

DOI: 10.1002/open.201100002

Semi-Synthesis and Analysis of Chemically Modified Zif268 Zinc-Finger Domains

 Friederike Fehr,^[a] André Nadler,^[a] Florian Brodhun,^[b] Ivo Feussner,^[b] and Ulf Diederichsen^{*,[a]}

Total synthesis of proteins can be challenging despite assembling techniques, such as native chemical ligation (NCL) and expressed protein ligation (EPL). Especially, the combination of recombinant protein expression and chemically addressable solid-phase peptide synthesis (SPPS) is well suited for the redesign of native protein structures. Incorporation of analytical probes and artificial amino acids into full-length natural protein domains, such as the sequence-specific DNA binding zinc-finger motifs, are of interest combining selective DNA recognition and artificial function. The semi-synthesis of the natural 90

amino acid long sequence of the zinc-finger domain of Zif268 is described including various chemically modified constructs. Our approach offers the possibility to exchange any amino acid within the third zinc finger. The realized modifications of the natural sequence include point mutations, attachment of a fluorophore, and the exchange of amino acids at different positions in the zinc finger by artificial amino acids to create additional metal binding sites. The individual constructs were analyzed by circular dichroism (CD) spectroscopy with respect to the integrity of the zinc-finger fold and DNA binding.

Introduction

The understanding of gene-regulation mechanisms aiming for early identification of morbid cells or false DNA sequences is one main focus in current postgenomic research towards medically relevant problems. Selective gene expression is mainly accomplished by the interaction of DNA with selective binding proteins containing specific DNA binding domains, such as helix-turn-helix (HTH) and zinc-finger motifs.^[1] It is assumed that in higher eukaryotes more than 15 000 zinc-finger motifs occur in 1000 proteins, which makes it one of the most abundant motifs.^[2] Zinc-finger motifs possess the ability to recognize DNA sequences selectively and are part of most transcription factors.^[3] They contain a zinc ion in tetrahedral coordination with cysteine and histidine residues, resulting in the eponymous simple $\beta\beta\alpha$ fold.^[4] Zinc-finger motifs are subdivided into classes with respect to the amino acid residues coordinating the structural zinc.^[5] The best investigated zinc-finger domain of the canonical Cys2His2 type belongs to the murine transcription factor Zif268 and serves as a useful model to analyze protein–DNA interactions.^[6] Since Zn^{2+} displays different properties within a wide range of proteins, its functions are classified into catalytically active, with a readily exchangeable water ligand at the coordination site and structure determining, where all coordination sites are occupied by donor atoms of the protein side chains.^[7] In zinc fingers, the metal contributes only to the $\beta\beta\alpha$ fold and plays no catalytic or functional role. In Zif268, the DNA binding motif is composed of a tandem repeat of three zinc fingers, each consisting of about 30 amino acids, serving as recognition units and binding to the major groove of B-DNA. Each of the three zinc fingers has the ability to recognize a sequence of three to four nucleobases through four conserved amino acids in the α -helical part and is linked by a short well-conserved sequence to the follow-

ing one. It is known that at least two zinc-finger units are required for selectivity.^[7,8]

Aim of the current work was the incorporation of an additional metal binding site into the zinc-finger motif allowing functionalization of a sequence-specific DNA binder placed in the major groove. A functional metal binding site placed in close proximity to the structural metal center might serve as spectroscopic probe or allows for cooperative effects with respect to catalytic activity.^[9] In particular, metal-based nucleases often contain two metal ions, such as Mg^{2+} , Mn^{2+} or Zn^{2+} , in close proximity that are involved in the cleavage of the DNA phosphodiester backbone.^[10] Despite numerous investigations on native metalloproteins and their functions to understand fundamental principles in chemistry and biology,^[11] little is known about strategies accomplishing artificial metal-containing, biologically active peptides or proteins.^[12] In this respect, the redesign of native proteins has an advantage over the de novo design of peptides, as it proceeds from a known scaffold.

[a] Dr. F. Fehr, Dr. A. Nadler, Prof. Dr. U. Diederichsen
Georg-August-Universität Göttingen
Institut für Organische und Biomolekulare Chemie
Tammannstrasse 2, 37077 Göttingen (Germany)
Fax: (+49)-551-39-2944
E-mail: udieder@gwdg.de

[b] Dr. F. Brodhun, Prof. Dr. I. Feussner
Georg-August-Universität Göttingen
Albrecht-von-Haller Institut für Pflanzenwissenschaften
Justus-von-Liebig-Weg 11, 37077 Göttingen (Germany)

© 2012 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

fold requiring only minor modulation.^[7,13] Zinc-finger motifs are also convenient from the synthetic aspect, because cysteine residues in a tandem repeat of a conserved amino acid sequence offer potential native chemical ligation sites in attractive intervals. Furthermore, by coordinating a metal ion, zinc fingers adopt a defined peptide fold. The semi-synthesis and characterization of the zinc-finger domain of the Zif268 transcription factor by expressed protein ligation (EPL) is reported. Besides the native binding domain, modifications within the sequence of the third zinc finger are investigated, such as single amino acid mutations, attached fluorophore and constructs with artificial metal-chelating amino acids.^[9b]

Results and Discussion

Strategy

The zinc-finger domain of the transcription factor Zif268 (Egr1) is a well-known and characterized motif, and for these reasons, was chosen as selective DNA binding motif. The sequence contains 90 amino acids and is hardly accessible by standard fluorenylmethyloxycarbonyl (Fmoc) solid-phase peptide synthesis (SPPS). A semi-synthetic approach based on EPL was applied to assemble the full-length peptide. Thereby, the first two zinc fingers (Zf12, amino acids 1–64) were expressed in *Escherichia coli*, whereas the third zinc finger (Zf3, amino acids 65–90) was synthesized by Fmoc SPPS. Native chemical ligation (NCL) yielded the desired sequence of 90 amino acids.

Generation and purification of Zf12

After amplification of the open reading frame of the zinc-finger domain (Zf12) by polymerase chain reaction (PCR), the desired DNA sequence was cloned in-frame into the pTXB1 expression vector. The Zf12 peptide was expressed in *E. coli* in fusion with a mini-intein (*GyrA* gene of *Mycobacterium xenopi*) and a chitin binding domain (CBD) from *Bacillus circulans*.^[14] After affinity chromatography using immobilized chitin beads, the Zf12 peptide was cleaved from the resin by thioesterification adding the sodium salt of 2-mercaptoethanesulfonic acid (MESNa; Figure 1). The thioester peptide **1** (Zf12) was eluted,

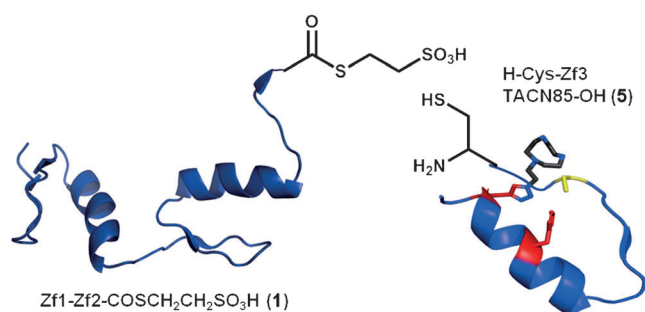


Figure 1. Image of the expressed C-terminal peptide thioester Zf12 (amino acids 1–64, left) and the peptide Zf3 (amino acids 65–91, right) with an N-terminal cysteine and a modification with an additional metal binding site at His 85 (modified Protein Data Bank (PDB) entry: IAAY^[6]).

and further purified by reverse phase (RP)-HPLC giving a final yield of 5 mg L⁻¹ cell culture. From tris-tricine-based sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), it was concluded that thioester **1** was eluted from the affinity column within the first fractions (Figure 2);^[15] the constitutional integrity of peptide **1** was confirmed by HRMS-ESI.

