© 1990 Oxford University Press

DNA sequence of the *putP* gene from *Salmonella typhimurium* and predicted structure of proline permease

Kenneth Miller and Stanley Maloy*
Department of Microbiology, University of Illinois, Urbana, IL 61801, USA

Submitted April 10, 1990

EMBL accession no. X52573

The complete DNA sequence of the *Salmonella typhimurium putP* gene has been determined and used to deduce the primary structure of proline permease, a Na$^+$/proline transport system (5). The *putP* gene encodes a 502 amino acid protein which is very hydrophobic (70% non-polar residues). Like many other integral membrane proteins, the apparent molecular weight of proline permease estimated from SDS-polyacrylamide gel electrophoresis (33 Kd) (2) is much smaller than predicted from the DNA sequence (54.3 Kd). Hydropathy analysis by the method of Kyte and Doolittle (4) together with turn potential predictions by the method of Chou and Fassman (1), predict that the protein has twelve membrane spanning regions (shown underlined).

The *putP* gene from *S. typhimurium* is very similar to its homolog in *E. coli* (6): the DNA sequences are 83% identical and the amino acid sequences are 94% identical with 4% of the differences conservative amino acid substitutions. The only nonconservative amino acid differences were outside of the predicted membrane spanning regions.

REFERENCES

1. Chou,P. and Fassman,G. (1978) *Adv. Enzymol.* **47**, 45.
2. Hahn,D., Myers,R., Kent,C. and Maloy,S. (1988) *Mol. Gen. Genet.* **213**, 125.
3. Hediger,M., Turk,E. and Wright,E. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5748.
4. Kyte,J. and Doolittle,R. (1982) *J. Mol. Biol.* **157**, 105.
5. Maloy,S. (1990) In Krulwich (ed.), *The Bacteria: Bioengergetics*. Academic Press, San Diego.
6. Nakao,T., Yamato,I. and Anraku,Y. (1987) *Mol. Gen. Genet.* **208**, 70.

```
                                                  M  A  I  S  T  P  M  L  V  T  F
ACAAAACGGCACGCTACGGTAGTGCCAATAAATAAATTTGGAGAACCTTGATGGCTATTAGCACACCGATGTTGGTGACATTC
 C  V  Y  I  F  G  M  I  L  I  G  F  I  A  W  R  S  T  K  N  F  D  D  Y  I  L  G  G
TGTGTCTATATTTTTGGCATGATATTGATTGGGTTTATCGCCTGGCGCTCAACCAAAAACTTTGATGACTATATTCTTGGCGG
 R  S  L  G  P  F  V  T  A  L  S  A  G  A  S  D  M  S  G  W  L  L  M  G  L  P  G
TCGCAGTCTGGGGCCGTTTGTTACGGCTTTATCAGCCGGCGCGTCGGATATGAGCGGCTGGCTGTTAATGGGGCTGCCTGGCG
 A  I  F  L  S  G  I  S  E  S  W  I  A  I  G  L  T  L  G  A  W  I  N  W  K  L  V  A
CTATCTTTCTGTCGGGGATCTCTGAAAGCTGGATCGCCATTGGCCTGACGTTAGGCGCATGGATTAACTGGAAGCTGGTGGCC
