SPPS-A2 described in section 6.3.1. Coupling of the Fmoc-O-benzyl-L-phosphoserine (**35**) building block as well as of the subsequent amino acid was performed according to protocol SPPS-A3. Coupling of the N-terminal NBD fluorophore was achieved manually. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **41** (12 mg, 0.04 mmol, 8.2%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 21.45 min.

ESI-MS *m/z*: 554.8 [M+6H]$^{6+}$, 665.5 [M+5H]$^{5+}$, 831.6 [M+4H]$^{4+}$.

ESI-HRMS: calculated for [$C_{135}H_{221}N_{52}O_{41}PS_2$]$^{6+}$ ([M+6H]$^{6+}$) = 554.6070, found = 554.6070, calculated for [$C_{135}H_{221}N_{52}O_{41}PS_2$]$^{5+}$ ([M+5H]$^{5+}$) = 665.3270, found = 665.3274, calculated for [$C_{135}H_{221}N_{52}O_{41}PS_2$]$^{4+}$ ([M+4H]$^{4+}$) = 831.4069, found = 831.4073.

Zf3Pser70FITC

42
$C_{153}H_{235}N_{48}O_{47}PS_3$
3626.02 g/mol
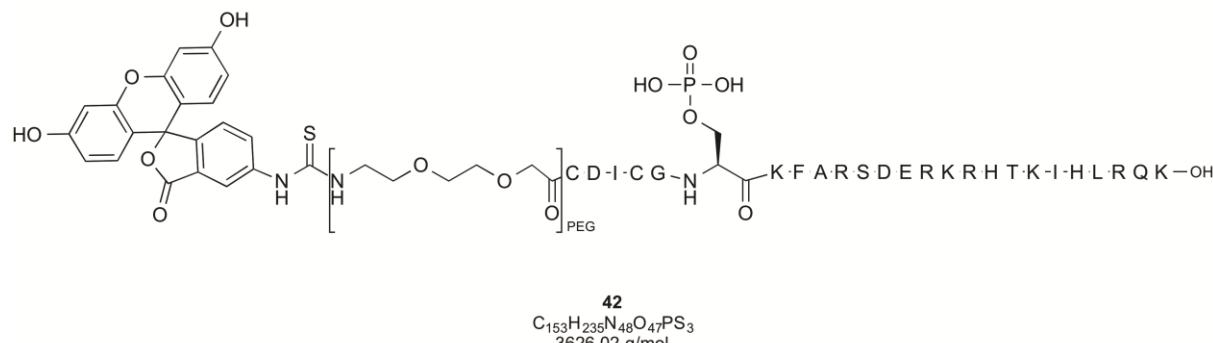
Peptide **42** was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Standard amino acids were coupled according to protocol SPPS-A2 described in section 6.3.1. The building block Fmoc-O-benzyl-L-phosphoserine (**35**) as well as the subsequent amino acid were coupled according to protocol SPPS-A3. Coupling of the N-terminal PEG spacer (8-(Fmoc-amino)-3,6-dioxaoctanoic acid) and of the FITC fluorophore was performed manually according to protocol SPPS-M1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **42** (23.0 mg, 0.63 mmol, 15.2%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 22.35 min.

ESI-MS *m/z*: 605.6 $[M+6H]^{6+}$, 726.5 $[M+5H]^{5+}$, 907.9 $[M+4H]^{4+}$.

ESI-HRMS: calculated for $[C_{153}H_{237}N_{48}O_{47}PS_3]^{6+}$ ($[M+6H]^{6+}$) = 605.2828, found = 605.2826, calculated for $[C_{153}H_{237}N_{48}O_{47}PS_3]^{5+}$ ($[M+5H]^{5+}$) = 726.1379, found = 726.1384, calculated for $[C_{153}H_{237}N_{48}O_{47}PS_3]^{4+}$ ($[M+4H]^{4+}$) = 907.4205, found = 907.4214.

Zf13Pser70

43
$C_{452}H_{729}N_{156}O_{134}PS_7$
10748.16 g/mol
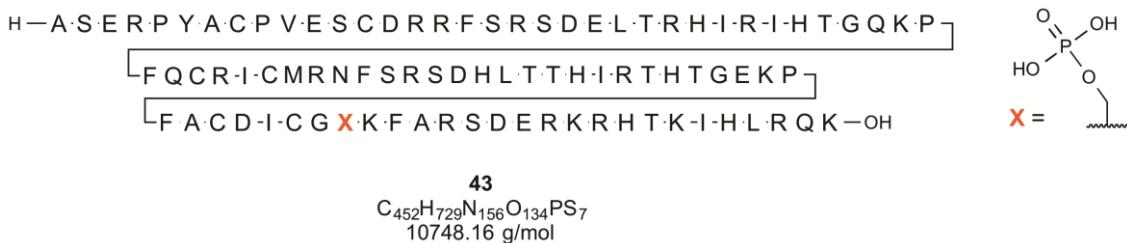
Peptide **43** was prepared by NCL between Zf3Pser70 (**36**) and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 1.18 mg (0.17 μmol, 22%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-50% B, 30 min, 1 mL/min): t_R = 25.35 min.

ESI-MS *m/z*: 717.5 $[M+15H]^{15+}$, 768.4 $[M+14H]^{14+}$, 827.4 $[M+13H]^{13+}$.

ESI-HRMS: calculated for $[C_{452}H_{729}N_{156}O_{134}PS_7]^{15+}$ ($[M+15H]^{15+}$) = 717.0927, found = 717.0936, $[C_{452}H_{729}N_{156}O_{134}PS_7]^{14+}$ ($[M+14H]^{14+}$) = 768.2416, found = 768.2522.

Zf13Pser75

44
$C_{455}H_{736}N_{159}O_{133}PS_7$
10817.27 g/mol
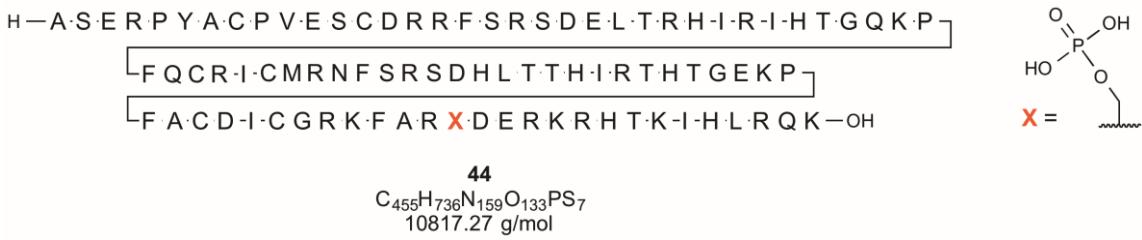
Peptide **44** was prepared by NCL between Zf3Pser75 (**37**) and the expressed Zf12-thioester domain **27** according to the method described in section 6.2.

Yield: 0.94 mg (0.09 μmol, 12%)

HPLC (RP-C18, analytical, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 20-50% B, 30 min, 1 mL/min): t_R = 25.12 min.

ESI-MS *m/z*: 721.8 [M+15H]$^{15+}$, 773.2 [M+14H]$^{14+}$, 832.6 [M+13H]$^{13+}$.

ESI-HRMS: calculated for [C$_{455}$H$_{736}$N$_{159}$O$_{133}$PS$_7$]$^{15+}$ ([M+15H]$^{15+}$) = 721.6973, found = 721.6981, [C$_{455}$H$_{736}$N$_{159}$O$_{133}$PS$_7$]$^{14+}$ ([M+14H]$^{14+}$) = 773.1751, found = 773.1756.

