Comparative Biochemistry and Physiology, Part B 156 (2010) 40–47

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology, Part B

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

N-glycosylation at non-canonical Asn-X-Cys sequence of an insect recombinant cathepsin B-like counter-defense protein

Yong Hun Chi [a,b], Yoon Duck Koo [a], Susie Y. Dai [c,d], Ji-Eun Ahn [a], Dae-Jin Yun [b], Sang Yeol Lee [b], Keyan Zhu-Salzman [a,e,*]

[a] Department of Entomology, Texas A&M University, College Station, TX 77843, USA
[b] Division of Applied Life Sciences, Environmental Biotechnology National Core Research Center, Gyeongsang National University, Jinju, Republic of Korea
[c] Office of the Texas State Chemist, Texas A&M University, College Station, TX 77843, USA
[d] Department of Veterinary Pathobiology, Texas A&M University, College Station, TX 77843, USA
[e] Vegetable and Fruit Improvement Center, Texas A&M University, College Station, TX 77843, USA

ARTICLE INFO

Article history:
Received 29 December 2009
Received in revised form 29 January 2010
Accepted 30 January 2010
Available online 6 February 2010

Keywords:
Cathepsin B
Counter-defense
N-glycosylation
Pichia
Site-directed mutagenesis

ABSTRACT

CmCatB, a cowpea bruchid cathepsin B-like cysteine protease, facilitates insects coping with dietary protease inhibitor challenge. Expression of recombinant CmCatB using a *Pichia pastoris* system yielded an enzymatically active protein that was heterogeneously glycosylated, migrating as a smear of ≥50 kDa on SDS-PAGE. Treatment with peptide:*N*-glycosidase F indicated that *N*-glycosylation was predominant. CmCatB contains three *N*-glycosylation Asn-X-Ser/Thr consensus sequences. Simultaneously replacing all three Asn residues with Gln via site-directed mutagenesis did not result in completely unglycosylated protein, suggesting the existence of additional atypical glycosylation sites. We subsequently investigated potential *N*-glycosylation at the two Asn-X-Cys sites (Asn^{100} and Asn^{236}) in CmCatB. Asn to Gln substitution at Asn^{100}-X-Cys on the background of the double mutation at the canonical sites (m1m2, Asn^{97}→Gln and Asn^{207}→Gln) resulted in a single discrete band on the gel, namely m1m2c1 (Asn^{97}→Gln, Asn^{207}→Gln and Asn^{100}→Gln). However, another triple mutant protein m1m2c2 (Asn^{97}→Gln, Asn^{207}→Gln and Asn^{236}→Gln) and quadruple mutant protein m1m2c1c2 were unable to be expressed in *Pichia* cells. Thus Asn^{236} appears necessary for protein expression while Asn^{100} is responsible for non-canonical glycosylation. Removal of carbohydrate moieties, particularly at Asn^{100}, substantially enhanced proteolytic activity but compromised protein stability. Thus, glycosylation could significantly impact biochemical properties of CmCatB.

1. Introduction

Like mammals, herbivorous insects need to obtain 10 essential amino acids from their diets to meet their nutritional requirements. They utilize digestive proteases to break down food proteins and make amino acids available for protein synthesis. Insect predation poses a strong selection pressure on plants to develop defense strategies, one of which is the production of protease inhibitors to impair insect digestive enzymes. Insects, on the other hand, have actively evolved adaptive mechanisms to evade natural plant defenses. To counteract plant protease inhibitors, insects mobilize genomic resources and upregulate inhibitor-insensitive enzymes to mitigate the negative impact. The cowpea bruchid, *Callosobruchus maculatus* (Coleoptera: Bruchidae) is a grain storage pest employing cysteine proteases as their major digestive enzymes (Murdock et al., 1987). Larvae feed and develop inside the seeds, fully responsible for grain damage. Emerging from the infested seeds, the adults no longer need to feed, but mate and lay eggs on the seed surface to start another feeding cycle. Each generation only takes about 40 days. When ingesting a diet containing the soybean cysteine protease inhibitor scN, the bruchids undergo broad transcriptomic changes in the larval midgut, including upregulation of a large number of genes encoding digestive enzymes (Chi et al., 2009). Such changes enable the bruchids to (i) overproduce existing, scN-sensitive digestive proteases to out-titer the inhibitors; (ii) increase expression of scN-insensitive proteases in the digestive tract; and (iii) activate proteases that hydrolyze and thus detoxify scN. The combined net outcome is that insects recover from amino acid deficiencies and resume normal feeding and development later on (Zhu-Salzman et al., 2003). Such tactics are commonly adopted by a variety of herbivorous insects to cope with natural plant defensive protease inhibitor proteins (Jongsma et al., 1995; Bown et al., 1997; Broadway, 1997; Jongsma and Bolter, 1997; De Leo et al., 1998; Cloutier et al., 2000; Mazumdar-Leighton and Broadway, 2001; Zhu-Salzman et al., 2003; Brunelle et al., 2004; Liu et al., 2004; Bayes et al., 2005; Koo et al., 2008).

