Protein Expression and Purification 59 (2008) 314–319

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Enhancing the specificity of the enterokinase cleavage reaction to promote efficient cleavage of a fusion tag

S. Hesam Shahravan [a], Xuanlu Qu [a], I-San Chan [a], Jumi A. Shin [a,b,*]

[a] Department of Chemistry, University of Toronto, Mississauga, Ont., Canada L5L 1C6
[b] Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ont., Canada M5S 3G9

ARTICLE INFO

Article history:
Received 15 January 2008
and in revised form 21 February 2008
Available online 5 March 2008

Keywords:
Enterokinase
Enteropeptidase
Histidine affinity tag
bZIP
Aryl hydrocarbon receptor (AhR)
C/EBP

ABSTRACT

In our work with designed minimalist proteins based on the bZIP motif, we have found our His-tagged proteins to be prone to inclusion body formation and aggregation; we suspect this problem is largely due to the His tag, known to promote aggregation. Using AhR6–C/EBP, a hybrid of the AhR basic region and C/EBP leucine zipper, as representative of our bZIP-like protein family, we attempted removal of the His tag with enterokinase (EK) but obtained the desired cleavage product in very small yield. EK is known for proteolysis at noncanonical sites, and most cleavage occurred at unintended sites. We manipulated experimental conditions to improve specificity of proteolysis and analyzed the cleavage products; no effect was observed after changing pH, temperature, or the amount of EK. We then suspected the accessibility of the EK site was impeded due to protein aggregation. We found that the easily implemented strategy of addition of urea (1–4 M) greatly improved EK cleavage specificity at the canonical site and reduced adventitious cleavage. We believe that this enhancement in specificity is due to a more "open" protein structure, in which the now accessible canonical target can compete effectively with adventitious cleavage sites of related sequence.

Protein purification is often facilitated by use of protein tags: hence, recombinant fusion proteins. The polyHis sequence is commonly used for chelation during immobilized metal–ion affinity chromatography. The His tag can be left on the protein, and often it is; however, the His tag may interfere with certain experiments, including those involving metals or *in vivo* administration, and it can lead to protein aggregation and inclusion body formation during bacterial expression [1]. Removal of the His tag may therefore be advantageous, and endoprotease sites are typically engineered between the desired protein and tag.

Efficient site-specific proteolysis of fusion proteins is critical, yet the most commonly used endoproteases—enterokinase, factor Xa, and thrombin—do not exhibit stringent sequence specificity and often cleave at nontarget sites [2]. Enterokinase (EK, also known as enteropeptidase)[1] is a membrane-bound serine protease found in the duodenum and initiates activation of pancreatic hydrolases by cleaving and activating trypsinogen. Although the canonical target site for enterokinase is DDDDK, it is known that EK does not exhibit high stringency in its specificity for this sequence [3–6]. For instance, Liew et al. showed that EK preferentially cleaved at an unexpected LKGDR site that was near the carboxyl terminus of N-terminal proCNP and more accessible than the internal canonical DDDDK sequence [5].

In our research on design of minimalist proteins, we express short proteins in *Escherichia coli* that are modeled after the basic region/leucine zipper (bZIP) family of transcription factors, a dimeric, α-helical motif comprising ~60 amino acids [7,8]. We express our bZIP proteins as fusions comprising an N-terminal 6×His tag with an intervening EK recognition site to facilitate tag removal. We encountered two significant problems during production of AhR6–C/EBP, a hybrid of the aryl hydrocarbon receptor (AhR) basic region and CCAAT/enhancer binding protein (C/EBP) leucine zipper that we use in this work as representative of our bZIP family of proteins (Fig. 1): limited protein solubility and inefficient EK proteolysis of the His tag. We suspected that solubility problems stemmed from the His tag, known to promote aggregation [1]. Previously, we found that by manipulation of temperature during expression and experimental handling (the temperature-leap tactic), we could maintain solubility of our His-tagged bZIP derivatives [9,10]. However, in order to minimize such experimental manipulations, we sought to remove the His tag by use of EK, but we could only obtain the desired cleavage product in vanishingly small yield.

We therefore manipulated experimental conditions to maximize the yield of the desired protein products from EK proteolysis.

* Corresponding author. Fax: +1 905 828 5425.
E-mail address: jumi.shin@utoronto.ca (J.A. Shin).

[1] *Abbreviations used:* AhR, aryl hydrocarbon receptor; EK, enterokinase; C/EBP, CCAAT/enhancer binding protein; ESI-MS, electrospray ionization mass spectrometry; DTT, dithiothreitol; PMSF, phenylmethanesulfonyl fluoride; GST-CaM, glutathione *S*-transferase-calmodulin fusion protein; NT-proCNP, N-terminal proCNP; hIL-2, human Interleukin-2.

