Multiple Genes Encoding Zinc Finger Domains Are Expressed in Human T Cells

Hans-Jürgen Thiesen

Proteins containing zinc finger domains have been implicated in developmental control of gene expression in *Drosophila, Xenopus,* mouse, and humans. Multiple cDNAs encoding zinc (II) finger structures were isolated from human cell lines of T-cell origin to explore whether zinc finger genes participate in the differentiation of human hematopoietic cells. Initial restriction analysis, genomic Southern blotting, and partial sequence comparisons revealed at least 30 nonoverlapping cDNAs designated c*Kox*(1-30) encoding zinc finger motifs. Analysis of c*Kox*1 demonstrated that *Kox*1 is a single-copy gene that is expressed in a variety of hematopoietic and nonhaematopoietic cell lines. c*Kox*1 encodes 11 zinc fingers that were shown to bind zinc when expressed as a β-gal-Kox1 fusion protein. Further analysis of the predicted amino acid sequence revealed a heptad repeat of leucines NH_2-terminal to the finger region, which suggests a potential domain for homo- or heterodimer protein formation. On the basis of screening results it was estimated that approximately 70 zinc finger genes are expressed in human T cells. Zinc finger motifs are probably present in a large family of proteins with quite diverse and distinct functions. However, comparisons of individual finger regions in c*Kox*1 with finger regions of c*Kox*2 to c*Kox*30 showed that some zinc fingers are highly conserved in their putative α-helical DNA binding region, supporting the notion of a zinc finger–specific DNA recognition code.

Received March 5, 1990; revised April 6, 1990

Differentiation is thought to be a process determined by genetic programming of individual cells, partially mediated by cell-cell contacts and regulated by hormones and cytokines. In the mammalian hematopoietic system, many decisions on lineage commitment and maturation must be made during differentiation from pluripotent stem cells to mature cells comprising multiple lineages (B and T lymphocytes, granulocytes, macrophages, mast cells, erythrocytes, and megakaryocytes) (Dexter and Spooner, 1987; Metcalf, 1989). In turn, each lineage is composed of numerous cell types. Thus, mature T lymphocytes can be further subdivided according to function (for example, helper or cytotoxic T cells) and phenotype (McMichael et al., 1987; Strominger, 1989). Such differentiation processes involve the expression of lineage- and stage-specific genes (Maniatis et al., 1987). In recent years it has become clear that the transcription of genes is regulated by DNA-protein interactions and depends on the interplay of common and cell-type/cell-lineage–specific DNA-binding proteins (Lichtsteiner et al., 1987; Mitchell and Tjian, 1989).

In *Drosophila,* analysis of developmental mutants led to the characterization of genes that play a pivotal role in the differentiation process. Many of these genes encode protein domains conserved in higher species, such as mouse and man (for example, homeobox: Manley and Levine, 1985; Gehring, 1987; Holland and Hogan, 1988; paired box: Bopp et al., 1986; Deutsch et al., 1988; and zinc finger: Chowdhury et al., 1987). The zinc finger domains first found in *Xenopus* transcription factor IIIA (TFIIIA) (Ginsberg et al., 1984) are typified by the *Drosophila* gap genes *Krüppel* (Rosenberg et al., 1986) and *hunchback* (Tautz et al., 1987). The zinc finger motif is a tandemly repeated 28-amino acid sequence analyzed in TFIIIA by Brown et al. (1985) and Miller et al. (1985). The conserved cysteines and histidines of one repeat (**YxCxxCxxxFxxxxxLxxHxxx HTGER/KP**) have been shown to chelate one zinc atom (Diakun et al., 1986), stabilizing the proposed finger structure (Miller et al., 1985). For TFIIIA and for the yeast transcription activators ADR1 and SWI5, the finger domain has been identified as the DNA-binding component (Blumberg et al., 1987; Stillman et al., 1988). A collection of 159 zinc finger motifs found in 25 translated genes, including the three human zinc finger genes SP1 (Kadonaga et al., 1987), GLI (Kinzler et al.,

Basel Institute for Immunology, Grenzacherstrasse 487, Postfach, CH-4005 Basel, Switzerland.

1043-4674/90/0204-0005$5.00/0

KEY WORDS: zinc finger genes/hematopoiesis/differentiation control/leucine zipper/DNA recognition code

A Kox1

```
        10                  30                  50                  70
GGGCGGTGGTGCTTTGTCTCCTCAGCACTCTGCTGTCACTCAAGGAAGTATCATCAAGAACAAGGAGGGCATGGATGCTA
 G  G  G  A  L  S  P  Q  H  S  A  V  T  Q  G  S  I  I  K  N  K  E  G  M  D  A  K

        90                 110                 130                 150
AGTCACTAACTGCCTGGTCCCGGACACTGGTGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTG
 S  L  T  A  W  S  R  T  L  V  T  F  K  D  V  F  V  D  F  T  R  E  E  W  K  L

       170                 190                 210                 230
CTGGACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGTTATCAGCT
L  D  T  A  Q  Q  I  V  Y  R  N  V  (M) L  E  N  Y  K  N  (L) V  S  L  G  Y  Q  (L)

       250                 270                 290                 310
TACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCTGGCTGGTGGAGAGAGAAATTCACCAAGAGACCC
 T  K  P  D  V  I  (L) R  L  E  K  G  E  E  P  W  L  V  E  R  E  I  H  Q  E  T  H

       330                 350                 370                 390
ATCCTGATTCAGAGACTGCATTTGAAATCAAATCATCAGTTTCCAGCAGGAGCATTTTTAAAGATAAGCAATCCTGTGAC
  P  D  S  E  T  A  F  E  I  K  S  S  V  S  S  R  S  I  F  K  D  K  Q  S  C  D

       410                 430                 450                 470
ATTAAAATGGAAGGAATGGCAAGGAATGATCTCTGGTATTTGTCATTAGAAGAAGTCTGGAAATGTAGAGACCAGTTAGA
I  K  M  E  G  M  A  R  N  D  L  W  Y  L  S  L  E  E  V  W  K  C  R  D  Q  L  D

       490                 510                 530                 550
CAAGTATCAGGAAAACCCAGAGAGACATTTGAGGCATCAGCTTATTCATACTGGAGAAAAACCCTATGAGTGTAAAGAAT
 K  Y  Q  E  N  P  E  R  H  L  R  H  Q  L  I  H  T  G  E  K  P  Y  E  C  K  E  C

       570                 590                 610                 630
GTGGAAAGTCTTTCAGCCGGAGTTCTCACCTCATTGGACATCAAAAGACCCATACTGGTGAGGAACCCTATGAATGTAAA
  G  K  S  F  S  R  S  S  H  L  I  G  H  Q  K  T  H  T  G  E  E  P  Y  E  C  K

       650                 670                 690                 710
GAATGTGGAAAATCCTTCAGCTGGTTCTCTCACCTTGTTACTCATCAGAGAACTCATACAGGAGACAAACTGTACACATG
E  C  G  K  S  F  S  W  F  S  H  L  V  T  H  Q  R  T  H  T  G  D  K  L  Y  T  C

       730                 750                 770                 790
TAATCAGTGTGGGAAATCTTTTGTTCATAGCTCTAGGCTTATTAGACACCAGAGGACACATACTGGACACAAACCCTATG
 N  Q  C  G  K  S  F  V  H  S  S  R  L  I  R  H  Q  R  T  H  T  G  H  K  P  Y  E

       810                 830                 850                 870
AATGTCCTGAATGTGGGAAATCTTTCAGACAGAGCACACATCTCATTCTGCATCAGAGAACCCATGTGAGAGTGAGGCCC
  C  P  E  C  G  K  S  F  R  Q  S  T  H  L  I  L  H  Q  R  T  H  V  R  V  R  P

       890                 910                 930                 950
TATGAATGCAATGAATGTGGAAAGTCTTACAGCCAGAGATCTCACCTTGTTGTGCATCATAGAATTCACACTGGACTAAA
Y  E  C  N  E  C  G  K  S  Y  S  Q  R  S  H  L  V  V  H  H  R  I  H  T  G  L  K

       970                 990                1010                1030
ACCTTTTGAGTGTAAGGATTGTGGAAAATGTTTTAGTCGAAGCTCTCACCTTTATTCACATCAAAGAACCCACACTGGAG
 P  F  E  C  K  D  C  G  K  C  F  S  R  S  S  H  L  Y  S  H  Q  R  T  H  T  G  E

      1050                1070                1090                1110
AGAAACCATATGAGTGTCATGATTGTGGAAAATCTTTCAGCCAGAGTTCTGCCCTTATTGTGCATCAGAGGATACACACT
  K  P  Y  E  C  H  D  C  G  K  S  F  S  Q  S  S  A  L  I  V  H  Q  R  I  H  T

      1130                1150                1170                1190
GGAGAGAAACCATATGAATGCTGTCAGTGTGGGAAAGCCTTCATCCGGAAGAATGACCTCATTAAGCACCAGAGAATTCA
G  E  K  P  Y  E  C  C  Q  C  G  K  A  F  I  R  K  N  D  L  I  K  H  Q  R  I  H

      1210                1230                1250                1270
TGTTGGAGGAGAGACCTATAAATGTAATCAATGTGGCATTATCTTCAGCCAGAACTCTCCATTTATAGTTCATCAAATAG
 V  G  G  E  T  Y  K  C  N  Q  C  G  I  I  F  S  Q  N  S  P  F  I  V  H  Q  I  A

      1290                1310                1330                1350
CTCACACTGGAGAGCAGTTCTTAACATGCAATCAATGTGGGACAGCGCTTGTTAATACCTCTAACCTTATTGGATACCAG
  H  T  G  E  Q  F  L  T  C  N  Q  C  G  T  A  L  V  N  T  S  N  L  I  G  Y  Q

      1370                1390                1410
ACAAATCATATTAGAGAAAATGCTTACTAATAAATATGGGAATTTTTCACAAAT
T  N  H  I  R  E  N  A  Y  *  *
```

