Structure of the mouse leukaemia inhibitory factor receptor gene: regulated expression of mRNA encoding a soluble receptor isoform from an alternative 5' untranslated region

Ian CHAMBERS^{*1}, Alison COZENS^{*}, Joanne BROADBENT^{*2}, Morag ROBERTSON^{*}, Muriel LEE[†], Meng LI^{*} and Austin SMITH^{*}

^{*}Centre for Genome Research, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JQ, Scotland, U.K., and [†]The MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, Scotland, U.K.

The low-affinity leukaemia inhibitory factor receptor (LIF-R) is a component of cell-surface receptor complexes for the multi-functional cytokines leukaemia inhibitory factor, ciliary neurotrophic factor, oncostatin M and cardiotrophin-1. Both soluble and transmembrane forms of the protein have been described and several LIF-R mRNAs have been reported previously. In order to determine the coding potential of LIF-R mRNAs we have isolated and characterized the mouse LIF-R gene. mRNA encoding soluble LIF-R (sLIF-R) is formed by inclusion of an exon in which polyadenylation signals are provided by a B2 repeat. This exon is located centrally within the LIF-R gene but is excluded from the transmembrane LIF-R mRNA by alternative splicing. The transmembrane receptor is encoded by 19 exons distributed over 38 kb. Two distinct 5' non-coding exons

have been identified, indicating the existence of alternative promoters. One of these is G/C rich and possesses a consensus initiator sequence as well as potential Sp1 binding sites. Expression of exon 1 from this promoter occurs in a wide variety of tissues, whereas expression of the alternative 5' untranslated region (exon 1a) is normally restricted to liver, the principal source of sLIF-R. During pregnancy expression of exon 1a becomes detectable also in the uterus. Expression of exon 1a increases dramatically during gestation and is accompanied by a similar quantitative rise in expression of sLIF-R mRNA. These findings establish that expression of LIF-R is under complex transcriptional control and indicate that regulated expression of the soluble cytokine receptor isoform may be due principally to an increase in the activity of a dedicated promoter.

INTRODUCTION

The biological actions of cytokines are initiated by binding to receptors present on the surface of responsive cells. These receptors can be grouped into families on the basis of shared structural features. The class I cytokine receptors possess one or in some cases two cytokine-receptor domains (CRDs) in their extracellular region [1,2]. A CRD is a composite of two fibronectin type-III-like domains possessing membrane-distal cysteine residues with conserved spacing and a membrane-proximal WSXWS motif [3]. Ligand specificity is considered to reside within the CRDs. The cytoplasmic domains of these proteins lack intrinsic enzymic activity but transduce signals via activation of associated molecules, including the Janus kinases and STAT factors [4]. Transmembrane receptor isoforms have also been described, which have altered signalling potential due to variations in their cytoplasmic domains [5–8]. Several cytokine receptors are produced in soluble as well as transmembrane forms, either through translation of alternatively spliced mRNA species or via proteolytic cleavage of the transmembrane receptor [9].

Ligands for the cytokine receptors have pleiotropic effects and exhibit overlapping biological activities. Both of these properties are exemplified by leukaemia inhibitory factor (LIF). The pleiotropic effects of LIF include stimulation of acute-phase protein

synthesis by hepatocytes, regulation of neuronal survival and differentiation, potentiation of proliferation of megakaryocyte precursors and stimulation of calcium release from bones [10]. This pleiotropy is in part due to the widespread expression of the molecules comprising the active LIF receptor complex, the low-affinity LIF receptor (LIF-R) and gp130 [11–14]. The fact that both LIF-R and gp130 are also components of the active signalling complexes for oncostatin M, ciliary neurotrophic factor and cardiotrophin-1 may explain the overlapping activities of these cytokines observed in assays *in vitro*. gp130 is also an obligate component of functional interleukin-6 and interleukin-11 receptors [1]. The physiological importance of the components of the LIF receptor complex has been demonstrated through the construction of inactive alleles of both genes by targeted mutagenesis. Homozygous *gp130* mutant mice exhibit embryonic lethality, possibly as a consequence of cardiac and haematopoietic defects [15], while homozygous *lif-r* mutants die perinatally [16]. Mortality in these latter animals may be due to severe losses of motor and other neurons [17].

A unique property of cytokines which act through the LIF-R/gp130 receptor complex is the ability to support the propagation of undifferentiated pluripotential embryonic stem (ES) cells [18–21]. The discovery of multiple LIF-R transcripts in ES cells prompted a study of the genomic organization of the LIF-R gene. We report here the structure of the complete coding

Abbreviations used: CRD, cytokine-receptor domain; dpc, day post coitum; LIF, leukaemia inhibitory factor; LIF-R, leukaemia inhibitory factor receptor; sLIF-R, soluble leukaemia inhibitory factor receptor; ES cell, embryonic stem cell; UTR, untranslated region; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; ORF, open reading frame; FN III, fibronectin type III; RACE, rapid amplification of cDNA ends; GCSF-R, granulocyte colony stimulating factor receptor.

¹ To whom correspondence should be addressed.

² Present address: The Randall Institute, King's College London, 26–29 Drury Lane, London WC2B 5RL, U.K.

The sequences of the exon 1 promoter and of 5' untranslated region 1a have been deposited with GenBank under accession numbers AF014933 and AF014934 respectively.

sequence of the mouse LIF-R gene. The identification of differentially expressed 5' non-coding exons points to the existence of alternative promoters.

MATERIALS AND METHODS

Oligonucleotides

The oligonucleotides used in this study are listed in Tables 1 and 2. All were purchased from Oswell DNA Service, University of Southampton, Southampton SO16 7PX, U.K.

