Nucleotide sequence of a gene for indole-3-acetamide hydrolase from *Bradyrhizobium japonicum*

Masami Sekine*, Kazutada Watanabe[1] and Kunihiko Syono[2]

Department of Pure and Applied Sciences, College of Arts and Sciences, The University of Tokyo, Meguro-ku, Tokyo 153 and [1]Department of Biochemistry, Tokyo Metropolitan Institute of Gerontology, Itabashi-ku, Tokyo 171, Japan

Submitted June 27, 1989 EMBL accession no. X15117

We have determined the nucleotide sequence of the bam gene, the genetic determinant of indole-3-acetamide hydrolase, an enzyme that catalyzes the conversion of indole-3-acetamide to indole-3-acetic acid in nitrogen-fixing bacterium Bradyrhizobium japonicum strain J1063 (1). The sequence analysis indicates that the bam locus contains an open reading frame of 465 amino acids which corresponds to a protein with a molecular weight of 50,266 daltons (Fig.1.). High degree of homology were found (Fig.2.) among the central region of the putative products of the bam gene, the iaaH gene from Pseudomonas savastanoi (2), and the tms2 gene from Agrobacterium tumefaciens (3).

```
   1  GAATTCCACG AGCATGGCAA ATGCTCCGAT GTGTGCGCTG TCAGATCGCT AGGTCAAAAG CCCAATCGCC ATGGCACAAT CTTGGCGGTG GTTTGTTCCA
                                                  M  V  R  G     R  H  R     S  R  D     P  Q  R  R     D  L  R  G  R  D
 101  CCATTCAACC AGTATATGGG TCAGGACAGT GGGGCCAAAG GTTCATGCAC GTCTCTCGCG CAGGCCCGTG GCAGAGCGTC GCGGCGACGC GCCATGCAGA
   R  R  S  A     S  R  T     D  A  R     R  Q  S  A     A  E  R     G  C  R     R  S  Q  R     G  S  A     E  G  C     A  C  G  R
 201  AACGCCCCTG TACACGCTCG CGTGATCTCC ATAACCTCCG CCGCGCAATT TTTTGGAGAT GCGGCGTGGC GAAGAACGAG GCGACGAAGA AGAAAAGTGT
   Q  T  G     P  R  A     A  R  S  A     C  C  T     A  C  P     S  R  S  R     R  H  S     T  T  E     G  R  P  H     F  N  G
 301  TGCAAGGAAA GTAGCGAGAC CTCGGTCAAG ACGAGCACCG CCCGCAAGGG GGCTGTTGCG AAGGCGGCCA AGAAGACGGT CAAGAAGCCG GCGCCGCGCA
 401  AGTCGGCGAC GGCGGCGCGT CCGAAAGGTC CGGTCTGGCA ATGGTCCGCG GTCGACACCG CAGCCGCGAT CCGCAACGGC GCGATCTCCG CGGTCGAGAC
                                                  M  V  R  G     R  H  R     S  R  D     P  Q  R  R     D  L  R  G  R  D
 501  CGTCGAAGCG CATCTCGAAC GGATGCGCGC CGTCAATCCG CGGCTGAACG CGGTTGTCGT CGATCTCAGC GAGGAAGCGC TGAAGGCTGC GCATGCGGCC
   R  R  S  A     S  R  T     D  A  R     R  Q  S  A     A  E  R     G  C  R     R  S  Q  R     G  S  A     E  G  C     A  C  G  R
 601  GACAAACAGG CCCAAGGGCG GCGCGCTCGG CCTGCTGCAC GGCGTGCCCA TCACGATCAA GGAGAATGTC GACTACCGAA GGCCGGCCGA ATTTCAACGG
     Q  T  G     P  R  A     A  R  S  A     C  C  T     A  C  P     S  R  S  R     R  H  S     T  T  E     G  R  P  H     F  N  G
 701  CTGCCCGCC AACAAGGATT TTGTCCGCCC GTCGACTCG CCCGTGGTGC ACAATCTCAA GAAAGCCGGC GCGATCTTCA TCGGCCTCAC CAACACGCCG
     V  P  A     N  D  F     V  A  P     S  D  S     P  V  H     N  L  K     K  A  G     A  I  V  I     G  L  T     N  T  P
 801  GAATTCTCCT TCCGCGGCTT CACCGACAAT CCGCTGCACG GGCTGACGGT GAACCCCTGG GACCCGAACA TCACCTGCGG CGGCTCTTCG GGGGCGCCCG
     E  F  S  F     R  G  F     T  D  N     P  L  H  G     L  T  L     N  P  W     D  P  N     I  T  C  G     G  S  F  G     G  A  G
 901  GCTCGGCGGT TGCGGCCGGC ATCGGCACCA TCGCCCATCG CAATGATATC GGCGGCTCGC TGCGCTGGCC GGCGCATTGC AACGGCGTCG CCACCATCAA
     S  A  V     A  A  G     I  G  T     I  A  H  G     H  D  I     G  G  S  L     R  W  P     A  H  C     N  G  V  A     T  I  K
1001  GCCGACCCAG GGCCGCATCC CCGCCTTCAA CGGAAGCGCC ACGGCCGAGC GGCCGATGCT GGCGCATCTG ATGTCGGCGC AAGGCCCGCT CGCCCGTCAC
     P  T  Q     G  R  I  P     A  F  N     G  S  A     T  A  E  R     P  H  L     A  H  L     H  S  A  Q     G  P  L     A  R  H
1101  GTCGGTGACG TCCGCCTCGC GCTCGATGTG ATGAGCCAGG CCGATCCGCG CGATCCCTGG TGGGTACCGG CGCCGCTGGC CGGGCCGAGG CCGAAGGGAC
     V  G  D  V     R  L  A     L  D  V     M  S  Q  A     D  P  R     D  P  W     W  V  P  A     P  L  A     G  P  R     P  K  G  P
1201  CGATCAAGGT CGCTCTGGCC AGGATCCCGG AGGATATGGA CGTCGATCCG TCCGTCCGCG CGGCGCTGCG TCAGGCCGCC GATCACCTGG AGCGTTCCGG
     I  K  V     A  L  A     R  I  P  E     D  H  D     V  D  P     S  V  R  A     A  L  R     Q  A  A     D  H  L  E     R  S  G
1301  CTATCGCGTG ACCGAGGTCG ATGTCCCCGA CATCGACGGG GTCTGGCAGA CCTGGTGCGA CATCATCACC AACGAGACCG TGGTGATGCA GGAGGCCGGC
     Y  R  V     T  E  V  D     V  P  D     I  D  G     V  W  Q  T     W  C  D     I  I  T     N  E  T  V     V  M  Q     E  A  G
1401  ATGCTGAAGG TGACGTCCGA GGACTTCCAC AAGGCCTGGG GTGGCATGAA GACCAAGGCC AATGTGCTCG ATCTCAAGGC CTGGATGCAG GCGACGGCCG
     M  L  K  V     T  S  E     D  F  H     K  A  W  G     G  M  K     T  K  A     N  V  L  D     L  K  A     W  M  Q     A  T  A  A
1501  CGCGCAACGG CCATATCCGC GCCTGGCAAT TGTTCTTCGA GGAGTATCCG GTCGTGCTGG CACCGACCAC GGTGAAGCCG ACGCCGGGAC CGCGGGACGA
     R  N  G     H  I  R     A  W  Q  L     F  F  E     E  Y  P     V  V  L  A     P  T  T     V  K  P     T  P  G  P     R  D  D
1601  CACCGTCAGT GCCGATCGCG TGAAGGAAAT CTTCTGGGGC GAGATCCGCT TCATCTCTGC CATCAACGTG CTGGGCCTGC CGGCCGCAGT GGTGCCGGTG
     T  V  S     A  D  R  V     K  E  I     F  W  G     E  I  R  F     I  S  A     I  N  V     L  G  L  P     G  A  V     V  P  V
1701  ACCCTGCATG ACGGCAAGCC GATCGGCGTG CAGCTCATCG CGGGGCGCTA TCGCGAGGAC CTCGCGCTCG ATGCGGCGGC CGCGATCGAG AAGCGTGCCG
     T  L  H  D     G  K  P     I  G  V     Q  L  I  A     G  R  Y     R  E  D     L  A  L  D     A  A  A     A  I  E     K  R  A  G
1801  GTGTGCTCGC CCACCGGCTC TGGGAGACGA TGGAGTAGGG TACTCTCACC CTCTCCCTTG TGGGAGAGGG TGGGCTGCCG CAAAGCGGCG AGACGGGTGA
     V  L  A     H  R  L     W  E  T  M     E  *
1901  GGGGTCTCTC CCCATAGGCG ATCATTTAGC CGTATGGATA GATACCCCTC ATCCGTTCGC GCCACCTTCT CCCGCAAGGG GAGAAGGGAA AGCACAGTCG
2001  CTGCCGCGTT GACTCGCAGG AATTC
```

