Nucleotide sequence and deduced amino acid sequence of *Mycobacterium leprae* gene showing homology to bacterial *atp* operon

Indira Nath and Suman Laal
Department of Biotechnology, All India Institute of Medical Sciences, New Delhi 110029, India

Submitted June 27, 1990 EMBL accession no. X53488

The genomic library of *Mycobacterium leprae* cloned in lambda gt11 (1) was screened with pooled sera from lepromatous leprosy patients. A 2.5 kb clone (LSR 2) was sequenced. The deduced amino acid sequence of one of the open reading frames exhibits significant homology with *atp* operon of *Rhodopseudomonas blastica* (2).

```
                                    M  F  E  R  F  T  D  R  A  R  R
   1  GGTGTAGGGAGAGCAGGTAACCGCCGATGTTCGAAAGATTTACCGATCGTGCCCGCAGGG
      V  V  V  L  A  Q  E  E  A  R  M  L  N  H  N  Y  I  G  T  E
  61  TGGTCGTCCTGGCACAGGAAGAGGCCCGGATGCTCAACCATAACTACATCGGCACCGAGE
      H  I  L  L  G  L  I  H  E  G  E  G  V  A  A  K  S  L  D  S
 121  ACATTTTGCTGGGTCTTACTCACGAAGGCAAAGGTGTCGCGGCGAAATCGTTGGATTCGT
      L  G  I  S  L  E  A  V  R  S  Q  V  E  D  I  I  G  Q  G  Q
 181  TGGGGATTTCACTTGAAGCCGTTCGCAGTCAGGTCGAAGATATTATCGGCCAGGGTCAGE
      Q  A  P  S  G  H  I  P  F  T  P  R  A  K  K  V  L  E  L  S
 241  AGGCGCCGTCGGGGCATATCCCGTTTACGCCTCGTGCCAAGAAGGTTCTTGAGCTGAGCT
      L  R  E  A  L  Q  L  G  H  N  Y  I  G  T  E  H  I  L  L  G
 301  TGCGTGAGGCGCTGCAGCTCGGCCATAATTACATCGGCACCGAGCACATTTTGCTGGGCC
      L  I  R  E  G  E  G  V  A  A  Q  V  L  V  K  L  G  A  E  L
 361  TTATCCGGGAGGGAGAGGGTGTGGCCGCCCAGGTGCTGGTCAAGTTGGGCGCCGAGCTGA
      T  R  V  R  Q  Q  V  I  Q  L  L  S  G  Y  Q  G  K  E  A  A
 421  CCCGAGTGCGCCAGCAGGTGATCCAGCTGCTAAGTGGCTACCAGGGCAAAGAAGCTGCTG
      E  A  G  T  G  G  R  G  G  E  S  G  S  P  S  T  S  L  V  L
 481  AAGCGGGCACTGGAGGCCGGGGTGGCGAATCCGGTAGTCCATCTACATCTCTGGTGCTCG
      D  Q  F  G  R  N  L  T  A  A  A  M  E  S  K  L  D  P  V  I
 541  ACCAGTTCGGTCGAAACCTGACGGCTGCTGCGATGGAGAGCAAGCTGGATCCGGTCATTG
      G  R  E  K  E  I  E  R  V  M  Q  V  L  S  R  R  T  K  N  N
 601  GCCGCGAAAAGGAAATCGAGCGGGTGATGCAGGTGCTGAGCCGGCGCACCAAGAACAACC
      P  V  L  I  G  E  P  G  V  G  K  T  A  V  V  E  G  L  A  Q
 661  CGGTGTTGATCGGCGAGCCTGGCGTCGGCAAGACCGCTGTCGTCGAGGGTCTTGCGCAGG
      A  I  V  H  G  E  V  P  E  T  L  K  D  K  Q  L  Y  T  L  D
 721  CCATCGTGCACGGTGAGGTTCCCGAGACGCTGAAAGACAAGCAGCTTTACACACTTGACC
      L  G  S  L  V  A  G  S  R  Y  R  G  D  F  E  E  R  L  K  K
 781  TGGGCTCGTTGGTAGCCGGCTCGCGCTACCGCGGTGACTTCGAGGAGCGCCTGAAGAAGG
      V  L  K  E  I  N  T  R  G  D  I  I  L  S  I  D  E  L  H  T
 841  TACTCAAGGAGATCAACACCCGCGGTGACATAATCTTGTTCATCGATGAGCTGCACACCC
