Nucleotide sequence of a gene from alkaliphilic *Bacillus firmus* RAB that is homologous to the *fpg* gene of *Escherichia coli*

D.Mack Ivey

Department of Biochemistry, Mount Sinai School of Medicine of CUNY, 1 G.Levy Place, New York, NY 10029, USA

Submitted August 28, 1990

EMBL accession no. X53930

A bacteriophage lambda clone that contains the genes for the ATP synthase of *B. firmus* RAB (1) has been shown by subcloning and sequencing to also contain a homologue of the formamidopyrimidine-DNA glycosylase (*fpg*) gene of *E. coli* (2). The *E. coli* enzyme catalyzes the removal of the imidazole ring-opened form of *N*7-methylguanine *in vitro*. The 824 bp gene of the alkaliphile is predicted to encode a protein of 276 amino acid residues ($M_r = 31,400$) with 59% similarity to and 41% identity with the gene product from *E. coli*.

ACKNOWLEDGEMENTS

Thanks are due to Charles Cheng for expert technical assistance. This work was supported by grant GM28454 from the National Institutes of Health.

REFERENCES

1. Ivey, D.M. and Krulwich, T.A. (1990) *Nucl. Acids Res.* **18**, 1296.
2. Boiteux, S., O'Connor, T.R. and Laval, J. (1987) *EMBO J.* **6**, 3177-3183.

```

1  CCGCCATGGTCGAGATGGCTCTTCGTTTGAAGAGGCAACCTAAGAGCAGACTGCTTTGCAAGTCATGATGAACATAATTTGAAG  90
91  TGACCGCCTGAAGAGGTGGACGAGATGATGAAAAAGTTCTCTGAAGTGATGGAGAATGCAGTAGAGCTGGCTGTTCCCTAAAAAGTTGATA  180
181 TATAGCCGGTGATACTTGGTATGATGCGAAGTAAGGAGAGAGGACAACATGCCTGAATTACCTGAAGTTGAAACGGTCAACCGGACGTTA  270
      M P E L P E V E T V K R T L
271 ACAGAATTAGTGATCGGTAAACGATAGCCGGAATAACCGTTAAGTGGGCAATATAATTAAGAGCCGGCGGATGTTTGGAGTTTGAA  360
      T E L V I G K T I A G I T V K W A N I I K E P A D V L E F E
361 ACACCTCTTATGAATCAACGATTCGGAGCATACGAAGAAGAGGGAAGTTTACTTTTGGAGTTCGATGATATCGTAATGGTTTCTCAC  450
      T L L M N Q T I R S I R R R G K F L L F E F D D I V M V S H
451 CTTGGAATGGAAGGTCGTTATGGTCTATATGAAAAAGAAGAGCCGCTGCCTCCTACTCATGTTATTTTCCATTTTACAGATGGGGA  540
      L R M E G R Y G L Y E K E E P L P P H T H V I F H F T D G E
541 GAGCTAAGGTATCAGGATGTGAGGAAGTTTGGCAGATGCATCTATTTCCAAAAGGCTCAGAAGAGAAGGTCCTGCCTTAGCTCATTTA  630
      E L R Y Q D V R K F G T M N L F P K G S E E K V L P L A N L
631 GGGGTAGAGCCATTTTCTGAGCAATTTACCTCAGAGCTTTTAATGAACGCATTTAGAAAACCAATCGGAAATAAAGGTTGCTCTCCTA  720
      G V E P F S E Q F T S E L L M N A F Q K T N R K I K V A L L
721 GACCAAAAACGGTGGTAGGGCTTGGAAACATCTATGTCGATGAAGCATTATTTAGAGCTCGTATTCATCCTGAGAGATTAGCACATTCT  810
      D Q K T V V G L G N I Y V D E A L F R A R I H P E R L A H S
811 TTAAGCAAAGAGGAAATGGCTGTATTGCATAAGGCGATTGTTCCACTTTAGAAGAAGCAGTTGAAATGGGAGGAAGCTCGATTAAAGTCC  900
      L S K E E M A V L H K A I V S T L E E A V E M G G S S I K S
901 TATGTGAATGGTCAAGGTGAAATGGGAATGTTTCAACAGAAATTAGGTGTGTACGGAAGAAAAACGAACCTTGCCGTCAGTGTGGTACG  990
      Y V N G Q G E M G M F Q Q K L G V Y G R K N E P C R Q C G T
991 GACATTTTAAAGACCGTTTGGGAGGCCGAGGCACTATTTTGTCCAAATTGCCAATTGTAACGTACATCGGATAGGCTTTTGCATATA  1080
      D I L K T V V G G R G T H F C P N C Q L *
1081 CTATCGTATCGTGTAGGTGGGAGGAGCTTATCAGCATGGTTGAGATTCTTCATTGCGCTCGCACTTGCTGTGAGCTTAGATAGTTT  1170
1171 CGGTGTAGGGTTAAGTATGGACTTAGGAAA  1200

```