Genomic organization, chromosomal mapping and promoter analysis of the mouse selenocysteine tRNA gene transcription-activating factor (*mStaf*) gene

Kazushige ADACHI, Masato KATSUYAMA, Shigeaki SONG and Takami OKA[1]

[1]Laboratory of Genetics and Physiology, Room 106, Building 8, National Institutes of Diabetes, Digestive and Kidney Disease, National Institutes of Health, Bethesda, MD 20892, U.S.A.

mStaf is a zinc-finger protein that activates the transcription of the mouse selenocysteine tRNA gene. The *mStaf* gene is approx. 35 kb long and split into 16 exons. All exon–intron junction sequences conform to the GT/AG rule. The transcription start site is located 83 bp upstream of the initiation codon. Chromosomal mapping localized the gene to mouse chromosome 7, region E3–F1. Sequence analysis of the proximal promoter region revealed several potential regulatory elements; these include the recognition elements of Sp1, Nkx, CP2, E2A, SIF (SIS-inducible factor), TFII-I and cAMP-responsive element (CRE), but no TATA sequences. Transfection experiments demonstrated that the 5′-flanking region (−1894 to +37) of the *mStaf* gene drives transcription in mouse NMuMG cells and that a construct containing a fragment from −387 to +37 showed the highest transcriptional activity. Deletion and mutation experiments suggested that four Sp1 sites played an important role for the basal promoter activity. Furthermore, electrophoretic mobility-shift assays demonstrated that Sp3 but not other Sp (specificity protein) family members binds to three of the Sp1 sites. Our present study suggests that Sp3 is involved in the basal transcriptional activation of the *mStaf* gene.

Key words: gene structure, promoter activity, Sp1, transcription start site.

INTRODUCTION

Selenocysteine tRNA (tRNASec) serves as a donor of selenocysteine in response to the specific UGA codon that normally functions as a stop codon [1–3]. The mouse *tRNASec* gene is transcribed by RNA polymerase III [4] and its transcription is controlled by 5′ regulatory elements (a TATA box, a proximal sequence element and an activator element) and an internal B-box element [5]. Recently, we have demonstrated that a transactivator, mStaf, binds to the activator element and enhances the transcription of the mouse *tRNASec* gene [6].

mStaf is a zinc-finger protein containing seven tandemly repeated C2-H2-type zinc-fingers and four repeated motifs at the N-terminus, which are important for DNA binding and transactivation, respectively [6]. In mammary glands, the DNA-binding activity of mStaf undergoes developmental changes, reaching maximal level during lactation *in vivo* and is stimulated by lactogenic hormones through the mitogen-activated protein kinase pathway *in vitro* [6,7]. The activities of two selenocysteine-containing enzymes (selenoproteins), type-I thyroxine 5′-deiodinase and glutathione peroxidase, increase in mammary glands during lactation [8,9] when the amounts of mStaf-binding activity and tRNASec increase. Thus it is possible that mStaf may be involved in stimulation of the biosynthesis of selenoproteins.

Two zinc-finger proteins, ZNF143 [10] and ZNF76 [11], have been identified as human homologues of mStaf [6,12], and shown to possess the capacity to transactivate not only the tRNASec promoter but other RNA polymerase III- and RNA polymerase II-type promoters by binding their activator elements [12]. These results raise the possibility that mStaf also may play a role in activating transcription of other types of genes in various tissues. In order to gain a better understanding of the regulatory mechanism of *mStaf* transcription, we report here the isolation and characterization of the *mStaf* gene.

MATERIALS AND METHODS

Library screening and sequencing

Approx. 1×10^6 plaques of the λFIX II vector containing female 129 SVJ mouse liver genomic phage library (Stratagene) were screened by hybridization with a 942-bp fragment of mStaf cDNA spanning the zinc-finger domain under the conditions described previously [6]. Phage DNA from three partially overlapping positive clones was isolated, digested with *Not*I restriction enzyme and subcloned into pBluescript II SK(+) (Stratagene) for further characterization. However, none of these clones contained the first exon or the 5′-flanking region. Therefore, approx. 2 kb of 5′-flanking region was cloned and sequenced using the GenomeWalker™ kit (Clontech) according to supplier's instructions using an antisense primer, 5′-CTCCGAGAAAA-TTTTCCAACAATTCC-3′, from the first exon. Then, mouse

Abbreviations used: CRE, cAMP-responsive element; DAPI, 4′,6′-diamidino-2-phenylindole; FISH, fluorescence *in situ* hybridization; 3′-RACE, rapid amplification of 3′-cDNA ends; tRNASec, selenocysteine tRNA; EMSA, electrophoretic mobility-shift assay.

[1] To whom correspondence should be addressed (e-mail okat@bdg8.niddk.nih.gov).

The nucleotide sequences reported have been submitted to GenBank Nucleotide Sequence Database under accession numbers AF145372 and AF145373.

