Volume 16 Number 10 1988 Nucleic Acids Research

Isolation of cDNAs encoding finger proteins and measurement of the corresponding mRNA levels during myeloid terminal differentiation

Antonio Pannuti, Luisa Lanfrancone[1], Anna Pascucci, Pier-Giuseppe Pelicci[1], Girolama La Mantia and Luigi Lania*

Dipartimento di Genetica, Biologia Generale e Molecolare, University of Naples, Via Mezzocannone 8, 80134 Naples and [1]Istituto di Clinica Medica I, University of Perugia, Policlinico Monteluce, 06100 Perugia, Italy

Received March 14, 1988; Accepted April 25, 1988 Accession nos X07289 and X07290

ABSTRACT

The finger motif is a tandemly repeated DNA-binding domain recently identified in the primary structure of several eukaryotic transcriptional regulatory proteins. It was first described for Xenopus TFIIIA, a factor required for transcription of 5S ribosomal genes and subsequently identified in regulatory proteins from other eukaryotic organisms including yeast, Drosophila and mammals. In this paper we report the isolation and characterization of two human cDNA clones both encoding for multifingered protein products. Transcriptional studies indicate that the two finger genes are expressed in a variety of human tissues. Moreover, upon in vitro induced terminal differentiation of human HL-60 and THP-1 myeloid cell lines the expression of both genes is drastically reduced. These data provide support for the existence of a human multigene family coding for regulatory finger proteins which are likely involved in the processes of cell differentiation and/or proliferation.

INTRODUCTION

Regulation of transcription in eukaryotes is mediated by the specific interaction between cellular proteins and regulatory cis-acting DNA sequences. The best characterized protein structure involved in DNA binding is the helix-turn-helix motif found in a variety of prokaryotic regulatory proteins (1), in the yeast mating type proteins (2) and in the homeo domain of several eukaryotic proteins (3). A new DNA binding motif, the zinc-finger motif, has been recently found in many eukaryotic transcription regulatory proteins. This DNA binding structure has been initially identified in the Xenopus RNA polymerase III transcription factor TFIIIA (4,5), and it has been postulated that in the basic structural unit pairs of Cys and His residues bind a Zn(II) ion and the resulting metalloprotein structures can be easily visualized as a finger of amino acids anchored by a Zinc ion. Similar motifs homologous to the fingers in TFIIIA have been subsequently identified in other genes, including Drosophila genes (6,7,8), in yeast regulatory proteins (9), in Xenopus (10) and in mammals (11,12,13,14). The fingers in these proteins have a consensus sequence of $Cys\text{-}X_2\text{-}Cys\text{-}X_3\text{-}Phe\text{-}X_5\text{-}Leu\text{-}X_2\text{-}His\text{-}X_3\text{-}His$, thus these motifs are referred to as Cys/Hys fingers, in addition the tandemly repeated finger units are

separated by a stretch of 7 to 8 amino acids (the H/C link). Isolation and comparison of cDNAs encoding potential finger proteins indicate that in contrast to the Cys/His group, there is another class of finger proteins, (the Cys/Cys group), that contains a variable number of conserved cysteines in appropriate positions for metal chelation (reviewed in 15,16).

Here we report on the isolation and characterization of two human cDNA clones both encoding a multifinger protein containing the Cys/His motif. Transcription studies show that the two human cDNA clones are expressed in a number of cell lines of different embryological derivation, moreover both sequences are expressed in the myeloid cell lines HL-60 and THP-1, and their expression is reduced after terminal differentiation induced in vitro by TPA or DMSO.

MATERIAL AND METHODS

Library screening and DNA sequencing

A λgt11 cDNA library was derived from human placenta (CAT# HL1008 Clontech Laboratories Inc.). Recombinant phages ($\cong$ 1.000.000) were transferred to nitrocellulose filters in duplicate and screened using a set of [^{32}P] end-labelled oligonucleotide CACAC(G/T)GGGGAGAA(G/A)CCCTA probes corresponding to a 20-nt consensus sequence derived from mkr1 finger regions (11). Hybridization to the probes was carried out at 65 °C for 16 hrs. in 5x SSPE (0.9M NaCl, 50mM Na_2HPO_4, pH7.4, 5mM EDTA), 5x Denhardt's solution, 0.1% SDS and 100µg /ml of denaured salmon sperm DNA. The specific activity of the end-labelled oligonuleotides was $\cong 10^8$ c.p.m./µg . Filters were washed at 37 °C in 4x SSPE, 0.2% SDS for 2 hrs., air-dried and autoradiographed for 24 hrs. Positive plaques were re-screened with the oligonucleotide set. DNAs from positive clones were digested with EcoR1 and subcloned in pGEM-3 plasmid for partial DNA sequencing and two clones (cHF.10 and cHF.12) were found to contain finger regions. Appropriate fragments of cHF.10 and cHF.12 were subcloned in pGEM-3 plasmid (Promega, Biotech) for restriction enzyme mapping and DNA sequencing. DNA sequencing was performed on both strands and all DNA sequences were independently determined by two different methods using the GemSeq K/RT system as recommended by Promega Biotech. Nucleotide and amino acid sequences were analyzed using the MicroGenie Sequence Analysis Program software (17).

