Phytochemistry 70 (2009) 465–472

Contents lists available at ScienceDirect

Phytochemistry

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

Production and recovery of recombinant propapain with high yield

Debi Choudhury, Sumana Roy, Chandana Chakrabarti, Sampa Biswas, J.K. Dattagupta *

Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700064, India

ARTICLE INFO

Article history:
Received 22 October 2008
Received in revised form 2 February 2009
Available online 9 March 2009

Keywords:
Propapain
Cysteine protease
High yield expression
Inclusion body
Refolding
Activation

ABSTRACT

Papain (EC 3.4.22.2), the archetypal cysteine protease of C1 family, is of considerable commercial significance. In order to obtain substantial quantities of active papain, the DNA coding for propapain, the papain precursor, has been cloned and expressed at a high level in *Escherichia coli* BL21(DE3) transformed with two T7 promoter based pET expression vectors – pET30 Ek/LIC and pET28a$^+$ each containing the propapain gene. In both cases, recombinant propapain was expressed as an insoluble His-tagged fusion protein, which was solubilized, and purified by nickel chelation affinity chromatography under denaturing conditions. By systematic variation of parameters influencing the folding, disulfide bond formation and prevention of aggregate formation, a straightforward refolding procedure, based on dilution method, has been designed. This refolded protein was subjected to size exclusion chromatography to remove impurities and around 400 mg of properly refolded propapain was obtained from 1 L of bacterial culture. The expressed protein was further verified by Western blot analysis by cross-reacting it with a polyclonal anti-papain antibody and the proteolytic activity was confirmed by gelatin SDS–PAGE. This refolded propapain could be converted to mature active papain by autocatalytic processing at low pH and the recombinant papain so obtained has a specific activity closely similar to the native papain. This is a simple and efficient expression and purification procedure to obtain a yield of active papain, which is the highest reported so far for any recombinant plant cysteine protease.

1. Introduction

Cysteine proteases are of immense biological importance and are found in almost all eukaryotes, bacteria and viruses. Among them, the cysteine proteases of plant origin, especially the enzymes extracted from the tropical plant *Carica papaya* (papain, caricain, chymopapain, papaya proteinase IV), are of considerable commercial significance because of their extensive use in various industrial processes (Rawlings and Barrett, 1994). Of these, Papain (EC 3.4.22.2) has been the primary source of experimental data in the superfamily and has been characterized most extensively from kinetic and structural point of view (Baker and Drenth, 1987; Polgár, 1989). It is the first protease whose crystal structure was determined (Drenth et al., 1968) and was later refined to 1.65 Å in 1984 by Kamphuis et al. (1984). The papain molecule is folded into two compact interacting domains, delimiting a cleft at the surface between these domains. The two catalytic residues, Cys25 and His159 are located at this interface on opposite domains of the molecule. Cys25 is part of an α-helix at the surface of one domain, while His159 is in a β-sheet at the surface of the other domain of the enzyme.

Papain has many uses and functions in a variety of industries: clarifying beer, meat tenderization, preservation of spices, contact lens cleaners, detergents, pet food palatability, digestive aids, blood stain remover, blood coagulant and cosmetics (Grzonka et al., 2007). Papain is also widely used for many medical and para-medical purposes such as to assist protein digestion in chronic dyspepsia, gastric fermentation, gastritis, removal of necrotic tissue, preparation of tyrosine derivatives for the treatment of Parkinsonism, preparation of tetanus vaccines, skin cleansing agents, acne treatment, etc. (Grzonka et al., 2007).

Since papain has such a variety of uses, it is desirable to obtain this protein in substantial quantities for both basic research and industrial use. The enzyme preparation from the plant source will depend on the climatic conditions for growth and methods used for its extraction and purification and hence should not be used for purposes where a highly pure homogeneous preparation is needed. To overcome these shortcomings, it is of much advantage to obtain a wild-type recombinant form of the enzyme. This is also a pre-requisite for introducing selective changes into the primary sequence so as to modify and tailor the properties of a protein by directed evolution or rational design to suit many of its sought-after commercial requirements as has been done by Liu et al. (2007) and Mosavi and Peng (2003) to increase protein solubility. Likewise, recombinant papain is planned to be prepared by us with an ultimate goal to have a stable papain-like cysteine protease through structure based (Guha Thakurta et al., 2004) site-directed mutagenesis.

* Corresponding author. Tel.: +91 33 2321 4986; fax: +91 33 2337 4637.
E-mail address: jibank.dattagupta@saha.ac.in (J.K. Dattagupta).

0031-9422/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2009.02.001

The papain precursor gene, prepropapain (Cohen et al., 1986) has been cloned and expressed earlier by other groups to varying levels, either in part or as a whole, in at least three different expression systems. The host organisms used included baculovirus/insect – *Autographa californica*/*Spodoptera frugiperda* Sf9 (Vernet et al., 1990), yeast – *Saccharomyces cerevisiae* (Vernet et al., 1993; Ramjee et al., 1996) and bacteria – *Escherichia coli* (Vernet et al., 1989; Cohen et al., 1990; Taylor et al., 1992). The prepropapain expression in *E. coli* was shown to be inhibited by the 26 amino acid signal peptide (Vernet et al., 1989) and expression of recombinant propapain (without signal peptide) in *E. coli* resulted in insoluble granular aggregates (inclusion bodies). Taylor et al. (1992) however, were able to renature the insoluble protein to a yield of ~3 mg of active processed papain per litre of *E. coli* culture. Similarly, expression of propapain in a baculovirus resulted in the production of ~0.3 mg/L of papain (Vernet et al., 1991). It was also pointed out that the baculovirus expression system is not ideal for a mutagenesis-based structure–function analysis of propapain (Vernet et al., 1993). Although Ramjee et al. (1996) were successful in expressing propapain in the soluble form in yeast, the yield obtained was only ~1.7 mg/L of culture. Thus, as mentioned above, each expression system has specific advantages and disadvantages in terms of yield, production costs, time required to generate the active protease and most importantly the researcher's application. In a review by Bromme et al. (2004), where a comparison of different expression systems for papain-like cysteine proteases was done, the *Pichia* system was found to be most efficient especially for Cathepsins. At the same time the bacterial expression system was recommended by them when an appropriate refolding protocol is available that allows the generation of active proteases.