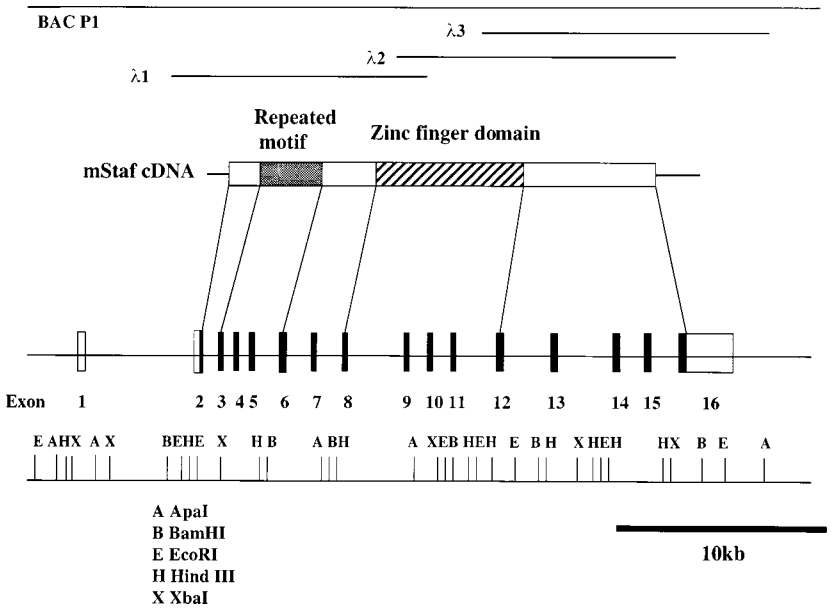

Figure 1 Structural organization and restriction map of the *mStaf* gene

Diagrammatic representation of the *mStaf* gene showing the resultant cDNA structure and the genomic clones in which the gene was found. Exon numbers are placed as they are arranged in the gene. For the genomic structure, solid boxes represent coding regions, and open boxes represent the 5′- and 3′-untranslated regions. For the cDNA structure, the shaded and hatched boxes represent the repeated motifs and the zinc-finger domain, respectively. The regions of genomic clones, λ1, λ2, λ3 and BAC P1 are shown at the top. The sites for the restriction enzymes, *Apa*I (A), *Bam*HI (B), *Eco*RI (E), *Hind*III (H) and *Xba*I (X), are shown at the bottom.

genomic BAC filters (Genome Systems) were screened using a 624-bp fragment of the 5′-flanking region. The filters were hybridized overnight in QuikHyb hybridization solution (Stratagene) and washed as described [6]. We obtained one positive clone and confirmed by sequencing and PCR analysis that it contained the first exon and 5′-flanking region.

The isolated clones were mapped by restriction-enzyme analysis. Boundaries between exons and introns were determined by DNA sequence and PCR analysis. Sequencing was carried out on both strands by the dideoxy chain-termination method.

Primer-extension analysis

The transcription start site of the *mStaf* gene was determined by primer extension with an end-labelled oligonucleotide primer complementary to +86 to +117 bp of the mStaf cDNA [6]. Total RNA was isolated from lactating-mouse mammary glands, and further purified using a mini-oligo-cellulose spin column kit (5Prime-3Prime) to prepare poly(A)$^+$ RNA. The primer was annealed to 3 μg of poly(A)$^+$ RNA in hybridization buffer (40 mM Pipes, pH 6.4, 1 mM EDTA, 0.4 M NaCl, 80% formamide) at 50 °C. The resulting DNA/RNA hybrids were precipitated by ethanol and dissolved in a reverse-transcription buffer (75 mM Tris/HCl, pH 8.3, 37.5 mM KCl, 4.5 mM $MgCl_2$, 2 μl of 0.1 M dithiothreitol, 1 μl of 10 mM dNTPs and 0.5 μl of Superscript II reverse transcriptase; Life Technologies). The reaction mixture was incubated at 50 °C for 60 min. The reaction products were electrophoresed on a 6% denaturing polyacrylamide gel. The start site was determined by comparison with sequencing reactions carried out on the genomic clone.

Rapid amplification of 3′ cDNA ends (3′-RACE)

3′-RACE was performed using the 3′-RACE System (Life Technologies). Briefly, the first-strand cDNA was synthesized from total mouse mammary gland RNA employing a 3′-RACE general primer supplied by the manufacturer. PCR was then performed using the cDNA as template and *mStaf* gene-specific primer (5′-TAAAACAGGAGGAGCCTTTCATC-3′) and a universal amplification primer supplied by the manufacturer. An amplified band was subcloned and sequenced to determine the polyadenylation site.

Chromosome mapping

Fluorescence *in situ* hybridization (FISH) was performed as described [13,14]. Lymphocytes were isolated from mouse spleen and cultured at 37 °C in RPMI 1640 medium supplemented with 15% fetal bovine serum, 3 μg/ml concanavalin A, 10 μg/ml lipopolysaccharide and 5×10^{-5} M β-mercaptoethanol. After 44 h, the cultured lymphocytes were treated with 0.18 mg/ml bromodeoxyuridine (Sigma) for an additional 14 h. The synchronized cells were washed and recultured at 37 °C for 4 h in α-minimal essential medium with thymidine (2.5 μg/ml; Sigma). Mouse chromosomal preparation was made by conventional method, as used for human chromosome preparation (hypotonic treatment, fixation and air-drying). DNA probe was biotinylated with dATP using the Gibco-BRL BioNick labelling kit as described by the manufacturer. The procedure for FISH detection was performed as described in [13]. Briefly, slides were baked at 55 °C for 1 h. After RNase A treatment, the slides were denatured in 70% formamide in 2 × SSC (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0) for 2 min at 70 °C followed by dehydration with ethanol. Probes were denatured at 75 °C for 5 min in a hybridization mixture consisting of 50% formamide and 10% dextran sulphate and mouse cot I DNA and prehybridized for 15 min at 37 °C. After overnight hybridization, slides were washed and signals were detected and amplified using the method of Heng et al. [13]. FISH signals and the DAPI (4′,6′-diamidino-2-phenylindole) banding pattern were recorded sep-

Table 1 Exon–intron organization of the *mStaf* gene

Part of the nucleotide sequence is shown with the corresponding encoded amino acid. Conserved nucleotides for donor and acceptor in introns are shown in bold.