Figure 1A. Nucleotide sequence and predicted amino acid sequence of c*Kox*1.

The finger region starts at nucleotide position 461 and ends at 1384. Methionine and leucine residues of the zipper region are encircled (amino acids 66 to 87). The negatively charged region (net charge −6) spans from amino acid position 89 to position 115.

1988), and TDF (Page et al., 1987), has been structurally compared by Gibson et al. (1988).

As an initial step in determining whether genes encoding zinc finger domains are involved in differentiation of hematopoietic cell lineages we have isolated and characterized 30 nonoverlapping cDNAs containing multiple zinc fingers derived from libraries of T-cell origin.

RESULTS AND DISCUSSION

*Structure of c*Kox1 *Isolated from Molt-4 cells*

We isolated a cDNA clone derived from Molt-4 cells, a human T-cell line, on the basis of cross-hybridization with the zinc finger region of mouse *Krüppel mkr*1 (Chowdhury et al., 1987). This clone was named c*Kox*1 for ***Krüppel* box**. On genomic blot analysis this cDNA clone detected a single-copy gene in the human genome (see below). Under the same hybridization conditions we did not detect cross-hybridization with mouse DNA, suggesting that the sequence is not highly conserved in evolution (data not shown). Sequence analysis (Fig. 1A) demonstrated an open reading frame of 462 amino acids potentially encoding a protein of at least 50 kD. The predicted protein sequence revealed the presence of 11 consecutive zinc fingers (Fig. 1B) which encoded the two highly conserved Cys/His residues necessary for zinc coordination with variations in finger 1 and finger 11. The H/C-link consensus sequence TGEKP is only present in three links (fingers 1, 7, and 8). Comparisons of individual fingers in Kox1 indicated that several amino acids are conserved between individual fingers (Fig. 1B). Besides the zinc finger consensus sequence (**YeCxxCgksFsxxshLixHqrtH**-tgekp) several amino acids are repeated in several fingers (small letters) with some variations in positions (underlined) thought to be involved in DNA binding (Gibson et al., 1988).

B

1. KCRDQLDKYQENPERHLR**H**QL**I**H**T**GEKP
2. **Y**E**C**KE**C**GKS**F**SRSSH**L**IG**H**QKT**H**TGEEP
3. **Y**E**C**KE**C**GKS**F**SWFSH**L**VT**H**QRT**H**TGDKL
4. **Y**T**C**NQ**C**GKS**F**VHSSR**L**IR**H**QRT**H**TGHKP
5. **Y**E**C**PE**C**GKS**F**RQSTH**L**IL**H**QRT**H**VRVRP
6. **Y**E**C**NE**C**GKSYSQRSH**L**VV**H**HRI**H**TGLKP
7. FE**C**KD**C**GKC**F**SRSSH**L**YS**H**QRT**H**TGEKP
8. **Y**E**C**HD**C**GKS**F**SQSSA**L**IV**H**QRI**H**TGEKP
9. **Y**E**C**CQ**C**GKA**F**IRKND**L**IK**H**QRI**H**VGGET
10. **Y**K**C**NQ**C**GII**F**SQNSPFIV**H**QIA**H**TGEQF
11. LT**C**NQ**C**GTALVNTSN**L**IGYQTN**H**IRENA

cons.:YeCxxCgksFsxxshLixHqrxHtgekp

Figure 1B. Compilation of 11 zinc fingers together with the consensus sequence predicted from c*Kox*1 nucleotide sequences (Fig. 1A).

The finger region starts at nucleotide 461 and ends at nucleotide 1385. Bold capitals show amino acids generally conserved in zinc finger motifs; small letters in the consensus sequence show amino acids that are repeated in several fingers of Kox1 with some variations in positions (underlined) thought to be involved in DNA binding.

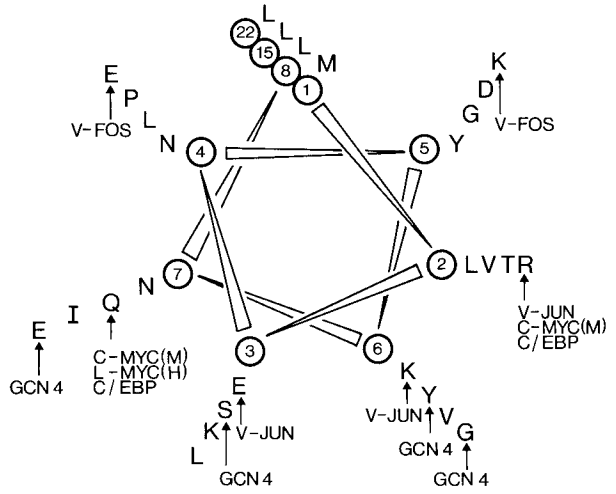

Figure 2. Comparison of heptad repeats of leucines present in Kox1 with leucine zipper structures from v-Fos, v-Jun, c-Myc, l-Myc, C/EBP, and GCN4.

Conserved amino acids within heptad repeats of v-Fos, v-Jun, c-Myc, l-Myc, C/EBP, and GCN4 are indicated by arrows.

Heptad Repeats of Leucines Present in Kox1

The analysis of transcription factors in yeast and higher eukaryotes established that transcription factors consist of DNA-binding domains, such as the homeobox (Desplan et al., 1988; Hoey and Levine, 1988) and basic regions (GCN4: Hope and Struhl, 1986), and an activating domain that can be negatively charged (GCN4: Hope et al., 1988; Ptashne, 1988) or may contain glutamines (SP1: Courey et al., 1988). A search of the coding region of c*Kox*1 for significant homologies revealed the presence of a structure resembling a leucine zipper from nucleotide 197 to nucleotide 262 (Fig. 2). However, this is followed by a negatively charged region (net charge of -6) from nucleotide position 268 to position 346 which is in contrast to the basic region seen in the analogous position in canonical leucine zipper structures. A leucine zipper, a heptad stretch of leucines separated by six amino acids, was first described in C/EBP and found to be present in the Myc family, in Jun, Fos, and GCN 4 (Landschulz et al., 1988). This region is involved in facilitating DNA-protein interaction as a result of homodimer formation in GCN4 (O'Shea et al., 1989) and heterodimer formation be-

tween c-Jun and c-Fos (Sassone-Corsi et al., 1988; Gentz et al., 1989; Turner et al., 1989). A comparison of the amino acids within the leucine stretches of Jun, Fos, Myc, and C/EBP with Kox1 reveals that additional amino acids are conserved (Fig. 2). Furthermore, a detailed analysis of the postulated leucine zipper is reminiscent of the coiled coil structure described in interfilament proteins, in which hydrophobic amino acids appear at every 3rd or 4th and 7th position (O'Shea et al., 1989). A detailed sequence comparison with the National Biomedical Research Foundation (NBRF) protein data bank with $Mx_6Lx_6Lx_6L$ from the Wisconsin GCG package being used as a mask revealed homologies to c-Myc in the rainbow trout, to human dystrophin, and to E4 and E7 from human papilloma viruses (Fig. 3). In this respect it is worth mentioning that E7 is suspected of forming oligomers or of being associated with one or more proteins (Smotkin and Wettstein, 1987). It is tempting to postulate that these interactions are determined by this methionine/leucine structure.

