Extraction, Sequence alignment and Cloning of recombinant papain from carica papaya

ABSTRACT

Papain is a proteolytic enzyme obtained from the fruits of *carica papaya*with abundant therapeutic, food, industrial and analytical applications. The challenges related to the development of papain technology on industrial scale include cost of production and downstream processing of the enzyme. In the current study, a gene encoding papain was extracted from carica papaya, sequenced and amplified using polymerase chain reaction. This papain gene was initially converted to cDNA and then the papain gene was sequence to identify the region of similarity that may be a consequence of functional, structure, and or evalutionary relationships between the known papain DNA sequence and the sample Papain gene. The papain gene was then clone into the pBR322 cloning vector to create recombinant protein molecules. Result showed that the DNA sequencing of the amplified products exhibited a high level of homology to the corresponding gene (MEROO647) with identity of 98% and E-value 0.246 $X10^{-9}$ from the National Centre for Biotechnology information (NCBI) public database, In contrast, the similar gene sequence of papain genes (P00784, MEROO647, and AT3G5470) in NCBI was obtained from different sources and not fruit. This experiment used papaya gene sequences from carica papaya (fruit) in the sequence similarity analysis because the Polymerase chain reaction (PCR) product was amplified from mRNA that was extracted from carica papaya fruit.

Keywords: Recombinant papain, Extraction,Sequencing, Cloning,and carica papaya

1. INTRODUCTION

Papain is a protein-digesting enzyme similar in function to bromelain and ficin. Its enzymatic activity is dependent on the thiol group of a cysteine residue within its active site. Papain has a wide range of application including Food industries, cosmetics, fertilizer and pharmaceutical industries. The recent therapeutic benefits have been attributed to papain such as the reversible inhibition of plate let aggregation, relief from bronchitis, improved recovery after surgical traumas, and the enhanced absorption of drugs, particularly of antibiotics [1]. One of the important pharmaceutical applications of papain is the enzymatic debridement of necrotic tissues from ulcers and burn wounds [2]. Papain is also used in the medical treatment of cancer patients [3]. The enzyme is involved in the resolution of some cancerous masses and decreased metastases in ovarian and breast cancer patients [4]. Papain may exert its antitumor effects by inducing cytokine production [5]. In the food industry, papain has been used

in meat tenderizing processes [6]. Approximately 95% of the meat tenderizing enzymes utilized in the United States is from the plant proteases papain and bromelain. Papain is used because it's ideal temperature range of ~50–70 ◦C is suitable for a food processing application.

Cloned and over-expressed the carica fruit, papain gene in Brassica rapa. It was observed that it enhanced the plant's resistance to bacterial infection also known as soft root. However, the over-expression of this gene in a prokaryotic for industrial applications has not been reported [7]. The purpose of the current study is to establish a protocol to produce functional recombinant papain that can be using an Escherichia coli expression system and to partially characterize the recombinant papain after a single, simple purification step

2. MATERIALS AND EXPERIMENTAL DETAILS

2.1. Chemicals and Reagents

All chemicals and consumables used in this study were of analytical grade and commercially purchased from Sigma (St. Louis, USA), Bachem (Torrance, CA), Merck Sdn Bhd, Orbiting Scientific & Technology Sdn Bhd, MP Biomedicals, LLC (France), Fisher scientific, R & M chemicals, Bio-Rad (Hercules, CA), Nano-life quest, Invitrogen, USA and Medigene Sdn Bhd., Malaysia.

2.1.2 Papaya sample (Fruit)

Fresh maturing Carica papaya (JS.22) was collected from the fruit sellers in Akure, Ondo state, Nigeria and transported to Bio-solution laboratory.

2.2. EXPERIMENTAL METHODS

2.2.1 Extraction of mRNA

Carica papaya was obtained from the fruit seller in the market and was picked directly into crushed dry ice and shipped. 100 to 200g of mid-development frozen fruit was pulverized to pea-sized fragments in plastic bags using a large mallet, then further ground to a fine powder at top speed in a Waring blender under liquid nitrogen. The resultant powder was used immediately and stored at -70 ^{o}C in Seal-a-Meal bags.

