cDNA sequence of the human cellular early growth response gene Egr-1

S.V.Suggs, J.L.Katzowitz, C.Tsai-Morris[1] and V.P.Sukhatme*[1]
Amgen, Thousand Oaks, CA 91320 and [1]The University of Chicago, Chicago, IL 60637, USA

Submitted April 13, 1990

EMBL accession no. X52541

The Egr family of cellular proteins is composed of four distinct members which contain closely related Cys_2-His_2 zinc finger motifs. Egr-1 (1, 2) [zif 268 (3), NGFIA (4), Krox 24 (5) and TIS 8(6)] is induced by diverse signals that initiate mitogenesis, differentiation and by neuronal excitation (2, 3, 4, 6). Egr-1 activates transcription via the sequence CGCCCCCGC (7, 8, unpublished observations), thus suggesting that it may play a broad role as a nuclear signal transducer. The human Egr-1 gene is located on chromosome 5q23−31 (2). Since interstitial deletions in this area occur frequently in patients with therapy related acute myelocytic leukemia (t-AML) (2) and since Egr-1 levels are high in differentiated myeloid cells (unpublished observations), it is possible that Egr-1 has a role in myeloid differentiation. We have therefore isolated and fully sequenced the human Egr-1 cDNA from a human 303 fibroblast library obtained by cross-hybridization to the murine Egr-1 clone OC3.1 (2). The overall nucleotide similarity between mouse and human sequences is 87% and 94% at the nucleotide and protein levels, respectively. The protein is particularly rich in proline, serine, and threonine residues that have been found in abundance in other transcription factors. Interestingly, the zinc finger domain (a.a. 340−419) critical in DNA binding is identical in mouse (2, 3, 5), rat (4), chicken (9) and man.

