Journal of Agricultural Science and Food Technology
Vol. 4 (2), pp. 32-40, March, 2018
ISSN: 2465-7522
Full Length Research Paper
http://pearlresearchjournals.org/journals/jasft/index.html

In Silico evaluation of bromelain from stem and fruit of pineapple (*Ananas comosus*)

Oluwasegun Victor Omotoyinbo[1], Ini-Obong Emmanuel Ekpenyong[1], Toluwase Hezekiah Fatoki [2], Morakinyo David Sanni[2*]

Accepted 19 March, 2018

[1]Department of Biological Sciences, Wesley University,Ondo State,Nigeria.
[2]Department of Biochemistry, Federal University of Technology, Akure,Ondo State, Nigeria.

ABSTRACT

Pineapple (*Ananas comosus*) is the leading edible member of the family *Bromeliaceae*, and has been used as a medicinal plant in several native cultures. The medicinal qualities of pineapple are attributed to the enzyme Bromelain which is abundant in the stem and fruit of the pineapple plant and can also be isolated in small amount from pineapple waste such as core, leaves, peel etc. In this study, *in silico* analysis of bromelain from pineapple stem and fruit was done using standard bioinformatics tools such as *Blastp, HMMER, Clustal O, OMA Browser, EMBOSS, Swiss model,* and *Hex*. The result showed specific protein from *Gossypium raimondii*, *Amborella trichopoda*, *Populus trichocarpa*, *Corchorus olitorius*, *Bromelia fastuosa* as homologs of bromelain present in pineapple. The theoretical isoelectric point of stem bromelain was 8.436, while that of fruit bromelain was 4.726. The phylogenetic analysis shows that fruit bromelain evolved slowly compared to stem bromelain. Heuristic search for distant homologs implicated Asian wild rice. Eudicots and rodent were among orthologs obtained. The modeled structure obtained revealed that both stem and fruit bromelain possessed secondary structures made up of mainly helixes, coils and beta strands.

Keywords: Bromelain, In silico, physicochemical properties, phylogeny, structure.

*Corresponding author. Email: moraksanni@yahoo.co.uk. Mobile: +234 806 925 6694

INTRODUCTION

Pineapple is the common name of *Ananas comosus* and it is the leading edible member of the family *Bromeliaceae*, grown in several tropical and subtropical countries including Philippines, Thailand, Indonesia, Malaysia, Kenya, India, and China. It has been used as a medicinal plant in several native cultures (Mondal et al*.,* 2011) and these medicinal qualities of pineapple are attributed to the enzyme Bromelain (EC 3.4.22.32). Bromelain has been chemically known since 1875 and is used as a phytomedicinal compound obtained from the crude extract of pineapple and contains among other compounds, different closely related proteinases, exhibiting various fibrinolytic, antiedematous, antithrombotic, and anti-inflammatory activities in vitro and in vivo (Pavan et al*.,* 2012).

Bromelain although present in all parts of the pineapple, has a high concentration in the pineapple stem which unlike the pineapple fruit which is normally used as food, the stem is a waste byproduct and thus inexpensive (Pavan et al*.,* 2012). A wide range of therapeutic benefits have been claimed for Bromelain, such as reversible inhibition of platelet aggregation, sinusitis, surgical traumas, thrombophlebitis, pyelonephritis angina pectoris, bronchitis, and enhanced absorption of drugs, particularly of antibiotics (Maurer, 2001; Pavan et al*.,* 2012). Bromelain acts on fibrinogen giving products that are similar, at least in effect, to those formed by plasmin (Pavan et al*.,* 2012). Experiment in mice showed that antacids such as sodium bicarbonate preserve the proteolytic activity of bromelain in the gastrointestinal tract (Hale, 2004). These and many more studies have been carried out indicating that Bromelain has useful phytomedical application. Existing evidence indicates that Bromelain can be a promising candidate for the development of future oral enzyme therapies for oncology patients as it can be absorbed in human intestines without degradation and without losing its biological activity (Chobotova et al., 2010). With the knowledge of sequencing, structural biology and bioinformatics revolutionising biomolecular science and millions of sequences in UniProt and tens of thousands of three dimensional (3D) structures in the PDB (Berman et al., 2000; Alderson et al., 2012), the design of enzymes have progressed from site-directed re-engineering of natural proteins towards *de novo* design and assembly of active enzymes. This has been possible over the decades with a plethora of software, much of it freely available, that allows sophisticated computations to be carried out on large datasets. However, good understanding of systems biology and metabolism depends on a detailed knowledge of enzyme function (Bairoch, 2000; Li et al., 2010). In this study, *in silico* analyses of stem bromelain (SB) and fruit bromelain (FB) from pineapple were done using standard bioinformatics tools.

MATERIALS AND METHODS

Sequence composition and homology analysis

The stem and fruit bromelain sequences were obtained from *UniprotKB/Swiss-prot* (http://www.uniprot.org) with ID: P14518 (BROM2_ANACO) for stem bromelain; 212 amino acids residues. The fruit bromelain with ID: O23791 (BROM1_ANACO); 351 amino acids residues. The homologs having minimum 50% identity were extracted by *Blastp* of both sequences from the Uniprot database, and they were stored in FASTA format for further analysis (Sanni et al., 2017).

Physicochemical analysis

The physicochemical analysis was done according to the method of Sanni et al. (2017). The sequence physicochemical statistics such as theoretical molecular weight, isoelectric point of P14518 and O23791 were obtained using *EMBOSS Pepstat* at default setting (www.ebi.ac.uk/Tools/seqstats/emboss_pepstats), while hydropathy plots for both sequences were obtained using *Protscale* from Expasy database using Kyte and Doolittle parameters at default setting (web.expasy.org/protscale).

Phylogenetic analysis

The phylogenetic analysis was done according to the method of Sanni et al., (2017). Multiple sequence alignment was carried out on the enzyme and its homologs by using *ClustalO* (www.ebi.ac.uk/Tools/msa/clustalo/), at default setting. The phylogenetic tree was obtained in cladogram and the tree data was then used to visualize the real phylogeny at *phylo.io* (http://phylo.io).

Local alignment analysis

The local alignment was carried out using *EMBOSS matcher* (www.ebi.ac.uk/Tools/psa/emboss_matcher). The FASTA format of the two sequences were run using P14518 and O23791 has query respectively against the sequence of a distant homolog A0A0E0QV61, obtained from interactive database searching of each of P14518 and O23791 sequences using *HMMER* (www.ebi.ac.uk/Tools/hmmer) (Sanni et al., 2017).

Evolution distance by orthologous and paralogous analysis

The sequence of P14518 and O23791 was independently retrieved from Uniprot, and entered to Blast of NCBI (Altschul, 1993). The query was first run against SWISS-PROT database to obtain the orthologs, while it was run against the non-redundant database to obtain the paralogs of the sequences. The BLAST Tree view for Fast Minimum Evolution was obtained.

Structural analysis

The tertiary structure of P14518 and O23791 were modeled separately using *Swiss Model* (https://swissmodel.expasy.org). The template with the highest quality and identity was selected for model building (Sanni et al., 2017).

