Molecular cloning and characterization of chemokine-like factor 1 (CKLF1), a novel human cytokine with unique structure and potential chemotactic activity

Wenling HAN*, Yaxin LOU†, Junmin TANG‡, Yingmei ZHANG*, Yingyu CHEN*, Ying LI*, Weifeng GU*, Jiaqiang HUANG*, Liming GUI*, Yan TANG‡, Feng LI‡, Quansheng SONG*, Chunhui DI*, Lu WANG†, Qun SHI*, Ronghua SUN†, Donglan XIA*, Min RUI†, Jian TANG† and Dalong MA*†[1]

*Laboratory of Medical Immunology, School of Basic Medical Science, Peking University, 38 Xueyuan Road, Beijing, 100083 China, †Center for Human Disease Genomics, Peking University, 38 Xueyuan Road, Beijing, 100083 China, and ‡Laboratory of Histology, School of Basic Medical Science, Peking University, 38 Xueyuan Road, Beijing, 100083 China

Cytokines are small proteins that have an essential role in the immune and inflammatory responses. The repertoire of cytokines is becoming diverse and expanding. Here we report the identification and characterization of a novel cytokine designated as chemokine-like factor 1 (CKLF1). The full-length cDNA of *CKLF1* is 530 bp long and a single open reading frame encoding 99 amino acid residues. CKLF1 bears no significant similarity to any other known cytokine in its amino acid sequence. Expression of CKLF1 can be partly inhibited by interleukin 10 in PHA-stimulated U937 cells. Recombinant CKLF1 is a potent chemoattractant for neutrophils, monocytes and lymphocytes; moreover, it can stimulate the proliferation of murine skeletal muscle cells. These results suggest that CKLF1 might have important roles in inflammation and in the regeneration of skeletal muscle.

Key words: interleukin 10, skeletal-muscle cells, splicing, suppression–subtractive hybridization.

INTRODUCTION

Cytokines consist of a broad group of small proteins that have a crucial role in the immune and inflammatory responses. In general, cytokines are induced transiently on cell activation and are regulated by many factors. Cytokines include interleukins, interferons, tumour necrosis factors, growth factors and chemokines [1]. Among them, the chemokines are a family of small, structurally related proteins. They have a major role in the recruitment and activation of leucocytes *in vivo*. Chemokines share a conserved four-cysteine motif in their amino acid sequences and can be divided into four families on the basis of the position of the first two cysteine residues: CXC (α-chemokine), CC (β-chemokine), C (γ-chemokine) and CX3C (δ-chemokine) [2].

It has been known for a long time that U937 cells, a human promonocytic cell line, can produce many kinds of cytokines and that interleukin 10 (IL-10) can inhibit the expression of cytokines [3,4]. On the basis of these two ideas, we devised a strategy to find novel cytokines. Using the suppression–subtractive hybridization (SSH) method [5], we have isolated cDNA species that could be inhibited by IL-10 in phytohaemagglutinin (PHA)-stimulated U937 cells and have identified a novel secreted cytokine chemokine-like factor 1 (CKLF1) containing a CC motif and having chemotactic effects on leucocytes. Nevertheless, this cytokine differs from classical chemokines in the following ways: (1) there is only one conserved CC motif in the mature protein and no additional C residue in its C-terminal region; (2) it has no significant similarity to other chemokines; (3) at least three potential mRNA splicing isoforms exist; and (4) it has stimulatory effects on skeletal muscle cells. CKLF might represent a new protein family that is different from classical chemokine; we therefore designated it as CKLF1, and its isoforms as CKLF2, CKLF3 and CKLF4.

MATERIALS AND METHODS

SSH

SSH was performed with Clontech PCR-Select™ cDNA Subtraction Kit (Clontech) described by Diatchenko et al. [5]. SSH is a powerful technique that combines the traditional subtractive hybridization method with suppression PCR. It can be used to compare two populations of mRNA and to obtain clones of genes that are expressed in one population but are lower or not expressed in the other. Both mRNA populations are converted into cDNA. We refer to the cDNA that contains specific (differentially expressed) transcripts as the 'tester' and the reference cDNA as the 'driver'. Tester and driver cDNA are hybridized; the hybrid sequences are then removed. The remaining non-hybridized cDNA species represent genes that are expressed in the tester population but are lower or not expressed in the driver. Furthermore, suppression PCR prevents undesirable amplification while the enrichment of target molecules proceeds. In the SSH procedure, human promonocytic cell line U937 cells were maintained in RPMI 1640 medium containing 10% (v/v) fetal bovine serum, 2 nM glutamine, 100 i.u./ml penicillin and 100 i.u./ml streptomycin; the U937 cells, stimulated with 10 μg/ml PHA (Sigma) for 8 h, were used as the tester, whereas the U937 cells with the same treatment in the presence of 100 i.u./ml IL-10 (Sigma) were used as the driver. cDNA was synthesized from 2 μg of mRNA as the tester

Abbreviations used: ANAE, acid α-naphthyl acetate esterase; CKLF, chemokine-like factor; EGFP, enhanced green fluorescent protein; EST, expressed sequence tag; HE, haematoxylin–eosin; IL-10, interleukin 10; PHA, phytohaemagglutinin; POX, peroxidase; RT–PCR, reverse-transcriptase-mediated PCR; SSH, suppression–subtractive hybridization; STCP-1, stimulated T-cell chemotactic protein 1; TARC, thymus and activation-regulated chemokine.

[1] To whom correspondence should be addressed (e-mail masong@mail.bjmu.edu.cn).

(without IL-10) or the driver (with IL-10) and was digested with *Rsa*I. Two types of adapter, provided by the manufacturer, were independently ligated to the tester cDNA species. First and second hybridizations were performed and the resulting annealed material was used as the PCR template. Hybridization mixture (1 μl; diluted 1:1000) was amplified for 27 primary PCR cycles, with the primers provided by the manufacturer; 15 cycles of the secondary PCR were then done with primary PCR products (diluted 1:10) as template. PCR reactions were performed with the Advantage PCR System (Clontech). PCR products were ligated into the pGEM-T easy vector (Promega) and transformed into XL-1 Blue bacteria; white colonies were screened. This led to the establishment of approx. 100 cDNA clones, of which 15, chosen at random, were sequenced with ALF Expressed Sequencer in accordance with the protocol (Pharmacia); the similarity of the sequences with known fragments was compared on the World Wide Web (National Center for Biotechnical Information, http://www.ncbi.nlm.nih.gov).

