Isolation, identification and in silico analysis of alpha-amylase gene of *Aspergillus niger* strain CSA35 obtained from cassava undergoing spoilage

Oghenetega J. Avwioroko<sup>ab</sup><sup>⁎</sup>, Akpovweheee A. Anigboto<sup>b</sup>, Nnanna N. Unachukwu<sup>a</sup>, Nyerhovwo J. Tonukari<sup>b</sup>

<sup>a</sup>Biochemistry Division, Department of Chemical Sciences, College of Natural Sciences, Redeemer's University, Ede, Osun State, Nigeria

<sup>b</sup>Department of Biochemistry, Faculty of Science, Delta State University, P.M.B. 1, Abraka, Nigeria

<sup>⁎</sup>E-mail addresses: joavwioroko@gmail.com, avwiorokoo@run.edu.ng (O.J. Avwioroko).

**ABSTRACT**

In this investigation, a gene (*CDF_Amyl*) encoding extracellular α-amylase in *Aspergillus niger* strain CSA35 associated with cassava spoilage was amplified using specific primers and characterized in silico. The gene had a partial nucleotide sequence of 968 bp and encoded a protein of 222 aa residues with a molecular weight and isoelectric point of 25.13 kDa and 4.17, respectively. Its catalytic site was located in the active site domain. BLASTp analysis showed that the protein primary sequence of the α-amylase gene had 98% and 99% homologies with the α-amylase of *A. niger* and *A. oryzae* RIB40, respectively. The gene is more closely related to α-amylase genes from fungi than to bacterial, plant, or animal α-amylase genes. Restriction mapping of the gene showed it can be digested with restriction enzymes like *NcoI*, *PstI*, *SmaI*, and *BcI* among others but not with *EcoRI* and *EcoRV*. Its protein product had a hydrophobicity score of 0.43 but no transmembrane helix. The *CDF_Amyl* gene is subcellularly localized in the secretory pathway, an indication of its release into extracellular space after secretion. Also, the 3D structure of the *CDF_Amyl* protein was barrel-shaped with domains characteristic of α-amylases. The encoded α-amylase *V*<sub>max</sub> is 6.90 U/mg protein and *K*<sub>m</sub> is 6.70 mg/ml. It was concluded that the unique characteristics of the *CDF_Amyl* gene and its deduced protein could find applications in biotechnological, food and pharmaceutical industries where cloning and further modification of this gene would be required for product development and improvement.

**Keywords:**

*Aspergillus niger*

Alpha-amylase gene

Bioinformatics characterisation

Restriction enzymes

Phylogenetics, evolution

1. Introduction

α-Amylases (EC 3.2.1.1) hydrolyse polysaccharides such as starch, glycolgen etc via random breakdown of their internal α - 1, 4-glucosidic bonds to produce oligosaccharides of various sizes [16,31]. They are among the most sought-after enzymes because of their high relevance in present-day biotechnology. α-Amylases have wide applications in starch processing, brewing industries, detergent, textile, pharmaceutical industries etc [31,41]. Although they can be obtained from almost all living organisms, the enzymes from microbial sources (basically bacteria and fungi) tend to be more suitable for industrial applications.

α-Amylases are grouped as family 13 of the glycosyl hydrolases. They have similar structures and catalytic sites and do not vary in their catalytic mechanism. A typical α-amylase has a three dimensional (3D) structure resembling a barrel shape with three characteristic domains – A, B and C [14,27]. Situated in loops at the C-termini of β-strands in the A domain are highly conserved residues of amino acid which participate in catalysis and binding of substrates [27]. The B domain is involved in binding of Ca<sup>2+</sup> and also participates in binding of substrate. The C-domain's role, a second protrusion in front of or behind the A-domain, however, is not fully known; but a study by Holm et al. [21] on genetic aberrations in the C-domain of *Bacillus stearothermophilus* α-amylase suggests that the domain plays a role in enzyme activity.

Amylases vary in molecular weights; usually from about 10–210 kDa [40]. The presence of carbohydrate moieties in some organisms' amylases tend to raise their molecular weights; for example, glycoproteins have been observed in *Aspergillus oryzae*, Lipomyces kononenkoae, *Bacillus stearothermophilus* and *Bacillus subtilis* strains. The isoelectric point (pI) of amylases ranges from 3.25 to 10.1 [12]. The pI refers to the pH at which an enzyme's overall electric charge is zero [10]. Knowledge of enzymes' pI values is used for electro focussing.