Exon		Exon–intron junction		
Number	Size (kb)	5′ donor	3′ acceptor	Intron size (kb)
1	40	CTCGGAG**gt**gttcgta	ttcttca**ag**GTAGAAAT	8.5
2	119	GTGGCAG**gt**aagccat AlaA . . .	ttatcat**ag**ATGGTGAC . . . spGly	0.9
3	93	TCTAAAG**gt**gagtgca LysA . . .	ttgttcc**ag**ACGGGAGA . . . spGly	0.8
4	84	AAAAGTA**gt**aagtatt SerA . . .	cttgagc**ag**GGGACAGT . . . rgAsp	0.7
5	82	TCGAAAG**gt**acgatag LysA . . .	ttccctc**ag**ACAGTTAT . . . spSer	2.0
6	197	AGCAAAG**gt**aaggcac AlaLys	ttactttt**ag**GTATCCAT ValSer	1.4
7	75	AATGCAG**gt**atcaaag MetGln	tttttcc**ag**ATTGTCTT IleVal	1.2
8	120	CCTTAAG**gt**ataggac LeuLys	tactaac**ag**GTCCATGA ValHis	3.0
9	76	GCAACAG**gt**aagaaca ThrG . . .	ttcttaa**ag**GTTATGGA . . . lyTyr	1.3
10	126	CATACAG**gt**accagtt ThrG . . .	gatcttt**ag**GAGAAAGG . . . lyGlu	1.3
11	180	CACACAG**gt**aatgact ThrG . . .	ttcattt**ag**GAGAAAAG . . . lyGlu	2.5
12	228	CCCCCAG**gt**gagtgct ProG . . .	aaatttc**ag**GTCAAGGT . . . lyGln	2.0
13	149	TCAACAT**gt**agtatct GlnHis	tgtttac**ag**GTCAACAT ValAsn	3.5
14	162	ACAGCAG**gt**aattatt GlnGln	tcacatc**ag**GTCGCAAT ValAla	1.2
15	144	AGTTCAG**gt**gagtgca ValGln	tggtttc**ag**CTTGGAGA LeuGly	1.2
16	1102 Poly(A)$^+$			

arately by taking photographs. Assignment of the FISH mapping data to chromosomal bands was achieved by superimposing FISH signals with DAPI-banded chromosomes [14].

Generation of luciferase reporter constructs

A fragment containing 1.9 kb of 5′-flanking region from the *mStaf* gene was generated by PCR amplification from the genomic clone. This promoter region (−1894/+37) was amplified by *Pfu* DNA polymerase (Stratagene) using the primers 5′-GTGGTTAACATGCCAGAGGCTGTTCACCTG-3′ and 5′-CCAGATCTCGAGAAAATTTTCCAACAATTCC-3′ (gAS*Bgl*II) to introduce a *Hpa*I site at the 5′ end and a *Bgl*II site at the 3′ end, respectively. The amplified fragment was cloned into pCR 2.1-TOPO (Invitrogen) after addition of 3′ A overhangs by *Taq* polymerase. The recombinant plasmid (pCR1894) was digested with *Hpa*I and *Bgl*II and the fragment containing the *mStaf* gene was cloned into the *Sma*I–*Bgl*II site of pGL-3 Basic (Promega) to generate a luciferase reporter plasmid, pmStaf-1894Luc. The construct was confirmed to have the expected sequence by DNA sequencing. The 5′ deletion constructs, pmStaf1021Luc, pmStaf598Luc, pmStaf448Luc and pmStaf-387Luc, were prepared by digestion of pCR1894 with *Bgl*II and *Hinc*II, *Stu*I, *Eco*ICRI or *Sma*I, respectively, followed by cloning into the *Sma*I–*Bgl*II site of pGL-3 Basic. The reporter plasmid, pmStaf328Luc, was prepared as follows; the −328/+37 region of the *mStaf* promoter was generated by PCR amplification with *Pfu* DNA polymerase using the primers 5′-CCCAGAATCGG-CATAATTCACTCGTCC-3′ and gAS*Bgl*II, to introduce a 5′ blunt end and a *Bgl*II site at the 3′ end. The amplified fragment was digested with *Bgl*II and the resultant fragment was cloned into the *Sma*I–*Bgl*II site of pGL-3 Basic to make pmStaf328Luc. The reporter plasmids pmStaf221Luc, pmStaf153Luc and pmStaf78Luc were generated in a similar way. Site-directed mutagenesis was performed using Quik Change™ Site-Directed Mutagenesis Kit (Stratagene). The following oligonucleotides and their complements were used (with mutated bases underlined; all 5′–3′): Sp1 site I, GGTTGCGGGGC-TTGGGGGTTAGGCGGCTGTTTCCTGTCCTGG; Sp1 site II, CAGCCCAATGGAGGGTAGGGGTTAGACTCAGAA-TTGCTATGC; Sp1 site III, CAATGACGCAGGTGAGGT-GGGTTAAGTTAGTGGCGGTTTCCTGC; and Sp1 site IV, CATAATTCACTCGTCCGTGGGGTTAAGTGGTCCGC-GATGAGC (see Figure 6 below). The entire promoter fragment was sequenced to verify desired mutations.

Transient transfection and luciferase assay

A mouse mammary epithelial cell line, NMuMG, was used for transient-transfection assays. Cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum containing 10 μg/ml insulin (Calbiochem) in six-well culture plates and grown to 70–80% confluence. Co-transfection experiments were carried out using LipofectAMINE Plus (Life Technologies) with 0.5 μg of reporter plasmids and 0.5 μg of pSV-β-gal (Promega) per well. After 24 h, cells were harvested

and assayed for luciferase and β-galactosidase activities. All experiments were carried out in duplicate and independently performed at least three times.

