Available online at www.sciencedirect.com
ScienceDirect
Analytical Biochemistry 356 (2006) 86–93
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

Protein purification via temperature-dependent, intein-mediated cleavage from an immobilized metal affinity resin

Kenneth V. Mills *, Katherine R. Connor, Deirdre M. Dorval, Katherine T. Lewandowski

Department of Chemistry, College of the Holy Cross, Worcester, MA 01610, USA

Received 10 March 2006
Available online 19 May 2006

Abstract

The intein that interrupts the DNA polymerase II DP2 subunit in *Pyrococcus abyssi* can be overexpressed in *Escherichia coli* and purified as an unspliced precursor. On in vitro incubation at 37 °C or higher, the intein mediates efficient protein splicing. Mutations can be introduced into an intein fusion protein that prevent the second and third steps of protein splicing. As a result, the intein fusion protein can facilitate temperature-dependent formation of a thioester linkage between the N-extein and intein. This thioester is susceptible to in vitro hydrolysis or thiolysis at temperatures of 40 °C or higher, and we have exploited this activity to generate a temperature-dependent protein purification scheme. Protein purification using this intein does not require the addition of exogenous thiols and is compatible with the use of immobilized metal affinity chromatography. The identity of the C-terminal residue of the N-extein has less influence on the cleavage reaction than in current purification systems in terms of premature in vivo cleavage and is complementary to current systems in terms of efficient in vitro cleavage.

Keywords: Protein purification; Intein; Protein splicing; *Pyrococcus abyssi*

Protein splicing is a posttranslational modification by which an intervening polypeptide, called an intein, directs its own excision from flanking polypeptides, called exteins, concomitant with the ligation of the exteins via a peptide bond [1]. The extein that is linked to the N terminus of the intein is called the N-extein, and the extein fused to the C terminus of the intein is called the C-extein.

Protein splicing occurs by a four-step mechanism [1]. The first step of protein splicing is an N–S or N–O acyl rearrangement of the peptide bond between the C-terminal residue of the N-extein and the first residue of the intein, which normally is Cys or Ser. This results in the formation of a linear thioester or ester. The second step of splicing is a transesterification reaction that results in the N-extein being linked via a thioester or ester bond to the side chain of the first residue of the C-extein, which normally is Cys, Ser, or Thr. The third step of splicing involves the coupling of peptide bond cleavage to cyclization of the C-terminal Asn of the intein, resulting in an excised intein with a C-terminal aminosuccinimide and exteins linked by an ester or thioester bond. Finally, the fourth step of splicing involves the facile conversion of the ester or thioester linking the exteins to a native peptide bond.

Mutations of conserved intein amino acids can interrupt the splicing process and divert the intein to facilitate side reactions [1]. Such side reactions have been used to generate protein purification systems [2–4]. In one type of system, a target protein for purification is used as the N-extein, an affinity domain is used as the C-extein, and the intein is modified by mutation such that the second or third step of splicing does not occur (Fig. 1). The precursor fusion protein is bound to an affinity resin via the C-terminal affinity domain, and the N-extein is eluted by hydrolysis or thiolysis of the thioester or ester bond linking the N-extein to the fusion protein. A second type of purification system uses an

* Corresponding author. Fax: +1 508 793 3530.
E-mail address: kmills@holycross.edu (K.V. Mills).

