© 1990 Oxford University Press

Nucleotide sequence between the *fadB* gene and the *rrnA* operon from *Escherichia coli*

Kenji Nakahigashi and Hachiro Inokuchi
Department of Biophysics, Faculty of Science, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Submitted September 24, 1990

EMBL accession no. X54687

We have determined the nucleotide sequence of a portion of clone #9Bll from the subset library (1), which corresponds to a region at about 86 min. on the genetic map of *E. coli*. This sequence fills the gap between the *fadB* gene (2) and the *rrnA* operon (3) which have already been sequenced. Open reading frames of 647 and 421 amino acids were identified. The polypeptide encoded by the longer sequence exhibits 29% homology with the prolidase (X-Pro dipeptidase) from a human source (4) and the proline aminopeptidase from *E. coli* (5), if the extra C-terminal region is excluded. The pepQ gene for X-Pro dipeptidase was mapped at a corresponding position on the chromosome of *Salmonella typhimurium* and this ORF appears to be the pepQ gene of *E. coli*. We have no information about the ORF of 421 amino acids.

ACKNOWLEDGMENT

This research was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Kohara,Y., Akiyama,K. and Isono,K. (1987) *Cell* **50**, 495–508.
2. Nakahigashi,K. and Inokuchi,H. (1990) *Nucl. Acids Res.* **18**, 4937.
3. de Boer,H.A., Gilbert,S.F. and Nomura,M. (1979) *Cell* **17**, 201–209.
4. Endo,F., Tanoue,A., Nakai,H., Hata,A., Indo,Y., Titani,K. and Matsuda,I. (1989) *J. Biol. Chem.* **264**, 4476–4481.
5. Yoshimoto,T., Tone,H., Honda,T., Osatomi,K., Kobayashi,R. and Tsuru,D. (1989) *J. Biochem.* **105**, 412–416.

```
GAACGATGTGCTAAGATGCGGAGACTTAAGGTCAAAAAAACAGAAGGGTAAAAAATGGAATCACTGGCCTCGCTCTATAAAAATCATATAGCTACCTTACAAGAACGGACTCGCGATGCGCTGGCGCGCTTCAAGCTGGATGCGTTACTT
                                                  pepQ    M  E  S  L  A  S  L  Y  K  N  H  I  A  T  L  Q  E  R  T  R  D  A  L  A  R  F  K  L  D  A  L  L
ATTCACTCCGGCGAGCTGTTCAATGTTTTTCTCGACGATCATCCCTATCCGTTTAAAGTGAACCCGCAATTCAAAGCGTGGGTGCCGGTAACTCAGGTGCCAAACTGCTGGTTGCTGGTGGATGGCGTGAACAAGCCGAAACTGTGGTTC
I  H  S  G  E  L  F  N  V  F  L  D  D  H  P  Y  P  F  K  V  N  P  Q  F  K  A  W  V  P  V  T  Q  V  P  N  C  W  L  L  V  D  G  V  N  K  P  K  L  W  F