1046-5928/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.02.015

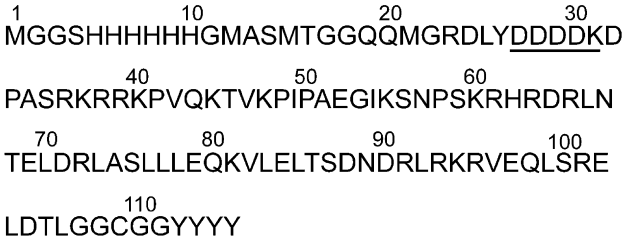

Fig. 1. Amino acid sequence of 6×His-tagged AhR6–C/EBP. The enterokinase recognition site is underlined. The AhR6–C/EBP basic region starts at Asp32, and the leucine zipper ends at Leu106. The His tag and EK cleavage site are at the N-terminus. At the C-terminus is the GGCGGYYYY sequence useful for spectroscopic evaluation and diazotization to solid support [23].

Because we suspected that the accessibility of the EK site was impeded, that is, a structural problem, we focused on use of denaturants as a way to open the structure, thereby essentially increasing the stoichiometry of the canonical recognition site over noncanonical, adventitious sites.

Materials and methods

DNA oligonucleotides were purchased from Operon Biotechnologies (Huntsville, AL). Enzymes were supplied by New England Biolabs (Pickering, ON). Native enterokinase (MW 150 kDa, obtained from calf intestine) was purchased from Roche Applied Science (Laval, QC). Reagents were supplied by EMD Chemicals (Gibbstown, NJ), Fisher Scientific Canada (Ottawa, ON), or Bioshop Canada (Burlington, ON). DNA sequencing was performed on an ABI 3730XL 96 capillary sequencer (Applied Biosystems) at the DNA Sequencing Facility in the Centre for Applied Genomics, Hospital for Sick Children (Toronto, ON). Electrospray ionization mass spectrometry (ESI-MS) was performed on a Micromass ZQ Model MM1 quadruple mass spectrometer (Waters) at the University of Toronto at Mississauga.

Preparation of AhR6–C/EBP protein

The gene for expression of AhR6–C/EBP was constructed by mutually primed synthesis, followed by polymerase chain reaction with terminal primers for gene amplification and purification by nondenaturing polyacrylamide gel electrophoresis. After restriction with BamHI and EcoRI, duplex DNA was cloned into protein expression vector pTrcHis B (Invitrogen, Carlsbad, CA) and sequenced; this vector expresses proteins with an N-terminal 6×His tag. The recombinant plasmid was transformed into *E. coli* strain BL21(DE3) (Stratagene, La Jolla, CA) by electroporation (Bio-Rad *E. coli* Gene Pulser).

Bacterial expression of AhR6–C/EBP was performed in LB medium containing 100 μg/mL ampicillin at 37 °C; induction was initiated at mid-log phase (OD_{600} ~0.6) with IPTG added to 1 mM final concentration. Cells were harvested after 3 h by centrifugation and lysed by sonication. The 6×His-tagged protein was purified on TALON cobalt metal–ion affinity resin (Clontech, Mountain View, CA) with the wash and elution buffers containing 8 M urea and 6 M guanidine–HCl, respectively, to promote protein solubility. The eluted protein was concentrated by centrifugation with a Centriplus centrifugal filter (Millipore, Bedford, MA). The protein was further purified by HPLC (Beckman System Gold) on a semipreparative reversed-phase C_4 column (Vydac, Hesperia, CA) with a gradient of acetonitrile–water plus 0.05% trifluoroacetic acid (v/v) at flow rate 4 mL/min; the gradient started at 10–25% acetonitrile over 15 min, followed by 25–55% acetonitrile over 60 min. Purity of the final product was confirmed by analytical HPLC on a C_4 column (Vydac) using the same gradient as above but flow rate 1 mL/min. Identity was verified by ESI-MS; calculated mass: 13034.6 g/mol; found: 13033.0 g/mol with the N-terminal Met cleaved during post-translational modification [11]. HPLC-purified protein was lyophilized and stored at −80 °C. The overall yield of purified protein was 2.5 mg/L LB culture.

Enterokinase cleavage of AhR6–C/EBP (typical EK reaction)

The enterokinase reactions (10 μL) were prepared according to the manufacturer's recommendations. We used the native form of EK supplied by Roche; in earlier experiments, we also used the light chain recombinant version of EK (New England Biolabs, data not shown); however, only with the native version could we obtain reproducible cleavage at the canonical EK target site, so we continued our studies with only the native form. 6.0 μg purified AhR6–C/EBP was incubated with 0.14 μg EK (EK: substrate ratio = 1:42) at 37 °C in 50 mM Tris, pH 7.6. The reactions were stopped by addition of an equal volume of Tricine Sample Buffer (Bio-Rad, Hercules, CA), which contains 125 mM dithiothreitol (DTT), followed by heating for 10 min at 95 °C, and analysis by 16.5% Tris–Tricine SDS–PAGE with molecular weight markers (Polypeptide SDS–PAGE Molecular Weight Standards, Bio-Rad, shown in Figs. 2–4 and 6). Coomassie stained gels were analyzed with LabImage (version 3.3.3) to calculate cleavage efficiencies. Purity of the native enzyme was assayed by analytical HPLC as above.