Though negatively charged regions in transcription factors have already been shown to activate transcription (Ma and Ptashne, 1987), it is unclear whether, in Kox1, the negatively charged region from amino acid position 89 to 115 and the leucine zipper region from amino acid position 66 to 87 form one functional unit for protein-protein interaction or whether they present two independent functional regions. Although several leucine zipper–containing transcription factors have been described, the product of *Kox*1 is the first example where a multiple zinc finger protein contains heptad repeats of leucines that might be involved in homodimer or heterodimer protein formation.

Expression of Kox1

It was established by Northern analysis (Fig. 4) that *Kox*1 is expressed in various hematopoietic and nonhematopoeitic cell lines. The highest expression compared to that of GAPDH was observed in U937, a myelomonocytic cell line. No expression has been detected thus far in HT29, a colon carcinoma cell line, or in OVCA, an ovarian carcinoma cell line. In addition to the longest transcript of c*Kox*1, shorter transcripts were visible that might represent alternatively spliced products of the *Kox* message or transcripts of related genes. Although the expression of *Kox*1 was not restricted to hematopoietic cells, the level of expression seemed to vary within hematopoietic lineages: *Kox*1 may be more strongly expressed in the myelomonocytic compartment (U937) in comparison to the lymphoid compartment (Molt-4, Jurkat, and Ramos).

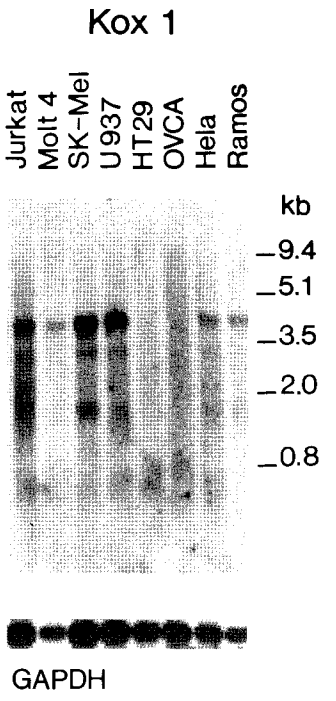

Figure 4. Northern blot probed with c*Kox*1.

Poly(A)$^+$ RNA selected from 100 μg of total RNA by mAP paper (Orgenics, Yavne, Israel) were run on 1% agarose-formaldehyde gel, transferred to GeneScreen (Dupont), and hybridized with c*Kox*1. RNA was obtained from cell lines indicated above the lanes. The GAPDH probe (Piechaczyk et al., 1984) was used to show that equal amounts of RNA were loaded in each sample. Phage λ digests (EcoRI, HindIII) were used as size markers.

Zinc Binding

Several amino acids are highly conserved within zinc finger regions (Gibson et al., 1988). In ADR1

```
MLENYKNLVSLGYQLTKPDVILRL     cKox1
|      |  |  ||      | |
MVGEMPALKDLVLQLEPSVLDLDL     E7(HPV: type 1a)
|      |  |  |      | |
MAAQLYVLLHLYLALYEKYPLLNL     E4(HPV: type 11)
|      |  |  | ||   | |
MAAQLYVLLHLYLALHKKYPFLNL     E4(HPV: type 6b)
|      || |  ||     |
MQTDEQRLVNLKEQLRRKSEHLKQK    cmyc (rainbow trout)
|      |  |  |      |
MEQRLGELQALQSSLQEQQSGL       dystrophin (human)
```

Figure 3. Methionine/leucine heptad repeats of Kox1 found in human papilloma virus, human dystrophin, and rainbow trout.

The methionine/leucine heptad motif from Kox1 ($Mx_6Lx_6Lx_6L$) is present in E4 (amino acids 11-34) and E7 (amino acids 1-24) in human papilloma virus, in c-Myc (amino acids 376-400) from rainbow trout, and in human dystrophin (amino acids 975-996). Vertical lines above individual sequences show identity to amino acids in Kox1.

(Parraga et al., 1988) and TFIIIA (Frankel et al., 1987) it has been demonstrated that one zinc atom coordinates the two Cys/His residues of one finger domain (Diakun et al., 1986). It is proposed that this coordination determines and stabilizes the putative zinc finger structure, thus reducing the degree of potential conformations and giving the zinc finger a more rigid structure independent of the amino acids at the non highly conserved positions (Parraga et al., 1988). A zinc blot performed on a β-gal-Kox1 fusion protein demonstrated the product of specific zinc binding to *Kox*1 (Fig. 5). A low level of zinc binding was observed with native β-galactosidase. Additional zinc blots with recombinant Kox1 consisting of the first five or last five zinc fingers demonstrated clear zinc binding in the absence of β-galactosidase sequences (data not shown).

*Isolation of Multiple Zinc Finger cDNAs (c*Kox*2-30)*

In a second screening procedure, with c*Kox*1 being used as a probe, 29 nonoverlapping cDNAs encoding zinc finger structures were isolated from T-cell libraries (c*Kox* 2-29 from Jurkat and c*Kox*30 from Molt-4 cells). Partial sequence analysis and restriction analysis of c*Kox*2 to c*Kox*30 indicated that these cDNAs are different from each other. Many zinc finger genes harbor EcoRI restriction sites within the zinc finger region. Therefore, we grouped the c*Kox* clones in Fig. 6 according to the position of their zinc finger regions. c*Kox*1 and c*Kox*30 encompass complete zinc finger regions. Most of the c*Kox* clones probably represent separate genes, because genomic blots with individual c*Kox* cDNAs as probes gave different hybridization patterns (Fig. 6). Genomic clones representing the c*Kox* family are currently being analysed to determine whether different *Kox* genes are linked to each other. However, preliminary data obtained by hybridizations with different stringencies indicate that *Kox* genes can be grouped into different subsets of genomic clones, indicating that zinc finger regions might have diverged through evolution from each other (H.-J. Thiesen, unpublished data). The fact that many fragments are detected on genomic Southern blots by c*Kox*8, by c*Kox*24, and by c*Kox*27 supports the notion that a large number of zinc finger genes are present in the human genome. The ratio of clones recovered from the Jurkat cDNA library only once to those recovered twice suggests that a lower limit for the number of zinc finger genes expressed in T cells is approximately 70. This is a minimum estimate based on the assumption that all genes are expressed in equal copy numbers and represented equally in amplified libraries. Studies of the expression of these *Kox* genes on Northern blots are complicated by this high number of different zinc finger genes and their low abundance (H.-J. Thiesen, in preparation).

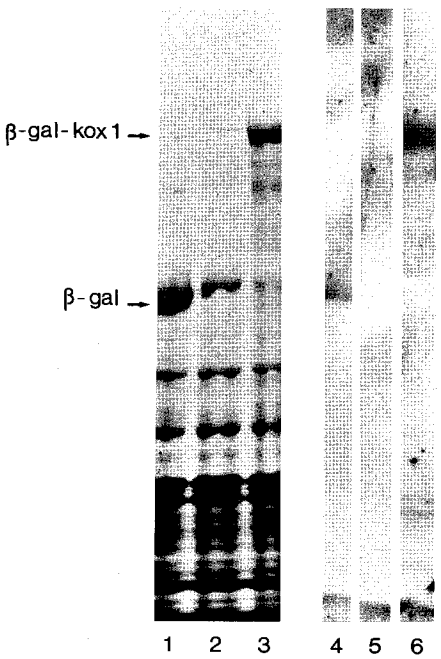

Figure 5. Zinc blot onto Kox1 β-gal-fusion protein.

c*Kox*1 was expressed as a β-gal-fusion protein in pUEX3, run on a 7.5% SDS-polyacrylamide gel, and electrotransferred to nitrocellulose. A zinc blot was then performed with ^{65}Zn (Schiff et al., 1988). Lanes: **1,** β-gal; **2,** β-gal-Kox1 (reverse); **3,** β-gal-Kox1; **4,** zinc blot of 1; **5,** zinc blot of 2; **6,** zinc blot of 3. Proteins in lanes 1 to 3 were stained with Coomassie brilliant blue.