* Corresponding author. Department of Entomology, Texas A&M University, College Station, TX 77843, USA. Tel.: +1 979 458 3357; fax: +1 979 862 4790.
E-mail address: ksalzman@tamu.edu (K. Zhu-Salzman).

1096-4959/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpb.2010.01.017

Cowpea bruchid cathepsin B-like protein (CmCatB) is a counter-defense protein belonging to the papain superfamily of cysteine proteases, as do the major cathepsin L-like digestive proteases of this insect, previously named CmCPs. *CmCatB* was among transcripts most strongly induced by dietary scN despite the fact that its expression was undetectable in the unchallenged insect (Moon et al., 2004; Chi et al., 2009). This is distinct from the major digestive CmCP enzymes which are constitutively expressed in unchallenged group. Expression of *CmCatB* is mainly associated with larval guts, the developmental stage and tissue type tied to food ingestion (Koo et al., 2008). The heterologously expressed recombinant protein exhibited cathepsin B-type cysteine protease activity. More interestingly, the enzymatic activity was unaffected by the presence of scN, in contrast to scN-susceptible CmCPs, explaining the scN-insensitive cysteine protease activity observed in dietary scN-challenged insects. The scN-insensitive enzymatic activity conferred by CmCatB is likely beneficial to cowpea bruchids as it may warrant digestion of dietary proteins to meet the needs for amino acids while major digestive CmCPs are inhibited. Presumably, the occluding loop in CmCatB, which does not occur in cathepsin L-like digestive enzymes, blocked the access of scN to its catalytic site cleft (Moon et al., 2004).

Although *Pichia pastoris*-expressed CmCatB showed specific cathepsin B activity, purified recombinant protein appeared to be heterogeneously glycosylated. The smeared appearance on SDS-PAGE became a discrete band after peptide:*N*-glycosidase F (PNGase F) treatment, indicating *N*-linked glycosylation (Koo et al., 2008). In this study, we attempted to produce a non-glycosylated mutant form of CmCatB by disrupting potential glycosylation sites. We found modification of Asn residues at both usual Asn-X-Ser/Thr and atypical Asn-X-Cys sequences. Furthermore, the carbohydrate moiety was found to impact enzymatic activity and protein stability.

2. Material and methods

2.1. Site-directed mutagenesis

Substitutions of Asn residues located at potential *N*-glycosylation sites of CmCatB with Gln were performed by site-directed mutagenesis (Weiner et al., 1994; Ishii et al., 1998). The sense and anti-sense primer sequences used to introduce specific mutations are shown in Table 1. PCR reactions were performed using 25 μM of each primer for 16 cycles of 30 s at 95 °C, 1 min at 55 °C, and 10 min at 68 °C, using previously constructed pCR2.1-CmCatB plasmid (Koo et al., 2008) as the template. Double (m1m2), triple (m1m2m3, m1m2c1, m1m2c2) and quadruple (m1m2c1c2) mutations were obtained by sequential point mutations. Following PCR amplification, the amplified products were subjected to *Dpn* I digestion to remove the parental DNA template, and subsequently transformed into *E. coli* DH5α competent cells. All constructs were subjected to DNA sequencing analyses to verify the entire coding sequences.

Table 1
Primers synthesized for site-directed mutagenesis.