Over the years, a number of application softwares or programmes have been developed and are being used for easy study of gene and...
Protein sequences in a new field of molecular biology presently referred to as bioinformatics or computational biology ([46]; [47]). These bioinformatics tools make characterisation of such nucleotide or protein sequences less stressful, time conservative and cost effective aside their highly precise or accurate outcomes. Some of these computational tools are utilized in an array of in silico analyses ranging from determination of evolutionary relatedness of genes or enzyme proteins (phylogeny), molecular weights, isoelectric points (pI), transmembrane helices, hydrophobicity, subcellular localization of proteins, three-dimensional (3D) structure of proteins, conserved domains, detection of introns, to in silico determination of restriction enzyme cleavage sites in gene sequences ([25,7,8]; Blum et al.[9]; [33,24]).

Food spoilage generally has been closely linked with the action of some microorganisms capable of degrading some nutrient materials in food substances[1,39]. Cassava (Manihot esculenta) and its products, being notable rich sources of carbohydrate, water and some minerals

Fig. 1. Partial nucleotide sequence and deduced amino acid sequence of the cassava degrading fungi amylase gene (CDF_Amyl). Four introns (non coding regions) were found in the CDF_Amyl gene sequence. The introns were denoted in lower case letters (italicized) while the exons (S) in upper case letters.

| Matched protein/organism | Max score | Total score | Query cover | E-value | Identity | Accession number |
|--------------------------|-----------|-------------|-------------|---------|----------|------------------|
| Alpha-amylase A type - 1/2 [Aspergillus oryzae RIB40] | 462 | 462 | 100% | 4e−159 | 99% | XP_003189619.1 |
| Chain A, Orthorhombic Crystal Structure Of Aspergillus niger Alpha-Amylase | 461 | 461 | 100% | 7e−159 | 98% | 2GUY_A |
| Alpha-amylase, putative [Aspergillus flavus NRRL3357] | 461 | 461 | 100% | 1e−158 | 98% | XP_002374124.1 |
| Alpha-amylase [Aspergillus kawachi] | 461 | 461 | 100% | 1e−158 | 98% | BAD01051.1 |
| Taka-amylase A (Taa-G1) precursor [Aspergillus oryzae] | 460 | 460 | 100% | 3e−158 | 98% | AAA32708.1 |
| Alpha-amylase [Aspergillus awamori] | 460 | 460 | 100% | 3e−158 | 98% | BAD06002.1 |
| Alpha amylase Precursor [Aspergillus awamori] | 459 | 459 | 100% | 4e−158 | 98% | Q02905.1 |
| Alpha Amylase Precursor [Aspergillus shirasuami] | 455 | 455 | 100% | 3e−157 | 98% | P30292.1 |
| Taka-amylase A precursor [Aspergillus oryzae] | 450 | 450 | 100% | 2e−154 | 97% | BAA00336.1 |
| Alpha-amylase [Aspergillus sojae] | 499 | 499 | 100% | 7e−154 | 95% | BAM28835.1 |
support the growth of a number of microbes, including fungi. Hence, the spoilage of cassava tubers or their processed products after a long period of exposure to air or dusts, are often times associated with visible growth of Aspergillus species on or around them with characteristic colour(s) and odour. These saprophytic fungi are believed to grow and survive on the cassava tubers by secreting starch hydrolyzing enzymes such as amylases, amidst others, thereby breaking down cassava starch to simple sugars for nutrition.

The present study amplified and characterized the α-amylase gene of Aspergillus niger strain CSA35 using in silico analyses. The fungi, which grow on decaying cassava tubers and peels, had been identified in a previous study. The characterisation of the fungal cassava starch-degrading α-amylase gene and its deduced protein sequence would facilitate the potential industrial applications of the α-amylase gene and its protein product.

2. Material and methods

2.1. Preparation of cassava flour

Peeled cassava (M. esculenta) tubers were processed into flour according to the method described by Oyewole and Sanni.

2.2. YPD agar preparation

Preparation of yeast extract-peptone dextrose (YPD) agar was carried out as described by Tonukari et al. [44]. Thereafter, it was slightly cooled and poured into plates to solidify.