Isolation and identification of cleavage products by HPLC and ESI-MS

The enterokinase reaction (100 μL) was prepared as above and stopped by addition of phenylmethanesulfonyl fluoride (PMSF) to a final concentration of 1 mM, followed by heating for 10 min at 95 °C. The stopped reaction was incubated with 30 mM DTT for ⩾30 min at room temperature, followed by analytical HPLC as above. The collected fractions were analyzed with ESI-MS at flow rate 40 μL/min with the capillary charged at +3200 V, source temperature 100 °C, and cone voltage 20 V. The mass range 600–1100 m/z was scanned every 1 s for 40 s. Uncalibrated data were processed using MassLynx (version 4.00.00). Molecular masses were analyzed by massXpert (version 1.6.0) to identify cleavage sites. Due to the observed discrepancy between the migration rates of the Bio-Rad molecular weight markers (in particular, the MW marker at 16.9 kDa) and the uncleaved AhR6–C/EBP in Figs. 2–4 and 6, the same typical enterokinase reaction with AhR6–C/EBP (MW 13,033 Da) as above was prepared at large scale to generate our own molecular weight standard (not shown in Figs. 2–4 and 6). All fragments were purified by HPLC, followed by confirmation of identity by ESI-MS before

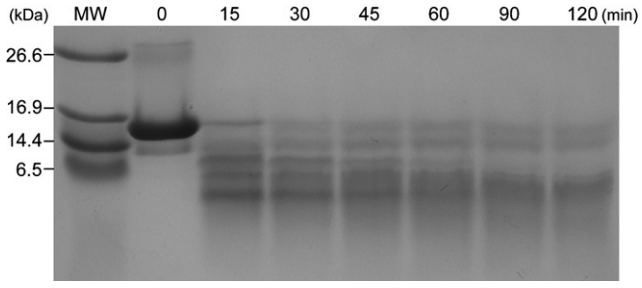

Fig. 2. Time-course analysis of cleavage products after incubation of AhR6–C/EBP with EK for the times indicated (top, *x*-axis, in minutes). Molecular-weight marker lane is at the left (MW, see Materials and methods), and apparent molecular weights are indicated in the *y*-axis legend. EK proteolysis of AhR6–C/EBP at the DDDDK recognition site should yield fragments of 9737 Da (band B in all SDS–PAGE figures) and 3315 Da (not visualized by SDS–PAGE).

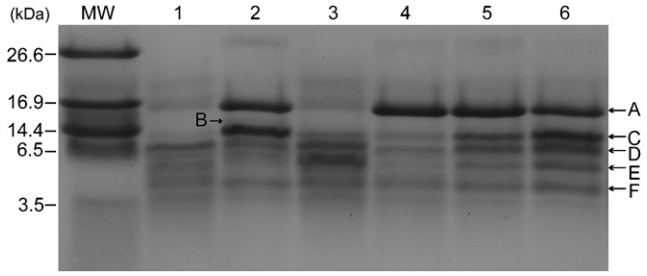

Fig. 3. EK cleavage reactions of AhR6–C/EBP with different reagents and under different experimental conditions. All reactions were stopped after 15 min incubation with EK. Molecular-weight marker lane (MW, see Materials and methods) and apparent molecular weights (*y*-axis legend) at left. Lane 1, EK reaction according to the manufacturer's recommendations (typical EK reaction); lane 2, typical EK reaction with 1 M urea; lane 3, typical EK reaction at pH 7.0; lane 4, typical EK reaction at pH 8.0; lane 5, typical EK reaction with one quarter of original EK amount; lane 6, typical EK reaction with half of original EK amount. The masses of the major bands as determined by ESI-MS: (A) 13,033 Da (intact AhR6–C/EBP); (B) 9737 Da; (C) 9-210 Da; (D) 8072 Da; (E) 7055 Da; (F) 4978 Da. Note that band B (9737 Da) is very faint; low yields of this cleavage product are obtained when the EK reaction contains only 1 M urea.

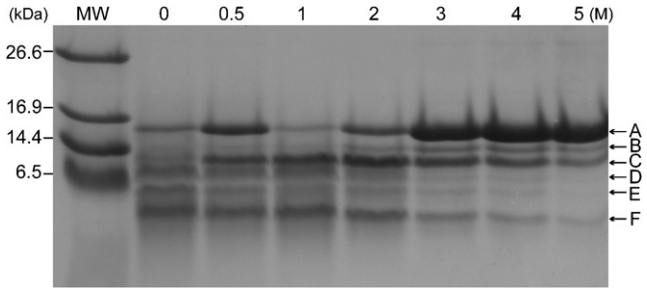

Fig. 4. EK cleavage reactions of AhR6–C/EBP under varying urea concentrations (top, *x*-axis). All reactions were stopped after 30 min incubation with EK. MW lane (see Materials and methods), *y*-axis legend, and masses of bands as determined by ESI-MS same as in Fig. 3. Note that the intensity of band B (9737 Da) increases as the concentration of urea is increased during EK cleavage.

loading onto SDS–PAGE. Our MW standard was therefore used to assign bands A–F in Figs. 3, 4 and 6.