RESULTS AND DISCUSSION

Sequence composition and homology analysis

Bromelain is a group of sulfhydryl proteolytic enzymes (Marshall and Golden, 2012) and encompasses variety of cysteine proteases (Tochi et al. 2008) when extracted from the stem and fruit of pineapple plant (Neta et al., 2012).The result from the Uniprot database and physicochemical statistics from pepstat showed that the two Bromelain sequences: P14518 (Stem Bromelain) and O23791 (Fruit Bromelain) had differences in amount of amino acid residues and molecular weight. Stem Bromelain has 212 residues with a molecular weight of 22.83 kDa which is consistent with findings of (Ritonja et al., 1989). Fruit bromelain having 351 residues and a molecular weight of 39.05 kDa, this molecular is higher than 31kDa reported by Yamada et al. (1976) and a range of 24.5-32 kDa (Grzonka et al., 2007; Lopes et al., 2009; Gautam et al., 2010). This difference may be because of post-translation modification within the cell after protein synthesis.

The FASTA format of the sequences of stem Bromelain with ID: P14518 (BROM2_ANACO) and fruit Bromelain

Table 1. Theoretical physiochemical properties of P14518 and O23791.

Theoretical physicochemical property	P14518	O23791
Amino acid Residues	212	351
Molecular weight	22830.93	39054.78
Average Residue Weight	107.693	111.267
Net Charge	4.5	-7.0
Isoelectric Point	8.4361	4.7261
A_{280} Molar Extinction Coefficients (reduced/cystine bridges)	48360/48735	75290/75790
A_{280} Extinction Coefficients 1mg/ml (reduced/cystine bridges)	2.118/2.135	1.928/1.941
Improbability of expression in inclusion bodies	0.923	0.778

with ID: O23791 (BROM1_ANACO) retrieved from the Uniprot database. The homologs of P14518 (BROM2_ANACO) and O23791 (BROM1_ANACO) having minimum of 50% identity include the following proteins; *Ananas comosus* (A0A199VSS3, A0A199V231, A0A199UL32, A0A199UKV8, A0A199W8N4, F1KD58, P80884, A0A199W9F2), *Gossypium raimondii* (A0A0D2UND8), *Amborella trichopoda* (U5D9T6), *Populus trichocarpa* (B9H2I7), Corchorus olitorius (A0A1R3KH25), *Bromelia fastuosa* (Q6Q2T4), *Ananas macrodontes* (P83443), *Erythranthe guttata* (A0A022QAF9).

Fastuosain is a novel cysteine proteinase from *Bromelia fastuosa* with an estimated molecular mass of 25 kDa. It has been purified from the “gravatá” fruit from south-eastern Brazil. The fruit is edible, and its juice is popularly used in the treatment of bronchitis and asthma. It has been functionally characterized, cloned, and sequenced by Cabral et al. (2006). Fastuosain, in a protocol of intraperitoneal injections along with the tumor cell challenge in C57Bl/6 mice, largely inhibited lung colonization, and this effect matched that of bromelain. Amborella is a monotypic genus of understory shrubs or small trees endemic to the main island, Grande Terre, of New Caledonia (Jérémie, 1982). The genus is the only member of the family *Amborellaceae* and the order *Amborellales* and contains a single species, *Amborella trichopoda* (Große-Veldmann et al*.,* 2011).

Physicochemical analysis

Theoretical physiochemical properties of P14518 and O23791 are shown in (Table 1), while the hydropathy plots of the bromelain were shown in (Figures 1a and b). Furthermore, the result from *pepstat* showed the isoelectric point of SB to be 8.4361 while that of FB was 4.7261 respectively. This result is close to experimental data on isoelectric point of SB and FB with 9.55 and 4.6 respectively as reported earlier (Murachi et al*.,* 1964 and Yamada et al*.,* 1976). In proteins the isoelectric point (pI) is defined as the pH at which a protein has no net charge. When the pH > pI, a protein has a net negative charge and when the pH < pI, a protein has a net positive charge. The pI varies for different proteins. The negative net charge on Fruit Bromelain shows an excess of negatively charged amino acids in its sequence, while a positive net charge on SB shows an excess of positively charged residues in its sequence as shown from the *pepstats* results. The combination from the effect of the isoelectric point and net charge suggests that SB is a basic protein while FB is an acidic protein, and this aligned with the findings of Omotoyinbo and Sanni, (2017).

It has been shown that protein family can indicate a region of parameter space in which a protein is likely to crystallize (Hennessy et al*.,* 2000) and that properties such as amino acid frequency, sequence length, grand average of hydrophobicity (GRAVY) (Kyte and Doolittle, 1982) and isoelectric point (pI) as well as other properties can determine a protein’s propensity to crystallize (Smialowski et al*.,* 2006; Chen et al*.,* 2007).

These properties help in the preparation of buffer of appropriate pH during protein extraction, isolation and characterization. Investigations have also indicated that the isoelectric point could be used to determine the pH at which a protein with an acidic isoelectric point is likely to crystallize (Kantardjieff and Rupp, 2004; Charles et al*.,* 2006). The isoelectric point determines a protein’s minimum solubility level due to protein-protein interactions being favoured over protein-water interactions (Gilliland, 1988; Luft et al*.,* 2011).

The result from the hydropathy plots (Kyte and Doolittle, 1982), shows that SB contains a higher proportion of hydrophobic/non-polar amino acids compared to FB. This indicates that SB might be a membrane protein or a protein that interacts well with membrane lipid bilayer compared to FB. One may also suggest that Stem Bromelain may be glycosylated, this is correct according to findings by Khan et al*.,* (2003). This results also shows that FB by having more hydrophilic residues may be more water soluble than SB. It can also be inferred that SB would also have more hydrophobic residues in the interior of its 3D structure than FB.

Phylogenetics analysis

The results for the multiple sequence alignment (MSA) revealed different of amino acid residues that were conserved during evolution of bromelain. In lane 1 of (Figure 2), it shows conserved D, E, D and W residues. In

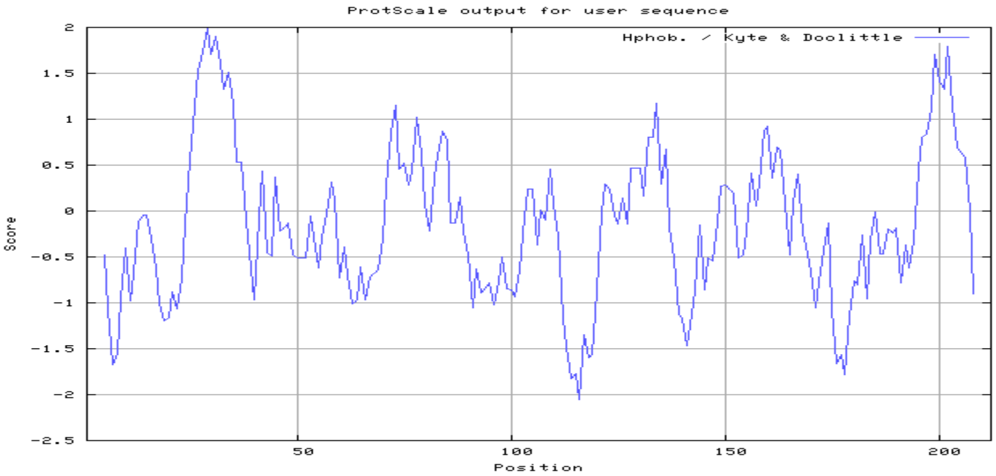


Figure 1a. Hydropathy plot for P14518. It shows that there is presence of many hydrophobic amino acid residues depicted by the peaks of the plot.