      L  V  G  A  G  A  A  E  G  A  I  D  A  A  S  I  L  K  P  K
 901  TGGTAGGTGCCGGTGCCGCCGAGGGGGCGATCGATGCCGCGTCGATCCTCAAGCCCAAGC
      L  A  R  G  E  L  Q  T  I  G  A  T  T  L  D  E  Y  R  K  Y
 961  TGGCCCGCGGTGAGCTGCAAACCATTGGCGCTACCACCCTCGACGAATATCGCAAGTACA
      I  E  K  D  A  A  L  E  R  R  F  Q  P  V  Q  V  G  E  P  T
1021  TCGAGAAAGACGCCGCCCTGGAGCGCCGCTTCCAGCCGGTGCAGGTGGGTGAGCCAACGG
      V  E  H  T  I  E  I  L  K  G  L  R  D  R  Y  E  A  H  H  R
1081  TGGAGCACACTATCGAGATTCTCAAGGGCCTGCGGGACCGCTACGAGGCCCACCATCGGG
      V  S  I  T  D  S  A  M  V  A  A  A  T  L  A  D  R  Y  I  N
1141  TGTCGATCACCGACTCGGCGATGGTGGCCGCCGCCACTCTGGCCGATCGCTACATCAATG
      D  R  F  L  P  D  K  A  I  D  L  I  D  E  A  G  A  R  M  R
1201  ACCGATTCCTCCCGGACAAGGCGATCGACCTGATTGACGAAGCAGGCGCCCGGATGCGGA
      I  R  R  M  T  A  P  P  D  L  R  E  F  D  E  K  I  A  E  A
1261  TTCGCCGAATGACCGCGCCACCAGATCTGCGTGAGTTCGACGAGAAGATCGCCGAAGCCC
      R  R  E  K  E  S  A  I  D  A  Q  D  F  E  K  A  A  S  L  R
1321  GCCGGGAGAAAGAGTCGGCGATCGACGCCCAGGATTTCGAGAAGGCGGCTAGCCTACGCG
      D  R  E  K  Q  L  V  A  Q  R  A  E  R  E  K  Q  W  R  S  G
1381  ACCGGGAGAAGCAACTGGTAGCTCAGCGTGCAGAACGCGAAAAGCAATGGCGTTCAGGCG
      D  L  D  V  I  A  E  V  D  D  E  Q  I  A  E  V  L  G  N  W
1441  ATCTCGATGTTATTGCGGAAGTTGACGACGAGCAAATTGCCGAGGTGCTGGGCAACTGGA
      T  G  I  P  V  F  K  L  T  E  A  E  T  T  R  L  L  R  M  E
1501  CCGGTATACCGGTGTTCAAGCTCACCGAGGCTGAGACTACCCGGCTGCTCCGTATGGAGG
      E  E  L  H  K  R  I  I  G  Q  E  D  A  V  K  A  V  S  K  A
1561  AGGAGCGTCACAAACGGATTATCGGCCAGGAGGATGCTGTTAAGGCCGTGTCCAAGGCGA
      I  R  R  T  R  A  G  L  K  D  P  K  R  P  S  G  S  F  I  F
1621  TCCGTCGTACCCGCGCCGGGCTAAAAGACCCCAAGCGCCCGTCGGGTTCGTTCATCTTCG
      A  G  P  S  G  V  G  K  T  E  L  S  K  A  L  A  N  F  L  F
1681  CCGGCCCGTCCGGTGTCGGTAAGACTGAACTGTCTAAGGCGCTGGCCAACTTTTTGTTTG
      G  D  D  D  A  L  I  Q  I  D  M  G  E  F  H  D  R  F  T  A
1741  GCGACGACGACGCGCTCATCCAGATCGATATGGGCGAGTTCCACGACCGGTTCACCGCGT
      S  R  L  F  G  A  P  P  G  Y  V  G  Y  E  E  G  G  Q  L  T
1821  CGCGGCTTTTTGGTGCTCCGCCGGGATACGTCGGCTACGAGGAGGGCGGTCAGCTCACCG
      E  K  V  R  R  K  P  F  S  V  V  L  F  D  E  I  E  K  A  H
1881  AGAAGGTGCGGCGCAAGCCGTTCTCGGTGGTGCTGTTCGACGAGATCGAAAAAGCGCATC
      Q  E  I  Y  N  S  L  L  Q  V  L  E  D  G  R  L  E  F
1941  AGGAGATCTATAACAGCCTGTTGCAGGTGCTTGAAGACGGGCGGCTGGAATTC
```

ACKNOWLEDGEMENT

This work was supported by the Department of Science and Technology, Government of India.

REFERENCES

1. Young *et al.* (1985) *Nature* **316**, 450–452.
2. Tybulewicz,V.L.J., Falk,G. and Walker,J.E. (1984) *J. Mol. Biol.* **179**, 185–214.