Protein Engineering, Design & Selection vol. 27 no. 12 pp. 481–488, 2014
Published online October 21, 2014 doi:10.1093/protein/gzu048

An optimized intein-mediated protein ligation approach for the efficient cyclization of cysteine-rich proteins

Katsiaryna Tarasava[1] and Eva Freisinger[1,2]

[1]Department of Chemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[2]To whom correspondence should be addressed.
E-mail: freisinger@chem.uzh.ch

Received March 14, 2014; revised September 16, 2014; accepted September 26, 2014

Edited by Niels Andersen

Head-to-tail backbone cyclization of proteins is a widely used approach for the improvement of protein stability. One way to obtain cyclic proteins via recombinant expression makes use of engineered Intein tags, which are self-cleaving protein domains. In this approach, pH-induced self-cleavage of the N-terminal Intein tag generates an N-terminal cysteine residue at the target protein, which then attacks in an intramolecular reaction the C-terminal thioester formed by the second C-terminal Intein tag resulting in the release of the cyclic target protein. In the current work we aimed to produce a cyclic analog of the small γ-E_c-1 domain of the wheat metallothionein, which contains six cysteine residues. During the purification process we faced several challenges, among them premature cleavage of one or the other Intein tag resulting in decreasing yields and contamination with linear species. To improve efficiency of the system we applied a number of optimizations such as the introduction of a Tobacco etch virus cleavage site and an additional poly-histidine tag. Our efforts resulted in the production of a cyclic protein in moderate yields without any contamination with linear protein species.

Keywords: cyclic proteins/cysteine/intein tags/metallothionein/TEV protease

Introduction

In backbone cyclized or circular proteins, the N- and C-termini are covalently joined by an amide bond. This protein class is widely found in nature (Trabi and Craik, 2002). The largest family of naturally occurring circular peptides are the well-known cyclotides, which are mainly observed in plants (Gran, 1973; Gustafson *et al.*, 1994; Witherup *et al.*, 1994; Saether *et al.*, 1995). More recently, a number of cyclic bacterial proteins distinct from cyclotides were discovered (Langdon *et al.*, 1998; Blond *et al.*, 1999; Gonzalez *et al.*, 2000; Abriouel *et al.*, 2001; Rebuffat *et al.*, 2004). Circular peptides can have numerous functions showing for example uterotonic (Gran, 1973), anti-HIV (Gustafson *et al.*, 1994), antimicrobial (Tam *et al.*, 1999*b*), and insecticidal activity (Wang *et al.*, 2009). The cyclized backbone is often crucial for protein functionality with the linear analogs showing decreased or even a loss of activity (Blond *et al.*, 1999) and increases the in vivo stability of circular peptides against proteolytic degradation (Iwai and Plückthun, 1999) as well as the stability against certain chemicals (Bogdanowich-Knipp *et al.*, 1999) and heat denaturation *in vitro* (Gonzalez *et al.*, 2000; Zhao *et al.*, 2010). An additional benefit for many applications is the reduced flexibility of the protein structure imposed by the fixation of the protein termini within the cycle, as protein ends usually have a higher entropy than the protein core. This can for example improve receptor binding affinities of circular proteins in comparison to the corresponding linear proteins (Luckett *et al.*, 1999; Korsinczky *et al.*, 2001). The versatile properties of circular peptides make them attractive for the development of, e.g. drugs or pesticides and hence everywhere where stability and target binding affinity are major determinants of efficacy (Craik *et al.*, 2002; Armishaw *et al.*, 2011). In addition, the generation of circular protein analogs provides a tool for manipulation and understanding of protein structure and folding mechanism as well as the investigation of splicing reactions (Camarero *et al.*, 2001; Heitz *et al.*, 2008).

Different synthetic strategies for the production of cyclic proteins can be found in the literature (White and Yudin, 2011) among them native chemical ligation and traceless Staudinger ligation that both afford head-to-tail cyclization via a native amide bond. Native chemical ligation uses an unprotected synthetic linear peptide with a C-terminal α-thioester and an N-terminal Cys residue that is required for the intramolecular transthioesterification reaction to yield the cyclic intermediate (Dawson *et al.*, 1994). The native amide bond is subsequently formed after a spontaneous S-N acyl rearrangement. The traceless Staudinger ligation abolishes the need for an N-terminal Cys residue and thus is applicable to any amino acid sequence (Tam and Raines, 2009). For the reaction, a bifunctional phosphinothiol is coupled to the C-terminus of a synthetic peptide resulting in a thioester and a free phosphino group that can undergo a Staudinger reaction with another peptide carrying an N-terminal azide group. The formed iminophosphorane intermediate further reacts with the thioester to afford the native amide bond after hydrolysis. Initially applied for the ligation of two peptide fragments, the traceless Staudinger ligation was recently also applied for the synthesis of cyclic peptides (Ha *et al.*, 2012).

Also recombinant techniques can be used for the preparation of cyclic peptides and proteins in moderate yields, which is especially attractive for larger ring sizes that are not accessible by synthetic methods. The main methods are expressed protein ligation (EPL) (Xu and Evans, 2001; Muir, 2003), protein trans splicing (PTS) (Tavassoli and Benkovic, 2007), trans-peptidation (Parthasarathy *et al.*, 2007) and genetic reprogramming (Ohta *et al.*, 2008). The first two approaches are most commonly used and are based on the use of engineered inteins. Inteins are self-cleaving protein domains involved in protein splicing. The mechanism of intein-mediated splicing includes four conserved steps: a spontaneous and reversible N→S acyl shift to a thioester involving the N-terminal Cys

residue of the intein, a reversible transesterification resulting from the attack of a second Cys located just next to the C-terminus of the intein on the thioester, intein excision due to succinimide formation at the intein C-terminal asparagine residue, and finally formation of a native amide bond between the two protein parts by a S→N acyl shift at the thioester (Xu and Perler, 1996; Xu and Evans, 2001). In nature this intein self-cleavage occurs spontaneously and does not require additional cofactors (Chong *et al.*, 1996). For the production of cyclic peptides a commercially available system, IMPACT^{TM}-TWIN (New England Biolabs, Inc., MA, USA), has been designed, in which the respective protein or peptide is sandwiched between two specifically engineered inteins, i.e. the N-terminal Ssp DnaB intein (intein-1) (Mathys *et al.*, 1999) and the C-terminal Mth RIR1 intein (intein-2) (Evans *et al.*, 1999*b*). To suppress self-cleavage of the inteins from the cyclic protein precursor during expression and initial purification, intein-1 lacks its N-terminal Cys residue while intein-2 is devoid of its C-terminal Asn. After transesterification of the thioester generated by N→S acyl shift at the N-terminus of intein-2 with small thiol compounds such as 1,4-dithiothreitol (DTT) or 2-mercaptoethanesulfonic acid (MESNA) and cleavage of intein-1 via temperature- or pH-induced succinimide formation at its C-terminal Asn residue, a linear protein intermediate with a N-terminal Cys residue and a C-terminal activated thioester is obtained, which undergoes fast intramolecular cyclization (Xu and Evans, 2001). Instead of intein-1 a protease cleavage site can also be used to generate the N-terminal Cys residue at the protein intermediate required for the ligation reaction (Camarero and Muir, 1999). Chitin binding domains (CBD) (Watanabe *et al.*, 1994) joined to each intein facilitate the affinity purification of the precursor protein by using a chitin resin.