Results and discussion

The reaction of AhR6–C/EBP with EK under the manufacturer's standard conditions (see Materials and methods) is rapid and somewhat nonspecific. In addition to the barely detectable intended cleavage after the canonical EK target sequence DDDDK yielding fragments of 9737 Da (band B in Figs. 3, 4, and 6, Fig. 5) and 3315 Da (too small to be visualized in SDS–PAGE, but shown in HPLC, Fig. 5), undesired adventitious cleavage occurs (Fig. 2). SDS–PAGE analysis of the reaction of AhR6–C/EBP with EK revealed that after just 15 min incubation, 82% of the protein had been cleaved at multiple sites. Prolonged incubation caused further degradation of these cleavage products, as indicated by their gradual disappearance during 2 h incubation. Thus, the prompt formation of these cleavage products, as well as their subsequent disappearance, indicates that, indeed, EK is a very active protease and cleaves our protein at sites other than the DDDDK target site.

We examined the purity of the enterokinase (the native form) by analytical HPLC to assess whether the adventitious cleavage of AhR6–C/EBP was due to proteases that might be copurified with the native enzyme (data not shown). We found that the EK used in our study was indeed pure and contained no detectable proteases. Therefore, as similarly observed by Choi et al. and Liew et al., we believe that degradation of AhR6–C/EBP by native enterokinase is the result of an intrinsic broad specificity of EK for various target sites rather than the presence of contaminating proteases [5,12].

Changing pH, temperature, or the amount of EK had no effect on cleavage pattern

We hypothesized that we could improve the specificity of the EK cleavage reaction of AhR6–C/EBP by manipulating the standard experimental conditions. We first tried changing the pH of the EK reaction buffer (Fig. 3). Lowering the pH from 7.6 to 7.0 had no appreciable influence on the reaction rate or cleavage pattern; however, increasing the pH to 8.0 had an inhibitory effect on EK activity, and after 15 min incubation with EK, 42% of our protein remained intact. We then examined whether lowering the reaction temperature would affect the cleavage specificity by decreasing EK activity. Lowering the temperature from 37 °C to 25 °C had no noticeable effect on the reaction rate or cleavage pattern (data not shown), similar to that observed by Menart and coworkers in their study on removal of affinity tags from oligomeric proteins [13].

We also tried reducing the amount of EK present in the reaction by adding one half or one quarter of the original EK amount: the rate of cleavage was reduced, but specificity was unaffected (Fig. 3). If we reduced the amount of EK and increased the incubation time from 2 h to 3 h, we observed the same additional cleavage products and their subsequent degradation, as above (data not shown). Contrary to the results of Liu et al., who reported that reducing the amount of EK could decrease nonspecific cleavage of human parathyroid hormone [14], we found that reducing the concentration of EK only affects the rate of proteolysis of AhR6–C/EBP, not the specificity.

Nonspecific cleavage due to inaccessibility of the target site

These results prompted us to adjust our hypothesis: we suspected that nonspecific proteolysis was due to inaccessibility of the EK target site, and hence, EK cleaved our protein at more accessible sequences, as it is known for cleaving at unexpected sites [3–6]. In our case, the probable reasons for the lack of accessibility to the target site are structural: protein dimerization and aggregation. The bZIP-like AhR6–C/EBP can homodimerize via the C/EBP leucine zipper coiled coil. Additionally, an unintended problem may stem from the observation that His-tagged proteins are prone to aggregation [1,15,16]. Thus, protein multimerization, whether by the normal dimerization mechanism or by misfolding and aggregation, may adversely affect EK's ability to bind to its target site.

We hypothesized that the presence of a denaturant would discourage secondary and tertiary, as well as quaternary, protein structure that shields EK from its target. Thus, we added urea to our cleavage reactions to reduce adventitious cleavage. Addition of urea caused a profound difference in our cleavage reactions (Fig. 3). After 15 min incubation of AhR6–C/EBP with EK in the presence of 1 M urea, we observed two major differences in the reaction pattern in comparison to the previously examined standard conditions: (i) a higher proportion of cleavage product at 9210 Da resulted (band C, Fig. 3); (ii) the cleavage product at 9737 Da, still only faintly detectable by SDS–PAGE, was produced in higher yield (band B, Fig. 3). We performed more cleavage reactions with even higher concentrations of urea; addition of over 1 M urea increased the amount of the 9737 Da product, as indicated by increased density at band B in Fig. 4. At 5 M urea, the denaturant began to inhibit EK activity, as indicated by the overall reduced amounts of cleavage products (Fig. 4, bands B–F).

We also analyzed EK cleavage of AhR6–C/EBP by use of other denaturants, including guanidine and SDS; both primarily inhib-