2.3. Growth of fungi

Aspergillus niger CSA35, previously identified using 18S rRNA sequence and reported to be associated with cassava spoilage in Nigeria [4], was inoculated into a sterile YPD agar in a Petri dish. It was allowed to grow and sporulate in the dark for one week and then subcultured in sterilized Cassava Starch Agar (CSA) which contained cassava flour (2%) as the main carbon source, NaNO₃ (1%) and agar powder (1.5%), for additional 7 days. To avoid growth of bacteria, 20% ampicillin (5 µl) was included in the subcultures. Thereafter, the hyphae were harvested.

2.4. Genomic DNA isolation

Genomic DNA was extracted from the fungi using a modified cetyl trimethyl ammonium bromide (CTAB) method and purified using ethanol precipitation. The presence and quality of the isolated genomic DNA was ascertained via agarose gel electrophoresis using 0.8% agarose gel stained with ethidium bromide (5 µg/µl). Images of bands obtained in the gel were viewed and captured using a UV-light automated Gel Documentation machine (Enduro, USA).

2.5. PCR amplification of α-amylase gene

Alpha amylase gene was amplified by polymerase chain reaction (PCR) method as described by Zidani et al. [49] using the specific forward and reverse primers (AmyFP 5′-CATCTGGATCACCCCCGTTA-3′ and AmyRP 5′-AGACTTACGAAGCGAACCGT-3′, respectively) and the isolated fungal genomic DNA as template. The PCR mixture (25 µl) contained, 1.25 µl each of forward and reverse primers (5 mM each), 1.25 µl of 2.5 mM dNTP mix (dATP, dGTP, dTTP, dCTP) (Promega), 10 × NH₄ buffer (1.25 µl), 1.0 µl of 50 mM MgCl₂ (Promega), 1.0 µl of 5 unit/µl of Taq DNA polymerase, 7.25 µl of dd·H₂O and 7.25 µl of 10 ng/µl of genomic DNA. The PCR program was performed on a thermal cycler (Biometra) system as follow: Stage 1 (x1): 94 °C for 5 min to achieve the initial DNA denaturation; Stage 2 (x 35): 94 °C for 30 s, for denaturation; 60 °C for 30 s, for annealing; and 72 °C for 5 min, for extension. Stage 3 (x1): 72 °C for 5 min and 10 °C for in infinity (∞). The presence of amplified α-amylase genes was established by agarose gel electrophoresis using 1.0% agarose gel stained with 3 µl of ethidium bromide (5 µg/µl) and viewed with Enduro Gel Doc machine.

2.6. Gene sequencing

The purified α-amylase amplicons were sequenced using an automated Applied Biosystem 3130xl Genetic Analyzer (ABI Sequencer, USA). The sequencing was carried out at the Bioscience Center, International Institute for Tropical Agriculture (IITA), Ibadan, Nigeria.

2.7. Bioinformatics characterisation of putative α-amylase gene

The nucleotide sequence of α-amylase obtained was translated to its corresponding protein primary sequence using the European Fig. 2. Evolutionary relationship between cassava-degrading fungal amylase (CDF_Amyl) and those in plants, fungi, bacteria, lower animals and human in NCBI GenBank.
Biotechnology Information (EBI) EMBOSS translation tool and National Center for Biotechnology Information (NCBI) BLASTX tool; the most suitable reading frame was selected and exons identified. Identifications of closely matched α-amylase sequences and conserved domains were carried out with the aid of NCBI BLASTp and CD tools available online at http://blast.ncbi.nlm.nih.gov/Blast.cgi and https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi respectively [30], while multiple sequence alignment and construction of phylogeny tree were done using EMBL EBI ClustalW2 and Phylogeny tools available at http://www.ebi.ac.uk/Tools. Prediction of CDF_Amyl molecular weight (Mw) and isoelectric point (pI) was done using ExPaSy - Compute pI/Mw tool available at http://web.expasy.org/compute_pi/, while the three-dimensional (3D) structure was determined using Phyre2 software available at http://www.sbg.bio.ic.ac.uk/phyre2/ [25]. The subcellular localization of gene protein product was predicted using MultiLoc2 software available at http://abi.inf.uni-tuebingen.de/Services/MultiLoc2 (Blum et al., 2009). In silico determinations of hydrophobic/hydrophilic nature and transmembrane helices of the deduced protein were carried out using the translated amino acid sequence of the gene as input data in specialized bioinformatics applications available at http://harrier.nagahama-i-bio.ac.jp/sosui/sosui_submit.html [20,32,33] and http://www.cbs.dtu.dk/services/TMHMM-2.0/ [24], respectively.

