Human SCF, a splice variant thereof its pharmaceutical use.

The present invention relates to a novel human stem cell factor (SCF) protein, DNA sequences coding for this protein, its use in therapy, particularly in *in vitro* fertilisation, as well as pharmaceutical formulations comprising such a protein.

Successful embryo implantation requires correct development of the pre-implantation embryo, resulting in a hatched blastocyst which is able to implant into receptive endometrium. A considerable body of data has been collected which supports the idea that soluble growth factors, if secreted by the uterine epithelium, act directly on the embryo to control this process (Pampfer, S. *et al*, Bioessays, **13**: 535-540 (1991); Tartakousky, B., and Ben Yair, E., *Development Biology*, **146**: 345-352 (1991); Anderson, E. D., *J. Cellular Biochem.*, **53**: 280-287 (1993); and Schultz, G. A. and Hevner, S., *Mutat. Res.*, **296**: 17-31 (1992)).

In addition, developing embryos have been shown to produce a variety of cytokines which may act in an autocrine fashion on the endometrium to influence its receptivity. Examples of growth factors shown to be produced by human embryos include IL-1, IL-6, CSF-1 and TNF-α (Zolti *et al*, *Fertil. Steril.*, **56** (1991) 265-272 and Witkin *et al*, *J. Reprod. Immunol.*, **19** (1991) 85-93). TNF-α has been shown to be present in culture medium of human embryos up to the morula stage, but not that from the blastocyst (Lachappelle *et al*, *Human Reproduction*, **8**: 1032-1038 (1993)). Production of cytokines by the embryo may therefore be regulated in a stage-specific manner.

Data on the possible direct effects of cytokines on embryos have come primarily from experiments in mice where many cytokines have been shown to affect the development of pre-implantation embryos *in vitro*. IFN-γ and CSF-1, at physiological concentrations, inhibit the number of embryos developing to the blastocyst stage (Hill *et al*, *J. Immunol.*, **139** (1987) 2250-2254). TNF-α has also been shown to have more subtle effects. Although TNF-α has no apparent effect on rates of blastocyst formation, it appears to specifically inhibit proliferation of cells contributing to the inner cell mass (ICM), which results in blastocysts with a reduced ICM (Pampfer *et al*, *Endocrinology*, **134**: 206-212 (1994)).

Other growth factors also have specific effects on ICM cells. For instance, insulin-like growth factors 1 and 2 stimulate ICM proliferation, whereas leukaemia inhibitory factor (LIF) inhibits their differentiation (Harvey *et al*, *Mol. Reprod. Dev.*, **31** (1992) 195-199).

It has been observed, in mouse systems, that embryos cultured *in vitro* lag in development compared to *in vivo* controls, and exhibit lower pregnancy rates after embryo transfer (Bowman, P. and McLaren, A., *J. Embryol. Exp. Morphol.*, **24**: 203-207 (1970)). Thus, a better understanding of the role of growth factors in development could lead to improved *in vitro* culture conditions and enhance the outcome in human IVF programs.

Stem cell factor (SCF) is a growth factor related in structure to CSF-1, and acts through the c-kit tyrosine kinase receptor. In bone marrow, SCF and CSF-1 act synergistically to promote proliferation and differentiation of stem cells into macrophage colonies.

EP-A-0423980 discloses the nucleic acid sequence of human SCF, and discusses potential uses of SCF in conditions requiring stimulation of cell proliferation, particularly blood cells.

In mouse, c-kit has been shown to be expressed throughout pre-implantation development (Arceci *et al* (1992)). We have now shown that the same is true in human embryos. At certain stages the human embryos also express SCF mRNA, suggesting that this growth factor may act in an autocrine fashion. This is in contrast to mouse, where no expression of SCF was detected in pre-implantation embryos (Arceci *et al* (1992)).

The full length SCF transcript consists of eight exons (Martin, F. H. *et al*, *Cell*, **63**: 203-211 (1990)), which paper also discloses a variant form of SCF. A splice-variant of SCF has also been described which arises by virtue of the loss of exon 6 (Flanagan *et al*, *Cell*, **63**: 1025-1035 (1991)).