Bioinformatics

A number of human ESTs were initially identified with an expressed sequence tag (EST) fragment obtained from the PHA-stimulated U937 subtracted library by running a BLASTN search against the public database of ESTs (dbEST). These ESTs were retrieved and a contig was assembled from overlapping ESTs by using EST Assembly Machine on the World Wide Web (http://www.tigem.it) [6]. The GenBank® accession numbers of the ESTs used for CKLF1 assembly were W38899, N95062, AA429945, AA987264, AA927461, N89912, AA516431, AA479657, AA455042, AA989129, AL044098, AF151058, NM_016951 and W52820. A full-length cDNA sequence was obtained and examined by reverse-transcriptase-mediated PCR (RT–PCR) in PHA-stimulated U937 cells. The putative cleavage site of the signal peptide was predicted by using Signal P server (http://www.cbs.dtu.dk/services/SignalP). The gene localization of *CKLF1* on the chromosome was analysed with the HTGS database (http://www.ncbi.nlm.nih.gov). The accession numbers of the chromosome fragments were AC010542 and AC018557.

Northern blot analysis

The expression of potentially interesting cDNA clones was studied by using Northern blot analysis. Total RNA was extracted with Trizol (Life Technologies) in accordance with the supplier's instructions. Approximately 20 μg of total RNA of tester or driver was fractionated on a 1.5% (w/v) agarose/formaldehyde gel, then transferred to GeneScreen-plus nylon membrane (DuPont–NEN, Boston, MA, U.S.A.) by capillary transfer. The probe for the *CKLF1* sequence used for Northern hybridization was the EST fragment obtained from SSH (bases 53–530 of the *CKLF1* cDNA). Labelling of the probe and hybridization were performed with the the protocol of the Random Primer Fluorescent Labeling Kit with anti-(fluorescent-HRP) (DuPont–NEN) provided by the company.

Oligonucleotides

The primers for amplifying the full-length cDNA sequence of CKLF1 and its variants were 5′-GCA AGA AGC GGG AAG CCG A-3′ (P1) and 5′-GGA AGA ATA CAG AAA TAT GTT TAA TAC-3′ (P2). Primers for amplifying the coding region of CKLF1 to construct pCDI-CKLF1 were 5′-ATG GAT AAC GTG CAG CCG AAA AT-3′ (P3) and 5′-TTA CAA AAC TTC TTT TTT TTC ATG C-3′ (P4). The 3′ end primer for amplifying the fragments to construct pEGFP-N1-CKLF1 (in which EGFP stands for enhanced green fluorescent protein), pEGFP-N1-CKLF2 and pEGFP-N1-CKLF4 was 5′-CGG GAT CCA AAA CTT CTT TTT TTT CAT GC-3′ (P5).

CKLF1 expression pattern analysis

The expression pattern of *CKLF1* across different tissues was analysed by using nested PCR with the one-strand cDNA libraries in the Multiple Tissue cDNA Panel kit (Clontech) by PCR primers P1/P2 and P3/P4. Template (1 μl) in a 25 μl reaction volume was amplified for 25 primary PCR cycles with primers P1 and P2; 20 cycles of secondary PCR were performed with 1 μl of primary PCR products as template in a 25 μl reaction system. A 10 μl sample of the second PCR products was analysed by SDS/PAGE [5% (w/v) gel].

Plasmid constructions

For analysis of the subcellular localization of CKLF1, CKLF2 and CKLF4, the full-length coding sequences were cloned in frame into the pEGFP-N1 expression vector (Clontech). The coding regions of *CKLF1*, *CKLF2* and *CKLF4* were amplified from the cDNA library of PHA-stimulated U937 cells, using primers P3 and P5, in which the stop codon was removed and a *Bam*H1 site was introduced at 3′ of the *CKLF1*, *CKLF2* or *CKLF4* coding region. The PCR products were blunted with Klenow enzyme and then cut with *Bam*HI. The pEGFP-N1 expression vector was cut with *Eco*RI, blunted with Klenow enzyme and then digested with *Bam*HI. After recovery from gels, the fragments were ligated into the treated vectors and the recombinant plasmids were designated pEGFP-N1-CKLF1, pEGFP-N1-CKLF2 and pEGFP-N1-CKLF4 respectively. With regard to the construction of pCDI-CKLF1, the coding region including the stop codon was amplified from the same cDNA library as described above with primers P3 and P4, then cloned into pGEM-T-easy vector (Promega). The insert was released by cutting with *Eco*RI and subcloned into the *Eco*RI site of the mammalian expression vector pCDI, which replaces the *Bgl*II–*Kpn*I fragment of pCDNA3 (Invitrogen) with the *Bgl*II–*Kpn*I fragment of pCI [7]. The sense pCDI-CKLF1 expression vector was selected by restriction enzyme analysis.

Transfection and transfectants

For transient transfections, COS-7 cells were maintained in Dulbecco's modified Eagle's medium with 10% (v/v) fetal calf serum and 4 mM glutamine. The cells were transfected with Superfect Transfection Reagent (Qiagen) in accordance with the instructions of the manufacturer. After 48 h the supernatants of the cells transfected with EGFP constructs were collected for Western blot analysis. The cells were washed twice in ice-cold PBS and then fixed in 4% (w/v) paraformaldehyde for 30 min. They were washed again in PBS and examined with a fluorescence microscope. The supernatants of the COS-7 cells transfected with pCDI and pCDI-CKLF1 were collected for bioassay 48 h after transfection.

Western blot analysis of CKLF1–EGFP, CKLF2–EGFP and CKLF4–EGFP in the supernatants of COS-7 cell culture

The supernatants of COS-7 cells transfected with the four different EGFP constructs were collected 48 h after transfection; 70 μl of each was subjected to SDS/PAGE [12.5% (w/v) gel] and transferred electrophoretically to nitrocellulose membrane. The membranes were reacted with EGFP polyclonal antibody (Clontech) and then with HRP-conjugated anti-IgG; the signal

was detected by enhanced chemiluminescence (Pierce, Rockford, IL, U.S.A.).

Chemotaxis assays

Human neutrophils, monocytes and lymphocytes were isolated from peripheral blood of young healthy volunteers as described previously [8]. Purified neutrophils were resuspended at 10^6 cells/ml in Hanks balanced salt solution; monocytes and lymphocytes were resuspended at 2×10^6 and 5×10^6 cells/ml respectively in RPMI 1640 medium (Life Technologies, Grand Island, NY, U.S.A.) supplemented with 0.5% low-endotoxin BSA (Sigma) and 20 mM Hepes. Cell migration was measured in a 48-well chemotaxis chamber (Neutroprobe; Cabin John, MD, U.S.A.) as described by Sarafi et al. [9]. The lower wells were filled with the supernatants of COS-7 cells transfected with pCDI or pCDI-CKLF1. All samples were assayed twice in triplicate. The PVDF-free polycarbonate membranes (Poretics, Livermore, CA, U.S.A.) for neutrophils were 3 μm in pore size and the incubation time was 30 min, and for monocytes and lymphocytes 5 μm pore-size membranes and 90 min were used.

