

gtccgtgggagaaggagtgaatccttaacgtggagcagctgttctgctggatgggatagctagttatcaatttctgcaattgaaatgacaaca -721
tctgagcttcaagacaccggggaagtccacctgaatgatcactccatccaggccacatatggttagaccccagcacaagctttaa -631
gcctatgggcattatccgaagcaagcagctgatttggggatgaatgttcgctcagcctcttattttttatggatgatgttctctatgcgt -541
ggacctgaaagatgccattctctctctcggcaggcaagttaatatttaatgaagttacggctacattcttaaaaaagaactcgacaaactc -451
tattttgaatgagtagaagaaacacagcattcttcatcggaatcatgaaatccatagaagaaacaacaagttggacagacttaccatgggga -361
aaagatccgaagaccggagaacgcggcttcgcttggaaatcaactatcgctatgtctaccgagagcacagcagagcagcagcagcagat -271
aacaaggctgacaaggacccccgggaaacaccggagatctcctaataataggattcagcgatgatgcccagactaaatcatgaatggagcc -181
ggccaagcagattagatctacttggcgtcgggctatcgcgaccacgataagagtgttattacacaatttaagggttgctctatctcgagc - 91
atcgtgcttgtgactaatgccaaccctcactgtaataataaagaggaagat**C**ggcagcattcaaaatgagggtgacagacttaatccttc - 1

ATGTTTCTTAGTACACTCTCTCACTGGCGGGCGTTCGTTGCCGGCGCTGCCATCCCCAATGGCCAGACGCTTCTCTCAATGACATTCCT 90
M F L S T L L S L A A V V A G A A A I P N G Q T L S L N D I P
.....
TACTATGTGAGCGGCATTCCTGTGTCAACTTTGCAAGGGTACAATGCCCTCTGCATATGCTGCTTTGACAAAGGGAATAGATTTGGTGCCA 180
Y Y V S G I P V S T L Q G Y N A S A Y A A L T K G I D L V P

TAACTGTCTTCTGTAACCTCTACCACGAACTTGGAGTCGCTGCATCGGACTATGTTGAACGCGATGATGTCTTCCAGCCGGCTTTT 270
L T V I P V T P T T N L E S L L S D Y V E R D D V F Q P A F

CTGCGTGCAGTCTATCTCACAGCTTCCACTGCTGATGACATTGACTCCCAACTGAGCAATTATGCGTCAATTCTCAAGTCTTCCGGCACC 360
L R A V Y L T A S T A D D I D S Q L S N Y A S I L K S S G T

GACGTGCTGCTGGTTGATTGAGAAGTACACACCCTTCGTCAGATTCCACAATCAGACGCGAGTTGACCAAAGAGCTGCCAGTGGGCGCT 450
D V L L V D S E V H T A S S D S T I T A Q L T K E L P S G P
.....
TATTTTGTCTCTTGTATACTGAGAGGTTTGTAGAGCGTACCGGTTGTACCTGACGACAACCTAGCTTTCATTCAAGCAGGAATCAGT 540
Y F V S L Y T G E V F R A Y R L Y P D D N L A F I Q A G I S

GACGAGAAGGAGGTTGCTGCCCCCTACCAGCCGTGACAGAAAACGCGATGACCAAAGACGTTGCCGTGCCCTCACGCTCTATTATACA 630
D E K G G V L P L P A V T E N A M T K D V A V P S R L Y Y T

CCGACCGCAGAAAAGCCATTAGCCGGTCTGAGGTAGGTGTCAAGGATATCTACCACGTTAAAGGCTCAAGACGAGTGGCGGCGAGTCGC 700
P T A E K P L A G L R L G V K D I Y H V K G L K T S G G S R

TCCTATTATTATTATACGGAACCTCAGAATGCTACTGCCCATCTATTAGAGACTGTTGGACTTAGGCGCGGCTTCTTGTGCGTAAACT 810
S Y Y Y L Y G T Q N V T A P S I Q R L L D L G A V F V G K T

GGGACCGTTCAGTTTGTAAACGGTGTGACCTACTGCCGACTGGGTGGATTTCCTACTGTCCATTCAACCcACGCGGAGAAGGATATCAG 900
G T V Q F A N G D R P T A D W V D F H C P F N P R G E G Y Q

GCAcCTAGCGGTTCTCTCCGGCTCAGGTGTGGCTATTGCAGCCTACGACTGGTTGGACCTGCTGCTGGTGTAGTGACACTGGCGGTTCA 990
A P S G S S S G S G V A I A A Y D W L D L A V G S D T G G S

ATCGGTTCCCCAGCTGCAGTTCAAGGATATATGGCAACAGGCCATCTACTGGCGCTATCTCTCTAGATCATGTCTTACCTCTCTCGCG 1080
M R S P A A V Q G I Y G N R P S T G A I S L D H V L P L S P