There has now been found a further, novel, splice-variant which appears to arise due to the inclusion of a novel exon consisting of 155 base pairs between exons 3 and 4. This also results in a frameshift, and codes for a species of SCF comprising 33 novel amino acids following exon 3, before terminating at an in frame stop codon which now appears in exon 4 due to the frameshift.

Thus, the present invention provides SCF which includes the following C-terminal sequence:

Glu Ile Cys Ser Leu Leu Ile Gly Leu Thr Ala Tyr Lys
Glu Leu Ser Leu Pro Lys Arg Lys Glu Thr Cys Arg Ala

THIS SHEET HAS BEEN REPLACED

Ile Gln His Pro Arg Lys Asp

or a sequence which is substantially homologous thereto.

Preferably, the novel SCF of the invention comprises the first 39 amino acids of full length SCF (not including any signal sequence) followed by the above-noted 57 new amino acids. In one embodiment the novel SCF of the invention has a sequence at positions 1-39 substantially homologous to that shown in figure 2.

At the amino acid level, a protein sequence may be regarded as substantially homologous to another protein sequence if a significant number of the constituent amino acids exhibit homology. At least 40%, 50%, 60%, 70%, 80%, 90%, 95% or even 99%, in increasing order of preference, of the amino acids may be homologous.

Thus, the alternative splicing mechanism can result in the production of a novel SCF in human embryos. Therefore, the novel SCF of the invention can be used in the treatment of pre-implantation embryos to ensure correct differentiation and development prior to implantation in a subject.

In addition, the invention also provides a DNA sequence coding for a protein of the invention which sequence includes a sequence substantially homologous to:

GAA ATC TGT TCA TTG TTG ATA GGG CTG ACG GCC TAT AAG
GAA TTA TCA CTC CCT AAA AGG AAA GAA ACT TGC AGA GCA
ATT CAG CAT CCA AGG AAA GAC TGA

and includes all other nucleic acid sequences which, by

This Sheet Has Been Replaced

virtue of the degeneracy of the genetic code, also code for the given amino acid sequence or which are substantially homologous to such a sequence.

Sequences having substantial homology may be regarded as those which will hybridise to the nucleic acid sequence shown in figure 2 under stringent conditions (for example, at 35 to 65°C in a salt solution of about 0.9M).

DNA constructs comprising DNA sequences of the invention form another aspect of the present invention.

As discussed herein, the protein of the invention is useful in treating embryos to ensure correct development prior to implantation. SCF has been shown to act by binding to the transmembrane receptor c-kit. Furthermore, we have shown that human embryos express c-kit throughout most stages of pre-implantation embryo development.

Thus, in further aspects, the present invention provides:

(a) a method for ensuring the correct development of a pre-implantation embryo which comprises the step of administering the SCF of the present invention to a pre-implantation embryo (and preferably a human embryo); and

(b) a method for ensuring the correct development of a human pre-implantation embryo which comprises the step of administering SCF to a human pre-implantation embryo. In this method, the SCF used can be any of the naturally occurring forms, including previously described variants (Martin *et*

al, supra and Flanagan *et al, supra*), as well as the novel variant described herein.

In addition, the invention also provides the use of SCF in the manufacture of a medicament for use in ensuring correct development in human pre-implantation embryos. Again, any form of SCF can be used to produce a suitable medicament.

The medicament is preferably presented in the form of a pharmaceutical formulation comprising the protein of the invention together with one or more pharmaceutically acceptable carriers and/or excipients. Such pharmaceutical formulations form a yet further aspect of the present invention. Such pharmaceutical formulations represent one way in which the SCF can be used in the methods described above.

The invention will now be described by means of the following examples, which examples should not be construed as in any way limiting the present invention. The examples refer to the following figures which show:

Figure 1: the sequence of the novel exon and the predicted amino acid sequence;

Figure 2: the sequence of human SCF;

Figure 3: agarose gel showing the products of nested RT-PCR amplification on RNA from human embryos. Each panel shows the products of amplification with primers specific for different cDNA targets. Amplified cDNAs from different embryos were loaded in each

lane. Lanes are labelled according to cDNA labels in Table 1 (below). Additional samples were: lane p, first trimester trophoblast; lane q, cDNA from 200 BeWO cells; lane r, 10 ng human genomic DNA; and lane s, no input cDNA, as a negative control. DNA molecular weight markers were a 123 base pair ladder loaded in lane i. The sizes of the expected PCR products are shown in bp.