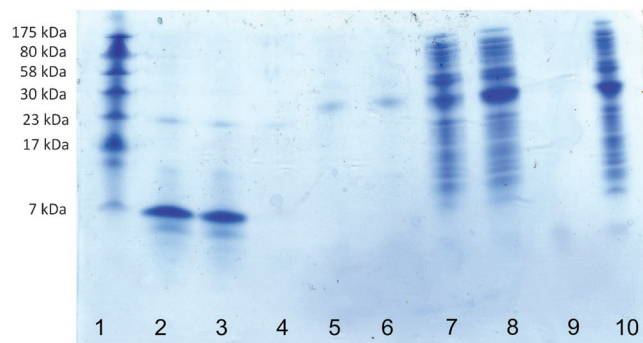


Figure 2. Tris-tricine SDS-PAGE (10–20%) of Zf12 (8 kDa); lanes: 1 = molecular weight standard; 2 = fraction 2 (Zf12); 3 = fraction 1 (Zf12); 4 = cleavage; 5 = flow through (FT); 6 = FT 2; 7 = FT 1; 8 = lysate; 9 = inclusion bodies; 10 = pellet.

SPPS of Zf3 and Zf3 modifications

The syntheses of Zf3 and its variants were achieved by a combination of manual and automated standard microwave-assisted Fmoc-based SPPS on Wang resin. Hydroxybenzotriazole (HOBt) and *O*-benzotriazole-*N,N,N',N'*-tetramethyl-uronium-hexafluoro-phosphate (HBTU) were used as coupling reagents. Besides the natural sequence of Zf3, five variants were synthesized. Modification Zf3 A85 (**2**) contains histidine at position 85 exchanged for an alanine (Figure 3). His85 is one out of four ligands coordinating the structural zinc ion in the system. Mutation to alanine creates a vacant coordination site at the metal center offering the potential for catalytic activity.^[16] This variant was further modified by introduction of a 5(6)-carboxyfluorescein fluorophore (Zf3 A85L89, **3**) to the second last C-terminal amino acid Lys89. Therefore, peptide **3** lacks His85 and contains 5(6)-carboxyfluorescein coupled to the lysine side chain. During SPPS, the lysine side chain was allyloxycarbonyl (alloc)-protected, and coupling to the fluorophore was activated by HOBt, *N,N'*-diisopropylcarbodiimide (DIC), and *N*-ethyl-diisopropylamine (DIPEA). By using the zinc finger containing a fluorophore, zinc-finger ligation could be monitored by UV, and this modified zinc finger will also be valuable for following its interactions with DNA by Förster resonance energy transfer (FRET).^[17]

In variant Zf3 Pra70 (**4**), the arginine residue at position 70 was exchanged with the artificial amino acid propargylglycine (Pra). This amino acid with its acetylene side chain permits the introduction of various functionalities into the protein by [2+3] cycloaddition with azides.^[18] Azide-functionalized moieties, such as fluorophores or metal chelators, can thereby be introduced at position 70 of the zinc finger. Native Arg 70 maintains a phosphate contact to the DNA backbone, which offers

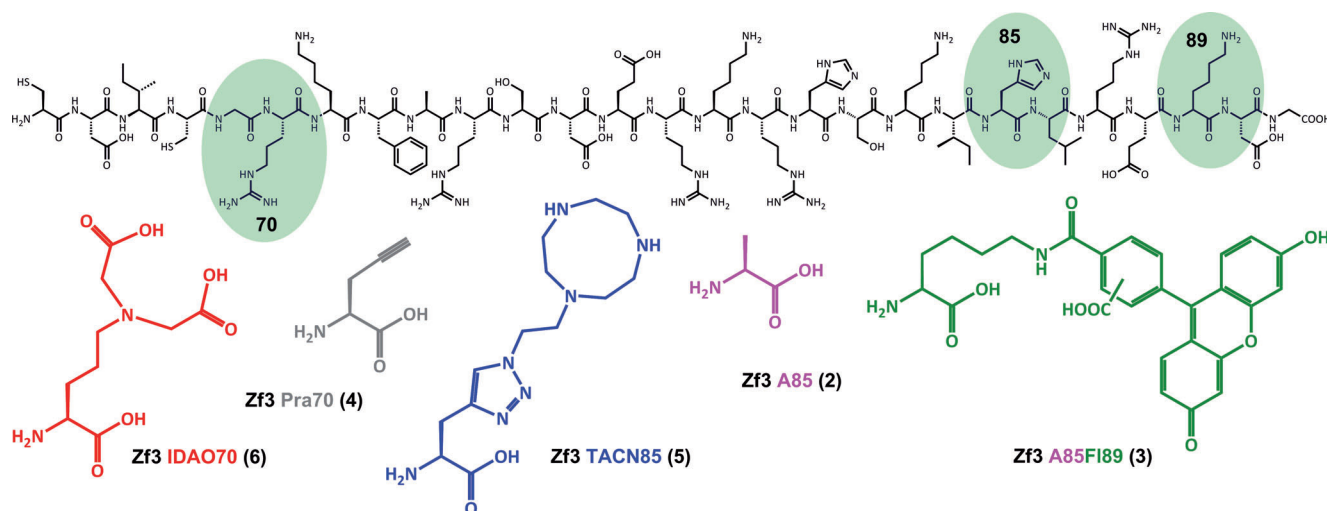


Figure 3. Zinc-finger 3 (Zf3) fragment prepared by SPPS. Exchanged amino acids Arg70, His85 and Lys89 are indicated together with newly introduced amino acids alanine (Zf3 A85, 2), alanine and 5(6)-carboxyfluoresceine (Zf3 A85 L89, 3), propargylglycine (Zf3 Pra70, 4), triazole-linked triazacyclononane (Zf3 TACN85, 5), and iminodiacetic acid ornithine (Zf3 IDAO70, 6).^[12]

the possibility to introduce a broad range of functional groups close to a potential DNA cleavage site.^[6] Furthermore, with Zf3 TACN85 (5), a triazole was introduced as an imidazole mimic of the metal-chelating amino acid His85 alkylated by a 1,4,7-triazacyclononane (TACN) group to create a second metal binding site. Preparation of Fmoc-propargylglycine-OH by [2+3] cycloaddition as well as an azide-functionalized *N,N*-bis-Boc-protected TACN was reported previously.^[9b] The appropriate Fmoc-TACN-OH building block was incorporated in the peptide by SPPS under standard HOBt/HBTU conditions. The triazole in Zf3 TACN85 (5) mimics the histidine and additionally creates a second metal binding site in close proximity to the natural one.^[19]