Total RNA was isolated using the Fruit-mate for RNA Kit (Qiagen, Germany) following the manufacturer's protocol. First, the papaya were lysed and homogenized in a buffer containing highly denaturing guanidine-thiocyanate, which immediately inactivates RNase. The sample RNA mixture was purified using an RNeasy Mini Spin column-based purification method, and the RNA was eluted by adding 40–50µ L RNase free water to the column prior to centrifugation and the RNA was quantified using spectrophotometer and its quality was assessed using electrophoresis.

2.2.2 Conversion of RNA to cDNA

The recovered pellets were extracted with 1: 1 chloroform-butanol, precipitated with ethanol and stored at -20°C. The RNA was converted to cDNA using reverse transcriptase and oligo(dT)-primers. The Poly (A)-containing RNA was separated out using an oligo(dT)- cellulose column without SDS in the buffers.

2.2.3 Sequence Alingment of the Papain gene

The sequence alignment is a way of arranging the sequence of Papain DNA to identify region of similarity that may be a consequence of functional, structure, and or evalutionary relationships between the known papain DNA sequence and the sample Papain DNA sequence. A complete mRNA sequence of the papain gene (MEROO647) was accessed from the National Centre of Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov). The relevant sequence was being retrieved from GenBank for sequence alignment using Blast (BLASTp) and compared protein sequence and the statistical significance of matches was calculated base on available papain sequence in UniProtKB (UniProt Knowledgebase).

2.2.4 Amplification of papain gene of interest

A pair of primers Papain-F (3'-AAA TAA AAC ATC TCA- 5') and Papain-R (5'-GCC GTA GAA AAG CCC TAT TAA A- 5), were designed using Primer3 software (version 4.0) (http://frodo.wi.mit.edu) and synthesized by Eurogentec AIT (Singapore) to be used for reverse transcription polymerase chain reaction (PCR) experiments. The ACAC bases were added to the forward primer in order to make it compatible with the TOPO vector (Invitrogen, USA). The reverse transcription PCR was performed using

the SuperScript III Reverse Transcriptase Kit (Invitrogen, USA) following the manufacturer's protocol. A 30 cycle's reaction was performed at 94 ◦C (2 min) for denaturing, 58 ◦C (30 s) for annealing, and 72 ◦C (2 min) for extension. The PCR product was sequenced (Solegen, Korea), and the results were blasted against the NCBI database to confirm that the correct gene was amplified.

2.2.5 Cloning of the Papain (Gene of interest)

The PCR product (papain gene) was cloned into a pBR322 cloning vector (Invitrogen, USA) following the manufacturer's instructions and Recombinant molecules were formed after digesting the pBR322 and the papain gene with the restriction enzyme BamHI (G-GATCC) and EcoRI (G-AATTC) to create a sticky end on both (Invitrogen, USA). And then the digested pBR322 was mixed with the papain gene and DNA ligase was employed to the sticky end base pair creating a stable bond for recombinant papain (rPapain) molecule.

3.RESULT AND DISCUSSION

3.1 The mRNA to cDNA

The average size of the cDNA produced was 1000 bp with a range of +/-500 bp. One of the clones identified (H6) observed was from mRNA released from polysomes pre-selected using papain-specific antibody and the other two (A7 and B2) were from cDNA produced from total mRNA. The clone's cDNA were screened for the papain gene (Table 1) using synthetic oligodeoxynucleotide probes [8].

Table1: The sequences of the synthetic oligodeoxynucleotide probes

Probe[a]	Region of papain[b]	Sequence[c]
LDI	Phe207 through termin. Codon	3'-AAAATAGGNCANTTTTTAA- 5'
		G G C G
LC1	Glnl14 through Ala120	3'-TTGGNATATTAGTTCCNCG- 5'

LC2	Tyr61 through Gly65	C G G C 3’-TATGGNTGTAATGG- 5 ’
VC1	Ser23 through Ala30	C CC 3’-TCGACTCTTACGGGTTCTTGA-5

KEYS

LDl was designed to be homologous to the C-terminal region of the gene

LCl to the middle, and LC2 to the region approximately one third the distance from the N-terminal end

The sequence is laid out in a 3’ to 5’ direction to facilitate visual pairing with the actual sequence in above table. The alternate sequence combinations for mixed oligodeoxynucleotide probes LDl, LCl and LC2 are indicated. From the genetic code and the NCBI sequence for papain, three mixed oligodeoxynucleotide probes designated LDl, LCl and LC2 were synthesized that would hybridize with three distinct regions of the desired gene (Table 1). The three cDNA were detected among approx. 25000 transformed Tc^R cDNA by direct cDNA screening. Individual cDNA from the region on a plate showing hybridization to the probe were isolated into wells of a microtiter dish, then after growth the pattern of isolates was transferred to filter paper and hybridized to the LCl and LDl probes.