Preparation of nuclear extracts and electrophoretic mobility-shift assays (EMSAs)

Nuclear extracts were prepared from NMuMG cells according to the method described previously [6]. EMSAs were performed using double-stranded oligonucleotides corresponding to the $-46/-12$, $-86/-52$, $-190/-155$ or $-312/-277$ regions of the *mStaf* gene as described previously [6]. Briefly, 6 μg of various nuclear extracts was incubated with the [γ-^{32}P]ATP-labelled oligonucleotide in a binding buffer (14 mM Hepes, pH 7.9, 12 % glycerol, 90 mM NaCl, 2.5 mM $MgCl_2$ and 1 mM dithiothreitol) for 30 min at room temperature. The resulting mixture was subjected to electrophoresis on a 4 % polyacrylamide gel in 0.25 × Tris-borate buffer. Competition experiments were performed using unlabelled wild or mutated types of competitors in the indicated molar excess (see Figure 6 below). Antibody supershift experiments were performed incubating reaction mixtures with antiserum against Sp1, Sp3 or Sp4 (Santa Cruz Biotechnology) at 25 °C for 1 h.

RESULTS

Cloning and sequencing of the *mStaf* gene

Three positive clones, λ1, λ2 and λ3, were isolated from a mouse genomic λFIX-II library and characterized by restriction mapping, subcloning and sequencing. The three clones overlapped partially and spanned from exon 2 to exon 16 but did not include exon 1. Therefore, mouse BAC genomic filters were screened with a probe for the 5′-untranslated region and one positive clone, designated BAC P1, was isolated. Sequencing showed that this clone contained exon 1 and the 5′-flanking region as well as exons 2–16. These clones and the exon organization are shown in Figure 1. The *mStaf* gene spans approx. 35 kb and consists of 16 exons and 15 introns. All exon–intron boundaries begin with GT at the 5′ end and terminate with AG at the 3′ end, conforming to the GT/AG rule (Table 1) [15,16]. The mStaf protein is composed of an N-terminus containing four repeated motifs (activation domain), zinc-fingers (DNA-binding domain) and a C-terminal region of unknown function [6]. The activation domain and the DNA-binding domain are encoded by exons 3–6 and exons 8–13, respectively.

Chromosomal location of the *mStaf* gene

Under the conditions used, the detection efficiency was 68 % for the probe (among 100 mitotic figures examined, 68 of them showed hybridization signals on one pair of chromosomes). Assignment of the FISH mapping data to chromosomal bands was achieved by superimposing FISH signals with DAPI-banded chromosomes. Based on the summary of 10 photos, the detailed position was determined to be mouse chromosome 7, region E3-F1 (Figure 2, upper panel). There were no other positive loci detectable under the conditions used; therefore, this gene was mapped to mouse chromosome 7 (Figure 2, lower panel).

Determination of the transcription start site and polyadenylation site

Primer-extension analysis was performed to determine the transcription start site of the *mStaf* gene. Poly(A)$^+$ RNA from lactating mammary glands was used for this assay. As shown in Figure 3, an extended fragment of 117 bp was obtained with the mammary-gland-derived RNA (Figure 3, lane 1) but not with mouse tRNA (Figure 3, lane 2). We repeated the experiments

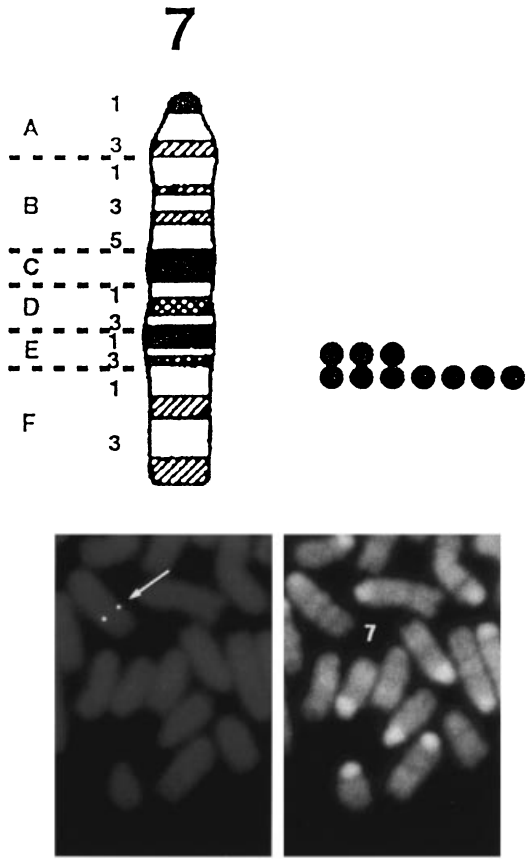

Figure 2 Chromosomal localization of the *mStaf* gene by FISH mapping

(Upper panel) Diagram of FISH-mapping results. Each dot represents the double FISH signals detected on mouse chromosome 7. (Lower panel) An example of FISH mapping. Left, the FISH signals on mouse chromosome; right, the same mitotic image stained with DAPI to identify mouse chromosome 7.