The zinc-finger variant Zf3 IDAO70 (6) contains an ornithine analogue, iminodiacetic acid ornithine (IDAO), which is functionalized with two acetic acid groups at its amine side chain, as an artificial metal-chelating amino acid.^[9b] Arg70 was exchanged with the Fmoc-IDAO-OH building block that was manually coupled, protected as bis-*tert*-butylester under 1-hydroxy-7-azabenzotriazole (HOAt)/HATU coupling conditions in threefold excess by using a microwave-assisted protocol.

Native chemical ligation

The full-length zinc-finger proteins were generated by native chemical ligation of the expressed Zf12 thioester peptide 1 and the respective synthetic Zf3 peptide with an N-terminal cysteine residue.^[20] These unfolded zinc fingers easily form disulfide bridges, and reduced cysteines are crucial for the ligation step. Therefore, the peptides were treated with the tris(2-carboxyethyl)phosphine (TCEP), as a reducing agent, carefully balancing the pH. Because reduction with TCEP only occurs at a pH of about 4, and the optimal pH for NCL is in the range of 7.8 to 8, Zf3 peptides were dissolved in a millimolar range (threefold excess with respect to Zf12) under argon in 10 mM phosphate buffer (pH 4) containing 20 mM TCEP and 6 M guanidinium hydrochloride (Gnd-HCl) (1 v/v). The reduction was

complete after 2 h, as confirmed by ESI-MS. The expressed Zf12 was solved in a millimolar range in 500 mM phosphate buffer and 6 M Gnd-HCl (1.2 v/v) at pH 8. The two solutions were mixed subsequently, and the resulting pH was kept between 7.8 and 8. Conversion into the ligation products was followed by analytical RP-HPLC. The respective Zf13 ligation product was obtained after 6–12 h, purified by RP-HPLC, and subsequently analyzed by ESI-MS. Because disulfide formation was indicated by ESI-MS, reduction was performed at pH 4 by reaction with 20 mM of TCEP under argon for 2 h. The pH was then adjusted to 8 to avoid amine protonation and allow Zn²⁺ coordination. As an example, the ESI-MS spectrum for the reduced modified zinc-finger construct Zf13 A85L89 (3) is presented in Figure 4.

Structural analysis and DNA interaction of Zf13 constructs

CD measurements were used to study the influence of the modifications on the structure of Zf3 peptides. The zinc-in-

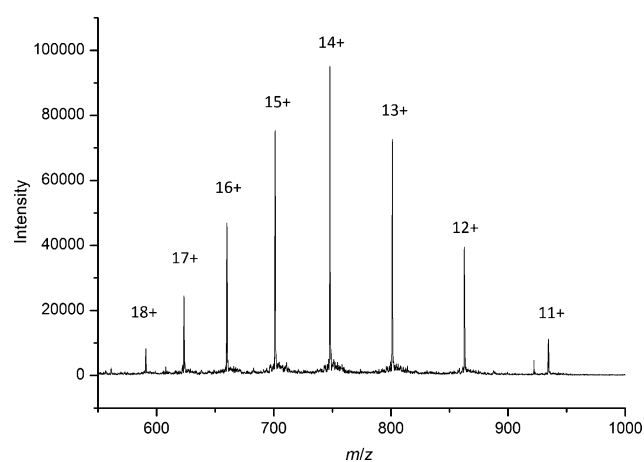


Figure 4. ESI-MS spectrum of the ligation product Zf13 A85 L89 (3).

duced protein folding can be monitored by CD spectroscopy, mainly by following a hypsochromic shift of the negative maximum close to 200 nm.^[21] All reduced Zf3 analogues were analyzed in absence and presence of ZnSO₄. The reported blueshift was observed for all analogues upon addition of an excess Zn²⁺; exemplarily, in Figure 5 spectra for native peptide Zf3 and the analogue containing an additional metal binding site, Zf3 TACN85 (5) are shown. In Zf3 TACN85 (5), the metal-coordinating His85 is replaced by a triazole that is also likely to be involved in zinc coordination, and as such, the CD spectra of Zf3 and Zf3 TACN85 (5) are quite similar. Also, mutant Zf3 A85 (2), lacking the histidine metal coordination site is still able to provide the correct fold (data not shown), indicating that the typically ββα zinc-finger fold is not significantly affected by the Zf3 mutations.^[16a]

CD spectra were also recorded for the full-length Zf13 analogues. As indicated for Zf13 TACN85 and Zf13 IDAO70 analogues, the negative maximum at 200 nm undergoes a slightly larger hypsochromic shift of about 6 nm after Zn²⁺ addition. In the case of Zf13 IDAO70, the shift of the negative maximum around 200 nm was accompanied with a strong negative Cotton effect and a shoulder at 222 nm, which is also typical of the zinc-finger fold (Figure 6).^[21]

The interaction of zinc-finger constructs with double-stranded (ds) DNA was further investigated by means of CD spectroscopy. The 28-base pair (bp) DNA sequence (5'-CTTACGCCACGCTGATCACACACACAC-3') contains the binding site of Zif268. CD spectra of the ds DNA and Zf13 IDAO70 were recorded separately in phosphate buffer and as a complex in the presence of zinc (Figure 7). The CD spectrum of the Zf13 IDAO70–ds DNA complex shows a significant decrease in the negative maximum at 205 nm and an increased and shifted positive maximum at 278 nm, similar to the spectrum reported previously for DNA binding of the native Zif268.^[8b,22] DNA binding of the zinc finger Zf13 IDAO70 seems not to be affected by introduction of an additional metal binding site.

