Biochem. J. (1993) **294**, 271–278 (Printed in Great Britain)

Osteoblast-specific factor 2: cloning of a putative bone adhesion protein with homology with the insect protein fasciclin I

Sunao TAKESHITA, Reiko KIKUNO, Ken-ichi TEZUKA* and Egon AMANN†
Hoechst Japan Limited, Pharma Research Laboratories, 1-3-2, Minamidai, Kawagoe, Saitama, 350 Japan

A cDNA library prepared from the mouse osteoblastic cell line MC3T3-E1 was screened for the presence of specifically expressed genes by employing a combined subtraction hybridization/differential screening approach. A cDNA was identified and sequenced which encodes a protein designated osteoblast-specific factor 2 (OSF-2) comprising 811 amino acids. OSF-2 has a typical signal sequence, followed by a cysteine-rich domain, a fourfold repeated domain and a C-terminal domain. The protein lacks a typical transmembrane region. The fourfold repeated domain of OSF-2 shows homology with the insect protein fasciclin I. RNA analyses revealed that OSF-2 is expressed in bone and to a lesser extent in lung, but not in other tissues. Mouse OSF-2 cDNA was subsequently used as a probe to clone the human counterpart. Mouse and human OSF-2 show a high amino acid sequence conservation except for the signal sequence and two regions in the C-terminal domain in which 'in-frame' insertions or deletions are observed, implying alternative splicing events. On the basis of the amino acid sequence homology with fasciclin I, we suggest that OSF-2 functions as a homophilic adhesion molecule in bone formation.

INTRODUCTION

Tissue organization during development requires the segregation of cells into different lineages and stage- and differentiation-specific intercellular adhesion of cells. There is, consequently, a great variety of specific cell–cell and cell–matrix interactions underlying these biological phenomena. Many different recognition molecules or cell surface adhesion receptors have been identified in recent years, which can be categorized into four major classes: (i) the cadherins, (ii) the immunoglobulin superfamily adhesion receptors, (iii) the selectins and (iv) the integrins (reviewed in ref. [1]). Adhesion receptors are also probably important in the development and remodelling of bone. The cellular and molecular mechanisms of bone formation and regeneration, however, are as yet only poorly understood. Among known bone adhesion receptors, two subtypes of integrins are expressed in osteoclasts: the vitronectin receptor [2,3] and the receptor for type I collagen [4,5]. Osteoblasts express the type-I collagen receptor and the fibronectin receptor [6], but little, if any, vitronectin receptor [7]. Several proteins containing the tripeptide recognition sequence RGD (Arg-Gly-Asp) [8,9], which is recognized by the vitronectin receptor [10,11], are constitutively expressed in osteoblasts and are found to be incorporated into the bone extracellular matrix, where osteoclasts bind to them. Among these proteins are osteopontin [12,13], thrombospondin [14,15], fibronectin [16] and laminin [17]. The receptor profile of osteoclasts differs from other haemopoietic cell types and may reflect a specialized role in bone. In contrast, the integrins on osteoblasts do not seem to differ radically from other stromal or fibroblastic cells. It has been speculated that unidentified homophilic receptors should be expressed on osteoclasts and osteoblasts during early stages of mesenchymal differentiation and pattern formation in the embryo [7]. Such homophilic receptors, however, have not yet been described for bone cells.

Here we describe the isolation and characterization of mouse and human cDNA encoding a new potential bone adhesion protein, which we name osteoblast-specific factor 2 (OSF-2) (a previously described factor was named OSF-1, see ref. [18]). Computer analysis of the deduced primary amino acid sequence of OSF-2 revealed a complex protein structure with a characteristic fourfold-repeated domain. A similar structure has been reported for the insect protein fasciclin I, a protein implicated in neuronal cell–cell adhesion [19–23] and sequence similarity between OSF-2 and fasciclin I within the repeat domain was detected. The expression of OSF-2 mRNA in the mouse calvarial osteoblastic cell line MC3T3-E1 is regulated by osteotrophic factors. OSF-2 is highly conserved between mouse and human and we speculate that this protein acts as a homophilic adhesion molecule in bone formation.

MATERIALS AND METHODS

MC3T3-E1 subtraction library construction and screening

MC3T3-E1 and NIH3T3 poly(A)$^+$ RNA was isolated as described [18]. Solution hybridization of MC3T3-E1 cDNA and photobiotinylated poly(A)$^+$ RNA from NIH3T3 cells was performed according to the 'subtractor II' protocol (Invitrogen). The double-stranded subtracted MC3T3-E1 cDNA was ligated with *Eco*RI/*Not*I adaptors (Pharmacia) and cloned into λgt10 (Stratagene). Recombinant phages of the subtracted MC3T3-E1 library (1×10^4) were rescreened by differential plaque hybridization (first round screening) employing cDNA probes prepared

---

Abbreviations used: OSF-2, osteoblast-specific factor 2; poly(A)$^+$, polyadenylated; FCS, fetal calf serum; EGF, epidermal growth factor; 1,25-(OH)$_2$D$_3$, 1,25-dihydroxyvitamin D$_3$; TGF-β_1, transforming growth factor β_1; IGF-I, insulin-like growth factor I; PDGF, platelet-derived growth factor; PGE$_2$, prostaglandin E$_2$; PTH, parathyroid hormone; hOSF-2pl, human OSF-2 from the placenta library screen; hOSF-2os, human OSF-2 from the osteosarcoma library screen; 1 × SSC, 0.15 M NaCl/0.015 M sodium citrate.

* Present address: Meikai University School of Dentistry, Department of Oral Anatomy, Sakado, Saitama 350-02 Japan.

† To whom correspondence should be addressed.

The sequence data for mouse OSF-2, human OSF-2os and human OSF-2pl will appear in the EMBL databank under the accession numbers D13664, D13666 and D13665 respectively.

from mRNA isolated from MC3T3-E1 and NIH3T3 cells as described [18]. MC3T3-E1-specific clones (155) were plaque-isolated, and cDNA inserts from these clones were amplified by PCR [24]. The PCR fragments were isolated, radiolabelled and used as probes to perform MC3T3-E1 and NIH3T3 RNA dot-blot analysis (second round screening). PCR fragments giving positive hybridization signals with MC3T3-E1 RNA, but negative signals with NIH3T3 RNA, were recloned into pUC118 (Takara Shuzo Company) or pHSG398 [25].

Screening of human cDNA libraries

The human osteosarcoma tissue cDNA library (in λZapII) was a gift from Michael Kiefer (Chiron Co., Emeryville, CA, U.S.A.) [26]. A human placenta cDNA library (in λgt11) was purchased from Clontech. Using the mouse OSF-2 cDNA insert of MC163 as a hybridization probe, both human cDNA phage libraries were screened: 72 positive clones from the placenta library and 31 positive clones from the osteosarcoma library were obtained. The seven plaques with the strongest signals from the placenta library and the five plaques with the strongest signals from the osteosarcoma library were amplified and their respective *Eco*RI or *Bgl*II inserts were isolated. The largest insert of each type was cloned into pUC118 and pHSG398 respectively, and the resulting plasmids were named pKOT133 and pKOT158.