0003-2697/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2006.04.055

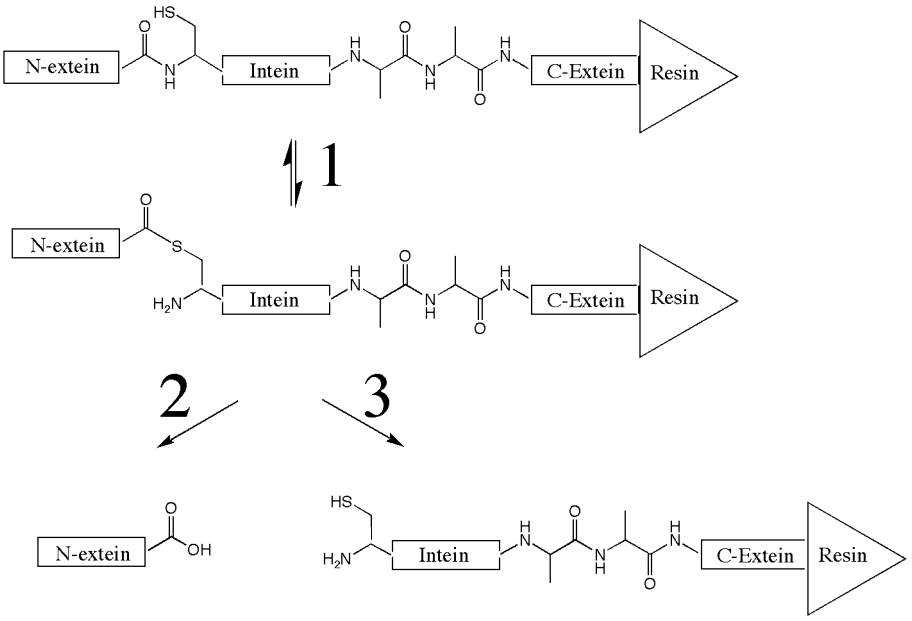

Fig. 1. Schematic representation of the affinity purification method. The N-extein is the *E. coli* maltose-binding protein, and the C-extein contains a poly-His tag. The N-extein is the protein of interest for purification in this scheme. The precursor extein–intein–extein fusion is bound to the immobilized metal affinity chromatography resin, and the resin is washed extensively. The intein is prevented from splicing by mutation to Ala of both the C-terminal residue of the intein and the N-terminal residue of the C-extein. The intein directs an N–S acyl rearrangement of the peptide bond linking the N-extein and the intein to a thioester (1). The thioester is susceptible to cleavage by hydrolysis at either 40 °C for 16 h or 55 °C for 5 h, and this liberates the N-extein from the resin (2) while the intein–C-extein fusion remains bound to the resin (3).

affinity domain as the N-extein and the target protein for purification as the C-extein. The exteins are interrupted by an intein modified by mutation such that the first two steps of splicing are prevented but the third step, Asn cyclization coupled to peptide bond cleavage, can be triggered by a change in temperature or pH. The result is elution of the C-extein from the resin while the N-extein–intein fusion remains bound to the resin.

The intein that interrupts the DNA polymerase II (PolII)[1] DP2 subunit in *Pyrococcus abyssi* can be overexpressed in *Escherichia coli* as a fusion protein with maltose-binding protein (MBP) as the N-extein and a poly-His tag as the C-extein [5]. This fusion protein can be overexpressed at 20 °C, isolated as an unspliced precursor, and induced to splice by incubation at elevated temperatures [5]. Mutation of the C-terminal Gln of the intein and the N-terminal Cys of the C-extein to Ala results in an intein fusion protein that can be purified and induced to undergo in vitro, temperature-dependent thiolysis or hydrolysis of the peptide bond between the N-extein and intein (i.e., N-terminal cleavage) [6]. In this article, we show that the intein fusion protein can be purified as an unspliced precursor regardless of the identity of the C-terminal residue of the N-extein. The intein fusion protein can facilitate liberation of a functional N-extein from an immobilized metal affinity column by N-terminal cleavage at neutral pH without the need for exogenous thiols.

Materials and methods

Plasmid preparation

Plasmid pPabPolQACA encodes protein QACA, an in-frame fusion of *E. coli* MBP to the seven C-terminal residues of the *P. abyssi* PolII N-extein to the 185 residues of the intein to the six N-terminal residues of the C-extein and a hexahistidine tag. The C-terminal Gln of the intein is changed to Ala, and the N-terminal Cys of the C-extein is changed to Ala in QACA. Preparation of plasmid pPabPolQACA was described previously [5,6].

To generate mutations of the C-terminal residue of the N-extein, 20 plasmids were generated and labeled pPabQACAX, in which X indicates the one-letter amino acid code of the substituted amino acid. To generate these plasmids, pPabPolQACA was digested with *Nsi*I and *Xma*I in the presence of shrimp alkaline phosphatase (Roche Diagnostics, Mannheim, Germany). The digested vector was ligated with the annealed and phosphorylated oligonucleotide pair QACAn1XU (5′-TGCTGCAGAGAGGAGAxxxTGCTT T) and QACAn1XL (5′-CCGGAAAGCAxxxTCTCCTC TCTGCAGCATGCA), in which codon xxx is replaced with an appropriate codon for each amino acid for each ligation.

The DNA sequences of all plasmids were verified by sequencing at Macrogen (Seoul, South Korea).

[1] *Abbreviations used:* PolII, DNA polymerase II; MBP, maltose-binding protein; Bis–Tris propane, 1,3-bis-[tris(hydroxymethyl)methylamine] propane; MALDI–TOF MS, matrix-assisted laser/desorption ionization time-of-flight MS.

Protein expression and purification

Plasmid-encoded proteins were overexpressed in *E. coli* BL21(DE3) (Novagen, Madison, WI, USA). The cultures were grown at 37 °C with shaking to a culture density (A_{600}) of approximately 0.7, and expression was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 0.4 mM and incubation at 20 °C overnight with shaking. The cells were harvested by centrifugation and resuspended in buffer A (50 mM Bis–Tris propane [pH 7.5] and 250 mM NaCl) supplemented with 100 μM phenylmethylsulfonylfluoride, 10 mM $MgCl_2$, 12 U/ml benzonase nuclease (Novagen), 1 mM 4-(2-aminoethyl)benzenesulfonylfluoride HCl, 0.8 μM aprotinin, 0.05 mM bestatin, 0.015 mM E-64, 0.02 mM leupeptin, and 0.01 mM pepstatin A (Calbiochem, La Jolla, CA, USA). The cells were disrupted by passage through a French pressure cell, and cellular debris was removed by ultracentrifugation at 300,000*g* for 30 min at 4 °C.