GCTCTGGATACAGCGGGCGTCTTTGCCCGAAGTGCCTCACTATGGTCCCATACTGTGCAAGCTTGGTATCTCTCATCTCCAGCACAATTT 1170
A L D T A G V F A R S A S L W S H T V Q A W Y P H L Q H N F

ACGTCCTTCCCTCGGCAGCTGCTCCTAGCCGGTGGTGGATGGGATGGTAAAGGCATCAGTCCCAGGCCCATCAGAGTCTTACCACATTC 1260
T S F P R Q L L L A G G G W D G K G I S P E A H Q S L T T F

ACACGTGGGCTTGGAGCATTCCTCGGAACAAACCATACCAATGTCGACGTGTGCGAGCGATGGCTTGACACACTCTCCACCACACCA 1350
T R G L E A F L G T N H T N V D V S Q R W L D T H S P T T P

AGCCTGGAAGAGATGCTCAACCTGACCTATGCCACACTTACTTCTGTGGATCAGTTCAACCACCTAGCCGTCCCTCTCTTTGCTGACTAT 1420
S L E E M L N L T Y A T L T S V D Q F N H L A V P L F A D Y

AAAGCCGTCCACCGCGGTCGTCAGCCTTTCATTAAACCCCGcCCaTtAGCGGTTGGCAGTGGGgCcAGGCGAATGGCGGAAACAcTCG 1530
K A V H R G R Q P F I N P G P L A R W Q W G Q A N G G N T S

TACGATGCAGCTCTGCGCAACATGACTAcTTTCCGAAACTGGTGGGAGAAGTCCGGGTATGGTCAGTCCGATAATGCCTCTTGTCTCCAG 1620
Y D A A L R N M T T F R N W W E K S G Y G Q S D N A S C S R

TCGCTTTTCTGTCAGTGTATTTCGGTCCGACCACTGACTACCGTAACCAATATTATGAGGCGCCACTACACCCCACTGGGATTCTCG 1710
S L F V S V Y S V G T T D Y R N Q Y Y E A P T T P P L G F S

ATCGGACGCATCGCGGTATTAGGTGGAGCACCTGAGGTTGTTGTTCTGTGGGAGAGTCCCCATAAATAGTACTATCTCTTTGCAGACC 1800
I G R I A V L G G A P E V V V P V G E S P Y N S T I S L Q T

GAGTATTTCCCGCTCAGTGTTCGCTGCGATGGCGGAGGATGTGACCATGTTCTGGCTTCTTGGTCTGCGCTTGAGAAGAAGGGC 1890
E Y L P V S V A L Q M A R G C D H V L A S L V A G L E K K G

GTCCTCCGACCTGTGAGTACCAGCTCTCGCTTACTCTAAatgaaaaatcacagatgaagaggcggtatcaagctagggtaaagtacaa 1980
V L R P V S T G S R L Y S *

agtagagtatgtcggttcaacaaatggagcagactagatgtacatatatcatagaaggataatgacaatttctgctatat**A**ctcaa 2070
tatcaacgagagaaaaggaagttcatgctgactgactactgaagataccaagcatggagaggatctcgttaattagaaacatatgttcc 2160
acgagaataatattcaccagcaacagttgtatgaactgcgatgactctgctgagccatgaataaaggatgatggctacggca 2250
tactggtctgacccttggcatccgttcccggtgggtcaggggttttccatccatcggtttactgttgaagtcaccgcatctcg 2340
tcccggtgctgctccttctcagtttctctactgctcaggcaatcggtcaattggataaacttcacaaatttggttgaagcttcccccg 2430
ccgagcaaaagtccaatgcccgcagctatcctgcccactaccaccaaggtcccccttgaccggtgaggttcttgaagataaagcattggga 2520
tctccccgatagctcattgcctgtggagcctaactcaatacacatgacaacccccaatgcgattaccgagctatgacaacgctgt 2610
cgataggccgagagctgtgacgaataactcatgcccctacacattctgttgccttaatacacagctcgggcaaggatcgggatgttccgat 2700

FIG. S2. Nucleotide and deduced amino acid sequences of the peptidoglutaminase-asparaginase from *Aspergillus sojae*. Residue 1 is the initiation methionine, and the asterisk indicates the stop codon. The transcription initiation and polyadenylation sites determined by RACE are marked in bold. Amino acid residues determined by sequencing of the purified AsGahA are underlined. N-terminal amino acids of polypeptide A or C, and polypeptide B are indicated by the solid and hollow triangles, respectively. The signal peptide sequence predicted by Signal-P (<http://www.cbs.dtu.dk/services/SignalP/>) is indicated by a dotted line.