RNA dot blot and Northern blot

Total RNAs from mouse tissues were purified by the guanidinium thiocyanate method [27]. Heat-denatured total RNA (1 μg) was dotted on to nylon membrane filters (Biodyne, Pall), and hybridization with ^{32}P-labelled cDNA fragments was performed. Northern-blot analysis with 1 μg of cytoplasmic RNA/lane was performed using a 1.2% formaldehyde/agarose gel. After electrophoresis, RNA was blotted on to a nylon filter, and hybridization with the randomly primed mouse OSF-2 cDNA probe was performed using standard procedures [27]. Genomic Southern-blot analysis was performed as described [27].

DNA sequencing and protein analysis

The nucleotide sequences of the cDNA inserts were determined by the dideoxy chain-termination method [28] using the automatic DNA sequence analyser, model 373A from Applied Biosystems. The nucleotide sequence of the entire coding region was determined by sequencing both strands. Nucleotide and amino acid sequence homology search was performed through the GeneBank DNA database (release 71.0) and NBRF protein database (release 32.0), utilizing the FastA and TfastA programs of GCG sequence analysis software package [29].

Cell cultures

MC3T3-E1 cells [30] were grown in the presence of 10% fetal calf serum (FCS) and were seeded at a density of 2.5×10^5 cells in 10 cm diameter culture dishes. After 3 days, cells were washed with PBS and cultured in serum-free α modification of Eagle's medium (αMEM) (purchased from Flow laboratories) for 24 h. FCS, growth factors, or hormones were added to the cultures as described later in the legends to Figures 3 and 6. After 24 h, cytoplasmic RNA was extracted by standard procedures [27].

RESULTS

Cloning of mouse OSF-2 cDNA

Employing a subtraction hybridization/differential screening approach, 22 cDNA clones were isolated from the mouse calvarial osteoblastic cell line MC3T3-E1, which hybridized with MC3T3-E1 RNA, but not with NIH3T3 RNA. These cDNA clones could be classified into six groups according to their mouse tissue RNA dot-blot hybridization pattern and were partially sequenced (not shown). The largest group consisted of 15 individual clones with cDNA inserts having similar restriction enzyme patterns (the other groups contained cDNA encoding different osteoblast-specific proteins and will be described elsewhere). The clone with the longest insert, MC163, was sequenced. The entire nucleotide sequence and the deduced amino acid sequence of mouse OSF-2 is shown in Figure 1. The 3187 bp insert from MC163 consists of an 18 bp 5′ untranslated region, an open reading frame spanning 2436 bp and a 3′ untranslated region of 733 bp, lacking a poly(A) stretch. The translation initiation site is deduced as follows: (i) there is no ATG codon further upstream from the assigned start codon, (ii) termination codons are present in all the other possible reading frames within the 2436 bp-long assigned reading frame and (iii) no longer 5′ sequences could be identified in other sequenced OSF-2 cDNA clones. The consensus sequence 5′-GGCACC-3′ often preceding the ATG initiation codon in vertebrate mRNA [32], however, is missing. The mouse OSF-2 open reading frame encodes a protein of 811 amino acids, with an M_r of 90254 (all M_r values are given for unmodified proteins). The protein has a typical N-terminal signal sequence of 23 amino acids and one possible N-glycosylation site, but lacks a typical transmembrane domain. Six of 12 cysteine residues are located in the 84-amino acid residue long region (cysteine-rich region) next to the putative signal sequence. Computer amino acid sequence comparison of mouse OSF-2 with itself revealed the protein structure shown in Figure 2. A fourfold internal repeat composed of approx. 130 amino acids each follows the cysteine-rich region and precedes the 178-residue-long C-terminal region. Each of the four internal repeat units contains two particularly conserved regions of 13 and 14 amino acids each (see below). Whereas the overall sequence similarity between the repeated units is relatively low (23.2% amino acid identity on average), the sequence similarity between the two particularly conserved regions is high (61.5% and 45.2% amino acid identity on average).

mRNA detection, gene copy determination and tissue-specific expression of mouse OSF-2

Northern-blot analysis revealed the presence of a single mRNA band of approx. 3.4 kb in MC3T3-E1 cells (Figure 3), which corresponds well to the size of the cloned cDNA of 3187 bp. Addition of FCS to the growth medium represses the amount of detectable OSF-2 mRNA (Figure 3; see also below). Southern-blot analysis employing different washing stringencies showed that the OSF-2 gene is present as a single-copy gene in the mouse genome (Figure 4).