REFERENCES

1. Sukhatme,V.P. *et al.* (1987) *Oncogene Res.* **1**, 343−355.
2. Sukhatme,V.P. *et al.* (1988) *Cell* **53**, 37−43.
3. Christy,B.A. *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7857−7861.
4. Milbrandt,J. (1987) *Science* **237**, 797−799.
5. Lemaire,P. *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4691−4695.
6. Lim,R.W. *et al.* (1987) *Oncogene* **1**, 263−270.
7. Christy,B.A. and Nathans,D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8737−8741.
8. Cao,X. *et al.* (1990) *Mol. Cell. Biol.* In press.
9. Simmons,D., personal communication.

```
CCGCAGAACTTGGGGAGCCGCCGCCGCCATCCGCCGCCGCAGCCAGCTTCCGCCGCCGCAGGACCGGCCCCTGCCCCAGCCTCCGCAGCCGCGGCGCGTCCACGCCCGCCCGCGCCCAGG   120
GCGAGTCGGGGTCGCCGCCTGCACGCTTCTCAGTGTTCCCCGCGCCCCGCATGTAACCCGGCCAGGCCCCCGCAACGGTGTCCCCTGCAGCTCCAGCCCCGGGCTGCACCCCCCCGCCCC   240
                                M  A  A  A  K  A  E  M  Q  L  M  S  P  L  Q  I  S  D  P  F  G  S  F  P  H  S  P  T  M  D    30
GACACCAGCTCTCCAGCCTGCTCGTCCAGGATGGCCGCGGCCAAGGCCGAGATGCAGCTGATGTCCCCGCTGCAGATCTCTGACCCGTTCGGATCCTTTCCTCACTCGCCCACCATGGAC   360
 N  Y  P  K  L  E  E  M  M  L  L  S  N  G  A  P  Q  F  L  G  A  A  G  A  P  E  G  S  G  S  N  S  S  S  S  S  S  G  G  G    70
AACTACCCTAAGCTGGAGGAGATGATGCTGCTGAGCAACGGGGCTCCCCAGTTCCTCGGCGCCGCCGGGGCCCCAGAGGGCAGCGGCAGCAACAGCAGCAGCAGCAGCGGGGGCGGT   480
 G  G  G  G  G  G  S  N  S  S  S  S  S  S  T  F  N  P  Q  A  D  T  G  E  Q  P  Y  E  H  L  T  A  E  S  F  P  D  I  S  L   110
GGAGGCGGCGGGGGCGGCAGCAACAGCAGCAGCAGCAGCAGCACCTTCAACCCTCAGGCGGACACGGGCGAGCAGCCCTACGAGCACCTGACCGCAGAGTCTTTTCCTGACATCTCTCTG   600
 N  N  E  K  V  L  V  E  T  S  Y  P  S  Q  T  T  R  L  P  P  I  T  Y  T  G  R  F  S  L  E  P  A  P  N  S  G  N  T  L  W   150
AACAACGAGAAGGTGCTGGTGGAGACCAGTTACCCCAGCCAAACCACTCGACTGCCCCCCATCACCTATACTGGCCGCTTTTCCCTGGAGCCTGCACCCAACAGTGGCAACACCTTGTGG   720
 P  E  P  L  F  S  L  V  S  G  L  V  S  M  T  N  P  P  A  S  S  S  S  A  P  S  P  A  A  S  S  A  S  A  S  Q  S  P  P  L   190
CCCGAGCCCCTCTTCAGCTTGGTCAGTGGCCTAGTGAGCATGACCAACCCACCGGCCTCCTCGTCCTCAGCACCATCTCCAGCGGCCTCCTCCGCCTCCGCCTCCCAGAGCCCACCCCTG   840
 S  C  A  V  P  S  N  D  S  S  P  I  Y  S  A  A  P  T  F  P  T  P  N  T  D  I  F  P  E  P  Q  S  Q  A  F  P  G  S  A  G   230
AGCTGCGCAGTGCCATCCAACGACAGCAGTCCCATTTACTCAGCGGCACCCACCTTCCCCACGCCGAACACTGACATTTTCCCTGAGCCACAAAGCCAGGCCTTCCCGGGCTCGGCAGGG   960
 T  A  L  Q  Y  P  P  P  A  Y  P  A  A  K  G  G  F  Q  V  P  M  I  P  D  Y  L  F  P  Q  Q  Q  G  D  L  G  L  G  T  P  D   270
ACAGCGCTCCAGTACCCGCCTCCTGCCTACCCTGCCGCCAAGGGTGGCTTCCAGGTTCCCATGATCCCCGACTACCTGTTTCCACAGCAGCAGGGGGATCTGGGCCTGGGCACCCCAGAC  1080
 Q  K  P  F  Q  G  L  E  S  R  T  Q  Q  P  S  L  T  P  L  S  T  I  K  A  F  A  T  Q  S  G  S  Q  D  L  K  A  L  N  T  S   310
CAGAAGCCCTTCCAGGGCCTGGAGAGCCGCACCCAGCAGCCTTCGCTAACCCCTCTGTCTACTATTAAGGCCTTTGCCACTCAGTCGGGCTCCCAGGACCTGAAGGCCCTCAATACCAGC  1200
