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 present research, the responsible for gene encoding papain enzyme was obtained from the carica papaya plant, 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 classify base on functional, structure, and phylogenic 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. This showed that the DNA sequencing of the extracted genes exhibited an acceptable level of similarity to the corresponding gene (MEROO647) with identity of 98% and E-value 0.246 $X10^{-9}$ from the public database NCBI (National Centre for Biotechnology information) In contrast, the similarpapain genes sequence (P00784, MEROO647, and AT3G5470) in NCBI extracted from other sources in this study, the papaya gene sequences was obtained from carica papaya (fruit) and in the sequence similarity analysis because the product was amplified from mRNA that was extracted from carica papaya fruit.

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

1. Introduction

“Papain is like other form of enzymes such as bromelain and ficin that are protein-digesting enzyme. There enzymatic activity is based 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 application of purified proteins in therapeutics has been known for a long period of time. Hormones such as insulin and human growth hormones, and other proteins were extracted and purified from blood and other tissues before the emergence of recombinant DNA technologythe ability to genetically engineer *Escherichia coli* and other organisms to produce large quantities of therapeutically desired molecules, has led to the development of several therapeutic proteins” [1, 3]. “Some of these proteins (e.g. erythropoietin) are greatly successful drugs with annual worldwide sales of several billion dollars”[4].

Genetic and protein engineering are one of the vital tools employed for the industrial production of recombinant proteins. Such proteins have greater applications in enzyme, pharmaceutical and the agricultural industries. Over 200 approved recombinant proteins including albumin, growth hormone and insulin among others, are on the FDA list [6]. The recombinant DNA technology had boosted enzyme industries in several means: (i) the use of fermentation to produce enzymes of animals and plants origin; (ii) the use of industrial organisms to produce enzymes from microorganisms that are hard to propagate; (iii) enhancing productivity of enzymes with the help of manifold copies of genes, competent signal sequences and very strong promoters; (iv) using a safe host to produce functional enzymes from[7]. Using recombinant DNA technology, currently, enzymes can be custom-made to meet the requirements of the users or of the process. Moreover, protein engineering is being employed to generate enzymes of greater quality[6]. In addition, directed evolution serves as a novel method for enhancing enzymes production [8].

2. MATERIALS AND METHODS

2.1. Chemicals and Reagents

All chemicals and reagents used in this study were of analytical grade and commercially purchased.

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. METHODS

2.2.1 Extraction of mRNA

Thecarica papaya fruit, at its mid-development stage, was picked and immediately placed in dry ice that had been crushed to a powder. The fruit was then shipped under these condition, 100 to 200 grams of the frozen fruit were pulverized into fragments approximately the size of peas using a large mallet and plastic bags. These fragments were then further ground into a fine powder at high speed in a Waring blender

that had been cooled with liquid nitrogen. The resulting powder was utilized immediately and stores at a temperature of 70^{o}Cin seal-a-meal bags.

Total RNA was isolated using the Fruit-mate (RNA Kit) as described by the manufacturers. To begin with, the carica papaya was subjecting to a lysis and homogenization process using a buffer that contained guanine-thiocyanate, a substance that denatures and inactivates RNase instantly. The sample RNA mixture was purified using Mini Spin column-based purification method, and the sample was eluted by adding 40–50µ L RNase free water to the column prior to centrifugation and 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 gene (RNA) was converted to cDNA using reverse transcriptase and primers. The Poly (A)-containing RNA was separated out using cellulose column without SDS in the buffers.

2.2.3 Papain gene Sequencing

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 relationships between the known papain DNA sequence and the sample Papain DNA sequence. mRNA sequence of the papain gene (MEROO647) 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 Basic Local Alignment Search Tool (BLAST) and compared protein sequence and the statistical significance of matches was calculated base on available papain sequence (UniProt Knowledgebase).

2.2.4 Amplification of papain gene

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 (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. 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 and the results were blasted against the NCBI database to confirm that, if 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 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 sticky end on both. 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.