In this paper, we report an improved protocol to obtain high level expression of papain precursor in *E. coli* using two T7 promoter based vectors – pET28a$^+$ and pET30 Ek/LIC. The protein was expressed as inclusion bodies and we could obtain ~400 mg of properly folded propapain per litre of *E. coli* culture by designing a scheme for refolding protocol. This refolded propapain could be converted to active mature papain by autocatalytic processing at low pH.

2. Results

2.1. PCR amplification, cloning and expression of the propapain coding region

The 1035 bp cDNA encoding the prepropapain extracted by reverse transcription (RT) and polymerase chain reaction (PCR) from leaf polyA$^+$ RNA was cloned (Fig. 1a) and the nucleotide sequence determined using appropriate internal primers. The sequence was found to be identical to that reported by Cohen et al. (1986) (GenBank accession no. M15203).

The propapain coding region was amplified from the above prepropapain clone by primers designed in such a way so as to attach an ATG codon at the 5′ end of the gene. The 78 bp (26 amino acids) signal peptide was removed before cloning the propapain gene because it has been observed that this fragment is toxic to bacteria and hence, hinders heterologous expression of papain-like cysteine proteases in *E. coli* (Vernet et al., 1989). The wild-type propapain gene was successfully cloned in *E. coli* pET30 Ek/LIC and pET28a$^+$ expression vectors (Fig. 1b). The recombinant protein expressed in pET30 Ek/LIC contained an Enterokinase cleavage site (DDDDK) and six His residues, and the one expressed in pET28a$^+$ had a Thrombin cleavage site (GLVPRGS) and six His residues at the N-terminal to simplify the purification procedures. These constructions were confirmed by DNA sequencing of the respective transformed recombinant plasmids.

E. coli strains BL21(DE3) and BL21(DE3)pLysS were tested for expression of propapain from both pET30 Ek/LIC and pET28a$^+$ clones. It was observed that the level of expression for BL21(DE3) was higher in both cases, hence all the subsequent processes were done with the protein expressed from this strain. Culture samples withdrawn at hourly intervals and after overnight induction of both the clones were analyzed by SDS–PAGE and the results indicated that the optimum time for expression was 5 h, since neither the cell density nor the amount of expressed protein increases after this. It was also observed that the level of expression of both clones was quite high and the newly synthesized protein represented the major fraction of the cell lysate and had the expected size for

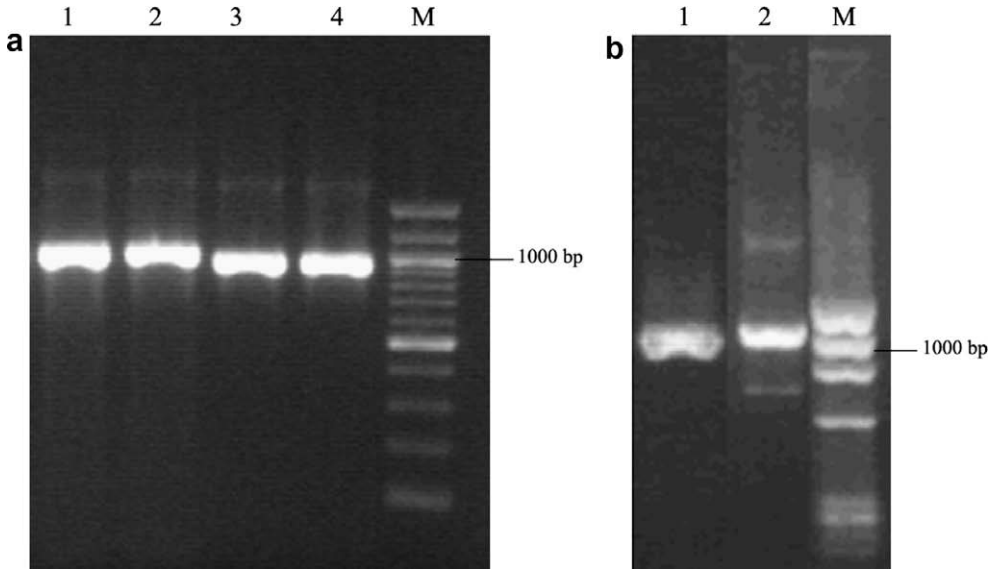

Fig. 1. Cloning of prepropapain and propapain cDNA from leaves of *Carica papaya*. (a) Prepropapain cDNA extracted from *Carica papaya* leaves and propapain cDNA amplified from prepropapain cDNA clone (visualized with ethidium bromide in 1% agarose gel). Lanes 1–2, PCR product amplified from cDNA with prepropapain-specific primers; lanes 3–4, PCR product amplified from prepropapain clone with propapain-specific primers; M, molecular mass marker. (b) Cloning of propapain in pET30 Ek/LIC and pET28a$^+$ expression vectors. Lane 1, pET28a$^+$ clone – PCR product with papain gene-specific forward and reverse primers; lane 2, pET30 Ek/LIC clone – PCR product with pET30 upstream and LIC reverse primer; M, molecular mass marker.

propapain containing respective fusion-tag sequences (~42 kDa: Fig. 2a). In addition, optimization of the IPTG concentration for maximum yield was done. The optimal IPTG concentration for induction was found to be 0.5 mM for pET28-propap and 1 mM for pET30-propap, respectively. Since, the level of protein expression, purification/refolding protocol and final yield of the active protein is nearly same for both the expression systems, the subsequent results are for the protein expressed in pET30 Ek/LIC (pET30-propap) using BL21(DE3) as expression host.

