

DNA sequence of a β -glucosidase from *Ruminococcus albus*

Kunio Ohmiya, Masayuki Takano and Shoichi Shimizu
 School of Agriculture, Nagoya University, Chikusa, Nagoya 464-01, Japan

Submitted December 23, 1989

EMBL accession no. X15415

β -Glucosidase is an essential component of the cellulose-degrading enzyme system of cellulolytic microorganisms. The properties of the enzyme have been greatly elucidated using recombinant DNA technology. A β -glucosidase gene from *Ruminococcus albus* has been cloned and its restriction map has been already determined (1). However, further information concerning the gene properties is not available. This paper describes the DNA sequence of the gene encoding β -glucosidase of *R. albus* which has been cloned and expressed in *E. coli* using

pBR 322. The nucleotide sequence of the DNA fragment was determined to be 3158 bp. The largest potential coding sequence was 2841 bp long (encoding a protein of 947 amino acids).

REFERENCE

1. Honda, H., Saito, T., Iijima, S. and Kobayashi, T. (1988) *Enzyme Microb. Technol.* **10**, 559-562.

```

FILE NAME: RALBGLT.FIG
GAATTCATCAAGGTGTGATGTTGATTATACCTTCGTGAAGTTTGTGAAGAAGCTGCAGGCTCGAAGCCAGGTATGGTGATATTCACACATAATTATACG 100
ATAACCGTAAAGCTGACGAAGAAACAGTACTGGGACTGATGCACAAAAGCTGAATCCCTGTGACGTTTCTTGTCAATTGACCTGACTGTGATACGGAG 200
GTAAGATATGATAAAGCTTGTATGGAACGAATATCTCGAAAAGCGCAGAGGTAACCGCTGAGGGCGCAGTGTCTTGTGAACAACCGGCTCTGGCC 300
M I K L D W N E Y L E K A A E V N A E G A V L L V N N G V L P
CTGGATAAAAATGCCGTTACGCAGGTTTTCGGACGTATACAGCTGGATTATAAAAAGCGGTACGGCTCTGGCGGAATGGTGAATGTTGCAAAAGGTC 400
L D K N A V T Q V F G R I Q L D Y Y K S G T G S G G M V N V A K V T
CGGGAATAACCGATGGCCTTATAGAAGCAGGTGCAAACTCAATGAGGATGTGCTGAAGGCTTACAAGGACTATGTTGCTGAACATCCCTACGATTACGG 500
G I T D G L I E A G A K L N E D V L K A Y K D Y V A E H P Y D Y G
CGAGGGCTGGGGCGGAGCCCTGGTGTCCAGGAGGATGCCCTTGTATGACAGCCTTGTAAAAAGGGCGGCTGAGAGTTCGATACAGCGATATGATT 600
E G W G G E P W C Q E E M P L D D S L V K R A A E S S D T A I C I
ATAGGACGCACCGCAGGCGAGGAACAGGACAACAGCTGCAAGGCGAGTTCCTATCTGCTGACAGAGCGGTAAAAAGGCTATTCTGCGCAAGTAAAGGATA 700
I G R T A G E E Q D N S C K A G S Y L L T D G E K A I L R K V R D N
ATTTCAGCAAAATGGTGACTGCTCAATGTGGCAATATAATCGACATGGGCTTATCGACGAATCTCACCGATGCTAATGATATGATGGCAGGG 800
F S K M V I L L N V G N I I D M G F I D E F S P D A V M Y V W Q G
TGATATGACAGGCGGTACAGGCCTCAAGGGTGTCTGGGTGAGGTATCTCCCTGCGGCAAGCTCCCGGATACTATCGCTTATGATATCACAGACTAT 900
G M T G G T G T A R V L L G E V S P C G K L P D T I A Y D I T D Y
CCCTCTGACAAAAATTTCCACAACAGGGATGGATATCTATGCTGAAGATATCTCTGGGATACAGATACTTTGATACCTTTGCAAAAGCAGGGTAA 1000
P S D K N F H N R D V D I Y A E D I F V G Y R Y F D T F A K D R V R
GATTCCTCCCTCGGATACGGACTTAGCTATACGAGTTGAGATAAGTCCGAGGGCAGAAAGACTGATGACGGTGTGTGATAACCTGCAAAAGGTAAGAA 1100
F P F G Y G L S Y T Q F E I S A E G R K T D D G V V I T A K V K N
TATCGGACGCGGAGGCAAGGAGTCTGTCAGGATACCTTGAAGGCGCAACTGTAAGCTTGGCAAGGCTGCGGTGTGCTTTGCGGATTTGAAAG 1200
I G S A A G K E V V Q V Y L E A P N C K L G K A A R V L C G F E K
ACAAAGGTACTCGCAGCAATGAAGAACAGCAGCTGACGATAGAGTCCAGGCGGTGATATAGCTTCTACGATGACAGCGGCATTACAGGAAATGCCCT 1300