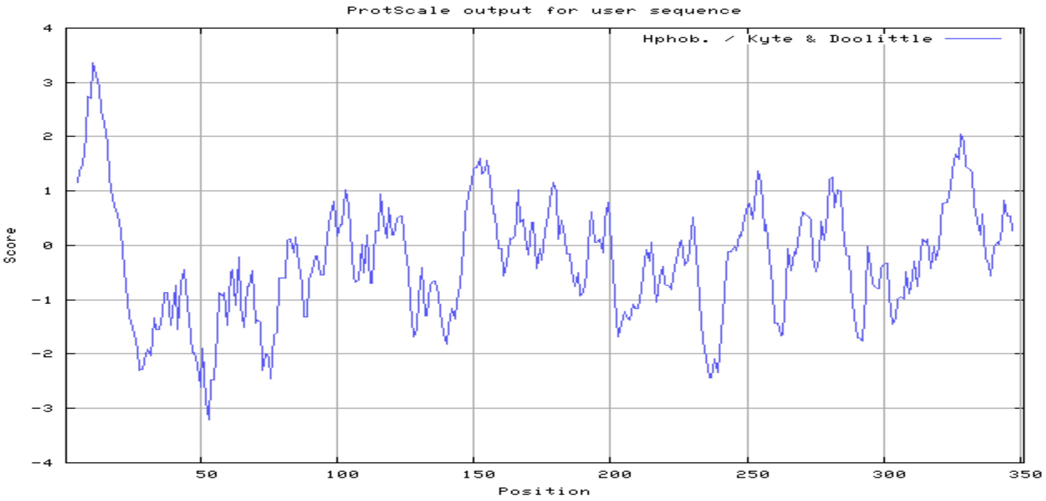
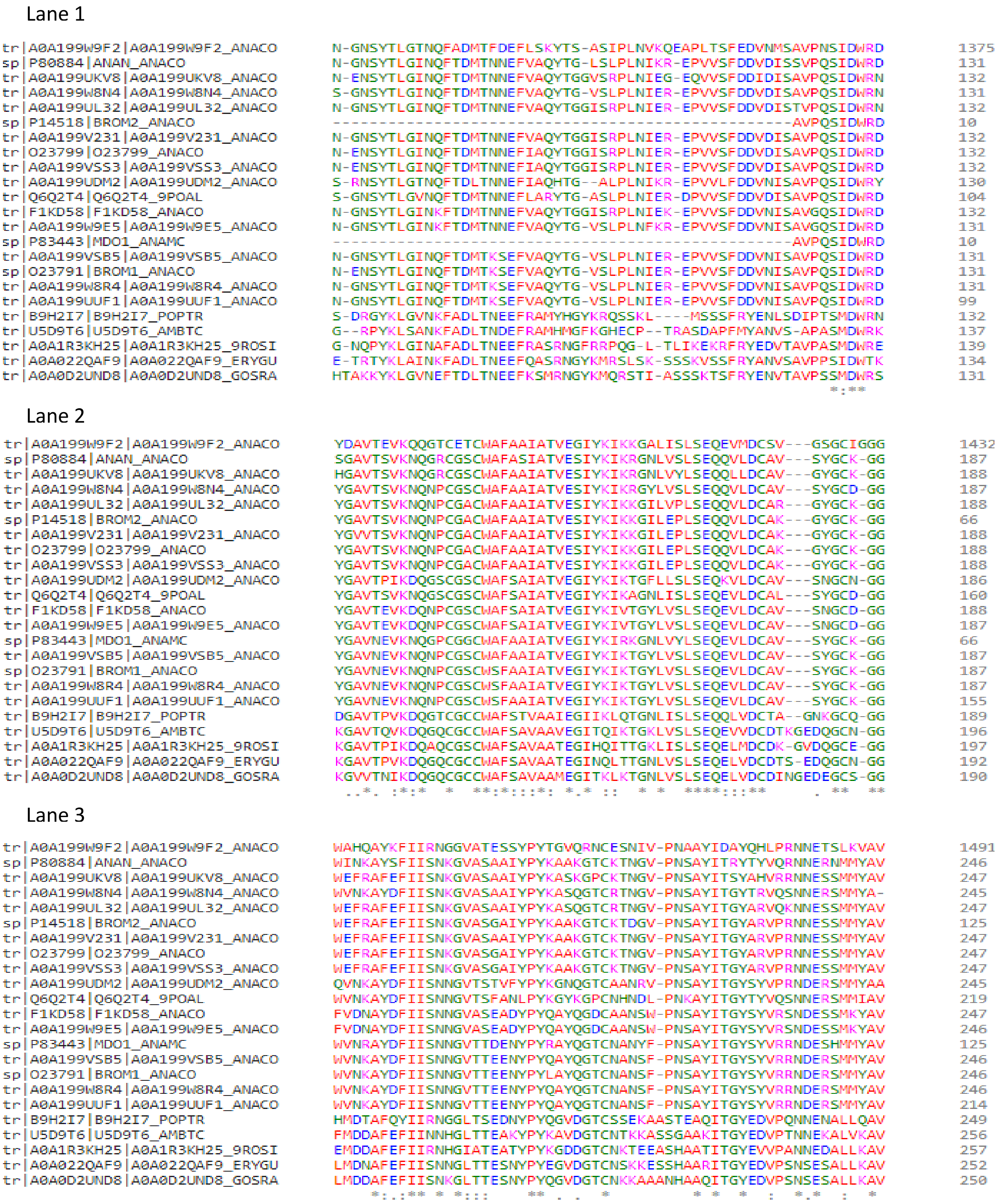


Figure 1b. Hydropathy plot for 023791. It shows that the hydrophobic regions are lesser than those found in P14518 above.

lane 2, it shows conserved C, V, Q, W, E, I, E, Q, D, G, and G residues. In lane 3 it shows conserved F, I, I, I, P, N, E, and A residues. The result also shows some semi-conserved residues in all the three lanes. Amino acids that are conserved are those most critical to the function of the protein. Thus, looking for evolutionarily conserved patches of amino acids in a 3D protein structure is a good way to locate functional sites (Pettit et al*.,* 2007).The phylogeny of bromelain from P14518 and O23791 and their homologs is shown in (Figure 3). A phylogenetic tree, also known as a phylogeny, is a diagram that depicts the lines of evolutionary descent of different species, organisms, or genes from a common ancestor (Baum, 2008). The result from phylogenetic analysis shows that FB has evolved slowly compared to SB. This may be due to modifications, mutations or generally consequences of evolution.