Mice

All tissues were from MF1 outbred albino mice. Pregnant females were obtained by natural matings of mice maintained on a cycle of 14 h of light and 10 h of darkness. The females were caged with stud males overnight, mice with a vaginal plug the following morning were considered 0.5 days pregnant at noon and tissues were prepared from mice at various days post coitum (dpc) thereafter. Blastocysts were flushed from the uteri of pre-implantation females to check for pregnancy and conceptuses were removed from the uteri of post-implantation females.

cDNA isolation

RT-PCR

A fragment of mouse LIF-R cDNA was obtained by a reverse transcription-PCR of liver RNA using oligonucleotides B and C (Table 1), directed against residues 196–201 and 509–514 of the deduced mouse LIF-R protein sequence [22] and cloned into pGEM2 to give p980.

cDNA library screens

The insert from p980 was radioactively labelled by random hexamer primed DNA synthesis in the presence of [α - 32 P]dCTP (3000 Ci/mmol) and hybridized to an ES cell cDNA library prepared from D3 ES cell RNA using the vector λ ZAP II. Three cDNAs were isolated; p2, p3 and p4, containing sequences of nt 255–3282, 614–1661 and 369–1130 respectively. In this report, nucleotides are numbered relative to the A of the initiation codon, previously referred to as nt 72 [14]. A second ES cell cDNA library prepared in λ ZAP was screened using a reconstructed transmembrane LIF-R cDNA as a probe. This resulted in the isolation of several cDNAs, three of which were characterized in detail. pJ2 extended from –271 to 2237, pJ16 from –48 to 2347 and pJ4 from 1858 to 4045.

Table 1 Oligonucleotides for cDNA cloning and probe generation

Primer	Oligonucleotide sequence
B	GGGATCCARCAYTGGAAYTGGAC
C	GGAATTCAYTTCCARAANGTYTT
E	GGGAATTCGTATTACGAATCCAGG
G	AGGATCTCGGAGTCTCTGG
O	CAACGTGAATTTACTCTGAAAGCC
56	GGTTGTGAGCTCTTCTAGATGG (T) 17
57	CTGCCCTCTGGACTGGAGAAA
59	GGTTGTGAGCTCTTCTAGATGG
81	CCCATGTGCCGCATATTT

Table 2 Oligonucleotides used for sequencing, primer extension and intron size determination

Primer	Oligonucleotide sequence
PR144	TTCGGGCGACACGAGTTT
PR145	AAACTCGTGTGCGCCCGAA
64	GCTGACTGAACAGCAAGG
54	CTGTCTTTAATGCAAATATCT
49	GGGTACAAGACTTGAAAT
O	CAACGTGAATTTACTCTGAAAGCC
F	CAGGTTCCATTCTTGTACCC
P	TTCCAGAGACTCCCGGGATCC
107	AGGAGATGTTCTTCAGCG
B	GGGATCCARCAYTGGAAYTGGAC
55	AAAGACCACCGTTCCGTA
47	ATGTAGCTGGAATCCAGG
48	AAGTCCTCAATCCTGTG
68	TCAGGAAAATCGGCAGTAT
V	CGAAGTCGGATCATGAGGAGC
75	GTGAAGGACATCAATTCACAG
76	GTCTACAGCAACATGGTAAG
106	GAATGCCACAATCAGAGG
52	TCCAGTAGACGATTAGATTT
105	TGAGGATCGAAGATCTCC
60	CGTGTTCAATTGAACGAGG
61	GCTCTACTGTGATGTCGT
104	ACGACATCACAGTAGAGC
111	GCCAGTTCTTCTACGTATCC
43	CTGGGAATTGAACTCAGG
X	AGATACAACTTTTACCTCTATGGG
42	GAAATCTATCCAGGCCCTA
I	ACTGAGCTCTTCCCGGGATG
41	GTCGCGCCAAATTTACC
40	CTCAGTGTCTTCTGGGAT
110	ATCCCAGAAGACACTGAG
S	GACAGCCACAGCCACTGGG
69	CTGTGGGATTGATTATTGC
51	GCAGACGCTCTTCTGAAA
50	AGGAAACATTCTACCCAG
103	CGCATCTGTGGCTTATAG
Px1a	TGGGCAGGAGCCAATGCCACGGAGGGG
T3	ATTAACCCTCACTAAAG

5'Rapid amplification of cDNA ends (RACE)

Adult liver RNA (10 μ g) was used as a template for first strand synthesis primed by 10 ng of oligonucleotide E (Table 1). After homopolymer tailing in the presence of dATP [23], two rounds of PCR were performed. In the first, 75% of tailed cDNA was amplified with oligonucleotides 56 and G (Table 1) for 40 cycles at 94 °C for 30 s; 60 °C for 30 s; 72 °C for 105 s. Two percent of the cDNA was then used for PCR with oligonucleotides 59 and O (Table 1) for 35 cycles at 94 °C for 30 s; 60 °C for 30 s; 72 °C for 60 s. cDNA was then digested with *Bcl*I and *Sac*I and ligated into *Bam*HI- and *Sac*I-digested pBluescript II SK+ to give p5'.

Genomic DNA analysis

Exon positions were determined by Southern-blot analysis of restriction enzyme digests of λ clones using cDNA probes. Intron sizes and exon/intron boundary sequences were determined by PCRs using exon specific primers (Table 2) and 30 cycles at 94 °C for 30 s; 55 °C for 30 s; 72 °C for 60 s. PCR products were cloned directly into pGEM-T (Promega) and the ends were sequenced. The boundaries between exons and introns

were identified on the basis of sequence divergence between the cDNA and gene, such that an intervening sequence matching the splice-site consensus at both ends could be removed precisely to generate the cDNA sequence.