The supernatant was added to 1 ml of Talon metal affinity resin (Clontech, Palo Alto, CA, USA) that was prewashed with buffer A. The supernatant/resin mixture was incubated with gentle shaking for 20 to 30 min at 4 °C. The resin was then washed batchwise with 2 × 10 ml of buffer A and transferred to a gravity flow column. The resin was washed with 10 column volumes of buffer B (50 mM Bis–Tris propane [pH 7.5] and 350 mM NaCl) and 10 column volumes of buffer B supplemented with 10 mM imidazole. Two alternate procedures were followed at this point: one to develop a protein purification system based on induced N-terminal cleavage and a second to isolate uncleaved fusion protein to examine in vitro cleavage.

For the protein purification system, the resin was washed again with 10 column volumes of buffer B. The resin was then transferred with 500 μl of buffer B to a screw-top, 2.0-ml centrifuge tube. The resin was incubated in a water bath at 40 °C for 16 h or at 55 °C for 5 h. The resin was then transferred to a gravity flow column, an additional 500 μl of buffer B was added, and the 1-ml elution was collected. If needed, the eluted protein was concentrated using an Ultrafree-0.5 centrifugal filter device (5000 MW cutoff) as directed by the manufacturer (Millipore, Bedford, MA, USA). The resin was then washed with 10 column volumes of buffer B, and remaining protein was eluted with three 500-μl fractions of buffer B supplemented with 100 mM imidazole (buffer BI).

To purify precursor fusion protein, after washing with buffer B supplemented with 10 mM imidazole, proteins were eluted with three 750-μl fractions of buffer BI. To induce N-terminal cleavage of the precursor protein in solution, each protein in buffer BI was supplemented with 2.0 mM Tris(2-carboxyethyl)phosphine, 10 mM EDTA, 136 mM Bis–Tris propane (pH 7.5) (resulting in a final pH of 7.8), and 100 mM DL-1,4-dithiothreitol. This cleavage mixture was incubated at 40 °C for 16 h. Cleavage was terminated by the addition of 3× SDS–PAGE buffer (New England Biolabs, Ipswich, MA, USA) supplemented with DL-1,4-dithiothreitol.

The concentration in each protein fraction was determined by the Bradford method [7].

Protein analysis

Proteins were analyzed by SDS–PAGE. Precast 4 to 20% gradient Tris–glycine gels (PAGEr Gold, Cambrex BioScience, Rockland, ME, USA) were used via the Laemmli method [8]. Protein samples were mixed with 3× SDS–PAGE sample buffer supplemented with 40 μM DL-1,4-dithiothreitol and boiled for 3 min before loading. Broad-range protein markers (New England Biolabs) were included on each gel. Gels were stained with Gradipore Colloidal Coomassie Stain (Gradipore, Frenchs Forest, Australia) and imaged using a UMax Powerlook 1000 scanner. Densitometric analysis was performed using NIH Image J software. Matrix-assisted laser/desorption ionization time-of-flight MS (MALDI–TOF MS) analysis was performed at the University of Massachusetts Proteomics and Mass Spectrometry Facility (Worcester, MA, USA).

The percentage of N-cleavage was determined by correcting the densitometry data by relative molecular mass and using the equation $M/(M + MIH) \times 100$, where M is defined as the amount of the N-terminal product of the N-terminal cleavage reaction and MIH as the amount of uncleaved precursor. The extent of cleavage off the resin was determined by correcting the densitometry data by relative molecular mass and using the equation $IH/(IH + MIH) \times 100$, where IH is defined as the amount of the C-terminal product of the N-terminal cleavage reaction.

Results

Purification of an N-extein by cleavage from an intein-containing fusion protein

Protein QACA was overexpressed in *E. coli*. The cells were disrupted via passage through a French pressure cell, and the protein was bound to immobilized metal affinity resin. The resin was washed thoroughly and then incubated at 40 °C for 16 h or at 55 °C for 5 h as described in Materials and methods. The elution from the resin was analyzed by SDS–PAGE (Fig. 2). A single band that was consistent in size with the N-extein (M, 43.7 kDa), which is the N-terminal product of cleavage of the bond linking the N-extein and intein–C-extein fusion, was observed (Fig. 1, step 2, and Fig. 2, lanes 1 and 3). The eluted maltose-binding domain retains the ability to bind to amylose resin (data not shown). The extent of N-terminal cleavage on the resin can be estimated by washing the resin and then eluting the remaining protein in buffer BI. On analysis of this elution fraction by SDS–PAGE, two bands were observed. One band corresponded in size with the full-length precursor fusion protein (MIH, 66.7 kDa), and the other band corresponded in size with the C-terminal product of the N-terminal cleavage reaction (IH, 22.9 kDa) (Fig. 2, lanes 2 and 4).