Cells

The following human cell lines were used: THP-1 (acute promonocytic leukaemia); MOLT-4 (T-cell lymphoma); HL-60 (acute promyelocytic leukaemia); NT2/D1 (teratocarcinoma); CRL1220 (normal fibroblasts); AV3 (Amnion); HTB10 (neuroblastoma); HTB144 (chorion carcinoma); CLL220.1 (colon carcinoma); HTB96 (osteogenic sarcoma); HTB19 (breast cancer); HeLa (cervical carcinoma); H5 (EBV infected B lymphocytes). Cell lines were obtained from the American Type Cell Collection.

$5x10^5$/ml HL-60 cells were induced to differentiate to granulocytes by the addition of 1.25% DMSO to the culture medium (RPMI 1640 + 10% foetal calf serum). Degree of differentiation was judged by the percentage of cells able to reduce the dye nitro blue tetrazolium (NBT), which is a histochemical marker of more mature myeloid cells (18) (85% positive cells at day 5 of culture). $1x10^6$/ml THP-1 cells were induced to differentiate along the

monocytic pathways by the addition of $5x10^{-8}$ M TPA to the culture medium (RPMI 1640 + 10% foetal calf serum). The following features of macrophage differentiation were analyzed: 1) capacity to adhere to plastic dishes (90% adherent cells at day 3 of culture); 2) phagocytosis (45% phagocytic cells at day 3 of culture); 3) appearance of α-naphtyl (non-specific) esterase (19).

RNA isolation and Northern analysis.

Total RNA was extracted from the various cell lines by the guanidine isothiocyanate method (20). Fifteen micrograms aliquots were electrophoresed on 1% agarose-formaldheyde gels, transferred to nitrocellulose filters. The blots were hybridized with the cHF.10 or cHF.12 cDNA clones under high stringency conditions (50% formamide, 5x SSPE, 1x Denhardt's, 0.1% SDS and 100 µg/ml salmon sperm DNA). The final washing was at 60°C in 0,2x SSC and 0.5% SDS. The DNA probes were labelled by the multiprime labelling kit (Amersham) to a specific activity of $5x10^{8}$ c.p.m./µg . The filters were autoradiographed at -70 C with an intensifying screen for 1 week.

RESULTS

Isolation of human cDNA clones containing Finger regions

The finger motif for DNA binding of TFIIIA is based on a tetrahedrically co-ordinated zinc ion that allows the folding of tandemly repeated DNA binding structures (4,5,15). The finger loops are connected by a short stretch of amino acids (H/C-link) and it appears that the H/C-links rather than the finger loops are the evolutionary most conserved portions of the finger domain (21). We selected a consensus sequence from the H/C-link regions of the mouse mkr1 finger gene (11) and used the corresponding 20-mer CACAC(G/T)GGGGAGAA(G/A)CCCTA oligonucleotides as probes for low stringency screening of a human placenta cDNA library. In a group of six independently isolated and partially sequenced cDNA clones, two clones cHF.10 and cHF.12 revealed the presence of a finger structure (data not shown). The complete nucleotide and predicted amino acid sequences of cHF.10 and cHF.12 clones are shown in Fig.1. Translation of the longest open reading frame in both cHF.10 and cHF.12 cDNAs yields multifingered proteins containing 10 and 6 finger repeats, respectively (Fig.1). The two clones hybridize to unique fragments on Southern blots of genomic DNAs digested with several restriction endonucleases (data not shown) and identify different transcripts on Northern blots of human RNAs (see below). These findings suggest that the two cDNAs are derived from two different transcripts.

The cHF.10 clone is 1720 nucleotides long and it contains a long open reading frame that begins at position 181 and extends to an in-frame TAA termination codon at position 1687. Two in-frame ATG initiation codons are found at positions 190 and 214. Only the ATG codon at position 214 is surrounded by sequences which are in good agreement with the consensus (CCG/ACCATGG) for eukaryotic initiation of translation (22), suggesting that it represents the translation start site of the cHF.10 protein. The termination codon is followed by a 3' untranslated region of 36 nucleotides, which lacks a polyadenylation signal and poly(A) tail indicating that the cHF.10 does not contains the entire HF.10 mRNA. The reading frame opened by the ATG at position 214 would code for a polypeptide of 491 amino acids with

a calculated relative molecular mass of 55,045 (Mr 55K). This protein appears to be divided into two major domains: the amino terminal region up to amino acid position 215 which is rich in Glu and Lys (24 %) and folds into a globular domain including a putative glycosylation site (23); and the carboxy terminal region that consists of a cluster of 10 repeated finger domains.

The nucleotide sequence of clone cHF.12 is shown in Fig.1B. The longest open reading frame begins at the third nucleotide of the sequence and terminates at position 587 with a TGA termination codon. A 1274 bp long 3' untranslated region contains a non-canonical polyadenylation signal (ATTAAA) located 20bp upstream from the poly(A) tail. Similarly to the cHF.10 clone,