DNA sequencing

DNA sequences were obtained using a Sequenase II kit (Amersham). Ambiguities due to sequence compressions were resolved using dITP and deazaGTP. DNA sequences were determined by sequencing both DNA strands at least once.

RNA preparation

RNA was prepared as described [24], except that the second nucleic acid precipitation was achieved by adjusting the solution to 2 M LiCl and incubating at 4 °C for 16 h in order to eliminate DNA contamination. Poly(A)⁺ RNA was selected using Oligotex (Qiagen), according to the manufacturer's instructions.

Northern-blot analysis

RNA was separated on a 0.66 M formaldehyde, 0.8% (w/v) agarose gel alongside RNA molecular-mass markers (Life Technologies) and was blotted on to nylon membrane (Boehringer). Hybridizations were performed using either DNA fragments radioactively labelled by random hexamer primed DNA synthesis [25] in the presence of [α -³²P]dCTP (3000 Ci/mmol; Amersham) or antisense RNA [26] prepared in the presence of [α -³²P]CTP (800 Ci/mmol; NEN). DNA/RNA hybridizations were performed as described [27], with the final wash being in 40 mM sodium phosphate, pH 7.2/0.1% (w/v) SDS at 65 °C. RNA/RNA hybridizations were according to the method of Krumlauf et al. [28], with the final wash in 0.1 × SSC (1 × SSC = 0.15 M NaCl/0.015 M sodium citrate)/1% (w/v) SDS at 80 °C. The probe used in Figure 2 (Bi) is a *BgIII*–*BstEII* cDNA fragment. Plasmids for riboprobe synthesis were prepared as follows: 5' untranslated region (UTR) 1; pJ2 was digested with *PstI*, diluted and re-ligated. DNA was then digested with *EcoRI* and transcribed with T3 RNA polymerase. 5'UTR 1a; p5' was digested with *XbaI* and *MvaI* and the cDNA fragment was cloned into *SpeI*-digested pGEM-T. DNA was restricted with *NotI* and transcribed with T7 RNA polymerase. Exon 16, a subclone in pBluescript II SK⁺ extending to the *NdeI* site at 2393, was digested with *HindIII* and transcribed with T7 RNA polymerase.

RNase protection

Antisense-RNA probes, prepared as described above (Northern blot), were used for ribonuclease protection assays as described [29], except that the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe was synthesized from an *AccI*-digested template in the presence of 6.25 μ l [α -³²P]CTP; 350 000 c.p.m. was used per sample. Templates for LIF-R probe synthesis were produced as follows: Exon 15+16; a cDNA fragment was amplified by PCR using oligonucleotides 57 and 81 (Table 1) and was cloned into pGEM-T. DNA was then digested with *NcoI* and transcribed with SP6 RNA polymerase. Exon 1+2; pJ2 was digested with *NcoI* and *XbaI*, the ends were repaired with Klenow and re-ligated. DNA was then digested with *HindIII* and transcribed by T3 RNA polymerase. Exon 1a+2; p5' was digested with *NcoI* and *Sall*, the ends were repaired with Klenow and re-ligated. DNA was then digested with *XbaI* and transcribed with T7 RNA polymerase.

Primer extension

Oligonucleotide was labelled by incubation with T4 polynucleotide kinase in the presence of [γ -³²P]ATP (5000 Ci/mmol; Amersham) as described [30]. Primer (50 000 c.p.m.) was then hybridized to 50 μ g of total RNA, primer extension was performed [30] and the products were separated by denaturing PAGE (8% gel).

RESULTS

Structure of the LIF-R gene

A LIF-R cDNA (p980) corresponding to residues 196–514 of the murine LIF-R protein sequence [22] was isolated by a PCR using degenerate oligonucleotides. This cDNA was used to screen two ES cell cDNA libraries, which resulted in the isolation of overlapping cDNAs covering the complete open reading frame (ORF) of LIF-R (see the Materials and methods section for details). Subsequent screening of three genomic libraries prepared from mouse strain 129 DNA with cDNA probes yielded genomic clones covering the entire ORF of LIF-R (Figure 1).

The transmembrane LIF-R ORF is encoded by 19 exons within approx. 38 kb (Figure 1). The translation initiation codon is within exon 2, which also encodes the secretory signal sequence. The 3'UTR of a soluble LIF-R mRNA [12,14] was localized to a single exon (exon 16) situated between the exons encoding the second and third fibronectin type III (FN III) repeats by a series of PCRs using genomic DNA. The intron/exon boundaries of the LIF-R gene were sequenced and the lengths of the introns were determined by PCR. All of the introns had a typical splice donor and acceptor sequence, conforming to the consensus sequences (Table 3; [31]). The coding potential of the exons and the intron phasing [32] of the LIF-R gene conform to the patterns observed previously for cytokine receptor genes [33]. This supports the notion that the LIF-R gene evolved from the same ancestral gene that gave rise to other members of the cytokine receptor family.

