Volume 219, number 2, 339–342 FEB 04935 July 1987

Nucleotide sequence of the α-amylase gene (ALP1) in the yeast *Saccharomycopsis fibuligera*

Tetsuya Itoh, Ichiro Yamashita and Sakuzo Fukui

Department of Fermentation Technology, Faculty of Engineering, Hiroshima University, Shitami, Saijo-cho, Higashi-Hiroshima 724, Japan

Received 20 April 1987

The complete nucleotide sequence of the secretable α-amylase gene *ALP1* from the yeast *Saccharomycopsis fibuligera* has been determined. The *ALP1* DNA hybridized to a polyadenylated RNA of 2.0 kilobases. A single open reading frame encodes a 494-amino acid protein which is highly homologous with α-amylase (Taka-amylase) of a fungus *Aspergillus oryzae*.

Yeast; α-Amylase; DNA sequence

1. INTRODUCTION

α-Amylase (EC 3.2.1.1) plays an important role in the liquefaction of starchy molecules and is widely distributed among various organisms from microorganisms to vertebrates and plants. Several yeasts including *Saccharomycopsis fibuligera* secrete α-amylase extracellularly [1–3], but the brewing yeast *Saccharomyces cerevisiae* does not. For the rationalization of the fermentation process, introduction of α-amylase genes into *S. cerevisiae* would be of great value, since transformed yeast cells would produce ethanol directly from starchy materials. Recently, we have succeeded in cloning and isolating the expression in *S. cerevisiae* of a putative secretable α-amylase gene (*ALP1*) from *Sa. fibuligera* [3].

In this paper, we describe the complete nucleotide sequence of *ALP1* gene. The deduced protein sequence was highly homologous with α-amylase (Taka-amylase) from *Aspergillus oryzae*, suggesting that *ALP1* is a structural gene for α-amylase of *Sa. fibuligera*.

Correspondence address: I. Yamashita, Department of Fermentation Technology, Faculty of Engineering, Hiroshima University, Shitami, Saijo-cho, Higashi-Hiroshima 724, Japan

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00683

2. MATERIALS AND METHODS

2.1. *DNA sequence analysis*

DNA was sequenced from M13 subclones by the dideoxy chain termination method of Sanger et al. [4].

2.2. *Northern blot analysis*

Cells of *Sa. fibuligera* strain HUT7212 [3] were cultured in 100 ml of YPS medium (1% yeast extract, 2% polypeptone and 3% soluble starch) to an absorbance at 660 nm of 1.0, harvested, and washed with sterile water. RNA was extracted from the cells as described by Jensen et al. [5]. Poly(A)$^+$ RNA was isolated with oligo(dT) cellulose as recommended by the supplier (Collaborative Research Inc., Waltham, MA). The RNA (10 μg per lane) was fractionated in agarose (1%) gels as described by McMaster and Carmichael [6], transferred to nitrocellulose paper, and hybridized as described by Thomas [7] with a

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50 © 1987 Federation of European Biochemical Societies