TATCTGCCGGTTGATTACTGGCACAACGTCGAACCGCTGCCGACCTCTTTCTGGACTGAAGATGTGGAAGTGATCGCGCTGCCGAAAGCCGATGGCATTGGTAGCCTGCTGCCTGCTGCGCGCGGCAATATCGGTTATATCGGTCCGGTG
Y  L  P  V  D  Y  W  H  N  V  E  P  L  P  T  S  F  W  T  E  D  V  E  V  I  A  L  P  K  A  D  G  I  G  S  L  L  P  A  A  R  G  N  I  G  Y  I  G  P  V
CCGGAACGTGCGCTGCAACTGGGTATTGAGGCCAGCAACATCAACCCGAAAGGGGTTATCGACTACCTGCATTACTATCGCTCCTTCAAAACTGAGTACGAACTGGCCTGTATGCGTGAAGCGCAGAAAATGGCGGTCAACGGTCACCGC
P  E  R  A  L  Q  L  G  I  E  A  S  N  I  N  P  K  G  V  I  D  Y  L  H  Y  Y  R  S  F  K  T  E  Y  E  L  A  C  M  R  E  A  Q  K  M  A  V  N  G  H  R
GCGGCAGAAGAAGCGTTCCGTTCTGGCATGAGCGAGTTCGATATCAACATTGCCTATCTGACCGCGACCGGTCATCGTGATACCGACGTACCTTACAGCAACATTGTGGCGCTTAACGAACACGCTGCGGTACTGCATTACACCAAACTG
A  A  E  E  A  F  R  S  G  M  S  E  F  D  I  N  I  A  Y  L  T  A  T  G  H  R  D  T  D  V  P  Y  S  N  I  V  A  L  N  E  H  A  A  V  L  H  Y  T  K  L
GACCATCAGGCACCGGAAGAGATGCGCAGCTTCCTGCTGGATGCCGGGGCAGAATATAACGGCTATGCGGCTGACCTGACCCGTACCTGGTCGGCAAAAAGTGACAACGACTACGCGCAGCTGGTGAAGGACGTTAATGATGAACAACTG
D  H  Q  A  P  E  E  M  R  S  F  L  L  D  A  G  A  E  Y  N  G  Y  A  A  D  L  T  R  T  W  S  A  K  S  D  N  D  Y  A  Q  L  V  K  D  V  N  D  E  Q  L
GCGCTGATCGCCACCATGAAAGCAGGCGTCAGCTATGTGGATTACCACATCCAGTTCCATCAGCGCATCGCCAAACTGCTGCGTAAACATCAAATCATCACCGATATGAGTGAAGAAGCGATGGTCGAAAACGATCTTACCGGGCCGTT
A  L  I  A  T  M  K  A  G  V  S  Y  V  D  Y  H  I  Q  F  H  Q  R  I  A  K  L  L  R  K  H  Q  I  I  T  D  M  S  E  E  A  M  V  E  N  D  L  T  G  P  F
ATGCCGCATGGTATCGGCCATCCGCTGGGCCTGCAGGTGCATGACGTCGCTGGTTTTATGCAGGATGATAGCGGTACGCACCTCGCGGCACCGGCAAAATATCCGTACCTGCGCTGCACCCGTATTCTCCAGCCGGGCATGGTGTTAACC
M  P  H  G  I  G  H  P  L  G  L  Q  V  H  D  V  A  G  F  M  Q  D  D  S  G  T  H  L  A  A  P  A  K  Y  P  Y  L  R  C  T  R  I  L  Q  P  G  M  V  L  T
ATCGAACCGGGTATCTACTTCATCGAATCGCTACTGGCACCGTGGCGTGAAGGGCAGTTCAGCAAGCACTTCAACTGGCAGAAAATTGAAGCACTGAAACCGTTCGGCGGCATTCGTATCGAAGACAACGTGGTGATCCACGAAAACAAC
I  E  P  G  I  Y  F  I  E  S  L  L  A  P  W  R  E  G  Q  F  S  K  H  F  N  W  Q  K  I  E  A  L  K  P  F  G  G  I  R  I  E  D  N  V  V  I  H  E  N  N
GTGGAAAACATGACCCGGGAAACTGGCGTGATGGAAAGCTGGTTAATTCCTGCGGCACCGGTCACGGTCGTTGAAGAGATCAAAAAGAGCCGCTTCATAACGATGTTGGCGCATACCGATGGCGTTGAGGCGGCGAAAGCGTTTGTTGAA
V  E  N  M  T  R  E  T  G  V  M  E  S  W  L  I  P  A  A  P  V  T  V  V  E  E  I  K  K  S  R  F  I  T  M  L  A  H  T  D  G  V  E  A  A  K  A  F  V  E