However, there are several reports on this system that show a surprisingly low protein yield or no protein production at all (Mathys *et al.*, 1999; Wood *et al.*, 2000; Cui *et al.*, 2006). The main proposed reason is premature intein cleavage during protein expression, probably provoked by cellular thiol compounds. For some inteins, the C-terminal thioester formation is also strongly dependent on the nature of the C-terminal residue of the target protein. (Xu *et al.*, 2000).

The research focus of our laboratory is on the investigation of metallothioneins (MTs). MTs are small proteins with a high percentage of Cys residues allowing them to coordinate metal ions via formation of metal-thiolate clusters (Vasak *et al.*, 1981; Good and Vasak, 1986*a*,*b*; Messerle *et al.*, 1992; Freisinger, 2010). MTs are abundant in all phylae of life and play roles in metal ion tolerance and homeostasis. Recent reports show that MTs can also participate in the scavenging of reactive oxygen species (Gobel *et al.*, 2000; Colangelo *et al.*, 2004). In its metal-free form MTs are generally proposed to adopt a highly flexible random coil structure mainly due to the absence of secondary structural elements (Boulanger *et al.*, 1983) emphasizing the importance of metal ion binding for the formation of distinct three-dimensional structures. Our attempts to produce backbone cyclized MT species were driven by mainly two objectives, the further stabilization of the three-dimensional structure by reducing the flexibility of the protein termini in order to facilitate crystallization and structure solution by nuclear magnetic resonance (NMR) as well as enhancement of metal ion binding affinity and specificity by reduction of metal ion binding site flexibility.

In this work we describe the production of a backbone cyclized form of the γ-domain from the seed-specific E_c-1 MT isolated from *Triticum aestivum* (common bread wheat). This protein is assumed to participate in Zn(II) storage for the germination process (Hanleybowdoin and Lane, 1983; Lane *et al.*, 1987) and forms two metal ion binding domains, i.e. γ- and β_E-E_c-1, which coordinate two and four Zn(II) ions, respectively (Peroza *et al.*, 2009*a*), and whose structures have been determined with NMR spectroscopy (Peroza *et al.*, 2009*b*; Loebus *et al.*, 2011). The two metal ions in γ-E_c-1 are arranged in a Zn_2Cys_6 cluster (Loebus *et al.*, 2011). To achieve backbone cyclization, we used a modified version of the IMPACT^{TM}-TWIN system that we optimized for our purpose and show that the high Cys content of MTs is not a hurdle for efficient cyclization.

Materials and methods

Vector construction

A nucleotide sequence consisting of a 5′ NdeI restriction site, a sequence encoding for a *Tobacco etch virus* (TEV) protease cleavage site, the *Escherichia coli* codon usage optimized γ-E_c-1 sequence (Peroza and Freisinger, 2007), and a 3′ SpeI restriction site was constructed *de novo* from commercially obtained primer sequences in three PCR steps (see Supplementary Material). In addition, a second sequence including the codons for a N-terminal 6xHis tag was prepared from the first construct using an additional primer.

Both final PCR products were digested with NdeI and SpeI and ligated into the double digested and dephosphorylated pTWIN2 vector (New England Biolabs Inc., Supplementary Fig. S1). In this way plasmids lacking the N-terminal intein-1 are obtained, i.e. plasmid pTWIN2cut-gEc1 encoding for the TEV-γE_c1-intein2-CBD construct and plasmid pTWIN2cut-6H-gEc1 encoding for the 6xHis-TEV-γE_c1-intein2-CBD. The original pTWIN2 plasmid and all restriction enzymes were from New England Biolabs, Inc.

Expression of fusion proteins

Escherichia coli (BL 21 line) cells transformed with plasmids pTWIN2cut-gEc1 or pTWIN2cut-6H-gEc1 were grown in Luria-Bertani (LB) medium (Roth AG, Arlesheim, Switzerland) supplemented with 100 μg ml^{-1} ampicillin (Roth AG) at 37°C. Protein expression was induced at OD_{600} ~0.6 with 500 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Biosolve Chimie, Dieuze, France) and media were supplemented with $ZnCl_2$ to a final concentration of 100 μM. Cells were harvested after 3 h at 30°C by centrifugation (4000 r/min, 15 min) and the pellets were frozen in liquid nitrogen for lysis and future purification.

Purification of the 6xHis-TEV-γE_c1-intein2-CBD construct

Cell pellets were resuspended in Buffer A (20 mM Na phosphate, pH = 7.2, 300 mM NaCl, 5% glycerol) and lysed by sonication. Cell debris was removed by centrifugation and filtration through a 0.22 μm filter, and the soluble cell extract was applied to a HisTrap HP column (GE Healthcare Europe GmbH, Glattbruck, Switzerland) that was pre-equilibrated with Buffer A. After washing with Buffer A supplemented with 30 mM imidazole, the fusion protein was eluted with Buffer A containing 400 mM imidazole. The elution fraction

was supplemented with 0.3 mM tris(2-carboxyethyl)phosphine (TCEP) and immediately subjected to a desalting and buffer exchange step with size-exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) that was pre-equilibrated with Buffer B (20 mM Na phosphate, pH = 7.2, 100 mM NaCl). Fractions containing the fusion construct were combined and used for TEV protease cleavage.

Fusion protein cleavage with TEV protease

The purified 6xHis-TEV-γE_c1-intein-2-CBD fusion protein was incubated with TEV protease (expressed and purified according to (Tropea *et al.*, 2009)) in a ratio of 10 : 1 based on OD_{280} in Buffer B additionally containing 0.3 mM TCEP for 1 h at 4°C and another 0.5 h at room temperature with constant shaking.