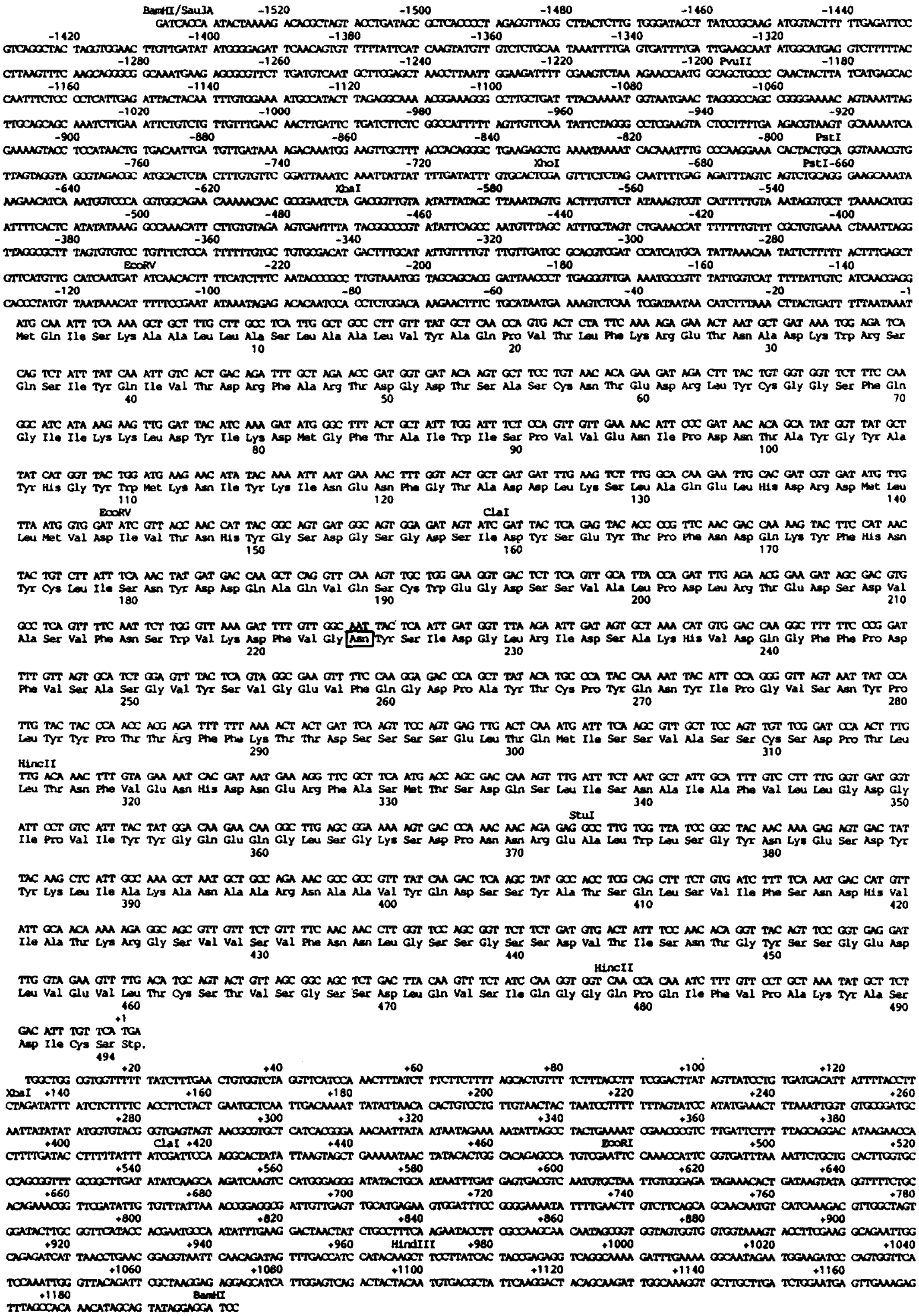

Fig.1. Sequence of the *ALP1* gene. Potential *N*-glycosylation site is boxed.

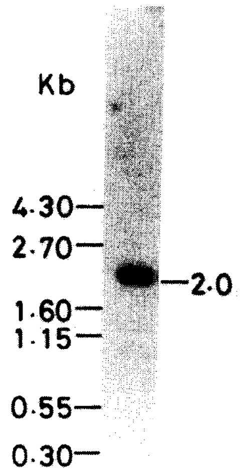

Fig.2. Northern blot analysis of the polyadenylated RNA from *Sa. fibuligera*.

nick-translated ^{32}P-labelled M13mp11 containing a *Stu*I-*Pst*I fragment of *ALP1*. Restriction fragments derived from both M13mp11 and the probe DNAs were used as size markers.