TCGGTGCGGGCAGAACACCCCGATGCCCGGCACCATTGCGTGGCGTGGGTCGCGGGTGCGCCGGATGATTCTCAACAGCTGGGTTTCTCTGACGACGGGGAGCCGGCGGGAACGGCAGGTAAACCGATGCTCGCCCAGCTAATGGGCAGC
S  V  R  A  E  H  P  D  A  R  H  H  C  V  A  W  V  A  G  A  P  D  D  S  Q  Q  L  G  F  S  D  D  G  E  P  A  G  T  A  G  K  P  M  L  A  Q  L  M  G  S
GGCGTCGGGGAAATTACCGCTGTGGTAGTGCGCTACTACGGCGGCATATTGCTTGGCACCGGTGGGTTAGTGAAAGCGTATGGCGGCGGCGTGAATCAGGCCGTGCGCCAGCTAACGACCCAACGCAAGACGCCATTAACCGAATATACT
G  V  G  E  I  T  A  V  V  V  R  Y  Y  G  G  I  L  L  G  T  G  G  L  V  K  A  Y  G  G  G  V  N  Q  A  V  R  Q  L  T  T  Q  R  K  T  P  L  T  E  Y  T
TTGCAATGTGAATATCATCAGTTAACCGGCATTGAAGCGTTGCTGGGGCAGTGTGACGGCAAAATTATCAACAGTGATTATCAGGCATTCGTTCTGCTGCGGGTGGCGCTTCCGGCGGCGAAAGTGGCTGAATTTTCCGCAAAGCTGGCG
L  Q  C  E  Y  H  Q  L  T  G  I  E  A  L  L  G  Q  C  D  G  K  I  I  N  S  D  Y  Q  A  F  V  L  L  R  V  A  L  P  A  A  K  V  A  E  F  S  A  K  L  A
GATTTTAGCCGTGGTTCATTGCAATTGTTAGCGATTGAAGAATAATCCCACTTCGTTTTGCAGAACTAAGGAAGCGGCAGAGATGCATTTTCGCGCCATTACCCGAATCGTTGGACTACTGGTCATCTTATTTTCAGGGACCATGATTAT
D  F  S  R  G  S  L  Q  L  L  A  I  E  E  *                                ORF421   M  H  F  R  A  I  T  R  I  V  G  L  L  V  I  L  F  S  G  T  M  I  I
CCCTGGGCTGGTAGCACTCATCTACCGGGATGGAGCGGGCCGCGCTTTTACCCAGACCTTTTTTGTCGCCCTCGCCATTGGCTCTATGCTGTGGTGGCCGAACCGCAAAGAGAAAGGCGAACTTAAATCCCGTGAGGGGTTTCTGATAGT
P  G  L  V  A  L  I  Y  R  D  G  A  G  R  A  F  T  Q  T  F  F  V  A  L  A  I  G  S  M  L  W  W  P  N  R  K  E  K  G  E  L  K  S  R  E  G  F  L  I  V
GGTGCTGTTCTGGACCGTGCTGGGTAGCGTCGGTGCGCTCCCTTTTATCTTCTCGGAAAGCCCGAACCTCACGATTGCGCTTTTTTGAATCTTTCTCTGGCCTGACCACTACGGGAGCCACTACGCTGGTGGGGCTGGATTCGCT
V  L  F  W  T  V  L  G  S  V  G  A  L  P  F  I  F  S  E  S  P  N  L  T  I  T  D  A  F  F  E  S  F  S  G  L  T  T  T  G  A  T  T  L  V  G  L  D  S  L
CCCTCACGCCATCCTCTTTTATCGCCAGATGCTGCAATGGTTTGGCGGGATGGGGATCATCATCGTGTTGGCGGTTGCGATACTGCCTATCCTCGGCGTGGGTGGGATGCAGCTCTATCGCGCAGAAATGCCCGGCCCGCTGAAAGATAA
P  H  A  I  L  F  Y  R  Q  M  L  Q  W  F  G  G  M  G  I  I  I  V  L  A  V  A  I  L  P  I  L  G  V  G  G  M  Q  L  Y  R  A  E  M  P  G  P  L  K  D  N
CAAAATGCGCCCGCGAATTGCGGAAACGGCGAAAACCCTGTGGTTGATTTATGTCTTGCTGACCGTCGCCTCTGCGCTGGCGTTGTGGTTTGCCGGAATGGATGCCTTTGATGCCATCGGCCATAGCTTTGCGACTATCGCTATTGGCGG
K  M  R  P  R  I  A  E  T  A  K  T  L  W  L  I  Y  V  L  L  T  V  A  S  A  L  A  L  W  F  A  G  M  D  A  F  D  A  I  G  H  S  F  A  T  I  A  I  G  G