Characterization of LIF-R transcripts

Northern-blot analysis of ES cell RNA with the p980 cDNA probe detected hybridization to transcripts of approx. 11, 4.4 and 3 kb (Figure 2A). To elucidate the coding potential of these transcripts, defined fragments of LIF-R cDNA were used to probe a panel of Northern blots. RNAs from liver, placenta and a parietal endoderm-like cell line were used, since LIF-R mRNA expression is elevated in these sources relative to ES cells. Hybridization with a probe derived from sequences encoding the FN III domains detected transcripts of the same sizes as seen in ES cell RNA previously, but the relative amounts varied (Figure 2Bi). In liver, an additional transcript of 2.8 kb was observed. Interestingly, the steady state level of the 2.8 kb mRNA was significantly higher than that of the other LIF-R RNAs. These observations indicate that expression of LIF-R is under complex regulatory control.

A similar pattern of hybridization was observed using probes containing the full ORF of the transmembrane LIF-R or sequences derived from either CRD1 or CRD2 (results not shown). This indicates that the RNAs detected each contain sequences derived from both CRDs and from the FN III repeats.

Significantly, a probe derived from the cytoplasmic domain of LIF-R hybridized to the 11 kb and 4.4 kb RNAs, but did not hybridize to the 3 or 2.8 kb RNAs (results not shown). This suggests that these latter RNAs encode soluble forms of LIF-R. This was tested by hybridization with a probe derived from exon

Figure 2 Northern-blot analysis of LIF-R RNAs

(A) Poly(A)⁺ RNA (5 µg) from the ES cell line EFC-1 [50] was analysed using a probe (cDNA p980) encoding extracellular sequences extending from the 3' end of CRD1 to the 3' end of CRD2. The positions of molecular-mass markers are indicated on the left. (B) Poly(A)⁺ RNA (500 ng per lane) prepared from: 1, adult male liver; 2, placenta at 12.5 dpc; 3, parietal endoderm-like cell line was subjected to electrophoresis in multiple panels of a single gel and was analysed by hybridization. The probes used were: i, extracellular domain covering exons 13, 14, 15, 17 and 18; ii, soluble LIF-R-specific 3'UTR (exon 16); iii, 5'UTR 1; and iv, 5'UTR 1a. Further details of the probes is given in the Materials and methods section. The positions of molecular-mass markers are indicated on the left of each panel.

but not to RNA of 2.8 kb (Figure 2Biii). Conversely, the exon 1a probe hybridized to the 2.8 but not to the 3 kb RNA (Figure 2Biv).

Figure 3 Nucleotide sequences of the alternative 5'UTRs of LIF-R and the sequence upstream of exon 1

(A) Sequences of 5'UTR 1, obtained from an ES cell cDNA library screen and 5'UTR 1a, obtained by 5'RACE using liver RNA. The exon 2 sequence is shown in lower case. (B) Sequence of the exon 1 promoter. Nucleotides are numbered relative to the A of the translation initiation codon (+1) within exon 2 (lower case). The sites of transcription initiation are marked by asterisks. The 5' end of the longer ES cell cDNA lies between these two positions at -271. The initiator sequence (Inr) and potential Sp1 sites are underlined and a polypurine stretch is overlined.

Further experiments were conducted to define the limits of the alternative 5'UTRs more precisely. For analysis of exon 1, an RNase protection assay was performed using a probe derived from genomic DNA which extended upstream from the *Pst*I site at -98 nt. In agreement with the Northern blot (Figure 2) and RNase protection (Figure 4) data, RNAs present in both liver and placenta protected the probe from digestion (Figure 5). The major protected fragment had a size of approx. 164 nt, with a minor fragment of 174 nt also being detected. This indicates that the 5' ends of the corresponding RNAs lie at positions -266 and -276 relative to the A of the initiator codon. The sequence in this region does not conform to that expected of a 3' splice site. A TATA box is not present, however, between -287 and -279, the sequence CGCAGTCT shows a 6/7 match to the consensus initiator sequence, CTCANTCT [35]. In addition, two Sp1 sites are present between -340 and -359 with their central cytosine residues separated by a single helical turn. Therefore it is likely that these positions mark the transcription initiation sites of exon 1.

The λ clone containing the exon 1 promoter (λ 5') does not overlap with λ 7, which contains exon 2. Restriction analysis of

Figure 4 RNase protection indicates preferential usage of 5'UTRs by distinct tissues

Riboprobes extending from the same position within exon 2 upstream into 5'UTR 1 (Probe 1) or 5'UTR 1a (Probe 1a) were synthesized and used to protect placental or liver RNAs from RNase digestion. The identity of the probes used and the source of RNA is indicated above each lane. The structure of the probes and of the various protected fragments is illustrated below the autoradiograph. The positions of migration of fragments of the probe protected from RNase digestion by hybridization to RNA containing sequences from exon 1 and exon 2 (exon 1/exon 2), exon 1a and 2 (exon 1a/exon 2) and exon 2 alone (exon 2) are indicated on the right. Lanes marked M contain sequencing reactions produced by single ddNTPs using characterized templates. Lanes marked tRNA are from assays performed using yeast tRNA as a control for complete digestion of the probe. The position of the input probes is indicated by the arrowheads.

the two clones indicates that the exon 1 promoter lies at least 14 kb upstream of exon 2. As neither $\lambda 7$ nor $\lambda 5'$ hybridized with an oligonucleotide specific for exon 1a, we were unable to position exon 1a relative to exons 1 and 2. To determine the length of exon 1a, a primer extension assay was performed using an exon-1a-specific oligonucleotide. In accordance with the Northern-blot (Figure 2) and RNase-protection (Figure 4) analyses, a primer extension product was obtained using liver but not placental RNA (Figure 5). The size of this product indicates that the 5' end of the exon 1a RNA lies 73 or 74 nt upstream of the A of the initiation codon. This agrees, to within 1 nt, with the length of exon 1a as determined by 5'RACE. Assuming a