DNA injection and preparation of muscle sections

All experiments involving animals were approved by the Peking University Animal Care and Use Committee. BALB/c mice 6–8 weeks old were injected with 100 μl of pyrogen-free physiological saline containing 100 μg of pCDI-CKLF1 or pCDI into the anterior tibial muscle, and an electric pulse (40 ms, 80 V) was generated to enhance the DNA uptake [10]. After 10 days the mice were killed and the injection sites were excised. Parts of the muscle biopsies were processed by routine histology, stained with Feulgen, acid α-naphthyl acetate esterase (ANAE), peroxidase (POX) and haematoxylin–eosin (HE) and observed microscopically. Other parts of muscle biopsies were prepared for the extraction of total RNA with Trizol (Life Technologies) for RT–PCR.

RESULTS

Development of differentially expressed cDNA sequences

With the use of the SSH method we have isolated genes whose expression could be inhibited by IL-10 in PHA-stimulated U937 cells. Screening the sequences of the first 15 clones containing inserts of more than 200 bp in the GenBank Database revealed that six clones including *CKLF1* might be fragments of novel genes (results not shown). Northern blot analysis of *CKLF1* demonstrated that it could be inhibited by IL-10 and that its mRNA was approx. 0.6 kb long (Figure 1).

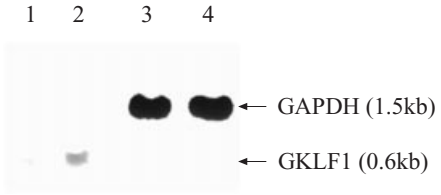

Figure 1 Northern blot analysis of CKLF1 expression

Lanes 2 and 4, total RNA from PHA-stimulated U937 cells; lanes 1 and 3, total RNA from U937 cells with the same treatment in the presence of IL-10. CKLF1 expression in PHA treated U937 cells can be partly inhibited by IL-10 at transcription level (lanes 1 and 2). Abbreviation: GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

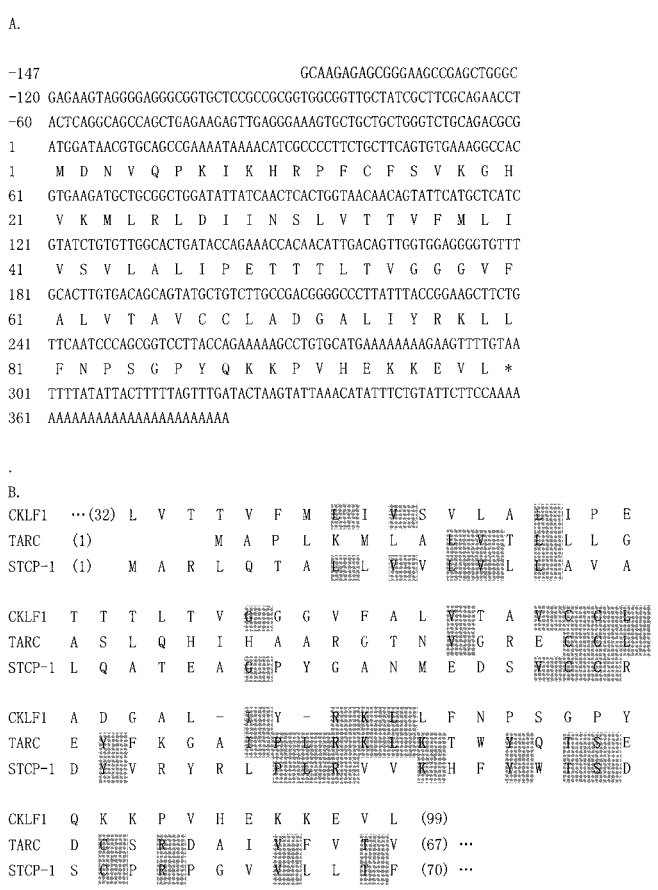

Figure 2 Nucleotide sequence and predicted amino acid sequence of human *CKLF1* and sequence alignment of human CKLF1 with TARC and STCP-1

(**A**) Nucleotide and deduced amino acid sequences of *CKLF1*. The nucleotide sequence of the message strand is numbered in the 5′-to-3′ direction. The predicted amino acid sequence is shown below the nucleotide sequence. The stop codon is indicated by an asterisk. (**B**) Alignment of the deduced amino acid sequence of CKLF1 with the CC-chemokines TARC and STCP-1. Conserved amino acids are shaded.

Isolation and characterization of the full length of *CKLF1* cDNA and encoding protein sequences

The *CKLF1* fragment isolated from the PHA-stimulated U937 cDNA library by the SSH method included a poly(A) tail and a single polyadenylation signal of a rare type (ATTAAA) that is also found in human eotaxin and TARC (thymus and activation-regulated chemokine) [11,12]. This suggests that the fragment contains a complete 3′ untranslated region [13,14]. There is an open reading frame encoding 99 residues with an ATG initiation codon at nt 96–98 whose surrounding sequence (GCGATGG) matches Kozak's rule [15]. It was a novel sequence after a search of the NR database in GenBank, but when a BLAST search of the EST database was conducted, some EST fragments derived from different cDNA libraries were identical with, or of high similarity to, our EST fragment. With the use of the EST Assembly Machine [6] we extended 52 nt beyond the 5′ untranslated region of the original *CKLF1* cDNA sequence. Northern blot analysis indicated that *CKLF1* had the expected molecular size. To verify the EST extension, we successfully amplified the full-length cDNA sequence of *CKLF1* from the PHA-stimulated U937 cell cDNA library with the primers P1 and P2

```
CKLF-2    - MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITG -50
CKLF-4    - MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITG -50
CKLF-1    - MDNVQPKIKHRPFCFSVKGHVKMLRL---------------------------- -26
CKLF-3    - MDNVQPKIKHRPFCFSVKGHVKMLRL---------------------------- -26

CKLF-2    - FEVTVILFFILLYVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALI -100
CKLF-4    - FEVTVILFFILLYVLRLDRLMKWLFWPLL---------------------- -79
CKLF-1    - ------------------------------DIINSLVTTVFMLIVSVLALI -47
CKLF-3    - ------------------------------------------------------ -26

CKLF-2    - PETTTLTVGGGVFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKE -150
CKLF-4    - ----------VFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKE -118
CKLF-1    - PETTTLTVGGGVFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKE -97
CKLF-3    - ----------VFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKE -65

CKLF-2    - VL -152
CKLF-4    - VL -120
CKLF-1    - VL -99
CKLF-3    - VL -67
```

Figure 3 Protein sequences of CKLF1, CKLF2, CKLF3 and CKLF4

The underlined sequences show the putative transmembrane regions of CKLF2 and CKLF4.