6.9 Development of zinc-finger-based peptidyl metal sensors

N-(2-Azidoethyl)-5-(dimethylamino)naphthalene-1-sulfonamide

45
C_2H_6BrN
123.98 g/mol

46
$C_2H_6N_4$
86.10 g/mol

47
$C_{12}H_{12}ClNO_2S$
269.74 g/mol

48
$C_{14}H_{17}N_5O_2S$
319.38 g/mol
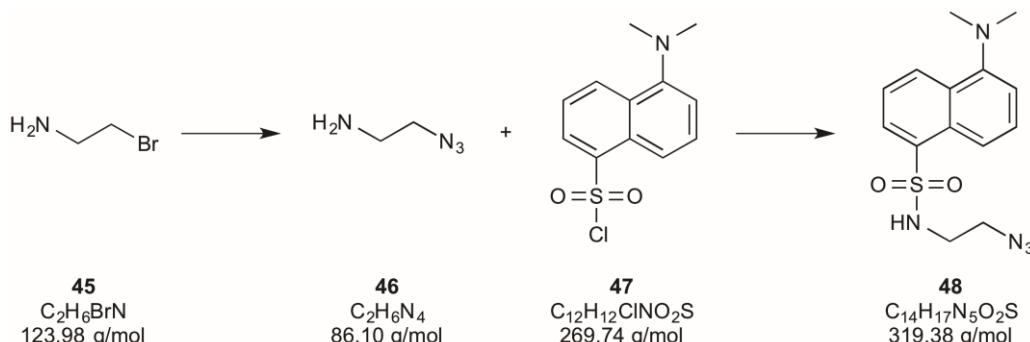
Sodium azide (0.51 g, 7.83 mmol, 2.11 eq) was added to a solution of 2-bromoethylamine hydrobromide (0.76 g, 3.71 mmol, 1.00 eq) in water (6.3 mL) and the reaction mixture was allowed to stir at 75 °C for 20 h. An aqueous solution of NaOH (1 M, 1.1 mL) was added and the organic compounds were extracted with chloroform (3 x 25 mL). The combined organic extracts were washed with brine (2 x 30 mL), dried over $MgSO_4$ and concentrated *in vacuo*. Compound **46** (0.32 g, 3.68 mmol, 1.10 eq) was dissolved in DCM (8 mL) and dansyl chloride (0.90 g, 3.35 mmol, 1.00 eq) was added. The mixture was allowed to stir at room temperature for 2 h. The solvent was removed under reduced pressure and the crude product was purified by flash column chromatography on silica gel (EtOAc/DCM, 1:8). Compound **48** (0.91 g, 2.85 mmol, 85%) was obtained as a yellow oil.

^{1}H-NMR (300 MHz, CDCl$_3$): δ = 2.87 (s, 6H), 3.04 (dt, $^3J_{HH}$ = 6.2 Hz, 5.8 Hz, 2H), 3.29 (t, $^3J_{HH}$ = 5.8 Hz, 2H), 5.04 (t, $^3J_{HH}$ = 6.2 Hz, 1H), 7.18 (d, $^3J_{HH}$ = 7.4 Hz, 1H), 7.51 (dd, $^3J_{HH}$ = 8.5 Hz, 7.4 Hz, 1H), 7.57 (dd, $^3J_{HH}$ = 8.5 Hz, 7.7 Hz, 1H), 8.27-8.20 (m, 2H), 8.54 (d, $^3J_{HH}$ = 8.5 Hz, 1H) ppm.

ESI-MS: 320.1 [M+H]$^+$, 342.1 [M+Na]$^+$.

ESI-HRMS: calculated for [C14H18N5O2S]$^+$ ([M+H]$^+$) = 320.1176, found = 342.0988,
calculated for [C14H17N5O2SNa]$^+$ ([M+Na]$^+$) = 342.0995, found = 342.0988.

Zf3 Wildtype

H−F A C D·I·C G R K F A R S D E R K R H T K·I·H L R Q K−OH

56

$C_{141}H_{233}N_{51}O_{37}S_2$

3298.86 g/mol

The native sequence of Zf3 was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.1 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Coupling was performed according to the Fmoc/*tert*-butyl protocol SPPS-A2 described in section 6.3.1. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail containing TFA/H_2O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **56** (148 mg, 4.48 mmol, 74.1%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 10-60% B, 30 min, 3 mL/min): t_R = 18.03 min.

ESI-MS *m/z*: 551.0 $[M+6H]^{6+}$, 661.0 $[M+5H]^{5+}$, 825.9 $[M+4H]^{4+}$.

ESI-HRMS: calculated for $[C_{141}H_{233}N_{51}O_{37}S_2]^{6+}$ ($[M+6H]^{6+}$) = 550.7992, found = 550.7993, calculated for $[C_{141}H_{233}N_{51}O_{37}S_2]^{5+}$ ($[M+5H]^{5+}$) = 660.7576, found = 660.7575.

Zf3Pg85

H−F A C D·I·C G R K F A R S D E R K R H T K·I−N...L R Q K−OH

57

$C_{140}H_{231}N_{49}O_{37}S_2$

3256.81 g/mol
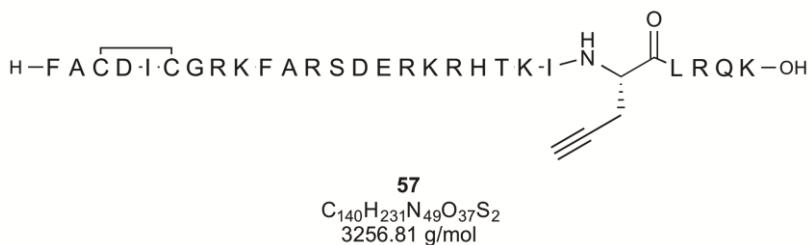
Peptide **57** was synthesized on a pre-loaded Fmoc-L-Lys(Boc)-Wang resin (0.34 mmol/g) at a scale of 0.05 mmol using an automated peptide synthesizer (Liberty Blue, CEM). Coupling was performed according to the Fmoc/*tert*-butyl protocol SPPS-A2 described in section 6.3.1. The histidine residue at position 85 was exchanged for Fmoc-L-propargylglycine-OH. Cleavage of the peptide from the resin was achieved by the addition of a cleavage cocktail

containing TFA/H$_2$O/EDT/TIS (94/2.5/2.5/1) and incubation for 2 h. The peptide was purified by RP-HPLC and subsequently lyophilized. Peptide **57** (31 mg, 0.95 mmol, 29.8%) was obtained as a white solid with a disulfide bond between the cysteine residues.

HPLC (RP-C18, semi-preparative, A (99.9% H$_2$O, 0.10% TFA), B (79.9% ACN, 20.0% H$_2$O, 0.10% TFA), gradient 0-60% B, 30 min, 3 mL/min): t_R = 23.41 min.

ESI-MS *m/z*: 652.6 [M+5H]$^{5+}$, 815.4 [M+4H]$^{4+}$, 1087.3 [M+3H]$^{3+}$.

ESI-HRMS: calculated for [C$_{140}$H$_{233}$N$_{49}$O$_{37}$S$_2$]$^{5+}$ ([M+5H]$^{5+}$) = 652.3532, found = 652.3544, calculated for [C$_{140}$H$_{233}$N$_{49}$O$_{37}$S$_2$]$^{4+}$ ([M+4H]$^{4+}$) = 815.1897, found = 815.1896.