Figure 5 Mapping the 5' ends of exons 1 and 1a

Upper panel: an RNase protection assay was performed to determine the position of the 5' end of exon 1 using an antisense riboprobe corresponding to nt -98 to -455 (see Figure 3). placental and liver RNAs were from tissue obtained at 14.5 and 10.5 dpc respectively. The major bands were estimated to be 164 and 174 nt in length (shown on the left of the panel), by comparison with the mobility of products of the sequencing reaction (M). Lower panel: primer-extension analysis to determine the extent of exon 1a using the primer Px1a (see Table 2) complementary to nt -24 to -50 of exon 1a. Sequencing of pBluescript KS with a T3 primer allowed determination of product length; the band present in the fourth lane of the sequencing ladder is 51 nt (shown on the left of the panel). RNA samples were prepared from tissues removed at the indicated dpc, as indicated above each lane.

poly(A) tail length of 250 nt [36], the predicted sizes of sLIF-R mRNAs produced by utilization of exon 1 and exon 1a are approx. 2950 and 2750 nt, respectively. These sizes are in good

agreement with the sizes of sLIF-R mRNAs estimated from Northern-blot analysis. This indicates that the difference in the lengths of the alternative 5'UTRs is responsible for the size difference observed between sLIF-R mRNAs of 2.8 and 3 kb.

Differentially-regulated expression of alternative 5'UTRs of LIF-R mRNA

The RNase-protection (Figure 4) and primer-extension (Figure 5) analyses indicated that exon 1 was utilized in both placenta and liver, but that exon 1a utilization occurred only in liver and that the level of expression increased between 10.5 and 15.5 days of pregnancy. To investigate this more thoroughly, RNase protection assays were performed on samples obtained at various stages of pregnancy using the riboprobes described in Figure 4. In analysis of liver and placental samples the exon 1/2 probe was used and uterine samples were analysed with the exon 1a/2 probe (Figure 6). Note that in these experiments protection of a probe fragment corresponding to exon 2 by the exon 1/2 probe reflects expression of exon 1a. Likewise, protection of exon 2 by the exon 1a/2 probe is a reflection of exon 1 expression. A clear example of this is provided by the 3.5 dpc uterus sample. Analysis with the exon 1/2 probe demonstrates that LIF-R expression is solely due to exon 1 utilization, as no exon 2 signal is detectable. Analysis with the alternative exon 1a/2 probe demonstrates that the same RNA sample protects only the exon 2 fragment of the probe. In the female liver, expression of LIF-R mRNA from exon 1 is relatively constant during pregnancy, whereas utilization of exon 1a begins to increase between 6.5 and 8.5 days of pregnancy, reaching a maximum at around day 15. Quantification by phosphorimage analysis indicated that expression of exon 1 increased by at most 3.5-fold at 15.5 days of pregnancy. In contrast, expression of exon 1a is elevated 15-fold at this stage. The quantitative elevation in the expression of exon 1a was confirmed by an RNase protection assay using the exon 1a/2 probe (results not shown). Expression of exon 1 and 1a in the uterus follows a similar pattern to that seen in the liver, with expression of exon 1a being 10–15-fold higher at 14.5 days of pregnancy than in non-pregnant animals. In contrast, placental expression of LIF-R mRNA is exclusively due to exon 1 utilization.

Expression of mRNAs encoding soluble and transmembrane LIF-R isoforms

The quantitative changes in expression of exon 1a during pregnancy are significant in the context of the changing expression of sLIF-R reported to occur during pregnancy [37]. Therefore, the relative levels of expression of mRNAs encoding soluble and transmembrane isoforms of LIF-R were examined by RNase protection assays using a probe which distinguishes between these two classes of LIF-R mRNA. Hepatic expression of sLIF-R mRNA increased significantly during pregnancy (Figure 7). The increase was evident by day 10.5 and maximal

Figure 6 Time course of expression of alternative 5' UTRs in liver, placenta and uterus during gestation

The LIF-R probes are described in the legend to Figure 4, and were included in hybridizations alongside a GAPDH (GAP) probe. Top panel: RNase protection by the exon 1+2 probe. Protected fragments corresponding to exon 1+2 and exon 2 are indicated by arrowheads, the latter gives a measure of the expression of exon 1a (see text for details). The uterus sample is from tissue dissected at 3.5 dpc. Liver and placental tissue was removed at the dpc indicated above the panel. Bottom panel: RNase protection by the exon 1a+2 probe. Protected fragments corresponding to exon 1a+2 and exon 2 are indicated by arrowheads. Abbreviation: non-oes, non-oestrus.

Figure 7 Time course of expression of RNAs encoding soluble and transmembrane forms of LIF-R during gestation

RNase protection of RNA samples from the tissues and at the dpc indicated above the panel was performed using a sLIF-R (LIF-R) probe in combination with a probe for GAPDH (GAP). The sLIF-R probe, described in detail in the Materials and methods section, contained sequences complementary to those in exons 15 and 16. Therefore hybridization of the probe to sLIF-R mRNA protects a fragment of the probe corresponding to exon 15 + 16 from RNase digestion, whereas hybridization of the probe to transmembrane LIF-R mRNA protects only a fragment of the probe corresponding to exon 15. The positions of migration of these products are indicated by exon 15 + 16 and exon 15 respectively. Lanes marked C and A contain sequencing reactions produced by ddCTP and ddATP respectively, using a characterized template.

levels were attained by day 15, which represented a 15-fold increase over the expression in liver of non-pregnant animals (Figure 7). In contrast, transmembrane LIF-R mRNA shows only a modest increase in expression. Significantly, sLIF-R mRNA expression in the uterus also increased during gestation, reaching a peak at a similar time as in liver. Despite this qualitative similarity, it is apparent from a comparison of the strengths of the GAPDH signals (Figure 7) that the level of expression of sLIF-R in the uterus is very much lower than in the liver. Interestingly, the lowest levels of sLIF-R mRNA observed in the uterus occurred at day 3.5, just before the time of implantation.