3.RESULT AND DISCUSSION

3.1 The mRNA to cDNA

The size of the cDNA obtained was 1000 bp with at a range of +/-500 bp. One of the identified (H6) clone observed was from mRNA released from polysomes pre-selected using papain-specific antibody produced from total mRNA. Table 1 shows the clone's cDNA screened for the papain gene using synthetic single –stranded sequence of DNA[8].

Table 1: The sequences of the synthetic single –stranded sequence of DNA

Probe	Region of papain complementary Sequence	Sequence
LDI	Phe207 through termin. Codon	3'-AAAATAGGNCANTTTTTAA- 5' G G C G
LC1	Glnl14 through Ala120	3'-TTGGNATATTAGTTCCNCG- 5' C G G C
LC2	Tyr61 through Gly65	3'-TATGGNTGTAATGG- 5 ' C CC
VC1	Ser23 through Ala30	3'-TCGACTCTTACGGGTTCTTGA-5

The alternate sequence combinations for mixed single-stranded DNALDl, LCl and LC2 are indicated table 1. From the genetic code and the NCBI sequence for papain, three mixed probe

DNAdesignated LDl, LCl and LC2 were synthesized that would hybridize with three distinct regions of the desired gene (Table 1). Each cDNA from the region on a plate showing hybridization to the probe DNA 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 LDlsingle-stranded DNA.

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 started at 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 segment of D14058.1

Table 2: 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 GCA TTT GCG GTG GAC GAT	300/2nd	102	86	0.999/85
D14058.1	TGTAGGAGCCCGTAGGTCATCT TTCTCTCTGTATTCTCGAGCCATCT	100/3th	135	79	0.988/79
D14059.1	GCCAAAAACCAACATGAAGCAT TGCTCATGGCACTTCCCAG	300/4th	95	52	0.995/83

3.3 Sequence alignment

The finding from a Blast search of the sequenced PCR product for papain gene showed that the amplified fragment matched the following 98%, 96% and 92% similar to the mRNA sequences of the papain gene in the NCBI database, MEROO647, P00784, and AT3G5470 respectively (Table 1). The papain gene isolated from the papaya alignment result was not 100% similar to the sequence in the public 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 available 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.

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 clone amplified fragment were successfully inserted into the pBR322 vector. Through a double digestion of the plasmid DNA using BamHI and EcoRI, a positive clone was identified, showing fragment of 820 bp, while EcoRI cut the TOPO vector DNA once at 673 bp. Enzymatic processes in industries may necessitate high-temperature drying reaction to enhance productivity. The issues discussed lend support to the idea that recombinant papain hold promise for various industrial uses. The method outlined for producing recombinant papain in this research offers advantages in term of process efficiency, scalability, and adaptability to specific application where commercially available papain is currently utilized.

4. Conclusion

Papain gene has been successfully isolated, sequenced, and cloned into the pBR322 cloning vector. The polar and hydrophobic amino acids encoded by the upstream sequence do not follow the typical distribution pattern found in signal peptides. Instead, hydrophobic residues appear to be scattered throughout the pro-sequence, suggesting that the polypeptide is unlikely to be transported out of the cell

during production. The presence of papain in the plant's latex is likely due to the natural breakdown of laticifer cells during fruit maturation. It will be intriguing to observe whether the cloned proprotein can be converted to active papain through self-activation or by adding extracts from maturing papaya fruit.

ACKNOWLEDGEMENTS

We would like to thank the Federal University of kashere, Gombe state through the Tertiary Education Trust-Fund (TETFUND) for providing the study fellowship 2021/2022. And my appreciation goes to the Bio-soluton laboratory and Central laboratory Federal university Wukari for support and guidance throughout the bench work

AUTHORS' CONTRIBUTION

This work was carried out in collaboration among all authors. Authors MSM,JOE and MMADDesigned the study. Authors MSM and VKP performed the Bioinformatics' analysis. Authors MSM, APD, MU wrote the protocol and wrote the first draft of the manuscript. Authors MSM,UUM and HRA manage the literature search of the work. All authors read and approved the final manuscript.

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 AcadSci U S A 1991;88:8972–6.