6.9.1 Fluorophore attachment by means of CuAAC

General procedure

The fully protected and resin bound peptide **57** (1 eq) was placed in a syringe (BD) equipped with a polyethylene frit. Anhydrous and argon flushed DMF (3 mL) was added and the resin was swollen for 2 h. The azide modified fluorophore (2 eq), CuI (1.5 eq) and sodium ascorbate (1.5 eq) were dissolved in DMF (2 mL) and subsequently added to the syringe which was subsequently agitated at room temperature for 2 h. The solvent was removed and the resin was excessively washed with NMP (5 x 3 mL), DMF (5 x 3 mL) and DCM (5 x 3 mL). The N-terminal Fmoc protecting group was cleaved off by the addition of piperidine (20% in DMF, 5 mL, 2 x 15 min) and the resin was repeatedly washed with NMP (5 x 3 mL), DMF (5 x 3 mL) and DCM (5 x 3 mL). The resin was dried under reduced pressure overnight before the peptide was liberated by the addition of a cleavage cocktail containing TFA/EDT/H$_2$O/TIS (94:2.5:2.5:1) and incubation for 90 min. The solution was concentrated in a nitrogen stream and the crude peptide was precipitated by the addition of ice-cold diethyl ether. After centrifugation (9.000 rpm, 15 min), the supernatant was removed, the crude peptide was dissolved in ACN/H$_2$O (1:4) and purified by RP-HPLC (RP-C18, preparative) with subsequent lyophilization.

The peptide was placed in a glass vial and the MeOBzl protecting groups were removed by the addition of a cleavage cocktail containing TFMSA/TFA/EDT/thioanisole (10:86:2:2) and incubation for 2.5 h. The solution was concentrated and the crude peptide was precipitated by the addition of ice-cold diethyl ether and subsequently centrifuged (9.000 rpm, 15 min). The supernatant was removed, the peptide was dissolved in ACN/H$_2$O (1:4) and purified by RP-HPLC (RP-C18, semi-preparative) to yield the fluorescently labeled Zf3 domains.

Zf3DNS

58
$C_{154}H_{248}N_{54}O_{39}S_3$
3576.20 g/mol
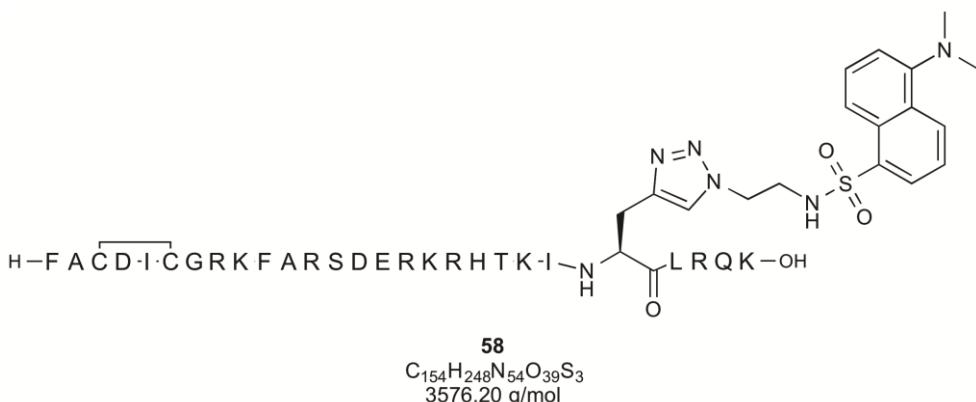
HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 0-70% B, 30 min, 3 mL/min): t_R = 20.36 min.

ESI-MS *m/z*: 596.7 [M+6H]$^{6+}$, 715.8 [M+5H]$^{5+}$, 894.7 [M+4H]$^{4+}$.

ESI-HRMS: calculated for $[C_{154}H_{248}N_{54}O_{39}S_3]^{5+}$ ([M+5H]$^{5+}$) = 715.7722, found = 715.7711, $[C_{154}H_{248}N_{54}O_{39}S_3]^{4+}$ ([M+4H]$^{4+}$) = 894.4634, found = 894.4624.

Zf3COUM

59
$C_{153}H_{245}N_{53}O_{39}S_2$
3515.10 g/mol
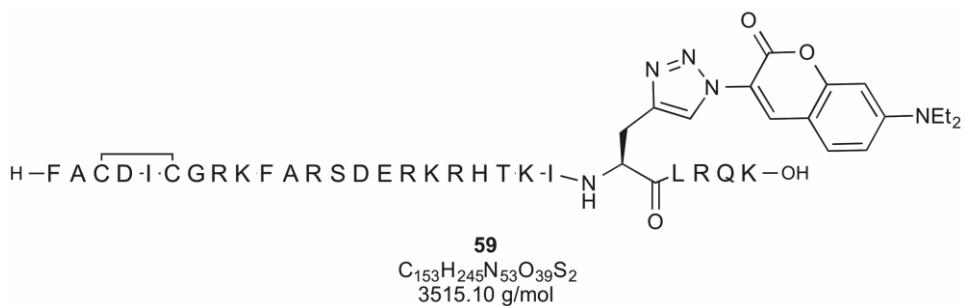
HPLC (RP-C18, semi-preparative, A (99.9% H_2O, 0.10% TFA), B (79.9% ACN, 20.0% H_2O, 0.10% TFA), gradient 15-70% B, 30 min, 3 mL/min): t_R = 20.23 min.

ESI-MS *m/z*: 586.5 [M+6H]$^{6+}$, 703.6 [M+5H]$^{5+}$, 879.5 [M+4H]$^{4+}$.

ESI-HRMS: calculated for $[C_{153}H_{245}N_{53}O_{39}S_2]^{5+}$ ([M+5H]$^{5+}$) = 703.5725, found = 703.5729.

Appendix

Derivation of Equation 5.3

$$K_d = \frac{[M] \cdot [L]}{[ML]}$$

$$[L] = c_L - [ML]$$

$$[M] = c_M - [ML]$$

with the total metal concentration c_M and the total ligand concentration c_L.

$$K_d = \frac{(c_M - [ML] \cdot (c_L - [ML])}{[ML]}$$

$$K_d[ML] = (c_M - [ML])(c_L - [ML])$$

$$K_d[ML] = c_L c_M - c_M[ML] - c_L[ML] + [ML]^2$$

$$[ML]^2 - (K_d + c_L + c_M)[ML] + c_M c_L = 0$$

Quadratic equation:

$$\Delta = (K_d + c_L + c_M)^2 - 4c_M c_L$$

$$[ML] = \frac{K_d c_L c_M - \sqrt{(K_d + c_L + c_M)^2 - 4c_M c_L}}{2}$$

$$[ML] = \frac{K_d c_L c_M + \sqrt{(K_d + c_L + c_M)^2 - 4c_M c_L}}{2}$$

Equation 5.3 is obtained after the insertion of the experimental absorbance values:

$$A = A_{min} - \left(\frac{A_{min} - A_{max}}{2c_L}\right)\left(K_d + c_L + c_M - \sqrt{(K_d + c_L + c_M)^2 - 4c_L c_M}\right)$$