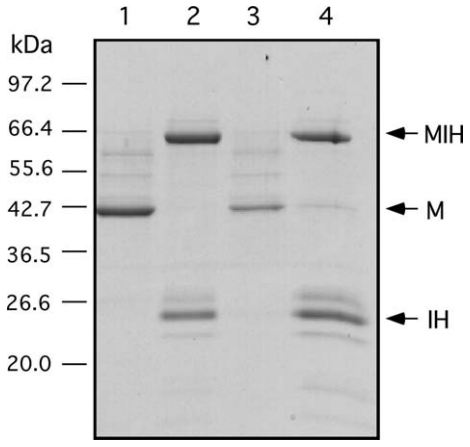

Fig. 2. SDS–PAGE analysis of the cleavage of the N-extein from the resin-linked precursor. Protein QACA was isolated from immobilized metal affinity resin. In each lane, 2.0 μg of total protein was loaded and SDS–PAGE was performed. Lanes 1 and 3 consist of protein isolated from the column flow-through after incubation on the resin at 40 °C for 16 h and at 55 °C for 5 h, respectively. Lanes 2 and 4 consist of the protein from the 100-mM imidazole elution fractions of the samples from lanes 1 and 3, respectively, following washing of the column. The bands labeled MIH, M, and IH correspond to uncleaved precursor protein and the N- and C-terminal products of the cleavage reaction, respectively.

The extent of N-terminal cleavage was 55% after 16 h incubation at 40 °C and was 71% after 5 h incubation at 55 °C. However, less total N-extein was isolated after the 5-h, 55 °C incubation in buffer B as determined by Bradford analysis, suggesting that the purified protein might be less soluble after incubation at the higher temperature.

Influence of the identity of the C-terminal residue of the N-extein on the cleavage reaction

To determine the influence of the identity of the C-terminal residue of the N-extein on the N-terminal cleavage reaction, 20 separate fusion proteins that vary only in the identity of that amino acid were generated as described in Materials and methods. These proteins also vary from protein QACA in that the amino acid four residues upstream from the N-terminal splice junction was changed from Lys to Glu as a result of including a new site for restriction digest screening during mutagenesis. Each protein was purified to determine whether any in vivo cleavage occurs. On SDS–PAGE analysis, a single band corresponding to the size of each of the precursor fusion proteins (*MIH*, 66.7 kDa) was observed (Fig. 3A).

To determine the extent of the N-terminal cleavage reaction facilitated by intein fusion proteins that are not resin bound, each precursor fusion protein was incubated at 40 °C for 16 h in the cleavage mixture described in Materials and methods. The cleavage mixture was analyzed by SDS–PAGE (Fig. 3B), and protein bands corresponding to the size of the precursor fusion protein (*MIH*, 66.7 kDa) and the N-terminal (*M*, 43.7 kDa) and C-terminal (*IH*, 22.9 kDa) products of N-terminal cleavage were observed. The extent of N-terminal cleavage is reported as the average of two trials. More than 90% cleavage was observed if the C-terminal amino acid of the N-extein was Ala, Arg, Asn, Asp, Cys, Gly, His, Phe, Ser, Thr, Trp, or Tyr. N-terminal cleavage between 80 and 90% was observed if the amino acid in that position was Leu, Lys, or Met. Cleavage between 60 and 75% was observed with Gln or Val. Cleavage was least efficient with Glu (46%), Ile (33%), or Pro (15%).

To correlate the efficiency of the N-terminal cleavage in solution with cleavage of resin-bound fusion protein, each QACA mutant was bound to a separate aliquot of immobilized metal affinity resin. The resin was washed and then incubated at 40 °C for 16 h. The protein elutions were analyzed by SDS–PAGE (Fig. 4). In Fig. 4, the elution from the 16-h incubation is analyzed in lane 1 for each sample, and the elution in buffer BI after washing is analyzed in lane 2.

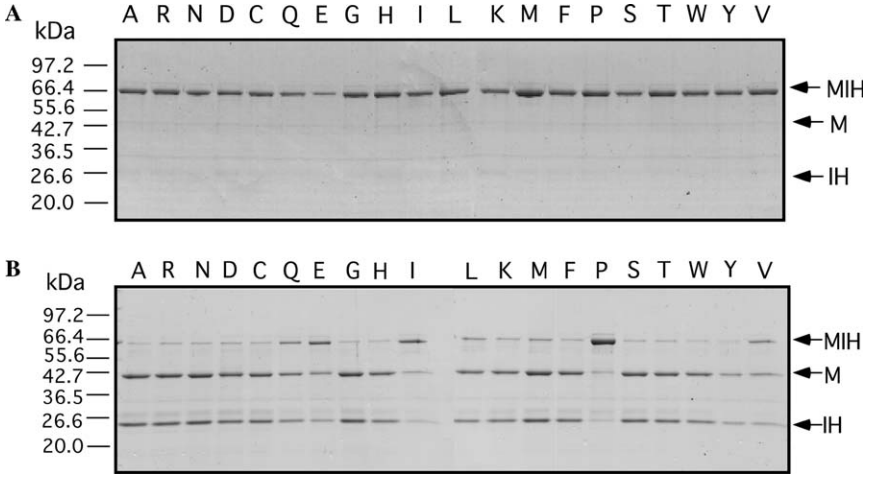

Fig. 3. SDS–PAGE analysis of N-terminal cleavage of protein QACA with variable C-terminal residues of the N-extein. (A) QACA precursor proteins were purified as described. In each lane, 1.5 μg of total protein was loaded and SDS–PAGE was performed. The one-letter amino acid code of the C-terminal residue of the N-extein for each QACA protein is given above the corresponding lane. (B) Each QACA variant was incubated under the "cleavage conditions" as described in Materials and methods for 16 h at 40 °C. Each lane consists of 20 μl of the cleavage mixture, and the one-letter amino acid code of the C-terminal residue of the N-extein is given above the corresponding lane. SDS–PAGE was performed as described.