3. Results and discussion

Genes in their native forms generally comprise protein-coding nucleotide sequences called exons and non protein-coding sequences referred to as introns (de Conti et al., 2013). During posttranslational
modifications, however, introns are spliced leaving exons that encode synthesis of functional proteins (de Conti et al. [13]). In this study, we observed the presence of four (4) introns and five (5) exons in the *Aspergillus niger* CSA35 α-amylase gene (CDF_Amyl) (Fig. 1). The gene has a partial nucleotide sequence of 926 bp and a deduced protein sequence of 222 amino acids (Fig. 1). Most fungi have been reported to harbor at least one fungal-type α-amylase gene that is evolutionarily intron-rich (da Lage et al. [11]).

The protein-protein BLAST (BLASTp) result obtained showed that the cassava starch-degrading partial α-amylase gene (CDF_Amyl) shares 98–99% identity with α-amylase proteins from *Aspergillus oryzae* RIB40, *A. niger*, *A. flavus* NRRL3357, *A. kawachii* (Table 1) and several other *Aspergillus* species and strains already deposited in NCBI Genbank. This is an indication that the amplified alpha amylase gene from *Aspergillus niger* CSA35 actually encodes α-amylase protein. Preliminary biochemical analysis of the encoded α-amylase protein secreted by the fungus showed that it has a maximum specific activity (Vmax) of 6.90 U/mg protein and a Michaelis constant (Km) of 6.70 mg/ml using soluble starch as substrate [5].

The evolutionary relationship (phylogeny) between the cassava starch-degrading fungal partial α-amylase gene (CDF_Amyl) and α-amylase genes in plants (rice, maize, banana, and cassava plant), other fungi, bacteria, animals (Chimpazee, goat) and human saliva is as shown in Fig. 2. Phylogenies help to provide necessary information on biological diversity, for structuring classifications as well as for proper understanding of events that occurred during evolution [7,8]. The α-amylase gene amplified from the *Aspergillus niger* CSA35, associated with cassava spoilage in Nigeria, is more closely related to α-amylase genes from fungi than to bacterial, plant, or animal α-amylase genes (Fig. 2), as expected in the absence of horizontal gene transfers [23].

The amino acid residues present in the conserved domain of the fungal partial amylase gene (CDF_Amyl) are as shown in Fig. 3a. α-
Amylase genes generally possess catalytic regions that comprise aspartate (D), glutamate (E) and histidine (H) residues that are involved in the enzymes’ mechanism of starch hydrolysis [26]. In the present research work, the amino acid residues present in the conserved domain of the CDF_Amyl gene include histidine (H), arginine (R), aspartate (D), and glutamate (E) amongst others (Fig. 3a). The catalytic site is present in the active site region (Fig. 3b). These findings agree with the report of Liu et al. [26] and affirm that the amylase gene (CDF-Amyl) belonged to the superfamily of α-amylases.

A typical structural characteristic of α-amylase proteins is their barrel-shape appearance [27]. The barrel shape is made up of curly polar α-helices in the outer parts of the protein and β-sheets close to the middle of the protein. The 3D-protein structure of the cassava starch-degrading fungal amylase (CDF_Amyl) reported in this study (Fig. 4) is identical to that of α-amylase isolated from Aspergillus oryzae [27] and also agrees with the typical α-amylase structure described by de Souza and Magalhães [14] and a number of others already deposited in NCBI genbank.

As clearly depicted by the endonuclease restriction linear map shown in Fig. 5, a large array of restriction enzymes (such as PstI, Smal, NcoI, BclI etc) can be used to cut the cassava-degrading fungal α-amylase gene (CDF_Amyl, 926 bp) when considering cloning and gene expression experiments. However, restriction enzymes like EcoRI and EcoRV will be unable to cut the CDF_Amyl gene (Fig. 5). These findings are in agreement with those of Yu [48] who observed that PstI and Smal can cut alpha-amylase gene (αAmy3 gene enhancer) isolated from rice (Oryza sativa). They are, however, in contrast to the report of Puspasari et al. [36] who digested raw starch-degrading α-amylase from Bacillus aquimaris MKSC6.2 using EcoRI and EcoRV during a gene expression study. The choice of restriction enzyme to use when considering digestion of the partial amylase gene (CDF_Amyl) in molecular biology studies, however, will not only depend on the choice of restriction enzyme but also on a number of other factors such as the size of gene fragment required, the ability of the choice restriction enzyme to cut both the potential vector of interest and the amylase gene (CDF_Amyl), as well as the goal or nature of the research being considered, etc [18,29,43].