A

```
                                                                                              51
                                         GAATTCGGGCCAGCCTCCTGCTGCAGAGCCAGTGAACTCAGGTCGGGCTTC
                                                                                             170
     TCAGCTGCGCACATAGGCAGTACTCATCTTGGCCCTGGGAAGAAACTCAAGAAGAAGCTTTTGAAACATAAAGCTTGGATGGGGTTTGACCTCTGCAGGGCAGCGCAGCTATAGGAGTT
                                                                                             270
     CCCCTGCTGAGCAGAGAAGATGACTGCAGAATTGAGAGAAGCC ATG GCC CTA GCC CCA TGG GGC CCA GTG AAG GTG AAA AAG GAG GAG GAA GAA GAA GAA
                                                 Met Ala Leu Ala Pro Trp Gly Pro Val Lys Val Lys Lys Glu Glu Glu Glu Glu Glu
                                                                                             360
     AAC TTC CCA GGT CAG GCA TCC AGC CAA CAA GTG CAC TCC GAG AAC ATC AAA GTC TGG GCC CCA GTG CAG GGT CTT CAG ACA GGC CTT GAT
20   Asn Phe Pro Gly Gln Ala Ser Ser Gln Gln Val His Ser Glu Asn Ile Lys Val Trp Ala Pro Val Gln Gly Leu Gln Thr Gly Leu Asp
                                                                                             450
     GGA TCA GAA GAG GAA GAA AAG GGT CAG AAC ATA TCC TGG GAT ATG GCG GTA GTC CTG AAA GCA ACT CAG GAG GCA CCT GCT GCT TCA ACC
50   Gly Ser Glu Glu Glu Glu Lys Gly Gln Asn Ile Ser Trp Asp Met Ala Val Val Leu Lys Ala Thr Gln Glu Ala Pro Ala Ala Ser Thr
                                                                                             540
     CTT GGC AGC TAC TCA TTA CCA GGG ACT CTG GCC AAG AGT GAG ATA CTG GAG ACT CAT GGG ACC ATG AAC TTT CTA GGT GCT GAA ACC AAG
80   Leu Gly Ser Tyr Ser Leu Pro Gly Thr Leu Ala Lys Ser Glu Ile Leu Glu Thr His Gly Thr Met Asn Phe Leu Gly Ala Glu Thr Lys
                                                                                             630
     AAC CTA CAG TTA CTG GTT CCA AAA ACT GAG ATA TGT GAG GAA GCT GAA AAA CCC CTC ATC ATA TCA GAA AGA ATC CAG AAA GCT GAT CCT
110  Asn Leu Gln Leu Leu Val Pro Lys Thr Glu Ile Cys Glu Glu Ala Glu Lys Pro Leu Ile Ile Ser Glu Arg Ile Gln Lys Ala Asp Pro
                                                                                             720
     CAA GGA CCT GAG TTA GGA GAA GCT TGT GAA AAG GGA AAC ATG TTA AAG AGG CAG AGA ATA AAG AGA GAA AAG AAA GAT TTC AGA CAA GTG
140  Gln Gly Pro Glu Leu Gly Glu Ala Cys Glu Lys Gly Asn Met Leu Lys Arg Gln Arg Ile Lys Arg Glu Lys Lys Asp Phe Arg Gln Val
                                                                                             810
     ATA GTG AAT GAC TGT CAC TTA CCT GAA AGC TTC AAA GAA GAG GAA AAC CAG AAA TGT AAG AAA TCT GGA GGA AAA TAT AGC CTT AAT TCT
170  Ile Val Asn Asp Cys His Leu Pro Glu Ser Phe Lys Glu Glu Glu Asn Gln Lys Cys Lys Lys Ser Gly Gly Lys Tyr Ser Leu Asn Ser
                                                                                             900
     GGC GCT GTT AAA AAT CCA AAA ACC CAG CTT GGA CAA AAG CCT TTT ACG TGT AGC GTG TGT GGG AAA GGA TTT AGT CAG AGT GCA AAC CTC
200  Gly Ala Val Lys Asn Pro Lys Thr Gln Leu Gly Gln Lys Pro Phe Thr Cys Ser Val Cys Gly Lys Gly Phe Ser Gln Ser Ala Asn Leu
                                                                     (1)                                                 990
     GTT GTG CAT CAG CGA ATC CAC ACT GGA GAG AAA CCC TTT GAA TGT CAT GAG TGT GGG AAG GCC TTC ATT CAG AGT GCA AAC CTC GTT GTG
230  Val Val His Gln Arg Ile His Thr Gly Glu Lys Pro Phe Glu Cys His Glu Cys Gly Lys Ala Phe Ile Gln Ser Ala Asn Leu Val Val
                                                         (2)                                                            1080
     CAT CAG AGA ATC CAC ACT GGA CAG AAA CCT TAT GTT TGC TCA AAA TGT GGG AAA GCC TTC ACT CAG AGT TCA AAT CTG ACT GTA CAT CAA
260  His Gln Arg Ile His Thr Gly Gln Lys Pro Tyr Val Cys Ser Lys Cys Gly Lys Ala Phe Thr Gln Ser Ser Asn Leu Thr Val His Gln
                                                     (3)                                                                1170
     AAA ATC CAC TCC TTA GAA AAA ACT TTT AAG TGC AAT GAA TGT GAG AAA GCC TTT AGT TAC AGC TCA CAA CTT GCT CGG CAC CAG AAA GTC
290  Lys Ile His Ser Leu Glu Lys Thr Phe Lys Cys Asn Glu Cys Glu Lys Ala Phe Ser Tyr Ser Ser Gln Leu Ala Arg His Gln Lys Val
                                             (4)                                                                        1260
     CAC ATT ACG GAA AAA TGC TAT GAA TGT AAT GAA TGT GGG AAA ACA TTT ACT AGG AGC TCA AAC CTC ATT GTC CAC CAG AGG ATC CAC ACT
320  His Ile Thr Glu Lys Cys Tyr Glu Cys Asn Glu Cys Gly Lys Thr Phe Thr Arg Ser Ser Asn Leu Ile Val His Gln Arg Ile His Thr
                                     (5)                                                                                1350
     GGG GAG AAG CCC TTT GCC TGT AAC GAC TGT GGC AAA GCC TTT ACC CAG AGT GCA AAT CTT ATT GTA CAT CAG CGA AGC CAT ACT GGT GAG
350  Gly Glu Lys Pro Phe Ala Cys Asn Asp Cys Gly Lys Ala Phe Thr Gln Ser Ala Asn Leu Ile Val His Gln Arg Ser His Thr Gly Glu
                             (6)                                                                                        1440
     AAG CCA TAT GAG TGT AAA GAG TGT GGG AAA GCC TTT AGT TGT TTT TCA CAC CTT ATT GTG CAC CAG AGA ATT CAC AGT GGA GAT CTT CCT
380  Lys Pro Tyr Glu Cys Lys Glu Cys Gly Lys Ala Phe Ser Cys Phe Ser His Leu Ile Val His Gln Arg Ile His Ser Gly Asp Leu Pro
                     (7)                                                                                                1530
     TAC GTG TGT AAT GAA TGT GGG AAG GCC TTC ACA TGT AGC TCA TAC CTA CTT ATT CAT CAG AGA ATT CAT AAT GGA GAA AAA CCT TAC ACA
410  Tyr Val Cys Asn Glu Cys Gly Lys Ala Phe Thr Cys Ser Ser Tyr Leu Leu Ile His Gln Arg Ile His Asn Gly Glu Lys Pro Tyr Thr
             (8)                                                                                                        1620
     TGT AAT GAG TGT GGG AAG GCC TTC AGA CAG AGG TCG AGC CTC ACC GTG CAC CAG AGA ACC CAC ACT GGG GAG AAG CCC TAT GAA TGT GAG
440  Cys Asn Glu Cys Gly Lys Ala Phe Arg Gln Arg Ser Ser Leu Thr Val His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Glu Cys Glu
     (9)                                                                                                          (10)  1710
     AAG TGT GGT GCA GCT TTC ATT TCC AAC TCA CAC CTC ATG CGA CAC CAT AGA ACC CAT CTT GTT GAA TAA CAAGTAAGGAAGAGGAAGACCTCCCCC
470  Lys Cys Gly Ala Ala Phe Ile Ser Asn Ser His Leu Met Arg His His Arg Thr His Leu Val Glu ***

     GCCCGAATTC
```