Conclusion

The semi-synthesis of the natural 90-amino-acid long sequence of the zinc-finger domain of Zif268 is described. Zif268 was modified in the third subunit by a semi-synthetic approach, using expressed protein ligation (EPL) of the chemically synthesized segment with the protein, obtained by recombinant protein expression. Six Zf13 variants were obtained containing additional metal-chelating side chains, a fluorophore, and an acetylene side chain, allowing further postsynthetic modifications by [2+3] cycloaddition. The integrity of the zinc-finger fold was confirmed by CD analysis of the modified zinc fingers. Furthermore, evidence for DNA binding of the full-length con-

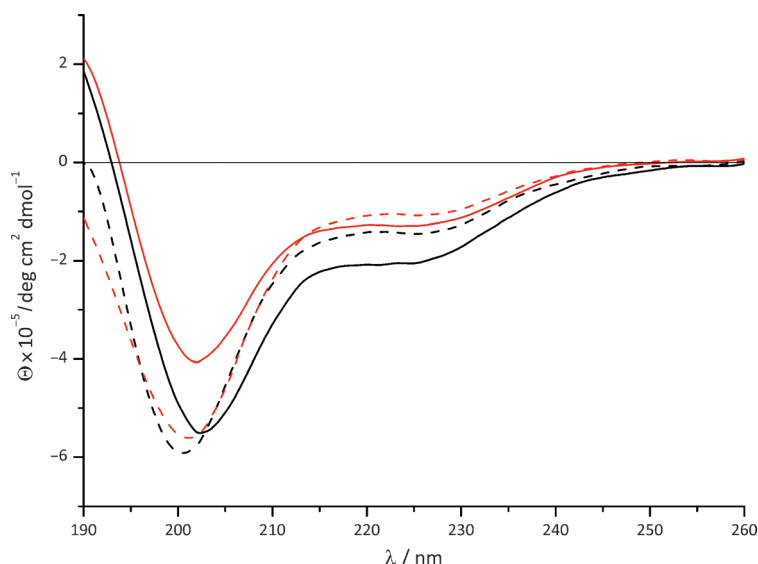


Figure 5. CD spectra of native Zf3 (17 μM; black) and modified Zf3 TACN85 (26 μM; red) in the absence (---) and presence (—) of 65 μM ZnSO₄.

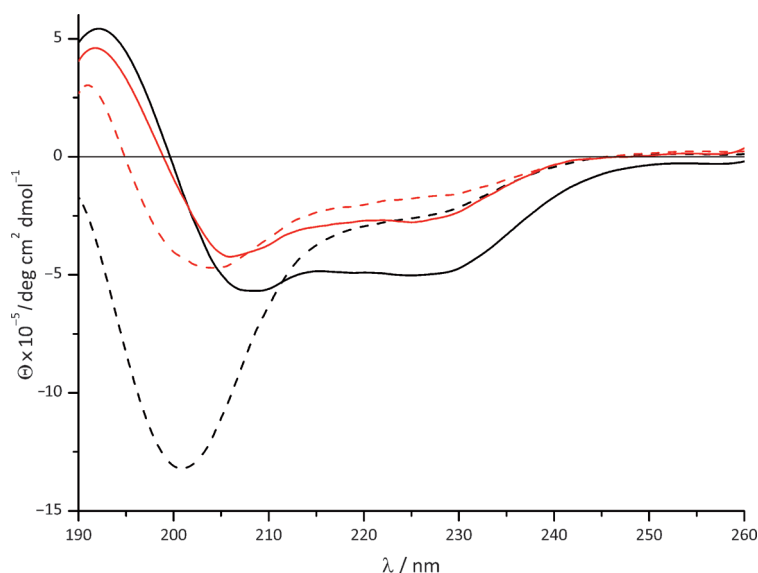


Figure 6. CD spectra of Zf13 IDAO70 (14 μM; black) and Zf13 TACN85 (8.5 μM; red) in the absence (---) and presence (—) of 130 μM Zn²⁺.

struct Zf13 IDAO70 was obtained by CD spectroscopy. Zif268 provides a scaffold for a selective DNA binding peptide that can be subjected to chemical modifications, such as the incorporation of additional metal binding sites or fluorophores without losing its DNA-binding potential.

Experimental Section

Materials: All *N*-α-Fmoc-protected amino acids were purchased from Nova Biochem (Darmstadt, Germany), GL Biochem (Shanghai, China), Fluka (Taufkirchen, Germany), Bachem (Bubendorf, Switzerland), Merck (Darmstadt, Germany), and IRIS Biotech (Marktredwitz,

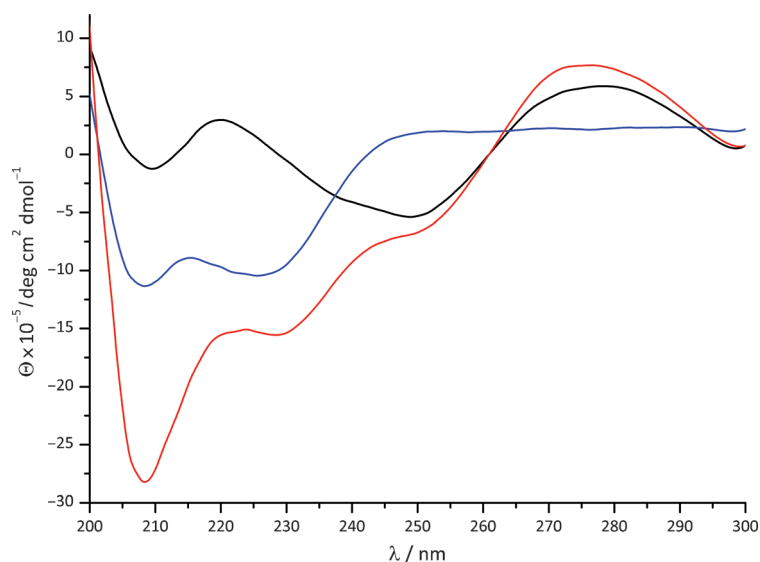


Figure 7. CD spectra of ds DNA (6.5 μM ; —), Zf13 IDAO70 (21 μM ; —), and the Zf13 IDAO70–ds DNA complex (3 μM ; —) in phosphate buffer in the presence of excess zinc (130 μM).

Germany). All other chemicals were purchased from Sigma–Aldrich (Taufkirchen, Germany), ABCR (Karlsruhe, Germany), Acros Organics (Geel, Belgium), Alfa Aesar (Karlsruhe, Germany), New England Biolabs (Ipswich, USA), Macherey&Nagel (Düren, Germany). Preloaded Gly or Lys-*p*-benzyloxybenzyl alcohol resin (Wang resin, 0.3 mmol g^{-1}), *N*-hydroxybenzotriazole (HOBt), *O*-benzotriazole-*N,N,N',N'*-tetramethyl-uronium-hexafluorophosphate (HBTU) were purchased from GL Biochem (Shanghai, China). *N,N*-Diisopropylcarbodiimide (DIC) and *N*-ethyldiisopropylamine (DIPEA) were obtained from IRIS Biotech (Marktredwitz, Germany). 5(6)-Carboxyfluorescein (CF), 2-mercaptoethane sulfonate sodium (MESNa) and *N*-methyl-2-pyrrolidinone (NMP) were purchased from Fluka (Taufkirchen, Germany). Trifluoroacetic acid (TFA) was bought from Roth (Karlsruhe, Germany).

ESI-MS: Data were obtained with a Finnigan instrument (type LCQ) or Bruker spectrometers (types Apex-Q IV 7T, HCT ultra, and micrOTOF API). High-resolution (HR) spectra were obtained with the Bruker Apex-Q IV 7T or the Bruker micrOTOF.