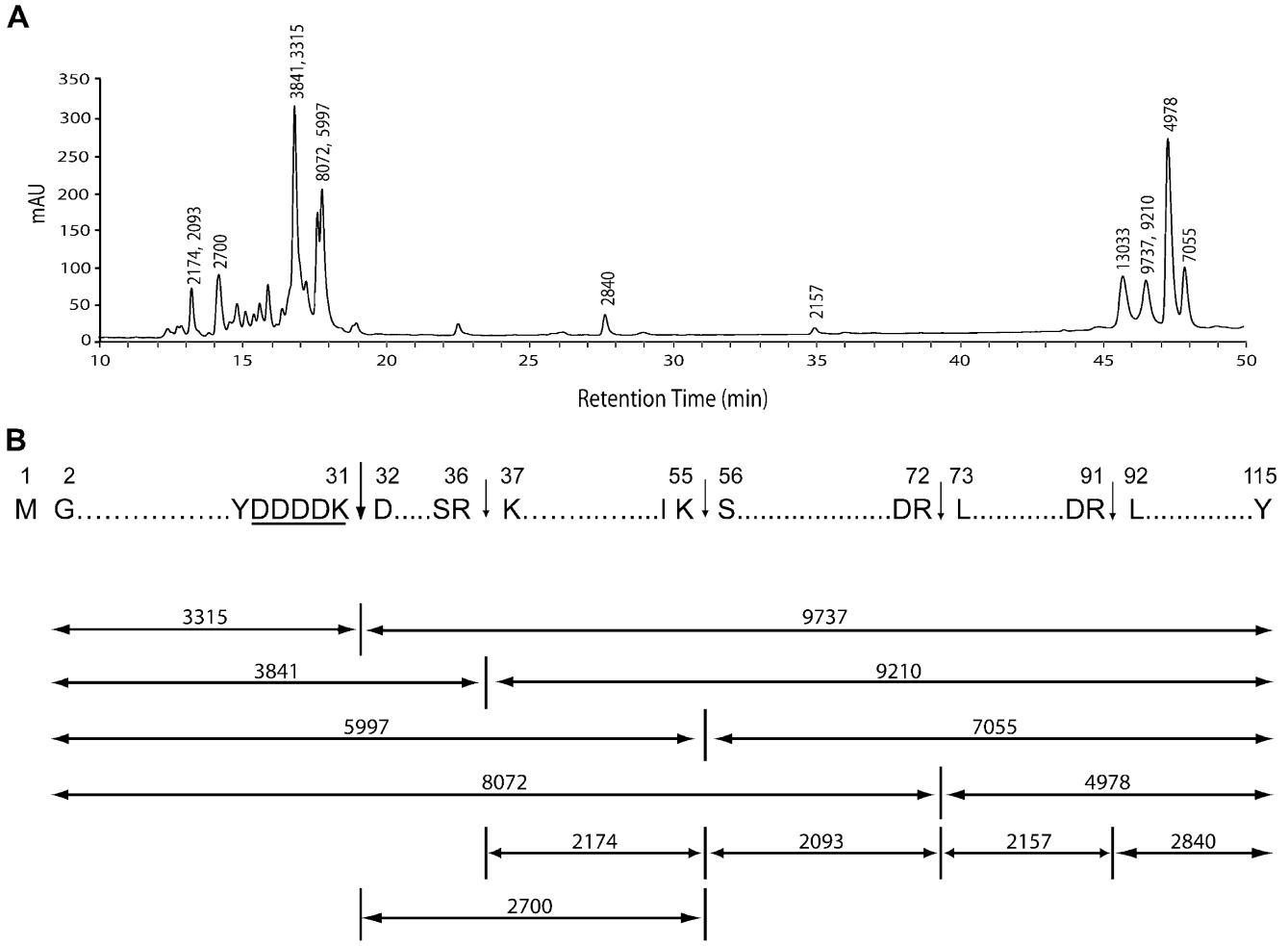

Fig. 5. (A) HPLC analysis of a *typical EK reaction* (see Fig. 3) of AhR6–C/EBP after 5 min incubation with EK monitored at 220 nm. Peaks are labeled with their associated masses, as determined by ESI-MS. (B) Enterokinase cleavage sites in AhR6–C/EBP. Top. Large arrow indicates canonical EK target site after DDDDK, and small arrows indicate secondary cleavage sites after Arg36, Lys55, Arg72, and Arg91. Bottom. Masses (Da) of HPLC-purified cleavage products.

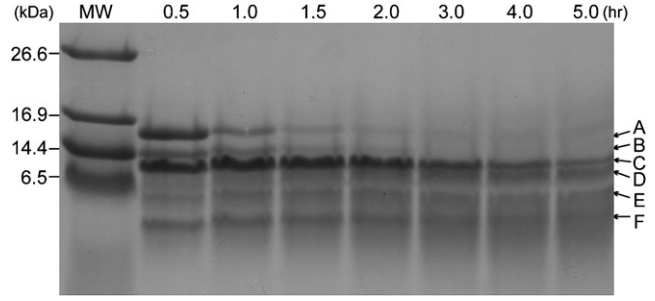

Fig. 6. Time-course analysis of cleavage products after incubation of AhR6–C/EBP with EK with 3 M urea for the times indicated (top, *x*-axis, in hours). MW lane (see Materials and methods), *y*-axis legend, and masses of bands as determined by ESI-MS same as in Fig. 3.

ited EK proteolysis at the canonical site rather than improving specificity of cleavage (data not shown). In addition, we performed an unfolding experiment to test whether we could achieve comparable results as observed upon addition of urea to the EK reaction by just manipulating temperature. We prepared a typical EK reaction (as specified in Materials and methods) but without EK and urea. This reaction was heated for 5 min at 90 °C followed by slow cooling to room temperature over 3 h. We then added EK and found that this temperature tactic had no effect on the specificity of the EK reaction (data not shown). Had this temperature tactic worked, we still believe that using urea is a more favorable method for achieving specificity of cleavage at higher yields, as it requires significantly less time.