Local alignment analysis

A distant homolog with Uniprot/Swissprot ID: A0A0E0QV61, an uncharacterized protein from *Oryza rufipogon* (Brown beard rice) (Asian wild rice), which has 1,893 amino acid residues, was obtained by HMMER, a hidden markov model algorithm which has great predictive power. The extent of similarity and identity of this distant homolog with P14518 and O23791 are shown in (Table 2).The result from Uniprot shows that the both Bromelain sequences have a taxonomical lineage as follows: cellular organisms › Eukaryota › Viridiplantae ›



Lane 1

```
tr|A0A199W9F2|A0A199W9F2_ANACO      N-GNSYTLGTNQFADMTFDEFLSKYTS-ASIPLNVKQEAPLTSFEDVNMSAVPNSIDWRD      1375
sp|P80884|ANAN_ANACO                N-GNSYTLGINQFTDMTNNEFVAQYTG-LSLPLNIKR-EPVVSFDDVDISSVPQSIDWRD      131
tr|A0A199UKV8|A0A199UKV8_ANACO      N-ENSYTLGINQFTDMTNNEFVAQYTGGVSRPLNIEG-EQVVSFDDIDISAVPQSIDWRN      132
tr|A0A199W8N4|A0A199W8N4_ANACO      S-GNSYTLGINQFTDMTNNEFVAQYTG-VSLPLNIER-EPVVSFDDVDISAVPQSIDWRN      131
tr|A0A199UL32|A0A199UL32_ANACO      N-GNSYTLGINQFTDMTNNEFVAQYTGGISRPLNIER-EPVVSFDDVDISTVPQSIDWRN      132
sp|P14518|BROM2_ANACO               ---------------------------------------------------AVPQSIDWRD      10
tr|A0A199V231|A0A199V231_ANACO      N-GNSYTLGINQFTDMTNNEFVAQYTGGISRPLNIER-EPVVSFDDVDISAVPQSIDWRD      132
tr|O23799|O23799_ANACO              N-ENSYTLGINQFTDMTNNEFIAQYTGGISRPLNIER-EPVVSFDDVDISAVPQSIDWRD      132
tr|A0A199VSS3|A0A199VSS3_ANACO      N-ENSYTLGINQFTDMTNNEFIAQYTGGISRPLNIER-EPVVSFDDVDISAVPQSIDWRD      132
tr|A0A199UDM2|A0A199UDM2_ANACO      S-RNSYTLGTNQFTDLTNNEFIAQHTG--ALPLNIKR-EPVVLFDDVNISAVPQSIDWRY      130
tr|Q6Q2T4|Q6Q2T4_9POAL              S-GNSYTLGVNQFTDMTNNEFLARYTG-ASLPLNIER-DPVVSFDDVDISAVPQSIDWRD      104
tr|F1KD58|F1KD58_ANACO              N-GNSYTLGINKFTDMTNNEFVAQYTGGISRPLNIEK-EPVVSFDDVNISAVGQSIDWRD      132
tr|A0A199W9E5|A0A199W9E5_ANACO      N-GNSYTLGINKFTDMTNNEFVAQYTG-VSLPLNFKR-EPVVSFDDVNISAVGQSIDWRD      131
sp|P83443|MDO1_ANAMC                ---------------------------------------------------AVPQSIDWRD      10
tr|A0A199VSB5|A0A199VSB5_ANACO      N-GNSYTLGINQFTDMTKSEFVAQYTG-VSLPLNIER-EPVVSFDDVNISAVPQSIDWRD      131
sp|O23791|BROM1_ANACO               N-ENSYTLGINQFTDMTKSEFVAQYTG-VSLPLNIER-EPVVSFDDVNISAVPQSIDWRD      131
tr|A0A199W8R4|A0A199W8R4_ANACO      N-GNSYTLGINQFTDMTKSEFVAQYTG-VSLPLNIER-EPVVSFDDVNISAVPQSIDWRD      131
tr|A0A199UUF1|A0A199UUF1_ANACO      N-GNSYTLGINQFTDMTKSEFVAQYTG-VSLPLNIER-EPVVSFDDVNISAVPQSIDWRD      99
tr|B9H2I7|B9H2I7_POPTR              S-DRGYKLGVNKFADLTNEEFRAMYHGYKRQSSKL----MSSSFRYENLSDIPTSMDWRN      132
tr|U5D9T6|U5D9T6_AMBTC              G--RPYKLSANKFADLTNDEFRAMHMGFKGHECP--TRASDAPFMYANVS-APASMDWRK      137
tr|A0A1R3KH25|A0A1R3KH25_9ROSI      G-NQPYKLGINAFADLTNEEFRASRNGFRRPQG-L-TLIKEKRFRYEDVTAVPASMDWRE      139
tr|A0A022QAF9|A0A022QAF9_ERYGU      E-TRTYKLAINKFADLTNEEFQASRNGYKMRSLSK-SSSKVSSFRYANVSAVPPSIDWTK      134
tr|A0A0D2UND8|A0A0D2UND8_GOSRA      HTAKKYKLGVNEFTDLTNEEFKSMRNGYKMQRSTI-ASSSKTSFRYENVTAVPSSMDWRS      131
                                                                                          *:**
```

Lane 2

```
tr|A0A199W9F2|A0A199W9F2_ANACO      YDAVTEVKQQGTCETCWAFAAIATVEGIYKIKKGALISLSEQEVMDCSV---GSGCIGGG      1432
sp|P80884|ANAN_ANACO                SGAVTSVKNQGRCGSCWAFASIATVESIYKIKRGNLVSLSEQQVLDCAV---SYGCK-GG      187
tr|A0A199UKV8|A0A199UKV8_ANACO      HGAVTSVKNQGRCGSCWAFAAIATVESIYKIKRGNLVYLSEQQLLDCAV---GYGCK-GG      188
tr|A0A199W8N4|A0A199W8N4_ANACO      YGAVTSVKNQNPCGSCWAFAAIATVESIYKIKRGYLVSLSEQQVLDCAV---SYGCD-GG      187
tr|A0A199UL32|A0A199UL32_ANACO      YGAVTSVKNQNPCGACWAFAAIATVESIYKIKKGILVPLSEQQVLDCAR---GYGCK-GG      188
sp|P14518|BROM2_ANACO               YGAVTSVKNQNPCGACWAFAAIATVESIYKIKKGILEPLSEQQVLDCAK---GYGCK-GG      66
tr|A0A199V231|A0A199V231_ANACO      YGVVTSVKNQNPCGACWAFAAIATVESIYKIKKGILEPLSEQQVLDCAK---GYGCK-GG      188
tr|O23799|O23799_ANACO              YGAVTSVKNQNPCGACWAFAAIATVESIYKIKKGILEPLSEQQVLDCAK---GYGCK-GG      188
tr|A0A199VSS3|A0A199VSS3_ANACO      YGAVTSVKNQNPCGACWAFAAIATVESIYKIKKGILEPLSEQQVLDCAK---GYGCK-GG      188
tr|A0A199UDM2|A0A199UDM2_ANACO      YGAVTPIKDQGSCGSCWAFSAIATVEGIYKIKTGFLLSLSEQKVLDCAV---SNGCN-GG      186
tr|Q6Q2T4|Q6Q2T4_9POAL              YGAVTSVKNQGSCGSCWAFSAIATVEGIYKIKAGNLISLSEQEVLDCAL---SYGCD-GG      160
tr|F1KD58|F1KD58_ANACO              YGAVTEVKDQNPCGSCWAFSAIATVEGIYKIVTGYLVSLSEQEVLDCAV---SNGCD-GG      188
tr|A0A199W9E5|A0A199W9E5_ANACO      YGAVTEVKDQNPCGSCWAFSAIATVEGIYKIVTGYLVSLSEQEVLDCAV---SNGCD-GG      187
sp|P83443|MDO1_ANAMC                YGAVNEVKNQGPCGGCWAFAAIATVEGIYKIRKGNLVYLSEQEVLDCAV---SYGCK-GG      66
tr|A0A199VSB5|A0A199VSB5_ANACO      YGAVNEVKNQNPCGSCWAFAAIATVEGIYKIKTGYLVSLSEQEVLDCAV---SYGCK-GG      187
sp|O23791|BROM1_ANACO               YGAVNEVKNQNPCGSCWSFAAIATVEGIYKIKTGYLVSLSEQEVLDCAV---SYGCK-GG      187
tr|A0A199W8R4|A0A199W8R4_ANACO      YGAVNEVKNQNPCGSCWSFAAIATVEGIYKIKTGYLVSLSEQEVLDCAV---SYGCK-GG      187
tr|A0A199UUF1|A0A199UUF1_ANACO      YGAVNEVKNQNPCGSCWSFAAIATVEGIYKIKTGYLVSLSEQEVLDCAV---SYGCK-GG      155
tr|B9H2I7|B9H2I7_POPTR              DGAVTPVKDQGTCGCCWAFSTVAAIEGIIKLQTGNLISLSEQQLVDCTA--GNKGCQ-GG      189
tr|U5D9T6|U5D9T6_AMBTC              KGAVTQVKDQGQCGCCWAFSAVAAVEGITQIKTGKLVSLSEQEVVDCDTKGEDQGCN-GG      196
tr|A0A1R3KH25|A0A1R3KH25_9ROSI      KGAVTPIKDQAQCGSCWAFSAVAATEGIHQITTGKLISLSEQELMDCDK-GVDQGCE-GG      197
tr|A0A022QAF9|A0A022QAF9_ERYGU      KGAVTPVKDQGQCGCCWAFSAVAATEGINQLTTGNLVSLSEQELVDCDTS-EDQGCN-GG      192
tr|A0A0D2UND8|A0A0D2UND8_GOSRA      KGVVTNIKDQGQCGCCWAFSAVAAMEGITKLKTGNLVSLSEQELVDCDINGEDEGCS-GG      190
                                     ..*. :*:*  *  **:*:::*: *.* ::  * *  ****:::**     . **  **
```