On-column cyclization reaction

The mixture after TEV cleavage was slowly, i.e. < 1 ml min^{-1} applied to a column ($d = 10$ mm, $L = 40$ mm) containing chitin resin (New England Biolabs) that was pre-equilibrated with Buffer B. Cleavage of the intein-2 was initiated by addition of two column volumes of Buffer B4 (20 mM Tris–HCl, pH = 8.5, 300 mM NaCl, 50 mM MESNA). After overnight incubation at 4°C with gentle agitation using a rocking platform shaker, the circular product was eluted from the column and concentrated to 2 ml using a stirred ultrafiltration cell and a membrane with a molecular weight cut-off of 1 kDa (Millipore AG, Zug, Switzerland).

Purification of the TEV-γE_c1-intein-2-CBD construct

After cell lysis and preparation of the soluble cell extract as described above for the 6xHis construct, but using Buffer B2 (20 mM Tris–HCl, pH = 7.0, 150 mM NaCl), further purification was either started with a fusion protein capture step using a chitin column or the cell extract was directly applied to a size exclusion column.

For the purification starting with the capture step, 15 ml of cell extract were again slowly loaded on a chitin column pre-equilibrated with Buffer B2 followed by a washing step with 80 ml B2. Intein cleavage was induced by addition of Buffer B4 (20 mM Tris–HCl, pH = 8.5, 150 mM NaCl, 50 mM MESNA) and overnight incubation at 4°C. The eluted fraction (15 ml) was concentrated to 4 ml using the stirred ultrafiltration cell and loaded on a HiLoad 16/600 Superdex 30 pg size exclusion column (GE Healthcare) pre-equilibrated with 25 mM Tris–HCl, pH = 7.5, 10 mM NaCl. The fraction containing the precursor protein was incubated with TEV protease in the ratio of 3 OD_{280} of fusion protein construct to 1 OD_{280} of TEV protease overnight at 4°C with gentle mixing in the same buffer. For the separation of linear and cyclic protein species, first potentially present disulfide bridges were reduced by incubation with 100 mM DTT for 40 min at 50°C and subsequently the metal clusters and hence the three-dimensional structure of the proteins were broken up by acidification of the solution to pH = 2 with HCl followed by a final separation step using a Superdex Peptide 3.2/300 size exclusion column (GE Healthcare).

For the purification without a prior capture step, 5 ml of the clear cell extract obtained from 1/3 l of cell culture was directly applied to a HiLoad 16/600 Superdex 200 pg column that was pre-equilibrated with 20 mM ammonium acetate, pH 7.5. The fraction containing the fusion construct was subjected to TEV cleavage using a ratio of 10 OD_{280} of construct to 1 OD_{280} of TEV protease and incubated for 4 h at 4°C, 1 h at 20°C and 0.5 h at 37°C. Subsequent on-column intein cleavage using a chitin column was performed as described for the His-tagged construct followed by size exclusion chromatography using the Superdex Peptide 3.2/300 column (10 mM ammonium acetate, pH 7.5).

Preparation of apo-cyc-γ-E_c-1

For the demetallation of the cyclic γ-E_c-1 form, i.e. cyc-γ-E_c-1, the cyclic peptide was incubated with 30 mM TCEP for 40 min at acidic pH (30 mM TCEP reduce the pH to 2–3) and applied to a Superdex Peptide 3.2/300 column, pre-equilibrated with 10 mM HCl and using a flow rate of 0.4 ml min^{-1}.

Mass spectrometry

Protein samples in 10 mM ammonium acetate, pH 7.5, or in 10 mM HCl for the analysis of the metal-bound or apo-form, respectively, were injected into a quadropole time-of-flight (TOF) Ultima API spectrometer (Waters, UK). Scans were accumulated and processed by the software MassLynx 3.5 (Micromass). Spectra were deconvoluted using the maximum entropy algorithm (MaxEnt1 in MassLynx 3.5). Electrospray parameters were capillary 2.8 V, cone 60 V and source temperature 80°C.

Protein quantification

Cyc-γ-E_c-1 concentrations were determined by quantification of thiolate groups with 2,2′-dithiodipyridine (2-PDS) at pH 4.0 (Pedersen and Jacobsen, 1980). 2-PDS reacts with free reduced thiol groups leading to the formation of thiopyridinone which can be quantified via its absorption at 343 nm (ε_{343} 7600 M^{-1} cm^{-1}). Protein concentrations were calculated based on six Cys residues per protein molecule.

SDS–PAGE

Protein separation was performed in 14% Tris-tricine polyacrylamide gels under denaturing conditions (Schägger and Vonjagow, 1987; Schägger, 2006). For visualization of proteins with a high Cys content, the protein samples were incubated with 4 mM monobromobimane (mBBr) for 20 min in the dark before denaturation with SDS loading buffer (Meloni *et al.*, 2005). After running the gel, unreacted mBBr was removed by incubation of the gel in 50% methanol for 1 h. Visualization of the bands based on fluorescence detection of formed mB-thioethers (mBSR) was performed on the UV-transilluminator Bio Vision 3026 using a filter for ethidiumbromide detection. As protein molecular mass marker Fermentas PageRuler^{TM} Prestained Protein Ladder (Thermo Fisher Scientific, Inc., Rockford, IL, USA) was used.

Results and discussion

Cyclization of proteins increases their stability and reduces their flexibility which should promote structural investigations. In addition, restrictions imposed on the conformational flexibility of a metal-binding protein might influence the environment of the coordination sites in such a way that higher specificity for a certain metal ion is achieved. However, also the opposite can be true, i.e. disturbance of metal ion binding by imposing unfavorable steric strain on the metal ion

coordinating ligands. To optimize protein cyclization for our needs we first attempted the cyclization of a small MT of known structure, i.e. the γ-domain of the wheat MT E_c-1. To ensure backbone cyclization without steric strain, we added an additional six amino acids long linker (GAAGAG) to the C-terminus of the γ-domain based on the NMR structure of the linear protein (Loebus *et al.*, 2011). The structure of this cyclic construct obtained in the here described way was also subsequently determined by NMR spectroscopy (Fig. 1A) and revealed no changes compared with the linear form (Tarasava *et al.*, 2013).