B

```
                                                                                                                    80
            GA ATT CGG GAT TCG GGT GAA TGT AGT GAT TGT GGG AAA ACC TTC AGC TGT AGC TCT GCC CTC ATT CTG CAT CGG AGG ATC
               Ile Arg Asp Ser Gly Glu Cys Ser Asp Cys Gly Lys Thr Phe Ser Cys Ser Ser Ala Leu Ile Leu His Arg Arg Ile
                                        (1)                                                                         170
CAC ACG GGG GAG AAA CCC TAT GAA TGT AAT GAG TGT GGG AAG ACC TTC AGC TGG AGC TCC ACC CTC ACC CAC CAT CAG AGA ATC CAC ACT
 27 His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Glu Cys Gly Lys Thr Phe Ser Trp Ser Ser Thr Leu Thr His His Gln Arg Ile His Thr
                                  (2)                                                                               260
GGT GAG AAA CCC TAC GCC TGC AAT GAA TGT GGG AAG GCC TTC AGC AGG AGC TCA ACC CTT ATT CAC CAT CAG AGA ATC CAC ACT GGA GAA
 57 Gly Glu Lys Pro Tyr Ala Cys Asn Glu Cys Gly Lys Ala Phe Ser Arg Ser Ser Thr Leu Ile His His Gln Arg Ile His Thr Gly Glu
                           (3)                                                                                      350
AAA CCC TAT GAA TGT AAT GAA TGT GGG AAA GCC TTC AGC CAG AGC TCA CAC CTC TAT CAG CAC CAG AGA ATC CAC ACT GGA GAG AAG CCC
 87 Lys Pro Tyr Glu Cys Asn Glu Cys Gly Lys Ala Phe Ser Gln Ser Ser His Leu Tyr Gln His Gln Arg Ile His Thr Gly Glu Lys Pro
                           (4)                                                                                      440
TAC GAA TGT ATG GAA TGT GGA GGA AAG TTT ACC TAC AGT TCA GGC CTT ATT CAG CAT CAA AGA ATC CAC ACC GGG GAG AAC CCC TAT GAA
117 Tyr Glu Cys Met Glu Cys Gly Gly Lys Phe Thr Tyr Ser Ser Gly Leu Ile Gln His Gln Arg Ile His Thr Gly Glu Asn Pro Tyr Glu
                  (5)                                                                                               530
TGT AGT GAG TGT GGG AAA GCC TTC AGG TAC AGC TCG GCT CTT GTT CGC CAT CAG AGA ATT CAC ACT GGA GAG AAG CCT TTG AAT GGG ATC
147 Cys Ser Glu Cys Gly Lys Ala Phe Arg Tyr Ser Ser Ala Leu Val Arg His Gln Arg Ile His Thr Gly Glu Lys Pro Leu Asn Gly Ile
 (6)                                                                                                                629
GGC ATG AGC AAA AGC TCC CTC AGA GTT ACG ACC GAG TTA AAT ATC AGA GAG TCC ACG TGA AAGAGCCACACACCCATTTTCCTCACTTTCCCTGAGTCT
177 Gly Met Ser Lys Ser Ser Leu Arg Val Thr Thr Glu Leu Asn Ile Arg Glu Ser Thr ***
                                                                                                                   748
CAAGAGCTCTTGCCTTACCCTATAAATCTCAACAGCTTAGGATGTGTCCCTTTCAACTCAGACTTTTCATTTTAGAGAATGGGGCAGATGGGGCAAATCGTTGAATTTTCCCAGAAATC
                                                                                                                   867
ACACCAGCCTTAGAAAGCGTCAAGGCAAGTGGATGGCGTGCTGGGAATAGAAAGCAGCTCTGGGACCAGTTACCCCATTTAGGAAAGGAGTTTGCACTAAACTGTTTTTCTCACAGCAG
                                                                                                                   986
AGGAACCTTTCCAAGGTGGGGATGGAAGCACAGTCGGGACAGAATTCGATGGATTCCTTTAGTTGGAGTCCGCGTAGTCAGCACAGACAGCAGTGGAAAGGACGCTCTGGGTCCTGTTA
                                                                                                                  1105
TTTGCTAGGGAGGGTAAAGGGAGACTATTTCGGGCTACTGTTCCTAGTCCAGCTTTAAGTTTCGGTAAGAAACATGCTGTTTTGTTTCATGATTTCGTTAATTATGGAAATTTGGCATT
                                                                                                                  1224
GAGGGATTATTTTATTGAGGGTAGAAGAGATTCCAGAATCATCATCTGTGATGATGGTGTCCTTTAGGGCTCTTGGAGCAGCCAGACCATGTTTCCAAGAGAAACCTGGTGATATTGCC
                                                                                                                  1343
AGCAGACCCCCTGCCATCCCCCCAGTTGTCCTGGGGCTAATGGGCAAATCTGTCCAAACAGCTAGTAACCGGCTGTGAGGGAGAGGGTCAGAAGCACTTAGCGTTGGCCTCTGATTGCT
                                                                                                                  1462
GTCCTCTCTTGTCCTCTTCCCACTCCAATGATGAAAATGATTTTCTCTAAATGCCTGGGTAAGGATGCTTTCAAGGAGCTCACTTGGCCTGCTTGCCCTGCCCTCTCACCTCTGACACC
                                                                                                                  1581
CAGCCCCAGGAGCCAGACCACTCCTGCCTCCACCTCTGACTCTTCAGCAGCTGAGATTAATGCAGAGAAAGAGCAAAGCCTAAAAGGAGAAAAAAAAAACAAAAAAACAAAAAAACGAA
                                                                                                                  1700
GCCCAGAGAAGGCAGCTGTGCCTGAGCCTGGTAGCAGAGCTCGTTGCTGGGTGAGGATGGCACTCCCCGAGTTTTCCCAACCGACGGGATGACCTTCCATTGTGTTTTCCCAACCCTGC

TCATCCTGTAGCATAAATGAAGTGCACTATTAAACAGAATAGTTTTTCAGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCGAATTC
```

