CHARACTERIZATION of THE $\alpha$-AMYLASE GENE
FROM Bacillus sp. BBM1

CARACTERIZACIÓN DEL GEN DE $\alpha$-AMILASA DE LA CEPA NATIVA Bacillus sp. BBM1

Juliana MUÑOZ1, Mónica QUINTERO1, Pablo A. GUTIÉRREZ*

Received: 08 August 2010 Accepted: 21 September 2011

ABSTRACT

Starch degrading enzymes represent about 30% of the enzyme world and they are used in the production of glucose, maltose and oligosaccharides, which can be further processed to produce fructose and dextrose syrups. The resulting glucose can also be fermented for the production of ethanol, amino acids and organic acids. $\alpha$-amylases are also used as an alternative to the addition of malt in the beer industry, the improvement of flour in the baking industry, the removal of starch in the paper and textile industry, and as a detergent additive. In this paper, the complete nucleotide sequence of the $\alpha$-amylases BBM1 produced by the native strain Bacillus sp. BBM1 is reported. The deduced aminoacid sequence shows that this enzyme is translated as a 659 a.a. protein, which after the secretion cleaves to generate a 618 mature enzyme of 68 kDa. The BBM1 $\alpha$-amylases is transcribed as a monocistronic mRNA, as it is suggested by the presence of inverted repeat elements upstream and downstream of the protein coding region. The expression of the BBM1 $\alpha$-amylase is under the control of the AmyR1 allele from the AmyO operator sequence, which is recognized by the regulatory protein CcpA. In contrast to most $\alpha$-amylases, BBM1 seems to possess two additional carbohydrate-binding domains, which probably increase its efficiency in the degradation of granular starch. A homology model of the enzyme is presented and its interaction with calcium ions and substrate is analyzed.

Keywords: Starch, polysaccharides, glucan 1,4-alpha-Glucosidases, enzymology.

RESUMEN

Las enzimas degradadoras del almidón representan cerca del 30% del mercado mundial de enzimas y son utilizadas en la producción de glucosa, maltosa y oligosacáridos; los cuales pueden ser transformados posteriormente en jarabes de fructosa y dextrosa. La glucosa también puede ser utilizada en la producción de etanol, aminoácidos y ácidos orgánicos. La $\alpha$-amilasa también puede ser utilizada como una alternativa a la adición de malta en la industria de la cerveza, el mejoramiento de harinas y la remoción de almidón en la industria papelera y textil y como aditivo de detergentes. En este trabajo reportamos la secuenciación completa del gen codificante para la $\alpha$-amilasa BBM1 producida por la cepa nativa Bacillus sp. BBM1, incluyendo sus regiones reguladoras 3’ y 5’. La secuencia de aminoácidos corresponde a una proteína de 659 residuos que, luego de ser secretada y procesada post-traduccionalmente, da origen a una enzima madura.


* Corresponding author: paguties@unal.edu.co
INTRODUCTION

α-amylases (EC 3.2.1.1) are enzymes that catalyze the hydrolysis of the α-1,4 glycosidic bonds that are present in starch, glycogen and other polysaccharides. Starch is composed of two different glucose polymers: amylose, linked by α-1,4 bonds; and amylpectin that, in addition to α-1,4 bonds, has α-1,6 branch sites (1). Starch degrading enzymes represent about 30% of the enzyme world in commercial products such as Ultra-thin® (Valley Research/Diversa), Multiject AA 21L® (Genencor), Termamyl® and Liquozyme® (Novozymes) (2). The starch industry is the main user of α-amylases for the production of glucose, maltose and oligosaccharides, which can be further processed to produce fructose and dextrose syrups. The resulting glucose can also be fermented for the production of ethanol, amino acids and organic acids (3). α-amylases are also used as an alternative to the addition of malt in the beer industry, the improvement of flour in the baking industry, the removal of starch in the paper and textile industry, and as a detergent additive (4). Previously, we reported the purification and biochemical characterization of an extracellular α-amylase produced by Bacillus sp. BBM1, a native strain that has been isolated by our research group (5). This enzyme has an optimal temperature of 60°C, but it can be used with 80% of efficiency at temperatures up to 72°C. In this paper, the analysis of the complete nucleotide sequence of α-amylase BBM1, including the 5’ and 3’ intergenic regions, is reported. A homology model was constructed and used to understand the interactions of this enzyme with the ions and the substrate.