Protein sequences of two consecutive zinc fingers derived from each of c*Kox*2 to c*Kox*30 are presented in Fig 7. The amino acids of the zinc finger consensus sequence (**YxCxxCxxxFxxxxxLxxHxxxHTGEKP**) are generally highly conserved in this collection of Kox finger regions. The second cysteines are replaced by serine and tyrosine in two Kox24 fingers. One histidine is altered to tyrosine in one Kox26 finger. The hydrophobic residues phenylalanine and leucine are present in most c*Kox* fingers. Leucine is replaced by isovaline in Kox13 and by phenylalanine in Kox5 and Kox14. In Kox7 one leucine was detected instead of phenylalanine, which was replaced by tyrosine in Kox17 and Kox20. These hydrophobic residues are thought to stabilize the finger tip by forming a hydrophobic core (Gibson et al., 1988).

Comparison of Individual Kox Zinc Fingers

Although zinc finger proteins might have quite diverse functions, the zinc finger domain is thought to bind specifically to DNA or RNA, as indicated in extensive analyses of TFIIIA (Engelke et al., 1980; Fairall et al., 1986; Pieler et al., 1985, 1987). Rhodes and Klug (1986) suggested that the repetitive structure of zinc finger domains might be reflected in a repeated DNA target sequence. Structural models generated by comparing known structures of metalloproteins (Berg, 1988) and by doing computer modelling (Gibson et al.,

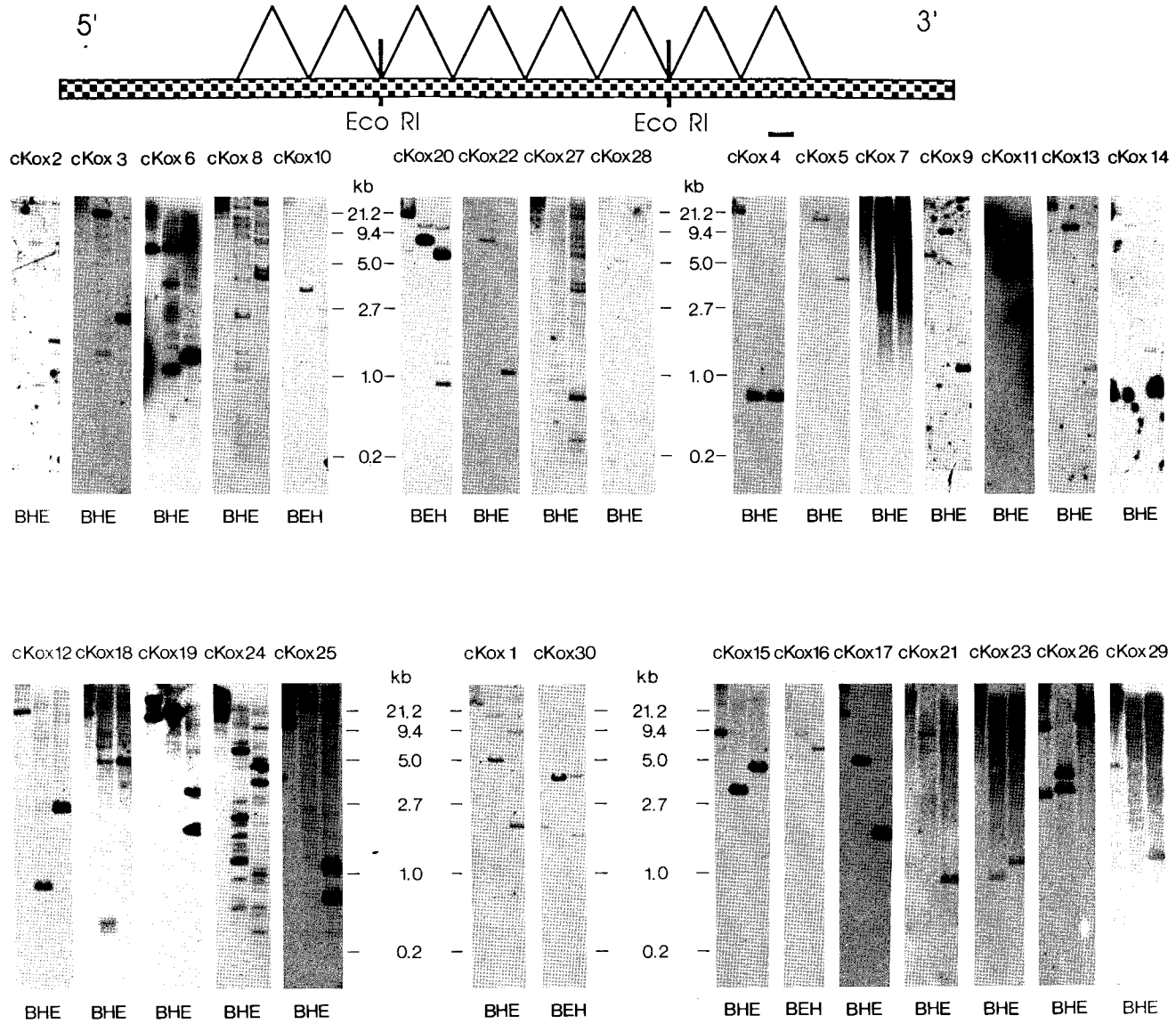

Figure 6. Genomic Southern blots probed with c*Kox*1 to c*Kox*30.

For the genomic Southern blots, 10 μg of Molt 4 DNA were digested with the restriction enzymes shown below the panels (BamH1, B; HindIII, H; and EcoRI, E), electrophoresed in 7% agarose gels, transferred to GeneScreen (Dupont) and hybridized with c*Kox* probes as shown above the panels. The c*Kox*1 blot was probed with a fragment containing the first 660 residues of c*Kox*1 (Fig. 1A). c*Kox* blots were grouped according to the location of the finger region in the cDNAs. The 3′ zinc fingers were located in c*Kox*2, 3, 8, 10, 12, 18, 19, 24, and 25, and the 5′ fingers were located in c*Kox*4, 5, 7, 9, 11, 13, 14, 15, 16, 17, 21, 23, 26 and 29. c*Kox*1 and c*Kox*30 encompass complete finger regions. c*Kox*20, 22, 27, and 28 contain only zinc finger motifs. Phage λ that had been cut with EcoRI/HindIII and pUC19 that had been cut with BamH1/AvaII were used as size markers.