RP-HPLC: All RP-HPLC analyses were performed on instruments from GE Healthcare or Jasco. For all semi-preparative/preparative RP-HPLC, we used a Pharmacia Äkta basic device (GE Healthcare) with a gradient of eluent A (0.1% TFA in H_2O) to eluent B (0.1% TFA in MeCN/ H_2O 8:2) with a flow rate of 3 mL min^{-1} (semipreparative)/10 mL min^{-1} (preparative) were performed. Preparative purification was performed on a Phenomenex column Jupiter, RP-C18, 250 \times 20 mm, 5 μm , 80 Å. For semi-preparative purification, a Phenomenex column Jupiter, RP-C18, 250 \times 10 mm, 5 μm , 80 Å was used. For analytical RP-HPLC, a semi-micro-HPLC system from Jasco was used, applying a gradient of eluent A (0.1% TFA in H_2O) to eluent B (0.1% TFA in MeCN) with a flow rate of 1 mL min^{-1} . In these cases, a Phenomenex column RP-C18, 250 \times 4.6 mm, 5 μm was used. UV detection was performed at 215 nm, 254 nm and, for the fluorophore-labeled peptides, at 444 nm.

Circular dichroism (CD) spectroscopy: CD spectra were recorded on a Jasco-810 spectropolarimeter equipped with a Jasco PTC4325 temperature controller. Prior to usage, the sample cell was flushed with nitrogen. For CD spectra recording zinc-finger derivatives,

1 mm quartz glass precision cells were used. The peptide concentrations were adjusted to between 5–30 μM for all derivatives. The spectra were recorded at 20 $^\circ\text{C}$ in a wavelength range of 260–190 nm for the zinc fingers and from 300–200 nm for the DNA binding studies with 1.0 nm bandwidth in continuous mode, 1.0 s response and a scan speed of 100 nm min^{-1} . Eight spectra were averaged. Spectra were background-corrected, smoothed (Savitzky–Golay), and depicted as molar ellipticity ($\Theta \times 10^{-5} \text{ deg cm}^2 \text{ dmol}^{-1}$).

Cloning of Zf12: The cDNA encoding the murine transcription factor Zif268 (Egr1) was purchased from ATCC (Wesel, Germany). After amplifying the desired DNA fragment of the two zinc-finger domains (632–827) by polymerase chain reaction (PCR) using the forward $5' \text{GGTGGT-GCTAGCGAACGCCATATGCTTGCCCTG}^{3'}$ and reverse $5' \text{GGTGGTTGCTCTTCCGAGCTCTTGTCTTAAATGGAT-TTTG}^{3'}$ primers from Sigma–Aldrich (Taufkirchen, Germany), and purification by agarose-gel electrophoresis, the fragment was ligated into the cloning vector pGEM-T from Promega (Mannheim, Germany). After digestion of the cloning vector pGEM-T and expression vector pTXB1 from New England Biolabs (Ipswich, USA), using the restriction enzymes SapI and NheI, the DNA insert and vector were subsequently ligated. To confirm an in-frame cloning of Zf12 (1), DNA sequencing was performed.

Protein expression of Zf12 in *E. coli*: *E. coli* ER2566 cells were transformed with the pTXB1/Zf12 (632–827) DNA plasmid and grown in lysogeny broth (LB) or 2 \times YT medium containing carbenicillin (100 $\mu\text{g mL}^{-1}$). A preparatory culture (5 mL) was used to inoculate the expression culture. The cells were cultivated at 200 rpm at 37 $^\circ\text{C}$, until the culture reached an optical density at 600 nm (OD_{600}) of 0.6–0.8. The culture was then induced by adding 0.4 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and incubated again overnight at 16 $^\circ\text{C}$. The cells were harvested by centrifugation at 4 $^\circ\text{C}$ at 9000 *g* and lysed in buffer A: 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic (HEPES) buffer (20 mM), NaCl (500 mM), tris(2-carboxyethyl)phosphine (0.1 mM), Tween20 (0.1%), pH 8, and 20 μM phenylmethylsulfonyl fluoride (PMSF) was added before incubating on ice for 30 min. Complete lysis was accomplished by pulsed sonication (5 \times 45 s). The soluble extracts of the protein were isolated by centrifugation at 23 000 *g* for 20 min at 4 $^\circ\text{C}$. Purification and isolation of Zf12 (1) was enabled by an attached chitin binding domain (CBD) as follows: a column (GE Healthcare, XK 26/20) filled with 100 mL chitin beads (New England Biolabs) was equilibrated at 4 $^\circ\text{C}$ with water (2 \times bed volumes) and buffer A (10 \times bed volumes). Loading of the cell lysate proceeded overnight at a flow rate of 1 mL min^{-1} (10 times). Unspecifically bound *E. coli* proteins were removed by washing with buffer A (3 \times bed volumes). The intein fusion protein of Zf12 was cleaved in the presence of buffer B: HEPES (20 mM), MESNa (250 mM, pH 6) over 70 h at 4 $^\circ\text{C}$. The peptide thioester was eluted with buffer A. The complete procedure was monitored at 4 $^\circ\text{C}$ by UV absorption (280 nm, Äkta Prime Plus, GE Healthcare). Analysis of the peptide thioester was carried out by SDS-PAGE (tris-tricine 10–20%), RP-HPLC on a C18 semi-preparative column (Phenomenex column Jupiter RP-C18, 250 \times 10 mm, 5 μm , 80 Å) with 0.1% TFA in H_2O (A) and 0.1% TFA in $\text{H}_2\text{O}/\text{MeCN}$ (20:80) (B) as eluting system with a 20 \rightarrow 50% gradient of B over 30 min at a flow rate of 3 mL min^{-1} , and HRMS-ESI.

Reduced Zf12 (1): $\text{H}_2\text{N-ASERPYPACPVESCDRRFRSDELTRHIR IHTGQKPFQCRICMRNFRSDHLTHTHRTHTGKPFPA-OH}$; $M_r =$

7794.8 ($C_{328}H_{520}N_{110}O_{98}S_7$); MS (ESI): $m/z = 650.57 [M+12H]^{12+}$, 709.62 $[M+11H]^{11+}$, 780.48 $[M+10H]^{10+}$; HRMS-ESI: m/z calcd for $C_{328}H_{520}N_{110}O_{98}S_7$: 709.3348 $[M+11H]^{11+}$, found: 709.3460; 780.1786 $[M+10H]^{10+}$, found: 780.1788.