For the cyclization of the γ-E_c-1 domain we chose the pTWIN2 vector of the IMPACT^{TM}-TWIN system (New England Biolabs) (Evans *et al.*, 1999*a*) as a basis. The vector contains an N-terminal-modified *Ssp* DnaB intein denoted herein as intein-1, (Mathys *et al.*, 1999) as well as a C-terminal modified *Mth* RIR1 intein denoted as intein-2 (Evans *et al.*, 1999*b*), each fused to a chitin-binding domain (CBD) for purification. This system allows the generation of circular proteins after recombinant expression of the precursor protein in *E. coli*. Nevertheless, previous trials in our lab with this expression system revealed extensive cleavage of both inteins already during protein expression (unpublished data). Cleavage of intein-1 is not unexpected as the conditions for succinimide formation at the C-terminus of intein-1 are close to the cytoplasmic pH of 7 and ambient temperature. Cleavage of intein-1 and the concomitant formation of the free N-terminal Cys residue at the target protein already during protein expression can then facilitate preliminary cleavage of also intein-2 even without addition of thiol reagents (Gao *et al.*, 2012). Purification of the concomitantly formed circular peptides directly from the cell lysate is however difficult.

Hence, to prevent premature formation of the N-terminal Cys we replaced intein-1 with a TEV protease cleavage site (Tolbert and Wong, 2002) generating the TEV-γE_c1-intein2-CBD construct (Fig. 1B, top). In contrast to other proteases TEV protease tolerates a number of different amino acids, among them Cys, in the P1′ position, hence the position directly following the protease recognition site (Kapust *et al.*, 2002). In this way an N-terminal Cys residue is generated at the target protein after cleavage with TEV protease. Nevertheless, even in the absence of a free N-terminal Cys residue still a large amount of cleaved intein-2 was observed during protein purification with the chitin column (Supplementary Fig. S2). Presumably, this is triggered by an internal thiol compound, i.e. a Cys residue (see discussion below), and hence cannot be completely avoided. Nevertheless, the yield of full-length precursor protein (Fig. 1B, top) was slightly increased by keeping the expression time to a minimum, i.e. 3 h at 30°C. The TEV-γE_c1-intein2-CBD construct was separated from the cell extract using a chitin column. Elution from the column and hence cleavage of intein-2 by thioester formation was induced by MESNA. After incubation with TEV protease and purification by size exclusion chromatography, a mixture of linear and circular γ-domains was obtained (Fig. 2). Apart from a probable unfavorable protein conformation for efficient cyclization, the reason for the formation of linear species can also be the sensitivity of the C-terminal thioester for hydrolysis at neutral or even slightly alkaline pH, competing with the cyclization reaction via transesterification or occurring even before formation of the N-terminal Cys-residue by TEV protease cleavage.

The separation of the linear γ-E_c-1 domain from its circular form is challenging. Naturally, most of their biophysical properties such as the isoelectric point and the degree of hydrophobicity are the same, and their hydrodynamic radii are too small and too similar to enable efficient separation by size exclusion chromatography. One method to isolate a backbone-cyclized protein from its linear contaminant is the addition of a short C-terminal 3xHis tag to the target protein in the fusion protein precursor, which will be included in the cyclic protein after the successful ring closure reaction (Iwai and Plückthun,

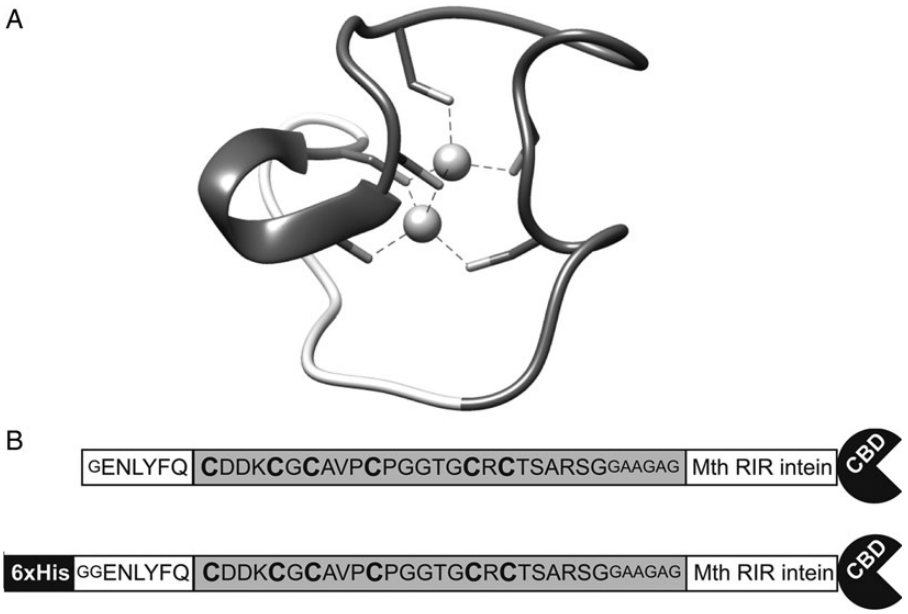

Fig. 1. (A) NMR solution structure of the cyclic Cd_2-γ-E_c-1 domain of a wheat MT (pdb entry 2MFP) in ribbon representation (Tarasava *et al.*, 2013). The additional amino acid linker (GAAGAG) is shown in light gray, Cd(II) ions as light gray spheres, and the Cys residues in stick mode. (B) Amino acid sequences and schematic representation of the two different protein precursors used in this study. The N-terminal white box contains the TEV protease cleavage site (ENLYQF), the gray box shows the sequence of the γ-E_c-1 with Cys residues highlighted in large bold letters. Small size letters indicate additional amino acids added as linkers to enable cyclization without steric strain (gray box) or to facilitate cloning and TEV cleavage (small letters in white box).

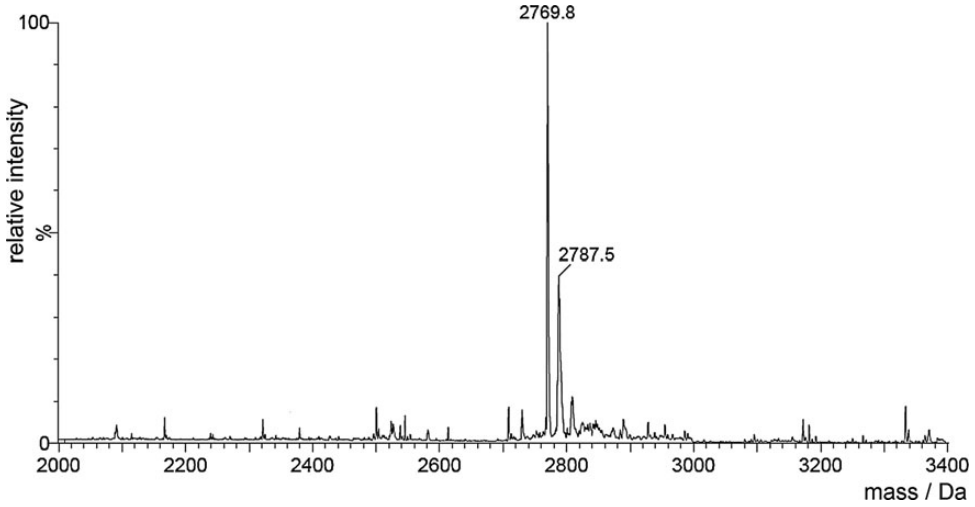

Fig. 2. ESI-MS spectrum of the final product obtained after purification and cyclization of the TEV-γE_c1-intein2-CBD construct (Fig. 1B, top) showing the presence of a mixture of cyclic (calc. 2770.1 Da) and linear γ-E_c-1 (calc. 2788.1 Da) species.