 Y  Q  S  Q  L  I  K  P  S  R  M  R  K  Y  P  N  R  P  S  K  T  P  P  H  E  R  P  Y  A  C  P  V  E  S  C  D  R  R  F  S   350
TACCAGTCCCAGCTCATCAAACCCAGCCGCATGCGCAAGTATCCCAACCGGCCCAGCAAGACGCCCCCCCACGAACGCCCTTACGCTTGCCCAGTGGAGTCCTGTGATCGCCGCTTCTCC  1320
 R  S  D  E  L  T  R  H  I  R  I  H  T  G  Q  K  P  F  Q  C  R  I  C  M  R  N  F  S  R  S  D  H  L  T  T  H  I  R  T  H   390
CGCTCCGACGAGCTCACCCGCCACATCCGCATCCACACAGGCCAGAAGCCCTTCCAGTGCCGCATCTGCATGCGCAACTTCAGCCGCAGCGACCACCTCACCACCCACATCCGCACCCAC  1440
 T  G  E  K  P  F  A  C  D  I  C  G  R  K  F  A  R  S  D  E  R  K  R  H  T  K  I  H  L  R  Q  K  D  K  K  A  D  K  S  V   430
ACAGGCGAAAAGCCCTTCGCCTGCGACATCTGTGGAAGAAAGTTTGCCAGGAGCGATGAACGCAAGAGGCATACCAAGATCCACTTGCGGCAGAAGGACAAGAAAGCAGACAAAAGTGTT  1560
 V  A  S  S  A  T  S  S  L  S  S  Y  P  S  P  V  A  T  S  Y  P  S  P  V  T  T  S  Y  P  S  P  A  T  T  S  Y  P  S  P  V   470
GTGGCCTCTTCGGCCACCTCCTCTCTCTCTTCCTACCCGTCCCCGGTTGCTACCTCTTACCCGTCCCCGGTTACTACCTCTTATCCATCCCCGGCCACCACCTCATACCCATCCCCTGTG  1680
 P  T  S  F  S  S  P  G  S  S  T  Y  P  S  P  V  H  S  G  F  P  S  P  S  V  A  T  T  Y  S  S  V  P  P  A  F  P  A  Q  V   510
CCCACCTCCTTCTCCTCTCCCGGCTCCTCGACCTACCCATCCCCTGTGCACAGTGGCTTCCCCTCCCCGTCGGTGGCCACCACGTACTCCTCTGTTCCCCCTGCTTTCCCGGCCCAGGTC  1800
 S  S  F  P  S  S  A  V  T  N  S  F  S  A  S  T  G  L  S  D  M  T  A  T  F  S  P  R  T  I  E  I  C                         543
AGCAGCTTCCCTTCCTCAGCTGTCACCAACTCCTTCAGCGCCTCCACAGGGCTTTCGGACATGACAGCAACCTTTTCTCCCAGGACAATTGAAATTTGCTAAAGGGAAAGGGGAAAGAAA  1920
GGGAAAAGGGAGAAAAAGAAACACAAGAGACTTAAAGGACAGGAGGAGGAGATGGCCATAGGAGAGGAGGGTTCCTCTTAGGTCAGATGGAGGTTCTCAGAGCCAAGTCCTCCCTCTCTA  2040
CTGGAGTGGAAGGTCTATTGGCCAACAATCCTTTCTGCCCACTTCCCCTTCCCCAATTACTATTCCCTTTGACTTCAGCTGCCTGAAACAGCCATGTCCAAGTTCTTCACCTCTATCCAA  2160
AGAACTTGATTTGCATGGATTTTGGATAAATCATTTCAGTATCATCTCCATCATATGCCTGACCCCTTGCTCCCTTCAATGCTAGAAAATCGAGTTGGCAAAATGGGGTTTGGGCCCCTC  2280
AGAGCCCTGCCCTGCACCCTTGTACAGTGTCTGTGCCATGGATTTCGTTTTTCTTGGGGTACTCTTGATGTGAAGATAATTTGCATATTCTATTGTATTATTTGGAGTTAGGTCCTCACT  2400
TGGGGGAAAAAAAAAAAAAAAAAGCCAAGCAAACCAATGGTGATCCTCTATTTTGTGATGATGCTGTGACAATAAGTTTGAACCTTTTTTTTTGAAACAGCAGTCCCAGTATTCTCAGAGC  2520
ATGTGTCAGAGTGTTGTTCCGTTAACCTTTTTGTAAATACTGCTTGACCGTACTCTCACATGTGGCAAAATATGGTTTGGTTTTTCTTTTTTTTTTTTGAAAGTGTTTTTTCTTCGTCCT  2640
TTTGGTTTAAAAAGTTTCACGTCTTGGTGCCTTTTGTGTGATGCCCCTTGCTGATGGCTTGACATGTGCAATTGTGAGGGACATGCTCACCTCTAGCCTTAAGGGGGGCAGGGAGTGATG  2760
ATTTGGGGGAGGCTTTGGGAGCAAAATAAGGAAGAGGGCTGAGCTGAGCTTCGGTTCTCCAGAATGTAAGAAAACAAAATCTAAAACAAAATCTGAACTCTCAAAAGTCTATTTTTTTAA  2880
CTGAAAATGTAAATTTATAAATATATTCAGGAGTTGGAATGTTGTAGTTACCTACTGAGTAGGCGGCGATTTTTGTATGTTATGAACATGCAGTTCATTATTTTGTGGTTCTATTTTACT  3000
TTGTACTTGTGTTTGCTTAAACAAAGTGACTGTTTGGCTTATAAACACATTGAATGCGCTTTATTGCCCATGGGATATGTGGTGTATATCCTTCCAAAAAATTAAAACGAAAATAAAGTA  3120
GCTGCGATTGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

* To whom correspondence should be addressed