2.2. Inclusion bodies: isolation, purification and solubilization

Majority of the recombinant protein was recovered in the pellet from high-speed centrifugation of the cell lysate, suggesting that propapain was expressed predominantly in inclusion bodies. We routinely obtained 6–7 g insoluble pellet from the lysate of 1 L of bacterial culture after 5 h of induction. Thorough washing of the pellet with 1–2 M Urea and Triton X-100, followed by washing with distilled water contributed to a significant increase in the

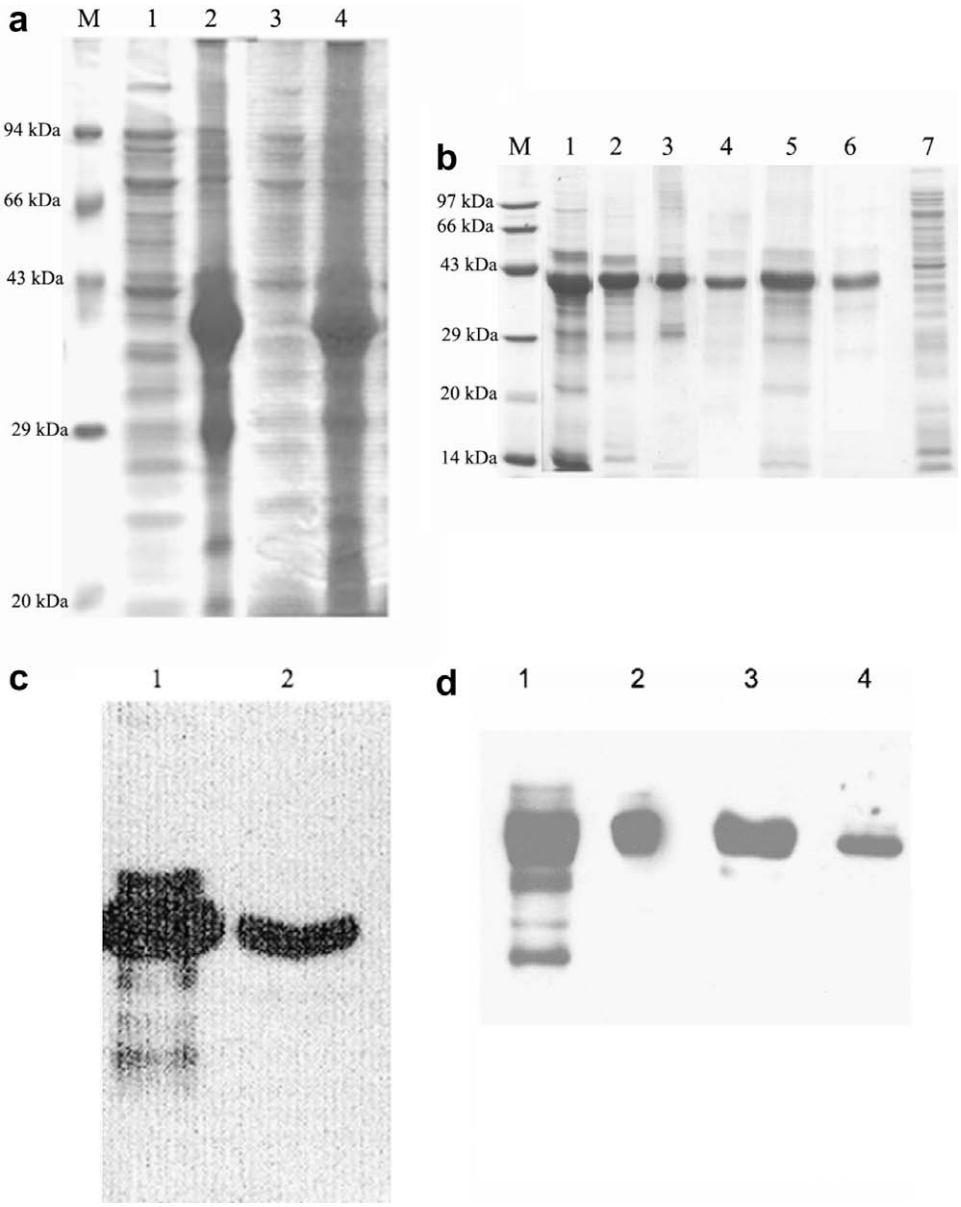

Fig. 2. SDS–PAGE analysis of expression, purification and refolding of recombinant propapain from *E. coli*. (a) Expression of propapain in pET30 Ek/LIC and pET28a$^+$ in BL21(DE3) cells. Lanes 1 and 3, total cell extract of bacteria containing pET28-propap and pET30-propap, respectively, prior to induction with IPTG; lanes 2 and 4, total cell extract of bacteria containing pET28-propap and pET30-propap respectively, 5 h post-induction with IPTG; M, protein molecular mass standards. Each lane represents pellet obtained from 1.5 ml of bacterial culture boiled with SDS–PAGE sample buffer, loaded on to 15% SDS–PAGE gel, run under reducing condition and stained with Coomassie blue. (b) Purification and refolding of recombinant propapain from inclusion bodies (15% gel). Lanes 1, 2 and 7 represent protein from 1.5 ml of bacterial culture. Lane 1, insoluble fraction of bacterial cell lysate; lane 2, inclusion bodies after Urea/Triton X-100/distilled water wash; lane 3, denatured propapain eluted from Ni–NTA column (~30 μg protein); lane 4, Ni–NTA fractions re-purified on Sephacryl S-200 column (~15 μg protein); lane 5, refolded and concentrated propapain (~40 μg protein); lane 6, refolded propapain dialyzed against 20 mM Tris–HCl (~10–15 μg protein); lane 7, soluble fraction of bacterial cell lysate and M, protein molecular mass standards. (c) Western blot analysis of recombinant propapain with polyclonal anti-His antibody. Lanes 1 and 2, solubilized inclusion bodies from ~500 μl of bacterial culture and purified refolded pro-papain (~5 μg protein), respectively. The amounts of protein indicated in this figure and in (d) below are the quantities of protein loaded on to the SDS–PAGE gels prior to blotting. (d) Western blot analysis of the recombinant precursor produced from *E. coli* cells transformed with BL21(DE3) with polyclonal anti-papain antibody. Lane 1: solubilized inclusion bodies from ~500 μl of bacterial culture, lane 2: Ni–NTA fractions of purified denatured propapain (~5 μg protein), lane 3: refolded propapain (~10 μg protein), lane 4: concentrated propapain (~2.5 μg protein).

purity of inclusion bodies (Fig. 2b, lane 2). The purified inclusion bodies were optimally solubilized with 8 M Urea. Following denaturing, the sample was purified by Ni–NTA affinity chromatography and the recombinant protein eluted as a single major band of about 42 kDa together with a few high and low molecular weight bands as assessed by SDS–PAGE (Fig. 2b, lane 3).

To eliminate these extra high and low molecular weight impurities, the protein eluted from Ni–NTA column was again purified by gel filtration chromatography on a Sephacryl S-200 column which yielded a highly purified protein as assessed by the single band on SDS–PAGE, as seen in Fig. 2b, lane 4.