1999). Treatment of the mixture after the cyclization reaction with an exopeptidase will remove the 3xHis tag from the linear form, while the integrity of the cyclic peptide is retained and it can be captured with an immobilized metal ion affinity chromatography (IMAC) column. This method, however, is not applicable for MTs as the additional 3xHis tag will remain in the final cyclic product. His residues are, however, also good ligands for, e.g. Zn(II) ions and hence can interfere with the metal-thiolate cluster formation that is crucial for correct folding of the MT. Another method that has been described is the selective precipitation of the linear peptide form by heat denaturation due to the usually increased stability of cyclic peptides compared with their linear forms (Iwai and Plückthun, 1999). However, as MTs are generally very temperature stable probably owing to the lack of secondary structural elements this approach is also not feasible here.

Due to the difficulties in separating the linear and cyclic forms efficiently from each other, a procedure is needed that excludes formation of the linear form concurrent to protein cyclization. To prevent hydrolysis of the C-terminal thioester resulting in the linear form, exposure time of the thioester to the aqueous buffer should be kept to a minimum. Accordingly, TEV protease cleavage and hence generation of the reactive N-terminal Cys residue have to be performed before intein cleavage. This approach can be accomplished in three different ways. The easiest way, which should produce rather pure cyclic proteins in a single step, would be to perform TEV cleavage of the precursor protein bound to the chitin column. However, after elution, mainly uncleaved TEV-γ-E_c-1 species were observed, most likely caused by low accessibility of the cleavage site when bound to the chitin column (Supplementary Fig. S3). Accordingly, at least for the here presented system performance of TEV cleavage in batch mode is crucial. The most simplistic approach would be direct incubation of the cleared cell extract with TEV protease, which however led to massive precipitation and inactivation of the protease due to buffer incompatibility. Hence a pre-capturing step for the γ-E_c-1 precursor protein from the cell lysate seems to be essential, but the intein-tag cannot be used for this as elution of the intact precursor protein from the chitin column requires denaturing conditions and would further complicate the consecutive steps.

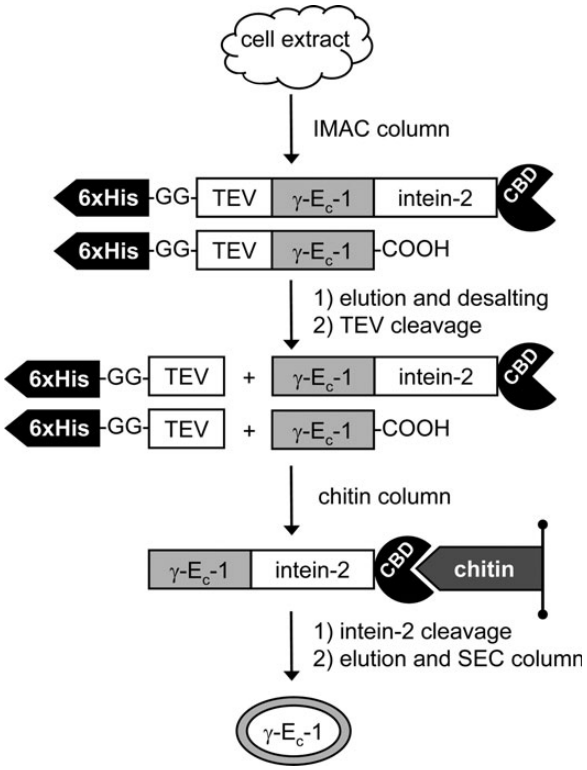

Fig. 3. Scheme of cyclic γ-E_c-1 production using the 6xHis-TEV-γE_c1-intein2-CBD construct (Fig. 1B, bottom). See the detailed description in the text.

Hence an additional 6xHis purification tag N-terminal of the TEV protease cleavage site was introduced (Fig. 1B, bottom) (Tolbert and Wong, 2002) and the purification and cyclization performed as illustrated in Fig. 3 and described in the following. Samples taken during the different purification steps were analysed with SDS–PAGE (Fig. 4). The clarified cell lysate was passed over an IMAC column to efficiently capture the full-length precursor protein via its 6xHis tag while the prematurely cleaved intein tag is found in the flow-through (Fig. 4, Lane 3). The precursor protein is eluted from the IMAC column with

imidazole and subjected to TEV protease cleavage in batch mode after buffer exchange with size exclusion chromatography (Lane 6). Again premature cleavage of the intein tag is observed (Lanes 5 and 6). Subsequently, the cleaved protein now bearing the N-terminal Cys residue is captured with a chitin column, getting rid at the same time of linear γ-E_c-1 species with cleaved intein-2 tags. On-column induction of intein self-cleavage and thioester formation followed by cyclization is induced with MESNA. The cyclic protein is then eluted from the chitin column and is free of any contaminations with its linear form according to SDS–PAGE (Lane 9). The purity was further checked with size exclusion chromatography. As the three-dimensional structures of the cyclic and the linear forms are closely similar owing to the metal-thiolate clusters that dictate the three-dimensional structures of the proteins, the protein fraction was acidified to pH 2 which provokes metal ion release from the proteins. The hydrodynamic radii of the apo-forms obtained in this way should be significantly different. Nevertheless, subsequent performance of size exclusion chromatography at acidic pH reveals a single symmetric elution peak (Supplementary Fig. S4), which is another strong indication for the purity of the cyclic protein. Finally, this was confirmed by an ESI-MS spectrum taken with the purified sample (Fig. 5). The total yield of purified apo-form of cyclic γ-E_c-1 was 0.7 mg per liter of cell culture. For comparison, the final yield of linear γ-E_c-1 prepared using the glutathione S-transferase tag was 4 mg per liter of cell culture (Loebus *et al.*, 2011).