Figure 1.

Nucleotide sequences of the human cDNA clones cHF.10 (panel A) and cHF.12 (panel B) together with the translation of their longest open reading frame. The ten finger motifs of the cHF.10 and the six motifs of the cHF.12 proteins are numbered, a presumptive glycosylation site is indicated by the broken line in the cHF.10 sequence. (***) indicates the in-frame termination codons. The non-canonical ATTAAA polyadenylation signal of cHF.12 cDNA is underlined.

the multifingered region is located at the carboxy terminus of the protein. From the extimate size of HF.12 mRNA (≅3500 nucleotides, see below), the 1.8kb insert in HF.12 represents about one-half of the HF.12 mRNA.

Each of the repeat units of both cHF.10 and cHF.12 finger domains contain cysteine, hystidine and hydrophobic residues in the appropriate combination and localization to fit with the finger consensus sequence as initially identified by Klug and co-workers in Xenopus laevis TFIIIA (5,15), and subsequently found in a number of eukaryotic genes. Table 1 shows an alignment of the HF.10, HF.12 fingers with the consensus sequences derived from the mouse and human finger genes (11-12-13-14). Accordingly to the model originally proposed by Klug and collaborators (5) the crucial amino

Table 1. Amino acid sequence comparison of human and mouse finger motifs

	1	10	20	30
HF.10	C x x - - C G K a F x x s s x L x v H Q R i H	t g e K P y x		
HF.12	C x E - - C G K x F x x S S x L x x H Q R I H	T G E K P Y x		
Sp1	C x x(x)(x) C g K r f x r s d h L x x H x r t H	T G E x x F x		
TDF	C x x - - C x x x f x x x x x l x x H x x x H	x x x x x x x		
Mkr1	C x e - - C g K t F x x x s n L i x H q r i H	t g e k p y x		
Mkr2	C x e - - C g K a F x x x s s L t x H q r i H	T G E K P Y x		
Krox-20	C x x(x)(x) C x R x F s R S D e l t r H i r i H	T G x K P f x		
				H/C link

The amino acids sequneces are in the one letter code. For each gene, a consensus for the repeated motif is displayed: capital letters correspond to strictly conserved amino acids, lower case letters to amino acids converved over 60% and x no conservation. The interfinger H/C link is also indicated.

acids residues that specify the recognition of the target DNA sequences are located within the loop between Cys 6 and His 19 (the repeated units are numbered as in Table 1). In this region we found that Gly 7, Lys 8, Ser 13 amino acids residues are common in the finger of HF.10, HF.12, Sp1, mkr1 and mkr2 (Table 1). In addition, accordingly with the work of Schuh and collaborators (21), a stretch of six amino acids (TGEKPY/F) located between successive fingers (H/C link)is highly conserved.