SPPS of Zf3 and respective modifications: The peptides were synthesized by standard Fmoc/*tert*-butyl peptide synthesis on solid support on preloaded Wang resin, 0.3 mmol g^{-1} using a microwave-supported automated peptide synthesizer LibertyTM (CEM Cooperation, Matthews, NC, USA). The side chain protection groups were *tert*-butyl for aspartic acid, glutamic acid, serine, tyrosine; *tert*-butyloxycarbonyl (Boc) or allyloxycarbonyl (alloc) for lysine; triphenylmethyl (trityl) for cysteine, glutamine and histidine; and 2,2,4,6,7-pentamethyl-dihydrofurane-5-sulfonyl (PBF) for arginine. Resins were swollen in CH_2Cl_2 for 30 min. The coupling protocol was adjusted to the different needs of the amino acids coupled. Fmoc amino acids were used as 0.2 M solutions in *N*-methylpyrrolidone (NMP). Coupling was performed with 0.5 M HBTU/HOBt (5 equiv) in DMF, 0.2 M amino acids in NMP (5 equiv) and 2 M DIPEA (10 equiv) in NMP. Double coupling of the amino acids arginine, cysteine and aspartic acid was performed to enhance the coupling efficiency. Capping was performed after every coupling cycle using a solution of acetic anhydride (10%), DIPEA (5%), and HOBt (0.2%) in NMP. All reaction steps were performed under microwave conditions (20 W, 300 s at 75 °C) and N_2 mixing. Arginine was coupled under N_2 mixing (1500 s) at RT followed by microwave-assisted coupling (300 s) at 75 °C. To minimize racemization during coupling, cysteine derivatives were coupled for 300 s at 50 °C. Histidine was coupled without using the microwave supply for 60 min.^[23] The Fmoc-protecting group was removed by two deprotection cycles with 20% piperidine in NMP (30 s, 180 s). Final deprotection of the Fmoc group was performed for all peptides except the orthogonal Lys(alloc)-protected derivatives required for further fluorophore coupling. Prior to cleavage, the resins were washed with DMF/ CH_2Cl_2 (1:1, 5 × 5 mL), CH_2Cl_2 (3 × 5 mL), EtOH (3 × 5 mL), and CH_2Cl_2 (5 × 5 mL) and dried overnight. Cleavage from the resin and simultaneous removal of all side chain protecting groups was performed separately for all derivatives in a 10 mL syringe (Becton, Dickinson & Co., Franklin Lakes, NJ, USA) by adding TFA/ H_2O /triethylsilyl/1,2-ethanedithiol ($v/v/v/v = 95:2:1:2$, 10 mL g^{-1} resin) and shaking for 2 h at RT. The peptides were precipitated from ice-cold *tert*-butylmethylether, isolated by centrifugation (9000 rpm, 4 °C) and washed with ice-cold Et_2O (3 ×); the pellet was subsequently lyophilized from H_2O using a freeze dryer (Martin Christ GmbH, Osterode am Harz, Germany). The lyophilized peptides were purified by semi-preparative/preparative RP-HPLC on C18 columns (Phenomenex column Jupiter RP-C18, 250 × 10 mm, 5 μm , 80 Å with 0.1% TFA in H_2O (A) and 0.1% TFA in H_2O /MeCN (20:80) (B) as eluting system with a 20–50% gradient of B over 30 min at a flow rate of 3/10 mL min^{-1} and lyophilized again. The peptides were analyzed by HRMS-ESI.

Introduction of modifications: All modifications were accessible within the SPPS protocol. In the case of the attachment of the fluorophore instead of Fmoc-Cys(Trt)-OH the complementary Boc derivative was coupled as N-terminal amino acid. 5(6)-Carboxyfluorescein was introduced by standard coupling to the side chain of Lys89, which was orthogonally protected by an alloc group. Prior to cleavage from the resin, the alloc group was removed by treatment with $MeNH BH_3$ (30 equiv) and $[Pd(PPh_3)_4]$ (0.1 equiv) over 4 h in DMF at RT. 5(6)-Carboxyfluorescein was activated with HOBt and DIC, and coupled to the lysine side chain over 2 h at RT. Arg70 was exchanged for the synthesized artificial amino acid Fmoc-Pra-OH to allow Click chemistry and also for the artificial metal-chelat-

ing amino acid Fmoc-IDAO-OH, which was coupled manually using HOAt/HBTU coupling reagents. His85 was exchanged by the metal-chelating amino acid Fmoc-TACN-OH.^[12]

Analysis of Zf3 and its modifications: The purified reduced peptides were analyzed by RP-HPLC on C18 semi-preparative column (Phenomenex column RP-C18, 250 × 10 mm, 5 μm), with 0.1% TFA in H_2O (A) and 0.1% TFA in H_2O /MeCN (20:80) (B) as eluting system at a flow rate of 3 mL min^{-1} , and HRMS-ESI.

Zf3: H_2N -CDICGRKFARSDEKRRHTKIHRLRQK-OH; $M_r = 3137.69$ ($C_{131}H_{222}N_{50}O_{36}S_2$); RP-HPLC: gradient B: 20 → 50% over 30 min, $t_R = 12.0$ min; MS (ESI): $m/z = 785.4 [M+4H]^{4+}$, 1046.7 $[M+3H]^{3+}$; HRMS-ESI: m/z calcd for $C_{131}H_{222}H_{50}O_{36}S_2$: 523.7830 $[M+4H]^{4+}$, found: 523.7832; calcd: 628.3382 $[M+5H]^{5+}$, found: 628.3388; calcd: 785.1709 $[M+4H]^{4+}$, found: 785.1712.

Zf3 A85 (2): H_2N -CDICGRKFARSDEKRRHTKIALRQKDG-OH; $M_r = 3188.71$ ($C_{132}H_{225}N_{49}O_{39}S_2$); RP-HPLC: gradient B: 20 → 50% over 30 min, $t_R = 12.76$ min; MS (ESI): $m/z = 532.12 [M+6H]^{6+}$, 638.14 $[M+5H]^{5+}$, 797.42 $[M+4H]^{4+}$, 1063.23 $[M+3H]^{3+}$; HRMS-ESI: m/z calcd for $C_{132}H_{225}H_{49}O_{39}S_2$: 638.1392 $[M+5H]^{5+}$, found: 638.1391.

Zf3 A85 I89 (3): H_2N -CDICGRKFARSDEKRRHTKIALRQK (FL) DG-OH; $M_r = 3547.02$ ($C_{153}H_{235}N_{49}O_{45}S_2$); RP-HPLC: gradient B: 20 → 50% over 30 min, $t_R = 19.83$ min; MS (ESI): $m/z = 591.63 [M+6H]^{6+}$, 709.75 $[M+5H]^{5+}$, 886.94 $[M+4H]^{4+}$, 1182.25 $[M+3H]^{3+}$; HRMS-ESI: m/z calcd for $C_{153}H_{235}H_{49}O_{45}S_2$: 709.9493 $[M+5H]^{5+}$, found: 709.9492; calcd: 886.6835 $[M+4H]^{4+}$, found: 886.6832.

Zf3 Pra70 (4): H_2N -CDICGPRAKFARSDEKRRHTKIHRLRQKDG-OH; $M_r = 3193.69$ ($C_{134}H_{220}N_{48}O_{39}S_2$); RP-HPLC: gradient B: 20 → 50% over 30 min, $t_R = 10.37$ min; MS (ESI): $m/z = 532.95 [M+6H]^{6+}$, 639.34 $[M+5H]^{5+}$, 798.92 $[M+4H]^{4+}$, 1064.88 $[M+3H]^{3+}$; HRMS-ESI: m/z calcd for $C_{134}H_{220}H_{48}O_{39}S_2$: 639.1308 $[M+5H]^{5+}$, found: 639.1309; calcd: 798.6617 $[M+4H]^{4+}$, found: 798.6614.