Rather puzzling is the observation that intein cleavage apparently also takes place while the protein precursor is bound to the IMAC column although no additional thiol compounds are present (Fig. 4, Lanes 4 and 5). Tests with the precursor protein revealed that intein cleavage also occurs in pure water (data not shown). Hence it is probable that cleavage might be triggered by an internal thiol compound, i.e. a Cys residue. This mechanism is known as the 'thia zip cyclization' of Cys-rich peptides (Tam *et al.*, 1999*a*). Also internal Cys-residues can undergo reversible transthioesterifications with the C-terminal thioester and hence cause cleavage of the intein tag. However, only the thioester formed with the N-terminal Cys residue can rearrange via a S→N shift to form a stable peptide bond and hence to finally yield the end-to-end cyclized protein. The rate of this uncontrolled intein cleavage depends on the number and the pattern of Cys residues in the protein sequence (Tam *et al.*, 1999*a*). The thia-zip mechanism of cyclization and hence the premature cleavage of the intein are supported by the fact that cleavage is inhibited when the Cys residues are involved in disulfide bridge formation as observed for the recombinant production of cyclotides (Garcia *et al.*, 2011). Significant reduction of the amount of intein cleavage was also reported to occur in the presence of certain metal ions, e.g. 2 mM Zn(II) (Mills and Paulus, 2001). This inhibition is reversible upon Zn(II) removal with EDTA. Inhibition of intein cleavage can be easily explained by the

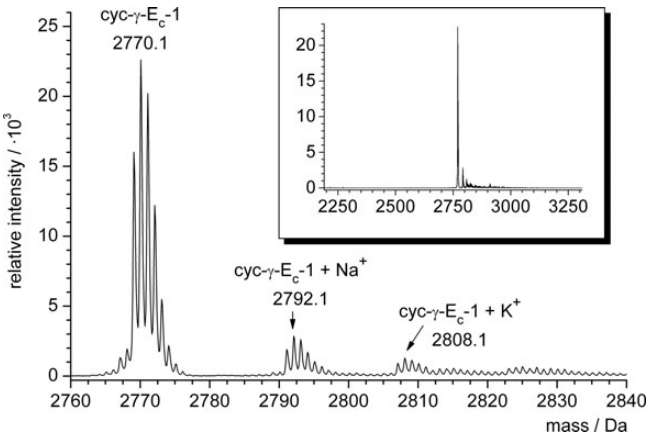

Fig. 5. Non-deconvoluted ESI-MS spectra of fraction eluted from the chitin column (Fig. 4., Lane 6) and further purified by size exclusion chromatography (Supplementary Fig. S4) showing the exclusive presence of cyclic γ-E_c-1 (calc. 2770.1 Da). The linear species would be expected at a molecular mass of 2788.1 Da. Additional small peaks belong to the sodium (calc. 2792.1 Da) and the potassium adduct (calc. 2808.1 Da) of the cyclic species. The inset shows the entire spectrum in the range of 2240–3260 Da revealing the purity of the sample.

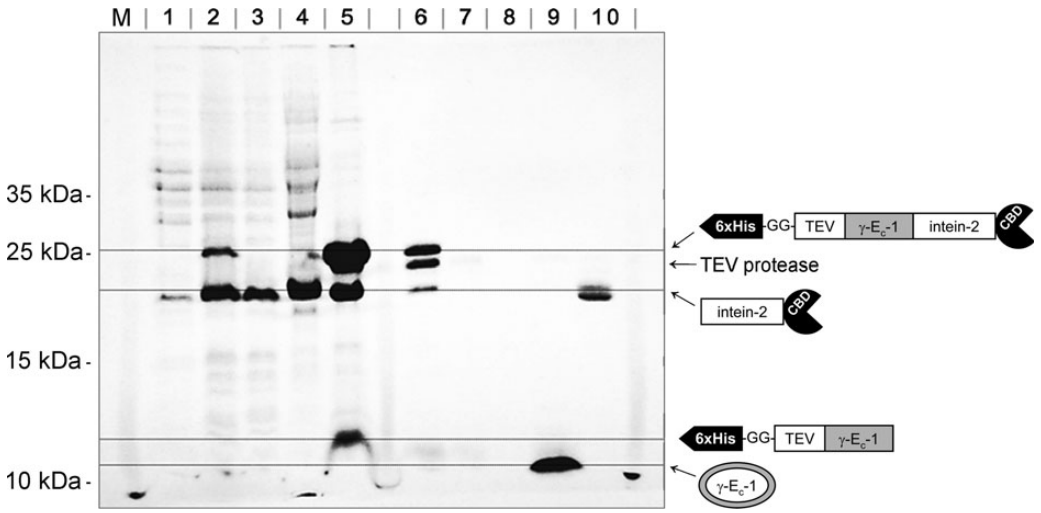

Fig. 4. Monitoring of cyclic γ-E_c-1 production by SDS–PAGE. For visualizing of proteins with high Cys content all samples were stained with mBBr and the gel visualized with fluorescence imaging. The silver stained gel can be found in the Supplementary Material (Supplementary Fig. S5). The single lanes show: the protein marker (M), non-induced soluble cell extract (1), soluble cell extract after induction and 3 h expression at 30°C (2), flow-through of IMAC column (3), washing of IMAC column (4), elution from IMAC column with imidazole (5), TEV incubation (6), flow-through of chitin column (7), washing of chitin column (8), elution from chitin column after incubation with MESNA (9), regeneration of chitin column with SDS (10).

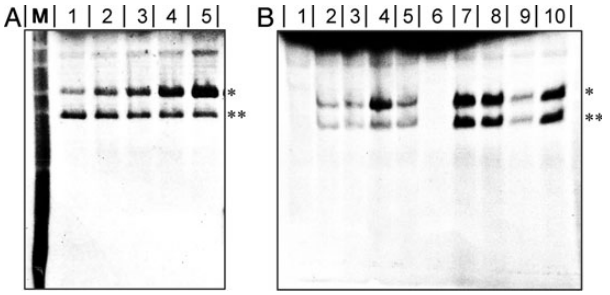

Fig. 6. mBBr stained SDS–PAGE showing the degree of intein cleavage depending on the amount of Zn(II) or Cd(II) added to the mixture. (A) Incubation of the size exclusion chromatography purified 6xHis-TEV-γE_c1-intein2-CBD construct (Fig. 1B, bottom) with 0, 1, 2, 4 and 8 mM Zn(II) (Lanes 1–5) showing the non-cleaved (*) and the cleaved (intein-CBD, **) construct. M denotes the lane with the protein ladder. (B) Composition of intact and cleaved constructs in the insoluble (Lanes 2–5) and soluble (Lanes 7–10) total protein fractions after addition of 0, 0.1 and 1 mM Zn(II) or 0.1 mM Cd(II) to the growth media after induction of protein expression with IPTG. Insoluble and soluble cell extracts without induction are loaded to Lanes 1 and 6, respectively.