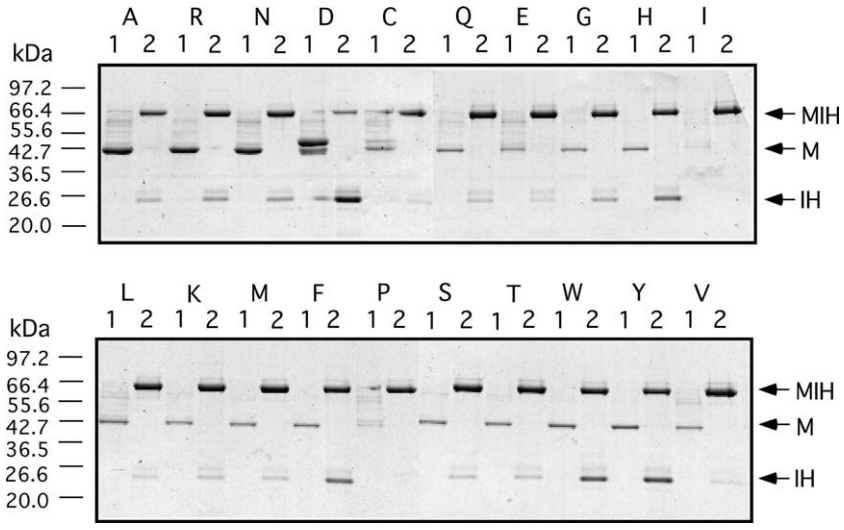

Fig. 4. SDS–PAGE analysis of cleavage of the N-extein from resin-linked precursor as a function of the identity of the C-terminal residue of the N-extein. QACA proteins were isolated from immobilized metal affinity resin. Pairs of lanes are labeled by the one-letter amino acid code for the C-terminal residue of the N-extein. SDS–PAGE was performed as described. Lane 1 of each pair consists of the protein isolated from the column flow-through after incubation on the resin at 40 °C for 16 h at pH 7.5. For each sample, 0.5 μg of total protein was loaded. Samples A, R, N, D, C, L, K, M, P, S, T, and V were concentrated as described to load the protein sample in a reasonable volume. Lane 2 of each pair consists of the protein from the 100-mM imidazole elution fractions of the samples from lane 1. Approximately 2 μg of total protein from the elution mixture was loaded.

For all samples except the fusion protein with Asp as the C-terminal residue of the N-extein, a single band was observed in lane 1, corresponding to the size of the cleaved N-extein (*M*, 43.7 kDa). To analyze the purity of the eluted N-extein and to load approximately 0.5 μg of protein in a volume conducive to SDS–PAGE, elutions in lane 1 for samples A, R, N, D, C, L, K, M, P, S, T, and V were concentrated using a centrifugal filtration device as described. For all samples, two bands were observed in lane 2, corresponding to the size of the precursor fusion protein (*MIH*, 66.7 kDa) and the C-terminal product of N-terminal cleavage (*IH*, 22.9 kDa).

The extent of N-terminal cleavage of the resin-bound fusion protein was estimated by analysis of lane 2 for each sample. N-terminal cleavage was between 50 and 70% for fusion proteins with His, Phe, Trp, or Tyr at the C terminus of the N-extein. N-terminal cleavage was between 40 and 50% with Ala, Arg, or Asn at that position; between 15 and 20% with Cys, Gln, Lys, Met, Ser, or Thr; between 5 and 10% with Glu, Leu, or Val; and less that 5% with Ile or Pro.

For the precursor fusion protein with Asp at the C terminus of the N-extein, two bands appeared in the elution fraction after the 16-h incubation (Fig. 4, lane D1). One band was consistent with the size of the N-extein (*M*, 43.7 kDa), but the other band was consistent with a protein of approximately 50 kDa. However, the elution from the resin using buffer BI (Fig. 4, lane D2) resulted in bands consistent with the sizes of the precursor protein (*MIH*, 66.7 kDa) and the C-terminal product of N-terminal cleavage (*IH*, 22.9 kDa). In fact, this resin-bound cleavage was the most efficient at 89%. Likewise, no anomalous results were obtained when the cleavage reaction was run in solution (Fig. 3B, lane D). Analysis by MALDI–TOF MS of the elution fraction after the 16-h incubation on the resin resulted in a single major peak with an *m/z* ratio of 43,612, which is within 0.2% of the predicted molecular mass of the N-extein (43,709.3 Da). No MS peak was detected near 50 kDa. After incubations of the 16-h elution fraction for 2 h at 25 or 37 °C in buffer B adjusted to pH 5.5, 7.0, or 9.0, or in buffer B supplemented with 100 mM DL-1,4-dithiothreitol, the results of SDS–PAGE analysis were similar to the result in lane D1 of Fig. 4.