A.

```
  Exon 1                        Intron 1                      Exon 2
CTG CGG CTG  gtgaggccgggc----  (5396 bp)  ---- tttttcttctag  GCA CTA ACT
 L   R   L                                                    A   L   T

  Exon 2                        Intron 2                      Exon 3
CCT TTG CTT  gtaagtgttcag----  (4774 bp)  ---- gtgtgtttttag  ATT ATC AAC
 P   L   L                                                    D   I   I

  Exon 3                        Intron 3                      Exon 4
GGT GGA GGG  gtaagtggaagt----  (2669 bp)  ---- ctttctttccag  GTG TTT GCA
 G   G   G                                                    V   F   A
```

B.

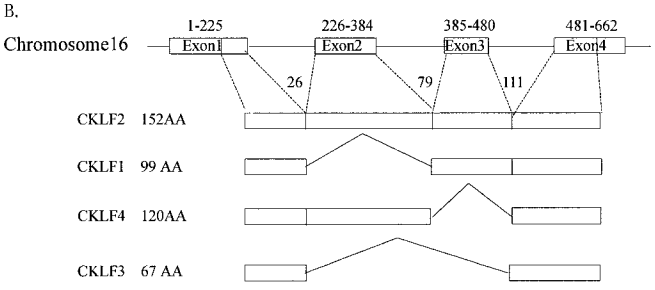

Figure 4 Sequences of the intron–exon junctions of the human *CKLF* gene and a sketch map of CKLF1, CKLF2, CKLF3 and CKLF4

(**A**) Sequences of the intron–exon junctions of the human *CKLF* gene. Exon sequences are shown in capitals; intron sequences are shown in lower-case letters. The length of each intron is shown in parentheses. (**B**) Different use of exons in CKLFs and a sketch map of CKLF1, CKLF2, CKLF3 and CKLF4. CKLF1 and its isoforms share common N-terminal and C-terminal ends. On the basis of the CKLF2 protein sequence, CKLF1, CKLF3 and CKLF4 have the selective use of residues 27–79, 27–111 and 80–111 respectively.

as described in the Materials and methods section. The full-length cDNA and predicted amino acid sequences of *CKLF1* (GenBank accession number AF096895) are shown in Figure 2(A).

The open reading frame of *CKLF1* cDNA encodes a highly basic and hydrophobic polypeptide of total 99 residues with a calculated molecular mass of 10.9 kDa. The deduced CKLF1 protein contained no putative N-glycosylation site, and no typical signal cleavage site was revealed by searching on the Signal P server [16,17], but it could be secreted from cells, as determined later by gene transfer assay and Western blotting. Some proteins without leader sequences have been discovered in other laboratories, including mammary-derived growth inhibitor, thio-

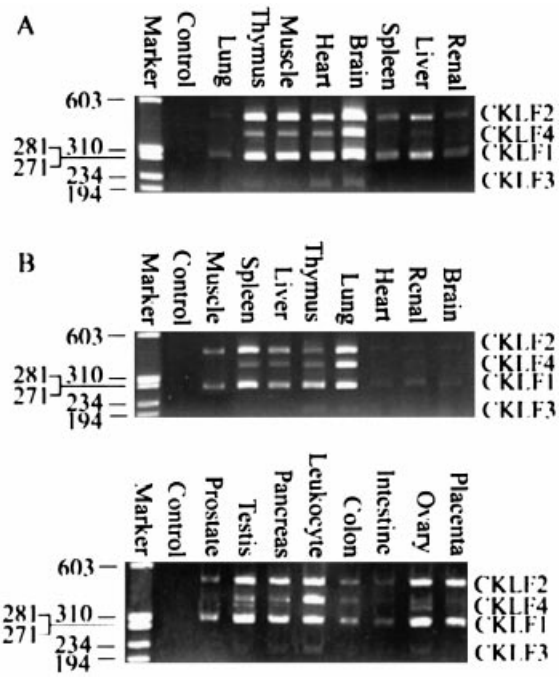

Figure 5 Expression of CKLF1 and its isoforms in human tissues

(**A**) Expression of CKLF forms by fetal tissues. (**B**) Expression of CKLF forms by adult tissues. Lengths in base-pairs are shown on the left.

redoxin, IL-1 and macrophage migration inhibitory factor 2 [18]. Hence CKLF1 might use another release mechanism. It was found that the amino acid sequence of CKLF1 had only limited similarity to that of the permease of *Caenorhabditis elegans* (AF025452) and shared no obvious similarity with other proteins. CKLF1 protein possesses two unique structural characteristics: (1) it contains three cysteine residues, the last two of which are organized in a pattern characteristic for the CC chemokine subfamily, and (2) it shares no significant similarity with other chemokines of CC family by BLAST searching. However, hand-operated alignment of CKLF1 with TARC and stimulated T-cell chemotactic protein 1 (STCP-1) [19], two of the members of CC chemokines located on chromosome 16, indicated the positioning of identical key residues near the CC motif (Figure 2B).

Cloning of three isoforms of CKLF1

With the RT–PCR technique we amplified the full cDNA sequence of *CKLF1* from the PHA-stimulated U937 cell cDNA library. To our surprise, three additional bands were found. After they had been cloned and sequenced, we designated them *CKLF2* (AF135380), *CKLF3* (AF135381) and *CKLF4* (AF145216). In U937 cells, the expression levels of *CKLF2*, *CKLF3* and *CKLF4* were lower than that of *CKLF1* (results not shown). This might be one of the reasons that there was only one *CKLF1* band in the Northern blot. The higher sensitivity of RT–PCR gives it the potential to detect the rare transcripts that are difficult to find by Northern blotting. DNA sequencing and open reading frame analysis revealed that *CKLF2*, *CKLF3* and *CKLF4* encoded 152, 67 and 120 residues respectively. They shared conserved N-terminal and C-terminal ends with CKLF1, but were structurally different within the region from residues 27 to 111 of CKLF2. All of them maintained the CC motif of chemokines. By the amino acid numbering of the longest isoform

