 © *Oxford University Press 1990*

# Sequence of the cDNA encoding ITF-2, a positive-acting transcription factor

**Paula Henthorn, Ruth McCarrick-Walmsley and Tom Kadesch**
Department of Human Genetics and the Howard Hughes Medical Institute, University of Pennsylvania School of Medicine, Philadelphia, PA 19104–6145, USA

Submitted December 18, 1989 Genbank accession no. M30314

Figure 1 shows the nucleotide sequence and corresponding amino acid sequence of cDNA E2–2, encoding the transcription factor ITF-2 (1).

## REFERENCE

1. Henthorn,P., Kiledjian,M. and Kadesch,T. (1990) *Science* in press.

```
         .         .         .         .         .         .         .         .         .
gaattcgggggggtagaagacagaagtagctcagggtcctgggggaatggaggacatccaagcccgtccaggaactatggagatgggact   90
E  F  G  G  V  E  D  R  S  S  S  G  S  W  G  N  G  G  H  P  S  P  S  R  N  Y  G  D  G  T
ccctatgaccacatgaccagcagggaccttgggtcacatgacaatctctctccaccttttgtcaattccagaatacaaagtaaaacagaa  180
P  Y  D  H  M  T  S  R  D  L  G  S  H  D  N  L  S  P  P  F  V  N  S  R  I  Q  S  K  T  E
aggggctcatactcatcttatgggagagaatcaaacttacagggttgccaccagcagagtctccttggaggtgacatggatatgggcaac  270
R  G  S  Y  S  S  Y  G  R  E  S  N  L  Q  G  C  H  Q  Q  S  L  L  G  G  D  M  D  M  G  N
ccaggaacccttttcgcccaccaaacctggttcccagtactatcagtattctagcaataatccccgaaggaggcctcttcacagtagtgcc  360
P  G  T  L  S  P  T  K  P  G  S  Q  Y  Y  Q  Y  S  S  N  N  P  R  R  R  P  L  H  S  S  A
atggaggtacagacaaagaaagttcgaaaagttcctccaggtttgccatcttcagtctatgctccatcagcaagcactgccgactacaat  450
M  E  V  Q  T  K  K  V  R  K  V  P  P  G  L  P  S  S  V  Y  A  P  S  A  S  T  A  D  Y  N
agggactcgccaggctatccttcctccaaaccagcaaccagcactttccctagctccttcttcatgcaagatggccatcacagcagtgac  540
R  D  S  P  G  Y  P  S  S  K  P  A  T  S  T  F  P  S  S  F  F  M  Q  D  G  H  H  S  S  D
ccttggagctcctccagtgggatgaatcagcctggctatgcaggaatgttgggcaactcttctcatattccacagtccagcagctactgt  630
P  W  S  S  S  S  G  M  N  Q  P  G  Y  A  G  M  L  G  N  S  S  H  I  P  Q  S  S  S  Y  C
agcctgcatccacatgaacgtttgagctatccatcacactcctcagcagacatcaattccagtcttcctccgatgtccactttccatcgt  720
S  L  H  P  H  E  R  L  S  Y  P  S  H  S  S  A  D  I  N  S  S  L  P  P  M  S  T  F  H  R
agtggtacaaaccattacagcacctcttcctgtacgcctcctgccaacgggacagacagtataatggcaaatagaggaagcggggcagcc  810
S  G  T  N  H  Y  S  T  S  S  C  T  P  P  A  N  G  T  D  S  I  M  A  N  R  G  S  G  A  A
ggcagctcccagactggagatgctctggggaaagcacttgcttcgatctattcttcagatcacactaacaacagcttttcatcaaaccct  900
G  S  S  Q  T  G  D  A  L  G  K  A  L  A  S  I  Y  S  S  D  H  T  N  N  S  F  S  S  N  P
tcaactcctgttggctctcctccatctctctcagcaggcacagctgtttggtctagaaatggaggacaggcctcatcgtctcctaattat  990
S  T  P  V  G  S  P  P  S  L  S  A  G  T  A  V  W  S  R  N  G  G  Q  A  S  S  S  P  N  Y
gaaggacccttacactctttgcaaagccgaattgaagatcgtttagaaagactggatgatgctattcatgttctccggaaccatgcagtg 1080
E  G  P  L  H  S  L  Q  S  R  I  E  D  R  L  E  R  L  D  D  A  I  H  V  L  R  N  H  A  V
ggcccatccacagctatgcctggtggtcatggggacatgcatggaatcattggaccttctcataatggagccatgggtggtctgggctca 1170
G  P  S  T  A  M  P  G  G  H  G  D  M  H  G  I  I  G  P  S  H  N  G  A  M  G  G  L  G  S
gggtatggaaccggccttctttcagccaacagacattcactcatggtggggacccatcgtgaagatggcgtggccctgagaggcagccat 1260
G  Y  G  T  G  L  L  S  A  N  R  H  S  L  M  V  G  T  H  R  E  D  G  V  A  L  R  G  S  H
tctcttctgccaaaccaggttccggttccacagcttcctgtccagtctgcgacttcccctgacctgaacccaccccaggacccttacaga 1350
S  L  L  P  N  Q  V  P  V  P  Q  L  P  V  Q  S  A  T  S  P  D  L  N  P  P  Q  D  P  Y  R
ggcatgccaccaggactacaggggcagagtgtctcctctggcagctctgagatcaaatccgatgacgagggtgatgagaacctgcaagac 1440
G  M  P  P  G  L  Q  G  Q  S  V  S  S  G  S  S  E  I  K  S  D  D  E  G  D  E  N  L  Q  D
acgaaatcttcggaggacaagaaattagatgacgacaagaaggatatcaaatcaattactagcaataatgacgatgaggacctgacacca 1530
T  K  S  S  E  D  K  K  L  D  D  D  K  K  D  I  K  S  I  T  S  N  N  D  D  E  D  L  T  P
gagcagaaggcagagcgtgagaaggagcggaggatggccaacaatgcccgagagcgtctgcgggtccgtgacatcaacgaggctttcaaa 1620
E  Q  K  A  E  R  E  K  E  R  R  M  A  N  N  A  R  E  R  L  R  V  R  D  I  N  E  A  F  K
gagctcggccgcatggtgcagctccacctcaagagtgacaagccccagaccaagctcctgatcctccaccaggcggtggccgtcatcctc 1710
E  L  G  R  M  V  Q  L  H  L  K  S  D  K  P  Q  T  K  L  L  I  L  H  Q  A  V  A  V  I  L
agtctggagcagcaagtccgagaaaggaatctgaatccgaaagctgcgtgtctgaaaagaagggaggaagagaaggtgtcctcggagcct 1800
S  L  E  Q  Q  V  R  E  R  N  L  N  P  K  A  A  C  L  K  R  R  E  E  E  K  V  S  S  E  P
ccccctctctccttggccggcccacaccctggaatgggagacgcatcgaatcacatgggacagatgtaaaagggtccaagttgccacatt 1890
P  P  L  S  L  A  G  P  H  P  G  M  G  D  A  S  N  H  M  G  Q  M  *
gcttcattaaaacaagagaccacttccttaacagctgtattatcttaaacccacataaacacttctccttaacccccatttttgtaatat 1980
aagacaagtctgagtagttatgaatcgcagacgcaagaggtttcagcattcccaattatcaaaaaacagaaaaacaaaaaaaaaaaaaa 2070
aaaacccgaattc                                                                              2083
```

**Figure 1.** The numbers at the right indicate the nucleotide positions. The flanking EcoR1 recognition sequences are derived from the lambda gt11 cloning vector.