MATERIALS AND METHODS

DNA extraction and PCR

Bacillus sp. BBM1 was isolated from a clay soil sample at Universidad Nacional de Colombia sede Medellín with the following characteristics: pH 5.4, 36 g/kg organic matter, 7 mg/kg NO₃, 16 mg/kg NH₄⁺, and 20 mg/kg phosphate. Bacillus sp. BBM1 grows optimally at 30°C and pH 7.0 in M9 minimal media and LB. A neighbor-joining analysis revealed a phylogenetic affinity of Bacillus sp. BBM1 with B. methylotrophicus, a recently characterized species isolated from the rhizosphere (5). Bacillus sp. BBM1 was grown in LB medium at 30ºC and DNA was purified with the DNeasy Blood & Tissue kit (QIAGEN), following the manufacturer’s protocol. A total of six primers were used for the amplification of the AmyE gene, as it is shown in figure 1 and table 1. Primers ycgB-F, ldh-R and ldh2-R annealed to flanking genes ycgB and ldh. Primers AmyE1I-F, AmyE2I-R and amy3-F annealed to the internal regions of AmyE. The 5’ and 3’ regions of AmyE were amplified with primers ycgB-F/AmyE2I-R and AmyE1I-F/ldh-R, respectively. Sequence gaps were completed with the primer combination amy3-F/ldh2-R. Each PCR reaction consisted of a total volume of 50 mL containing 0.05 u/ml of Taq DNA polymerase, 1X PCR buffer, 2.5 mM MgCl₂, 0.2 mM of each deoxynucleotidetriphosphate, 0.5 mM for each primer, and 50 ng/ml of template DNA. The PCR conditions were the following: initial denaturation at 94°C for 1 min, 30 cycles of amplification consisting of a 30 second denaturation step at 94°C, a 30 s annealing step, and a 3 min extension step at 72°C. The annealing temperature was set according to the primer melting temperature reported in table 1. DNA amplicons were visualized in 1% agarose gel electrophoresis in TBE buffer. DNA was sequenced by Macrogen Inc. (Seoul, Korea) and assembled with the CAP Contig assembly module from BIOEDIT (6). The complete sequence of the AmyE gene was deposited in GenBank with the accession code GU947099.
**Characterization of the α-amylase gene from Bacillus sp. BBM1**

Table 1. Primers used for sequencing the α-amylase gene of Bacillus sp. BBM1.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
<th>Tm (°C)</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>amyE1I-F</td>
<td>TGATGCAGGATATACACCCATTCAGAGC</td>
<td>50.9</td>
<td>2822</td>
</tr>
<tr>
<td>ldh-R</td>
<td>CACCGTATTGCCCCGTCC</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>ycgB-F</td>
<td>TGGTCGAGGATGTACGACCAGG</td>
<td>62.2</td>
<td>2231</td>
</tr>
<tr>
<td>amyE2I-R</td>
<td>GCCAGGTCAGCGTTGAATTCC</td>
<td>63.5</td>
<td></td>
</tr>
<tr>
<td>amy3-F</td>
<td>AGGACACCGCGGAGGTTGCAGC</td>
<td>66.1</td>
<td>1073</td>
</tr>
<tr>
<td>ldh2-R</td>
<td>GCCGCTGCCGATGCAGCG</td>
<td>67.6</td>
<td></td>
</tr>
</tbody>
</table>

**Sequence annotation**

The coding region for the AmyE gene from the Bacillus sp. BBM1 was identified by means of the BLASTX software (7). Transcription terminators were predicted with RNAfold (8). The Shine-Dalgarno sequence (SD) was identified through a sequence complementarity to the 3’ region of the 16S RNA of B. amyloliquefaciens FZB42 (NC_009725.1) and B. subtilis subsp. subtilis str. 168 (NC_000964.3), using a Perl routine written by the authors. The promoter region was identified with a custom designed program that used a Position Specific Scoring Matrix (PSSM) derived from the promoter regions of the α-amylase gene from B. subtilis, B. stearothermophilus, B. licheniformis and B. amyloliquefaciens (9).

**Phylogenetic Analysis**

A set of 136 non-redundant α-amylases sequences from fungi and bacteria were downloaded from the SwissProt/TrEMBL database (10). A multiple alignment was performed with CLUSTALX using the Gonnet substitution matrix and penalties of 10 and 0.2 for gap insertion and extension, respectively (11). Phylogenetic trees were built using MEGA4 (12). Sequences in the trees were labeled using their corresponding entry name at SwissProt/TrEMBL.

**Homology modeling**

The N-terminal domains (A, B and C) were modeled using the crystallographic structure of Bacillus subtilis α-amylase as a template (PDBid: 1BAG) (13). The template for the C-terminus was chosen with the iterative BLAST protocol PSI-BLAST (14) that identified the cyclodextrin glucanotransferases from Bacillus circulans and Geobacillus stearothermophilus as distant homologues of α-amylase BBM1 (15). The homology model was generated with MODELLER (16).