In order to examine the tissue-specific expression of OSF-2, RNA dot-blot analysis was performed. Figure 5 shows the result of this experiment. Strong hybridization signals are observed in calvarial osteoblast-enriched cells and in MC3T3-E1 cells, and a weaker signal is observed in lung. No expression of OSF-2 can be detected in brain, heart, kidney, liver, muscle, placenta, spleen, testis and thymus. This result indicates that OSF-2 is primarily expressed in bone.

```
CGGAGCTCAGGGCTGAAG  ATG GTT CCT CTC CTG CCC TTA TAT GCT CTG CTG    51
                    Met Val Pro Leu Leu Pro Leu Tyr Ala Leu Leu    11

CTG CTG TTC CTG TGT GAT ATT AAC CCT GCA AAT GCC AAC AGT TAC TAT    99
Leu Leu Phe Leu Cys Asp Ile Asn Pro Ala Asn Ala Asn Ser Tyr Tyr    27

GAC AAG GTC CTG GCT CAC AGC CGC ATC AGG GGT CGG GAT CAG GGC CCA   147
Asp Lys Val Leu Ala His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro    43

AAC GTC TGT GCC CTC CAG CAA ATT CTG GGC ACC AAA AAG AAA TAC TTC   195
Asn Val Cys Ala Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe    59

AGC TCC TGT AAG AAC TGG TAT CAA GGT GCT ATC TGC GGG AAG AAA ACC   243
Ser Ser Cys Lys Asn Trp Tyr Gln Gly Ala Ile Cys Gly Lys Lys Thr    75

ACT GTG CTA TAT GAA TGC TGC CCT GGC TAT ATG AGA ATG GAA GGG ATG   291
Thr Val Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met    91

AAA GGC TGC CCC GCA GTG ATG CCT ATT GAC CAT GTT TAT GGC ACG CTG   339
Lys Gly Cys Pro Ala Val Met Pro Ile Asp His Val Tyr Gly Thr Leu   107

GGC ATT GTG GGA GCC ACT ACC ACT CAG CAC TAC TCC GAT GTC TCG AAG   387
Gly Ile Val Gly Ala Thr Thr Thr Gln His Tyr Ser Asp Val Ser Lys   123

CTG AGA GAA GAG ATT GAA GGA AAA GGG TCA TAC ACG TAC TTC GCG CCG   435
Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Tyr Thr Tyr Phe Ala Pro   139

AGT AAC GAG GCT TGG GAG AAC CTG GAT TCT GAC ATT CGC AGA GGA CTG   483
Ser Asn Glu Ala Trp Glu Asn Leu Asp Ser Asp Ile Arg Arg Gly Leu   155

GAG AAC AAT GTC AAT GTT GAG CTA CTG AAT GCC TTA CAC AGC CAC ATG   531
Glu Asn Asn Val Asn Val Glu Leu Leu Asn Ala Leu His Ser His Met   171

GTT AAT AAG AGA ATG TTA ACC AAG GAC CTG AAA CAC GGC ATG GTT ATT   579
Val Asn Lys Arg Met Leu Thr Lys Asp Leu Lys His Gly Met Val Ile   187

CCT TCA ATG TAC AAC AAT CTG GGG CTT TTT ATT AAC CAT TAT CCC AAT   627
Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe Ile Asn His Tyr Pro Asn   203

GGG GTT GTC ACT GTG AAC TGT GCT CGA GTC ATC CAT GGG AAC CAG ATT   675
Gly Val Val Thr Val Asn Cys Ala Arg Val Ile His Gly Asn Gln Ile   219

GCC ACA AAT GGT GTC GTC CAT GTC ATT GAC CGT GTC CTG ACA CAA ATT   723
Ala Thr Asn Gly Val Val His Val Ile Asp Arg Val Leu Thr Gln Ile   235

GGT ACC TCC ATC CAA GAC TTC CTT GAA GCA GAA GAC GAC CTT TCA TCA   771
Gly Thr Ser Ile Gln Asp Phe Leu Glu Ala Glu Asp Asp Leu Ser Ser   251

TTT AGA GCA GCC GCC ATC ACC TCT GAC CTC TTG GAG TCC CTT GGA AGA   819
Phe Arg Ala Ala Ala Ile Thr Ser Asp Leu Leu Glu Ser Leu Gly Arg   267

GAT GGT CAC TTC ACG CTC TTT GCT CCC ACC AAT GAA GCT TTC GAG AAA   867
Asp Gly His Phe Thr Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys   283

CTG CCA CGA GGT GTC CTA GAA AGG ATC ATG GGA GAC AAA GTG GCT TCT   915
Leu Pro Arg Gly Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser   299

GAA GCT CTC ATG AAG TAC CAC ATC CTA AAT ACC CTC CAG TGC TCT GAG   963
Glu Ala Leu Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu   315

GCC ATC ACT GGA GGA GCC GTG TTT GAG ACC ATG GAA GGA AAC ACT ATT  1011
Ala Ile Thr Gly Gly Ala Val Phe Glu Thr Met Glu Gly Asn Thr Ile   331

GAG ATA GGG TGC GAA GGG GAC AGT ATC TCC ATT AAC GGA ATC AAG ATG  1059
Glu Ile Gly Cys Glu Gly Asp Ser Ile Ser Ile Asn Gly Ile Lys Met   347

GTG AAC AAG AAA GAC ATT GTG ACT AAG AAT GGT GTC ATC CAC CTG ATT  1107
Val Asn Lys Lys Asp Ile Val Thr Lys Asn Gly Val Ile His Leu Ile   363

GAT GAA GTC CTC ATT CCT GAT TCT GCC AAA CAA GTT ATT GAG CTG GCT  1155
Asp Glu Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu Leu Ala   379

GGA AAA CAG CAA ACC ACT TTC ACC GAC CTG GTA GCC CAA TTA GGC TTG  1203
Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln Leu Gly Leu   395

GCA TCC TCT CTG AAG CCA GAT GGA GAG TAC ACC TTA TTA GCA CCT GTG  1251
Ala Ser Ser Leu Lys Pro Asp Gly Glu Tyr Thr Leu Leu Ala Pro Val   411

AAC AAT GCG TTC TCT GAT GAC ACT CTG AGC ATG GAC CAA CGC CTT CTT  1299
Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met Asp Gln Arg Leu Leu   427

AAG CTA ATT CTG CAA AAT CAC ATA TTG AAA GTA AAA GTT GGC CTT AGC  1347
Lys Leu Ile Leu Gln Asn His Ile Leu Lys Val Lys Val Gly Leu Ser   443

GAC CTC TAC AAT GGA CAG ATA CTG GAA ACC ATT GGA GGC AAA CAA CTC  1395
Asp Leu Tyr Asn Gly Gln Ile Leu Glu Thr Ile Gly Gly Lys Gln Leu   459

CGA GTC TTT GTG TAT CGG ACG GCT ATC TGC ATA GAA AAC TCA TGC ATG  1443
Arg Val Phe Val Tyr Arg Thr Ala Ile Cys Ile Glu Asn Ser Cys Met   475

GTG AGA GGA AGC AAG CAG GGA AGG AAT GGT GCC ATT CAC ATA TTC CGA  1491
Val Arg Gly Ser Lys Gln Gly Arg Asn Gly Ala Ile His Ile Phe Arg   491

GAA ATC ATC CAA CCA GCA GAG AAA TCC CTG CAC GAC AAG CTG CGG CAA  1539
Glu Ile Ile Gln Pro Ala Glu Lys Ser Leu His Asp Lys Leu Arg Gln   507

GAC AAG CGC TTT AGC ATC TTC CTC AGC CTC CTT GAA GCT GCA GAT TTG  1587
Asp Lys Arg Phe Ser Ile Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu   523

AAA GAT CTC CTG ACA CAG CCC GGA GAT TGG ACC TTG TTT GCA CCA ACC  1635
Lys Asp Leu Leu Thr Gln Pro Gly Asp Trp Thr Leu Phe Ala Pro Thr   539

AAT GAT GCC TTC AAG GGA ATG ACT AGC GAA GAA AGG GAG CTT CTG ATT  1683
Asn Asp Ala Phe Lys Gly Met Thr Ser Glu Glu Arg Glu Leu Leu Ile   555

GGG GAT AAA AAT GCT CTC CAA AAC ATC ATT CTT TAT CAC CTG ACC CCA  1731
Gly Asp Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro   571

GGG GTT TAT ATT GGA AAG GGA TTC GAA CCC GGA GTC ACT AAT ATC CTG  1779
Gly Val Tyr Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu   587

AAG ACC ACA CAG GGA AGC AAA ATC TAT CTG AAA GGA GTA AAC GAA ACG  1827
Lys Thr Thr Gln Gly Ser Lys Ile Tyr Leu Lys Gly Val Asn Glu Thr   603

CTT CTA GTG AAT GAG TTG AAG TCC AAA GAA TCT GAC ATC ATG ACG ACA  1875
Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met Thr Thr   619

AAT GGT GTC ATC CAC GTC GTG GAC AAA CTC CTC TAT CCA GCA GAT ATT  1923
Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro Ala Asp Ile   635

CCA GTT GGA AAT GAT CAG CTC TTG GAA TTA CTG AAC AAA CTG ATA AAA  1971
Pro Val Gly Asn Asp Gln Leu Leu Glu Leu Leu Asn Lys Leu Ile Lys   651

TAC ATC CAA ATC AAG TTT GTT CGT GGC AGC ACC TTC AAA GAA ATC CCC  2019
Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr Phe Lys Glu Ile Pro   667

ATG ACT GTC TAT AGA CCT GCA ATG ACG AAG ATC CAA ATT GAA GGT GAT  2067
Met Thr Val Tyr Arg Pro Ala Met Thr Lys Ile Gln Ile Glu Gly Asp   683

CCC GAC TTC AGG CTG ATT AAA GAA GGC GAA ACG GTG ACA GAA GTG ATC  2115
Pro Asp Phe Arg Leu Ile Lys Glu Gly Glu Thr Val Thr Glu Val Ile   699

CAC GGA GAG CCA GTC ATT AAA AAG TAC ACC AAA ATC ATA GAT GGA GTT  2163
His Gly Glu Pro Val Ile Lys Lys Tyr Thr Lys Ile Ile Asp Gly Val   715

CCT GTT GAA ATA ACT GAA AAA CAG ACT CGG GAA GAA CGA ATC ATT ACA  2211
Pro Val Glu Ile Thr Glu Lys Gln Thr Arg Glu Glu Arg Ile Ile Thr   731

GGT CCT GAG ATA AAA TAT ACC AGG ATT TCC ACA GGA GGT GGA GAA ACA  2259
Gly Pro Glu Ile Lys Tyr Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr   747

GGA GAG ACC TTG CAG AAA TTC TTG CAA AAA GAG GTC TCC AAG GTC ACA  2307
Gly Glu Thr Leu Gln Lys Phe Leu Gln Lys Glu Val Ser Lys Val Thr   763

AAG TTC ATT GAA GGT GGC GAT GGT CAC TTA TTT GAA GAT GAG GAG ATT  2355
Lys Phe Ile Glu Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile   779

AAA AGA CTG CTT CAG GGA GAC ACA CCT GCA AAG AAG ATA CCA GCC AAC  2403
Lys Arg Leu Leu Gln Gly Asp Thr Pro Ala Lys Lys Ile Pro Ala Asn   795

AAA AGG GTT CAA GGG CCT AGA AGA CGA TCA AGA GAA GGC CGT TCT CAG  2451
Lys Arg Val Gln Gly Pro Arg Arg Arg Ser Arg Glu Gly Arg Ser Gln   811

TGAAAACCCAGAGGCCAGACCACAGAGTTTATATAATCCTAAATCAACGATCTGATTTTAAGG  2514
***
GAAATTGTAAGAGCCACCACACTGACTTCAGAATCTGAAATGACAACCAACAGAAGCCAATCT  2577
TCAAGCAAGTCCAAACACAGAGTTCATGTCTTTGTTTCTGCATGAGAAATATAAGAAAATGAT  2640
AGCTAGTCTCCTGTGGGGTAGGAACTGAGGAAATATAGGACCATGCAGGGATTTTATCTCAAT  2703
GAGAAAACTTCTGATTAAAGTAGAATCCACCAAAGAACATCATTGTGACTGGGTCCATACAGC  2766
TAAGTCTTTGCACAGTAAAAACCTTCCGCCTCAGGAAGAGGCTGGAAAAACCCAAAGCACACA  2829
GTTACCTTTCCAGGGGAGGCTAAGGTATCAAAAGGGGTGTTCAGTTATACAACATGCAAACAA  2892
ACCTACCAAATTACGAACAGTGGTGTTACATATTTCTCATGCAATGTGGGTTTCCTGCTAAAT  2955
TTTGTTATTTTTACACTTGATTTATATCCTCGAGATGATTGTCATAAGCTTCTTGCAATACAA  3018
ATGTTTTCTCTCAAACATTTCAATAAAACCATTCTTCAGGTATAAAGAGAATTACTGCAGAGT  3081
TGGTAATTCAGAAAACTCAAGGTTTAAGTTAAAAGTGAGTTTAGACTTTGGAATAGGACTTCA  3144
TACCTTTTTTTATTGTTAACAAGTACTCAATAAAGTAAACTGA                      3187
```