2.3. Folding and activation

After gel filtration, only the fractions showing single bands at desirable positions on SDS–PAGE gel were pooled and refolded by rapid dilution into a refolding buffer at pH 8.6, yielding folded propapain (Fig. 2b, lane 5). The refolding procedure for propapain was systematically standardized for optimum temperature, pH, ionic strength of buffer, requirement for various additives like redox pair, arginine, etc. It was observed that the denatured protein folded best at pH ~8.5 and at 4 °C, while pH values above 8.8 and below 7.0 were not suitable for folding. Also, the folding process required addition of 0.5 M arginine, 1 mM GSH, 0.1 mM GSSG and 15% glycerol. It was also observed that the protein concentration needed to be around 50 μg/ml. A concentration above 100 μg/ml resulted in a large amount of insoluble protein and low yield of the refolded protein. An additional purification of the refolded protein by size exclusion chromatography helped in removing impurities like improperly folded protein from the properly folded propapain. Thus 1.0 g of protein was purified by Ni–NTA chromatography from 1 L of bacterial culture, which in turn provided 400 mg of refolded propapain. This refolded propapain could be autocatalytically activated to mature active papain at low pH. The purification steps and yield of propapain from *E. coli* BL21(DE3) are summarized in Table 1.

Expression of propapain in *E. coli* BL21(DE3) harboring pET30-propap plasmid and its subsequent isolation, purification and refolding were established by Western blot analysis. By using either the anti-His (Fig. 2c) or anti-papain blots (Fig. 2d), a distinct single band was observed. This confirmed that the expressed protein is a papain precursor tagged with Histidine.

The optimal conditions for *in vitro* activation of the refolded papain precursor (propapain) to mature papain were determined by systematic manipulation of the reaction mixture, temperature and time. The highest rates of activation were observed between pH 3.8 and 4.5, while the rates were greatly reduced at pH higher than 6.

The requirement for reducing conditions during the activation process was investigated with 20 mM DTT or 20 mM cysteine. It was observed that the propapain did not convert to papain even after prolonged incubation (12 h) with DTT whereas in presence of 20 mM cysteine, the activation occurred between 30 and 40 min at 50 °C.

In addition to study the effect of reducing agents, the time course and temperature required for activation were determined by incubating the reaction mixture for 0, 2, 5, 15, 20, 30, 40, 50, 60, 90, 120 min and overnight (12 h) at 20, 37 and 50 °C. It was observed that no activation occurred at 20 °C even after prolonged incubation. On the other hand, at 37 °C the activation of propapain to papain occurred after 2 h of incubation. The best result was obtained at 50 °C, where maximum conversion of propapain to papain occurred between 30 and 40 min of incubation (Fig. 3a and b).

The process of activation of propapain to mature papain was also monitored by SDS–PAGE gelatin gel zymograms. As can be seen from Fig. 4, the mature papain appeared as a single band when the activation was complete.

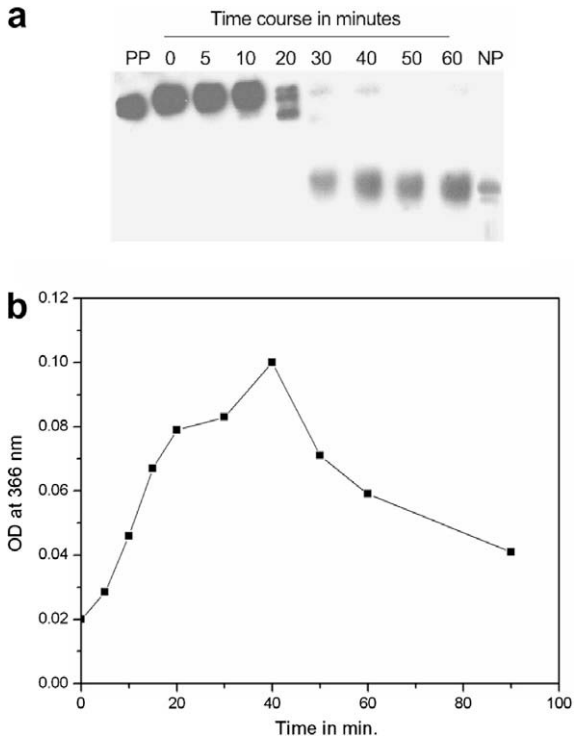

Fig. 3. Time course of activation of propapain to mature papain in presence of 20 mM cysteine at 50 °C. (a) Western blot profile of activation of 2 mg/ml propapain to mature papain for 0–60 min; lane PP: refolded propapain before activation; lane NP: SDS–PAGE of commercially available native papain to indicate the molecular weight position of the mature papain. (Lanes PP, 0–60 contain ~10–15 μg of protein loaded onto the SDS–PAGE gels prior to blotting. Lane NP contains ~5 μg of protein. (b) Activity profile of mature papain converted from propapain (25 μg) during activation.

Table 1
Purification of recombinant propapain from *E. coli* inclusion bodies[a].

Purification step	Protein (g)	Estimated propapain (g)	Yield per step (%)	Overall yield[b] (%)
Total cell lysate	4.7	–	–	–
Inclusion body	3.0	2.7	64	–
Ni–NTA affinity chromatography	2.5	1.9	83	63.3
Refolded using dilution method	1.3	1.2	52	40
Dialyzed protein	0.965	0.965	74	32
Gel filtration chromatography (Sephacryl S-100)	0.945	0.390	97 (40[c])	13

[a] The total cell lysate was obtained from 1 L of bacterial culture containing 6–7 g wet weight.
[b] Calculated based on the amount of protein solubilized using 8 M Urea.
[c] Figure in the bracket indicates the percent of properly folded propapain obtained from gel filtration experiment.

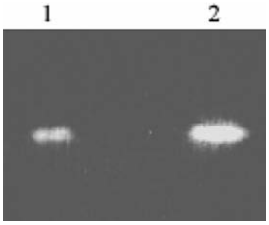

Fig. 4. Gelatin SDS–PAGE zymogram of the purified recombinant wild-type papain. Lanes 1 and 2 correspond to 5 μg and 10 μg of activated propapain, respectively. SDS gel (12%) was copolymerized with 0.1% gelatin and developed as described in Section 5. The gels were finally stained with Coomassie blue. Areas of proteolysis appear as clear regions in the gel.