1988) suggest that each finger contains an α-helical region as previously predicted by Brown and Argos (1986); this region seems to start around the third position prior to the conserved leucine and end around the second conserved histidine, and thus spans three to four helical turns. A similar placement of an α-helix was identified in ADR1 (Parraga et al., 1988). In addition, it is claimed that amino acids at positions 3 and 2 prior to the conserved leucine, and both amino acids located between the leucine and the first conserved histidine, determine sequence specificity (Gibson et al., 1988). Recently, the three-dimensional solution structure of a single zinc finger derived from the *Xenopus* protein Xfin was determined by nuclear magnetic resonance (NMR); this work showed that the zinc finger is an independently folded domain with a compact globular structure, and suggested that the side-chains of the amino acids on the exposed face of the well-defined helix may participate in DNA binding (Lee et al., 1989). Mutations of the yeast regulatory protein ADR1 also suggest that the position

KOX 2.3
F E C N E C G E A F S H K S A L T L H Q R T H T G E K P
Y Q C N A C G E T F Y Q K S D L T K H Q R T H T G Q K P

KOX 3.3
Y K C S E C G K C F C R K S T L T T H L R T H T G E K P
Y E C N E C G K F F S R L S Y L T V H Y R T H S G E K P

KOX 4.5
F K C D E C G K A F R W I S R L S Q H Q L I H T G E K P
Y K C N K C T K A F G C S S R L I R H Q R T H T G E K P

KOX 5.5
Y Q C D A C G K G F S R S S D F N I H F R V H T G E K P
Y K C E E C G K G F S R A S N L L A H Q R G H T G E K P

KOX 6.3
Y E C K Q C H K T F S F S S S L R E H E T T H T G E K P
Y E C K Q C G K T F S F S S S L Q R H E R T H N A E K P

KOX 7.5
Y E C K Q C G K L L S H R S S F R R H M M A H T G D G P
H K C T V C G K A F D S P S V F Q R H E R T H T G E K P

KOX 8.3
Y K C E E C G K S F I L S S H L T T H K I I H T G E K P
Y K C K E C G K A F N Q S S T L M K H K I I H T G E K P

KOX 9.5
H V C N V C G K A F S Y S S V L R K H Q I I H T G E K P
Y R C S V C G K A F S H S S A L I Q H Q G V H T G D K P

KOX 10.3
Y E C N K C G K F F R Y C F T L N R H Q R V H S G E R P
Y E C S E C G K F F V D S C T L K S H Q R V H T G E R P

KOX 11.5
F Q C T I C K K A F L R S S D F V K H Q R T H T G E K P
C K C D Y C G K G F S D F S G L R H H E K I H T G E K P

KOX 12.3
Y E C N E C G K A F V G N S P L L R H Q K I H T G E K P
Y E C N E C G K S F G R T S H L S Q H Q R I H T G E K P

KOX 13.5
Y K C K Q C G K A F T R S T T L P V H E R T H T G V N A
D E C K E C G N A F S F P S E I R R H K R S H T G E K P

KOX 14.5
Y E C P V C W K A F S Q K S Q L I I H Q R T H T G E K P
Y A C T E C G K A F R E K S T F T V H Q R T H T G E K P

KOX 15.5
Y Q C D E C G R C F S Q S S H L I Q H Q R T H T G E K P
Y Q C S E C G K C F S Q S S H L R Q H M K V H K E E K P

KOX 16.5
F E C N E C G R C F T S K R N L L D H H R I H T G E K P
Y Q C K E C G K A F S I N A K L T R H Q R I H T G E K P

KOX 17.5
. . . N S G L I N H Q R I H T G E K P
Y E C V Q C G K S Y S Q S S N L F R H Q R R H N A E K L

KOX 18.3
Y E C S E C G K A F S L N S N L V L H Q R I H T G E K P
H E C N E C G K A F S H S S N L I L H Q R I H S G E K P

KOX 19.3
F E C T E C G K F F Y V K A Y L M V H Q K T H T G E K P
Y E C K E C G K A F S Q K S H L T V H Q R M H T G E K P

KOX 20.3
Y E C N E C E K A Y P R K A S L Q I H Q K T H S G E K P
F K C S E C G K A F T Q K S S L S E H Q R V H T G E K P

KOX 21.5
Y G C H E C G K S F S E K S T L T Q H Q R T H T G E K P
Y E C H E C G K T F S F K S V L T V H Q K T H T G E K P

KOX 22.5
F K C V E C G K G F S R R S A L N V H H K L H T G E K P
Y N C E E C G K A F I H D S Q L Q E H Q R I H T G E T P

KOX 23.5
Y E C S E C G K S F R Q R S G L I Q H R R L H T G E R P
Y E C S E C G K S F S Q S A S L I Q H Q R V H T G E R P

KOX 24.5
F Q C I E S G K S F N C S S L L K K H Q I T H L E E K Q
C K C D V Y G K V F N Q K R Y L A C H R R S H I D E K P

KOX 25.3
Y E C N E C G K A F S Q S S H L Y Q H Q R I H T G E K P
Y E C M E C G G K F T Y S S G L I Q H Q R I H T G E N P

KOX 26.5
Y K C K E C G E A F S Q S S T L T K H I K V Y T G E K P
Y T C K D C R K A F S Q S S S L T Q H Q R V H T G K K P

KOX 27.5
Y K C E E C G K A F S R S S N L T K H K K I H I E K K P
Y K C E E C G K A F K W S S K L T E H K I T H T G E K P

KOX 28.3
Y E C K E C G K T F S R A S Y L V Q H S R I H T G K K P
Y E C K E C G K A F S S G S Y L V Q H Q R I . . .

KOX 29.5
Y E C S E C G R S F S K S S A L I S H Q R I H T G E K P
Y E C A E C G K S F S K S S T L A N H Q R T H T G E K P

KOX 30.5
Y E C Q E C G K S F R Q K G S L T L H E R I H T G Q K P
F E C T H C G K S F R A K A N L V T H Q R I H T . . . .

Figure 7. Two consecutive zinc finger domains from cDNAs c*Kox*2 to c*Kox*30.

Predicted amino acid sequences have been derived from the 3′ or 5′ end of the cDNA clones as indicated by a 3 or 5 in the numbering of each c*Kox* clone. Thus, Kox2.3 zinc finger sequences are derived from the 3′ end of the c*Kox*2 clone.

just prior to the leucine might contribute to DNA recognition (Blumberg et al., 1987). This region is thought to interact with nucleotides in the major groove of the DNA (Berg, 1989).

Each individual finger derived from c*Kox*1 (total 11) was compared to two consecutive fingers derived from c*Kox*2 to k*Kox*30 (total 58) to evaluate whether single DNA-binding regions of Kox1 appear in other zinc finger proteins (Fig. 8). The triplets SSH, SSR, and SSA appeared in several other Kox zinc fingers. It is questionable whether the SS doublet confers any specificity because of the low selectivity of serines in terms of hydrogen bonding to nucleotides (Saenger, 1984). In further comparisons identical amino acid doublets were

Kox8	YKCEECGKSFIL **SSH** L TT HKIIHTGEKP
Kox1.F7	FECKDCGKCFSR **SSH** L YS HQRTHTGEKP
Kox1.F2	YECKECGKSFSR **SSH** L IG HQKTHTGEEP
Kox15	YQCDECGRCFSQ **SSH** L IQ HQRTHTGEKP
Kox25	YECNECGKAFSQ **SSH** L YQ HQRIHTGEKP
Kox1.F8	YECHDCGKSFSQ **SSA** L IV HQRIHTGEKP
Kox29	YECSECGRSFSK **SSA** L IS HQRIHTGEKP
Kox9	YRCSVCGKAFSH **SSA** L IQ HQGVHTGDKP
Kox1.F4	YTCNQCGKSFVH **SSR** L **IR** HQRTHTGHKP
Kox4	YKCNKCTKAFGC **SSR** L **IR** HQRTHTGEKP
Kox18	HECNECGKAFSH SSN L **IL** HQRIHSGEKP
Kox1.F5	YECPECGKSFRQ STH L **IL** HQRTHVRVRP
consensus	YeC..CgksFs. **SS.** L **I.** HqrtHtgekp

Figure 8. Comparison of individual zinc fingers of Kox1 with zinc fingers of Kox2 to Kox30.

The Kox fingers (Kox1.F) are listed only if they contain high conservations of amino acids within the putative DNA binding region (xxxLxxH), implying similar specificities in DNA binding. Identical triplets and doublets are presented in bold characters.

noticed located between leucine and histidine. The Kox1 fingers 2, 4, and 8, when compared with fingers from Kox4, 9, 15, 18, and 29, differ only in the first position prior to leucine and in the first position prior to histidine (ssxlixH). It is worth asking whether every third position in these two consecutive helical turns determines specific binding to DNA or whether mainly identical triplets and doublets confer similar specificities for DNA recognition. If one compares the fourth finger of Kox1 with a finger in Kox4, one finds a total identity in the putative DNA-binding regions, suggesting that these two fingers should recognize the same DNA target sequence according to the models generated by Berg (1988) and Gibson et al. (1988). With the identification of Kox target sequences, finger-swap experiments will be the appropriate method to evaluate the existence of a zinc finger specific recognition code.