Table 1

Human embryo cDNAs and controls

name	stage of development
a	2 cell
b	3 cell
c	4 cell
d	6 cell
e	8 cell
f	morula
g	blastocyst
h	culture supernatant for a to g
j	three pooled blastocysts
k	culture supernatant for j
l	2 x 6 cell and 1 x 8 cell
m	culture supernatant for l
n	1 x 4 cell and 1 x 6 cell
o	culture supernatant for n

samples a to h are from the same donor.

Figure 4: primers used for RT-PCR, outer pain A and B, inner pain C and D.

EXAMPLE 1: **Embryo Culture and RNA Extraction**

Crypopreserved human embryos which had been fertilised as part of an IVF program were used in this study. These embryos had been donated for research purposes by the parents and this study complied with the requirements of the Human Embryology and Fertilisation Authority, and the local ethical committee. Frozen embryos were thawed and cultured in Earles balanced salts medium supplemented with 0.4% human serum albumin (Armour Pharmaceuticals UK), until the required developmental stage, then flash frozen in liquid nitrogen in 5 μl of culture fluid (and thus lysed by ice crystals). An identical volume of culture supernatant was frozen as a control. Any remaining cumulus cells were removed during routine handling.

Total RNA from first trimester trophoblast was isolated by the method of Chomsczynski and Sacchi, *Anal. Biochem.*, **162**: 156-159 (1987) in which frozen tissue is homogenised in 5 ml of buffer containing 4 M guanidinium thiocyanate (Gibco BRL Livingston, Scotland), 25 mM sodium citrate pH 7.0, 0.5% sarcosyl and 0.1 M 2-mercaptoethanol. The lysate was acidified by the addition of 0.5 ml of 2 M sodium acetate pH 4, and phenol-chloroform extracted using 5 ml of buffer saturated phenol and 1 ml chloroform-isoamylalcohol (49:1 v/v). The suspension was placed on ice for 15 minutes and centrifuged at 10,000 g for 20 minutes at 4°C. The aqueous phase containing RNA was precipitated, washed twice in 70% ethanol, dried and resuspended in TE (10 mM Tris-HCl pH 7.4 and 1 mM EDTA). The concentration of RNA was determined spectrophotometrically at 260 nm.

RNA was prepared from single human embryos using a scaled down protocol based on the above procedure. To assist precipitation of the RNA 100 μg of carrier yeast tRNA (Gibco BRL, Livingston, Scotland) was added at the homogenisation step. The remaining details are as described above, except that all the volumes were 50 fold less and the whole procedure was carried out in 400 μl Eppendorf tubes.

EXAMPLE 2: **Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)**

cDNA was synthesised from half the total RNA from each embryo using AMV reverse transcriptase (Super RT, HT Biotech, Cambridge, UK). 3-5 micrograms of RNA was primed with oligo dT (Pharmacia), according to the manufacturers instructions for 60 minutes at 42°C. PCR amplification of the cDNA preparations was performed as previously described (Sharkey, A. *et al*, *Molecular Endocrinol.*, 6: 1235-1241 (1992)) with a Hybaid Omnigene DNA thermal cycler in a final volume of 30 μl using 1 U of Taq DNA polymerase (Cetus, Emeryville, CA) and 10 μM of each of the pair of external primers (see Figure 4) in the manufacturer's recommended buffer. The following cycle profile was used:- 30s at 95°C, 30s at X°C, 30s at 72°C for 30 cycles, where X is the annealing temperature for each pair of cytokine primers, as shown below.

	External Primers (°C)	Internal Primers (°C)
SCF	54	54
HistRNA	52	59
c-kit	56	56

Oligonucleotide primers

Oligonucleotide primers for SCF, c-kit and Histidyl-t RNA synthetase were synthesised on a Cruachem PS250 DNA synthesiser. Primer sequences were designed from published nucleotide sequences (see Figure 4), such that amplification of any contaminating genomic DNA would result in a differently sized product from the cDNA species.