DISCUSSION

The structure of the LIF-R gene, both in terms of the coding potential of individual exons and the phasing of introns is analogous to other members of the cytokine receptor gene

family. However, two features of LIF-R are shared only with a subset of other family members and their combination within a single gene is, thus far, unique to LIF-R. These are the duplication of the CRD and the interposition of a tandem array of three FN III domains between the CRDs and the transmembrane domain. The CRD duplication appears to have occurred within the LIF-R gene itself, based upon comparison of the sequences of the CRDs of several cytokine receptors [38]. A tandem array of similarly located FN III domains is present in gp130 and in the granulocyte colony stimulating factor receptor (GCSF-R) [22]. Of these, the intron phasing of the FN III repeats of GCSF-R has been reported [39] and is identical to that seen in the present study for LIF-R. This suggests that the FN III domains were inserted into an ancestral LIF-R-GCSF-R gene with subsequent divergence rather than that these domains were individually inserted into the two genes separately.

The chromosomal location of the GCSF-R and LIF-R genes is also pertinent to the evolution of cytokine receptors. The LIF-

R gene has been mapped to the proximal region of chromosome 15, both by interspecific backcross analysis [40], and fluorescent *in situ* hybridization analysis (results not shown). The receptors for growth hormone, IL-7 and prolactin also map to this region [40], suggesting that this cytokine receptor cluster evolved by tandem duplication of an ancestral gene. However, the gene for GCSF-R, which is more closely related to LIF-R than any of the above genes, does not map to this region [41]. Therefore the LIF-R and GCSF-R genes appear to have evolved from an ancestral gene by a mechanism involving dispersal to remote genomic sites. This suggests that other cytokine receptor genes may exist in the vicinity of the GCSF-R gene.

The intron phasing of the LIF-R gene also suggests a potential for exon skipping, whereby the first exon of the CRD1 could be spliced to the second exon of CRD2, without altering the reading frame. This would produce an mRNA encoding a novel receptor isoform containing a single CRD which may exhibit altered ligand binding specificity. Further receptor isoforms could be generated by analogous splicing of the second or third exons of CRD1 to the third or fourth exons of CRD2 respectively. Exon skipping in the region of the FN III domains C-terminal to the CRDs could also result in the synthesis of distinct proteins. Minor transcripts seen on Northern blots ([12]; I. Chambers and A. Smith, unpublished work) may possess such alternative structures. However, further analysis will be required to determine if such exon skipping actually occurs.

From an analysis of the structure of LIF-R mRNAs by Northern blotting, four major classes of transcript were detected having sizes of approx. 11 kb, 4.4 kb, 3 kb and 2.8 kb. Both the 11 kb and the 4.4 kb mRNAs can apparently give rise to the transmembrane receptor, since they both hybridize to a cytoplasmic domain probe (results not shown). In fact it appears that these mRNAs differ only in the extent of their 3'UTRs, since a probe located 2 kb downstream of the transmembrane LIF-R stop codon hybridized to the 11 kb but not the 4.4 kb RNA (results not shown). mRNA encoding the transmembrane receptor is produced by splicing exon 15 to exon 17, whereas sLIF-R mRNA is formed by splicing exon 15 to exon 16. The 2.8 and 3 kb mRNAs encode sLIF-R since they both hybridize to an exon-16-specific probe (Figure 2). This latter exon contains a translation terminator three codons 3' to the splice site. In addition, exon 16 contains a B2 repetitive element aligned in the same transcriptional orientation as the LIF-R gene. B2 elements possess polyadenylation signals and these are active in the formation of the sLIF-R mRNA 3' end (results not shown). The fact that B2 repeats are rodent specific may account for the lack of detection of sLIF-R transcripts in humans [22]. However, it should be noted that soluble cytokine receptors can also be generated by proteolytic cleavage of transmembrane isoforms [9]. Whether such a mechanism is involved in the generation of a soluble form of LIF-R remains to be seen.

A major finding of our studies is the presence of alternative 5'UTRs in LIF-R transcripts. Expression analysis of LIF-R indicates that the promoter giving rise to exon 1 is utilized in many tissues, whereas the promoter driving exon 1a is restricted in its activity and is silent in all tissues examined of the non-pregnant animals, with the exception of liver (Figure 6 and results not shown). During pregnancy, expression of exon 1a increases in the liver and also becomes detectable in the uterus. Quantification by phosphorimage analysis indicates that the peak levels of expression in these two tissues are 10–15-fold higher in pregnant than in non-pregnant animals. Despite these increases, expression of exon 1 is relatively constant. The similarity in the time course of induction in these two tissues suggests that the exon 1a promoter may be responding to a

systemically acting effector(s). The nature of this putative effector is unknown but the reproductive hormones are likely candidates.

The increases in LIF-R mRNA during pregnancy could be due to a change in the expression of LIF-R mRNA within individual cells or may reflect the expansion of a discrete population of cells expressing a constant level of LIF-R mRNA. Discrimination between these possibilities will require determination of the cellular sites of LIF-R mRNA expression. This could be achieved by an extension of the *in situ* hybridization analysis which has shown that expression of LIF-R mRNA within the uterus at 4.5 days of pregnancy is restricted to the luminal epithelium [42]. Alternatively, histochemical staining of tissues from mice carrying a *lacZ* gene integrated at the *lifr* locus may be used to the same end [17].