Zf3 TACN85 (5): H_2N -CDICGRKFARSDEKRRHTKITACNLRQK-OH; $M_r = 3234.8$ ($C_{136}H_{235}N_{53}O_{35}S_2$); RP-HPLC: gradient B: 5 → 40% over 30 min, $t_R = 21.85$ min; MS (ESI): $m/z = 648.36 [M+5H]^{5+}$, 809.95 $[M+4H]^{4+}$, 1079.27 $[M+3H]^{3+}$.

Zf3 IDAO70 (6): H_2N -CDICGIDAOKFARSDEKRRHTKIHRLRQK-OH; $M_r = 3154.67$; RP-HPLC: gradient B: 20 → 50% over 30 min, $t_R = 13.5$ min; MS (ESI): $m/z = 631.94 [M+5H]^{5+}$, 789.67 $[M+4H]^{4+}$, 1052.55 $[M+3H]^{3+}$; HRMS-ESI: m/z calcd for $C_{132}H_{219}H_{47}O_{39}S_2$: 631.5312 $[M+5H]^{5+}$, found: 631.5316; calcd: 789.1622 $[M+4H]^{4+}$, found: 789.1618.

Ligation of Zf13 and its modifications: The N-terminal Cys65 of the different Zf3 peptides had to be available for nucleophilic attack at the thioester of **1** during the ligation step. Therefore, the disulfide bridges in Zf3 and the other constructs were first reduced. The peptide (3 mM) was dissolved in phosphate buffer (10 mM, pH 4, TCEP 20 mM) and stirred overnight at RT under inert conditions. Complete reduction of the peptide was confirmed by ESI-MS. Subsequently, peptide thioester **1** (1 mM) was added in phosphate buffer (500 mM, pH 8) and the pH of the solution was adjusted to pH 7.8–8. As determined by RP-HPLC, the conversion of the two peptides into the ligation product was observed after stirring for 6–12 h at RT under inert conditions.

Analysis of the ligation products: The ligation products were analyzed by RP-HPLC on an analytical C18 column (Phenomenex column RP-C18, 250 × 4.6 mm, 5 μm) with 0.1% TFA in H_2O (A) and 0.1% TFA in MeCN (B) as eluting system with a 20–50% gradient of B over 30 min at a flow rate of 1 mL min^{-1} and ESI-MS.

Zf13: H₂N-ASERPYPVESCDDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLLTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLRQKG-OH; M_r = 10794.4 (C₄₅₇H₇₃₂N₁₆₀O₁₃₁S₇); RP-HPLC: gradient B: 20→50 in 30 min, t_R = 18.1 min; ESI-MS: m/z = 721.3 [M+15H]¹⁵⁺, 771.9 [M+14H]¹⁴⁺, 830.9 [M+13H]¹³⁺, 900.1 [M+12H]¹²⁺, 981.7 [M+11H]¹¹⁺, 1080.0 [M+10H]¹⁰⁺.

Zf13 A85: H₂N-ASERPYPVESCDDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLLTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLRQKDG-OH; M_r = 10843.4 (C₄₅₈H₇₄₁N₁₅₉O₁₃₄S₇); RP-HPLC: gradient B 20→40% over 40 min, flow rate = 0.3 mL min⁻¹, t_R = 27.5 min; MS (ESI): m/z = 603.3 [M+18H]¹⁸⁺, 678.4 [M+16H]¹⁶⁺, 723.5 [M+15H]¹⁵⁺, 834.8 [M+13H]¹³⁺, 904.4 [M+12H]¹²⁺, 986.7 [M+11H]¹¹⁺, 1085.2 [M+10H]¹⁰⁺, 1205.7 [M+9H]⁹⁺, 1401.1 [M+8H]⁸⁺, 1356.3 [M+7H]⁷⁺; HRMS: m/z calcd for C₄₅₈H₇₄₁N₁₅₉O₁₃₄S₇: 678.3456 [M+16H]¹⁶⁺, found: 678.3536; calcd: 1085.248 [M+10H]¹⁰⁺, found: 1085.237; calcd: 1205.720 [M+9H]⁹⁺, found: 1205.703; calcd: 1356.309 [M+7H]⁷⁺, found: 1356.291.

Zf13 A85 L89: H₂N-ASERPYPVESCDDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLLTTHIRTHTGEKPFACDICGRKFARSDERKRHTKIHLRQK (FL) DG-OH; M_r = 11215.5 (C₄₈₀H₇₅₃N₁₅₉O₁₄₀S₇); RP-HPLC: gradient B: 20→50% over 30 min, t_R = 18.62 min; MS (ESI): m/z = 701.3 [M+16H]¹⁶⁺, 747.9 [M+15H]¹⁵⁺, 801.3 [M+14H]¹⁴⁺, 862.8 [M+13H]¹³⁺, 934.6 [M+12H]¹²⁺, 1019.5 [M+11H]¹¹⁺, 1121.2 [M+10H]¹⁰⁺, 1245.6 [M+9H]⁹⁺, 1401.1 [M+8H]⁸⁺, 1601.1 [M+7H]⁷⁺, 1867.6 [M+6H]⁶⁺.

Zf13 PRA70: H₂N-ASERPYPVESCDDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLLTTHIRTHTGEKPFACDICGPRAKFARSDERKRHTKIHLRQKDG-OH; M_r = 10848.4 (C₄₆₀H₇₃₆N₁₅₈O₁₃₄S₇); RP-HPLC: gradient B: 27→50% over 30 min, t_R = 10.56 min; MS (ESI): m/z = 640.1 [M+17H]¹⁷⁺, 679.8 [M+16H]¹⁶⁺, 725.0 [M+15H]¹⁵⁺, 776.8 [M+14H]¹⁴⁺, 836.3 [M+13H]¹³⁺, 905.5 [M+12H]¹²⁺, 987.3 [M+11H]¹¹⁺, 1085.9 [M+10H]¹⁰⁺.

Zf13 TACN85: H₂N-ASERPYPVESCDDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLLTTHIRTHTGEKPFACDICGRKFARSDERKRHTKITACNLRQK-OH; M_r = 10893.5 (C₄₆₂H₇₅₁N₁₆₃O₁₃₀S₇); RP-HPLC: gradient B: 20→50% over 30 min, t_R = 18.9 min; MS (ESI): m/z = 574.5 [M+19H]¹⁹⁺, 606.4 [M+18H]¹⁸⁺, 642.0 [M+17H]¹⁷⁺, 682.0 [M+16H]¹⁶⁺, 727.4 [M+15H]¹⁵⁺, 779.3 [M+14H]¹⁴⁺, 891.1 [M+13H]¹³⁺, 908.9 [M+12H]¹²⁺, 991.4 [M+11H]¹¹⁺, 1090.4 [M+10H]¹⁰⁺.