3. RESULTS AND DISCUSSION

Fig.1 shows the complete nucleotide sequence of a 4223-base-pair-long DNA fragment which is essential for α-amylase production in *S. cerevisiae*. A unique open reading frame of 1495 base pairs was found. A corresponding peptide sequence consisting of 494 amino acids is also shown in fig.1. The length of the coding region is in reasonable agreement with the observation that *ALP1*-specific mRNA was 2.0 kilobases in size (fig.2). The predicted protein carries one potentially glycosylated asparagine residue [8] (boxed in fig.1) and a hydrophobic amino-terminal segment of ~20

```
Sf  (  1)  MQISKAALLASLAALVYAQPVTLFKRETNADKWRSQSIYQIVTDRFARTDGDTSASCNTE  ( 60)
Ao  (  1)                              A-P-- -------FLL---------S-T-T---A  ( 33)
                                                                   1

Sf  ( 61)  DRLYCGGSFQGIIKKLDYIKDMGFTAIWISPVVENIPDNTAYGYAYHGYWMKNIYKINEN  (120)
Ao  ( 34)  -QK----TW----D-----QG--------T--TAQL-QDC---D--T---QTD--SL---  ( 93)
            *  1                                 *    *    *
                                S1
Sf  (121)  FGTADDLKSLAQELHDRDMLLMVDIVTNHYGSDGSGDSIDYSEYTPFNDQKYFHNYCLIS  (180)
Ao  ( 94)  Y-------A-SSA--E-G-Y----V-A--M-Y--A-S-V---VFK--SS-D---PF-F-Q  (153)
                                        *                          2
                                                               S2
Sf  (181)  NYDDQAQVQSCWEGDSSVALPDLRTEDSDVASVFNSWVKDFVGNYSIDGLRIDSAKHVDQ  (240)
Ao  (154)  --E--T--ED--L--NT-S----D-TKDV-KNEWYD--GSL-S----------TV---QK  (213)
            *        2   * ▼                       CHO          *  **
                         S3
Sf  (241)  GFFPDFVSASGVYSVGEVFQGDPAYTCPYQNYIPGVSNYPLYYPTTRFFKTTDSSSSELT  (300)
Ao  (214)  D-W-GYNK-A---CI---LD-----------VMD--L---I---LLNA--S-SG-MDD-Y  (273)
                           ****       3
                             S4     ▼
Sf  (301)  QMISSVASSCSDPTLLTNFVENHDNERFASMTSDQSLISNAIAFVLLGDGIPVIYYGQEQ  (360)
Ao  (274)  N--NT-K-D-P-S---GT-------P----Y-N-IA-AK-VA--II-N--L-I--A----  (333)
                    3              **

Sf  (361)  GLSGKSDTNNREALWLSGYNKESDYYKLIAKANAARNAAVYQDSSYATSQLSVIFSNDHV  (420)
Ao  (334)  HYA-GN--A----T-----PTD-EL-----S---I--Y-ISK-TGFV-YKNPY-KDDTT   (391)
                 *    *                                     M ← → C

Sf  (421)  IATKRGSVVSV    FNNLGSSG SSDVTISNTGYSSGEDLVEVLTCSTVSGSSDLQVSI  (475)
Ao  (392)  --MRK-TDG-QIVTILS-K-A--D-YTLSL-GAS-TA-QQ-T--IG-T--TVG--GN-PV  (452)
                                                        4

Sf  (476)    QGGQPQIFVP AKYA SDICS                                     (494)
Ao  (453)  PMA--L-RVLY-TE-L-G-K---DSS                                  (478)
                               4
```

Fig.3. Comparison of amino acid sequence between *Sa. fibuligera* (Sf) and *A. oryzae* (Ao) α-amylases. Identical amino acid residues to those in the top sequence are indicated by dashes. Asterisks and arrowheads indicate substrate-binding and catalytic residues in Taka-amylase, respectively. Pairs of cysteine residues are indicated by boxes with identical numbers. The *N*-glycosylation site is indicated by -CHO. Four highly homologous segments (S1–S4) are also boxed. The main (M) and C-terminal (C) domains of *A. oryzae* α-amylase are indicated by hooked arrows.

amino acids which resembles signal sequences found in a wide variety of secretory protein precursors [9].