Addition of denaturant reduces adventitious enterokinase cleavage of AhR6–C/EBP

In order to identify the cleavage products, we used the manufacturer's standard conditions for reaction of EK with AhR6–C/EBP, isolated the adventitious cleavage products by reversed-phase HPLC (Fig. 5), and identified their masses by ESI-MS, which allowed determination of the precise locations of EK cleavage sites in AhR6–C/EBP. EK proteolyzes our protein after Lys31 (its canonical DDDDK site), Arg36, Lys55, Arg72, and Arg91, which explains the total degradation of our protein observed after extended incubation with enterokinase (Fig. 2).

The largest cleavage product with mass 9737 Da, which appeared in EK reactions containing more than 1 M urea (Fig. 4, band B), is the result of cleavage after the DDDDK target site. Hence, addition of urea to our cleavage reactions indeed facilitated cleavage of AhR6–C/EBP at the intended EK target site. We then incubated AhR6–C/EBP with EK in reaction buffer containing 3 M urea for 5 h (Fig. 6). Comparison of these results with those obtained earlier without urea (Fig. 2) demonstrates the much slower

disappearance of the cleavage fragment at 9210 Da, formed after EK cleavage at Arg36, as indicated by band C in Fig. 6, thereby indicating inhibition of further digestion after Lys55, Arg72, and Arg91. We note that under all tested conditions, the adventitious 9210 Da product appears to be in higher yield than the desired 9737 Da product, as qualitatively visualized by SDS–PAGE; under HPLC analysis, we were unable to separate fully the 9210 Da and 9737 Da peaks (Fig. 5), and therefore, we were unable to assess quantitatively the amounts of these species.

Relative affinities of EK for its cleavage sites in AhR6–C/EBP

Analysis of the SDS–PAGE of cleavage reactions containing 3 M urea (Fig. 6) further increases our understanding of the relative affinities of EK for its different cleavage sites in AhR6–C/EBP. Enterokinase cleaves our protein at sites preceded by either Lys or Arg. Light et al. have reported that the minimum sequence for EK cleavage requires a basic amino acid (Lys or Arg) at the P1 position, which directly precedes the cleavage site, and an acidic amino acid (Asp or Glu) at the P2 site, which precedes P1; they also reported that the rate of hydrolysis of a cleavage site increases when more acidic residues occupy the P3–P5 positions [17]. Although all of the cleavage sites observed in AhR6–C/EBP (–DK^{31}, –SR^{36}, –IK^{55}, –DR^{72}, and –DR^{91}) have either Lys or Arg in the P1 position (Fig. 5), not all of the cleavage sites have an acidic amino acid in the P2 position (–SR^{36}, –IK^{55}); hence, the requirement for an acidic residue at P2 is not as stringent as that for a basic residue at P1 in AhR6–C/EBP.

Moreover, SDS–PAGE analysis of EK reactions containing 3 M urea shows that within the first 30 min of incubation of AhR6–C/EBP with EK (Fig. 6), the cleavage product present in the largest amount is that resulting from cleavage at –SR^{36} (Fig. 6, band C, 9210 Da). Serine is a polar, uncharged amino acid, yet cleavage at –SR^{36} is the fastest of all the sites (–IK^{55}, –DR^{72}, and –DR^{91}). Regarding AhR6–C/EBP, EK proteolysis does not require an acidic residue in the P2 position; additionally, any site containing a basic residue at P1 and an acidic residue at P2 is not necessarily an actual cleavage site for EK. For example, we were unable to see any cleavage of our protein at –DR^{65} (Fig. 1), although we did observe cleavage at –DR^{72} and –DR^{91} cleavage sites. A probable explanation is that in AhR6–C/EBP, –DR^{65} is preceded by four basic residues that may inhibit proteolysis [17].

Modification of the EK cleavage site as an alternative approach to enhance cleavage efficiency and specificity is protein-specific

Others have examined the effects of residues downstream from the canonical EK site in their efforts to enhance the specificity of enterokinase cleavage. Hosfield and Lu investigated the influence of the first downstream residue after the canonical site on the EK cleavage efficiency of their glutathione *S*-transferase-calmodulin fusion protein (GST–CaM) [18]. They found that EK is permissive regarding the amino acid immediately downstream of the EK recognition site in their protein. Of the residues occupying this position, smaller amino acids allow for more efficient cleavage, whereas those that create steric hindrance in the binding pocket result in lower cleavage efficiency. Similarly, Liew et al. studied the effect of addition of an SRLLR motif immediately downstream of the EK target site on cleavage efficiency of their thioredoxin-fused N-terminal proCNP (NT-proCNP) [19]. They reported steady increase in the rate of hydrolysis of their protein at the DDDDK site upon incremental additions of SRLLR motifs downstream from the DDDDK site. In our case, we took a different approach to address the problem of proteolytic specificity. By addition of denaturant, we were able to achieve increased EK cleavage specificity of AhR6-C/EBP without any modifications to the protein construct. This experimental modification is easier to implement and may be more generally applicable.