Derivation of Equation 5.7

$$K_d^{M1} = \frac{[M1] \cdot [L]}{[M1L]} \qquad K_d^{M2} = \frac{[M2] \cdot [L]}{[M2L]}$$

$$[M_{Zn}L] = \frac{[M_{Zn}] \cdot [M_{Co}L]K_d^{Co}}{[M_{Co}]K_d^{Zn}}$$

Quadratic equation obtained after solution with respect to $M_{Zn}L$:

$$\left(K_d^{Co} - K_d^{Zn}\right)[M_{Zn}L]^2 - \left(c_{Zn}K_d^{Co} + c_L K_d^{Co} + c_{Co}K_d^{Zn} - c_L K_d^{Zn}\right)[M_{Zn}L] + c_{Zn}c_L K_d^{Co} = 0$$

$$[M_{Zn}L] = \left(\frac{1}{2(K_d^{Co} - K_d^{Zn})}\right)\left(c_{Zn}K_d^{Co} + c_L K_d^{Co} + c_{Co}K_d^{Zn} - c_L K_d^{Zn}\right.$$

$$\left. - \sqrt{c_{Zn}K_d^{Co} + c_L K_d^{Co} + c_{Co}K_d^{Zn} - c_L K_d^{Zn} - 4c_{Zn}c_L K_d^{Co}(K_d^{Co} - K_d^{Zn})}\right)$$

$$A = A_{min} - \left(\frac{A_{min} - A_{max}}{2 \cdot c_L \cdot (K_d^{Co} - K_d^{Zn})}\right)$$

$$= \left(\frac{1}{2(K_d^{Co} - K_d^{Zn})}\right)\left(c_{Zn}K_d^{Co} + c_L K_d^{Co} + c_{Co}K_d^{Zn} - c_L K_d^{Zn}\right.$$

$$\left. - \sqrt{c_{Zn}K_d^{Co} + c_L K_d^{Co} + c_{Co}K_d^{Zn} - c_L K_d^{Zn} - 4c_{Zn}c_L K_d^{Co}(K_d^{Co} - K_d^{Zn})}\right)$$

References

[1]   G. Thiel, G. Cibelli, *J. Cell. Physiol.* **2002**, *193*, 287–292.

[2]   M. Bibikova, K. Beumer, J. K. Trautman, D. Carroll, *Science* **2003**, *300*, 764–764.

[3]   R. M. Gupta, K. Musunuru, *J. Clin. Invest.* **2014**, *124*, 4154–4161.

[4]   F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang, P. D. Gregory, *Nat. Rev. Genet.* **2010**, *11*, 636–646.

[5]   C. Cavanagh, M. Morell, I. Mackay, W. Powell, *Curr. Opin. Plant Biol.* **2008**, *11*, 215–221.

[6]   E. L. Hegg, J. N. Burstyn, *Coord. Chem. Rev.* **1998**, *173*, 133–165.

[7]   J. Sumaoka, Y. Yamamoto, Y. Kitamura, M. Komiyama, *Curr. Org. Chem.* **2007**, *11*, 463–475.

[8]   W. Kameshima, T. Ishizuka, M. Minoshima, M. Yamamoto, H. Sugiyama, Y. Xu, M. Komiyama, *Angew. Chem.* **2013**, *125*, 13926–13929.

[9]   J. H. Laity, B. M. Lee, P. E. Wright, *Curr. Opin. Struct. Biol.* **2001**, *11*, 39–46.

[10] N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809–817.

[11] R. S. Brown, C. Sander, P. Argos, *FEBS Lett.* **1985**, *186*, 271–274.

[12] G. Parraga, S. J. Horvath, A. Eisen, W. E. Taylor, L. Hood, E. T. Young, R. E. Klevit, *Science* **1988**, *241*, 1489–1492.

[13] G. Mardon, S. W. Luoh, E. M. Simpson, G. Gill, L. G. Brown, D. C. Page, *Mol. Cell. Biol.* **1990**, *10*, 681–688.

[14] J. M. Berg, *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 405–421.

[15] B. Christy, D. Nathans, *Proc. Natl. Acad. Sci.* **1989**, *86*, 8737–8741.

[16] A. M. Rich, E. Bombarda, A. D. Schenk, P. E. Lee, E. H. Cox, A. M. Spuches, L. D. Hudson, B. Kieffer, D. E. Wilcox, *J. Am. Chem. Soc.* **2012**, *134*, 10405–10418.

[17] L. Nekludova, C. O. Pabo, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6948–6952.

[18] A. Klug, *Annu. Rev. Biochem.* **2010**, *79*, 213–231.

[19] S. A. Wolfe, L. Nekludova, C. O. Pabo, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 183–212.

[20] J. C. Miller, M. C. Holmes, J. Wang, D. Y. Guschin, Y.-L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, K. A. Kim, et al., *Nat. Biotechnol.* **2007**, *25*, 778–785.

[21] R. D. Johnson, M. Jasin, *Biochem. Soc. Trans.* **2001**, *29*, 196–201.

[22] E. E. Perez, J. Wang, J. C. Miller, Y. Jouvenot, K. A. Kim, O. Liu, N. Wang, G. Lee, V. V. Bartsevich, Y.-L. Lee, et al., *Nat. Biotechnol.* **2008**, *26*, 808–816.

[23] G. M. Cooper, The Cell: A Molecular Approach, 2nd Edition, Boston **2000**.

[24] C. A. Pritsos, *Chem. Biol. Interact.* **2000**, *129*, 195–208.

[25] D. Sheehan, G. Meade, V. M. Foley, C. A. Dowd, *Biochem. J.* **2001**, *360*, 1–16.

[26] E. Klussmann, W. Rosenthal, *Kidney Int.* **2001**, *60*, 446–449.

[27] A. J. Barrett, N. D. Rawlings, J. F. Woessner, *Handbook of Proteolytic Enzymes*, Academic Press, **2012**.

[28] A. J. Barrett, J. K. McDonald, *Biochem. J.* **1986**, *237*, 935–935.

[29] R. D. Schmid, R. Verger, *Angew. Chem. Int. Ed.* **1998**, *37*, 1608–1633.

[30] T. Nishino, K. Morikawa*, Oncogene* **2002**, *21*, 9022-9032.

[31] E. Willerslev, A. J. Hansen, H. N. Poinar, *Trends Ecol. Evol.* **2004**, *19*, 141–147.

[32] C. Liu, M. Wang, T. Zhang, H. Sun, *Coord. Chem. Rev.* **2004**, *248*, 147–168.

[33] A. Garen, C. Levinthal, *Biochim. Biophys. Acta* **1960**, *38*, 470–483.

[34] A. Volbeda, A. Lahm, F. Sakiyama, D. Suck, *EMBO J.* **1991**, *10*, 1607–1618.

[35] J. M. Falcone, H. C. Box, *Biochim. Biophys. Acta BBA-Protein Struct. Mol. Enzymol.* **1997**, *1337*, 267–275.

[36] A. J. Kirby, F. Hollfelder, *From Enzyme Models to Model Enzymes*, Royal Society Of Chemistry, **2009**.

[37] R. A. Anderson, W. F. Bosron, F. S. Kennedy, B. L. Vallee, *Proc. Natl. Acad. Sci.* **1975**, *72*, 2989–2993.

[38] A. Warshel, *J. Biol. Chem.* **1998**, *273*, 27035–27038.

[39] C. Liu, L. Wang, *Dalton Trans* **2009**, 227–239.

[40] C. M. Dupureur, *Metallomics* **2010**, *2*, 609.