Lane 3

```
tr|A0A199W9F2|A0A199W9F2_ANACO      WAHQAYKFIIRNGGVATESSYPYTGVQRNCESNIV-PNAAYIDAYQHLPRNNETSLKVAV      1491
sp|P80884|ANAN_ANACO                WINKAYSFIISNKGVASAAIYPYKAAKGTCKTNGV-PNSAYITRYTYVQRNNERNMMYAV      246
tr|A0A199UKV8|A0A199UKV8_ANACO      WEFRAFEFIISNKGVASAAIYPYKASKGPCKTNGV-PNSAYITSYAHVRRNNESSMMYAV      247
tr|A0A199W8N4|A0A199W8N4_ANACO      WVNKAYDFIISNKGVASAAIYPYKASQGTCRTNGV-PNSAYITGYTRVQSNNERSMMYA-      245
tr|A0A199UL32|A0A199UL32_ANACO      WEFRAFEFIISNKGVASAAIYPYKASQGTCRTNGV-PNSAYITGYARVQKNNESSMMYAV      247
sp|P14518|BROM2_ANACO               WEFRAFEFIISNKGVASGAIYPYKAAKGTCKTDGV-PNSAYITGYARVPRNNESSMMYAV      125
tr|A0A199V231|A0A199V231_ANACO      WEFRAFEFIISNKGVASAAIYPYKAAKGTCKTNGV-PNSAYITGYARVPRNNESSMMYAV      247
tr|O23799|O23799_ANACO              WEFRAFEFIISNKGVASGAIYPYKAAKGTCKTNGV-PNSAYITGYARVPRNNESSMMYAV      247
tr|A0A199VSS3|A0A199VSS3_ANACO      WEFRAFEFIISNKGVASGAIYPYKAAKGTCKTNGV-PNSAYITGYARVPRNNESSMMYAV      247
tr|A0A199UDM2|A0A199UDM2_ANACO      QVNKAYDFIISNNGVTSTVFYPYKGNQGTCAANRV-PNSAYITGYSYVPRNDERSMMYAA      245
tr|Q6Q2T4|Q6Q2T4_9POAL              WVNKAYDFIISNNGVTSFANLPYKGYKGPCNHNDL-PNKAYITGYTYVQSNNERSMMIAV      219
tr|F1KD58|F1KD58_ANACO              FVDNAYDFIISNNGVASEADYPYQAYQGDCAANSW-PNSAYITGYSYVRSNDESSMKYAV      247
tr|A0A199W9E5|A0A199W9E5_ANACO      FVDNAYDFIISNNGVASEADYPYQAYQGDCAANSW-PNSAYITGYSYVRSNDESSMKYAV      246
sp|P83443|MDO1_ANAMC                WVNRAYDFIISNNGVTTDENYPYRAYQGTCNANYF-PNSAYITGYSYVRRNDESHMMYAV      125
tr|A0A199VSB5|A0A199VSB5_ANACO      WVNKAYDFIISNNGVTTEENYPYQAYQGTCNANSF-PNSAYITGYSYVRRNDERSMMYAV      246
sp|O23791|BROM1_ANACO               WVNKAYDFIISNNGVTTEENYPYLAYQGTCNANSF-PNSAYITGYSYVRRNDERSMMYAV      246
tr|A0A199W8R4|A0A199W8R4_ANACO      WVNKAYDFIISNNGVTTEENYPYQAYQGTCNANSF-PNSAYITGYSYVRRNDERSMMYAV      246
tr|A0A199UUF1|A0A199UUF1_ANACO      WVNKAYDFIISNNGVTTEENYPYQAYQGTCNANSF-PNSAYITGYSYVRRNDERSMMYAV      214
tr|B9H2I7|B9H2I7_POPTR              HMDTAFQYIIRNGGLTSEDNYPYQGVDGTCSSEKAASTEAQITGYEDVPQNNENALLQAV      249
tr|U5D9T6|U5D9T6_AMBTC              FMDDAFEFIINNHGLTTEAKYPYKAVDGTCNTKKASSGAAKITGYEDVPTNNEKALVKAV      256
tr|A0A1R3KH25|A0A1R3KH25_9ROSI      EMDDAFEFIIRNHGIATEATYPYKGDDGTCNKTEEASHAATITGYEVVPANNEDALLKAV      257
tr|A0A022QAF9|A0A022QAF9_ERYGU      LMDNAFEFIISNNGLTTESNYPYEGVDGTCNSKKESSHAARITGYEDVPSNSESALLKAV      252
tr|A0A0D2UND8|A0A0D2UND8_GOSRA      LMDDAFEFIISNKGLTTESNYPYQGVDGTCNKKAAANHAAQITGYEDVPSNSESALLKAV      250
                                        *:.:** * *:::    ** . .  *         * *  *  :  *.*  :  *
```

Figure 2. Segments of multiple sequence alignment of P14518, O23791 and their homologs. (Red= small, hydrophobic, aromatic, not Y. Blue= acidic. Magenta= basic. Green= hydroxyl, amine, amide, basic. Gray= others) “*”= identical, “:”= conserved substitutions (same colour group), “.”= semi-conserved substitution (similar shapes).