Comparative studies revealed amino acid sequence homology between *Sa. fibuligera* and *A. oryzae* α-amylases as shown in fig.3. From X-ray analysis of Taka-amylase, it has been found that Glu-230 and Asp-297 residues (indicated by arrowheads in fig.3) play a central role for enzymatic activity, and that other amino acid residues indicated by asterisks in fig.3 serve for binding of substrate [10]. These residues are well conserved in the *ALP1* protein. Moreover, Taka-amylase carries 4 pairs of cysteine residues for disulfide bonds (boxed with numbers in fig.3) and one *N*-glycosylated asparagine residue (*N*-glycoside was indicated by -CHO in fig.3). Such amino acid residues were also conserved in the *ALP1* protein. Kusunoki et al. [11] have reported that Taka-amylase consists of two globular domains, main and C-terminal ones (indicated by hooked arrows in fig.3). The function of the C-terminal domain has not yet been understood. The amino acid sequence of the C-terminal region of *ALP1* protein showed very little homology with that of Taka-amylase.

A comparison of *ALP1* protein with other α-amylases from *Bacillus* sp. [12–14], human salivary glands and pancreas [15] and barley [16] showed only 4 short homologous segments (S1–S4 in fig.3) which are proposed to be essential for the catalytic reaction of Taka-amylase.

These results suggest that *ALP1* is a structural gene for α-amylase of *Sa. fibuligera*.

ACKNOWLEDGEMENTS

This work was partly supported by a grant-in-aid for scientific research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

[1] De Mot, R. and Verachtert, H. (1985) Appl. Environ. Microbiol. 50, 1474–1482.

[2] Wilson, J.J. and Ingledew, W.M. (1982) Appl. Environ. Microbiol. 44, 301–307.

[3] Yamashita, I., Itoh, T. and Fukui, S. (1985) Agric. Biol. Chem. 49, 3089–3091.

[4] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.

[5] Jensen, R., Sprague, G.F. jr and Herskowitz, I. (1983) Proc. Natl. Acad. Sci. USA 80, 3035–3039.

[6] McMaster, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835–4838.

[7] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.

[8] Struck, D.K., Lennarz, W.J. and Brew, K. (1978) J. Biol. Chem. 253, 5786–5794.

[9] Keil, G. (1981) Annu. Rev. Biochem. 50, 317–348.

[10] Matsuura, Y., Kusunoki, M., Harada, W. and Kakudo, M. (1984) J. Biochem. 95, 697–702.

[11] Kusunoki, M., Matsuura, Y., Tanaka, N. and Kakudo, M. (1983) J. Jap. Soc. Starch Sci. 30, 141–148.

[12] Yamazaki, H., Ohmura, K., Nakayama, A., Takeichi, Y., Otozai, K., Yamasaki, M., Tamura, G. and Yamane, K. (1983) J. Bacteriol. 156, 327–337.

[13] Gregory, L.G., Mainzer, S.E., Rey, M.W., Lamsa, M.H., Kindle, K.L., Carmona, C. and Requadt, C. (1986) J. Bacteriol. 166, 635–643.

[14] Takkinen, K., Pettersson, R.F., Kalkkinen, N., Palva, I., Söderlund, H. and Kääriäinen, L. (1983) J. Biol. Chem. 258, 1007–1013.

[15] Nakamura, Y., Ogawa, M., Nishide, T., Emi, M., Kosaki, G., Himeno, S. and Matsubara, K. (1984) Gene 28, 263–270.

[16] Rogers, J.C. and Milliman, C. (1983) J. Biol. Chem. 258, 8169–8174.