[41] N. H. Williams, B. Takasaki, M. Wall, J. Chin, *Acc. Chem. Res.* **1999**, *32*, 485–493.

[42] C. He, S. J. Lippard, *J. Am. Chem. Soc.* **2000**, *122*, 184–185.

[43] A. Vogel, F. Spener, B. Krebs, in *Encycl. Inorg. Bioinorg. Chem.*, John Wiley & Sons, Ltd, **2011**.

[44] T. Klabunde, N. Sträter, B. Krebs, H. Witzel, *FEBS Lett.* **1995**, *367*, 56–60.

[45] G. Schenk, N. Mitić, G. R. Hanson, P. Comba, *Coord. Chem. Rev.* **2013**, *257*, 473–482.

[46] S. Albedyhl, M. T. Averbuch-Pouchot, C. Belle, B. Krebs, J. L. Pierre, E. Saint-Aman, S. Torelli, *Eur. J. Inorg. Chem.* **2001**, *2001*, 1457–1464.

[47] M. Jarenmark, M. Haukka, S. Demeshko, F. Tuczek, L. Zuppiroli, F. Meyer, E. Nordlander, *Inorg. Chem.* **2011**, *50*, 3866–3887.

[48] A. Neves, M. Aires de Brito, V. Drago, K. Griesar, W. Haase, *Inorganica Chim. Acta* **1995**, *237*, 131–135.

[49] A. E. Roberts, G. Schenk, L. R. Gahan, *Eur. J. Inorg. Chem.* **2015**, *2015*, 3076–3086.

[50] B. Bauer-Siebenlist, F. Meyer, E. Farkas, D. Vidovic, S. Dechert, *Chem. – Eur. J.* **2005**, *11*, 4349–4360.

[51] L. R. Gahan, S. J. Smith, A. Neves, G. Schenk, *Eur. J. Inorg. Chem.* **2009**, *2009*, 2745–2758.

[52] S. Bosch, P. Comba, L. R. Gahan, G. Schenk, *Inorg. Chem.* **2014**, *53*, 9036–9051.

[53] J. H. Burckhalter, R. I. Leib, *J. Org. Chem.* **1961**, *26*, 4078–4083.

[54] R. G. Hanshaw, E. J. O'Neil, M. Foley, R. T. Carpenter, B. D. Smith, *J. Mater. Chem.* **2005**, *15*, 2707.

[55] K. J. Oberhausen, J. F. Richardson, R. M. Buchanan, W. Pierce, *Polyhedron* **1989**, *8*, 659–668.

[56] X. Jiang, L. Huang, X. Liu, *Microchim. Acta* **2012**, *180*, 219–226.

[57] V. R. Correia, A. J. Bortoluzzi, A. Neves, A. C. Joussef, M. G. M. Vieira, S. C. Batista, *Acta Crystallogr. Sect. E Struct. Rep. Online* **2003**, *59*, m464–m466.

[58] H. Jiang, E. J. O'Neil, K. M. DiVittorio, B. D. Smith, *Org. Lett.* **2005**, *7*, 3013–3016.

[59] M. Minakawa, H.-M. Guo, F. Tanaka, *J. Org. Chem.* **2008**, *73*, 8669–8672.

[60] E. W. Payne, *J. Chem. Educ.* **1992**, *69*, A40.

[61] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239–254.

[62] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942–15943.

[63] F. Mancin, P. Tecilla, *New J. Chem.* **2007**, *31*, 800.

[64] C. A. Bunton, S. J. Farber, *J. Org. Chem.* **1969**, *34*, 767–772.

[65] J. R. Morrow, W. C. Trogler, *Inorg. Chem.* **1988**, *27*, 3387–3394.

[66] A. Tauber, H. P. Schuchmann, *Ultrason. Sonochem.* **2000**, *7*, 45–52.

[67] S. Torelli, C. Belle, I. Gautier-Luneau, J. L. Pierre, E. Saint-Aman, J. M. Latour, L. Le Pape, D. Luneau, *Inorg. Chem.* **2000**, *39*, 3526–3536.

[68] I. Bertini, *Biological Inorganic Chemistry: Structure and Reactivity*, University Science Books, **2007**.

[69] S. Iuchi, N. Kuldell, *Zinc Finger Proteins: From Atomic Contact to Cellular Function*, Springer US, **2005**.

[70] J. Kyte, *Structure in Protein Chemistry*, Garland Science, **2006**.

[71] B. A. Krizek, B. T. Amann, V. J. Kilfoil, D. L. Merkle, J. M. Berg, *J. Am. Chem. Soc.* **1991**, *113*, 4518–4523.

[72] M. Owska, K. Pawlas, W. Bal, M. Owska, K. Pawlas, W. Bal, *Bioinorg. Chem. Appl. Bioinorg. Chem. Appl.* **2010**, *2010, 2010*, e725153.

[73] H. M. Levy, N. Sharon, E. M. Ryan, D. E. Koshland, *Biochim. Biophys. Acta* **1962**, *56*, 118–126.

[74] O. Iranzo, A. Y. Kovalevsky, J. R. Morrow, J. P. Richard, *J. Am. Chem. Soc.* **2003**, *125*, 1988–1993.

[75] M. Zhao, H.-L. Wang, L. Zhang, C. Zhao, L.-N. Ji, Z.-W. Mao, *Chem. Commun.* **2011**, *47*, 7344.

[76] D. Jantz, J. M. Berg, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 7589–7593.

[77] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

[78] W. Chan, P. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, OUP Oxford, **2000**.

[79] S. A. Palasek, Z. J. Cox, J. M. Collins, *J. Pept. Sci.* **2007**, *13*, 143–148.

[80] E. Frérot, J. Coste, A. Pantaloni, M.-N. Dufour, P. Jouin, *Tetrahedron* **1991**, *47*, 259–270.

[81] A. Krezel, R. Latajka, G. D. Bujacz, W. Bal, *Inorg. Chem.* **2003**, *42*, 1994–2003.

[82] J. Adamczyk, W. Bal, A. Krężel, *Inorg. Chem.* **2015**, *54*, 596–606.

[83] A. D. Frankel, J. M. Berg, C. O. Pabo, *Proc. Natl. Acad. Sci.* **1987**, *84*, 4841–4845.

[84] S. M. Kelly, T. J. Jess, N. C. Price, *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2005**, *1751*, 119–139.

[85] E. Kopera, T. Schwerdtle, A. Hartwig, W. Bal, *Chem. Res. Toxicol.* **2004**, *17*, 1452–1458.

[86] J. A. Kornblatt, *PLOS ONE* **2009**, *4*, e6196.

[87] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci.* **1998**, *95*, 6705–6710.

[88] B. L. Nilsson, M. B. Soellner, R. T. Raines, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91–118.

[89] F. Fehr, *Synthesis and Analyses of Modified Zif268 Zinc Finger Domains*, Göttingen, **2011**.

[90] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249–289.

[91] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.

[92] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci.* **1999**, *96*, 10068–10073.

[93] G. S. Beligere, P. E. Dawson, *Pept. Sci.* **1999**, *51*, 363–369.

[94] R. Kothinti, N. M. Tabatabai, D. H. Petering, *J. Inorg. Biochem.* **2011**, *105*, 569–576.

[95] C. Arndt, S. Koristka, H. Bartsch, M. Bachmann, *Methods Mol. Biol. Clifton NJ* **2012**, *869*, 49–53.