Zf13 IDAO70: H₂N-ASERPYPVESCDDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLLTTHIRTHTGEKPFACDICGIDAOKFARSDERKRHTKIHLRQK-OH; M_r = 10811.4 (C₄₅₈H₇₃₇N₁₅₇O₁₃₄S₇); RP-HPLC: gradient B: 20→50% over 30 min, t_R = 18.74 min; MS (ESI): m/z = 570.2 [M+19H]¹⁹⁺, 601.8 [M+18H]¹⁸⁺, 637.1 [M+17H]¹⁷⁺, 676.9 [M+16H]¹⁶⁺, 721.9 [M+15H]¹⁵⁺, 773.4 [M+14H]¹⁴⁺, 832.8 [M+13H]¹³⁺, 902.1 [M+12H]¹²⁺, 983.9 [M+11H]¹¹⁺, 1082.2 [M+10H]¹⁰⁺.

Acknowledgements

Generous support from the Deutsche Forschungsgemeinschaft (DFG) (IRTG 1422) is gratefully acknowledged.

Keywords: expressed protein ligation • metalloproteins • peptide synthesis • proteins • zinc-finger domains

- [1] a) A. Klug, *J. Mol. Biol.* **1999**, *293*, 215–218; b) P. H. von Hippel, J. D. McGhee, *Annu. Rev. Biochem.* **1972**, *41*, 231–300; c) P. J. Mitchell, R. Tijan, *Science* **1989**, *245*, 371–378; d) J. D. McGhee, P. H. von Hippel, *J. Mol. Biol.* **1974**, *86*, 469–489.
- [2] G. M. Rubin, M. D. Yandell, J. R. Wortman, G. L. Gabor, Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, P. W. Li, R. Apweiler, W. Fleischmann, J. M. Cherry, S. Henikoff, M. P. Skupski, S. Misra, M. Ashburner, E. Birney, M. S. Boguski, T. Brody, P. Brokstein, et al., *Science* **2000**, *287*, 2204–2215.
- [3] J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, Mark Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, Peter Amanatides, R. M. Ballew, D. H. Huson, Jennifer Russo Wortman, Qing Zhang, C. D. Kodira, X. H. Zheng, Lin Chen, Marian Skupski, et al., *Science* **2001**, *291*, 1304–1351.
- [4] a) H. Arakawa, H. Nagase, M. Ogawa, M. Nagata, T. Fujiwara, E. Takahashi, S. Shin, Y. Nakamura, *Cytogenet. Cell Genet.* **1995**, *70*, 235–238; b) J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **1985**, *4*, 1609–1614.
- [5] a) S. S. Krishna, *Nucleic Acids Res.* **2003**, *31*, 532–550; b) S. Iuchi, *Cell. Mol. Life Sci.* **2001**, *58*, 625–635.
- [6] M. Elrod-Erickson, M. A. Rould, L. Nekudova, C. O. Pabo, *Structure* **1996**, *4*, 1171–1180.
- [7] A. I. Anzellotti, N. P. Farrell, *Chem. Soc. Rev.* **2008**, *37*, 1629–1651.
- [8] a) N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809–817; b) M. Elrod-Erickson, C. O. Pabo, *J. Biol. Chem.* **1999**, *274*, 19281–19285.
- [9] a) J. Matthews, F. Loughlin, J. Mackay, *Curr. Opin. Struct. Biol.* **2008**, *18*, 484–490; b) A. Nadler, C. Hain, U. Diederichsen, *Eur. J. Org. Chem.* **2009**, 4593–4599.
- [10] W. N. Lipscomb, N. Sträter, *Chem. Rev.* **1996**, *96*, 2375–2433.
- [11] a) S. Negi, M. Imanishi, M. Matsumoto, Y. Sugiura, *Chem. Eur. J.* **2008**, *14*, 3236–3249; b) K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson, *Nature* **2009**, *460*, 823–830; c) Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855–862.
- [12] a) G. Guillena, G. Rodriguez, G. van Koten, *Tetrahedron Lett.* **2002**, *43*, 3895–3898; b) K. Johnsson, R. K. Allemann, H. Widmer, S. A. Benner, *Nature* **1999**, *365*, 530–532.
- [13] a) K. L. Harris, S. Lim, S. J. Franklin, *Inorg. Chem.* **2006**, *45*, 10002–10012; b) H. W. Hellinga, *Curr. Opin. Struct. Biotech.* **1996**, *7*, 437–441; c) Y. Lu, S. M. Berry, T. D. Pfister, *Chem. Rev.* **2001**, *101*, 3047–3080; d) W. DeGrado, C. Summa, V. Pavone, F. Nastro, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
- [14] pTXB1 vector was purchased from New England Biolabs (Frankfurt, Germany), catalogue no. N67075.
- [15] H. Schägger, *Nat. Protoc.* **2006**, *1*, 16–22.
- [16] a) A. Nomura, Y. Sugiura, *Inorg. Chem.* **2002**, *41*, 3693–3698; b) A. Nomura, Y. Sugiura, *Inorg. Chem.* **2004**, *43*, 1708–1713; c) R. J. Simpson, E. D. Cram, R. Czolij, J. M. Matthews, M. Crossley, J. P. Mackay, *J. Biol. Chem.* **2003**, *278*, 28011–28018.
- [17] a) E. A. Jares-Erijman, T. M. Jovin, *Nat. Biotechnol.* **2003**, *21*, 1387–1395; b) A. Giannetti, L. Citti, C. Domenici, L. Tedeschi, F. Baldini, M. B. Wabueyle, T. Vo-Dinh, *Sens. Actuators B* **2006**, *113*, 649–654.
- [18] a) C. W. Tornøe, C. C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064; b) H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2001**, *113*, 2056–2075; *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021; c) T. L. Mindt, H. Struthers, L. Brans, T. Anguelov, C. Schweinsberg, V. Maes, D. Tourwe, R. Schibli, *J. Am. Chem. Soc.* **2006**, *128*, 15096–15097.
- [19] M. Gajewski, B. Seaver, C. S. Esslinger, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4163–4166.
- [20] P. E. Dawson, T. W. Muir, I. C. Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
- [21] a) A. D. Frankel, J. M. Berg, C. O. Pabo, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4841–4845; b) S. Negi, M. Itazu, M. Imanishi, A. Nomura, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **2004**, *325*, 421–425; c) E. Kopera, T. Schwerdtle, A. Hartwig, W. Bal, *Chem. Res. Toxicol.* **2004**, *17*, 1452–1458.
- [22] G. S. Beligere, P. E. Dawson, *Biopolymers* **1999**, *51*, 363–369.
- [23] B. Bacsa, S. Bösze, C. O. Kappe, *J. Org. Chem.* **2010**, *75*, 2103–2106.

Received: September 7, 2011

Published online on January 2, 2012