**RESULTS AND DISCUSSION**

**Sequence analysis**

Bacillus sp. BBM1 has a phylogenetic affinity with B. methylotrophicus, a species closely related to B. amyloliquefaciens and B. subtilis. In these species, the α-amylase gene is flanked by ycgB, a gene of unknown function, and ldh, coding for lactate dehydrogenase (17, 18). Based on the complete genome sequence Bacillus amyloliquefaciens FZB42 (NC_009725.1) and Bacillus subtilis subsp. subtilis str. 168 (NC_000964.3), a set of primers for the amplification and sequencing of 2971 nt (comprising the AmyE gene and the 3’ and 5’ intergenic regions) were designed (see table 1 and figure 1).

![Figure 1](attachment:image.png)

Figure 1. Genetic context of the AmyE gene and location of the primers used in this work.

The BLASTX analysis showed that α-amylase BBM1 has 99 and 98% of similarity with its homologues in Bacillus spB-5 (gb:ADB81848.1) and B. amyloliquefaciens FZB42 (gb:ABS72721.1), respectively. The expressed protein has a secretion signal that spans the first 27 residues (MFEKRFKT-SLLPLFAGF), and a cleavage site at position 41 (NK /VT) that gives rise to a mature protein of 618 amino acids with a predicted molecular weight of 68051.7 Da, as it is shown in figure 2. According to our predictions, the ribosomal binding site (RBS) corresponds to the AATAAGGAGT sequence located eight nucleotides upstream of the start codon.
In order to prove that the expression of the *AmyE* gene is coupled to the transcription of the *YcgA-YcgB* operon, an analysis of attenuators in the 5' intergenic region was performed. Secondary structure predictions suggest the presence of a 93 nt sequence downstream of the *YcgB* termination codon, which can form a hairpin structure with a free energy of -14.80 kcal/mol. This fact implies an independent transcription from the *YcgA-YcgB* operon (shown in figures 2 and 3). Another termination signal with a free energy of -28.40 kcal/mol was detected 85 nt downstream the termination codon, suggesting that the transcription of *AmyE* is not coupled to *ldh* either. Therefore, the translation of *AmyE* BBM1 is probably caused by a monocistronic mRNA, a fact that can be observed in other α-amylases from the genus *Bacillus* (19-22).
Figure 3. Predicted transcription termination signals flanking the AmyE BBM1 gene.

The analysis of the promotor region is important for understanding the regulatory mechanisms involved in the regulation of the protein expression (23-25). In this case, the -35 promotor region has the sequence TTGATA, and it is identical to the one present in the α-amylase hyperproducing strain of B. subtilis (22). The -10 sequence corresponds to TTGAAA, which is slightly different from the consensus TANAAT observed in other α-amylase genes. In B. subtilis, the expression of α-amylase is regulated in cis by the operator sequence amyO (TGT/AAANCGNTNA/TCA), which is recognized by protein factor CcpA that is a repressor of the LacI-GaR family (26). In our case, we found the sequence TGTAAGCGTTAACA, which is identical to the amyR1 allele present in B. subtilis 168 (27). In amylase overproducing strains such as B. amyloliquefaciens and Bacillus natto IAM 121, the amyO sequence has a CG substitution at position 7, which presents a weaker affinity with the repressor (28). This condition suggests that the production of α-amylase by Bacillus sp. BBM1 is under a tighter control, and that it is produced at lower levels than the corresponding hyperproducing strains.

Comparison with other microbial α-amylases

Using the UPGMA clustering method, we found that α-amylases can be divided into five groups as it is shown in figure 4:

- **Group 1**: Mostly fungal α-amylases.
- **Group 2**: Bacterial α-amylases from the genera Aliivibrio, Streptococcus, Bacillus and the Enterobacteriaceae family.
- **Group 3**: Proteobacterial α-amylases from the genera Vibrio, Aeromonas and Xanthomonas.
- **Group 4**: Bacterial α-amylases from the phyla Proteobacteria and Actinobacteria.
- **Group 5**: Bacterial α-amylases from the genera Clostridium, Streptococcus and Bacillus.

With the exception of B. megaterium (AMY_BACME) and B. circulans (AMY_BACCI), all α-amylases from the Bacillus genus are members of the groups 2 or 5. α-amylase BBM1 belongs to B.amyloliquefaciens FZB42, B. amyloliquefaciens TB2, and B. subtilis X-23. Group 5 α-amylases have an approximate size of 660 residues, in contrast to enzymes from group 2, which have an average size of 520 residues.