[96] A. Harwood, in *Nucleic Acid Protoc. Handb.* (Ed.: R. Rapley), Humana Press, **2000**, pp. 73–75.

[97] B. D. Hames, *Gel Electrophoresis of Proteins: A Practical Approach*, OUP Oxford, **1998**.

[98] E. T. Kool, J. C. Morales, K. M. Guckian, *Angew. Chem. Int. Ed.* **2000**, *39*, 990–1009.

[99] A. Sreedhara, J. A. Cowan, *JBIC J. Biol. Inorg. Chem.* **2001**, *6*, 337–347.

[100] P. K. Mohapatra, P. K. Khopkar, *Polyhedron* **1989**, *8*, 2071–2076.

[101] B. K. Takasaki, J. Chin, *J. Am. Chem. Soc.* **1994**, *116*, 1121–1122.

[102] M. Komiyama, N. Takeda, Y. Takahashi, H. Uchida, T. Shiiba, T. Kodama, M. Yashiro, *J Chem Soc Perkin Trans 2* **1995**, 269–274.

[103] S. J. Franklin, *Curr. Opin. Chem. Biol.* **2001**, *5*, 201–208.

[104] M. Komiyama, N. Takeda, H. Shigekawa, *Chem. Commun.* **1999**, 1443–1451.

[105] C. F. Baes, R. E. Mesmer, *The Hydrolysis of Cations*, New York : Wiley, **1976**.

[106] J. Sumaoka, T. Igawa, K. Furuki, M. Komiyama, *Chem. Lett.* **2000**, *29*, 56–57.

[107] H. Shigekawa, H. Ikawa, R. Yoshizaki, Y. Iijima, J. Sumaoka, M. Komiyama, *Appl. Phys. Lett.* **1996**, *68*, 1433–1435.

[108] H. Shigekawa, M. Ishida, K. Miyake, R. Shioda, Y. Iijima, T. Imai, H. Takahashi, J. Sumaoka, M. Komiyama, *Appl. Phys. Lett.* **1999**, *74*, 460–462.

[109] A. Fujimori, *Phys. Rev. B* **1983**, *28*, 4489–4499.

[110] A. Bianconi, A. Marcelli, H. Dexpert, R. Karnatak, A. Kotani, T. Jo, J. Petiau, *Phys. Rev. B* **1987**, *35*, 806–812.

[111] Y. Aiba, M. Komiyama, *Polym. J.* **2012**, *44*, 929–938.

[112] C. S. Chen, M. Mrksich, S. Huang, G. M. Whitesides, D. E. Ingber, *Science* **1997**, *276*, 1425–1428.

[113] M. Komiyama, *Int. J. Mol. Sci.* **2013**, *14*, 3343–3357.

[114] M. E. Branum, A. K. Tipton, S. Zhu, L. Que, *J. Am. Chem. Soc.* **2001**, *123*, 1898–1904.

[115] Y. Yamamoto, A. Uehara, T. Tomita, M. Komiyama, *Nucleic Acids Res.* **2004**, *32*, e153–e153.

[116] W. Kameshima, T. Ishizuka, M. Minoshima, M. Yamamoto, H. Sugiyama, Y. Xu, M. Komiyama, *Angew. Chem.* **2013**, *125*, 13926–13929.

[117] E G Krebs, J. A. Beavo, *Annu. Rev. Biochem.* **1979**, *48*, 923–959.

[118] I. Johansson, M. Karlsson, V. K. Shukla, M. J. Chrispeels, C. Larsson, P. Kjellbom, *Plant Cell* **1998**, *10*, 451–459.

[119] J. V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, M. Mann, *Cell* **2006**, *127*, 635–648.

[120] D. E. Petrillo, D. R. Mowrey, S. P. Allwein, R. P. Bakale, *Org. Lett.* **2012**, *14*, 1206–1209.

[121] T. Wakamiya, K. Saruta, J. Yasuoka, S. Kusumoto, *Chem. Lett.* **1994**, *23*, 1099–1102.

[122] M. Mann, S.-E. Ong, M. Grønborg, H. Steen, O. N. Jensen, A. Pandey, *Trends Biotechnol.* **2002**, *20*, 261–268.

[123] T. J. Attard, N. M. O'Brien-Simpson, E. C. Reynolds, *Int. J. Pept. Res. Ther.* **2009**, *15*, 69–79.

[124] J. M. Collins, M. J. Collins, *ChemInform* **2008**, *39*.

[125] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, S. Duhr, *ASSAY Drug Dev. Technol.* **2011**, *9*, 342–353.

[126] M. C. Jecklin, S. Schauer, C. E. Dumelin, R. Zenobi, *J. Mol. Recognit.* **2009**, *22*, 319–329.

[127] S. A. I. Seidel, P. M. Dijkman, W. A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J. S. Joseph, P. Srinivasan, P. Baaske, A. Simeonov, *Methods* **2013**, *59*, 301–315.

[128] Y. Yamamoto, M. Mori, Y. Aiba, T. Tomita, W. Chen, J.-M. Zhou, A. Uehara, Y. Ren, Y. Kitamura, M. Komiyama, *Nucleic Acids Res.* **2007**, *35*, e53.

[129] M. Jerabek-Willemsen, T. André, R. Wanner, H. M. Roth, S. Duhr, P. Baaske, D. Breitsprecher, *J. Mol. Struct.* **2014**, *1077*, 101–113.

[130] S. Haldar, A. Chattopadhyay, in *Fluoresc. Methods Study Biol. Membr.* (Eds.: Y. Mély, G. Duportail), Springer Berlin Heidelberg, Berlin, Heidelberg, **2012**, pp. 37–50.

[131] M. J. Geisow, *Exp. Cell Res.* **1984**, *150*, 29–35.

[132] K. H. Scheller, T. H. Abel, P. E. Polanyi, P. K. Wenk, B. E. Fischer, H. Sigel, *Eur. J. Biochem. FEBS* **1980**, *107*, 455–466.

[133] Y. Yamamoto, T. Igawa, J. Sumaoka, M. Komiyama, in *Nucleic Acids Symp. Ser.*, Oxford Univ Press, **2000**, pp. 231–232.

[134] J. Otsuki, T. Akasaka, K. Araki, *Coord. Chem. Rev.* **2008**, *252*, 32–56.

[135] T. H. Scheuermann, S. B. Padrick, K. H. Gardner, C. A. Brautigam, *Anal. Biochem.* **2016**, *496*, 79–93.

[136] NanoTemper Technologies GmbH, *Monolith NT.115 User Manual*, München, **2011**.

[137] R. A. Reichle, K. G. McCurdy, L. G. Hepler, *Can. J. Chem.* **1975**, *53*, 3841–3845.

[138] Y.-B. Choi, S. A. Lipton, *Neuron* **1999**, *23*, 171–180.

[139] P. Paoletti, P. Ascher, J. Neyton, *J. Neurosci.* **1997**, *17*, 5711–5725.

[140] T. Terashima, M. Kawabe, Y. Miyabara, H. Yoda, M. Sawamoto, *Nat. Commun.* **2013**, *4*, 2321.

[141] G. M. Sanders, G. A. Kassavetis, E. P. Geiduschek, *Proc. Natl. Acad. Sci.* 1994, *91*, 7703–7707.

[142] Y. G. Kim, J. Cha, S. Chandrasegaran, *Proc. Natl. Acad. Sci.* **1996**, *93*, 1156–1160.