The bioinformatics analysis also showed that the extracellular enzyme (alpha amylase) secreted by Aspergillus niger CSA35 has an isoelectric point (pl) of 4.17 and a molecular weight (Mw) of 25.13 kDa (Table 2). Molecular weights of amylases vary from about 10–210 kDa but most of these enzymes are in the range of 50–60 kDa. Amylases from Bacillus caldolyticus and Chloroflexus aurantiacus have the lowest (10 kDa) and highest (210 kDa) molecular weights, respectively [19,40]. The pl values are quite useful for understanding of enzyme behaviour against different purification procedures at different pH values. Amylase isoelectric points are generally in the range of 3.25–10.1 [12]. The pl value obtained in this study is also in accord with the report of Ahmad et al. [2] who opined that alpha amylase from Aspergillus oryzae is an acidic enzyme with optimum pH of 4.5 and low isoelectric point (pl = 4.2). The knowledge of the isoelectric point and the Mw would be of great advantage to laboratories or industries when considering small or large scale extraction of the CDF_Amyl protein.

In silico analysis of the subcellular localization of the α-amylase protein in A. niger CSA35 revealed that the enzyme (CDF_Amyl) is predominantly found in the secretory pathway within the cassava degrading fungi (Fig. 6). Soluble proteins are synthesized in the secretory pathway and thereafter delivered into the extracellular space where they are utilized for cellular functions [17]. Our report is in agreement with the reports of de Souza and Magalhães [14] and Deb et al. [15] that α-amylases are mainly extracellularly localized.

Many studies have shown that the amino acid sequence of proteins play a crucial role in determining solubility or hydrophobicity of expressed proteins [22,32,33]. In this study, it was observed that the CDF_Amyl protein had a hydrophobicity score of ~ 0.43 which indicates that the amino acid sequence is of a soluble protein (Fig. 7). Our present finding that the cassava starch-degrading fungal α-amylase (CDF_Amyl) is a soluble protein also agrees with the fact that the enzyme is localized in the secretory pathway (Fig. 6), since proteins secreted into extracellular space are mostly hydrophilic in nature [38].

The results of the in silico analysis also indicate that the amylase
The protein (CDF_Amyl) under study is not a transmembrane protein (Fig. 8). This is also in agreement with the previous findings of this study that the alpha amylase (CDF_Amyl) from A. niger CSA35 is not hydrophobic but a secreted, soluble protein. According to Phillips et al. [35], transmembrane proteins are typically hydrophobic in nature since they span through the non-polar phospholipid bilayers of membranes and a large proportion of their structures are embedded in the lipids. Separation of such proteins from biomembranes is usually difficult except with use of nonpolar solvents, detergents, and/or denaturing agents [37]. However, the CDF_Amyl protein, being a non-transmembrane and hydrophilic protein, is not bound to the fungal cell membrane but rather secreted into the fungi growth medium. This characteristic can be exploited during industrial isolation and purification of the cassava starch-degrading fungal α-amylase to obtain the functional protein in its intact or un-denatured form [45].

4. Conclusion

The present study findings strongly affirmed that the amplified amylase gene from A. niger CSA35 is a member of the α-amylase superfamily. The unique characteristics of the CDF_Amyl gene and its deduced protein could find applications in biotechnological, food and pharmaceutical industries where cloning and further modification of this gene would be required for product development and improvement.

Acknowledgements

The authors greatly appreciate all research personnel at the African Research Laboratories (ARL), Otorho-Agbon, Delta State and Bioscience Center, IITA, Ibadan, Nigeria for their kind assistance during the research.

Conflict of interests

The authors do not declare any conflict of interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Transparency document

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.bbrep.2018.03.006.

References

[1] C.O. Adebayo, B.I. Aderiye, O.B. Akpor, Assessment of bacterial and fungal spoilage of some Nigerian fermented and unfermented foods, Afri. J. Food Sci. 8 (3) (2014) 140–147.
[2] R. Ahmad, M. Mohsin, T. Ahmad, M. Sardar, Alpha amylase assisted synthesis of TiO2 nanoparticles: structural characterization and application as antibacterial agents, J. Hazardous Mater. 283 (2015) 171–177.