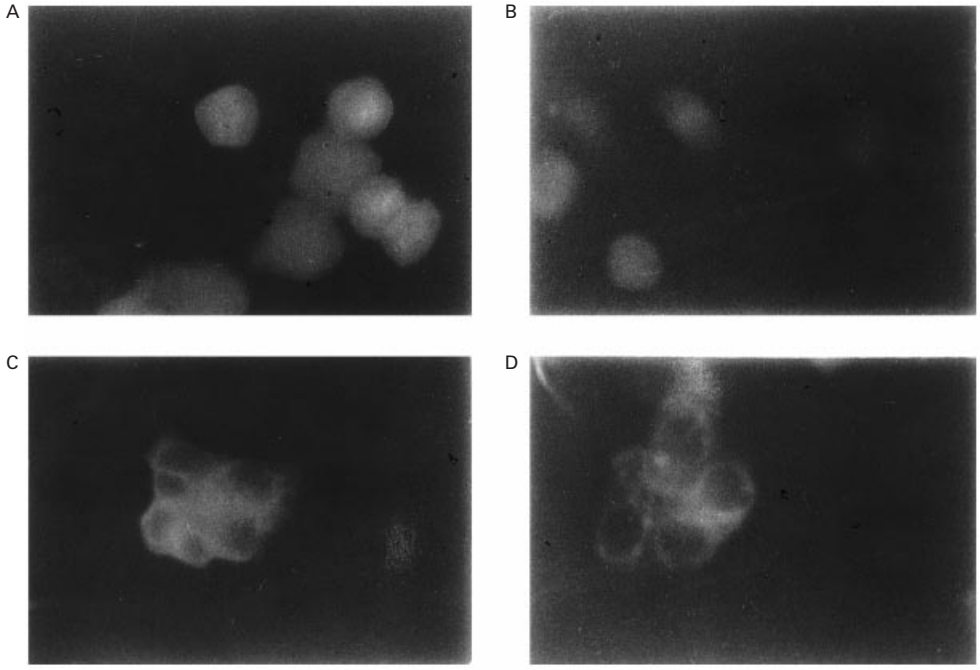

Figure 6 Analysis of the subcellular localization of EGFP, CKLF1–EGFP, CKLF2–EGFP and CKLF4–EGFP fusion proteins

For each experiment, EGFP was positioned in frame with the C-terminus of CKLF1, CKLF2 or CKLF4 and expressed in COS-7 cells by transient transfection. Fluorescence was examined with a fluorescence microscope. (**A**) Control EGFP transfectants. (**B**) CKLF1–EGFP transfectants. (**C**) CKLF2–EGFP transfectants. (**D**) CKLF4–EGFP transfectants. Magnification × 400.

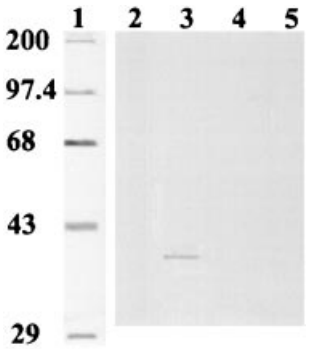

Figure 7 Detection of CKLFs/EGFP in the supernatants of transfected COS-7 cells using a polyclonal antibody recognizing EGFP

Lane 1, protein molecular-mass standards (high range; Gibco BRL); lane 2, EGFP; lane 3, CKLF1/EGFP; lane 4, CKLF2/EGFP; lane 5, CKLF4/EGFP.

CKLF2, it seems that *CKLF1*, *CKLF3* and *CKLF4* are the RNA splicing forms of *CKLF2* at amino acid residue positions 27–79, 27–111 and 80–111 respectively (Figure 3).

CKLF gene location and structure

Comparing the cDNA sequence of *CKLF2* with the working draft of the human genome produced by the Human Genome Project, we found that two clones located on human chromosome 16 had high similarity to the *CKLF2* cDNA sequence by searching the HTGS database. The putative *CKLF* gene consists of four exons and three introns. The sequences of intron–exon junctions matched the consensus sequence of eukaryotic splice junctions (Figure 4A) [19]. Alignment of the *CKLF* gene with the cDNA sequence showed that exons 1, 2, 3 and 4 encode amino acids 1–26, 27–79, 80–111 and 112–152 of CKLF2 protein respectively (Figure 4B). CKLF1, CKLF2, CKLF3 and CKLF4 have exons 1 and 4 in common but selective usage of exons 2 and 3. Therefore it can be concluded that they are alternative RNA splicing variants.

Distribution of CKLF1 mRNA in human tissues

CKLF1 mRNA had diverse variants in PHA-stimulated U937 cells. RT–PCR analysis with the Multiple Tissue cDNA Panel kit (Clontech) revealed that *CKLF1* and its variants had higher expression levels in human spleen, lung, testis, ovary, peripheral blood leucocyte, placenta, pancreas, fetal brain, fetal skeletal muscle, fetal thymus and heart compared with that in human skeletal muscle, liver, thymus, colon, prostate, fetal spleen and fetal liver; in other human tissues, including human brain, renal, heart and intestine and fetal lung and renal, the expression of *CKLF* forms was barely detected. For most tissues the expression levels of CKLF1 and CKLF2 were similar and higher than that of CKLF4, whereas the expression level of CKLF3 was the lowest among the CKLF isoforms (Figure 5).

Subcellular localization of CKLF1, CKLF2 and CKLF4

To investigate the subcellular localization of CKLF1, CKLF2 and CKLF4, we generated fusion cDNA species in which the coding sequence for EGFP was positioned in frame with the C-terminus of CKLF1, CKLF2 and CKLF4. These fusion constructs and the control vector were transfected into COS-7 cells. The distribution of fluorescence in living cells was examined with a fluorescence microscope. Cells expressing control EGFP exhibited bright green fluorescence that could be seen throughout

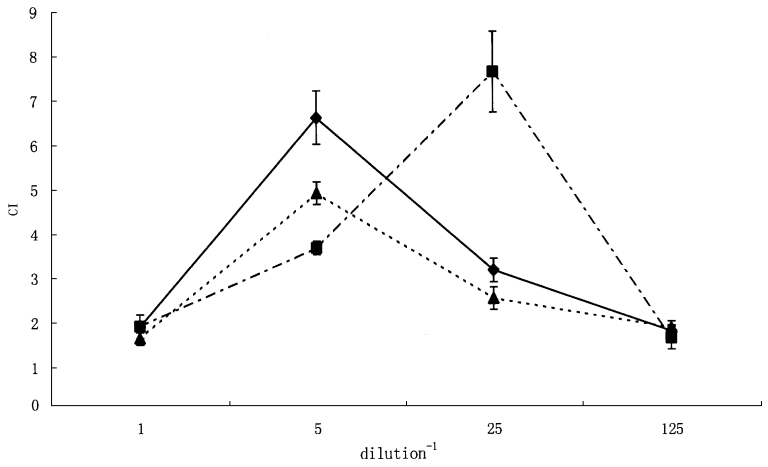

Figure 8 Chemoattraction of leucocyte subsets by CKLF1

Human peripheral blood lymphocytes (▲), monocytes (◆) and neutrophils (■) were exposed in a micro-chamber to increasing dilutions of supernatants from a transfected COS-7 cell culture; the numbers of cells that migrated to the membrane were determined. Migrated cells were counted under a microscope at × 400 magnification in five randomly selected fields per well. Abbreviation: CI, chemotactic index.