2.4. Kinetics of recombinant activated papain

Azocasein was used as the substrate to examine the proteolytic activity of recombinant papain and, as shown in Table 2, the specific activity of recombinant papain is comparable (of the same order) with commercially available native papain.

3. Discussion

In this paper we present an efficient method for obtaining large amounts of refolded recombinant propapain overexpressed in *E. coli* as inclusion bodies. This procedure led to a production of about 400 mg of refolded propapain from 1 L of bacterial culture. The yields achieved in this work are the highest to date for any recombinant papain-like plant cysteine protease. Since the cysteine proteases present in the *C. papaya* plant are commercially important members of the cysteine protease superfamily (papain, caricain, chymopapain, papaya proteinase IV), several attempts have been made to clone and express these proteases especially using bacterial expression systems. Among these, the pro-forms of caricain (Revell et al., 1993), chymopapain (Taylor et al., 1999) and papain (Taylor et al., 1992) have been expressed using T7 promoter based pET expression vector pET3a which does not carry any fusion-tag. The yield of expression for procaricain is not reported. However, the expression of prochymopapain and propapain in *E. coli* resulted in yields of about 8–12 mg/L and ~3 mg/L of the recombinant mature chymopapain and papain, respectively, both of which are low compared to that reported here.

In the present work we have successfully expressed the protein using two pET expression vectors pET30 Ek/LIC and pET28a$^+$. We chose these vectors because both of them have His-tags, which help in purification of the target protein. Also the Ek/LIC system eliminates the restriction enzyme digestion step and thereby simplifies the cloning process.

It is known that overexpression of papain-like cysteine proteases generally results in inclusion bodies which can be solubilized with high concentrations of denaturants in presence of reducing agents. We found that extensive washing with small amounts of denaturants and multiple washes with distilled water resulted in >90% purity of the inclusion bodies. Overnight solubilization of the inclusion bodies at 4 °C yielded the best results. We also noticed that the presence of a reducing agent during the entire solubilization experiment decreased the yield of the protein. Hence, the Ni–NTA purified protein was reduced by incubating with DTT for 1 h prior to refolding.

We have also standardized the refolding protocol so as to obtain the maximum amount of refolded propapain. *In vivo* folding of proteins is rapid and is assisted by folding catalysts, such as disulfide isomerase, and molecular chaperones, which inhibit misfolding and aggregation (Woycechowsky and Raines, 2000). *In vitro* refolding of proteins that contain disulfide bonds is usually difficult. Low refolding yields are often seen due to improper disulfide bond formation, aggregation and instability of folding intermediates and products (Goldberg et al., 1991; De Bernardez et al., 1999; Raman et al., 1996). Correct refolding is dependent on multiple parameters including pH, temperature, composition of the refolding buffer, redox conditions and type of denaturing agents. We therefore, screened each of these parameters to optimize the conditions that would result in a high yield of the correctly folded product with minimum loss of starting material. Introduction of L-arginine, 10:1 ratio of GSH:GSSG and glycerol to the refolding medium increased the amount of refolded product. Another aspect which was given special attention to, was the initial concentration of the unfolded protein. At higher protein concentrations, non-specific aggregation predominates over first-order folding (Jaenicke and Rudolph, 1989). In case of propapain, an initial concentration of up to 50 μg/ml of the denatured protein resulted in the highest proportion of the refolded product.

Processing of propapain to mature papain is triggered by low pH and is stimulated by high temperatures. These two factors probably induce a structural transition required for intramolecular cleavage. Vernet et al. (1991) showed that complete removal of the proregion may possibly be carried out by the active intermediate molecule during processing of the papain precursor expressed in baculovirus/insect cells (Vernet et al., 1991). In our study, during processing of propapain to papain we observed two intermediates in the molecular weight range of 30–38 kDa at 20 min incubation at 50 °C (Fig. 3a) indicating that processing of propapain to papain occurs in a stepwise manner. Similar results have been obtained earlier by Sanderson et al. (2000) for activation of the cysteine proteinase of *Leishmania mexicana* expressed in *E. coli* (Sanderson et al., 2000). We also observed that although a distinct single band without any intermediates corresponding to papain is observed after 60 min of incubation at low pH, the highest activity of the processed papain is observed at 40 min of incubation (Fig. 3b) indicating that longer incubation of the precursor at high temperature and low pH probably leads to depletion of its catalytic activity.

Table 2
Specific activities of recombinant folded papain and commercial papain.

Protein	Concentration	Specific activity (units/mg)[a]
Recombinant papain	1 mg/ml propapain[b]	125.94
Commercial papain	2 mg/ml	221.73

[a] One enzyme unit is defined as the amount of soluble protease required to release 1 μg of soluble azopeptides/min.

[b] For recombinant papain the concentration of mature papain could not be estimated, hence the concentration of propapain used for activation to mature papain is indicated.

4. Conclusion

In conclusion, the protocol given in this study describes the production of active papain using *E. coli* as expression host. It is simple, resulting in high yield of the active enzyme whose specific activity is closely similar to its native counterpart.

5. Experimental

5.1. RNA isolation and cDNA synthesis

Total RNA was extracted from the young leaves of *C. papaya* plant using the RNAqueous-4PCR Kit (Ambion, USA) in accordance with the manufacturer's instructions. The yield and purity of the RNA was quantified spectrophotometrically. First-strand cDNA was synthesized by reverse transcription using 200 U of RevertAid

M-MuLV reverse transcriptase (Fermentas, USA) with 5 μg of total RNA, 1 μg oligo(dT)18 primer, 1 mM dNTPs and 20 U RNAse inhibitor in a total volume of 20 μl.

5.2. Construction of expression plasmid

Prepropapain-encoding DNA was amplified by PCR using *Taq* DNA Polymerase (Promega, USA) from the cDNA pool with primers derived from the published cDNA sequence of prepropapain (Cohen et al., 1986). The amplified product was ligated to the pTZ57R/T vector included in the T/A cloning kit (Fermentas, USA) and transformed into *E. coli* XL1Blue competent cells. Recombinant clones carrying the prepropapain gene were screened by α-complementation of lacZ gene and plasmid DNA was extracted with the QIAprep Spin Miniprep kit (Qiagen, USA). Appropriate insertion of prepropapain gene was verified by PCR and sequenced in both directions with forward and reverse M13 primers using the MegaBACE™1000 sequencing system (Amersham Biosciences, USA).