Genes Encoding Zinc Finger Motifs Expressed in Human T Cells

In this report, I have described the existence of at least 30 nonoverlapping cDNAs (c*Kox*1 to c*Kox*30) encoding the zinc finger motif consensus sequence Y/FxCxxCxxxFxxxxxLxxHxxx HTGEK/RP, most of them representing distinct genes. These cDNAs were derived from libraries generated from human T-cell lines. We expect that some of the *Kox* genes might be related to differentiation processes during the generation of mature T lymphocytes. In my laboratory we are currently exploiting the advantages of working with individual T-cell clones that represent different subsets of differentiated T cells to ask whether zinc finger genes differ in their expression pattern when thymocytes differentiate to peripheral T cells. Preliminary data seem to indicate that *Kox*27 is predominantly expressed in the B- and T-cell lineage (data not shown). Evidence for the involvement of zinc finger genes in hematopoietic development comes from the murine system where an ectopic expression of a zinc finger gene (*Evi*1) induced by retroviral insertion prevented Il-3–dependent myeloid cell lines from terminally differentiating (Morishita et al., 1988). Furthermore, the zinc finger gene cHF.12, whose finger region is identical to that in Kox25, was reported to be downregulated during myelomonocytic differentiation (Pannuti et al., 1988).

Recently, Kongsuwan et al. (1988) reported the detection of more than 20 transcription factors encoding homeobox domains in hematopoietic cells. It now appears that at least two families of transcription factors encompassing two highly conserved DNA-binding motifs (homeobox and zinc fingers) are present in hematopoietic cells, and it seems likely that some of them are involved in the commitment and differentiation of these cells.

Comparison of Zinc Finger Genes

Sequence comparisons with published zinc finger structures revealed that the *Kox* genes (with the possible exceptions of c*Kox*10 and c*Kox*25) are new representatives of the human zinc finger family. Partial sequences available to us indicate identity between c*Kox*10 and HPFp3 (Bellefroid et al., 1989) and between c*Kox*25 and cHF.12 (Pannuti et al., 1988). However, since we do not have complete cDNA sequences we cannot rule out some sequence divergence outside of the homologous regions we have identified. We estimate that approximately 70 zinc finger genes are expressed in human T cells. Thirty nonoverlapping cDNAs (c*Kox*1 to c*Kox*30) have been reported in this paper, supporting the notion that zinc finger motifs have been exploited frequently during evolution. Thus far, about 20 human zinc finger genes have been identified as cDNAs or genomic clones (Kadonaga et al., 1987; Kinzler et al., 1988; Page et al., 1987; Joseph et al., 1988; Pannuti et al., 1988; Ruppert et al., 1988; Sukhatme et al., 1988; and Bellefroid et al., 1989). We have not obtained any zinc finger genes that code for only two zinc fingers as in ADR1 (Hartshorne et al., 1986) or three as in SP1 (Kadonaga et al., 1987) and NGR-1a (Milbrandt, 1987). In addition, the zinc finger repeats of the *Kox* family did not encode four amino acids between the conserved cysteine residues identified in SP1 and in the GLI family (Ruppert et al., 1988). The alternative spacing of three and four residues between the histidine residues observed in TDF (Page et al., 1987) were also not found in our zinc finger collection.

Does a Zinc Finger–Specific Recognition Code Exist?

Comparison of individual Kox1 fingers (Fig. 1B) with fingers from the proteins encoded by the c*Kox* family (Fig. 7) established that different zinc finger proteins can have identical or related DNA binding regions if one assumes that these fingers bind similar sequences with the same affinity or the same sequence with different affinities. If the postulated α-helical region can be shown to be the DNA recognition site, it will be interesting to see whether individual positions in the α-helix (every third position, for example, the consensus in Fig. 8) or whether the neighboring amino acids (pairs or triplets) are important in determining the specificity. It is possible that careful analysis of zinc finger structures might lead to the discovery of a zinc finger–specific recognition code. Perhaps finger swap experiments could be designed to see whether a recognition code can be extracted by determining target preferences for conserved doublets, such as IR and IL, or conserved triplets, such as SSH, SSA, and SSR (Fig. 8). In this laboratory we envisage the existence of a zinc finger–specific DNA recognition code on the basis of the following assumptions: (i) Individual fingers have the potential to bind DNA (Parraga et al., 1988); (ii) conserved H/C links position α-helical regions of zinc fingers in the major groove of a distorted B-form of DNA (Gottesfeld et al., 1987); (iii) amino acids at specific positions within the postulated α-helix of the zinc finger determine the sequence-specific DNA binding (Blumberg et al., 1987); (iv) the specificity of multiple fingers may be determined by knowing the specificities of each individual finger; and (v) knowing the position of the putative α-helix in regard to the DNA should make it possible to determine specific amino acid–nucleotide interactions.

The existence of a DNA recognition code has been postulated previously but not yet substantiated (Seeman et al., 1976; Pabo and Sauer, 1984). Besides the lack of appropriate methods, mutations within DNA-binding motifs run the risk of destroying the overall conformation necessary for DNA recognition. Summarizing current knowledge, one can view the zinc finger domain as a rigid structure where changes of amino acids that are not highly conserved do not dramatically influence the overall structure determined by zinc coordination and the hydrophobic interactions between phenylalanine and leucine. Each individual finger domain found in zinc finger proteins can be considered as one mutant of the zinc finger consensus. Thus, the zinc finger structure can be interpreted as a passe-partout: Depending on the amino acids in the putative recognition helices, specific amino acids interact with the DNA with variable affinities. The highly conserved H/C link provides the appropiate spacing to position the α-helical region in the major groove of DNA. These considerations suggest that zinc finger genes may bind to quite diverse target sequences. Each finger may be considered to have the potential for high-, medium-, and low-affinity binding.

CONCLUSIONS

Thirty nonoverlapping cDNAs (c*Kox*1 to c*Kox*30) encoding zinc finger domains have been isolated from human T-cell lines. Screening results indicate that a high number of zinc finger proteins are expressed in human tissues. Sequence comparisons and Southern blots with the *Kox* genes as probes indicate that various zinc finger genes can be further classified into subfamilies. We are interested in determining whether some *Kox* genes are clustered in the genome and whether the expression pattern of some of them is restricted to specific cell lineages within the hematopoietic differentiation system. Taking c*Kox*1 as one representative of the zinc finger family, we have demonstrated that Kox1 is expressed in various hematopoietic and nonhematopoietic cell lines and that the finger region of Kox1 incorporates zinc and binds to double-stranded DNA (H.-J. Thiesen, unpublished data). It is also possible that c*Kox*1 encodes a protein domain outside the finger region that facilitate protein-protein interactions.

In view of the high number of zinc finger genes and the postulated complexity of gene regulation, it can be assumed that nuclear transcription factors constitute a network of factors with activating and repressing functions. It will now be important to identify target genes of the c*Kox* family in order to evaluate the influence of the Kox proteins on gene expression. Thus far, target sites and target genes are known for SP1 (Dynan and Tjian, 1983) and GLI (Kinzler and Vogelstein, 1990) in humans, for Zif268 (Christy and Nathans, 1989) in mouse, for ADR1 (Blumberg et al., 1987) and SWI5 (Stillman et al., 1988) in yeast, for Krüppel (Stanojevic et al., 1989) and suppressor of Hairy-wing protein in *Drosophila* (Spana et al., 1988) and for TFIIIA (Engelke et al., 1980) in *Xenopus laevis*.

Since the genomes from *Drosophila* to human harbor hundreds of zinc finger genes with multiple consecutive fingers we believe that the finger structure presents a highly conserved protein domain that might be described as a passe-partout for DNA binding. If certain rules for zinc finger–specific DNA-protein interactions could be generated, it should be possible, theoretically, to determine consensus target sites for zinc finger proteins and, even more exciting, to generate zinc finger domains for desired DNA target sites.

We have been able to demonstrate sequence-specific binding of Kox1 protein to double-stranded oligonucleotides selected by the target detection assay (TDA) (H.-J. Thiesen, in preparation). This TDA method has

been designed to determine target sites for putative DNA-binding proteins by using randomized oligonucleotides in combination with the amplification steps of the polymerase chain reaction (Thiesen, in press). Further characterization of zinc finger genes and their expression patterns should help to elucidate their mechanisms of action.