Discussion

Protein purification by use of the P. abyssi PolII intein

Protein splicing mediated by the intein that interrupts the PolII DP2 subunit of *P. abyssi* is temperature-dependent in an intein fusion protein with *E. coli* MBP as the N-extein and a poly-His tag as the C-extein [5]. If the C-terminal Gln of the intein and the N-terminal Cys of the C-extein are changed to Ala via site-directed mutagenesis, the intein fusion protein can be isolated as a functional precursor and cleavage at the N terminus can be induced by an increase in temperature, either with or without the addition of exogenous thiols [6]. Based on this behavior, the *P. abyssi* PolII intein is a good candidate for use in a protein purification system in which the target proteins of interest are stable at 40 °C. Previous intein-mediated protein purification systems that are based on controlled N-terminal splice junction cleavage have found wide use in biotechnology, but each has certain limitations, including difficulty in purifying uncleaved precursor, the requirement of exogenous thiol reagents, and the influence of extein amino acids on cleavage efficiency [2,4]. The *P. abyssi* PolII intein should find use as a complement to these purification systems.

Benefits and limitations of previously described intein-based purification protocols using controlled cleavage at the intein N terminus

The intein that interrupts the VMA gene in *Saccharomyces cerevisiae*, modified by mutation, was the first intein used to generate a protein purification system based on cleavage at the N-terminal splice junction [9]. This system is marketed by New England Biolabs as the pTYB or pCYB series of vectors [2,4]. The intein fusion protein is stable at 4 °C and temperatures up to 25 °C, and N-terminal cleavage can be induced by the addition of exogenous thiols or hydroxylamine at a broad range of pH values [9]. However, it has been reported that the intein can promote significant in vivo N-terminal cleavage if the N-extein ends in Asp, Glu, Arg, His, or Thr and that in vitro N-terminal cleavage can be inefficient if the N-extein ends with Asn, Cys, or Pro [2,4].

A second protein purification system is based on a modified version of the *Mycobacterium xenopi* gyrase A intein [10]. The wild-type intein, when expressed in *E. coli* as a fusion protein between *E. coli* MBP and paramyosin, promotes mostly splicing on expression at 19 °C. However, expression at 37 °C results in the isolation of mostly unspliced precursor. The precursor can be induced to promote N-terminal cleavage by incubation with 50 mM DL-1,4-dithiothreitol. The identity of the C-terminal residue of the N-extein has an influence on the efficiency of the cleavage. N-exteins that end with Asp promote in vivo cleavage, and N-exteins that end with Ser, Pro, or Glu preclude in vitro cleavage [2,4]. This system is marketed by New England Biolabs as the pTXB series of vectors [11].

The *Methanobacterium thermoautotrophicum* RIR1 intein has found use as part of the New England Biolabs pTWIN and pTRB series of vectors [12]. This mini-intein has a Pro as the native C-terminal residue of the N-extein. Although splicing is not efficient with Pro in this position, efficiency is improved on mutation of this residue to Ala [12]. The intein promotes efficient thiol-induced cleavage when the C-terminal Pro of the N-extein is changed to Gly and the C-terminal Asn of the intein is changed to Ala. The flexibility of the C-terminal residue of the N-extein has not been explored, and some in vivo cleavage is reported at low expression temperatures [4].

Another N-terminal cleavage system is based on a fusion protein using the *Synechocystis* sp. PCC6803 DnaB mini-intein [13], also promoted by New England Biolabs as the pTSB vector series. If the C-terminal Asn of the intein is mutated to Ala, significant in vivo cleavage occurs at the N-terminal splice junction, and this limits the yield of full-length uncleaved precursor for purification. However, the intein can promote N-terminal cleavage by hydrolysis, and it is unclear whether some reported in vivo cleavage actually occurs during SDS–PAGE sample preparation. Interestingly, replacement of the C-terminal Asn of the intein with Leu, Lys, or Pro reduces in vivo cleavage. However, in vitro cleavage of those modified inteins can proceed only by thiolysis (and not by hydrolysis) and at elevated pH for the Asn-Pro intein. N-terminal cleavage of the *Synechocystis* sp. PCC6803 DnaB mini-intein fusion requires Gly as the C-terminal residue of the N-extein because substitution to Ala, Leu, or Pro prevents N-terminal cleavage [13]. It has also been suggested that purification using this intein fusion system should be performed at 4 °C and within 4 to 6 h and that buffers of pH 6.0 should be used to prevent cleavage during purification [4].