Figure 1 Nucleotide and predicted amino acid sequences of mouse OSF-2

The nucleotide sequence of the cDNA insert from clone pMC163 was determined from both strands. The deduced amino acid sequence is displayed below the DNA sequence. Numbering of amino acids is relative to the N-terminal methionine (position 1). The solid arrowhead indicates the putative signal sequence-cleavage site inferred as described by von Heijne [31]. The potential *N*-glycosylation site is underlined. The boxed areas indicate the weakly conserved fourfold-repeat domains of OSF-2 and correspond to domains R1, R2, R3 and R4 in Figure 2. The poly(A) addition site is doubly underlined. The EMBL databank accession number is D13664.

Regulation of the expression of OSF-2 mRNA in MC3T3-E1 cells by several osteotrophic factors

A variety of growth factors, hormones and cytokines are involved in the regulation of bone turnover (reviewed in [35,36]). We examined whether some of these osteotrophic factors could regulate OSF-2 mRNA expression in MC3T3-E1 cells. Identically seeded MC3T3-E1 cells were treated with different factors for 24 h. Cytoplasmic RNA was extracted and RNA dot-blot analysis was performed (Figure 6). FCS, epidermal growth

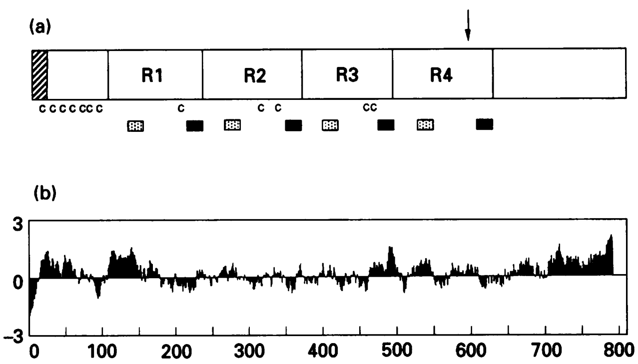

Figure 2 Domain structure (a) and hydropathy profile (b) of mouse OSF-2

The shaded area at the N-terminus indicates the signal sequence. R1, R2, R3 and R4 indicate the fourfold-repeating domains, followed by the C-terminal domain. C denotes the position of cysteines and the arrow indicates a potential *N*-glycosylation site. Stippled and solid boxes show the two particularly well conserved regions (corresponding to the boxed areas in Figure 8). The hydropathy profile (**b**) was calculated by the method of Kyte and Doolittle [33] using a window of 19 residues. Values above the centre line indicate hydrophilic regions and below the line hydrophobic regions.

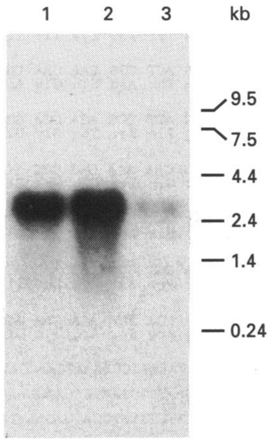

Figure 3 Northern-blot analysis of mouse OSF-2 mRNA

Cytoplasmic RNA was isolated from MC3T3-E1 cells, which were pretreated as follows: lane 1, cells constantly grown in the presence of 10% FCS; lane 2, cells were kept for 2 days in the absence of FCS; lane 3, cells were incubated in the absence of FCS for 24 h and subsequently grown in the presence of 10% FCS for 24 h. Gel electrophoresis, blotting and hybridization were carried out as described in the Materials and methods section. The size standards are indicated.

factor (EGF) and 1,25-dihydroxyvitamin D_3 [1,25-$(OH)_2D_3$] decreased the OSF-2 mRNA levels. As serum is known to contain many growth-associated factors, including EGF and 1,25-$(OH)_2D_3$, the serum-induced down-regulation of OSF-2 mRNA is probably caused by these factors. Transforming growth factor-β_1 (TGF-β_1)-treated cells contained slightly elevated OSF-2 mRNA levels. 17β-Oestradiol, insulin-like growth factor I (IGF-I), platelet-derived growth factor (PDGF), prostaglandin E_2 (PGE_2), parathyroid hormone (PTH) and retinoic acid, however, had no effect. The result of this experiment suggests that OSF-2 mRNA expression is at least partly regulated by several osteotrophic factors.