Figure 4. UPGMA clustering of microbial α-amylases.
Figure 5. Sequence and structure comparison between *B. subtilis* (BSU), *B. licheniformis* (BLA), and *Bacillus* sp. BBM1 α-amylases.

Most α-amylases comprise three domains: Domain A has a TIM barrel structure and corresponds to the catalytic center. Domain B originates from a protuberance in the region that connects the third helix and the strand from domain A. The structure of domain B is very variable and its function seems to be associated with substrate binding and thermostability. Domain C has a b-sandwich structure, but the information regarding its function is scarce (29). To understand
the interaction of the α-amylase BBM1 with the ions and the substrate, a homology model was built. Domains A, B and C were modeled using the structure B. subtilis α-amylase as template (PDBid: 1BAG). The selection of templates for the C-terminal extension was more complex because no structures with a significant similarity to BBM1 were found using a regular BLAST search. However, an iterative search using the PSI-BLAST software found a homology between α-amylase BBM1 and the cyclodextrin glycosiltransferases (CGTases) from Bacillus circulans and Geobacillus stearothermophilus (15); this alignment included the C-terminal extension (data not shown). This fact is not surprising as both CGTases and α-amylases belong to the family 13 of glycosyl hydrolases.

A comparison between the BBM1 model and the structure of B. licheniformis amylase (group 2) is shown in figure 5. According to our model, the substrate binding site of α-amylase BBM1 can accommodate at least five glucose units between domains A (Blue) and B (green), as it is shown in figure 5B. α-amylases interact with the substrate through multiple binding sites in the enzyme surface, as it is shown in figure 5A. Black and white circles correspond to polar and hydrophobic interactions respectively. The only difference between the binding sites of α-amylase BBM1 and its homolog in B. subtilis is the substitution F103S, which interacts with the first glucose unit (30, 31). Catalytic residues Asp174, Glu206 y Asp267 are completely conserved (shown in figure 5A, with asterisks). Glu206 is responsible for the protonation of the glycosidic oxygen. Asp 174 is involved in the following step and its function consists in performing a nucleophilic attack on C1 on the sugar position at subsite -1 (29). The restitution of the active site is achieved by water activated by deprotonated Glu206. Asp267 is an essential residue involved in raising the pKa of Glu206 (29).

Most α-amylases have calcium-binding sites that are important for the structural stability of the enzyme. B. subtilis α-amylase has three calcium ions associated to its structure, while B. licheniformis has an extra sodium ion that forms an anionic triad together with two calcium ions (shown in figure 5B in green) (32). The binding of the first calcium ion is due to interactions with Asn99, Thr135, Asp144 and His178 that are located on the interface between domains A and B, as it is shown in figure 5A. The second calcium ion is located at the protein surface, and it involves interactions with conserved residues Gly167 and D169. The third site presents an E87K substitution that is detrimental to calcium binding.

The most interesting feature of our model is that it suggests the presence of two additional starch-binding domains (shown in figure 5B in orange and red). Starch binding domains have a variable size ranging from 30 to 200 residues, and they are present in approximately 10% of the starch degrading enzymes. It has been proposed that starch-binding domains improve the affinity with granular starch, and that they allow the hydrolysis of otherwise inaccessible glycosidic bonds (33, 34). This approximation effect increases the effective concentration of the enzyme and, therefore, the hydrolysis rate (35). These domains are of great interest in biotechnology as they can be used for the production of fusion proteins, and also as a model for the study of sugar-protein interactions (36, 37). Novamyl®, a maltogenic amylase from B. stearothermophilus, is the only domain-five α-amylase which 3D structure has been solved so far (30). The effect of these extra domains and of the calcium ions on substrate binding and catalysis will be further investigated in a following research work on the recombinant enzyme.

CONCLUSION

α-amylase BBM1 is an inducible enzyme controlled by the CcpA protein, and it is transcribed as monocistronic mRNA. This protein is secreted as a mature enzyme of 619 residues, and its structure is stabilized by the calcium ions located between domains A and B. The sequence analysis suggests that α-amylase BBM1 might have two additional starch-binding domains in contrast to most conventional amylases.

ACKNOWLEDGEMENTS

We would like to thank Prof. Mauricio A. Marín from the Laboratory of Cell and Molecular Biology, and Olga Inés Montoya from the Microbial Biotechnology group at Universidad Nacional de Colombia sede Medellín for their help at different stages of this work. This project was funded by the DIME (Dirección de Investigaciones de la Universidad Nacional de Colombia Sede Medellín, Grant 20101007304).
REFERENCES