We believe that the problem of inefficient cleavage in AhR6–C/EBP at the canonical EK site stems from its inaccessibility, and thus, adventitious sites of related sequence can outcompete the canonical site for EK proteolysis. His-tagged AhR6–C/EBP can homodimerize via the C/EBP leucine zipper and is prone to inclusion body formation, both of which can adversely affect EK's ability to bind to its target site. Others have also observed similar phenomena of adventitious cleavage and have speculated that this may result from inaccessibility of the EK target site during proteolysis [5,20]. For instance, Liew et al. observed less efficient cleavage at the adventitious LKGDR sequence when they moved it from the C-terminus of NT-proCNP to a more internal location [5].

Enhancement of cleavage specificity at the DDDDK site achieved by methods involving replacing residues upstream or downstream of the canonical site is protein-specific, and hence, may not be universally applied to all fusion proteins. As an example, Hosfield and Lu reported that in their GST-CaM fusion protein, 84% cleavage at the DDDDK site resulted when Asp occupied the position immediately after the canonical site [18]. In our case, the EK target site in AhR6–C/EBP is also immediately followed by Asp (Fig. 1). Yet, until standard experimental conditions were altered by addition of denaturant, little to no cleavage was observed after the canonical EK target site (Figs. 2 and 3).

Thus, we suspect that such methods involving enhancement of EK cleavage by modifying the target site and flanking regions depend on protein structure. AhR6–C/EBP may be dimerized, even aggregated potentially, thereby restricting access to the EK target; opening this structure by use of denaturant can alleviate this problem. In contrast, Hosfield and Lu's GST–CaM fusion may not have a tendency to dimerize or form aggregates, and hence, this structural issue may not be the reason for their EK cleavage inefficiency.

Others have also observed that protein structure can be responsible for successful proteolysis. Kim et al. studied protease cleavage efficiency of glucagon-fused human interleukin-2 by modifying the sequence upstream from the canonical EK site [20]. Human Interleukin-2 (hIL-2) contains one disulfide linkage and is folded into a bundle-shaped protein comprising four α-helices [21]. Heterologous expression of hIL-2 involves formation of heterogeneous aggregates by nonspecific intermolecular disulfide linkages [22]. Kim et al. carried out expression of hIL-2 by dissolving insoluble aggregates by a pH-shift and performed subsequent EK cleavage without any denaturing agent in the reaction buffer [20]. They reported low cleavage efficiency, which they speculated to be due to the EK recognition site being sterically hindered. In order to enhance the cleavage efficiency, they inserted the negatively charged DDDD sequence upstream from the EK target site. Contrary to their expectations, this modification to the target site did not enhance EK cleavage specificity; rather, they observed increased adventitious cleavage of hIL-2, and they concluded that hIL-2 was inherently an inefficient substrate for EK [20]. Zhang et al. prepared recombinant 6×His-tagged hepcidin, a highly disulfide-bonded peptide containing eight Cys residues and four intermolecular disulfide bonds [16]. They used 8 M urea to promote solubility and prevent precipitation of hepcidin resulting from disulfide-linked multimerization during protein preparation, and their EK reaction buffer contained 2 M urea. After a 16 h EK cleavage reaction, they obtained intact, tag-free hepcidin in ~20% yield [16].

With hIL-2, Kim et al. did not use denaturant; rather, they tried to increase EK recognition by modifying the target site [20]. For hIL-2, this approach was not successful, although for GST–CaM and NT-proCNP, this type of strategy was effective [18,19]. Zhang et al. used high concentrations of urea with hepcidin and successfully obtained desired EK cleavage product [16]. We surmise that an open, accessible recognition site for EK is critical for efficient

proteolysis, and that the success of different strategies for promoting open protein structure can be case-dependent.

Addition of urea to our EK reactions of bZIP-like AhR6–C/EBP not only promotes the desired cleavage after the canonical EK site, but also inhibits additional proteolytic cleavage at adventitious sites: we suspect that this is achieved when the canonical site is as accessible as any other potential cleavage site, and is therefore better able to compete with these adventitious sites for recognition by EK. Addition of denaturant to the EK reaction buffer may therefore be a general strategy for discouraging protein structure that shields the intended cleavage site from EK recognition. However, use of higher amounts of urea (~5 M) had an inhibitory effect on the activity of EK during cleavage of AhR6–C/EBP (Fig. 4). Therefore, promoting accessibility of the canonical EK target site can increase proteolytic specificity and cleavage yield, and general strategies that promote a more open structure should be useful for preparation of proteins requiring endoprotease treatment.

Acknowledgments

We thank Alevtina Pavlenco for technical assistance and are grateful for funding from NIH (R01 GM069041), Premier's Research Excellence Award (PREA), Canadian Foundation for Innovation/Ontario Innovation Trust (CFI/OIT), and the University of Toronto. The authors declare no competing financial interests.