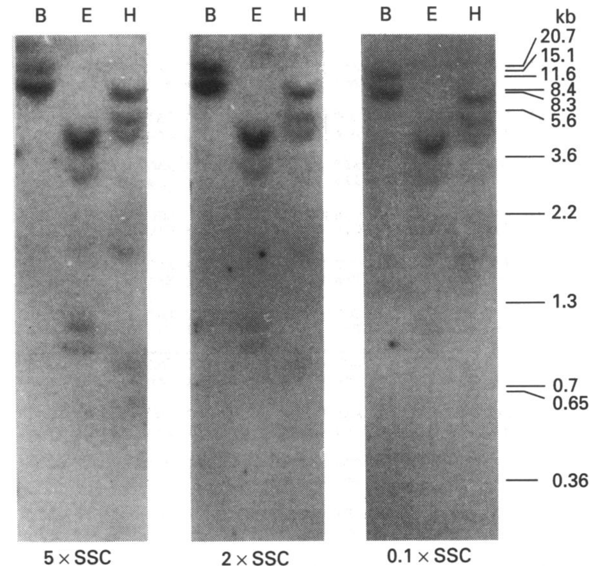

Figure 4 Genomic Southern-blot analyses of mouse OSF-2

DNA (10 μg) isolated from mouse liver was digested completely with *Bam*HI (B), *Eco*RI (E) and *Hin*dIII (H). Fragments were separated by electrophoresis on a 0.8% agarose gel, blotted on to a nylon filter and hybridized with a mouse OSF-2 cDNA probe. Hybridization was carried out for 24 h at 42 °C in a solution consisting of 6 × SSC, 50 mM NaH_2PO_4, 10 × Denhardt's solution, 1% SDS, denatured salmon sperm DNA at 100 μg/ml and 30% formamide. Washing was performed at 60 °C in the indicated SSC solution, which contained 0.1% SDS. When washed with 5 × SSC, 2 × SSC and 0.1 × SSC, it is assumed that hybridizing DNA fragments have sequence similarities of about 60%, 65% and 90% respectively [34]. Sizes of M_r markers are indicated on the right.

Cloning of human OSF-2 cDNA

Screening of human placental and osteosarcoma cDNA libraries under stringent conditions with the mouse OSF-2 cDNA as a probe resulted in a large number of positive clones (see the Materials and methods section). The longest hybridization positive insert from each library was subcloned into plasmid vectors: pKOT133 encodes human OSF-2 from the placenta library screen (hOSF-2pl) and pKOT158 encodes human OSF-2 from the osteosarcoma library screen (hOSF-2os). Determination of the DNA sequences of these clones revealed two different cDNA forms of human OSF-2 cDNAs (not shown). The 3077 bp insert of pKOT133 consists of a 27 bp 5′ untranslated region, a coding region of 2340 bp and a 3′ untranslated region of 710 bp. The 3213 bp insert of pKOT158 consists of an 11 bp 5′ untranslated region, a coding region of 2511 bp and a 3′ untranslated region of 691 bp. Both sequences lack a poly(A) stretch but the polyadenylation signal is present in both cDNAs. The hOSF-2pl open reading frame encodes a protein of 779 amino acids with an M_r of 87037 and the hOSF-2os open reading frame encodes a protein of 836 amino acids with an M_r of 93331. Figure 7 shows the alignment of the deduced amino acid sequences of mouse and human OSF-2 proteins. Compared with mouse OSF-2, hOSF-2os has an insertion of 27 amino acids and hOSF-2pl has a deletion of 57 amino acids within the C-terminal domain. Partial sequencing of three other independently derived clones from both libraries (two from the placenta and one from the osteosarcoma library) revealed a second site (residues 785–812 in Figure 7) within the C-terminal domain at which variations are observed. All differences at the two variable sites constitute in-frame deletions or insertions, strongly suggesting that the isolated

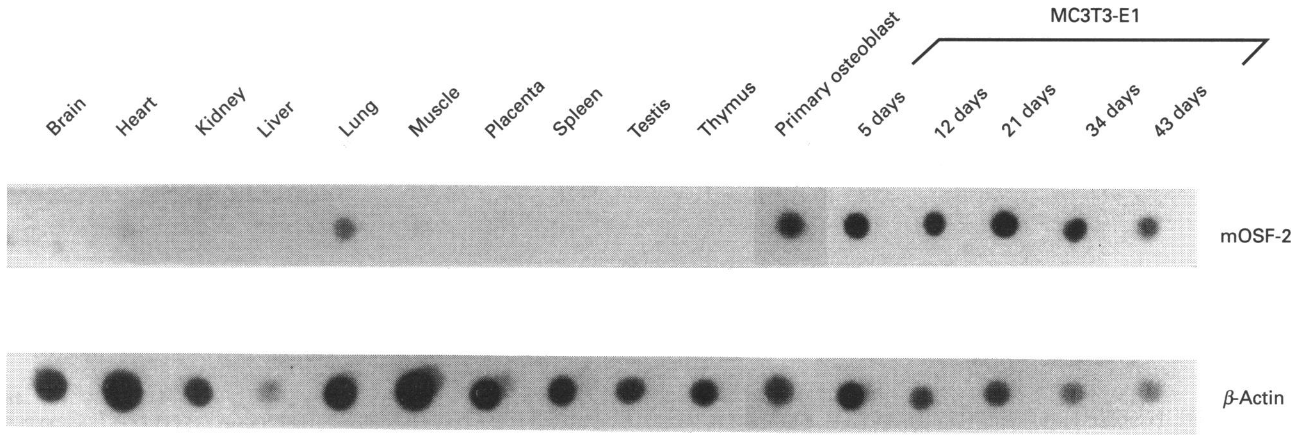

Figure 5 Tissue-specific expression of mouse OSF-2 mRNA

Total RNA was isolated from the indicated tissues by the guanidinium thiocyanate method. Cytoplasmic RNA was isolated from mouse calvarial osteoblast-enriched cells and from MC3T3-E1 cells grown for the indicated times. RNA was analysed by dot filter hybridization. The filters were hybridized with the randomly primed mouse OSF-2 cDNA probe isolated from pMC163 or with a β-actin genomic DNA probe.