MATERIAL AND METHODS

Cell Culture and DNA and RNA Extraction

The cell lines Jurkat and Molt-4 (T-cell lines), Ramos (B-cell line), SK-Mel (melanoma cell line), U937 (myelomonocytic cell line), HT29 (colon carcinoma cell line), OVCA (ovarian carcinoma cell line), and HeLa (adenocarcinoma cell line) were obtained from the American Type Culture Collection and grown as recommended. Total RNA was isolated from cell lines according to the guanidium thiocyanate procedure of Chirgwin and collaborators (1979), as modified by Maniatis and collaborators (1982). Poly(A)$^+$ RNA was selected by mAP paper (Orgenics, Yavne, Israel), using 100 μg of total RNA.

cDNA Library Screening

A pEX-Molt-4 cDNA library (Almendral et al., 1987) was prehybridized (1 h) and hybridized with the finger region of *mkr*1 (kindly supplied by P. Gruss) (>10^9 cpm/μg of DNA) at 50°C in 0.5 M sodium phosphate, pH 7.2, 7% SDS, and 1 mM EDTA for 20 h. Washing was done in 40 mM sodium phosphate, pH 7.2, 1% SDS, for 30 min (2×) at 50°C. Positive clones appearing in duplicate were rescreened, isolated, and transferred to pUC19 plasmids. We obtained the clones c*Kox*2-29 by screening a Jurkat cDNA library (purchased from Clontech) and cKox30 by screening a Molt-4 library with c*Kox*1 (hybridization as done for the *mkr*1 probe). Positive inserts were cloned into Bluescript and nucleotides encoding finger regions were determined by sequencing.

DNA Sequencing and Sequence Analysis

The *Kox* cDNAs were cloned and subcloned in Bluescript and pUC19. Both strands were sequenced using universal and internal primers by the Sequenase™ procedure (USB). Nucleotide and amino acid sequence analyses were carried out using the Wisconsin GCG Sequence Analysis Software Package.

Southern and Northern Blots

Genomic Southern blots were done with 10 μg of genomic Molt-4 DNA that was digested with BamH1, EcoRI, or HindIII, eletrophoresed on 0.7% agarose, and transferred to GeneScreen membrane (Dupont). The DNA was then UV cross-linked, baked for 2 h at 80°C, and prehybridized (2 h) and hybridized (24 h) in 0.5 M sodium phosphate, pH 7.2, 7% SDS, and 1 mM EDTA with ^{32}P-oligo-labeled c*Kox*1-30 at 67°/69°C. As c*Kox* probes, complete inserts of isolated *Kox* cDNA were used. In the case of c*Kox*1, a probe containing the first 660 nucleotides (Fig. 1A) was used. Washings were done in 40 mM Na_iPO_i, pH 7.2, and 1% SDS at 67°/70°C twice for 30 min. For the Northern blots, poly(A)$^+$ RNA was extracted from 100 μg of total RNA with the use of mAP paper (Orgenics, Yavne, Israel) and then loaded on a 1% agarose-formaldehyde gel, electrophoresed, and transferred to nitrocellulose. Hybridization was done as above. Filters were washed twice at 65°C in 0.1 × SSC containing 1% SDS for 30 min. Kodak XAR5 film was exposed for 1 to 7 days at −70°C in the presence of an intensifying screen. For rehybridization with GAPDH the filters were stripped of the probe by washing for 1 h at 80°C in 1 mM Tris (pH 8.0), 0.1 mM EDTA, and 0.05% SDS.

Zinc Blotting

c*Kox*1 was expressed as a fusion protein in pUEX 3 (Bressan and Stanley, 1987), an overnight culture was diluted 1:100 in LB-medium (final volume 5 ml), grown at 28°C to an optical density of 0.2 (600 nm), and then induced at 42°C for 2 h. The *E. coli* pellet was taken up in 400 μl of SDS-Laemmli sample buffer and 25 μl was run on a 7.5% SDS-polyacrylamide gel (Laemmli, 1970). One half of the gel was stained with Coomassie brilliant blue, the other half transferred to nitrocellulose and a zinc blot performed according to Schiff et al., 1988.

Accession Numbers

The nucleotide sequence data reported here will appear in the EMBL, GenBank, and DDBJ nucleotide sequence databases under the following accession numbers:

Human *Kox1* cDNA to human *Kox10* cDNA, numbers X52332 to X52341.

Human *Kox11* cDNA to human *Kox20* cDNA, numbers X52342 to X52351.

Human *Kox21* cDNA to human *Kox30* cDNA, numbers X52352 to X52361.

Acknowledgments

I would like to thank P. Gruss for kindly supplying the *mkr*1 probe, R. Bravo for the Molt-4 library, K. Stanley for pUEX plasmids, B. Wicki for technical assistance, H. Kiefer for synthesis of sequencing oligonucleotides, and A. Tunnacliffe and S.R. Bauer for critical reading of the manuscript. The Basel Institute for Immunology was founded and is supported by F. Hoff-

mann-La Roche & Co. Ltd., CH-4005 Basel, Switzerland.

REFERENCES

Almendral JM, Huebsch D, Blundell PA, MacDonald-Bravo H, Bravo R (1987): Cloning and sequence of the human nuclear protein cyclin: homology with DNA-binding proteins. Proc Natl Acad Sci USA 84:1575-1579

Bellefroid EJ, Lecocq PJ, Benhida A, Poncelet DA, Belayew A, Martial JA (1989): The human genome contains hundreds of genes coding for zinc finger proteins of the Krüppel type. DNA 8:377-387

Berg J (1988): Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc Natl Acad Sci USA 85:99-102

Berg J (1989): DNA binding specificity of steroid receptors. Cell 57:1065-1068

Blumberg H, Eisen A, Sledziewski A, Bader D, Young ET (1987): Two zinc fingers of a yeast regulatory protein shown by genetic evidence to be essential for its function. Nature 328:443-445

Bopp D, Burri M, Baumgartner S, Frigerio G, Noll M (1986): Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. Cell 47:1033-1040

Bressan GM, Stanley KK (1987): pUEX, a bacterial expression vector related to pEX with universal host specificity. Nucleic Acids Res 15:10056

Brown RS, Sander C, Argos P (1985): The primary structure of transcription factor TFIIIA has 12 consecutive repeats. FEBS Lett 186:271-274

Brown RS, Argos P (1986): Fingers and helices. Nature 324:215

Chirgwin JM, Przybyla AE, McDonald RJ, Rutter WJ (1979): Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294-5299

Chowdhury K, Deutsch U, Gruss P (1987): A multiple family encoding several "finger" structures is present and differentially active in mammalian genomes. Cell 48:771-778

Christy B, Nathans D (1989): DNA binding site of the growth factor inducible protein Zif268. Proc Natl Acad Sci USA 86:8737-8741

Courey JA, Tjian R (1988): Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell 55:887-898

Desplan C, Theis J, O'Farrell PH (1988): The sequence specificity of homeodomain-DNA interaction. Cell 54:1081-1090

Deutsch U, Dressler GR, Gruss P (1988): Pax 1, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. Cell 53:617-625

Dexter TM, Spooner E (1987): Growth and differentiation in the hematopoietic system. Annu Rev Cell Biol 3:423-441

Diakun GP, Fairall L, Klug A (1986): EXAFS study of the zinc-binding sites in the protein transcription factor IIIA. Nature 324:689-691

Dynan WS, Tjian R (1983): Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell 32:669-680

Engelke DR, Ng S-Y, Shastry BS, Roeder RG (1980): Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell 19:717-728

Fairall L, Rhodes D, Klug A (1986): Mapping of the sites of protection on a 5S RNA gene by the Xenopus transcription factor IIIA. A model for the interaction. J Mol Biol 192:577-591

Frankel AD, Berg JM, Pabo CO (1987): Metal-dependent folding of a single zinc finger from transcription factor IIIA. Proc Natl Acad Sci USA 84:4841-4845

Gehring WJ (1987): Homeo boxes in the study of development. Science 236:1245-1252