A fragment of the original cDNA encoding the prodomain and the mature domain of papain was PCR-amplified with KOD HiFi DNA polymerase (Novagen, USA) from the prepropapain clone so as to delete the first 26 codons (the signal peptide) using the following primers: Forward – 5′-GACGACGACAAG**ATG**GATTTTTCTA TTGTGGGT-3′ and Reverse – 5′-GAGGAGAAGCCCGGTCAGTTTTTAA CAGGATAGAATG-3′. The 5′ end of the primers incorporated the ligation-independent cloning (LIC) sequences (underlined). The amplified product was ligated into the expression vector pET-30 Ek/LIC (Novagen, USA) after being treated with LIC-qualified T4 DNA polymerase. This created a fusion with an N-terminal His-tag including 43 amino acid residues (MHHHHHHSSGLVPRGSGM KETAAAKFERQHMDSPDLGTDDDDK) and generated a plasmid named pET30-propap which was then transformed into *E. coli* strain DH5α, and the positive clones were confirmed by colony PCR with gene-specific/vector-specific primers and DNA sequencing.

Simultaneously, the propapain gene was also cloned into pET28a$^+$ expression vector (Novagen, USA). For this, the propapain fragment was PCR-amplified using KOD HiFi DNA polymerase from the prepropapain clone with the primers: Forward – 5′-GGTCGGGATCC**ATG**GATTTTTCTATTGTGGG-3′ and Reverse – 5′-GTGGTGCTCGAGTCAGTTTTTAACAGG-3′. To facilitate cloning in pET28a$^+$, sense and antisense primers were designed to include BamH1 and Xho1 sites (underlined sequences). The vector and the insert were digested with BamH1 and Xho1, ligated and the ligation products transformed into *E. coli* strain DH5α. The clone created was named pET28-propap. An N-terminal tag of 34 amino acid residues including six histidines was added to the gene product (MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGS) and the correct insertion was confirmed by DNA sequencing, restriction digestion and colony PCR with gene-specific/vector-specific primers.

5.3. Expression of fusion protein

E. coli expression strains BL21(DE3) and BL21(DE3)pLysS were transformed with both pET30-propap and pET28-propap recombinant plasmids, respectively, in order to express the propapain protein. The successfully transformed *E. coli* were picked up from a single colony and grown overnight at 37 °C in Luria Bertani (LB) medium (0.5% yeast extract, 1% Bactotryptone, 1% NaCl), supplemented with 34 μg ml^{-1} kanamycin for BL21(DE3) strain or kanamycin and 25 μg ml^{-1} chloramphenicol for BL21(DE3)pLysS strain. The culture mixture was then inoculated to fresh LB medium (1:50 dilution) and grown at 37 °C under continuous shaking, until the absorbance at 600 nm reached 0.6–0.8. To optimize the culture conditions, propapain expression was induced by adding 0.1–1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) to the transformed *E. coli* cultures and incubated at 20 or 37 °C for a period of 1, 2, 3, 4, 5, 6, or 20 h, respectively. The degree of expression was evaluated by SDS–PAGE. After optimization of small-scale expression, large-scale expression of propapain was carried out as follows. Bacteria containing pET30-propap or pET28-propap were grown to mid-log phase at 37 °C, IPTG (1 mM for pET30-propap and 0.5 mM for pET28-propap) was added, and growth continued for 5 h at 37 °C. After induction, the bacterial cells were harvested by centrifugation at 6000g for 10 min. This pellet was stored at either −20 °C or at −80 °C till further use.

5.4. Isolation and solubilization of inclusion bodies

The bacterial pellets obtained from 1 L cultures were resuspended in 50 ml of 50 mM Tris/HCl, pH 8.0, containing 150 mM NaCl, 5 mM ethylene diamine tetraacetic acid (EDTA), 2.5 mM dithiothreitol (DTT) and 1 mM phenylmethanesulfonylfluoride (PMSF) and pelleted again by centrifugation at 5000g for 20 min. This pellet was resuspended in 25 ml 50 mM Tris/HCl, pH 8.0, containing 200 mM NaCl, 2.5 mM EDTA, 2.5 mM DTT, 1% Triton X-100, 5% Sucrose, 1 mM Benzamidine–HCl, 1 μg/ml Pepstatin and 1 mM PMSF and subjected to sonication (10 × 30 s). After centrifugation at 25,000g for 45 min, the inclusion body pellet was washed once with 50 mM Tris/HCl, pH 8.0, containing 200 mM NaCl, 2.5 mM EDTA, 2.5 mM DTT, 1 M Urea, 1 mM Benzamidine–HCl, 1 μg/ml Pepstatin and 1 mM PMSF, twice with distilled water and resuspended in 50 mM Tris/HCl, pH 8.0, containing 300 mM NaCl, 8 M Urea, 10 mM Imidazole, and 1 mM PMSF (5 ml/g of inclusion body pellet) with or without 5 mM βME. The inclusion bodies were allowed to solubilize overnight at room temperature or 4 °C with constant stirring on a magnetic stirrer.

5.5. Purification and refolding of recombinant propapain

The solubilized inclusion body solution was centrifuged at 20,000g for 45 min to remove the unsolubilized material. The supernatant was bound to Ni–NTA (Novagen, USA) matrix pre-equilibrated with 50 mM Tris/HCl, pH 8.0, containing 300 mM NaCl, 7 M Urea and 10 mM Imidazole and the protein was purified under denaturing conditions by metal chelation chromatography according to the His-Bind kits manual (Novagen). Proteins were eluted with five column volumes of 50 mM Tris/HCl, pH 8.0, containing 300 mM NaCl, 7 M Urea and 250 mM Imidazole.

The Ni–NTA fractions were re-purified by size exclusion chromatography on a Sephacryl S-200 (GE Healthcare, USA) column equilibrated with 50 mM Tris/HCl, pH 8.0, containing 300 mM NaCl, 5 M Urea. The protein was eluted with the same buffer. The fractions were checked for appropriate protein size and purity by SDS–PAGE and the peak fractions were pooled and protein concentration estimated by Bradford dye binding assay (Bradford, 1976) using BSA as standard.