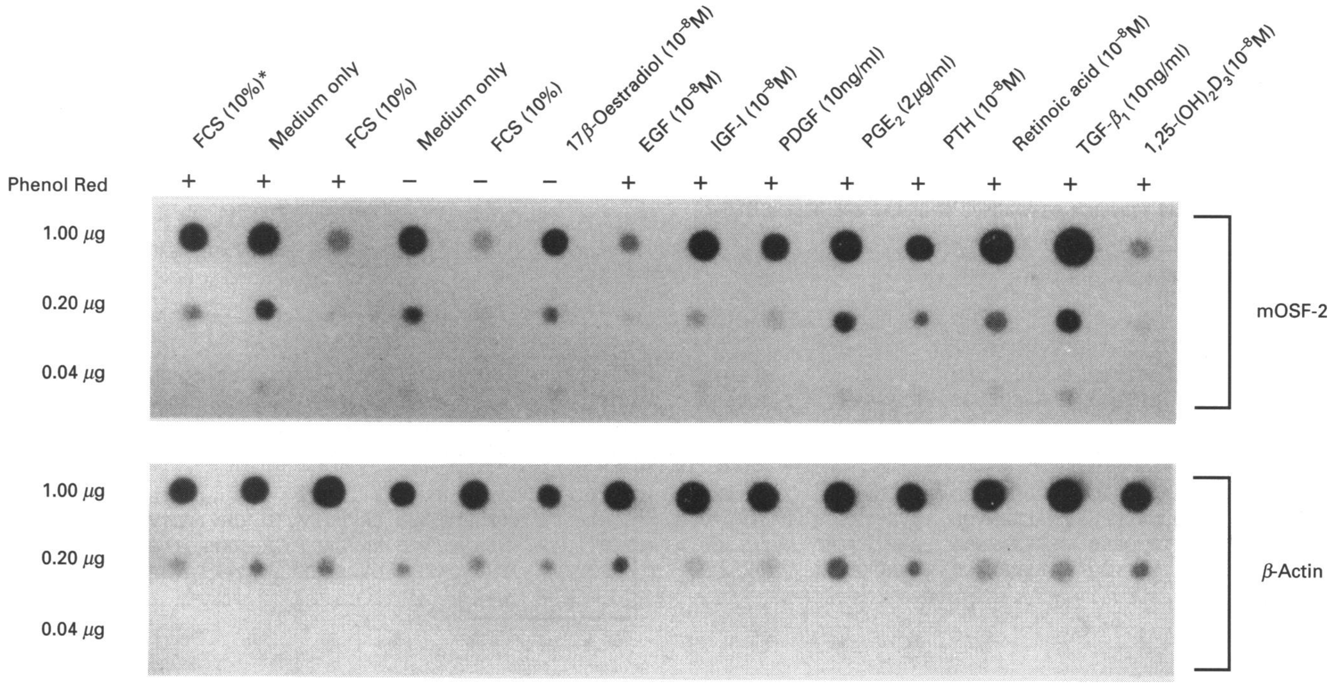

Figure 6 Regulation of mouse OSF-2 (mOSF-2) mRNA expression in MC3T3-E1 cells

Cells were seeded, preincubated and treated with the indicated factors at the indicated concentrations for 24 h. RNA was extracted and dot-blot hybridization was performed as described in the Materials and methods sections. 'FCS(10%)*', cells were treated as described in the legend to Figure 3, lane 1; 'medium only', cells were treated as described in the legend to Figure 3, lane 2; 'FCS(10%)*', cells were treated as described in the legend to Figure 3, lane 3. Hybridization probes were the same as described in the legend to Figure 4. The presence (+) or absence (−) of Phenol Red in the commercial αMEM is indicated (this compound has 17β-oestradiol-like activity [37] and was therefore omitted from some of the tests).

cDNAs reflect individual alternative splicing events. Except for this pattern of insertions or deletions, the sequences of the human OSF-2 cDNAs are identical. OSF-2 is highly conserved between mouse and human. Comparison of the deduced amino acid sequences between the two species shows an identity of 89.2% for the entire protein and 90.1% for the mature form. Compared with other regions in the mature protein, the C-terminal regions are a little less conserved, showing 85.5% identity.

```
                                                                   60                                                          120
mOSF-2    MVPLLPLYALLLLFLCDINPANANSYYDKVLAHSRIRGRDQGPNVCALQQILGTKKKYFSSCKNWYQGAICGKKTTVLYECCPGYMRMEGMKGCPAVMPIDHVYGTLGIVGATTTQHYSD
hOSF-2os  .I.F..MFS....LI--V..I...NH...I..............................T.....KKS...Q........................L..................R...
hOSF-2pl  .I.F..MFS....LI--V..I...NH...I..............................T.....KKS...Q........................L..................R...

                                                                  180                                                          240
mOSF-2    VSKLREEIEGKGSYTYFAPSNEAWENLDSDIRRGLENNVNVELLNALHSHMVNKRMLTKDLKHGMVIPSMYNNLGLFINHYPNGVVTVNCARVIHGNQIATNGVVHVIDRVLTQIGTSIQ
hOSF-2os  A............F..........D...........S..............I..........N..I..........................I...........................
hOSF-2pl  A............F..........D...........S..............I..........N..I..........................I...........................

                                                                  300                                                          360
mOSF-2    DFLEAEDDLSSFRAAAITSDLLESLGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHILNTLQCSEAITGGAVFETMEGNTIEIGCEGDSISINGIKMVNKKDIVTKNGVI
hOSF-2os  ..I.................I..A...........................F.......................S.M.......L.........D...TV..............N....
hOSF-2pl  ..I.................I..A...........................F.......................S.M.......L.........D...TV..............N....

                                                                  420                                                          480
mOSF-2    HLIDEVLIPDSAKQVIELAGKQQTTFTDLVAQLGLASSLKPDGEYTLLAPVNNAFSDDTLSMDQRLLKLILQNHILKVKVGLSDLYNGQILETIGGKQLRVFVYRTAICIENSCMVRGSK
hOSF-2os  ....Q................................A.R.....................V....................NE.......................V.......EK...
hOSF-2pl  ....Q................................A.R.....................V....................NE.......................V.......EK...

                                                                  540                                                          600
mOSF-2    QGRNGAIHIFREIIQPAEKSLHDKLRQDKRFSIFLSLLEAADLKDLLTQPGDWTLFAPTNDAFKGMTSEERELLIGDKNALQNIILYHLTPGVYIGKGFEPGVTNILKTTQGSKIYLKGV
hOSF-2os  ..............K.......E..K......T...........E...........V.............K.I..R.................F.....................F..E.
hOSF-2pl  ..............K.......E..K......T...........E...........V.............K.I..R.................F.....................F..E.

                                                                  660                                                          720
mOSF-2    NETLLVNELKSKESDIMTTNGVIHVVDKLLYPADIPVGNDQLLELLNKLIKYIQIKFVRGSTFKEIPMTVYR---------------------------PAMTKIQIEGDPDFRLIKEGE
hOSF-2os  .D................................T.........I......................V...TTKIITKVVEPKIKVIEGSLQPIIKTEG.TL..VK...E.E........
hOSF-2pl  .D................................T.........I......................V...K------------------------------------------------

                                                                  780                                                          838
mOSF-2    TVTEVIHGEPVIKKYTKIIDGVPVEITEKQTREERIITGPEIKYTRISTGGGETGETLQKFLQKEVSKVTKFIEGGDGHLFEDEEIKRLLQGDTPAKKIPANKRVQGPRRRSREGRSQ
hOSF-2os  .I........I..................E........................E...K.L..E..T............................VR.LQ...K...S...L......
hOSF-2pl  ---------.I..................E........................E...K.L..E..T............................VR.LQ...K...S...L......
```