The analysis of 5'UTR utilization in the liver and uterus during gestation highlights the correlation between exon 1a utilization and the expression of sLIF-R mRNA. sLIF-R synthesis requires that exon 15 be spliced to exon 16, whereas synthesis of transmembrane LIF-R requires splicing of exon 15 to exon 17. Splicing between exons 15 and 17 cannot occur if the transcript has undergone polyadenylation in response to signals at the 3' end of exon 16. Whether polyadenylation or splicing is the key determinant in the process of selection of exon 16 for inclusion in the mRNA remains to be determined. However, the inclusion of exon 1a in the transcript appears to contribute to the selection of exon 16. This could be due to a direct role of the exon 1a sequences or may be a consequence of a qualitative difference between the transcription complexes assembled at the two promoters.

A significant regulatory consequence of the correlated expression of exon 1a and exon 16 is that by initiating transcription at promoter 1a rather than promoter 1, a cell can produce sLIF-R without incurring a concomitant rise in the cell surface density of the transmembrane isoform.

The efficiency of translation of LIF-R may also be affected by the presence of the alternative 5'UTRs. Exon 1 has a higher GC content (74%) and a greater length (248 nt) than exon 1a (65% GC; 51 nt). Both of these factors may indicate an increased potential for secondary structure within exon 1 relative to exon 1a. Exon 1 also harbours an upstream AUG. As both upstream AUG codons and increased secondary structure have been shown to decrease translation of a downstream ORF [43,44], it is possible that exon 1a may allow more efficient translation initiation than exon 1. Utilization of exon 1a may, therefore, lead to an increase in the synthesis of sLIF-R, which is distinct either quantitatively or kinetically from that which could be achieved by simply modulating the expression of exon 1.

The role of sLIF-R *in vivo* is presently unclear. Assays *in vitro* demonstrate that sLIF-R is able to block the activity of murine LIF [37,45]. Although the levels of sLIF-R required are extremely high (in the order of $\mu\text{g/ml}$), these levels have been reported to be present in the serum of non-pregnant females and to increase during pregnancy [37]. However, the significance of these observations has not been determined. LIF is required for blastocyst implantation [46], with a peak in LIF mRNA expression in the endometrial glands of the uterus occurring at 3.5 days of pregnancy, before implantation [47,48]. The presence of transmembrane LIF-R transcripts in the uterus before and throughout gestation, suggests that the uterus is capable of responding to signals mediated by LIF family members at all stages. Interestingly, no exon 1a expression is detectable in the uterus at the time of implantation (Figure 6) and, as a consequence, sLIF-R mRNA expression is minimal (Figure 7). This is consistent with the notion that lowered sLIF-R levels may be required to facilitate the efficient action of LIF on nearby target cells. sLIF-

R mRNA levels do not begin to increase in either the uterus or liver until after 4.5 days of pregnancy and reach maximal levels between days 12 and 15, when LIF mRNA expression has declined to background levels [47]. These time courses may suggest that, rather than blocking the systemic action of LIF, sLIF-R fulfils another role by possibly acting as a transport protein for LIF and affecting its half-life in the circulation, as has been suggested for a soluble form of IL-4 receptor [49]. Alternatively, sLIF-R can bind cardiotrophin-1 [21] and so may regulate the availability of cardiotrophin-1 or other members of the LIF family of cytokines. Clarification of the role *in vivo* of sLIF-R will require deletion of portions of the LIF-R gene by homologous recombination. This could be addressed either by deletion of exon 16 or by deletion of exon 1a. The latter approach would specifically prevent increased synthesis of sLIF-R during pregnancy.

In conclusion, the data presented here demonstrate that expression of transmembrane and soluble isoforms of LIF-R can be uncoupled by transcription from alternative promoters. This allows the synthesis of receptor isoforms with separate physiological functions to be independently regulated. We anticipate that further examples of cytokine receptor isoforms, synthesized by the use of dedicated promoters, will emerge as more of the corresponding genes are subjected to close scrutiny.

We thank Peter Mountford for the EFC-1 RNA, Hitoshi Niwa and Peter Rathjen for cDNA libraries, Andrew Reaume and Andrew Smith for genomic libraries, Jenny Nichols for help with dissections, Xiao-Qing Wei for performing one of the primary genomic library screens, Frank Johnston, Graham Brown and colleagues for valuable photographic assistance and John Bishop for critical reading of the manuscript. This work was supported by the Biotechnology and Biological Sciences Research Council of the U.K.