Expression of Finger-containg genes in various cell lines and during myeloid differentiation

To determine the size and the distribution of the cHF.10 and cHF.12 transcripts a number of RNAs extracted from different human cell lines, were analyzed by Northern blotting. The cell lines were representative of various human tissues of different embryological derivation. cHF.10 and cHF.12 clones hybridize to a 2.6 Kb and a 3.5 kb mRNAs, respectively (Fig.2). The relative amount of the cHF.10 and cHF.12 transcripts varied among the different RNA samples. Rehybridization of Northern blots to the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe (24) revealed that a comparable amount of RNA was loaded in each lane (data not shown). These results suggest that the two finger genes cHF.10 and cHF.12 are variably expressed in all human tissues tested.

Next, we sought to determine whether the expression of the human finger genes was regulated during the process of in vitro differentiation, a cell

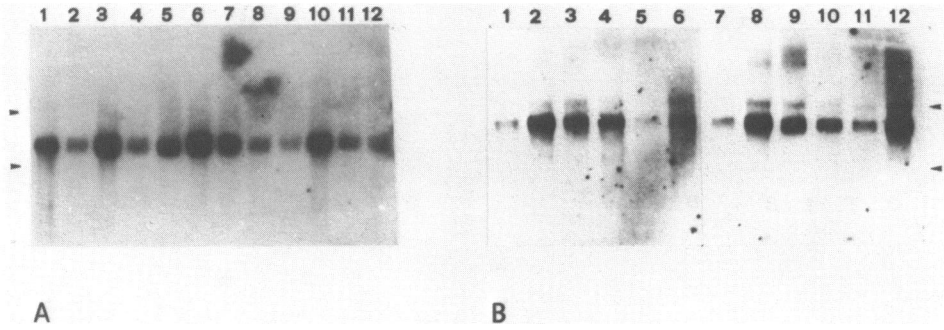

Figure 2.

Northern blot analysis of RNA isolated from various human cell lines hybridized to cHF.10 (panel A) and cHF.12 (panel B) cDNA probes. The RNA samples are: 1, EBV infected B lymphocytes (H5) ; 2, T-cell lymphoma (MOLT-4); 3, acute promyelocytic leukaemia (HL-60); 4, teratocarcinoma (NT2/D2); 5, normal fibroblasts (CRL1220); 6, Amnion (AV3); 7, neuroblastoma (HTB10); 8, chorion carcinoma (HTB144); 9, colon carcinoma (CLL220.1); 10, osteogenic sarcoma (HTB96); 11, breast cancer (HTB19); 12, cervical carcinoma (HeLa). Fifteen micrograms aliquots were electrophoresed on 1% agarose-formaldheyde gels, transferred to nitrocellulose filters. The blots were hybridized with the cHF.10 (A) or cHF.12 (B) cDNA clones under high stringency conditions. In overexposed autoradiograms of total RNA, some hybridization to both 28S and 18S RNA is detectable. Ribosomal RNAs (28S and 18S) are indicated by arrowheads for size markers.

process which typically requires a cascade of tightly regulated events. The rationale for these experiments was based on the facts that finger proteins are candidates for different transcriptional regulatory factors and that the finger genes mkrl, TFIIIA and Kruppel are thought to be implicated in various processes of cell differentiation (7,11,25). HL-60 and THP-1 are human acute leukaemia cell lines with phenotypic features of immature myeloid cells (26-27) capable of undergoing terminal differentiation upon treatment with a variety of chemical inducers (28). In both cell lines the process of in vitro induced terminal differentiation is accompanied by the arrest of cell proliferation (28). Upon treatment with dimethylsulfoxide (DMSO) or 12-O-tetradecanoylphorbol-13-acetate (TPA), HL-60 and THP-1 cells differentiate along the granulocytic or macrophage lineages, respectively (28). Fig. 3 shows a Northern blot analysis of total RNA isolated from untreated, DMSO-treated HL-60 cells and TPA-treated THP-1 cells and hybridized with cHF.10 (panel A) and cHF.12 (panel B) cDNA probes. The relative amount of the 2.6 kb cHF.10 transcript decreased after DMSO or TPA treatment in both cell lines (Fig.3A). cHF.12 hybridizing 3.5 kb trascript showed a similar reduction during the terminal differentiation of both HL-60 and THP-1-treated cells (Fig. 3b). Rehybridization of the Northern blots to the GAPDH probe revealed that comparable amounts of total RNA were present in each lane. We interpret these results as evidence that the expression of