Because of the small amount of material, two pairs of primers were used for each target cDNA, in a nested PCR protocol. One thirtieth of the cDNA products were amplified using Amplitaq (Cetus), in the manufacturers recommended buffer. Following 30 cycles of PCT using the external primer pair, one fiftieth of the first round reaction was transferred to a fresh tube containing the inner primer pair, and subjected to a further 30 rounds of amplification. As negative control, an equal volume of the culture fluid in which the embryo was grown was extracted and subjected to RT-PCR in the same way. Also, 200 cells of the BeWo cell line (ECACC No 86082803) were extracted as positive control.

The primers used in this study are shown in Figure 4, together with the size of the expected product. The identity of each product was confirmed by cloning and sequencing as described previously (Sharkey *et al*, *Mol. Endocrinol.* (1992)). To ensure that the product detected resulted from amplification of cDNA rather than contaminating genomic DNA, primers were chosen to cross intron/exon boundaries. Ten nanograms of genomic DNA was also subjected to PCR at the same time as the cDNA to verify no product of the expected size resulted from

genomic DNA.

RESULTS

The technique of RT-PCR was applied to total RNA extracted from human embryos produced by *in vitro* fertilisation. Embryos were cultured to the appropriate stage, then quick-frozen in liquid nitrogen. Stored embryos were thawed and total RNA extracted. In order to produce detectable RT-PCR product from total RNA extracted from a single embryo, a nested PCR protocol was employed in which the cDNA was subjected to two sets of PCR amplification with an external primer pair, followed by an internal pair. Primers were based on published cDNA sequences and designed to span intron-exon boundaries so that amplification of contaminating genomic DNA could be readily distinguished from cDNA products.

Initially, cDNA from each embryo was tested with primers for histidyl tRNA synthetase (HistRS) to confirm successful RNA isolation and reverse transcription. The primers used gave rise to weak products of greater than 400 bp from genomic DNA and 110 bp from cDNA derived from HistRS mRNA. Transcripts for Hist RS were detected in mRNA from embryos at all stages of development, as well as in decidua and the choriocarcinoma cell line BeWo, used as positive controls (Figure 3, lanes p and q respectively). No product was detected in an equal volume of embryo culture supernatant extracted and subjected to RT-PCR in the same way, indicating that there was no contamination of the culture with extraneous cDNA or RNA.

Examples of similar RT-PCR analysis with primers for SCF

THIS SHEET HAS BEEN REPLACED

and c-kit are shown in Fig 3. Stocks of cDNA were reverse transcribed from each RNA sample on two separate occasions and the PCR assays were repeated twice on each cDNA stock. The results are shown in Figure 3, which displays the pattern of expression of c-kit and SCF during pre-implantation development. The identity of the PCR fragments of the correct size was confirmed by sequencing of the cloned PCR products. In cases where novel sized products were seen, these were also cloned and sequenced.

For SCF, the predicted fragment is 966 bp. However, the SCF transcripts appeared to show stage-specific differences in size. Upon cloning and sequencing, the new product appeared to arise due to an alternative splicing event which inserts a new exon between exons 3 and 4. The predicted sequence of the novel transcript is shown in Figure 1. The novel splicing pattern also involves a frameshift, giving a total of 57 new amino acids, before an in frame stop codon in exon 4.

In a similar analysis using primers specific for c-kit, the receptor for SCF showed that c-kit was expressed at most stages of human pre-implantation embryo development. This suggests that the embryo has the ability to respond to SCF throughout this period.

DISCUSSION

Many growth factors have been shown to influence the development of cultured pre-implantation mammalian embryos (for review see Anderson, E. D., *J. Cellular Biochem.*, **53**: 280-287 (1993) and Schultz, G. A. and Hevner, S., *Mutat. Res.*, **296**: 17-31 (1992)).