The proteins from the major peak eluted from the gel filtration column was reduced by adding DTT to a final concentration of 10 mM and incubated at 37 °C for 1 h. This reduced solution was filtered through a 0.45 μm filter to remove any sediment. Refolding of recombinant propapain was then assessed using a series of buffers differing in pH, concentration of redox couple (reduced glutathione (GSH) to oxidized glutathione (GSSG)), aggregation suppressors (L-arginine, NaCl, Triton X-100 and Urea) and cosolvents (glycerol and sucrose) so as to get maximum amounts of refolded protein. Finally, a rapid dilution method was adapted – the denatured protein was diluted 100 times by mixing with the optimized refolding buffer (50 mM Tris/HCl, pH 8.6, 500 mM NaCl, 5 mM EDTA, 1 mM GSH, 0.1 mM GSSG, 500 mM arginine and

10–15% glycerol) maintaining the protein concentration at 10–50 μg/ml and stirring overnight at 4 °C. Precipitated proteins were removed by centrifugation at 10,000g for 20 min at 4 °C and soluble renatured proteins were concentrated at 4 °C on a 200 ml Amicon stirred cell ultrafiltration unit using an YM-10 membrane, dialyzed against 50 mM Tris–HCl, pH 8.0 containing 150 mM NaCl and 2 mM EDTA. Further purification of the refolded protein was carried out by applying the protein to a size exclusion chromatography column (Sephacryl S-100, GE Healthcare, USA) equilibrated with 50 mM Tris–HCl, pH 8.0 containing 150 mM NaCl and 2 mM EDTA. The protein elution was monitored at 280 nm and the fractions were checked for proteolytic activity by gelatin gel SDS–PAGE as described in a subsequent section. The fractions showing proteolytic activity were pooled, concentrated and stored at −20 °C for further use. This propapain remains stable for months till activated to mature papain for further use. Protein concentration of the refolded propapain was determined by Bradford dye binding assay (Bradford, 1976) and from the absorbance at 280 nm, using the extinction coefficient ($A_{280\ nm}$ = 1.9 ml mg^{-1} cm^{-2}) calculated combining the vector-specific and the propapain amino acid sequence together (Gill and von Hippel, 1989).

5.6. Western blot analysis

Expression of the correct gene product was checked by Western blotting initially using anti-His antibody as both the clones contain His-tag. Subsequently, expression of propapain was confirmed by cross-reacting the protein with anti-papain antibody.

Primarily after separation on 12% SDS–PAGE gel, the proteins were transferred to a nitrocellulose membrane (Amicon Hybond) for 1.5 h at 4 °C at 300 mA in 25 mM Tris, 192 mM glycine and 20% (v/v) methanol. After electrotransfer, membranes were stained with Ponceau S solution for protein detection. For immunodetection of propapain, the membranes were blocked for 1 h at room temperature with 3% BSA in PBS containing 0.1% Tween-20 (PBST) and 5 mM sodium azide, incubated overnight at 4 °C with the same solution containing rabbit anti-His antibody (1:8000 dilution). The membranes were then washed thoroughly with PBST to remove excess primary antibody and then incubated for 1 h with goat anti-rabbit IgG conjugated with Horseradish Peroxidase. Unbound secondary antibody was removed by vigorous washing with PBST. Western blotting with anti-papain antibody was performed in the same manner as above except that the primary antibody used was polyclonal goat anti-papapain antibody (1:2000 dilution, Abcam, UK) and the secondary antibody was rabbit anti-goat IgG (1:10,000 dilution Genei, India) conjugated with Horseradish Peroxidase. Both the blots were detected using chemiluminiscent substrates.

5.7. Activation and processing of pro-papain to active papain

Complete processing of propapain to papain as a function of time and reducing agents was optimized by incubating the refolded propapain in 100 mM Na-acetate, pH 5.0, 2 mM EDTA at 20, 37 or 50 °C either in absence or presence of 20 mM DTT or 20 mM L-cysteine. At designated time points, 10 μl aliquots of the sample were quenched by adding 1 μl of 10 μM 1-[L-*N*-(*trans*-epoxysuccinyl)leucyl]amino-4-guanidinobutane (E-64). To each of these inhibited aliquots, 0.5 vol of SDS–PAGE sample buffer without 2-mercaptoethanol (βME) was added and the sample was boiled for 10 min. Before loading on to the gel, 0.5 vol of sample buffer with 5 μM βME was added to each sample. The reaction was monitored by SDS–PAGE on a 15% resolving gel by Coomassie blue staining and Western blotting using anti-papain antibody as described previously. Activation was also monitored by gelatin SDS–PAGE and by assay using azocasein substrate.

For gelatin SDS–PAGE, 12% gels containing 0.1% gelatin were prepared by mixing gelatin with the resolving gel solution. The activated enzyme was diluted with nonreducing SDS–PAGE sample buffer, incubated for 20 min at 37 °C, and separated by SDS–PAGE at 4 °C. After electrophoresis, the gels were washed with 2.5% Triton X-100 in 20 mM Tris/HCl, pH 8.0 for 30 min. After this, the gel was washed extensively with distilled water to remove the detergent completely. The gel was then incubated in activation buffer (100 mM Na-acetate, pH 5.5, 1 mM EDTA, 2 mM DTT) at 37 °C for 30 min. Following this, the gel was incubated overnight at 37 °C in the activation buffer without DTT to allow refolding and protein digestion. The gel was then stained with Coomassie brilliant blue R-250 and proteinase activity detected by appearance of clear bands on a dark blue background.

5.8. Azocasein assay

The azocasein assay was used to test for proteolytic activity of recombinant papain. Hydrolysis of azocasein by recombinant papain was determined by absorbance increase at 366 nm against time in 100 mM Na-acetate, pH 6.5, containing 2 mM EDTA. Assays were initiated by addition of 25 μl activated recombinant papain to 975 μl of reaction mixture containing 20 μl of 2% azocasein, 2 mM EDTA and 20 mM cysteine. After incubating the solution at 37 °C for designated time intervals ranging from 0 to 30 min, the reaction was stopped by addition of 500 μl of ice-cold trichloroacetic acid. Released azo-peptides were detected by measuring absorbance at 366 nm in a Thermo Nicolat Evolution 100 spectrophotometer using the specific absorption coefficient ($A^{1\%}_{366}$ = 40) for azocasein solution (Chen et al., 2003). One enzyme unit is defined as the amount of protease required to release 1 μg of soluble azopeptides/min. Specific activity of a commercially available native papain (Merck, USA) was measured in the similar manner as mentioned above for comparison with recombinant papain except that a pre-incubation step of 5 min with 20 mM cysteine was included to activate the native protein.

