# The nuclear autoantigen La/SS-associated antigen B: one gene, three functional mRNAs

Daniel GRÖLZ and Michael BACHMANN*

Institut für Physiologische Chemie, Johannes-Gutenberg Universität, Duesbergweg 6, D-55099 Mainz, Germany

Transcription of the gene encoding for the nuclear autoantigen La resulted in three mRNA forms. A promoter switching combined with an alternative splicing pathway replaced exon 1 with either exon 1′ or exon 1″. The exon 1″ donor splice site was located 4 nts downstream of the exon 1′ donor splice site. All three La mRNA forms were expressed in all the tissues analysed including peripheral blood lymphocytes, liver, fetal spleen, cultured primary endothelial cells, and mouse LTA cell lines permanently transfected with the human La gene. Both the exons 1′ and 1″ had unusual structures. They contained GC-rich regions and an oligo(U)-tail of 23 uridine residues. Moreover, they encoded for three open reading frames upstream of the La protein reading frame. In spite of this unusual structure, when exon 1′ or exon 1″ La mRNAs were expressed in transfected mouse LTA cells, both La mRNAs were translated to nuclear La protein, indicating that all La mRNA forms are functional mRNAs.

## INTRODUCTION

Sera of patients with autoimmune diseases such as systemic lupus erythematosus (SLE) or primary Sjögren's syndrome (pSS) frequently develop autoantibodies to nuclear RNA binding proteins [1]. One of the targets of such autoantibodies is the so-called La (SS-B) protein. The La protein can be found associated with all primary RNA polymerase III transcripts including precursor molecules of ribosomal 5 S RNA, the tRNAs, some 4.5 S RNAs, as well as some of the U1 and U6 small nuclear RNAs [2–4].

Common to all primary RNA polymerase III transcripts is the 3′-terminal oligo(U)-tail, which is transcribed during the transcription termination step. These oligo(U)-tails were shown to be a binding site for the La protein [5]. In most cases the binding of the La protein to the primary RNA polymerase III transcripts is only transient, as their respective oligo(U)-tails are removed after transcription during an unknown processing step. However, some cytoplasmic small La RNAs retain their oligo(U)-tails even after transport to the cytoplasm [6]. In addition to the oligouridylated RNA polymerase III transcripts, an association of the La protein with some non-oligouridylated RNAs has been reported, especially for some viral RNAs, including the leader RNAs of the vesicular stomatitis virus and rabies virus [7,8].

The La protein is assumed to be involved in transcription termination of RNA polymerase III and internal initiation of translation of at least the polio virus mRNA [9–12].

Most recently five La cDNAs were isolated, when a cDNA library made from peripheral blood lymphocytes (PBL) of a patient with pSS was screened with her own anti-La serum [14]. In two of these five La cDNAs exon 1 was replaced with an alternative 5′-end, which started with a 5′-oligo(dT)-tail. Genomic analysis revealed that these La cDNAs were derivatives of shortened alternative transcripts of the La gene. Genomic analysis revealed that an alternative promoter site was located in the intron between exons 1 and 2, which served as initiation site for transcription of the alternative exon 1′. During further genomic studies three La pseudogenes were identified and sequenced (EMBL accession numbers X91336, X91337, X91338; [13]). All pseudogenes were processed retropseudogenes. Two of the pseudogenes were derivatives of exon 1 La mRNAs, and one was a derivative of an exon 1′ La mRNA. One of the exon 1 pseudogenes lacked exon 3, and the exon 1′ pseudogene lacked 172 nts of the exon 1′ sequence. These differences were the results of further alternative splicing pathways.

This was the impetus to look for further alternative transcripts. Here we report the identification of a further alternative La mRNA form. In addition, we show that both alternative transcripts represent translatable mRNAs in spite of their unusual 5′-terminus.

## MATERIALS AND METHODS

### Materials

Materials were obtained as follows: *Bst*EII, *Eco*O109, *Eco*RI, *Kpn*I, and T7-sequencing kit from Pharmacia (Freiburg, Germany); *Sal*I, *Nhe*I, and *Nco*I from MBI Fermentas (St. Leon-Rot, Germany); QIAprep-spin kit and QIAEX from Qiagen (Hilden, Germany); Dulbecco' modified Eagle's medium (DMEM), Opti-MEM medium, Lipofectamine, and Taq polymerase from Gibco BRL Life Technologies (Eggenstein, Germany); *Bgl*II, pfu polymerase, pBluescript SK(−) from Stratagene GmbH (Heidelberg, Germany); CDP-Star Tropix, pCI-neo, and pGEM-T vector systems from Promega (Serva, Heidelberg, Germany); *Bst*XI, *Hind*III, Taq-buffer (10 ×), DNA molecular-mass marker VI, positively charged nylon membrane (1209272), blocking reagent, RNA/DNA-digoxigenin (DIG) labelling kit from Boehringer Mannheim (Mannheim, Germany);