Advantages of a protein purification system using the P. abyssi PolII intein

A protein purification system based on the *P. abyssi* PolII intein has advantages in comparison with the established protocols. The *P. abyssi* PolII intein fusion protein can be overexpressed at 20 °C and purified as uncleaved precursor regardless of the identity of the C-terminal residue of the N-extein (Fig. 3A). The modified *P. abyssi* PolII intein does not promote notable in vivo cleavage, as noted by the absence of band *IH* (the intein–C-extein fusion) in Fig. 3A. N-terminal cleavage of the intein, although more efficient when promoted by an exogenous thiol (Fig. 3B and [5,6]), can be induced at elevated temperature (Figs. 2 and 4 and [6]), likely by hydrolysis of the thioester formed during the first step of protein splicing. This advantage will allow a purification system based on this intein to be used to isolate target proteins sensitive to treatment by thiols or thioester-reactive nucleophiles. Such a purification system can also make use of popular metal affinity resins. Although purification by N-terminal cleavage is inefficient when the N-extein ends with Asp, Cys, Glu, Ile, Pro, and Val (Fig. 4), the *P. abyssi* PolII intein promotes efficient in vitro cleavage with N-exteins that have presented difficulties in existing protocols such as those with C-terminal Arg, Asn, His, Ser, or Thr [2,4]. We currently are exploring the cause of the unusual SDS–PAGE mobility of the N-extein fragment with C-terminal Asp that results only from cleavage of resin-bound fusion protein, so at this time we do not advise purification of N-exteins ending with Asp using this intein. The influence of other, upstream N-extein residues has not been explored thoroughly. However, in generation of the QACAX mutations, the Lys that is four residues upstream from the N-terminal splice junction was changed to Glu with little influence on N-terminal cleavage efficiency (Figs. 2 and 4).

The influence of the C-terminal residue of the N-extein on the extent of N-terminal cleavage (Figs. 3 and 4) could be explained by the intein-promoted use of catalytic strain to drive forward the unfavorable amide to thioester rearrangement step. For instance, crystal structure analysis of the *M. xenopi* GyrA intein shows that the peptide bond linking the N-extein and intein is in the *cis* conformation [14], and NMR evidence suggests loss of amide bond resonance that could be due to bond rotation [15]. A crystal structure of the *S. cerevisiae* VMA intein demonstrates main chain distortion at the N-terminal scissile bond [16].

These two inteins also display a considerable influence of the N-1 residue on N-terminal cleavage efficiency [10,17], suggesting that the local conformation near the scissile bond is very important for intein promotion of thioester formation and likely will affect the efficiency of protein purification systems based on induced N-terminal cleavage.

Protein purification using cleavage at the intein C terminus

In principle, the *P. abyssi* PolII intein could also be used to construct a protein purification system based on temperature-induced C-terminal cleavage. In such a purification scheme, the target protein for purification is fused to the C terminus of the intein, bound to an affinity resin by the N-extein, and liberated from the resin by C-terminal cleavage [2,4]. Intein fusion proteins of the *P. abyssi* PolII intein with mutations that prevent the first step of protein splicing can be purified as active precursors and induced by elevation of temperature to promote cleavage of the peptide bond at the intein C terminus by cyclization of the C-terminal Gln or Asn [5,6]. One such intein fusion protein was created by changing the N-terminal residue of the C-extein to Ala, suggesting that purified proteins liberated by temperature-induced C-terminal cleavage need not be limited by the requirement of Cys as their N-terminal residue [5,6].

Other purification schemes based on induced C-terminal cleavage have been described. However, one of these purification schemes is limited by the need for preinduced N-terminal cleavage, and others are limited by premature in vivo cleavage or dependence of cleavage efficiency on C-extein amino acids. The *S. cerevisiae* VMA intein has been used to create a protein purification system based on controlled C-terminal intein cleavage and is marketed by New England Biolabs as the pTYB11/12 vectors [18]. However, prior N-terminal cleavage of the intein induced by thiols is required to initiate C-terminal cleavage, so the target protein is coeluted with the excess thiol reagent and a small peptide. A second system, marketed by New England Biolabs as the pBSC vectors, is based on the *Synechocystis* sp. PCC6803 DnaB intein modified by mutation to prevent splicing [13]. In this system, C-terminal cleavage during purification is suppressed by elevation of the pH of the purification buffer to 8.5, with C-terminal cleavage induced by incubation at pH 7.0. This cleavage system is limited by dependence on the identity of the N-terminal residue of the C-extein of both in vivo cleavage (which prevents purification of uncleaved precursor for purification) and in vitro cleavage (which prevents efficient elution of target protein from the resin) [4,19]. A modified *M. xenopi* GyrA intein was used to purify paramyosin and thioredoxin from chitin resin by temperature-induced C-terminal cleavage [10]. A modified *M. thermoautotrophicum* RIR1 intein can be used to purify target proteins by temperature-induced C-terminal cleavage [12] and is marketed by New England Biolabs as the pBRC vector. However, in vivo cleavage can occur at low expression temperatures (15 °C), and in vitro cleavage is curiously inhibited by expression at higher temperature [4,12]. A minimal version of the *Mycobacterium tuberculosis* RecA intein has been subjected to genetic selection to produce a mini-intein with increased dependence of C-terminal cleavage activity on pH [20] and has been optimized for protein purification and for incorporation of proteins in fluidic platforms [21,22].