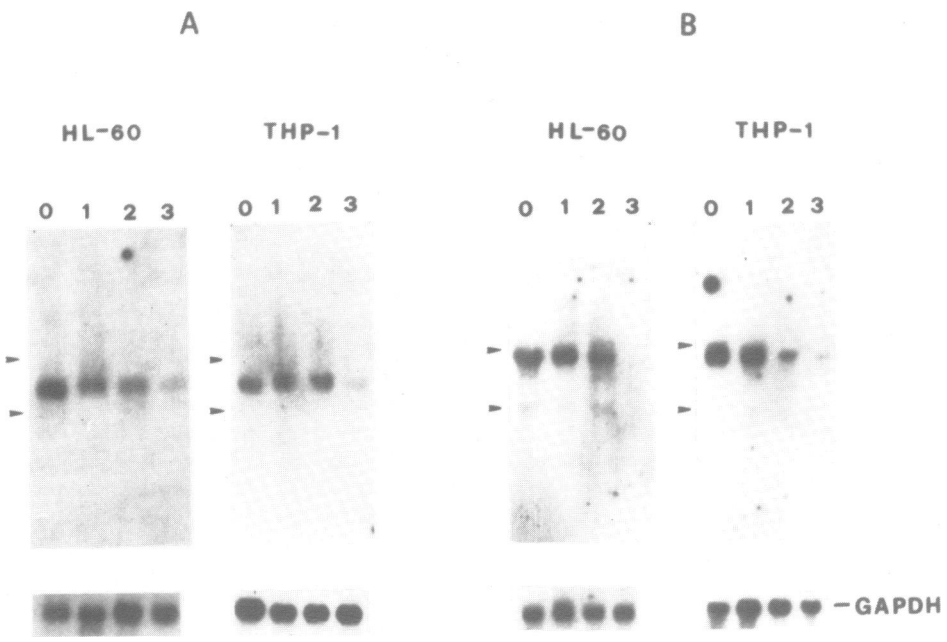

Figure 3.

Expression of cHF.10 (panel A) and cHF.12 (panel B) cDNA clones in undifferentiated (0) and differentiated (1-3) HL-60 and THP-1 myeloid cells. HL-60 were treated with DMSO for 5 days and RNAs extracted at day 1 (lane 1), day 3 (lane 2), and day 5 (lane 3) of culture in presence of the inducing agent. THP-1 cells were treated with TPA for 3 days and RNAs extracted at day 1 (lane 1), day 2 (lane 2) and day 3 (lane 3) of culture. The same blots were re-hybridized to the GAPDH probe (18). RNA extraction and Northern blot analysis were performed as described in the text. Ribosomal RNAs (28S and 18S) are indicated by arrowheads.

cHF.10 and cHF.12 transcripts is down regulated during the granulocyte and/or macrophage differentiation of human myeloid cells.

DISCUSSION

Since the first description of Zinc finger DNA binding domain in the 5S TFIIIA transcription factor in Xenopus (4,5), sequences for this protein domain have been found in genes from other species, including yeast, Drosophila and mammals (15,16,29). The importance of isolation and characterization of such genes lies in their potential capability of binding to nucleic acids and in their putative transcriptional regulatory activity. In this work, we have isolated and characterized two human cDNA both encoding multifinger proteins containing the Cys/His motif, and we have shown that the two genes are expressed in a number of different human cell lines and their expression is negatively regulated upon terminal

differentiation of immature myeloid cell lines.

The predicted amino acid sequences of cHF.10 and cHF.12 human clones indicate the presence of 10 and 6 finger repeats, respectively, and in both genes the multifingered regions are located at the carboxy terminus of the proteins. Since the cHF.12 clone represents about one-half of the HF.12 mRNA, we cannot exclude the presence of additional finger repeats. At present, we are in the process of isolating additional HF.12 clones to obtain a full-lenght cDNA.

The finger motifs of both HF.10 and Hf.12 encoded proteins contain the appropiate amino acids residues to fit with the consensus sequence of the Cys/His finger. In addition, comparison of human and mouse finger sequences shows the presence of highly conserved residues within the DNA-binding loop. As reported in Table 1, apart from the Phe 10 and Leu 16, we found that the amino acids residues Gly 7, Lys 8 and Ser 13 are highly conserved in the finger repeats of HF.10, HF.12, Sp1, mkr1 and mkr2 (11,13). Examination of the two human finger regions reveals that there is a highly conserved TGEKPYX sequence between contiguous fingers. Similar conserved sequence is also present between adjacent fingers encoded by several genes as: Kruppel, mkr1, mkr2, Krox-20, sp1 and Xfin (7,10,11,12,13). In contrast the Cys/His fingers in Drosophila hunchback and serendipity genes are not separated by a similar sequence (6,8). The partial conservation of this 7 amino acid stretch suggests that these amino acids residues are not strictlty required for the DNA binding of all finger proteins, but more likely they are important for the DNA-binding activity of some specific finger-containing proteins.

Transcription studies indicates that the HF.10 and HF.12 are expressed in a number of cell lines of different embryological derivation, and upon in vitro differentiation of human HL-60 and THP-1 myeloid cell lines the expression of both genes is negatively regulated. The fact that both human cDNA clones are expressed in immature myeloid cell lines and their expression is reduced in terminally differentiate cells, togheter with the potential DNA-binding capacity of the proteins, suggests that the two human finger genes may be involved in controlling gene expression during the process of in vitro myeloid terminal differentiation.