T K V L A P N E E Q T L T I E V T E R D I A S Y D D S G I T G N A F
TCGCATGGTAGAGGAGGAGAGTACACATTCTATGACGAGTGTGATGTCGCGAGTGCAGAGGATGCTTTGCTTTCACACTGGATTCTACCAAGGT 1400
A M W E E A G E Y T F Y A G S D V R S A K E C F A F T L D S T K V
CATCGACAGCTTGAACAGGCACTGGCAGCTTTACGCCTTTCAAGAGGATGGTTCGACCCGAGAGGGTCTTTCCTATGAGGATACCCCTCTTCAAG 1500
I E Q L E Q A L A P V T P F K R M V R T A E G L S Y E D T P L S K
GTTGACGAAGCTGCAGCAGACTTGGATATCTGCTGCGGAAACAGCATATACAGGTGATAAGGGTATAGCCCTTTCCGATGTGGCCATGGTAAGAACA 1600
V D E A A R R L G Y L P A E T A Y T G D K G I A L S D V A H G K N T
CCCTTGTAGTTCATAGCACAGCTTGATGACAATGACCTTAACCTGCTTGTGCGGCGGAGGGTATGTTTCTCCAAAGGTCCTCCCGTACCGGCTGC 1700
L D E F I A Q L D D N D L N C L V R G E G M C S P K V T P G T A A
GGCATTCGGAGGTGAGCAACCTGGAAGAAGTGGCATTACCTGACGGTGTCTTCCGACGGTCTTCCGGAATGCGTCTTGTATGATGATGATGATGAT 1800
A F G G V A K H L E E L G I P A G G C S D G P S G M R L D V G T K
GCATTAGCTGCTGACCTACCTGACCTACCTGCGGCAACATTCACAAGTCTCTTATCACAGGCTTTTACATATCTCGGATGGAGATGCGGGCAATA 1900
A F S L P N G T L I A A T F N K S L I T E L F T Y L G L E M R A N K
AAGTGGATTGTCTGTTGACCGGATGATGAATATCCACCGCACCCCTCAAGGAGTCTTATCACAGGCTTTTACAGGCTTTTACAGGCTTTTACAGG 2000
V D C L L G P G M N I H R H P L N G R N F E Y F S E D P F L T G T
AATGGCTGCGGCAGAACTTGAAGTCTGCACAGGCTGGGCGTTGAGGTTACTATAAAACATTTTTCGCGCAACAATCAGGAGACCAACAGACATTTTATC 2100
H A A A E L E G L H S V G V E G T I K K H F C A N N Q E T N R H F I
GACTCCGTAGCAAGTAGAGAGCGCTCCGTGATATACCTCAAAGGATTCAGATAGCTTACGAAAGCAAGGACGTTCTGTTATGACGACCTACG 2200
D S V A S E R A L R E I Y L K G F E I A V R K S K A R S V M T T Y G
GTAAGGTAAACGGTCTCGGACAGCAGGCGAGCTTACCTTAATACTATGACTCCGCAAGCAGTGGGGCTTTGACGGCTTTACCATGACTGACTGGT 2300
K V N G L W T A G S F D L N T M I L R K Q W H F D G F T M T D W H
GGCAATNATCAATGACCGGGCTGTGCTCCTGACAAGAACAATTTTCGGGCTATGGTACGTCGCGAGAATGATGTTATATGGTATGGCAGACGGTAA 2400
A N I N D R G C A P D K N N F A A M V R A Q N D V Y H C A D G E
AGCGGACGTGACAACTGTTATAGCGGCACTGGCAGACGGCAGACTTACCGGCGCAGAACTTCAGAGGTTGCAAGGAACATACTCTCATTATGATGACA 2500
S G S D N V I A A L A D G R L T R A E L Q R S A R H I L S F M H S T
CTCATGCTATGGCAAGAAAGCTTGGCGAGGACGAGGCTGTGAAGTTATCAACAAGCTGCCGAGACCGTTGATGACGGCGAGGCGCAGAGTTCCT 2600
H A M A R K L G E D E A V E V I N K P A E T V D D G E G D R V F L
GCTGGACGGACCTGACCATAGATAGTGGTGTAAAGCAGGAGAAATCTCGATTACAGCTTCACTGTAGATGGCACAGTTCAGTCCAGTACCGC 2700
L D G D L T I D M S G V K T E R N L D V S F T V D V A Q F G Q Y R
ATGGAAATGACAGCAAGCTCCACAGAGCGAGCTTGCACAGATGCCCGTACCGTATTCAGCATGGGTACTGCGATGGGCGCACTTACATGGAACGGCA 2800
M E M T A S S T Q S E L A Q M P T R S R Y T I F R L H D G E G T F T W N G T
CAGCGGAAAACCGGTGACCTTCGCGTGAAGAATGCCCATGTTTCCCGGATACATATTCAGGCTTCACTTCCGTTGCGGCGGACTGATATGGA 2900
G G K P V T F A V E E M P T C M F G L G G A L L D M D
TAAGATAGTATCAAAAGATAAGACCGCGGAGGCACAGGCTGTGCTGGTGGAGGATATCGAGAGATGGCTTCAACCGCAGACGCTTCCGTTGGTGA 3000
K I V F K K I R P A E A Q V C R L R I S E R W L Q T Q T Y F W L K
GCGAATTTCAAAAGTAAAGCTGCTTCGAGGACCGGCTTTCATGATAAAAGACGGGCGAGAACTTTCCCGTATGCGATATGACGCGACCGGAGGG 3100
A N F Q S K K L L R G R A Y R *
CGCAGTGAAGCAAGACGAAAGCTTATCGATGATAAGCTGTCAACATGAGAATT 3158
    
```