Fig.1. Nucleotide sequence of the bam gene region and deduced amino acid sequence of the bam gene product. The proposed ribosome-binding sequence of the bam gene is underlined.

```
bam    110 GAIVIGLTNTPEFSFRGFTDNPLHGLTLNPWDPNITCGGSSGGAGSAVAAGIGTIAHGNDIGGSLRWPAHCNGVATIKPTQGR 192
           **:* *  *  * **     **  *  **  *  ******  *** **  :  * * ***:* **   *:   :** **
iaaH   107 GAVVAGKHNHHELSFGVTSINPHNGTVGHPVAPGYCAGGSSGGSAAAVASGIVPLSVGTDTGGSIRIPAAFCGITGFRPTTGR 189
           **:     *** ****:** *   * * **  *     ****** ****** ::    :***** *:*:*** **; ***** **
tms2   109 GALPGASGNMHLLSFGITSNNYATGAVRHPWNPDLIPGGSSGGVAAAVASRLMLGGIGTDTGASVRLPAALCGVVGFRPTLGR 191
```

Fig.2. Amino acid sequence homology between the central region of the putative products of the bam gene from B. japonicum, the iaaH gene from P. savastanoi, and the tms2 gene from A. tumefaciens. Amino acid residues that are conserved (*) or that belong to the same group (:) between adjacents are indicated.

*To whom correspondence should be addressed at Department of Biochemistry, University of Washington, Seattle, WA 98195, USA

References
1. Sekine, M., Watanabe, K., and Syōno, K. (1989) J. Bacteriol. 171, 1718-1724.
2. Yamada, T., Palm, C.J., Brooks, B., and Kosuge, T. (1985) Proc. Natl. Acad. Sci. USA 82, 6522-6526.
3. Klee, H., Montoya, A., Horodyski, F., Lichtenstein, C., Garfinkel, D., Fuller, S., Flores, C., Peschon, J., Nester, E., and Gordon, M. (1984) Proc. Natl. Acad. Sci. USA 81, 1728-1732.