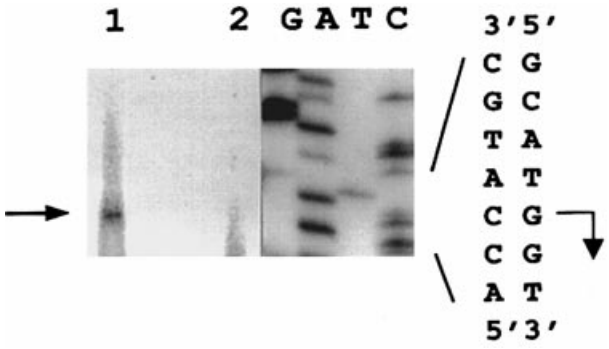

Figure 3 Primer-extension analysis of the transcription start site of *mStaf*

Primer-extension analysis was carried out to map the transcription start site. The extended transcript was obtained from mammary gland poly(A)$^+$ RNA (lane 1), whereas no fragment was detected from mouse tRNA (lane 2). A sequencing ladder of the corresponding genomic fragment was electrophoresed in parallel. The extended band and the nucleotide of the transcription start site are indicated by arrows on the left and right, respectively.

```
GGGCCACCGGGAGCTCCATGGGTTCCACATTAGCCCTGTAGGGAAAATGGCGCCCATCAC    -402

CCTATTTCCCTCCCGGGCAATCAAAAGGACGAACCTCCCTCCCTTGCTACAAAATGGCGC    -342

CTCCCGCTGGGGCCCCAGAATCGGCATAATTCACTCGTCCGTGGGGCGGAGTGGTCCGCG    -282
                        Nkx                  Sp1
ATGAGCAGTTTCGCCATCCAGTCAGGCACAGAAGCTTTGATTCCGTACCTGGCGCCAGCC    -222
                                                          CP2
ACTCATGAGCGGGTAGGAAGGTCAGGCAACAATGACGCAGGTGAGGTGGGCGTAGTTAGT    -162
                                      E2A          Sp1
GGCGGTTTCCTGCCCGTCAGCTAAATCTAGACCAATCGAGAGCAAGAGTCTTCCAGTGAG    -102
                  SIF     TFII-I
TGACGTAAAGACAGCCCAATGGAGGGTAGGGGCGGGACTCAGAATTGCTATGCGGGTTGC     -42
  CRE                          Sp1
GGGGCTTGGGGGCCGGGCGGCTGTTTCCTGTCCTGGTGCATGGTGGTCGGACGAAGGAAT     +18
          Sp1                          +1
TGTTGGAAAATTTTCTCGGAG                                            +40
```

Figure 4 Nucleotide sequence of the 5′-flanking region of the *mStaf* gene

Position +1 denotes the transcription start site. Possible regulatory elements are underlined.

several times and obtained the same results. Based on these results, the transcription start site of the *mStaf* gene was determined to be located 83 bp upstream of the initiation codon.

The size of exon 16 along with the polyadenylation site of the *mStaf* gene were determined by 3′-RACE and subsequent sequencing of the product. The cDNA primed at the polyadenylation site was synthesized from mouse mammary-gland RNA and then subjected to PCR using a specific *mStaf* primer. The PCR product was sequenced, and the size of exon 16 and the polyadenylation site were determined (Table 1).

Transcriptional activity of the *mStaf* promoter

Sequence analysis of the putative transcription-initiation region demonstrated the absence of a consensus TATA box but the presence of various potential regulatory sequences (Figure 4).

We generated a series of chimaeric *mStaf* promoter/luciferase reporter-gene constructs, as depicted schematically in Figure 5. These constructs were transfected into NMuMG cells. The construct containing −1894 to +37 of the 5′-flanking region (pmStaf1894Luc) produced approx. 160-fold higher luciferase activity than the control (pGL-3 Basic) level. The highest activity was observed with pmStaf387Luc. Constructs containing less than 387 bp upstream of the transcription start site showed a progressive loss of promoter activity and a construct of pmStaf-78Luc showed only 2.9-fold higher activity than pGL-3 Basic.

Sp1 often interacts with various promoters and has a role in constitutive transcription [17–21]. In fact, four Sp1-binding sites exist in the *mStaf* promoter at positions −33 to −24, −73 to −64, −175 to −166 and −299 to −290. The relevance of these Sp1 sites for promoter activity was investigated by examining reporter plasmids containing mutations in the Sp1 sites of the *mStaf* promoter. Mutants were generated from pmStaf387Luc by introducing mutations in a single or multiple Sp1 sites (Figure 6, upper panel). Luciferase analysis demonstrated that a single mutation in each Sp1 site alone had essentially no effect on the promoter activity (Figure 6, lower panel). However, mutation of Sp1-II and Sp1-III reduced the promoter activity by about 60%. Mutations of Sp1-I, Sp1-II and Sp1-III or Sp1-II, Sp1-III and Sp1-IV decreased promoter activity by 75 and 70%, respectively.

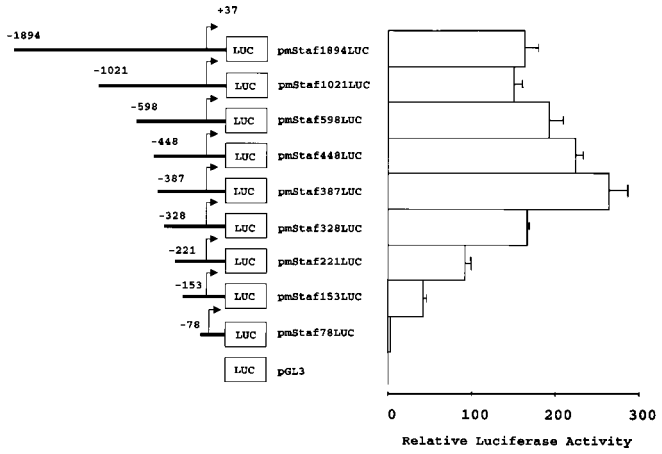

Figure 5 Deletion analysis of the *mStaf* gene promoter

A schematic presentation of the structures of luciferase reporter plasmids linked to various promoter sequences of *mStaf* is shown on the left. Relative luciferase activities of the deletion mutants are shown on the right. The promoter activity of each reporter construct is presented as a value relative to that of the pGL-3 Basic vector. Values are the means ± S.E.M. of at least three independent experiments.