[143] Y. Kitamura, J. Sumaoka, M. Komiyama, *Tetrahedron* **2003**, *59*, 10403–10408.

[144] Y.-Y. Jang, L. Cai, Z. Ye, *Discov. Med.* **2016**, *21*, 57–64.

[145] T. J. Cradick, in *Nov. Immunother. Approaches Treat. Cancer* (Ed.: P.D. Rennert), Springer International Publishing, **2016**, pp. 203–247.

[146] K. Abe, J. Izumi, M. Ohba, T. Yokoyama, H. Okawa, *Bull. Chem. Soc. Jpn.* **2001**, *74*, 85–95.

[147] M. Stefani, A. Caselli, M. Bucciantini, L. Pazzagli, F. Dolfi, G. Camici, G. Manao, G. Ramponi, *FEBS Lett.* **1993**, *326*, 131–134.

[148] V. D. Romanenko, V. P. Kukhar, *Chem. Rev.* **2006**, *106*, 3868–3935.

[149] K. Panigrahi, M. Eggen, J.-H. Maeng, Q. Shen, D. B. Berkowitz, *Chem. Biol.* **2009**, *16*, 928–936.

[150] D. B. Berkowitz, Q. Shen, J.-H. Maeng, *Tetrahedron Lett.* **1994**, *35*, 6445–6448.

[151] D. E. Koshland, *Angew. Chem. Int. Ed. Engl.* **1995**, *33*, 2375–2378.

[152] D. L. Purich, *Enzyme Kinetics: Catalysis & Control: A Reference of Theory and Best-Practice Methods*, Elsevier, **2010**.

[153] R. C. Dubey, *Advanced Biotechnology*, S. Chand Publishing, **2014**.

[154] D. Herschlag, *Bioorganic Chem.* **1988**, *16*, 62–96.

[155] T. Nakatsukasa, Y. Shiraishi, S. Negi, M. Imanishi, S. Futaki, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **2005**, *330*, 247–252.

[156] A. Klug, *Annu. Rev. Biochem.* **2010**, *79*, 213–231.

[157] G. K. Walkup, B. Imperiali, *J. Am. Chem. Soc.* **1996**, *118*, 3053–3054.

[158] C. Nadler, *Extending the Scope of Protein Synthesis by a Novel Auxiliary-Based Native Chemical Ligation Strategy*, Göttingen, **2013**.

[159] B. A. Krizek, D. L. Merkle, J. M. Berg, *Inorg. Chem.* **1993**, *32*, 937–940.

[160] P. F. Predki, B. Sarkar, *Environ. Health Perspect.* **1994**, *102*, 195–198.

[161] B. L. Vallee, D. S. Auld, *Acc. Chem. Res.* **1993**, *26*, 543–551.

[162] T.-L. Ho, *Chem. Rev.* **1975**, *75*, 1–20.

[163] A. P. de Silva, H. Q. N. Gunaratne, T. Gunnlaugsson, A. J. M. Huxley, C. P. McCoy, J. T. Rademacher, T. E. Rice, *Chem. Rev.* **1997**, *97*, 1515–1566.

[164] K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson, *Nature* **2009**, *460*, 823–830.

[165] R. M. Clegg, *Curr. Opin. Biotechnol.* **1995**, *6*, 103–110.

[166] G. K. Walkup, B. Imperiali, *J. Am. Chem. Soc.* **1997**, *119*, 3443–3450.

[167] M. J. Romero, R. Pedrido, A. M. González-Noya, M. Maneiro, M. I. Fernández-García, G. Zaragoza, M. R. Bermejo, *Dalton Trans.* **2012**, *41*, 10832.

[168] J. Hatai, S. Pal, G. P. Jose, S. Bandyopadhyay, *Inorg. Chem.* **2012**, *51*, 10129–10135.

[169] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

[170] C. O. Kappe, E. Van der Eycken, *Chem Soc Rev* **2010**, *39*, 1280–1290.

[171] A. Nadler, C. Hain, U. Diederichsen, *Eur. J. Org. Chem.* **2009**, 4593–4599.

[172] H. Struthers, T. L. Mindt, R. Schibli, *Dalton Trans* **2010**, *39*, 675–696.

[173] M. M. Yamashita, L. Wesson, G. Eisenman, D. Eisenberg, *Proc. Natl. Acad. Sci.* 1990, *87*, 5648–5652.

[174] R. J. Sundberg, R. B. Martin, *Chem. Rev.* **1974**, *74*, 471–517.

[175] J. Košmrlj, B. R. Buckley, Eds. , *Click Triazoles*, Springer, Heidelberg ; New York, **2012**.

[176] C. S. McKay, M. G. Finn, *Chem. Biol.* **2014**, *21*, 1075–1101.

[177] D. L. Merkle, M. H. Schmidt, J. M. Berg, *J. Am. Chem. Soc.* **1991**, *113*, 5450–5451.

[178] K. P. Carter, A. M. Young, A. E. Palmer, *Chem. Rev.* **2014**, *114*, 4564–4601.

[179] Y. H. Lau, P. J. Rutledge, M. Watkinson, M. H. Todd, *Chem. Soc. Rev.* **2011**, *40*, 2848.

[180] S. Huang, R. J. Clark, L. Zhu, *Org. Lett.* **2007**, *9*, 4999–5002.

[181] K. Varazo, F. Xie, D. Gulledge, Q. Wang, *Tetrahedron Lett.* 2008, *49*, 5293–5296.

[182] P. Ceroni, V. Balzani, in *Explor. Supramol. Syst. Nanostructures Photochem. Tech.* (Ed.: P. Ceroni), Springer Netherlands, Dordrecht, **2012**, pp. 21–38.

[183] Z. Xu, J. Yoon, D. R. Spring, *Chem. Soc. Rev.* **2010**, *39*, 1996.

[184] G. de Miguel, M. Wielopolski, D. I. Schuster, M. A. Fazio, O. P. Lee, C. K. Haley, A. L. Ortiz, L. Echegoyen, T. Clark, D. M. Guldi, *J. Am. Chem. Soc.* **2011**, *133*, 13036–13054.

[185] Z. Xu, Y. Xiao, X. Qian, J. Cui, D. Cui, *Org. Lett.* **2005**, *7*, 889–892.

[186] B. P. Joshi, J. Park, W. I. Lee, K.-H. Lee, *Talanta* **2009**, *78*, 903–909.

[187] B. N. G. Giepmans, S. R. Adams, M. H. Ellisman, R. Y. Tsien, *Science* **2006**, *312*, 217–224.

[188] I. A. Inverarity, A. N. Hulme, *Org. Biomol. Chem.* **2007**, *5*, 636.

[189] K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill, Q. Wang, *Org. Lett.* **2004**, *6*, 4603–4606.

[190] N. Balachander, C. N. Sukenik, *Langmuir* **1990**, *6*, 1621–1627.

[191] S. I. Presolski, V. P. Hong, M. G. Finn, *Curr. Protoc. Chem. Biol.* **2011**, *3*, 153–162.

[192] A. Rigo, A. Corazza, M. L. di Paolo, M. Rossetto, R. Ugolini, M. Scarpa, *J. Inorg. Biochem.* **2004**, *98*, 1495–1501.

[193] Y. Kim, S. H. Kim, D. Ferracane, J. A. Katzenellenbogen, C. M. Schroeder, *Bioconjug. Chem.* **2012**, *23*, 1891–1901.