Gentz R, Rauscher III FJ, Abate C, Curran T (1989): Parallel association of Fos and Jun Leucine zippers juxtaposes DNA binding domains. Science 243:1695-1699

Gibson TJ, Postma JPM, Brown RS, Argos P (1988): A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins. Protein Engineering 2:209-218

Ginsberg AM, King BO, Roeder RG (1984): Xenopus 5S gene transcription factor, TFIIIA: Characterization of a cDNA clone and measurement of RNA levels throughout development. Cell 39:479-489

Gottesfeld JM, Blanco J, Tennant LL (1987): The 5S gene internal control region is B-form both free in solution and in complex with TFIIIA. Nature 329:460-462

Hartshorne TH, Blumberg H, Young ET (1986): Sequence homology of the yeast regulatory protein ADR1 with Xenopus transcription factor TFIIIA. Nature 320:283

Hoey T, Levine M (1988): Divergent homeo box proteins recognize similar DNA sequences in Drosophila. Nature 332:858-861

Holland PWH, Hogan BLM (1988): Expression of homeo box genes during mouse development: a review. Genes & Dev 2:773-782

Hope IA, Struhl K (1986): Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885-894

Hope IA, Mahadevan S, Struhl K (1988): Structural and functional characterization of the short acidic transcriptional activation of yeast GCN4 protein. Nature 333:635-640

Joseph LJ, Le Beau MM, Jamieson GA, Acharya S, Shows JD, Rowley JD, Sukhatme VP (1988): Molecular cloning, sequencing, and mapping of EGR2, a human early response gene encoding a protein with "zinc-binding finger" structure. Proc Natl Acad Sci 85:7164-7168

Kadonaga JT, Carner KR, Masiar FR, Tjian R (1987): Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell 51:1079-1090

Kinzler KW, Ruppert JM, Bigner SH, Vogelstein B (1988): The GLI gene is a member of the Krüppel family of zinc finger proteins. Nature 332:371-374

Kinzler KW, Vogelstein B (1990): The GLI gene encodes a nuclear protein which binds specific sequences in the human genome. Mol Cell Biol 10:634-642

Kongsuwan K, Webb E, Housiaux P, Adams JM (1988): Expression of multiple homeobox genes within diverse mammalian haemopoietic lineages. EMBO J 7:2131-2138

Laemmli UK (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685

Landschulz WH, Johnson PF, McKnight SL (1988): The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science 240:1759-1764

Lee MS, Gippert GP, Soman KV, Case DA, Wright PE (1989): Three-dimensional solution structure of a single zinc finger DNA-binding domain. Science 245:635-637

Lichtsteiner S, Wuarin J, Schibler U (1987): The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell 51:963-973

Ma J, Ptashne M (1987): A new class of yeast transcriptional activators. Cell 51:113-119

Maniatis T, Fritsch EF, Sambrook J (eds) (1982): Molecular Cloning, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory

Maniatis T, Goodburn S, Fischer JA (1987): Regulation of inducible and tissue-specific gene expression. Science 236:1237-1245

Manley JL, Levine MS (1985): The homeobox and mammalian development. Cell 43:1-2

McMichael AJ (ed) (1987): Leucocyte typing III. Oxford, Oxford University Press

Metcalf D (1989): The molecular control of cell division, differentiation commitment and maturation in haemopoietic cells. Nature 339:27-30

Milbrandt J (1987): A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. Science 238:797-799

Miller J, McLachlan AD, Klug A (1985): Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J 4:1609-1614

Mitchell PJ, Tjian R (1989): Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245:371-378

Morishita K, Parker DS, Mucenski ML, Jenkins NA, Copeland NG, Ihle JN (1988): Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. Cell 54:831-840

O'Shea EK, Rutkowski R, Kim PS (1989): Evidence that the leucine zipper is a coiled coil. Science 243:538-542

Pabo CO, Sauer RT (1984): Protein-RNA recognition. Annu Rev Biochem 53:293-321

Page DC, Mosher R, Simpson EM, Fisher EMC, Mardom G, Pollack J, McGillivray B, de la Chappelle A, Brown LG (1987): The sex-determining region of the human Y chromosome encodes a finger protein. Cell 51:1091-1104

Pannuti A, Lanfrancone L, Pascucci A, Pelicci P-G, La Mantia G, Lania L (1988): Isolation of cDNAs encoding finger proteins and measurement of the corresponding mRNA levels during myeloid terminal differentiation. Nucleic Acids Res 16:4227-4237

Párraga G, Horvath SJ, Eisen A, Taylor WE, Hood L, Young ET, Klevit RE (1988): Zinc-dependent structure of a single-finger domain of yeast ADR1. Science 241:1489-1492

Piechaczyk M, Blanchard JM, Dani C, Panabieres F, El Sabouty S, Fort P, Jeanteur P (1984): Post-transcriptional regulation of glyceraldehyde-3-phosphate-dehydrogenase gene expression in rat tissues. Nucleic Acids Res 12:6951-6963

Pieler T, Hamm J, Roeder RG (1987): The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. Cell 48:91-100

Pieler T, Oei SL, Hamm J, Engelke U, Erdmann VA (1985): Functional domains of the Xenopus laevis 5S gene promoter. EMBO J 4:3751-3756

Ptashne M (1988): How eukaryotic transcriptional activators work. Nature 335:683-689

Rhodes D, Klug A (1986): An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA. Cell 46:123-132

Rosenberg UB, Schröeder C, Priess A, Kienlin A, Côte S, Riede I, Jäckle H (1986): Structural homology of the product of the Drosophila Krüppel gene with Xenopus transcription factor IIIA. Nature 319:336-339

Ruppert JM, Kinzler KW, Wong AJ, Bigner SH, Kao F-A, Law ML, Seuanez HN, O'Brien SJ, Volgelstein B (1988): The GLI-Krüppel family of human genes. Mol Cell Biol 8:3104-3113

Saenger, W (ed) (1984): Principles of Nucleic Acid Structure. New York, Springer-Verlag

Sassone-Corsi P, Ransone LJ, Lamph WW, Verma IM (1988): The direct interaction between fos and jun nuclear oncoproteins: role of the 'leucine zipper' domain. Nature 336:692-695

Schiff LA, Nibert ML, Fields BN (1988): Characterization of a zinc blotting technique: Evidence that a retroviral gag protein binds zinc. Proc Natl Acad Sci USA 85:4195-4199

Seeman NC, Rosenberg JM, Rich A (1976): Sequence-specific recognition of double helical nucleic acids by proteins. Proc Natl Acad Sci USA 7:804-805

Smotkin D, Wettstein O (1987): The major human papillomavirus protein in cervical cancers is a cytoplasmic phosphoprotein. J Virol 61:1686-1689

Spana C, Harrison DA, Corces VC (1988): The Drosophila melangaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. Genes Dev 2:1414-1423

Stanojevic D, Hoey T, Levine M (1989): Sequence-specific DNA-binding activities of the gap proteins encoded by hunchback and Krüppel in Drosophila. Nature 341:331-335

Stillman DJ, Bankier AT, Seddon A, Groenhout EG, Nasmyth KA (1988): Characterization of a transcription factor involved in mother cell specific transcription of the yeast HO gene. EMBO J 7:485-494

Strominger JL (1989): Developmental biology of T cell receptors. Science 244:943-950

Sukhatme VP, Cao X, Chang LC, Tsai-Morris CH, Stamenkovich D, Ferreira PCP, Cohen DR, Edwards SA, Shows TB, Curran T, LeBeau M, Adamson ED (1988): A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization. Cell 53:37-43

Tautz D, Lehmann R, Schnürch H, Schuh R, Seifert E, Kienlin A, Jones K, Jäckle H (1987): Finger protein of novel structure encoded by hunchback, a second member of the gap class of Drosophila segmentation genes. Nature 327:383-389

Thiesen H-J: Target detection assay (TDA): a general method to determine DNA binding sites for putative DNA binding proteins. (in press) in Lefkovits I, Pernis B (eds) Immunological Methods Vol. IV, Orlando, FL, Academic Press

Turner R, Tjian R (1989): Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cfos-cjun heterodimers. Science 243:1689-1694