Mutations of all four Sp1 sites reduced the promoter activity by as much as 96% compared with that of pmStaf387Luc. These results indicated that the four Sp1 sites are critical for the basal promoter activity of *mStaf*.

Sp3 interacts with Sp1 sites in the *mStaf* promoter

Because members of the Sp1 family share the same binding consensus sequence, some of them may play important roles in transcriptional regulation. Thus we carried out a series of EMSAs to determine which proteins in NMuMG cell nuclear extracts were interacting with the Sp1 sites. When the nuclear extracts were incubated with a radiolabelled oligonucleotide com-

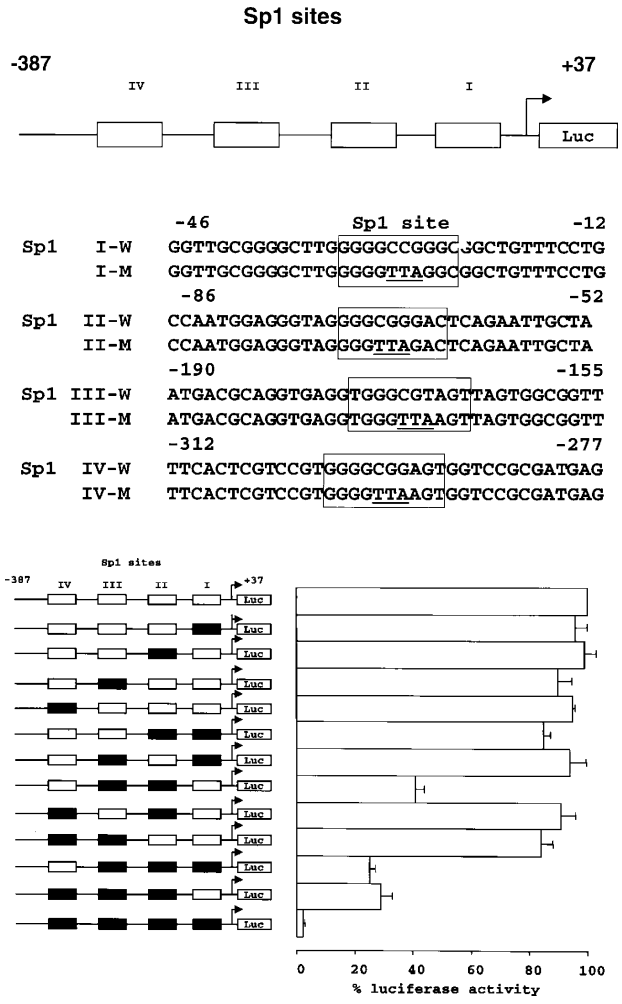

Figure 6 Effects of Sp1 site mutations on the *mStaf* gene promoter activity

(Upper panel) The locations of the four Sp1 sites within the *mStaf* promoter are shown. Sequences of the wild-type or mutated oligonucleotides used in EMSAs or to produce pmStaf387Luc variants are shown. Putative Sp1 sites are represented in bold and introduced mutations are underlined. (Lower panel) A schematic representation of the structures of pmStaf387Luc reporter construct and various Sp1 variants are shown on the left. Activity of pmStaf387Luc is set as 100%. Wild-type Sp1 sites and mutated Sp1 sites are indicated in white and black boxes respectively.

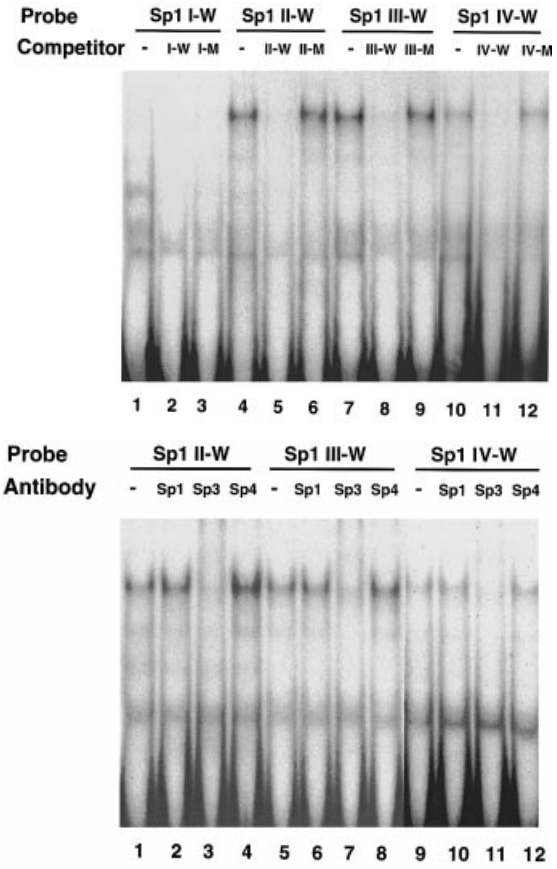

Figure 7 Binding of Sp3 protein to the Sp1 sites of the *mStaf* promoter

EMSA was performed with the indicated ^{32}P-labelled oligonucleotides, incubated with 6 μg of nuclear extract from NMuMG cells. (Upper panel) Competition experiments were performed with 20-fold molar excess of the indicated unlabelled oligonucleotides. (Lower panel) The specific antibody included in the reaction mixture is indicated.