 G  R  L  R  V  H  T  E  F  N  N  N  A  L  T  L  P  D  Y  F  T  G  R  F  E  D  K  S
GGGCGCCTGCGGGTGCATACCGAATTTAACAATAACGCGCTCACGCTGCCGGACTATTTTACCGGTCGGTTTGAGGATAAGAG
 R  V  L  R  I  I  S  A  L  V  I  L  L  F  F  T  I  Y  C  G  S  G  I  V  A  G  A
CCGAGTCCTGCGTATTATTTCCGCGCTGGTCATTCTGTTGTTTTTCACTATCTATTGCGGGTCAGGTATTGTCGCTGGGGCAC
 R  L  F  E  S  T  F  G  M  S  Y  E  T  A  L  W  A  G  A  A  A  T  I  I  Y  T  F  I
GACTGTTCGAAAGCACCTTCGGTATGAGCTATGAAACCGCACTGTGGGCGGGGGCCGCGGCAACCATTATTTATACCTTTATC
 G  G  F  L  A  V  S  W  T  D  T  V  Q  A  S  L  M  I  F  A  L  I  L  T  P  V  M  V
GGCGGGTTCCTTGCCGTTAGCTGGACGGATACCGTTCAGGCCAGCCTGATGATTTTTGCGTTAATCCTGACGCCGGTGATGGT
 I  V  G  V  G  G  F  S  E  S  L  E  V  I  K  Q  K  S  I  E  N  V  D  M  L  K  G
TATTGTTGGCGTAGGCGGTTTCAGCGAGTCGCTGGAGGTGATCAAGCAAAAGAGCATCGAGAATGTCGACATGCTCAAGGGGC
 L  N  F  V  A  I  I  S  L  M  G  W  G  L  G  Y  F  G  Q  P  H  I  L  A  R  F  M  A
TGAATTTTGTCGCTATTATTTCCCTGATGGGCTGGGGACTGGGTTACTTCGGTCAGCCGCATATCCTGGCGCGCTTTATGGCG
 A  D  S  H  H  S  I  V  H  A  R  R  I  S  M  T  W  M  I  L  C  L  A  G  A  V  A  V
GCGGACTCCCATCACAGCATTGTTCATGCGCGTCGTATCAGTATGACCTGGATGATTCTGTGTCTGGCGGGCGCGGTGGCGGT
 G  F  F  G  I  A  Y  F  N  N  N  P  A  L  A  G  A  V  N  Q  N  S  E  R  V  F  I
GGGCTTCTTTGGCATTGCGTACTTTAACAATAACCCCGCGCTGGCCGGGGCGGTGAACCAAAACTCAGAACGCGTATTTATTG
 E  L  A  Q  I  L  F  N  P  W  I  A  G  V  L  L  S  A  I  L  A  A  V  M  S  T  L  S
AACTGGCGCAGATCCTGTTTAACCCGTGGATTGCCGGTGTTCTGCTGTCTGCTATCCTGGCGGCGGTGATGTCGACGTTGAGC
 C  Q  M  L  V  C  S  S  A  I  T  E  D  L  Y  K  A  F  L  R  K  S  A  S  Q  Q  E  L
TGTCAGATGCTGGTATGCTCCAGCGCGATTACAGAAGATTTATATAAGGCTTTTCTGCGTAAAAGCGCCAGCCAGCAAGAGCT
 V  W  V  G  R  V  M  V  L  V  V  A  L  I  A  I  A  L  A  A  N  P  D  N  R  V  L
GGTATGGGTAGGGCGAGTGATGGTGCTGGTGGTAGCGCTGATCGCCATTGCGCTGGCGGCGAATCCCGATAACCGTGTGCTGG
 G  L  V  S  Y  A  W  A  G  F  G  A  A  F  G  P  V  V  L  F  S  V  M  W  S  R  M  T
GGCTGGTGAGCTACGCCTGGGCTGGATTCGGCGCGGCATTTGGACCTGTTGTTCTGTTTTCTGTGATGTGGTCGCGTATGACA
 R  N  G  A  L  A  G  M  I  I  G  A  V  T  V  I  V  W  K  Q  Y  G  W  L  D  L  Y  E
CGTAACGGCGCGCTGGCGGGAATGATTATTGGCGCGGTGACGGTTATCGTCTGGAAACAATATGGCTGGCTGGATCTGTATGA
 I  I  P  G  F  I  F  G  S  L  G  I  V  I  F  S  L  L  G  K  A  A  T  A  A  M  Q
GATTATCCCTGGCTTCATTTTCGGCAGCCTGGGGATCGTAATCTTTAGCCTGCTTGGCAAAGCGGCGACAGCAGCGATGCAGG
 E  R  F  A  K  A  D  A  H  Y  H  S  A  P  P  S  K  L  Q  A  E
AACGCTTTGCAAAAGCGGACGCGCATTATCATTCCGCGCCGCCGTCGAAGCTACAGGCGGAATAACCGACATGTCCGATAGCA

TTATTGCCATCGACATATTTTACAGGC
```

---

* To whom correspondence should be addressed