Mutations	Residues altered	Sequence (5′ to 3′)
m1	N^{97} to Q	Sense: GTCTCTGGTGGAGAGTATCAATCAACTAATGGCTGC
		Anti-sense: GCAGCCATTAGTTGATTGATACTCTCCACCAGAGAC
m2	N^{207} to Q	Sense: GGTTGGGGAATAGAACAAGGTACGTATCCTTACTGG
		Anti-sense: CCAGTAAGGATACGTACCTTGTTCTATTCCCCAACC
m3	N^{111} to Q	Sense: CTTCCAAGATGCCAACCATCATGCAAGACACTTTAC
		Anti-sense: GTAAAGTGTCTTGCATGATGGTTGGCATCTTGGAAG
c1	N^{100} to Q	Sense: GGAGAGTATCAATCAACTCAAGGCTGCATGTCATACCC
		Anti-sense: GGGTATGACATGCAGCCTTGAGTTGATTGATACTCTCC
c2	N^{236} to Q	Sense: CAAGATATGGCGAGGCAAACAAGAATGTGGAATTGAGG
		Anti-sense: CCTCAATTCCACATTCTTGTTTGCCTCGCCATATCTTG

The underlined bases indicate the mutated codons.

2.2. Expression of CmCatB mutated proteins in Pichia

For expression constructs, cDNAs encoding proproteins of *CmCatB* and its single, double, triple and quadruple mutations without the signal peptide and the stop codon were PCR-amplified for 35 cycles of 30 s at 95 °C, 1 min at 55 °C, and 1 min at 72 °C. The primers are: sense 5′-GTAGTACTCGAG**AAAAGAGAGGCTGAAGCT**CAACCAGAGCTG-GATTTCCTCTCA-3′, and anti-sense 5′-GTGCTATCTAGACGTGGTAGTC-CAGCCGTGA-3′. Restriction sites *Xho* I and *Xba* I (underlined) as well as sequences for Kex2 and Ste13 cleavage sites (bold faced) were incorporated into primers. After restriction digestion, the fragments were cloned into the pPIC6αA vector (Invitrogen, Carlsbad, CA, USA) in-frame with α-factor signal sequence and hexa-His tag, and the constructs were transformed into *E. coli* XL1-Blue cells. As designed by the manufacturer, the α-factor signal sequence was subsequently auto-processed by endogenous Kex2 and Ste13 in a two-step excision process (Invitrogen).

Once the sequences were confirmed, the constructs were linearized with *Sac* I and transformed into *P. pastoris* host strain X33 (Invitrogen). Transformants were selected on YPDB agar plates containing 1% yeast extract, 2% peptone, 2% glucose, and 2% agar and blasticidin at final concentration of 300 μg/mL. Since pPIC6α vector does not contain a yeast origin of replication, transformants can only be isolated if recombination occurs between the plasmid and the *Pichia* genome. Blasticidin-resistant *Pichia* colonies were streaked again followed by PCR screening for DNA integration using 5′AOX1 and 3′AOX1 primers as directed by the manufacturer's protocol. Selected colonies were incubated in 10 mL of BMGY medium (1% yeast extract, 2% peptone, 1.34% YNB, 100 mM potassium phosphate, pH 6.0, 4×10^{-5}% biotin and 1% glycerol) at 29 °C for 24 h with shaking at 250 rpm. These 10 mL seed cultures were transferred to 200 mL BMGY medium and continuously grown for approximately 24 h at 29 °C until the cells reached an OD_{600} of 2. The cells were harvested by centrifugation at 350 *g* for 5 min and the cell pellet was resuspended in 200 mL BMMY medium (BMGY with 0.1% methanol instead of glycerol) and further incubated at 29 °C for 72 h with shaking. Methanol was added every 24 h to a final concentration of 0.5% as the sole carbon source to induce protein expression.

2.3. Purification of CmCatB and mutated proteins from yeast culture

The three-day cultures were collected and the cells were removed by centrifugation at 2500 *g* for 10 min. Approximately 200 mL supernatant containing all secreted proteins was subjected to overnight ammonium sulfate precipitation (70% saturation, pH 7.0) at 4 °C. The solution was centrifuged for 30 min at 25,000 *g* and 4 °C. The protein pellet was dissolved in 100 mL binding buffer (50 mM sodium phosphate buffer, 500 mM NaCl, 5 mM imidazole, pH 7.9) and loaded on a Ni^{2+}-chelate affinity column (Amersham Pharmacia Biotech, Piscataway, NJ, USA). After the column was washed with binding and washing buffers (50 mM sodium phosphate buffer, 500 mM NaCl, 60 mM imidazole, pH 7.9), bound proteins were eluted in 25 mM sodium phosphate buffer, 250 mM NaCl, 500 mM imidazole, pH 7.9, and concentrated by adding ammonium sulfate to 70% saturation and subjected to SDS-PAGE. The Bradford protein assay was used to quantify proteins (Bradford, 1976). For immunoblot analysis, mouse penta-His monoclonal antibody (Qiagen Inc, Valencia, CA,