Sequence of the cDNA encoding ITF-1, 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 Pennyslvania School of Medicine, Philadelphia, PA 19104–6145, USA

Submitted December 18, 1989 Genbank accession no. M30313

Figure 1 shows the nucleotide sequence and corresponding amino acid sequence of cDNA E2–5, encoding the transcription factor ITF-1 (1).

REFERENCE

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

```
gaattcgggggaggtgagtgcctggcctggtgcggtccctctgctgtgcacagatgtgcagacgtgggccttggcatggtcagtgcacgc   90
E  F  G  G  G  E  C  L  A  W  C  G  P  S  A  V  H  R  C  A  D  V  G  L  G  M  V  S  A  R
acagcgccaggcaagagcggtgagcggggcgcctatgcctccttcgggagagacgcaggcgtgggcggcctgactcaggctggcttcctg  180
T  A  P  G  K  S  G  E  R  G  A  Y  A  S  F  G  R  D  A  G  V  G  G  L  T  Q  A  G  F  L
tcaggcgagctggccctcaacagccccgggcccctgtccccttcgggcatgaaggggacctcccagtactacccctcctactccggcagc  270
S  G  E  L  A  L  N  S  P  G  P  L  S  P  S  G  M  K  G  T  S  Q  Y  Y  P  S  Y  S  G  S
tcccggcggagagcggcagacggcagcctagacacgcagcccaagaaggtccggaaggtcccgccgggtcttccatcctcggtgtaccca  360
S  R  R  R  A  A  D  G  S  L  D  T  Q  P  K  K  V  R  K  V  P  P  G  L  P  S  S  V  Y  P
cccagctcaggtgaggactacggcagggatgccaccgcctacccgtccgccaagacccccagcagcacctatcccgccccccttctacgtg  450
P  S  S  G  E  D  Y  G  R  D  A  T  A  Y  P  S  A  K  T  P  S  S  T  Y  P  A  P  F  Y  V
gcagatggcagcctgcacccctcagccgagctctggagtcccccgggccaggcgggcttcgggcccatgctgggtgggggctcatccccg  540
A  D  G  S  L  H  P  S  A  E  L  W  S  P  P  G  Q  A  G  F  G  P  M  L  G  G  G  S  S  P
ctgcccctcccgcccggtagcggcccggtgggcagcagtggaagcagcagcacgtttggtggcctgcaccagcacgagcgtatgggctac  630
L  P  L  P  P  G  S  G  P  V  G  S  S  G  S  S  S  T  F  G  G  L  H  Q  H  E  R  M  G  Y
cagctgcatggagcagaggtgaacggtgggctcccatctgcatcctccttctcctcagcccccggagccacgtacggcggcgtctccagc  720
Q  L  H  G  A  E  V  N  G  G  L  P  S  A  S  S  F  S  S  A  P  G  A  T  Y  G  G  V  S  S
cacacgccgcctgtcagcggggccgacagcctcctgggctcccgagggaccacagctggcagctccggggatgccctcggcaaagcactg  810
H  T  P  P  V  S  G  A  D  S  L  L  G  S  R  G  T  T  A  G  S  S  G  D  A  L  G  K  A  L
gcctcgatctactccccggatcactcaagcaataacttctcgtccagcccttctacccccgtgggctccccccagggcctggcaggaacg  900
A  S  I  Y  S  P  D  H  S  S  N  N  F  S  S  S  P  S  T  P  V  G  S  P  Q  G  L  A  G  T
tcacagtggcctcgagcaggagcccccggtgccttatcgcccagctacgacgggggtctccacggcctgcagagtaagatagaagaccac  990
S  Q  W  P  R  A  G  A  P  G  A  L  S  P  S  Y  D  G  G  L  H  G  L  Q  S  K  I  E  D  H
ctggacgaggccatccacgtgctccgcagccacgccgtgggcacagccggcgacatgcacacgctgctgcctggccacggggcgctggcc 1080