Figure 7 Amino acid sequence comparison of mouse and human OSF-2

Amino acid sequences were aligned to give maximal identity. Dashes indicate amino acid deletions. The EMBL databank accession numbers for the cDNA encoding hOSF-2os and hOSF-2pl are D13666 and D13665 respectively.

```
mOSF-2 R1   .GIVGATTTQ H......YSD V......... ..SKLREEIE GKGSYTYFAP SNEAWENLDS D....IRRGL ENNVNVELLN ALHSHMVNKR
mOSF-2 R2   ....GTSIQD FLEAEDDLSS FRAAAIT... ..SDLLESLG RDGHFTLFAP TNEAFEKLPR GVLERIMGDK VASEALMKYH ILNTLQCSEA
mOSF-2 R3   ......SAKQ VIE....LAG KQQTTFTDLV AQLGLASSLK PDGEYTLLAP VNNAFSDDTL SMDQRLLKLI LQNHILKVKV GLSDLYNGQI
mOSF-2 R4   ....EKSLHD KLRQDKRFSI FLSLLEA... ..ADLKDLLT QPGDWTLFAP TNDAFKGMTS EERELLIGDK NALQNIILYH LTPGVYIGKG

dFas-1 R1   ..AAAADLAD KLRDDSELSQ FYSLLES... ..NQIANSTL SLRSCTIFVP TNEAFQRYKS KT..AHV... .......LYH ITTEAYTQKR
dFas-1 R2   ..NPNALKFL KNAEEFNVDN IGVRTYRSQV TMAKKESVYD AAGQHTFLVP VDEGFKLSAR SSVDGKV.ID GHVIPNTVIF TAAAQHDDPK
dFas-1 R3   IDTTVTQFLQ SFKENAENGA LRKFYEVIMD NGGAVLDDIN SLTEVTILAP SNEAWNSSNI NNVLRD...R NKMRQILNMH IIKDRLNVDK
dFas-1 R4   ....YTTVLG KLESDPMMSD TYKMGKFSHF ..NDQLNN.. TQRRFTYFVP RDKGWQKTEL DYPSAHKKLF MADFSYHSKS ILERHLAISD
Consensus   ---------- KL------S- ---------- ----L----- --G--T-FAP -NEAF----- ---------- ---------H IL--------


mOSF-2 R1   MLTKDLKHGM VIPS....MY NNLGLFIN.. .HYPNGVVTV NC...AR.VI HGNQIATNGV VHVIDRVLTQ I    ( 108 - 235 )
mOSF-2 R2   ..ITGGAVFE TMEG....NT IEIG...... ..CEGDSISI NG...IKMVN KKDIVTKNGV IHLIDEVLIP D    ( 236 - 370 )
mOSF-2 R3   LETIGGKQLR VF......VY RTAI...... ..CIENSCMV RG......SK Q....GRNGA IHIFREIIQP A    ( 371 - 497 )
mOSF-2 R4   FEPGVTNILK TTQG....SK IYL....... .KGVNETLLV NE...LK.SK ESDIMTTNGV IHVVDKLLYP A    ( 498 - 633 )

dFas-1 R1   LPNTVSSDMA GNPP....LY ITK....... .NSNGDIFVN NARIIPSLSV ETNSDGKRQI MHIIDEVLEP L    (  18 - 147 )
dFas-1 R2   ASAAFEDLLK VTVS....FF KQKNGKMYVK SNTIVGDAKH RVGVVLAEIV KANIPVSNGV VHLIHRPLMI .    ( 159 - 312 )
dFas-1 R3   IRQKNANLIA QVPT....VN NNTFLYFNVR GEGSDTVITV EGGGVNATVI QADVAQTNGY VHIIDHVLGV P    ( 313 - 466 )
dFas-1 R4   KEYTMKDLVK FSQESGSVIL PTFRDSLSIR VEEEAGRYVI IWNYKKINVY RPDVECTNGI IHVIDYPLLE E    ( 467 - 619 )
Consensus   ---------- ---------- ---------- ---------V N-------V- --D---TNGV IH-ID-VL-P -
```

Figure 8 Comparison of the conserved repeat sequences in mouse OSF-2 and *Drosophila* fasciclin I

Position number of the first and last residue of each domain is given in parentheses after each sequence; position number 1 corresponds to the first methionine. Gaps (.) were introduced to get maximum sequence similarity using the Pileup program of GCG [29]. Amino acid residues identical in at least four positions are indicated in the line of consensus sequence. Two particularly conserved regions are boxed. dFas-1, *Drosophila* fasciclin I [19]. R1, R2, R3 and R4 correspond to the fourfold repeat domains indicated in Figures 1 and 2.

Structural similarity of OSF-2 to the insect protein fasciclin I

Database homology search using the whole mouse OSF-2 sequence did not shown any significant sequence similarity to other known proteins, but by limiting the query sequence to the conserved regions, the search revealed sequence similarity to fasciclin I (Figure 8). Fasciclin I has a similar protein organization to OSF-2. It also has a fourfold repeat structure of similar size (approx. 150 amino acids each) with weak sequence similarity (7–15% identity) with each other including highly conserved amino acid 'regions' (up to 45% identity between regions) within the fourfold repeat structure [19]. The two particularly conserved regions (see Figure 8) found in OSF-2 are well conserved in fasciclin I. The amino acid sequence identity between these two particularly conserved regions of OSF-2 and fasciclin I is 41.4% and 43.3% on average respectively. In addition, weak sequence similarity is observed for the entire fourfold repeat structure of OSF-2 and fasciclin I (Figure 8). In contrast with fasciclin I [22], OSF-2 has no glycosylphosphatidylinositol lipid membrane anchor site.

DISCUSSION

Using a combined subtraction hybridization/differential screening approach to distinguish between mRNA expressed in the mouse osteoblastic cell line MC3T3-E1, but not in NIH3T3 cells, several new cDNAs have been isolated that are selectively expressed in MC3T3-E1 cells. One of these cDNAs encodes OSF-2. Mouse tissue RNA dot-blot analysis showed that OSF-2 is strongly expressed in bone and weakly in lung, but not in other tissues. Further experiments showed that the expression of OSF-2 is negatively regulated by the osteotrophic factors EGF and 1,25-$(OH)_2D_3$ and up-regulated by TGF-β_1. In contrast, 17β-oestradiol, IGF-I, PDGF, PGE_2, PTH and retinoic acid had no effect on the OSF-2 mRNA expression in MC3T3-E1 cells. Southern-blot analysis indicated that the OSF-2 gene is present as a single copy in the mouse genome.

Using the mouse OSF-2 cDNA as a probe, human osteosarcoma and placental libraries were screened. Many positive clones were isolated, indicating that OSF-2 cDNA is abundantly present in these libraries. One complete human OSF-2 cDNA from each library was characterized in detail. Except for a variation within the C-terminal domain, the deduced amino acid sequences of both human OSF-2 proteins are identical. Subsequently, three further human OSF-2 cDNA clones were analysed and a second site of variability within the C-terminal domain was identified. In total, five different forms of human OSF-2 were isolated. All differences constitute in-frame deletions or insertions, implying alternative splicing events. The reason for the high degree of variability is at present unknown, but it is interesting to note that all splicing events occur within the C-terminal domain, indicating that the other protein domains might be essential for the biological role of OSF-2.

Mouse and human OSF-2 are highly conserved. The common protein structure shows a typical signal sequence, followed by a cysteine-rich domain, a fourfold repeated domain and a C-terminal domain, but lacks a transmembrane region. Alignment of each repeat unit reveals that there are two particularly conserved regions of 13 and 14 amino acids each (Figure 8).