References

[1] M. Hammarström, N. Hellgren, S. Van den Berg, H. Berglund, T. Härd, Rapid screening for improved solubility to small human proteins produced as fusion proteins in *Escherichia coli*, Protein Sci. 11 (2002) 313–321.
[2] R.J. Jenny, K.G. Mann, R.L. Lundblad, A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa, Protein Expr. Purif. 31 (2003) 1–11.
[3] A. Light, H. Janska, Enterokinase (enteropeptidase): comparative aspects, TIBS 14 (1989) 110–112.
[4] C.-C. Hung, S.-H. Chiou, Expression of a kallikrein-like protease from the snake venom: engineering of autocatalytic site in the fusion protein to facilitate protein refolding, Biochem. Biophys. Res. Commun. 275 (2000) 924–930.
[5] O.W. Liew, J.P.C. Chong, T.G. Yandle, S.O. Brennan, Preparation of recombinant thioredoxin fused N-terminal proCNP: analysis of enterokinase cleavage products reveals new enterokinase cleavage sites, Protein Expr. Purif. 41 (2005) 332–340.
[6] A. Hillar, G. Otulakowski, H. O'Brodovich, Purification and characterization of a recombinant rae prohaptoglobin expressed in baculovirus-infected Sf9 insect cells, Protein Expr. Purif. 55 (2007) 246–256.
[7] K. Struhl, Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eucaryotic transcriptional regulatory proteins, Trends Biochem. Sci. 14 (1989) 137–140.
[8] W.H. Landschulz, P.F. Johnson, S.L. McKnight, The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins, Science 240 (1988) 1759–1764.
[9] A.R. Lajmi, T.R. Wallace, J.A. Shin, Short, hydrophobic, alanine-based proteins based on the bzip motif: overcoming inclusion body formation and protein aggregation during overexpression, purification, and renaturation, Protein Expr. Purif. 18 (2000) 394–403.
[10] G.H. Bird, A.R. Lajmi, J.A. Shin, Manipulation of temperature improves solubility of hydrophobic proteins and cocrystallization with matrix for analysis by MALDI-TOF mass spectrometry, Anal. Chem. 74 (2002) 219–225.
[11] G.H. Bird, J.A. Shin, MALDI-TOF mass spectrometry characterization of hydrophobic basic region/leucine zipper proteins, Biochim. Biophys. Acta (Protein Struct. Mol. Enzyme) 1597 (2002) 252–259.
[12] S.I. Choi, H.W. Song, J.W. Moon, B.L. Seong, Recombinant enterokinase light chain with affinity tag: expression from *Saccharomyces cerevisiae* and its utilities in fusion protein technology, Biotechnol. Bioeng. 75 (2001) 718–724.
[13] M. Kenig, S. Peternel, V. Gaberc-Porekar, V. Menart, Influence of the protein oligomericity on final yield after affinity tag removal in purification of recombinant proteins, J. Chromatogr. A 1101 (2006) 293–306.
[14] Q.H. Liu, J.P. Lin, M.Y. Liu, X.Y. Tao, D.Z. Wei, X.Y. Ma, S.L. Yang, Large scale preparation of recombinant human parathyroid hormone 1–84 from *Escherichia coli*, Protein Expr. Purif. 54 (2007) 212–219.
[15] P. Turner, O. Holst, E.N. Karlsson, Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration, Protein Expr. Purif. 39 (2005) 54–60.
[16] H. Zhang, Q. Yuan, Y. Zhu, R. Ma, Expression and preparation of recombinant hepcidin in *Escherichia coli*, Protein Expr. Purif. 41 (2005) 409–416.
[17] A. Light, H.S. Savithri, J.J. Liepnieks, Specificity of bovine enterokinase toward protein substrates, Anal. Biochem. 106 (1980) 199–206.
[18] T. Hosfield, Q. Lu, Influence of the amino acid residue downstream of Asp_4 Lys on enterokinase cleavage of a fusion protein, Anal. Biochem. 269 (1999) 10–16.
[19] O.W. Liew, P.C.J. Chong, Y.Z. Lim, C.X. Ang, Y.C.A. Lau, T.G. Yandle, S.O. Brennan, An SRLLR motif downstream of the scissile bond enhances enterokinase cleavage efficiency, Biochimie 89 (2007) 21–29.
[20] S.W. Kim, J.B. Kim, W.S. Lee, W.H. Jung, J.M. Ryu, H.W. Jang, Y.B. Jo, J.K. Jung, J.H. Kim, Enhanced protease cleavage efficiency on the glucagon-fused interleukin-2 by the addition of synthetic oligopeptides, Protein Expr. Purif. 55 (2007) 159–165.
[21] S.M. Liang, B. Allet, K. Rose, M. Hirschi, C.M. Liang, D.R. Thatcher, Characterization of human interleukin-2 derived from *Escherichia coli*, Biochem. J. 229 (1985) 429–439.
[22] D.Y. Kim, J. Lee, V. Saraswat, Y.H. Park, Glucagon-induced self-association of recombinant proteins in *Escherichia coli* and affinity purification using a fragment of glucagon receptor, Biotechnol. Bioeng. 69 (2000) 418–428.
[23] J.A. Shin, Specific DNA binding peptide derivatized solid support, Bioorg. Med. Chem. Lett. 7 (1997) 2367–2372.