L  D  E  A  I  H  V  L  R  S  H  A  V  G  T  A  G  D  M  H  T  L  L  P  G  H  G  A  L  A
tcaggtttcaccggccccatgtcgctgggtgggcggcacgcaggcctggttggaggcagccaccccgaggacggcctcgcaggcagcacc 1170
S  G  F  T  G  P  M  S  L  G  G  R  H  A  G  L  V  G  G  S  H  P  E  D  G  L  A  G  S  T
agcctcatgcacaaccacgcggccctccccagccagccaggcaccctccctgacctgtctcggcctcccgactcctacagtgggctaggg 1260
S  L  M  H  N  H  A  A  L  P  S  Q  P  G  T  L  P  D  L  S  R  P  P  D  S  Y  S  G  L  G
cgagcaggtgccacggcggccgccagcgagatcaagcgggaggagaaggaggacgaggagaacacgtcagcggctgaccactcggaggag 1350
R  A  G  A  T  A  A  A  S  E  I  K  R  E  E  K  E  D  E  E  N  T  S  A  A  D  H  S  E  E
gagaagaaggagctgaaggcccccccgggcccggaccagtacggacgaggtgctgtccctggaggagaaagacctgagggaccgggagagg 1440
E  K  K  E  L  K  A  P  R  A  R  T  S  T  D  E  V  L  S  L  E  E  K  D  L  R  D  R  E  R
cgcatggccaataacgcgcgggagcgggtgcgcgtgcgggatattaacgaggccttccgggagctggggcgcatgtgccagatgcacctc 1530
R  M  A  N  N  A  R  E  R  V  R  V  R  D  I  N  E  A  F  R  E  L  G  R  M  C  Q  M  H  L
aagtcggacaaagcgcagaccaagctgctcatcctgcagcaggccgtgcaggtcatcctggggctggagcagcaggtgcgagagcggaac 1620
K  S  D  K  A  Q  T  K  L  L  I  L  Q  Q  A  V  Q  V  I  L  G  L  E  Q  Q  V  R  E  R  N
ctgaatcccaaagcagcctgtttgaaacggcgagaagaggaaaaggtgtcaggtgtggttggagacccccagatggtgctttcagctccc 1710
L  N  P  K  A  A  C  L  K  R  R  E  E  E  K  V  S  G  V  V  G  D  P  Q  M  V  L  S  A  P
cacccaggcctgagcgaagcccacaaccccgccgggcacatgtgaaagggacccaggctcatgctctgaactctggctgagaggatgctg 1800
H  P  G  L  S  E  A  H  N  P  A  G  H  M  *
ctcaggagccagcacaggacacccccccaccccaccccaccatgtccccattacaccagagggccatcgtgacgtagacaggatgccaggg 1890
gcctgaccagcctcccccaatgctggggagcatccctgggcctggggccacacctgctgccctccctctgtgtggtccaagggcaagagt 1980
ggctggagccgggggactgtgctggtctgagccccacgaaggccttgggctgtgcgtccgaccctgctgcagaaccagcagggtgtcccc 2070
tcgggcccatctgtgtcccatgtcccagcacccaggcctctctccaggtctcctttctggtcttttgccatgagggtaaccagctcttc 2160
ccagctggctggggactgtcttgggtttaaaactgcaagtctcctaccctgggatcccatccagttccacacgaactagggcagtggtca 2250
ctgtggcacccaggtgtgggcctggctagctgggggccttcatgtgcccttcatgcccctccctgcattgaggccttgtggacccctggg 2340
ctggctgtgttcatccccgctgcaggtcgggcgtctccccccgtgccactcctgagactcccaccgttacccccaggagatcctggactg 2430
cctgactcccctccccagactggcttgggagcctgggccccatggtagatgcaagggaaacctcaaggccagctcaatgcctggtatctg 2520
cccccagtccaggccaggcggagggggaggggctgtccggctgcctctcccttctcggtggcttcccctgcgccctgggagtttgatctct 2610
taagggaacttgcctctccctcttgttttgctcctggccctgcccctaggtctgggtgggcagtggccccatagcctctggaactgtgcg 2700
ttctgcatagaattc                                                                              2715
```

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