prising the Sp1-II-W sequence, a major slowly migrating band was observed (Figure 7, upper panel, lane 4). This complex was specific because it was competed with by an unlabelled wild-type competitor, Sp1-II-W, but not by an unlabelled mutated competitor, Sp1-II-M (Figure 7, upper panel, lanes 5 and 6). When the nuclear extracts were incubated with radiolabelled Sp1-III-W or Sp1-IV-W, similar results to Sp1-II-W were obtained (Figure 7, upper panel, lanes 7–9 and 10–12). However, no specific bands were obtained for the nuclear extracts with a radiolabelled oligonucleotide, Sp1-I-W (Figure 7, upper panel, lanes 1–3). We investigated further to identify which member(s) of Sp1 family was involved in the DNA binding. The major slowly migrating band formed by radiolabelled Sp1-II-W was interrupted by addition of anti-Sp3 antibody but not by addition of anti-Sp1 or -Sp4 antibodies (Figure 7, lower panel, lanes 1–4). Similar results were obtained when nuclear extracts were incubated with radiolabelled Sp1-III-W or SP1-IV-W (Figure 7, lower panel, lanes 5–8 and 9–12).

DISCUSSION

The present study revealed the structure of the *mStaf* gene. This gene is approx. 35 kb long and split into 16 exons. Sequence analysis of exon–intron boundaries revealed that all junctions followed the normal consensus-sequence rules (Table 1) [15,16]. The sequence of the 5′-flanking region of the gene has several notable features. First, the upstream region contained no consensus TATA box. Similar findings have been reported for a variety of genes for transcription factors, such as the human *MAZ* [22] and the mouse *Kiz1/Limk1* [23]; for oncogenes, such as murine c-*Ki-ras* [24] and human c-*src* [25]; and for other kinds of genes, such as the human gene for insulin-like growth factor receptor [26] and the human *cyclin A*1 [27]. Secondly, the *mStaf* promoter has multiple Sp1-binding sites in the proximal upstream region (−33 to −24, −73 to −64, −175 to −166 and −299 to −290) as was often found in the case of TATA-less promoters [17,18].

In order to determine the elements that regulate the basal transcription of the *mStaf* gene, we constructed fusion genes consisting of various 5′-flanking regions of *mStaf* and the luciferase gene. Transient-transfection experiments using the

fusion genes demonstrated that the construct containing up to position −387 of the 5′-flanking region (pmStaf387Luc) has the strongest promoter activity. The data suggest that the region contains full basal promoter activity. The promoter activity was reduced with the stepwise deletion, which included the removal of putative regulatory elements, such as Sp1, cAMP-responsive element (CRE), TFII-I, SIF (SIS-inducible factor), E2A, CP2 and Nkx (Figures 4 and 5).

Among these elements, we focused on Sp1 sites to determine their contribution to the basal promoter activity of *mStaf*. It was reported that Sp1 sites facilitate binding of transcription factors, including members of the Sp1 family [28]. The Sp1 family includes Sp2, Sp3 and Sp4, which have been cloned and shown to be homologous to Sp1 [29]. This group possesses a DNA-binding domain consisting of three zinc fingers, which is critical for constitutive transcription of various TATA-less promoters [19–21,30]. Our mutation analysis demonstrated that the double mutation of Sp1-II and Sp1-III reduced the promoter activity of *mStaf* by 60% although the mutation of each Sp1 site itself has little effect (results not shown). Furthermore, mutations of three or all of the four Sp1 sites caused 70 or 96% decreases of the promoter activity, respectively. EMSAs were performed using nuclear extracts from NMuMG cells, demonstrating that Sp3 protein could bind to three Sp1 sites, Sp1-II, Sp1-III and Sp1-IV, in the proximal *mStaf* promoter, whereas other Sp (specificity protein) family members did not contribute to the binding. All these data indicate that the Sp3 protein is important for the basal promoter activity of *mStaf*.

Our deletion experiments also suggest that the regions between −387 and −328 bp and between −153 and −78 bp, where no Sp1 sites exist, contain potential regulatory elements (Figure 6). It is possible that other transcriptional activators also contribute to the activation of the *mStaf* promoter. Recent studies have revealed that the Sp1 family can act not only alone but also co-operatively with other transcription factors. For example, Sp1 and CRE-binding protein work together to activate mouse chromogranin A promoter [31]. Sp3 participates in activation of the leukocyte integrin gene, *CD11c*, by interaction with c-Jun [32]. Thus it is possible that Sp3 forms a transcription complex with other transcription factors to increase the promoter activity. It is important to identify these factors and their interaction in order to elucidate the exact nature of the transcription factors and the specificity of expression of the *mStaf* gene from this promoter region. Our previous studies indicated that the *mStaf* gene is expressed differently among various mouse tissues [6]. Similar results were obtained with ZNF143, the most homologous human counterpart of mStaf [12]. Such a difference may be due to the difference in the activity and/or interaction of these transcription factors among different tissues.

Based on the results of FISH analysis, *mStaf* is located at a locus on mouse chromosome 7, region E3-F1. The genes on mouse chromosome 7 are distributed on a variety of human chromosomes including chromosomes 10, 11, 15, 16 and 19 [33]. Disorders of both mouse chromosome 7 and human chromosome 11 are implicated in fetal overgrowth syndromes such as Beckwith–Wiedemann syndrome, which causes developmental and malignant disorders [34,35]. It has been suggested that these disorders are caused by imprinting errors in the insulin-like growth factor-2 (*Igf2*) gene located in these regions [34,35]. However, there is a report indicating that the abnormal expression of genes other than *Igf2* may induce many of the features associated with the disorders [36]. Furthermore, ZNF143 has been localized to human chromosome 11 p15.3-p15.4, a region associated with Beckwith–Wiedemann syndrome [10]. Therefore, it is of interest to investigate the possible involvement of mStaf in such a disease.