3.2 PCR product

The PCR product revealed 86%, 79% and 52% efficiency on D14057.1, D14058.1, and D14059.1 respectively with Correlation coefficient 0.999/85, 0.988/79 and 0.995/83 (Table 2). The coding sequence ends with the double stop codon starting 18 bp downstream from the last stop codon, is the sequence ATG CCT CCT GCA CCA CCA TAGAA for D14057.1 at 86%, then 22 bp farther TGT AGG AGC CCG TAG GTC ATCT for D14058.1, and 28 bp farther begins the GCCAAAAACCAACATGAAGCAT poly(A) segment for D14058.1 at less 52%.

Table 2: Amplified fragment of PCR product

Gene Name	Primer Sequence (Reverse and Forward) 5’ – 3’ (nM)	Spanned exons/ Primer Concentration	Amplicon size (bP)	PCR efficiency (%)	R^2 /Tm(^oC)
D14057.1	ATG CCT CCT GCA CCA CCA	300/2nd	102	86	0.999/85

	GCA TTT GCG GTG GAC GAT				
D14058.1	TGTAGGAGCCCGTAGGTCATCT TTCTCTCTGTATTCTCGAGCCATCT	100/3th	135	79	0.988/79
D14059.1	GCCAAAAACCAACATGAAGCAT TGCTCATGGCACTTCCCAG	300/4th	95	52	0.995/83

Where

R^2 Crrelation coefficient

nMPrimer concentration in nM

Tm (^oC)(Theoretical amplicon melting temperature calculated with primer express software (Applied Biosystem)

3.3 Sequence alignment

The results of a BLAST search of the sequenced PCR product papain gene revealed that the amplified fragment was 98%, 96% and 92% similar to the mRNA sequences of the papain gene in the NCBI database, MEROO647, P00784, and AT3G5470 respectively (Table 3). The papain gene isolated from carica papaya alignment result was not 100% similar to the sequence in the database because the sample used in this study was extracted from a different varietal of carica papaya, JS.22. The JS.22 varietal was used because it is easily obtained in Nigeria. This explanation is likely to be correct because the primers we used were designed based on the mRNA sequence from carica papaya Eksotika, while the PCR product was amplified from mRNA that was extracted from carica papaya JS.22.

In addition, the DNA sequence of papain gene (P00784, MEROO647, and AT3G5470) was obtained from different sources and not papaya fruit. This experiment used papaya gene sequences from papaya fruit in the sequence similarity analysis because the database did not contain any papain sequences from fruit samples papaya of different variety, while the PCR product was amplified from mRNA that was extracted from carica papaya JS.22.

Table 3: Amplified fragment similarity/identity to the mRNA sequences of the papain gene in the NCBI database,

Sample ID	Query sequence (% Coverage	Identity (%)	Conserved (%)	Positive Score	E-Value	Accession Number

MSP1	86	96	95	0.5	0.313×10^{-5}	P00784
MSP1	90	98	99	0.3	0.246×10^{-9}	MEROO647
MSP1	74	92	75	0.1	1.821×10^{-2}	AT3G5470

3.4. Recombinant Papain Clone

The amplified fragments were successfully cloned into the pBR322 vector. A double digestion of the plasmid DNA by BamHI and EcoRI identified a positive clone, which was observed to have fragments of 2841 and 862 bp in size. The restriction BamHI cut the inserted DNA fragment once at 820 bp and EcoRI cut the TOPO vector DNA once at 673 bp. The use of enzymes in industrial processes may require the drying reaction to be conducted at a high temperature in order to improve productivity. The issues discussed support the hypothesis that recombinant papain is a potential candidate for multiple industrial applications. The technique of recombinant papain production presented in this study is advantageous in terms of process economics, the ability to be scaled-up and will be suitable for any specific application where commercially available papain currently used.