formation of rather stable Cys thiolate-Zn(II) bonds, which block the Cys residues for the thia-zip reaction. Inhibition by Zn(II) is well applicable to our aim, the production of cyclic MTs, as Zn(II) is the natural co-factor for many MTs and in addition, also TEV protease is known to tolerate Zn(II) concentrations up to 5 mM without a significant reduction in activity (Tropea *et al.*, 2009). Incubation of the IMAC column purified 6xHis-TEV-γE_c1-intein2-CBD construct with increasing concentrations of Zn(II) indeed reveal a significant reduction of intein cleavage (Fig. 6A). To test reduction of the pronounced intein cleavage during protein expression, LB growth media, which already contain 10 μM Zn(II), were supplemented with additional Zn(II) or Cd(II) after induction of protein expression. While 0.1 mM Zn(II) or Cd(II) had no effect, 1 mM Zn(II) caused strong inhibition of cleavage (Fig. 6B). However, the reason for this inhibition is most likely some sort of denaturation of the precursor protein as it is now almost exclusively present in the insoluble fraction. If the intein can be re-solubilized and its activity restored remains to be evaluated.

Production of cyclic proteins with intein-based expression and purification systems such as the IMPACT^{TM} TWIN system has the general drawback of generating mixtures of cyclic and linear species that can be challenging to separate. The major problem is the continuous slow cleavage of the intein-tags during all steps. In an attempt to optimize the cyclization of the Cys-rich γ-E_c-1 domain of a wheat MT on the basis of the pTWIN2 vector, we replaced the N-terminal intein by a TEV cleavage site for the protection of N-terminal cysteine residue. However, the C-terminal intein is crucial for thioester formation. The here described optimized purification procedure aimed to constantly remove species with the preliminary cleaved intein tag at each stage of the process. In doing so we were able to produce a highly pure cyclic protein. To increase the yield the further reduction of self-cleavage by protein engineering or the addition of reversible inhibitors should be the focus of future investigations.

Supplementary Material

Supplementary Material are available at *PEDS* online.

Acknowledgements

We thank Dr. Serge Chesnov, Functional Genomics Centre Zurich, Dr. Laurent Bigler, and Dr. Igor A. Oleinich for the acquisition of MS spectra as well as Jovana Jakovleska for careful reading of the manuscript.

Funding

This work was supported by the University of Zurich (Forschungskredit to KT), the Swiss National Science Foundation (EF), the Faculty of Natural Sciences as well as the Department of Chemistry, University of Zurich.

References

Abriouel,H., Sanchez-Gonzalez,J., Maqueda,M., Galvez,A., Valdivia,E. and Galvez-Ruiz,M.J. (2001) *J. Colloid Interface Sci.*, **233**, 306–312.
Armishaw,C.J., Jensen,A.A., Balle,L.D., Scott,K.C.M., Sorensen,L. and Stromgaard,K. (2011) *Antioxid. Redox Signal.*, **14**, 65–76.
Blond,A., Peduzzi,J., Goulard,C., *et al.* (1999) *Eur. J. Biochem.*, **259**, 747–755.
Bogdanowich-Knipp,S.J., Chakrabarti,S., Williams,T.D., Dillmall,R.K. and Siahaan,T.J. (1999) *J. Pept. Res.*, **53**, 530–541.
Boulanger,Y., Goodman,C.M., Forte,C.P., Fesik,S.W. and Armitage,I.M. (1983) *Proc. Acad. Natl Sci. USA*, **80**, 1501–1505.
Camarero,J.A. and Muir,T.W. (1999) *J. Am. Chem. Soc.*, **121**, 5597–5598.
Camarero,J.A., Fushman,D., Sato,S., Giriat,I., Cowburn,D., Raleigh,D.P. and Muir,T.W. (2001) *J. Mol. Biol.*, **308**, 1045–1062.
Chong,S., Shao,Y., Paulus,H., Benner,J., Perler,F. and Xu,M. (1996) *J. Biol. Chem.*, **271**, 22159–22168.
Colangelo,D., Mahboobi,H., Viarengo,A. and Osella,D. (2004) *Biometals*, **17**, 365–370.
Craik,D.J., Simonsen,S. and Daly,N.L. (2002) *Curr. Opin. Drug Disc.*, **5**, 251–260.
Cui,C.X., Zhao,W.T., Chen,J.L., Wang,J.J. and Li,Q.Q. (2006) *Protein Expr. Purif.*, **50**, 74–81.
Dawson,P.E., Muir,T.W., Clarklewis,I. and Kent,S.B.H. (1994) *Science*, **266**, 776–779.
Evans,T.C., Benner,J. and Xu,M.Q. (1999a) *J. Biol. Chem.*, **274**, 18359–18363.
Evans,T.C., Benner,J. and Xu,M.Q. (1999b) *J. Biol. Chem.*, **274**, 3923–3926.
Freisinger,E. (2010) *Chimia*, **64**, 217–224.
Gao,M.M., Tong,Y., Tian,H., Gao,X.D. and Yao,W.B. (2012) *Appl. Microbiol. Biotechnol.*, **96**, 1283–1290.
Garcia,A.E., Tai,K.P., Puttamadappa,S.S., Shekhtman,A., Ouellette,A.J. and Camarero,J.A. (2011) *Biochemistry*, **50**, 10508–10519.
Gobel,H., van der Wal,A.C., Teeling,P., van der Loos,C.M. and Becker,A.E. (2000) *Virchows Arch.*, **437**, 528–533.
Gonzalez,C., Langdon,G.M., Bruix,M., Galvez,A., Valdivia,E., Maqueda,M. and Rico,M. (2000) *Proc. Acad. Natl Sci. USA*, **97**, 11221–11226.
Good,M. and Vasak,M. (1986a) *Biochemistry*, **25**, 8353–8356.
Good,M. and Vasak,M. (1986b) *Biochemistry*, **25**, 3328–3334.
Gran,L. (1973) *Acta Pharmacol. Toxicol.*, **33**, 400–408.
Gustafson,K.R., Sowder,R.C., Henderson,L.E., *et al.* (1994) *J. Am. Chem. Soc.*, **116**, 9337–9338.
Ha,K., Monbaliu,J.C.M., Williams,B.C., Pillai,G.G., Ocampo,C.E., Zeller,M., Stevens,C.V. and Katritzky,A.R. (2012) *Org. Biomol. Chem.*, **10**, 8055–8058.
Hanleybowdoin,L. and Lane,B.G. (1983) *Eur. J. Biochem.*, **135**, 9–15.
Heitz,A., Avrutina,O., Le-Nguyen,D., Diederichsen,U., Hernandez,J.F., Gracy,J., Kolmar,H. and Chiche,L. (2008) *BMC Struct. Biol.*, **8**, 54.
Iwai,H. and Plückthun,A. (1999) *FEBS Lett.*, **459**, 166–172.
Kapust,R.B., Tözser,J., Copeland,T.D. and Waugh,D.S. (2002) *Biochem. Biophys. Res. Commun.*, **294**, 949–955.
Korsinczky,M.L.J., Schirra,H.J., Rosengren,K.J., West,J., Condie,B.A., Otvos,L., Anderson,M.A. and Craik,D.J. (2001) *J. Mol. Biol.*, **311**, 579–591.
Lane,B., Kajioka,R. and Kennedy,T. (1987) *Biochem. Cell Biol.*, **65**, 1001–1005.
Langdon,G.M., Bruix,M., Galvez,A., Valdivia,E., Maqueda,M. and Rico,M. (1998) *J. Biomol. NMR*, **12**, 173–175.
Loebus,J., Peroza,E.A., Blüthgen,N., Fox,T., Meyer-Klaucke,W., Zerbe,O. and Freisinger,E. (2011) *J. Biol. Inorg. Chem.*, **16**, 683–694.
Luckett,S., Garcia,R.S., Barker,J.J., Konarev,A.V., Shewry,P.R., Clarke,A.R. and Brady,R.L. (1999) *J. Mol. Biol.*, **290**, 525–533.