Acknowledgements

We thank Nirmal Datta and Bikram Nath for their help during cloning and expression experiments. The work was partially supported by Council for Scientific and Industrial Research (No. 21/0653/06/EMR-II), Department of Science and Technology (No. SR/WOS-A/LS-161/2004) and Department of Biotechnology (No. BT/PRO0139/R&D/15/11/96), Government of India.

References

Baker, E.N., Drenth, J., 1987. In: Jurnak, F.A., McPherson, A. (Eds.), The Thiol Proteases: Structure and Mechanism in Biological Macromolecules and Assemblies, vol. 3. Wiley, New York, pp. 313–368.

Bradford, M.M., 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein–dye-binding. Anal. Biochem. 72, 248.

Bromme, D., Nallaseth, F.Z., Turk, B., 2004. Production and activation of recombinant papain-like cysteine proteases. Methods, 199–206.

Chen, C.-Y., Luo, S.-C., Kuo, C.-F., Lin, Y.-S., Wu, J.-J., Lin, M.T., Liu, C.-C., Jeng, W.-Y., Chuang, W.-J., 2003. Maturation processing and characterization of Streptopain. J. Biol. Chem. 278, 17336.

Cohen, L.W., Coghlan, V.M., Dihel, L.C., 1986. Cloning and sequencing of papain-encoding cDNA. Gene 48, 219.

Cohen, L.W., Fluharty, C., Dihel, L.C., 1990. Synthesis of papain in *Escherichia coli*. Gene 88, 263.

De Bernardez, C.E., Schwarz, E., Rudolph, R., 1999. Inhibition of aggregation side reactions during *in vitro* protein folding. Meth. Enzymol. 309, 217.

Drenth, J., Jansonius, J.N., Koekoek, R., Swen, H.M., Wolthers, B.G., 1968. Structure of papain. Nature 218, 929.

Gill, S.C., von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182, 319.

Goldberg, M.E., Rudolph, R., Jaenicke, R., 1991. A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme. Biochemistry 30, 2790.

Grzonka, Z., Kasprzykowski, F., Wiczk, W., 2007. Industrial Enzymes: Structure, Function and Applications. In: Polaina, J., MacCabe, A.P. (Eds.). Springer, The Netherlands, pp. 181–195.

Guha Thakurta, P., Biswas, S., Chakrabarti, C., Sundd, M., Jagannadham, M.V., Dattagupta, J.K., 2004. Structural basis of the unusual stability and substrate specificity of ervatamin C, a plant cysteine protease from *Ervatamia coronaria*. Biochemistry 43, 1532.

Jaenicke, R., Rudolph, R., 1989. Folding proteins. In: Creighton, T.E. (Ed.), Protein Structure, A Practical Approach. IRL Press, Oxford, pp. 191–224.

Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A., Drenth, J., 1984. Structure of papain refined at 1.65 Å resolution. J. Mol. Biol. 179, 233.

Liu, J-W., Hadler, K.S., Schenk, G., Ollis, D., 2007. Using directed evolution to improve the solubility of the C-terminal domain of *Escherichia coli* aminopeptidase P. Implications for metal binding and protein stability. FEBS J. 274, 4742.

Mosavi, L.K., Peng, Z., 2003. Structure-based substitutions for increased solubility of a designed protein. Protein Eng. 16, 739.

Polgár, L., 1989. Mechanisms of Protease Action. CRC Press, Boca Raton.

Raman, B., Ramkrishna, T., Rao, C.M., 1996. Refolding of denatured and denatured/reduced lysozyme at high concentrations. J. Biol. Chem. 271, 17067.

Ramjee, M.K., Petithory, J.R., McElver, J., Weber, S.C., Kirsch, J.F., 1996. A yeast expression/secretion system for the recombinant plant thiol endoprotease propapain. Protein Eng. 9, 1055.

Rawlings, N.D., Barrett, A.J., 1994. Meth. Enzymol. 244, 461.

Revell, D.F., Cummings, N.J., Baker, K.C., Collins, M.E., Taylor, M.A.J., Sumner, I.G., Pickersgill, R.W., Connerton, I.F., Goodenough, P.W., 1993. Nucleotide sequence and expression in *Eschericia coli* of cDNAs encoding papaya proteinase omega from *Carica papaya*. Gene 127, 221.

Sanderson, J.S., Pollock, K.G.J., Hilley, J.D., Meldal, M., Hilaire, P.S., Juliano, M.A., Juliano, L., Mottram, J.C., Coombs, G.H., 2000. Expression and characterization of a recombinant cysteine proteinase of *Leishmania mexicana*. J. Biochem. 347, 383.

Taylor, M.A.J., Pratt, K.A., Revell, D.F., Baker, K.C., Sumner, I.G., Goodenough, P.W., 1992. Active papain renatured and processed from insoluble recombinant propapain expressed in *Escherichia coli*. Protein Eng. 5, 455.

Taylor, M.A.J., Al-sheikh, M., Revell, D.F., Sumner, I.G., Connerton, I.F., 1999. cDNA cloning and expression of *Carica papaya* prochymopapain isoforms in *Escherichia coli*. Plant Sci. 145, 41.

Vernet, T., Tessier, D.C., Laliberté, F., Dignard, D., Thomas, D.Y., 1989. The expression in *Escherichia coli* of a synthetic gene coding for the precursor of papain is prevented by its own putative signal sequence. Gene 77, 229.

Vernet, T., Tessier, D.C., Richardson, C., Laliberté, F., Khouri, H.E., Bell, A.W., Storer, A.C., Thomas, D.Y., 1990. Secretion of functional papain precursor from insect cells. J. Biol. Chem. 265, 16661.

Vernet, T., Khouri, H.E., Laflamme, P., Tessier, D.C., Musil, R., Gour-Salin, B.J., Storer, A.C., Thomas, D.Y.J., 1991. Processing of papain precursor. J. Biol. Chem. 266, 21451.

Vernet, T., Chatellier, J., Tessier, D.C., Thomas, D.Y., 1993. Expression of functional papain precursor in *Saccharomyces cerevisiae*: rapid screening of mutants. Protein Eng. 6, 213.

Woycechowsky, K.J., Raines, R.T., 2000. Native disulfide bond formation in proteins. Curr. Opin. Chem. Biol. 4, 533.