Computer search revealed amino acid similarity between OSF-2 and fasciclin I, a homophilic adhesion protein involved in the neuron growth cone guidance during development of *Drosophila* and grasshopper embryos [19–23]. Fasciclin I and OSF-2 have a similar protein structure organization, characterized by a weakly conserved fourfold repeat structure. Each repeat has approx. 130 amino acids in OSF-2 and approx. 150 amino acids in fasciclin I. Within each of the fourfold repeat units, two particularly conserved regions are located in which the strong similarity between OSF-2 and fasciclin I is observed. Neither OSF-2 nor fasciclin I possess typical transmembrane domains. Based on these similarities and the amino acid sequence similarity, we suggest that OSF-2 and fasciclin I are derived from a common ancestor and have similar protein structures and functions.

Fasciclin I has a glycosylphosphatidylinositol lipid moiety [22] which facilitates membrane association (reviewed in [38,39]). We could not detect a potential glycosylphosphatidylinositol membrane anchor site at the corresponding region of OSF-2. However, when we analysed the C-terminal sequence of mouse OSF-2 by the methods of Berzofsky *et al.* [40] and Vogel and Jähnig [41], one site possibly involved in the formation of an amphipathic α-helix (residues 774–795 in Figure 7) and two sites possibly involved in the formation of β-strands (residues 806–815 and 819–829 in Figure 7) were found. These sites might be involved in the membrane association and may compensate for the lack of a glycosylphosphatidylinositol membrane anchor sequence.

The five different forms of human OSF-2 cDNA, probably the result of alternative splicing, may encode membrane-bound and secreted variants of the protein. This is conceivable, since all five splicing events occurred in the C-terminal domain leaving the fourfold repeat structure unchanged. Alternatively, the splicing may alter the binding specificity of OSF-2, as has been suggested in the case of fasciclin I [23]. The alternative splicing in fasciclin I, however, results in variation in the second repeat unit and therefore might not serve as a model to explain the biological relevance of alternative splicing of OSF-2. More experiments are needed to ascertain the detailed biological role of OSF-2.

We thank M. Kiefer for the osteosarcoma cDNA library, R. Iwase for the synthesis of oligodeoxynucleotides, S. Kawai and M. Okazaki for help with DNA sequencing, T. Hashimoto-Gotoh for his encouragement in the initial stage of this work, and E. Konz and H. B. Maruyama for support.

REFERENCES

1 Hynes, R. O. and Lander, A. D. (1992) Cell **68**, 303–322

2 Horton, M. A., Lewis, D., McNulty, K., Pringle, J. A. S. and Chambers, T. J. (1985) Cancer Res. **45**, 5663–5669

3 Horton, M. A. and Chambers, T. J. (1986) Br. J. Exp. Pathol. **67**, 95–104

4 Athanasou, N. A., Quinn, J. and McGee, J. D. (1989) in Leucocyte Typing IV (Knapp, W., ed), pp. 921–923, Oxford University Press, Oxford

5 Soligo, D., Parravicini, C. L., Luksch, R., Quirrici, N., Vago, L. and Lambertenghi-Deliliers (1989) in Leucocyte Typing IV (Knapp, W., ed.), pp. 1029–1032, Oxford University Press, Oxford

6 Argraves, W. S., Suzuki, S., Arai, H., Thompson, K., Pierschbacher, M. D. and Ruoslahti, E. (1987) J. Cell Biol. **105**, 1183–1190

7 Horton, M. A. and Davies, J. (1989) J. Bone Miner. Res. **4**, 803–808

8 Ruoslahti, E. and Pierschbacher, M. D. (1986) Cell **44**, 517–518

9 Ruoslahti, E. and Pierschbacher, M. D. (1987) Science **238**, 491–497

10 Pytela, R., Pierschbacher, M. D. and Ruoslahti, E. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5766–5770

11 Cheresh, D. A. and Spiro, R. C. (1987) J. Biol. Chem. **262**, 17703–17711

12 Oldberg, Å., Franzén, A. and Heinegård, D. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 8819–8823

13 Butler, W. T. (1989) Connect. Tissue Res. **23**, 123–136

14 Clezardin, P., Jouishomme, H., Chavassieux, P. and Marie, P. J. (1989) Eur. J. Biochem. **181**, 721–726

15 Gehron Robey, P., Young, M. F., Fisher, L. W. and McClain, T. D. (1989) J. Cell. Biol. **108**, 719–727

16 Helfrich, M. H., Nesbitt, S. A., Dorey, E. L. and Horton, M. A. (1992) J. Bone Miner. Res. **7**, 335–343

17 Vukicevic, S., Luyten, F. P., Kleinman, H. K. and Reddi, A. H. (1990) Cell **63**, 437–445

18 Tezuka, K., Takeshita, S., Hakeda, Y., Kumegawa, M., Kikuno, R. and Hashimoto-Gotoh, T. (1990) Biochem. Biophys. Res. Commun. **173**, 246–251

19 Zinn, K., McAllister, L. and Goodman, C. S. (1988) Cell **53**, 577–587

20 Elkins, T., Hortsch, M., Bieber, A. J., Snow, P. M. and Goodman, C. S. (1990) J. Cell Biol. **110**, 1825–1832

21 Elkins, T., Zinn, K., McAllister, L., Hoffmann, F. M. and Goodman, C. S. (1990) Cell **60**, 565–575

22 Hortsch, M. and Goodman, C. S. (1990) J. Biol. Chem. **265**, 15104–15109

23 McAllister, L., Rehm, E. J., Goodman, C. S. and Zinn, K. (1992) J. Neurosci. **12**, 895–905

24 Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988) Science **239**, 487–491

25 Takeshita, S., Sato, M., Toba, M., Masahashi, W. and Hashimoto-Gotoh, T. (1987) Gene **61**, 63–74

26 Kiefer, M. C., Joh, R. S., Bauer, D. M. and Zapf, J. (1991) Biochem. Biophys. Res. Commun. **176**, 219–225

27 Sambrook, J., Fritsch, E. F. and Maniatis, T. (eds) (1989) in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

28 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467

29 Deveraux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395

30 Kodama, H., Amagai, Y., Sudo, H., Kasai, S. and Yamamoto, S. (1981) Jpn. J. Oral Biol. **23**, 899–901

31 von Heijne, G. (1986) Nucleic Acids Res. **14**, 4683–4690

32 Kozak, M. (1987) Nucleic Acids Res. **15**, 8125–8148

33 Kyte, J. and Doolittle, R. F. (1982) J. Mol. Biol. **157**, 105–132

34 Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. **48**, 443–453

35 Simpson, E. (1984) Trends Biochem. Sci. **9**, 527–530

36 Centrella, M. and Canalis, E. (1985) Endocrinol. Rev. **6**, 544–551

37 Berthois, Y., Katzenellenbogen, J. A. and Katzenellenbogen, B. S. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2496–2500

38 Ferguson, M. A. J. and Williams, A. F. (1988) Annu. Rev. Biochem. **57**, 285–320

39 Low, M. G. and Saltiel, A. R. (1988) Science **239**, 268–275

40 Margalit, H., Spouge, J. L., Cornette, J. L., Cease, K. B., Delisi, C. and Berzofsky, J. A. (1987) J. Immunol. **138**, 2213–2229

41 Vogel, H. and Jähnig, F. (1980) J. Mol. Biol. **190**, 191–199

Received 21 December 1992/5 March 1993; accepted 22 March 1993