Mathys,S., Evans,T., Chute,I., Wu,H., Chong,S., Benner,J., Liu,X. and Xu,M. (1999) *Gene*, **231**, 1–13.
Meloni,G., Knipp,M. and Vašák,M. (2005) *J. Biochem. Biophys. Methods*, **64**, 76–81.
Messerle,B.A., Schaffer,A., Vasak,M., Kägi,J.H.R. and Wüthrich,K. (1992) *J. Mol. Biol.*, **225**, 433–443.
Mills,K.V. and Paulus,H. (2001) *J. Biol. Chem.*, **276**, 10832–10838.
Muir,T.W. (2003) *Annu. Rev. Biochem.*, **72**, 249–289.
Ohta,A., Yamagishi,Y. and Suga,H. (2008) *Curr. Opin. Chem. Biol.*, **12**, 159–167.
Parthasarathy,R., Subramanian,S. and Boder,E.T. (2007) *Bioconjugate Chem.*, **18**, 469–476.
Pedersen,A.O. and Jacobsen,J. (1980) *Eur. J. Biochem.*, **106**, 291–295.
Peroza,E.A. and Freisinger,E. (2007) *J. Biol. Inorg. Chem.*, **12**, 377–391.
Peroza,E.A., Al Kaabi,A., Meyer-Klaucke,W., Wellenreuther,G. and Freisinger,E. (2009a) *J. Inorg. Biochem.*, **103**, 342–353.
Peroza,E.A., Schmucki,R., Güntert,P., Freisinger,E. and Zerbe,O. (2009b) *J. Mol. Biol.*, **387**, 207–218.
Rebuffat,S., Blond,A., Destoumieux-Garzon,D., Goulard,C. and Peduzzi,J. (2004) *Curr. Protein Pept. Sci.*, **5**, 383–391.
Saether,O., Craik,D.J., Campbell,I.D., Sletten,K., Juul,J. and Norman,D.G. (1995) *Biochemistry*, **34**, 4147–4158.
Schägger,H. (2006) *Nat. Protoc.*, **1**, 16–22.
Schägger,H. and Vonjagow,G. (1987) *Anal. Biochem.*, **166**, 368–379.
Tam,A. and Raines,R.T. (2009) *Methods Enzymol.*, **462**, 25–44.
Tam,J.P., Lu,Y.A. and Yu,Q.T. (1999a) *J. Am. Chem. Soc.*, **121**, 4316–4324.
Tam,J.P., Lu,Y.A., Yang,J.L. and Chiu,K.W. (1999b) *Proc. Acad. Natl Sci. USA*, **96**, 8913–8918.
Tarasava,K., Johannsen,S. and Freisinger,E. (2013) *Molecules*, **18**, 14414–14429.
Tavassoli,A. and Benkovic,S.J. (2007) *Nat. Protoc.*, **2**, 1126–1133.
Tolbert,T.J. and Wong,C.H. (2002) Angew. Chem. Int. Ed. Engl., **41**, 2171–2174.
Trabi,M. and Craik,D.J. (2002) *Trends Biochem. Sci.*, **27**, 132–138.
Tropea,J.E., Cherry,S. and S.,W.D. (2009) In Doyle,S.A. (ed.), *High throughput protein expression and purification*. New York: Humana Press Inc., Vol. **498**, pp. 297–307.
Vasak,M., Kägi,J.H.R., Holmquist,B. and Vallee,B.L. (1981) *Biochemistry*, **20**, 6659–6664.
Wang,C.K., Hu,S.H., Martin,J.L., *et al*. (2009) *J. Biol. Chem.*, **284**, 10672–10683.
Watanabe,T., Ito,Y., Yamada,T., Hashimoto,M., Sekine,S. and Tanaka,H. (1994) *J. Bacteriol.*, **176**, 4465–4472.
White,C.J. and Yudin,A.K. (2011) *Nat. Chem.*, **3**, 509–524.
Witherup,K.M., Bogusky,M.J., Anderson,P.S., Ramjit,H., Ransom,R.W., Wood,T. and Sardana,M. (1994) *J. Nat. Prod.*, **57**, 1619–1625.
Wood,D.W., Derbyshire,V., Wu,W., Chartrain,M., Belfort,M. and Belfort,G. (2000) *Biotechnol. Prog.*, **16**, 1055–1063.
Xu,M.Q. and Evans,T.C. (2001) *Methods*, **24**, 257–277.
Xu,M.Q. and Perler,F.B. (1996) *EMBO J.*, **15**, 5146–5153.
Xu,M.Q., Paulus,H. and Chong,S.R. (2000) *Methods Enzymol.*, **326**, 376–418.
Zhao,Z.L., Ma,X., Li,L., Zhang,W., Ping,S.Z., Xu,M.Q. and Lin,M. (2010) *J. Microbiol. Biotechnol.*, **20**, 460–466.