[194] C. Kellenberger, H. Hietter, B. Luu, *Pept. Res.* **1994**, *8*, 321–327.

[195] J. E. Penner-Hahn, *Coord. Chem. Rev.* **2005**, *249*, 161–177.

[196] Y. Shi, R. D. Beger, J. M. Berg, *Biophys. J.* **1993**, *64*, 749.

[197] J. M. Berg, D. L. Merkle, *J. Am. Chem. Soc.* **1989**, *111*, 3759–3761.

[198] A. G. Marangoni, *Enzyme Kinetics: A Modern Approach*, John Wiley & Sons, **2003**.

[199] A. Miłoch, A. Krężel, *Metallomics* **2014**, *6*, 2015–2024.

[200] M. Homocianu, A. Airinei, D. O. Dorohoi, I. Olariu, N. Fifere, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2011**, *82*, 355–359.

[201] W. Y. Lin, H. E. Van Wart, *J. Inorg. Biochem.* **1988**, *32*, 21–38.

[202] X. Liu, N. Zhang, J. Zhou, T. Chang, C. Fang, D. Shangguan, *The Analyst* **2013**, *138*, 901–906.

[203] D. Maity, T. Govindaraju, *Chem Commun* **2012**, *48*, 1039–1041.

[204] A. Kosoy, C. Moller, D. Perdomo, J. Bubis, *J. Biochem. Mol. Biol.* **2004**, *37*, 260.

[205] L. M. Miller, Q. Wang, T. P. Telivala, R. J. Smith, A. Lanzirotti, J. Miklossy, *J. Struct. Biol.* **2006**, *155*, 30–37.

[206] H. Rao, *Metal Containing Peptides as Specific DNA Binders*, Göttingen, **2015**.

[207] T. Kawakami, H. Murakami, T. Kawakami, H. Murakami, *J. Nucleic Acids J. Nucleic Acids* **2012**, e713510.

List of Abbreviations

A	Adenine
Å	Ångström
aa	amino acid
ACEN	anthracene
Alloc	allyloxycarbonylchloride
Approx	approximately
APS	ammonium persulfate
aq	aqueous
Bn	benzyl
Boc	butyloxycarbonyl
BNPP	bis(4-nitrophenyl)phosphate
bp	base pairs
br	broad
Bu	butyl
Cas9	CRISPR associated protein 9
CBD	Chitin Binding Domain
C	cytosine
°C	degree celsius
CD	Circular Dichroism
cm	centimeter
COUM	coumarin
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
δ	chemical shift
d	doublet
DNS	dansyl
dd	doublet of doublet
DCM	dichloromethane

DIC	diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMF	dimethylformamide
DNA	deoxyribonucleic acid
ds	double strand
ε	extinction coefficient
EDT	1,2-ethanedithiol
EDTA	ethylenediaminetetraacetic acid
EPL	Expressed Protein Ligation
ESI-MS	Electron Spray Ionization Mass Spectroscopy
et al.	et alii
Et_3N	triethylamine
EtOAc	ethylacetate
EtOH	ethanol
Et_2O	diethyl ether
eq	equivalents
FAM	carboxyfluorescein
FITC	fluorescein isothiocyanate
Fig.	figure number
Fmoc	9-fluorenylmethoxycarbonyl
Fmoc-OSu	N-(9-fluorenylmethoxycarbonyl)succinimide
g	gram
G	guanine
GmdCl	guanidine hydrochloride
h	hours
HATU	1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxid hexafluorophosphate
HBTU	*N,N,N',N'*-tetramethyl-O-(1H-benzotriazol-1-yl)uraniumhexafluorophosphate

HFIP	hexafluoroisopropanol
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectroscopy
IMPACT	Intein Mediated Purification with an Affinity Chitin-Tag
IPL	Intein Mediated Protein Ligation
IPTG	isopropyl-beta-D-thiogalactopyranoside
IRTG	International Research Training Group
J	coupling constant
L	liter
LB	Luria-Bertani
M	multiplet
MeCN/ACN	acetonitrile
MeOH	methanol
MeSNa	sodium 2-mercaptoethanesulfonate
mg	milligram
MHz	megahertz
min	minute
mL	millilitre
mM	millimolar
MST	Microscale Thermophoresis
μM	micromolar
NaAsc	sodium ascorbate
NCL	Native Chemical Ligation
nm	nanometer
nM	nanomolar
NMM	N-methylmorpholine

NMP	N-methyl-2-pyrrolidone
NMR	Nuclear Magnetic Resonance
NCL	Native Chemical Ligation
PAGE	Polyacrylamide Gel Electrophoresis
PDB	Protein Data Base
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
Pd/C	palladium on activated charcoal
PE	polyethylene
PG	protecting group
ppm	parts per million
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
q	quartet
r_f	ratio of peptide to DNA
R_f	retardation factor
RNA	ribonucleic acid
RP	reverse phase
rpm	rounds per minute
r.t.	room temperature
s	singlet
sec	seconds
SPPS	Solid Phase Peptide Synthesis
ss	single strand
t	triplet
T	Thymine
Tacn	triazacyclononane
TALEN	Transcription Activator-Like Effector Nuclease
tBoc	*tert*-butoxycarbonyl
TCEP	tris(carboxyethyl)phosphine

TEMED	N,N,N′,N′-tetramethylethylenediamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIS	triisopropylsilane
TLC	Thin Layer Chromatography
TMS	trimethylsilyl
TosCl	tosylchloride
t_R	retardation time
Trt	trityl protecting group
UV	ultra-violet
V	volt
v/v	volume by volume
W	watt
wt%	weight percentage
w/v	weight by volume
Zf	zinc finger
ZFN	zinc finger nuclease

Acknowledgements

Curriculum Vitae

Personal Data

Name:	Florian Czerny
Date of birth:	22. January 1985
Place of birth:	Straubing, Bavaria, Germany
Nationality:	German

Work Experience

Since 05/2012

Research associate at the Institute of Organic and Biomolecular Chemistry at the Georg-August-University, Göttingen.
- *Teaching of different chemistry classes and tutorials*
- *Lab course leader and teaching assistant in the modules "organic chemistry for biochemistry students", "basic organic chemistry" and "advanced organic chemistry"*
- *Mentoring of students during their Bachelor- and Master thesis*

Academic Education

Since 05/2012

PhD thesis in the working group of Prof. Dr. Ulf Diederichsen at the Institute of Organic and Biomolecular Chemistry at the Georg-August-University Göttingen within the framework of the IRTG 1422 "metal sites in biomolecules".

Title: "*Development of zinc-finger-based artificial restriction endonucleases and fluorescent peptidyl metal sensors*".

04/2015 – 06/2015

Visiting scholar in the working group of Prof. Dr. Sofi Elmroth at the Department of Biochemistry and Structural Biology at the University of Lund (Sweden).

05/2011 – 02/2012

Diploma thesis in the working group of Prof. Dr. Ulf Diederichsen at the Institute of Organic and Biomolecular Chemistry at the Georg-August-University, Göttingen.

Title: "*Synthesis and modification of the zinc finger protein Zif268 by inserting a second metal binding site*".

10/2005 – 05/2012

Studies in chemistry at the Georg-August-University, Göttingen

Education

06/2005	General matriculation standard at the Otto-Hahn Gymnasium in Göttingen

Languages & IT

Languages	German (native speaker), English (business fluent)
IT	MS Windows, Mac OSX, MS Office, SciFinder, ChemOffice, Origin, PyMOL