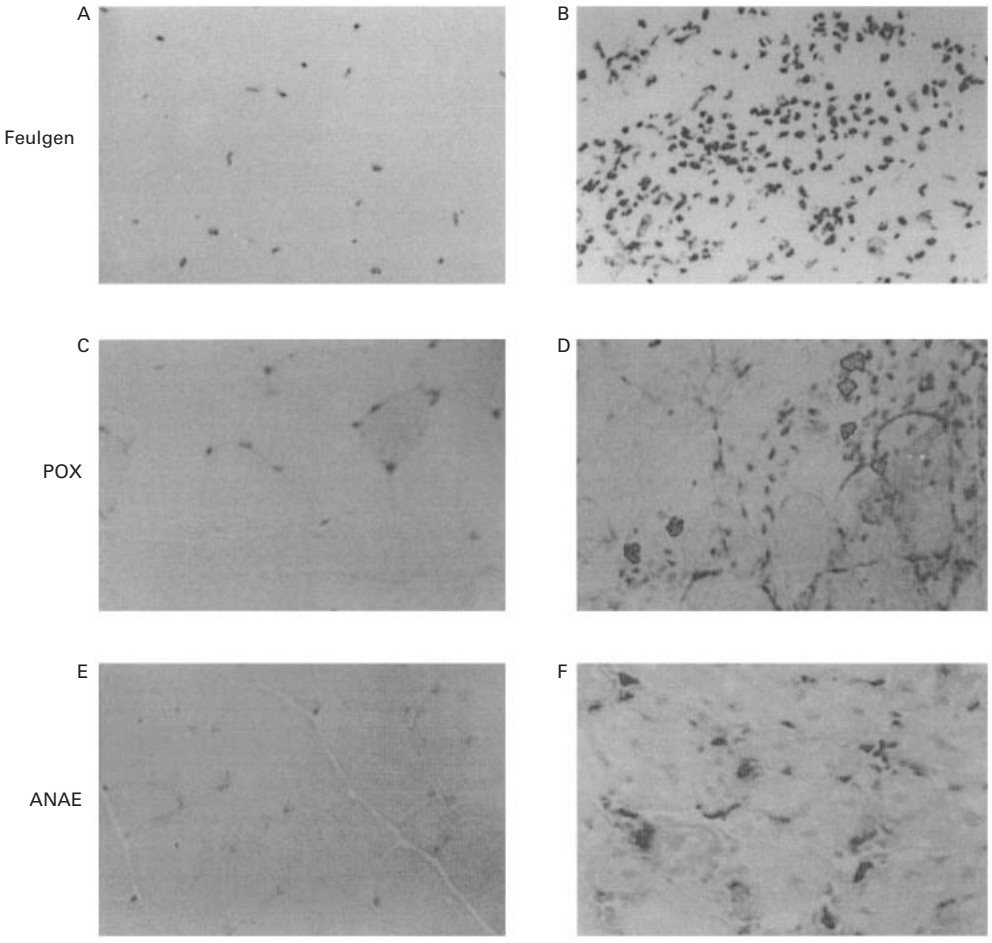

Figure 9 Infiltration of leucocytes in mouse muscle after intramuscular injection with pCDI-CKLF1

Injections were performed with pCDI plasmid (100 μg) in 100 μl of pyrogen-free physiological saline (**A**, **C**, **E**) or 100 μg of pCDI-CKLF1 plasmid in the same buffer (**B**, **D**, **F**). Typical histological sections of the muscle biopsies 10 days after injection are shown. Leucocyte recruitment can be seen in (**D**) and (**F**). Magnification × 200.

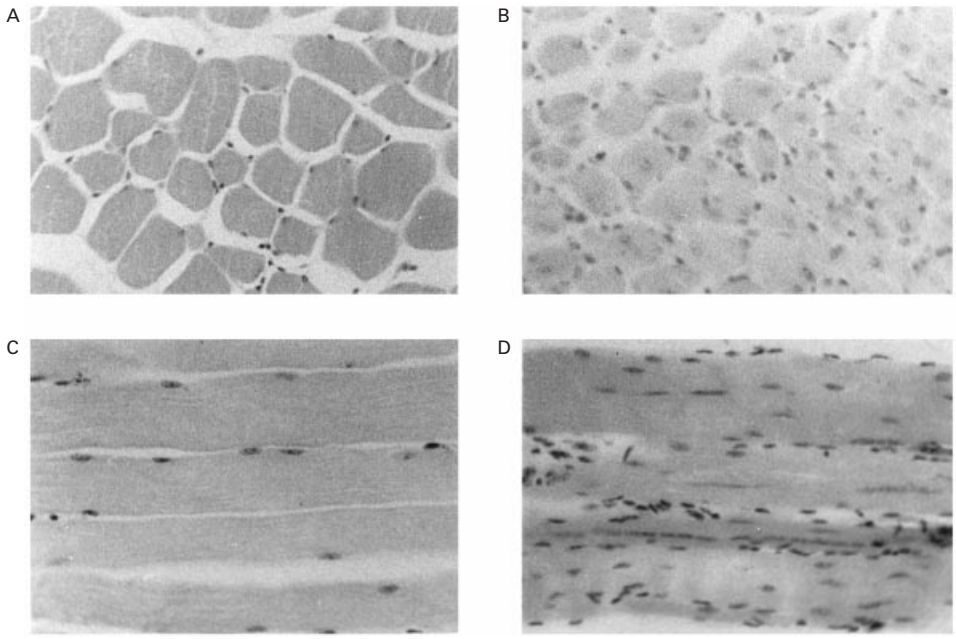

Figure 10 Effect of expression of CKLF1 on anterior tibial muscle of BALB/c mice

Cross-sections (**A**, **B**) and vertical sections (**C**, **D**) stained with HE 10 days after injection with naked DNA. Muscle fibre regeneration was evident in CKLF1-expressing muscle by the presence of central nuclei (**B**, **D**), which was absent from control muscle (**A**, **C**). Magnification $\times 200$.

the transfected cells; there was no specific subcellular localization (Figure 6A). In cells expressing CKLF1–EGFP, weak fluorescence was detected throughout the cells (Figure 6B). On the basis of the structural characteristics of CKLF1, it can be supposed that much of the CKLF1–EGFP protein was secreted into the cell culture supernatant. In contrast, CKLF2–EGFP and CKLF4–EGFP are predominantly located on the cell periphery, while only limited CKLF2 and CKLF4 directed fluorescence were detected in other cellular locations (Figures 6C and 6D). HEK-293 cells yielded identical results, confirming that the subcellular localizations of CKLF1, CKLF2 and CKLF4 were not unique to COS-7 cells (results not shown).

CKLF1 is expressed in the secreted pathway

From the subcellular localization study we speculated that CKLF1 was a secreted protein. To confirm this, the supernatants of COS-7 cells transfected with four different constructs were collected and subjected to SDS/PAGE. Western blot analysis verified a specific band with molecular mass of approx. 38 kDa (Figure 7), which is in accord with the size of the CKLF1–EGFP fusion protein. There was no obvious band in the lanes of control EGFP, CKLF2–EGFP or CKLF4–EGFP. Therefore we can conclude that CKLF1 is a secreted protein, although its signal cleavage site remains to be determined, whereas CKLF2–EGFP and CKLF4–EGFP could not be detected in the supernatants of cell culture. Moreover, computer analysis of CKLF2 and CKLF4 indicated that they contained a highly hydrophobic region of 27–79 amino acids and might be transmembrane proteins. After we had submitted the sequence of *CKLF2* to GenBank, two additional sequences were submitted from other laboratories, both with the same coding sequence as CKLF2. One of them, AF057306, is identified as a transmembrane proteolipid, indicating that CKLF2 is a transmembrane protein.