This Sheet Has Been Replaced

However, there is good evidence for species to species differences in expression of growth factor receptors in pre-implantation development. For instance, EGF mRNA is expressed in the pig embryo but has not been found at any stage in mouse pre-implantation embryos (Vaughan *et al*, *Development*, **116**: 663-669 (1992); Rapolee *et al*, *Science*, **241**: 1823-1825 (1988); and Watson, A. J. *et al*, *Biol. Reprod.*, **50**: 725-733 (1994)). Therefore the usefulness of these studies to researchers interested in factors controlling human pre-implantation development is limited. In addition, the specific growth factors and receptors investigated in such studies frequently have been chosen on an *ad hoc* basis. Both for ethical and practical reasons, such an approach is not suitable for use with human embryos. We have therefore used a nested RT-PCR method which has allowed us to screen for the expression of growth factor and receptor mRNAs in single human pre-implantation embryos. This method has been widely used over the last few years in other species since it is reliable, sensitive and economical in its use of embryo material.

RT-PCR with primers for Histidyl-tRNA synthetase was used on cDNA samples to confirm that cDNA had been successfully prepared from each embryo RNA sample. cDNA specific for this housekeeping gene was successfully detected in cDNA samples made even from a single 2-cell embryo, indicating that the method was sufficiently sensitive for this study.

SCF was expressed at the 2-cell stage, and then reappeared at the 6-cell stage. This is consistent with maternal expression followed by re-expression from the

THIS SHEET HAS BEEN REPLACED

embryos's genome at the 6-cell stage (Braude, P. *et al*, *Nature*, 332: 459-461 1988)). SCF transcripts appeared to show stage-specific differences in the transcript size. On cloning and sequencing, these were found to be due to alternative splicing of the primary transcript. Two of these variants were similar to those published previously (Martin *et al*, *supra* and Sharkey, A. *et al*, *Mol. Endocrinol.*, 6: 1235-1241 (1992)), and one was a novel form which predicts a species of SCF with 57 new amino acids at the carboxy terminus. Several variants of SCF are now known, some of which are membrane bound and bioactive. The species expressed by the pre-implantation embryo include those known to be bioactive, and indicates that various forms of SCF can act through c-kit expressed by the embryo, and can affect embryo development at this time.

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED

CLAIMS:

1. SCF which includes the following C-terminal sequence:

Glu Ile Cys Ser Leu Leu Ile Gly Leu Thr Ala Tyr Lys Glu Leu Ser Leu Pro Lys Arg Lys Glu Thr Cys Arg Ala Ile Gln His Pro Arg Lys Asp

or a C-terminal sequence which is substantially homologous thereto.

2. SCF as claimed in claim 1 wherein the C-terminal sequence specified in claim 1 begins at position 40.

3. SCF as claimed in claim 1 or claim 2 wherein the sequence at positions 1 to 39 is substantially homologous to that shown for positions 1 to 39 in figure 2.

4. A DNA sequence coding for a protein as defined in any one of claims 1 to 3.

5. A DNA sequence as claimed in claim 4 which sequence includes a sequence substantially homologous to:

GAA ATC TGT TCA TTG TTG ATA GGG CTG ACG GCC TAT AAG GAA TTA TCA CTC CCT AAA AGG AAA GAA ACT TGC AGA GCA ATT CAG CAT CCA AGG AAA GAC TGA.

6. A DNA consruct comprising a DNA sequence as defined in claim 4 or claim 5.

7. The use of a protein as defined in any one of claims 1 to 3 for use in medicine.

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED

8. The use as claimed in claim 7 wherein the protein is used to ensure the correct development of pre-implantation embryos.

9. The use of a protein as defined in any one of claims 1 to 3 in the manufacture of a medicament for use in ensuring the correct development of pre-implantation embryos.

10. A pharmaceutical formulation comprising a protein as defined in any one of claims 1 to 3 together with one or more pharmaceutically acceptable carriers and/or excipients.

11. A pharmaceutical formulation as claimed in claim 10 which is for use in ensuring the correct development of pre-implantation embryos.

12. A method of ensuring the correct development of pre-implantation embryos which comprises the step of administering to an embryo a protein as defined in any one of claims 1 to 3.

13. A method of ensuring the correct development of human pre-implantation embryos which comprises the step of administering to a human embryo SCF.

14. A method as claimed in claim 13 wherein the SCF is as defined in any one of claims 1 to 3.

This Sheet Has Been Replaced