In both cell lines the process of in vitro induced terminal differentiation is accompained by the arrest of cell proliferation (28). It is therefore possible that the low levels of HF.10 and HF.12 transcripts observed in the terminal differentiated HL-60 and THP-1 cells are due to expression of the genes in a small subpopulation of proliferating cells. Alternatively, the expression patterns of HF.10 and HF.12 mRNAs might be due to cell-type specificity and may not be related to proliferation. Current research is aimed at differentiating between these possibilities. In this respect, it is interesting to note that a mouse finger gene, has beeen recently identified (12), and it has been shown that the expression of Krox-20 mRNA is activated during G_0/G_1 transition of cultured cells.

The isolation and characterization of other human finger genes may contribute to elucidate the role of finger proteins in gene regulation. We are presently using the human finger cDNA to isolate additional distinct

human finger genes and preliminary results indicate that others members of the human finger gene family are differentially expressed upon in vitro induced differentiation. In conclusion, our data add further support to the hypothesis that the finger gene family may play a major role in controlling gene activity during cell differentiation and/or proliferation.
Hovewer, the characterization of the protein products encoded by these human genes will be necessary to demonstrate their function.

ACKNOWLEDGEMENTS

We would like to thank Dr. E. Boncinelli for providing the oligonucleotides used in this work and for advice in preparation of this manuscript. This work was paid for by grants from Progetti Finalizzati CNR "Oncologia" and "Ingegneria Genetica" to L.L. and by the Italian Association for Cancer Research (AIRC) to L.L. and P-G.P. Anna Pascucci is recipient of an AIRC fellowship.

*To whom correspondence should be addressed

REFERENCES

1. Pabo,C.O. and Sauer,R.T. (1984) Ann. Rev. Biochem. 53, 293-321.
2. Laughon,A. and Scott,M.P. (1984) Nature 310, 25-31.
3. Shepherd,J.C.W., McGinnis,W., Carrasco,A.E., De Robertis,E.M. and Gehring,W.J. (1984) Nature 310, 70-71.
4. Brown,R.S., Sander,C. and Argos,P.(1985) FEBS Lett. 186, 271-274.
5. Miller,J., McLachlan,A.D. and Klug,A.(1985) EMBO J.4, 1609-1614.
6. Vincent,A., Colot,H.V. and Rosbash,M. (1985) J. Mol. Biol. 186, 149-166.
7. Rosenberg,U.B., Schroder,C., Preiss,A., Kienlin,A., Cote,S., Riede,I. and Jackle,H. (1986) Nature 319, 336-339.
8. Tautz,D., Lehmann,R, Schnurch,H., Schuh,R. Seifert,E., Kienlin,A., Jones,K. and Jackle,H. (1986) Nature 327, 383-389.
9. Hartshorne,T.A., Blumberg,H. and Young,E.T. (1986) Nature 320, 285-287.
10. Ruiz i Altaba,A., Perry-O'Keefe,H. and Melton,D.A. (1987) EMBO J. 6, 3065-3070.
11. Chowdhury,K., Deutsch,U. and Gruss,P. (1987) Cell 48, 771-778.
12. Chavrier,P., Zerial,M., Lemaire,P., Almendral,J., Bravo,R. and Charnay,P. (1988) EMBO J. 7, 29-35.
13. Kadonaga,J.T., Carner, K.R., Masiarz, F.R. and Tjian, R. (1987) Cell 51, 1079-1090.
14. Page,D.C., Mosher,R., Simpson,E.M., Fisher,E.M.C., Mardon,G., Pollack,J., McGillivray,B., de la Chapelle,A. and Brown,L.G. (1987) Cell 51, 1091-1104.
15. Klug,A. and Rhodes,D. (1987) T.I.B.S. 12, 464-469.
16. Evans,R.M. and Hollenberg,S.M. (1988) Cell, 52, 1-3.
17. Queen,C. and Korn,L.J. (1984) Nucleic Acids Res. 12, 581-599.
18. Newburger,P.E., Chovaniec,M.E., Grenberger,J.S. and Cohen,H.J. (1979) J. Cell. Biol. 82, 315-322.
19. Tsuchiya,S, Kobayashi,Y., Goto,Y., Okumura,H., Nakae,S., Konno,T. and Tada,K.. (1982) Cancer Res. 42, 1530-1536.

20. Chirgwin,J.M., Przybyla,A.E., Macdonald,R.J. and Rutter,W.J.(1979) Biochemistry 18, 5294-5299.
21. Schuh,R., Aicher,W., Gaul,U., Cotè,S., Preiss,A., Maier,D., Seifert,E., Nauber,U., Schroder,C., Kemler,R. and Jackle,H.(1986) Cell 47, 1025-1032.
22. Kozak, M. (1986) Cell 47, 481-483.
23. Marshall,R.D. (1972) Ann. Rev. Biochem. 41, 673-702.
24. Piechaczyk,M., Blanchard.,J.M., Marty,L., Dani,C., Panabieres,F., El Sabouty,S., Fort,P. and Jeanteur,P. (1984) Nucleic Acids Res. 12, 6951-6963.
25. Ginsberg,A.M., King,B.O. and Roeder,R.G. (1984) Cell 39, 479-489.
26. Collins,S.J., Gallo,R.C. and Gallegher,R.E. (1977) Nature 270, 347-350.
27. Tsuchiya,S., Yamabe,M., Yamaguchi,Y., Kobayashi,Y., Konno,T. and Tada,K. (1980) Int. J. Cancer 26, 171-176.
28. Nilsson,K. and Totterman,T.H. (1984) Immunol. Clin. Sper. 1, 11-29.
29. Berg,J. (1986) Science 232, 485-487.