We thank Dr. Karen Usdin and Dr. Deborah M. Hinton for their discussion and critical reading of the manuscript.

REFERENCES

1 Stadtman, T. C. (1996) Annu. Rev. Biochem. **65**, 83–100
2 Baron, C. and Bock, A. (1995) in tRNA: Structure, Biosynthesis, and Function (Soll, D. and RajBhandary, V. L., eds.), pp. 529–606, American Society for Microbiology, Washington DC
3 Hatfield, D. L., Lee, B. J., Smith, W. E. and Oroszlan, S. (1990) in Progress in Molecular and Cellular Biology (Jeanteur, P., Kuchino, Y., Muller, W. E. G. and Plaine, P. L., eds.), pp. 115–138, Springer Verlag, Berlin
4 Lee, B. J., Kang, S. G. and Hatfield, D. (1989) J. Biol. Chem. **264**, 9696–9702
5 Meibner, W., Wanandi, I., Carbon, P., Krol, A. and Seifart, K. H. (1994) Nucleic Acids Res. **22**, 553–559
6 Adachi, K., Saito, H., Tanaka, T. and Oka, T. (1998) J. Biol. Chem. **273**, 8598–8606
7 Adachi, K., Tanaka, T., Saito, H. and Oka, T. (1999) Endocrinology **140**, 618–623
8 Burk, R. F. and Hill, K. E. (1993) Annu. Rev. Nutr. **13**, 65–81
9 Aceves, C. and Valverde, C. (1989) Endocrinology **124**, 2818–2820
10 Tommerup, N. and Vissing, H. (1995) Genomics **27**, 259–264
11 Ragoussis, J., Senger, P., Mockridge, I., Sanseau, P., Ruddy, S., Dudley, K., Sheer, D. and Trowsdale, J. (1992) Genomics **14**, 673–679
12 Myslinski, E., Krol, A. and Carbon, P. (1998) J. Biol. Chem. **273**, 21998–22006
13 Heng, H. H. Q., Squire, J. and Tsui, L.-C. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 9509–9513
14 Heng, H. H. Q. and Tsui, L.-C. (1993) Chromosoma **102**, 325–332
15 Breathnach, R. and Chambon, P. (1978) Annu. Rev. Biochem. **50**, 349–383
16 Stephens, R. M. and Schneider, T. D. (1992) J. Mol. Biol. **228**, 1124–1136
17 Azizkhan, J. C., Jensen, D. E., Pierce, A. and Wade, M. (1993) Crit. Rev. Eukaryotic Gene Exp. **3**, 229–254
18 Bird, A. P. (1986) Nature (London) **321**, 209–213
19 Briggs, M. R., Kadonaga, J. T., Bell, S. P. and Tjian, R. (1986) Science **234**, 47–52
20 Courey, A. J. and Tjian, R. (1988) Cell **55**, 887–898
21 Kadonaga, J. T. and Tjian, R. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 1566–1570
22 Song, J., Murakami, H., Tsutsui, H., Tang, X., Matsuura, M., Itakura, K., Kanazawa, I., Sun, K. and Yokoyama, K. K. (1998) J. Biol. Chem. **273**, 20603–20614
23 Bernard, O., Burkitt, V., Webb, G. C., Bottema, C. D., Nicholl, J., Sutherland, G. R. and Matthew, P. (1996) Genomics **35**, 593–596
24 Hoffman, E. K., Trusko, S. P., Murphy, M. and George, D. L. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 2705–2709
25 Bonham, K. and Fujita, D. J. (1993) Oncogene **8**, 1973–1981
26 Evans, T., DeChiara, T. and Efstatiadis, A. (1988) J. Mol. Biol. **199**, 61–81
27 Muller, C., Yang, R., Beck-von-Peccoz, L., Idos, G., Verbeek, W. and Koeffler, H. P. (1999) J. Biol. Chem. **274**, 11220–11228
28 Hagen, G., Muller, S., Beato, M. and Suske, G. (1994) EMBO J. **13**, 3843–3851
29 Kingsley, C. and Winoto, A. (1992) Mol. Cell. Biol. **12**, 4251–4261
30 Lu, J., Lee, W., Jiang, C. and Keller, E. B. (1994) J. Biol. Chem. **269**, 5391–5402
31 Hocker, M., Raychowdhury, R., Plath, T., Wu, H., O'Cornnor, D. T., Wiedenmann, B., Rosewicz, S. and Wang, T. C. (1998) J. Biol. Chem. **273**, 34000–34007
32 Noti, J. D. (1997) J. Biol. Chem. **272**, 24038–24045
33 Kataoka, M., Yoshiyama, K., Matsuura, K., Hijiya, N., Higuchi, Y. and Yamamoto, S. (1997) J. Biol. Chem. **272**, 18209–18215
34 Giannoukakis, N., Deal, C., Paquette, J., Goodyer, C. G. and Polychronakos, C. (1993) Nat. Genet. **4**, 98–101
35 Eggenschwiler, J., Ludwig, T., Fisher, P., Leighton, P. A., Tilghman, S. M. and Efstratiadis, A. (1997) Genes Dev. **11**, 3128–3142
36 McLaughlin, K. J., Kochanowski, H., Solter, D., Schwarzkopf, G., Szabo, P. E. and Mann, J. R. (1997) Development **124**, 4897–4904

Received 30 July 1999/25 October 1999; accepted 26 November 1999