CTTCTCGACACATGATGCCAGTATCGGTTATTTCGATAGCCCGACTATTAACACTATCATTGCTATCTTCCTGCTGATCTCCGGCTGTAACTACGGTCTGCACTTTTCACTGTTAAGTGGGCGTAGTCTGAAGGTTTATTGGCGCGATCC
F  S  T  H  D  A  S  I  G  Y  F  D  S  P  T  I  N  T  I  I  A  I  F  L  L  I  S  G  C  N  Y  G  L  H  F  S  L  L  S  G  R  S  L  K  V  Y  W  R  D  P
GGAATTTCGCATGTTTATCGGCGTACAGTTTACGCTGGTGGTTATTTGTACCCTCGTACTGTGGTTTCATAATGTCTACAGTTCGGCGCTGATGACAATTAACCAGGCGTTTTTCCAGGTGGTGTCGATGGCGACAACCGCCGGGTTTAC
E  F  R  M  F  I  G  V  Q  F  T  L  V  V  I  C  T  L  V  L  W  F  H  N  V  Y  S  S  A  L  M  T  I  N  Q  A  F  F  Q  V  V  S  M  A  T  T  A  G  F  T
AACTGACAGCATTGCCCGCTGGCCGCTCTTTTTGCCGGTACTGCTTTTATGTTCAGCCTTTATCGGCGGTTGCGCCGGGTCAACGGGCGGTGGCCTGAAGGTGATCCGCATCCTGCTGCTGTTTAAGCAGGGGAACCGTGAGCTGAAACG
T  D  S  I  A  R  W  P  L  F  L  P  V  L  L  L  C  S  A  F  I  G  G  C  A  G  S  T  G  G  G  L  K  V  I  R  I  L  L  L  F  K  Q  G  N  R  E  L  K  R
ACTGGTGCATCCGAACGCCGTCTATAGCATTAAGCTGGGGAATCGCGCACTGCCGGAACGTATCCTCGAAGCCGTTTGGGGATTTTTCTCCGCCTATGCATTGGTGTTTATTGTCAGTATGCTGGCGATTATCGCACGGGCTGGATGACT
L  V  H  P  N  A  V  Y  S  I  K  L  G  N  R  A  L  P  E  R  I  L  E  A  V  W  G  F  F  S  A  Y  A  L  V  F  I  V  S  M  L  A  I  I  A  R  A  G  *
TTTCTGCCTTTGCGTCGGTTGTTGCGACATTGAATAACCTGGGGCCAGGGCTTGGCGTGGTTGCTGATAACTTTACCAGTATGAACCCGGTGGCTAAATGGATCCTGATTGCCAACATGCTGTTTGGTCGTCTCGAGGTCTTTACATTGC
TGGTGCTCTTTACCCCGACTTTCTGGCGTGAATGATGGAGTAATACGTGAAAACATTAATTCTTTTCTCAACAAGGGACGGACAAACGCGCGAGATTGCCTCCTACCTGGCTTCGGAACTGAAAGAACTGGGGATCCAGGCGGATGTCGC
CAATGTGCACCGCATTGAAGAACCACAGTGGGAAAACTATGACCGTGTGGTCATTGGTGCTTCTATTCGCTATGGTCACTACCATTCAGCGTTCCAGGAATTTGTCAAAAAACATGCGACGCGGCTGAATTCGATGCCGAGCGCCTTTTA
CTCCGTGAATCTGGTGGCGCGCAAACCGGAGAAGCGTACTCCACAGACCAACAGCTACGCGCGCAAGTTTCTGATGAACTCGCAATGGCGTCCCGATCGCTGCGCGGTCATTGCCGGGGCGCTGCGTTACCCACGTTATCGCTGGTACGA
CCGTTTTATGATCAAGCTGATTATGAAGATGTCAGGCGGTGAAACGGATACGCGCAAAGAAGTTGTCTATCCGATTGGGAGCAGGTGGCGAATTTCGCCCGAGAAATCGCCCATTTAACCGACAAACCGACGCTGAAATAAGCATAAAGA
                                                                                                                                                  16SrRNA   5'aaauaagcauaaaga

ATAAAAAATGCGCGGTCAGAAAATTATTTTAAATTTCCTCTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACAC
auaaaaaaugcgcggucagaaaauuauuuuaaauuuccucuugucaggccggaauaacucccuauaaugcgccaccacugacacggaacaacggcaaacac..
```