REFERENCES

- Kishimoto, T., Taga, T. and Akira, S. (1994) *Cell* **76**, 253–262
- Yamamori, T. and Sarai, A. (1994) *J. Physiol. (Paris)* **88**, 165–171
- Bazan, J. F. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6934–6938
- Stahl, N., Farrugella, T. J., Boulton, T. G., Zhong, Z., Darnell, J. E. and Yancopoulos, G. D. (1995) *Science* **267**, 1349–1353
- Nakamura, Y., Komatsu, N. and Nakauchi, H. (1992) *Science* **257**, 1138–1411
- Nakamura, Y. and Nakauchi, H. (1994) *Science* **264**, 588–589
- Nakamura, Y., Tokumoto, Y. and Nakauchi, H. (1996) *Biochem. Biophys. Res. Commun.* **218**, 205–209
- Fukunaga, R., Seto, Y., Mizushima, S. and Nagata, S. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8702–8706
- Rose-John, S. and Heinrich, P. C. (1994) *Biochem. J.* **300**, 281–290
- Hilton, D. J. (1992) *Trends Biochem. Sci.* **17**, 72–76
- Saito, M., Yoshida, K., Hibi, M., Taga, T. and Kishimoto, T. (1992) *J. Immunol.* **148**, 4066–4071
- Owczarek, C. M., Layton, M. J., Robb, L. G., Nicola, N. A. and Begley, C. G. (1996) *J. Biol. Chem.* **271**, 5495–5504
- Tomida, M., Yamamoto-Yamaguchi, Y. and Hozumi, M. (1993) *FEBS Lett.* **334**, 193–197
- Tomida, M., Yamamoto-Yamaguchi, Y. and Hozumi, M. (1994) *J. Biochem. (Tokyo)* **155**, 557–562
- Yoshida, K., Taga, T., Saito, M., Suematsu, S., Kumanogoh, A., Tanaka, T., Fujiwara, H., Hirata, M., Yamagami, T., Nakahata, T. et al. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 407–411
- Ware, C. B., Horowitz, M. C., Renshaw, B. R., Hunt, J. S., Liggit, D., Koblar, S. A., Gliniak, B. C., McKenna, H. J., Papayannopoulou, T., Thoma, B. et al. (1995) *Development* **121**, 1283–1299
- Li, M., Sendtner, M. and Smith, A. (1995) *Nature (London)* **378**, 724–727
- Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D. (1988) *Nature (London)* **336**, 688–690
- Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A. and Gough, N. M. (1988) *Nature (London)* **336**, 684–687
- Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T. and Kishimoto, T. (1994) *Mech. Dev.* **45**, 163–171
- Pennica, D., Shaw, K. J., Swanson, T. A., Moore, M. W., Shelton, D. L., Zionscheck, K. A., Rosenthal, A., Taga, T., Paoni, N. F. and Wood, W. I. (1995) *J. Biol. Chem.* **270**, 10915–10922
- Gearing, D. P., Thut, C. J., VandenBos, T., Gimpel, S. D., Delaney, P. B., King, J., Price, V., Cosman, D. and Beckman, M. P. (1991) *EMBO J.* **10**, 2839–2848
- Frohman, M. A., Dush, M. K. and Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8998–9002
- Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159
- Feinberg, A. P. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13
- Kreig, P. A. and Melton, D. A. (1987) *Methods Enzymol.* **155**, 397–415
- Church, G. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991–1995
- Krumlauf, R., Holland, P. W. H., McVey, J. H. and Hogan, B. M. L. (1987) *Development* **99**, 603–617
- Robertson, M., Chambers, I., Rathjen, P., Nichols, J. and Smith, A. (1993) *Dev. Genet.* **14**, 165–73
- Kingston, R. E. (1990) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K., eds.), pp. 4.8.1–4.8.3, Wiley, New York
- Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. and Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55**, 1119–1150
- Sharp, P. A. (1981) *Cell* **23**, 643–646
- Nakagawa, Y., Kosugi, H., Miyajima, A., Arai, K.-i. and Yokota, T. (1994) *J. Biol. Chem.* **269**, 10905–10912
- Ryskov, A. P., Ivanov, P. L., Kramerov, D. A. and Georgiev, G. P. (1983) *Nucleic Acids Res.* **11**, 6541–6558
- Azizkhan, J. C., Jensen, D. E., Pierce, A. J. and Wade, M. (1993) *Crit. Rev. Eukaryotic Gene Expression* **3**, 229–254
- Brawerman, G. (1981) *Crit. Rev. Biochem.* **10**, 1–38
- Layton, M. J., Cross, B. A., Metcalf, D., Ward, L. D., Simpson, R. J. and Nicola, N. A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8616–8620
- Shields, D. C., Harmon, D. L., Nunez, F. and Whitehead, A. S. (1995) *Cytokine* **7**, 679–688
- Seto, Y., Fukunaga, R. and Nagata, S. (1992) *J. Immunol.* **148**, 259–266
- Gearing, D. P., Druck, T., Huebner, K., Overhauser, J., Gilbert, D. J., Copeland, N. G. and Jenkins, N. A. (1993) *Genomics* **18**, 148–150
- Inazawa, J., Fukunaga, R., Seto, Y., Nakagawa, H., Misawa, S., Abe, T. and Nagata, S. (1991) *Genomics* **10**, 1075–1078
- Nichols, J., Davidson, D., Yoshida, K., Taga, T., Chambers, I. and Smith, A. (1996) *Mech. Dev.* **57**, 123–131
- Geballe, A. P. and Morris, D. R. (1994) *Trends Biochem. Sci.* **19**, 159–164
- Pelletier, J. and Sonenberg, N. (1985) *Cell* **40**, 515–526
- Tomida, M. (1995) *J. Biochem. (Tokyo)* **117**, 1228–1231
- Stewart, C. L., Kaspar, P., Brunet, L. J., Bhatt, H., Gadi, I., Kontgen, F. and Abbondanzo, S. J. (1992) *Nature (London)* **359**, 76–79
- Bhatt, H., Brunet, L. J. and Stewart, C. L. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11408–11412
- Smith, A. G., Nichols, J., Robertson, M. and Rathjen, P. D. (1992) *Dev. Biol.* **151**, 339–351
- Fernandez-Botran, F. and Vitetta, E. S. (1991) *J. Exp. Med.* **174**, 673–681
- Nichols, J., Evans, E. P. and Smith, A. G. (1990) *Development* **110**, 1341–1348