Chemotactic activities of recombinant CKLF1 *in vitro*

To investigate the biological activity of CKLF1 we constructed a mammalian expression vector pCDI-CKLF1 that contained the full length of the *CKLF1* coding region. Another construct was a control expression vector without any cDNA insert. Separated cultures of COS-7 cells were transfected with each of the two constructs; the supernatants from cell cultures were collected for chemotactic analysis. The chemotactic properties of recombinant CKLF1 on different cell populations were evaluated by using transwell migration assays. The chemotactic activity of CKLF1 is expressed as the chemotactic index (the ratio between the number of cells that migrated in the supernatants of pCDI-CKLF1-transfected cells and the number of cells that migrated in the supernatants of the control vector pCDI-transfected cells) [20]. The supernatants from the pCDI-CKLF1-transfected cells had chemotactic effects on human neutrophils, lymphocytes and monocytes compared with the control supernatants (Figure 8). We obtained a similar result with the *Drosophila melanogaster* S2 cells expression system, which indicated that the chemotactic activity of CKLF1 was not unique to the supernatants of COS-7 cells (results not shown).

Chemotactic potency of CKLF1 *in vivo*

The chemotactic properties of recombinant CKLF1 were further analysed after the administration of naked plasmid DNA *in vivo*. Purified endotoxin-free plasmids were injected into the anterior tibial muscles of BALB/c mice. At 10 days after injection the CKLF1 expression level was high when examined using the RT–PCR technique (results not shown). When the muscle cross-sections were stained with Feulgen, there were many more nuclei in the mice injected with pCDI-CKLF1 than with pCDI (Figures 9A and 9B). This suggested that CKLF1 could elicit leucocyte recruitment or stimulate muscle stem cell proliferation, or both, *in vivo*. To verify the chemotactic activity of CKLF1 *in vivo*,

some muscle cross-sections were stained for POX or ANAE, and then re-stained with HE [21]. As illustrated in Figures 9(D) and 9(F), many cells showed strong POX or ANAE activity, whereas in Figures 9(C) and 9(E) no positive cells are seen. It is known that POX-positive cells are mainly neutrophils and ANAE-positive ones are monocytes and T lymphocytes [22]. On the basis of these enzyme activities and a morphological study of other sections (results not shown), we conclude that CKLF1 can attract monocytes, neutrophils and lymphocytes, which is in agreement with the chemotactic effects of CKLF1 *in vitro*.

CKLF1 promotes morphological changes in skeletal muscle cells *in vivo*

After the muscle biopsies were fixed and sectioned, some of the muscle cross-sections and vertical sections were stained with HE. In addition to the infiltration of more cells in muscle expressing CKLF1, there were also morphological changes. Figure 10 illustrates the effect of CKLF1 expression on skeletal muscles morphologically. Central nuclei were visible within the muscle fibres in cross-sections of the CKLF1 group (Figure 10B); no central nuclei were observed in the control mice (Figure 10A). With regard to the vertical sections the nuclei were located in the middle of the muscle fibres, like a string of beads (Figure 10D). The existence of such changes in the nucleus indicated that the muscles were undergoing continuing regeneration that might be induced by satellite cell activation [23]. Further experiments are needed to confirm whether this activation is affected directly by CKLF1 or indirectly by the cytokines produced from the attracted leucocytes.

DISCUSSION

We have successfully cloned a novel cytokine, CKLF1, as well as its three splicing variants by a new approach. To find more cytokines, the strategy could also be used in other cytokine-producing cells in combination with different inhibitory factors of cytokine synthesis.

The results from transfections of COS-7 cells and HEK-293 cells with CKLF1–EGFP, CKLF2–EGFP and CKLF4–EGFP constructs indicate that CKLF1 is a secreted protein, whereas CKLF2 and CKLF4 are transmembrane proteins. The other variant, CKLF3, is still under study, although on the basis of a computer analysis we propose that CKLF3 is expressed in the same way as CKLF1. It is interesting that differentially spliced CKLFs can be either secreted or transmembrane isoforms. It has been reported that a novel CC chemokine, Eskine, is produced as two splice variants arising as a result of alternative 5′ exon usage [24]. One of them is a classical secreted chemokine; the other is targeted to the nucleus. With regard to CKLF variants, we consider that their cell localizations are caused by a selective use of a highly hydrophobic region of residues 27–79. Structurally, CKLF1 contains three cysteine residues, of which the last two are organized in a pattern characteristic of the CC chemokine subfamily and the first is likely to be located within the signal peptide. CKLF1 is different from a classical chemokine in that it has no additional cysteine residue in its C-terminal region. The *CKLF1* gene is located on human chromosome 16 and shows only very weak similarity to the CC chemokines TARC and STCP-1 [25], the only two known CC chemokines located on chromosome 16, whereas most CC chemokines are located on chromosome 17 [26]. Fractalkine, the sole CX3C chemokine and transmembrane chemokine, is also on chromosome 16 [25]. Because recombinant CKLF1 exhibits potent chemotactic activity on a broad spectrum of leucocytes both *in vitro* and *in vivo*, we consider that CKLF1 and its variants could be classified as a new chemokine family; there might be an evolutionarily conserved relationship between CKLFs, TARC, STCP-1 and fractalkine.

The expression of recombinant CKLF1 in muscles could stimulate the proliferation and differentiation of the skeletal muscle cells, which is compromised by the effect of insulin-like growth factor 1 expression in extensor digitorum longus muscles of young adult mice [27]. We propose that there might be three possible mechanisms by which CKLF1 could regulate muscle regeneration: first, CKLF1 itself could affect the proliferation and differentiation of skeletal muscle satellite cells in the same way as insulin-like growth factor 1; secondly, the leucocytes attracted by CKLF1 could produce different kinds of cytokine when activated, which could influence the growth of myogenic cells; thirdly, CKLF1 could have a synergistic action with some of the secreted cytokines to increase the rate of cellular proliferation and the formation of myotubes. Further experiments will be required to validate our hypothesis.

In summary, we have successfully isolated a novel cytokine CKLF1 and its three variants. Its expression level could be partly decreased by IL-10. Functional studies verified its chemotactic properties and its enhancing effects on the proliferation and differentiation of skeletal muscle cells.