The discovery of inteins, such as the *P. abyssi* PolII intein, that have unique structural elements and splice by noncanonical mechanisms has lent new insights into the means by which inteins promote the splicing reaction [6,23]. The *P. abyssi* PolII intein can facilitate temperature-dependent protein purification without the need for thiol reagents and with reduced influence of extein amino acids on cleavage efficiency. This suggests that such novel inteins also can be the basis for improved protein purification schemes and can serve as useful tools in protein-based biotechnology.

Acknowledgments

This material is based on work supported by the National Science Foundation under Grant 0320824 and NSF Career Grant MCB-0447647. We thank Robert Bellin (College of the Holy Cross) for useful conversations and a critical review of the manuscript.

References

[1] H. Paulus, Protein splicing and related forms of protein autoprocessing, Annu. Rev. Biochem. 69 (2000) 447–496.
[2] S. Chong, F.B. Perler, Intein-mediated protein purification, in: M. Blot (Ed.), Prokaryotic Genomics: Methods and Tools in Bioscience and Medicine, Birkhäuser Verlag, Berlin, 2003, pp. 173–193.
[3] I. Giriat, T.W. Muir, F.B. Perler, Protein splicing and its applications, Genet. Eng. (NY) 23 (2001) 171–199.
[4] M.Q. Xu, H. Paulus, S. Chong, Fusions to self-splicing inteins for protein purification, Methods Enzymol. 326 (2000) 376–418.
[5] K.V. Mills, J.S. Manning, A.M. Garcia, L.A. Wuerdeman, Protein splicing of a *Pyrococcus abyssi* intein with a C-terminal glutamine, J. Biol. Chem. 279 (2004) 20685–20691.
[6] K.V. Mills, D.M. Dorval, K.T. Lewandowski, Kinetic analysis of the individual steps of protein splicing for the *Pyrococcus abyssi* PolII intein, J. Biol. Chem. 280 (2005) 2714–2720.
[7] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding, Anal. Biochem. 72 (1976) 248–254.
[8] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[9] S. Chong, F.B. Mersha, D.G. Comb, M.E. Scott, D. Landry, L.M. Vence, F.B. Perler, J. Benner, R.B. Kucera, C.A. Hirvonen, J.J. Pelletier, H. Paulus, M.Q. Xu, Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, Gene 192 (1997) 271–281.
[10] M.W. Southworth, K. Amaya, T.C. Evans, M.Q. Xu, F.B. Perler, Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein, BioTechniques 27 (1999) 110–120.
[11] T.C. Evans Jr., J. Benner, M.Q. Xu, Semisynthesis of cytotoxic proteins using a modified protein splicing element, Protein Sci. 7 (1998) 2256–2264.
[12] T.C. Evans Jr., J. Benner, M.Q. Xu, The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*, J. Biol. Chem. 274 (1999) 3923–3926.

[13] S. Mathys, T.C. Evans, I.C. Chute, H. Wu, S. Chong, J. Benner, X.Q. Liu, M.Q. Xu, Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation, Gene 231 (1999) 1–13.

[14] T. Klabunde, S. Sharma, A. Telenti, W.R. Jacobs Jr., J.C. Sacchettini, Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing, Nat. Struct. Biol. 5 (1998) 31–36.

[15] A. Romanelli, A. Shekhtman, D. Cowburn, T.W. Muir, Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein–intein junction, Proc. Natl. Acad. Sci. USA 101 (2004) 6397–6402.

[16] B.W. Poland, M.Q. Xu, F.A. Quiocho, Structural insights into the protein splicing mechanism of PI–*Sce*I, J. Biol. Chem. 275 (2000) 16408–16413.

[17] S. Chong, K.S. Williams, C. Wotkowicz, M.Q. Xu, Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein, J. Biol. Chem. 273 (1998) 10567–10577.

[18] S. Chong, G.E. Montello, A. Zhang, E.J. Cantor, W. Liao, M.Q. Xu, J. Benner, Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step, Nucleic Acids Res. 26 (1998) 5109–5115.

[19] T.C. Evans Jr., J. Benner, M.Q. Xu, The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins, J. Biol. Chem. 274 (1999) 18359–18363.

[20] D.W. Wood, W. Wu, G. Belfort, V. Derbyshire, M. Belfort, A genetic system yields self-cleaving inteins for bioseparations, Nat. Biotechnol. 17 (1999) 889–892.

[21] J. Miao, W. Wu, T. Spielmann, M. Belfort, V. Derbyshire, G. Belfort, Single-step affinity purification of toxic and non-toxic proteins on a fluidics platform, Lab. Chip 5 (2005) 248–253.

[22] D.W. Wood, V. Derbyshire, W. Wu, M. Chartrain, M. Belfort, G. Belfort, Optimized single-step affinity purification with a self-cleaving intein applied to human acidic fibroblast growth factor, Biotechnol. Prog. 16 (2000) 1055–1063.

[23] K.V. Mills, F.B. Perler, The mechanism of interin-mediated protein splicing: variations on a theme, Protein Peptide Lett. 12 (2005) 751–755.