Streptophyta › Streptophytina › Embryophyta › Tracheophyta › Euphyllophyta › Spermatophyta › Magnoliophyta › Mesangiospermae › Liliopsida › Petrosaviidae › commelinids › Poales › Bromeliaceae › Bromelioideae › Ananas. The taxonomical lineage of Oryza rufipogon (Brownbeard rice) (Asian wild rice) is as follows: Eukaryota › Viridiplantae › Streptophyta › Embryophyta › Tracheophyta › Spermatophyta ›

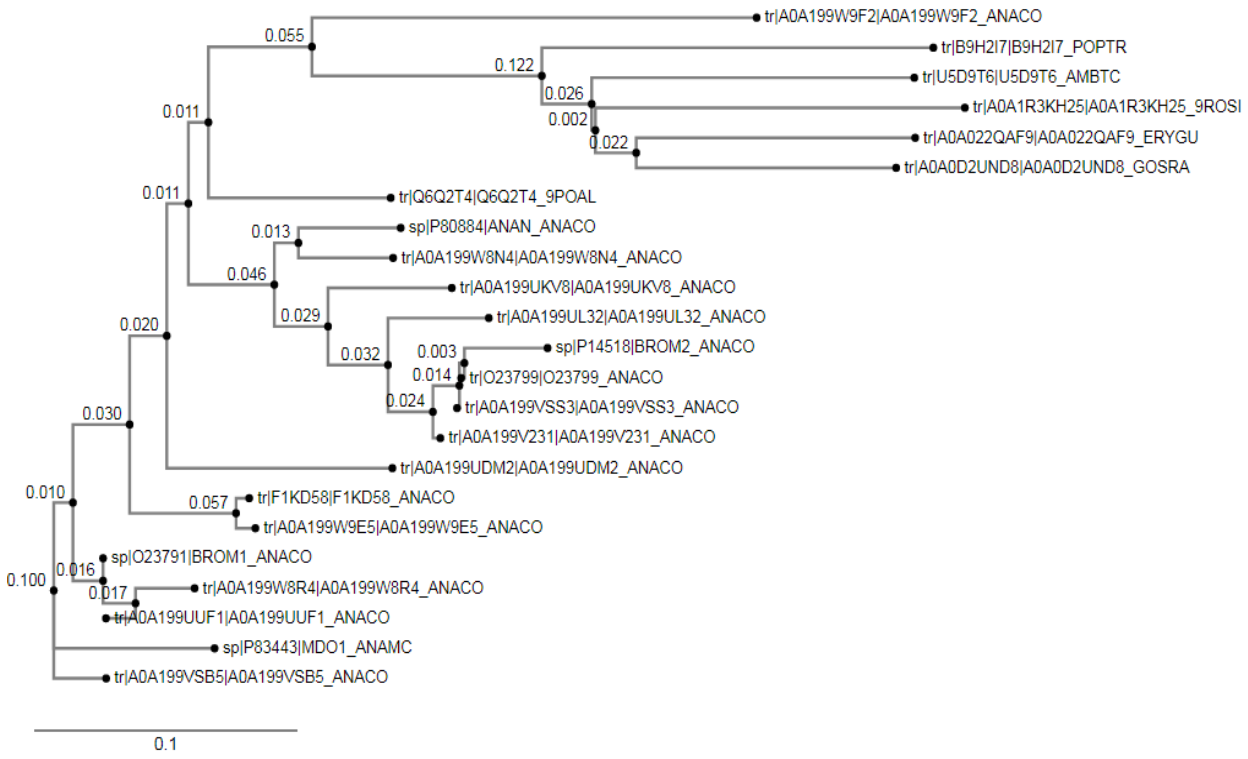


Figure 3. Phylogeny visualized at Phylo.io. It shows the degree of divergence of P14518, O23791, and their homologs from ancestors.

Table 2. Local alignment of P14518 and O23791 against A0A0E0QV61, using EMBOSS Matcher.

Parameters	P14518	O23791
Matrix	EBLOSUM62	EBLOSUM62
Gap penalty	14	14
Extend penalty	4	4
Length	208	318
Identity	102/208 (49.0%)	140/318 (44.0%)
Similarity	137/208 (65.9%)	192/318 (60.4%)
Gaps	6/208 (2.9%)	12/318 (3.8%)
Score	524	665

Magnoliophyta › Liliopsida › Poales › Poaceae › BOP clade › Oryzoideae › Oryzeae › Oryzinae › Oryza.

From the above it can be inferred that both are related and have same ancestors only diverging at the “Poales” stage/level. Rice is an angiosperm (flowering plant), with respect to the characteristics that it shares with all other angiosperms it thus can serve as a model for all flowering plants. Within the angiosperms, rice belongs to the large group known as the monocotyledons; rice thus provides a model for investigating characteristics that appear in monocots. Within the monocots, rice falls within the group known informally as the commelinidclade. Other members of this large group are palms (Arecaceae), pineapples (Bromeliaceae), gingers, and bananas (Zingiberales), as well as grasses, sedges, and rushes. Within the commelinids, rice falls in the order Poales, a member of the Poaceae, the grass family, while pineapple is a member of the Bromeliaceae, the Bromeliads, they are the only family within the Poales that has septal nectaries and inferior ovaries (Kellogg, 2009).

Evolution distance by orthologous and paralogous analysis

The tree view for fast minimum evolution obtained for the orthologous and paralogous analysis of O23791 is shown in (Figures 4a and b). The results of evolutionary distance show orthologs of fruit bromelain, which consists of enzymes present in other species that arose from the same ancestor that show the same function as itself.Results show that it is genetically similar to SB, and

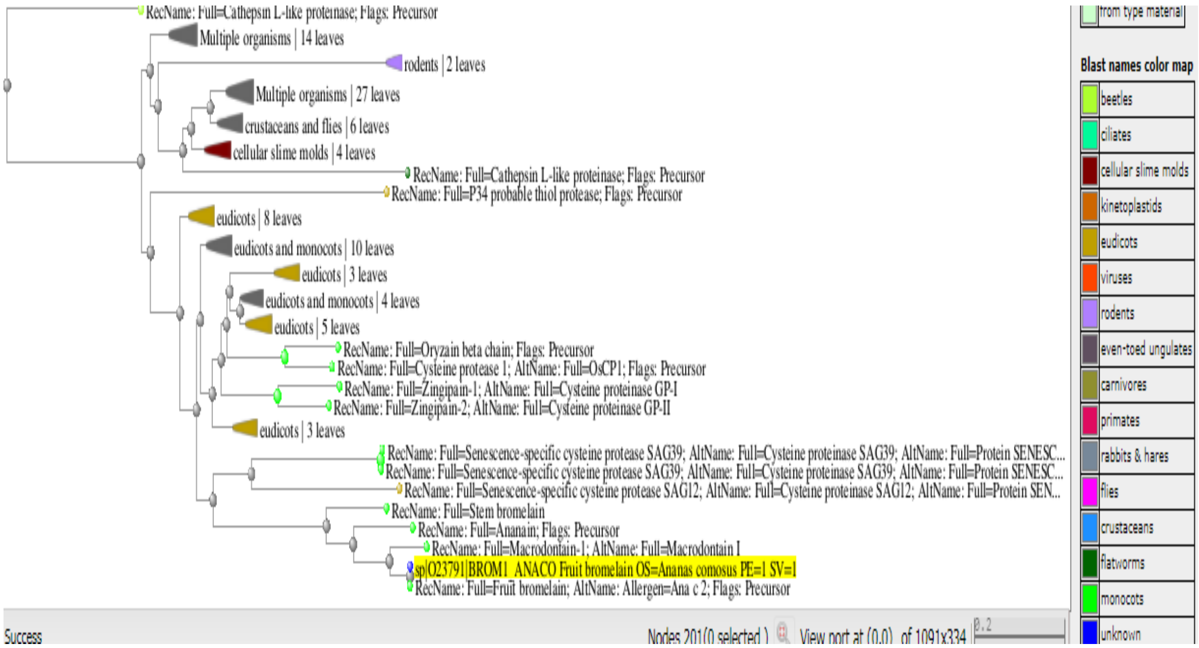


Figure 4a. BLAST Tree View of O23791 Orthologs which are similar enzymes in other species that arose by speciation from a common ancestor.