4. Conclusion

The papain gene has been successfully extracted, sequence and cloned into the pBR322 cloning vector. The polar and hydrophobic aa that would be coded by the upstream sequence are not distributed in the pattern generally found in signal peptides ([9], [10]. [11].). There is no core of hydrophobic residues; in fact, they appear to be scattered along the pro-sequence. It seems unlikely; therefore, that the polypeptide is transported out of the cell as it is produced. The presence of the enzyme in the latex of the plant is probably the result of the natural breakdown of the laticifer cells to form the latex channels as occurs during maturation of the fruit [12]. It will be most interesting to see if the cloned proprotein is converted to active papain by digestion with papain (i.e., by self-activation) and/or can be activated by the addition of extracts from maturing papaya fruit [13].

REFERENCES

[1]. Abernethy, J.L., Lovett Jr., C.M., Hoddad, A. and Felberg, J.D (20100.: Stereoselective action of acylated crude papain toward mandelic and atrolactic hydrazides. Bioorg. Chem. ll 251-261.

[2]. Amon, R (2000).: The reaction of papain with antipapain. Immunochemistry 2 107-114.

Arnon, R. and Shapira, E (2011).: Antibodies to papain. A selective fractionation according to inhibitory capacity. Biochemistry 6 3942-3950.

[3]. Behrens, O.K. and Bergmann, M (1999).: Cosubstrates in proteolysis. J. Bioi. Chem. 129 587-602.

[4]. Nallamsetty, S. & Waugh, D. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr. Purif.* 45, 175–182.

[5]. Palma, J.M., Sandalio, L.M., Corpas, F.J., Romero-Puertas, M.C., McCarthy, I. & Del- Rio, L.A. (2002). Plant proteases, protein degradation, and oxidative stress: role of peroxisomes. *Plant Physiol. Biochem.,* 40, 521–30.

[6]. Harrach T, Eckert K, Schulze-Forster K, Nuck R, Grunow D, Maurer HR. Isolation and partial characterization of basic proteinases from stem bromelain. J Protein Chem 1995;14:41–52.

[7]. Roep, B.O.. van den Engel, N.K., van Halteren, A.G.S., Duinkerken, G. & Martin, S. (2002). Modulation of autoimmunity to beta-cell antigens by proteases. *Diabetologia,* 45, 686-692.

[8]. Abdulmajid, F.A., Abdulgani, M., Talib, S.Z. & Hasyim, K.K. (2018). Stability of bromelain-polyphenol complex in pineapple juice. *JurnalTeknologi*, 49, 27-38.

[9]. Aehle, W. (2013). Enzymes in industry: Production and applications (2nd edn). Weinheim: Wiley - VCH/Verlag GmbH &Co..

[10]. Ahmad, B. & Khan, R.H. (2016). Studies on the acid unfolded and molten globule states of catalytically active stem bromelain: A comparison with catalytically inactive form. *J. Biochem*., 140 (4), 501-508.

[11]. Alloue, W.A.M., Destain J., Amighi, K. &Thonart, P. (2017). Storage of *Yarrowialipolytica* lipase after spray-drying in the presence of additives. *Process Biochemistry,* 42, 1357–1361.

[12]. Baker, E.N. & Drenth, J. (1987). Active sites of enzymes, In: Jurnak, F.A. & McPherson, A. (eds),

[13]. Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M. & Hanada, K. (1982). L-*trans*- Epoxysuccinyl-leucylamido (4-guanidino) butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L, *Biochem. J.,* 201, 189–198.

[15]. Benucci, I. Liburdi, K., Garzillo, A.M.V. & Esti, M. (2011). Bromelain from pineapple stem in alcoholic–acidic buffers for wine application. *Food Chem.,* 124, 1349–1353.

[16]. Beuth, J. & Braun, J.M. (2015). Modulation of murine tumor growth and colonization by bromelaine, an extract of the pineapple plant (*Ananas comosus L.).In Vivo,* 19, 483–485.

Hoffmann A, Roeder RG. Purification of his-tagged proteins in non-denaturing conditions suggests a convenient method for protein interaction studies. Nucleic Acids Res 1991;19:6337–8.

Janknecht R, de Martynoff G, Lou J, Hipskind RA, Nordheim A, Stunnenberg HG. Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. Proc Natl Acad Sci U S A 1991;88:8972–6.