We thank Dr Julia White (HUGO Nomenclature Committee, The Galton Laboratory, University College London, U.K.) for advice on naming CKLF; Professor Lingsong Li for assistance in the preparation of the manuscript; Dawu Huang and Yingcheng Zhong for the processing of photos; and Zhongyu Fan for typing the manuscript. This work was supported by Chinese High Tech Program (863) and State Key Program of Basic Research (973).

REFERENCES

1 Nicola, N. A. (ed.) (1995) Guidebook to Cytokines and their Receptors, Oxford University Press, Oxford.
2 Mantovani, A. (1999) The chemokine system: redundancy for robust outputs. Immunol. Today **20**, 254–257
3 Brantschen, S., Gauchat, J. F., de Weck, A. L. and Staller, B. M. (1989) Differential expression of cytokine mRNAs in human cell lines. Lymphokine Res. **8**, 163–172
4 de Waal Malefyt, R., Abrams, J. and Bennet, B. (1991) Interleukin-10 (IL-10) inhibits cytokine synthesis by human monocytes: an autoregulatory role of IL-10 produced by monocytes. J. Exp. Med. **174**, 1209–1220
5 Diatchenko, L., Lau, Y. F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Gurskaya, N., Sverdlov, E. D., Lukyanov, K. et al. (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. U.S.A. **93**, 6025–6030
6 Banfi, S., Guffanti, A. and Borsani, G. (1998) How to get the best of dbEST. Trends Genet. **14**, 80–81
7 Liu, H., Wang, Y., Zhang, Y., Song, Q., Di, C., Chen, G., Tang, J. and Ma, D. (1999) *TFAR19*, a novel apoptosis-related gene cloned from human leukemia cell line TF-1, could enhance apoptosis of some tumor cells induced by growth factor withdrawal. Biochem. Biophys. Res. Commun. **254**, 203–210
8 Youn, B.-S., Zhang, S. M., Lee, E. K., Park, D. H., Broxmeyer, H. E., Murphy, P. M., Locati, M., Pease, J. E., Kim, K. K., Antol, K. and Kwon, B. S. (1997) Molecular cloning of leukotactin-1: a novel human chemokine, a chemoattractant for neutrophils, monocytes, and lymphocytes and a potent agonist at CC chemokine receptors 1 and 3. J. Immunol. **159**, 5201–5209
9 Sarafi, M. N., Garcia-Zapeda, E. A., Maclean, J. A., Charo, I. F. and Luster, A. D. (1997) Murine monocyte chemoattractant protein (MCP)-5: a novel CC chemokine that is a structural and functional homologue of human MCP-1. J. Exp. Med. **185**, 99–109
10 Mir, L. M., Bureau, M. F., Gehl, J., Rangara, R., Rouy, D., Gailland, J. M., Delaere, P., Branellec, D., Schwartz, B. and Scherman, D. (1999) High-efficiency gene transfer into skeletal muscle mediated by electric pulses. Proc. Natl. Acad. Sci. U.S.A. **96**, 4262–4267
11 Garcia-Zepeda, E. A., Rothenberg, M. E., Ownbey, R. T., Cekestin, J., Leder, P. and Luster, A. D. (1996) Human eotaxin is specific chemoattractant for eosinophil cells and provides a new mechanism to explain tissue eosinophilia. Nat. Med **4**, 449–456

12 Imai, T., Yoshida, T., Baba, M., Nishimura, M., Kakizaki, M. and Yoshie, O. (1996) Molecular cloning of a novel T cell-directed CC chemokine expressed in thymus by signal sequence trap using Epstein-Barr virus vector. J. Biol. Chem. **271**, 21514–21521
13 Thekkumkara, T. J., Livingston, W., Kumar, R. S. and Sen, G. C. (1992) Use of alternative polyadenylation sites for tissue-specific transcription of two angiotensin-converting enzyme mRNAs. Nucleic Acids Res. **20**, 683–687
14 Yang, H. and Melera, P. W. (1994) A genetic polymorphism within the third poly(A) signal of the DHFR gene alters the polyadenylation pattern of DHFR transcripts in CHL cells. Nucleic Acids Res. **22**, 2694–2702
15 Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. **266**, 19867–19870
16 Nielsen, H., Brunak, S., Engelbrecht, J. and Von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10**, 1–6
17 Nielsen, H., Brunak, S. and Von Heijne, G. (1999) Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein Eng. **12**, 3–9
18 Senior, K. (2000) Novel drug development strategy targets protein 'export'. Drug Discov. Today **5**, 379–380
19 Chang, M., McNinch, J., Elias, III, C., Manthey, C. L., Grosshans, D., Meng, T., Boone, T. and Andrew, D. P. (1997) Molecular cloning and functional characterization of a novel CC chemokine, stimulated T cell chemotactic protein (STCP-1) that specifically acts on activated T lymphocytes. J. Biol. Chem. **272**, 25229–25237
20 Patel, V. P., Kreider, B. L., Li, Y., Li, H., Leung, K., Salcedo, T., Nardelli, B., Pippalla, V., Gentz, S., Thotakura, R. et al. (1997) Molecular and functional characterization of two novel human CC chemokines as inhibitors of two distinct classes of myeloid progenitors. J. Exp. Med. **185**, 1163–1172
21 Stevens, A. and Bancroft, J. D. (1977) Theory and Practice of Histological Techniques, Spottiswoode Ballantyne, London
22 Smith, H., Filipe, M. I. and Lake, B. D. (1983) The leukaemias. In Histochemistry in Pathology (Filipe, M. I., ed.), pp. 206–214, Churchill Livingstone, Edinburgh
23 Carlson, B. M. and Faulkner, J. A. (1983) The regeneration of skeletal muscle fibers following injury. Med. Sci. Sports Exercise **15**, 187–198
24 Baird, J. W., Nibbs, R. J. B., Komai-Koma, M., Connolly, J. A., Ottersbach, K., Clark-Lewis, I., Liew, F. Y. and Graham, G. J. (1999) Eskine, a novel β-chemokine, is differentially spliced to produce secretable and nuclear targeted isoforms. J. Biol. Chem. **274**, 33496–33503
25 Nomiyama, H., Imai, T., Kusuda, J. and Miura, R. (1998) Human chemokines fractalkine (SCYD1), MDC (SCYA22) and TARC (SCYA17) are clustered on chromosome 16q13. Cytogenet. Cell Genet. **81**, 10–11
26 Vaddi, K., Keller, M. and Newton, R. C. (1997) The Chemokine Facts Books, Academic Press, San Diego
27 Barton-Davis, E. R., Shoturma, D. I., Musaro, A., Rosenthal, N. and Sweeney, H. L. (1998) Viral mediated expression of insulin-like growth factor I blocks the ageing-related loss of skeletal muscle function. Proc. Natl. Acad. Sci. U.S.A. **95**, 15603–15607

Received 8 November 2000/9 March 2001; accepted 18 April 2001