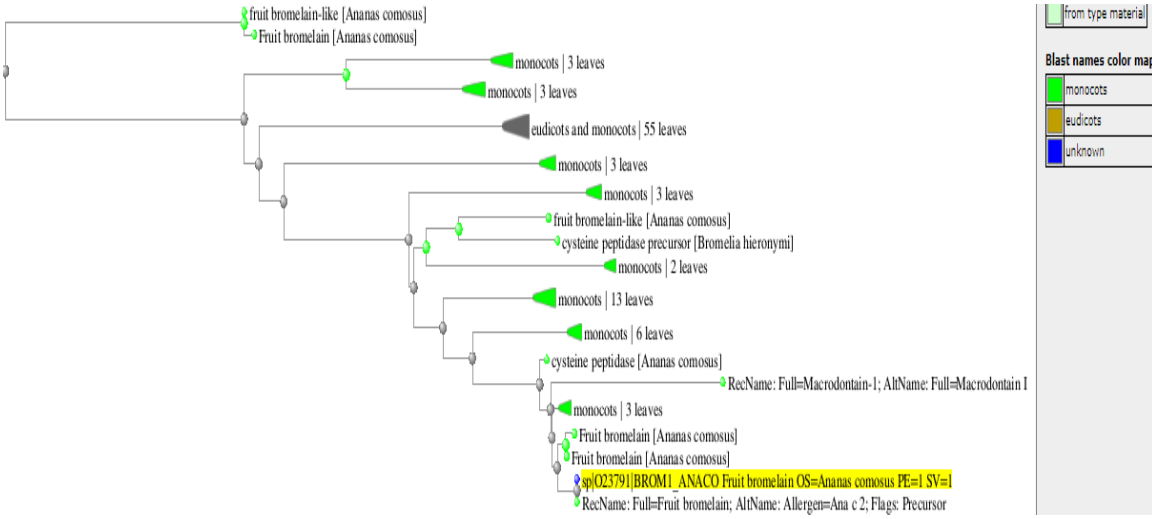


Figure 4b. BLAST Tree View of O23791 Paralogs which are similar enzymes that resulted from gene duplication events in evolution.

macrodontain-1, an enzyme from *Pseudananas macrodontes* (false pineapple). It seems clear that the *Bromeliaceae* endopeptidases are more closely related to each other than to other members of the papain family, suggesting relatively recent divergence (López et al., 2000). Orthologs, or orthologous genes, are genes in different species that originated by vertical descent from a single gene of the last common ancestor (Fitch, 1970). The result also shows paralogs of FB, which are homologous sequences that were produced by gene-duplication events (paralogs) (Koonin, 2005).

Structural analysis

The structure of P14518 used in this study was obtained from homology modelling of the 3D structure of Chymopapain (PDB ID: 1yal.1.A) retrieved from the Protein Data Bank (PDB) which had 52.15% identity with P14518, 218 amino acid residues, and was resolved at 1.7 Armstrong using X-ray Crystallography.The structure of O23791 used in this study was obtained from homology modelling of the 3D structure of Procaricain (PDB ID: 1pci.1.A) retrieved from the Protein Data Bank

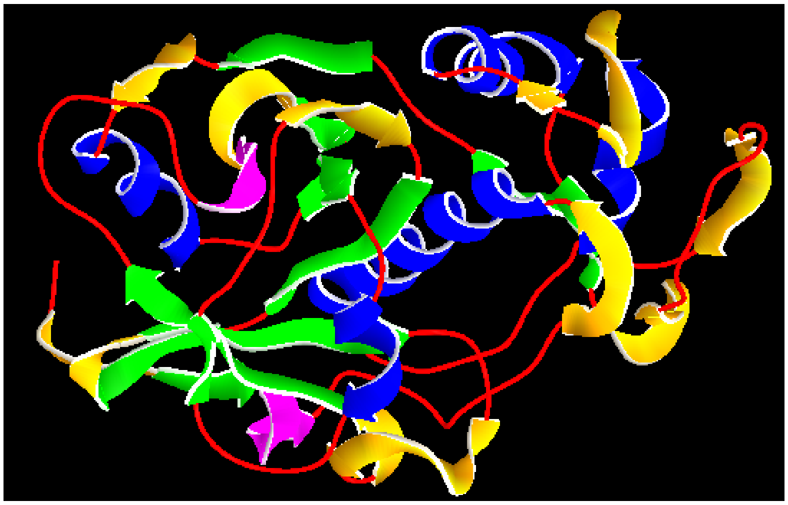


Figure 5a. Structure of stem bromelain (P14518) modeled from Swiss Model and viewed by Hex Software.

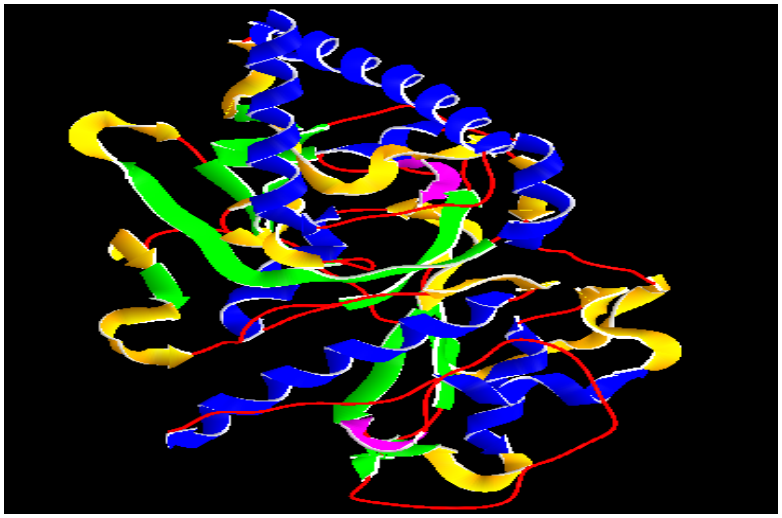


Figure 5b. Structure of fruit bromelain (O23791) modeled from Swiss Model and viewed and viewed by Hex software.

which had 41.21 % identity with O23791, with 325 amino acid residues, and was resolved at 3.2 Armstrong using X-ray Crystallography. The modelling was done using Swiss Model database, and the modelled structure were viewed using Hex Software version 8.0.0, as shown in (Figures 5a and b). The secondary structure of bromelain enzyme was made up mostly of helices (40 %), coils (39 %) and beta-strands (11 %) (Fatahiya et al*.,* 2016). This is also evident in the modelled structures obtained from Swiss Model.

Conclusion

Stem and fruit Bromelain were analysed *in silico* by using different bioinformatics tools and databases described above. It has been shown that Stem Bromelain is a glycoprotein while fruit Bromelain is not. Bromelain and pineapple in general has been shown to have a common ancestry with the Asian wild rice, and with the rice family. The phylogenetic analysis has shown that the amino sequences between the two enzymes are almost similar showing that they are closely related. Structural analysis revealed the fact that the Stem Bromelain is showing more similarity with chymopapain, while Fruit Bromelain shows similarity with Procaricain, molecular visualization of which has been represented.

REFERENCES

Alderson RG, De Ferrari L, Mavridis L, McDonagh JL, Mitchell JBO, Nath N (2012). Enzyme Informatics. Curr Top Med Chem. 12(17): 1911–1923.

Altschul SF (1993). "A protein alignment scoring system sensitive at all evolutionary distances*."* J. Mol. Evol*.* 36: 290-300.

Bairoch A (2000). The ENZYME database in 2000. Nucleic Acids Res. 28(1): 304–305.

Baum D (2008). Reading a Phylogenetic Tree: The Meaning of Monophyletic Groups*.* Nature Edu. 1(1): 190

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank. Nucleic Acids Res. 28(1):235–242.

Cabral H, Leopoldino AM, Tajara EH, Greene LJ, Faca VM, Mateus RP, Ceron CR, Judice, WAD, Juliano L, Bonilla-Rodriguez GO (2006). Preliminary functional characterization, cloning and primary sequence of Fastuosain, a cysteine peptidase isolated from fruits of Bromelia fastuosa. Protein Pept Lett*.* 13: 83–89.

Charles M, Veesler S, Bonneté F (2006). MPCD: a new interactive on-line crystallization data bank for screening strategies. Acta Crystallogr. D Biol. Crystallogr. 62: 1311–1318.

Chen K, Kurgan L, Rahbari M (2007). Prediction of protein crystallization using collocation of amino acid pairs. Biochem. Biophys. Res. Commun*.* 355: 764–769.

Chobotova K, Vernallis AB, Majid FAA (2010). Bromelain's activity and potential as an anti-cancer agent: current evidence and perspectives. Cancer Letters. 290(2): 148–156.

Fatahiya MT, Fadzilah AAM, Nurul BAK (2016). Structure prediction of Stem Bromelain from pineapples (*Ananas Comosus*) using procaricain enzyme as a modelling template. Int. J Appl Eng Res. 11(9): 6109-6111.

Fitch WM (1970). "Distinguishing homologous from analogous proteins". Systematic Zoology. 19 (2): 99–113.

Gautam SS, Mishra SK, Dash V, Goyal AK, Rath G (2010). Comparative study of extraction, purification and estimation of bromelain from stem and fruit of pineapple plant. Thai J Pharm. 34(1): 67-76.

Gilliland GL (1988). A biological macromolecule crystallization database: a basis for a crystallization strategy. J. Cryst. Growth, 90, 51–59.

Große-Veldmann B, Korotkova N, Reinken B, Lobin W, Barthlott W (2011). "Amborella trichopoda — Cultivation of the most ancestral angiosperm in botanic gardens*".* The J Botanic Gard Hort. 9: 143–155.

Grzonka Z, Kasprzykowski F, Wiczk W (2007). Cysteine proteases. In: Industrial Enzymes. Polaina J, MacCabe P, (Eds.), Springer, NY, pp. 181-195.

Hale LP (2004). Proteolytic activity and immunogenicity of oral bromelain within the gastrointestinal tract of mice. Int Immunopharmacology*.* 4(2): 255–264.

Hennessy D, Buchanan B, Subramanian D, Wilkosz PA, Rosenberg JM (2000). Statistical methods for the objective design of screening procedures for macromolecular crystallization. Acta Crystallogr. D Biol. Crystallogr. 56: 817–827.

Jérémie J (1982). "Amborellacées". In. Flore de La Nouvelle-Calédonie et Dépendances (in French). Aubréville A, Leroy JF (Eds.), Muséum National d'Histoire Naturelle.Paris, pp. 157–160.

Kantardjieff KA, Rupp B (2004). Protein isoelectric point as a predictor for increased crystallization screening efficiency. Bioinformatics. 20: 2162–2168.

Kellogg EA (2009). The evolutionary history of ehrartoideae, oryzeae, and Oryza. Rice. 2:1–14.

Khan R, Rasheedi S, Haq S (2003). Effect of pH, temperature and alcohols on the stability of glycosylated and deglycosylated stem bromelain. J Biosci. 28(6):709–714.

Koonin EV (2005). Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet. 39: 309–338.

Kyte J, Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105–132.

Li P, Dada JO, Jameson D, Spasic I, Swainston N, Carroll K, Dunn W, Khan F, Malys N, Messiha HL, Simeonidis E, Weichart D, Winder C, Wishart J, Broomhead DS, Goble CA, Gaskell SJ, Kell DB, Westerhoff HV, Mendes P, Paton NW (2010). Systematic integration of experimental data and models in systems biology. BMC Bioinformatics. 11(1): 582.

Lopes FLG, Junior JBS, Souza RR, Ehrhardt DD, Santana JCC, Tambourgi EB (2009). Concentration by Membrane Separation Processes of a Medicinal Product Obtained from Pineapple Pulp. Braz Arch Biol Technol. 52 (2): 457-464.

López LM, Sequeiros C, Natalucci CL, Brullo A, Maras B, Barra D, Caffini NO (2000). Purification and characterization of macrodontain I, a cysteine peptidase from unripe fruits of *Pseudananas macrodontes* (Morr.) harms (Bromeliaceae). Protein Expr Purif. 18(2):133-40.

Luft JR, Wolfley JR, Snell EH (2011). What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. Cryst. Growth Des. 11, 651–663

Maurer HR (2001). Bromelain: biochemistry, pharmacology and medical use. Cellular and Mol Life Sci. 58(9): 1234–1245.

Marshall JS, Golden KD (2012). Characterization of Bromelain from *Morinda citrifolia* (Noni). J. Sci. Res. 4 (2), 445-456.

Mondal S, Bhattacharya S, Pandey JN, Biswas M (2011). Evaluation of acute anti-inflammatory effect of Ananas Comosus leaf extract in Rats. Pharm Online*.* 3: 1312–1315.

Murachi T, Yasui M, Yasuda Y (1964). Purification and physical characterization of stem bromelain. Biochem . 3(1):48–55

Neta JLV, DaSilva LA, Lima AA, Santana JC, Leite NS, Ruzene DS, Silva DP, DeSouza RR (2012). Bromelain Enzyme from Pineapple: In Vitro Activity Study under Different Micropropagation Conditions. Appl Biochem Biotechnol. 168(2): 234-246.

Omotoyinbo OV, Sanni DM (2017). Characterization of Bromelain from Parts of Three Different Pineapple Varieties in Nigeria. Am. J. BioSci. 5(3): 35-41.

Pavan R, Jain S, Shraddha, Kumar A (2012). Properties and Therapeutic Application of Bromelain: A Review. Biotech. Res. Intl*.* 976203.

Pettit FK, Bare E, Tsai A, Bowie JU (2007). Hot Patch: A Statistical Approach to Finding Biologically Relevant Features on Protein Surfaces. J. Mol. Biol. 369(3):863–879.

Ritonja A, Rowan AD, Buttle DJ, Rawlings ND, Turk V, Barrett AJ (1989). Stem bromelain: amino acid sequence and implications for weak binding of cystatin. FEBS Lett 247: 419–424.

Sanni DM, Fatoki TH, Omotoyinbo OV (2017). Comparative Evaluation of Computational and Experimental Analysis of Polyphenol Oxidase from Cocoa (*Theobroma cacaoL*.). J Microbiol. Biotech. Res. 7(1): 18-25.

Smialowski P, Schmidt T, Cox J, Kirschner A, Frishman D (2006). Will my protein crystallize? A sequence-based predictor. Proteins. 62: 343–355.

Tochi BN, Wang Z, Xu SY, Zhang W (2008). Therapeutic application of pineappleprotease (Bromelain): A Review. Pak J Nutr. 7(4): 513–520.

Yamada F, Takahashi N, Murachi T (1976). Purification and characterization of a protease from pineapple fruit, fruit bromelain FA2. J Biochem. 79: 1223–1234.