Abbreviations used: BCIP, 5-bromo-4-chloro-3-indolyl phosphate; cLSM, confocal laser scanning microscopy; DIG, digoxigenin; DMEM, Dulbecco's modified Eagle's medium; EC, endothelial cells; FCS, fetal calf serum; Fs, fetal spleen; L, liver; mAb, monoclonal antibody; NBT, 4-Nitro Blue Tetrazolium chloride; ORF, open reading frame; PBL, peripheral blood lymphocytes; $P_p$, PBL of a patient with primary Sjögren's syndrome; $P_h$, PBL of a healthy donor; PS, pseudogene; pSS, primary Sjögren's syndrome; 5′-RACE, 5′-rapid amplification of cDNA ends; SLE, systemic lupus erythematosus; SS-B, SS-associated antigen B; UAP, universal amplification primer.

* To whom correspondence should be addressed.

agarose and NuSieve agarose from Biozym (Hameln, Germany); anti-human IgG ($\gamma$-chain specific) conjugated with alkaline phosphatase (A3150; 8 units/ml), and isopropyl $\beta$-D-thiogalactopyranoside (isopropyl $\beta$-D-thiogalactoside) from Sigma (St. Louis, MO, U.S.A.); the enhanced chemiluminescence/Western blotting detection reagents were from Amersham–Buchler (Braunschweig, Germany); 4-Nitro Blue Tetrazolium chloride (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) from Roth (Karlsruhe, Germany); PVDF membrane (IPVH 000 10; pore size 0.45 $\mu$m) was obtained from Millipore (Bedford, MA, U.S.A.).

### Anti-La antibodies and sera

The anti-La monoclonal antibody (mAb) SW5, which was found to be directed to the N-terminal domain of La protein [15], was originally described by Smith et al. [16] and kindly provided by Professor Dr. W. J. van Venrooij (University of Nijmegen, The Netherlands). As anti-La serum we used the serum of the patient (Ma) [12]. In accordance with the anti-La mAb SW5 this anti-La serum preferentially reacted with the N-terminus of the La protein [17].

### Southern blot analysis: preparation and labelling of the probes

For Southern blot analysis DIG-labelled probes were prepared. The probes were obtained by transcription of plasmids *in vitro* containing specific target sequences using T7 RNA polymerase. Before *in vitro* transcription the plasmid containing the respective insert was linearized using *Xho*I.

The exon 1-specific probe was constructed using a *Hae*III fragment containing 67 nts of the exon 1 (nts +25 to +92 according to Chambers et al. [18]) which was prepared from a La cDNA described by Pruijn et al. [19] and cloned into the *Hind*II sites of pBluescript SK(−).

For the exon 1′ probe a series of probes using different parts of exon 1′ as target sequence were made. All of them gave identical results in dot-blot analysis, *in situ* hybridization and Southern blot studies ([20,21]; results not shown). The exon 1′-specific probe, which was used for the Southern blot analysis presented here, was derived from a deletion clone containing the 5′-terminal oligo(dT)-tail of 22 dT residues and the 199 nts located downstream of the exon 1′ cloned into the *Bst*XI/*Xho*I sites of pBluescript SK(−).

Southern blot analysis was performed as described previously [14].

### Polymerase chain reaction

PCR was performed using a TC9600 Cycler (Perkin-Elmer, Überlingen, Germany). The 50 $\mu$l assay in 1 × Taq buffer contained 2 units of *Taq* polymerase, 1.5 mM $MgCl_2$, 5% (v/v) DMSO, 200 $\mu$M dNTP, 20 pmol of each primer, and 1 ng of DNA. Cycling was started by heating for 3 min to 95 °C; 40 cycles followed, each consisting of 30 s at 95 °C, 15 s at 57 °C, and 3 min at 72 °C. Then the temperature was held for 10 min at 72 °C and cooled down to 4 °C.

The PCR products were further analysed by agarose gel electrophoresis. PCR fragments were visualized by ethidium bromide staining. Isolated PCR fragments were either directly sequenced or sequenced after subcloning into pGEM-T [14].

### DNA sequence analysis

DNA was prepared using the QIAprep-spin kit and concentrated by ethanol precipitation to a final concentration of about 1 $\mu$g/$\mu$l in 10 mM Tris/HCl/1 mM EDTA, pH 8, buffer. DNA sequencing was performed as described previously [14].

### 5′-Rapid amplification of cDNA ends (5′-RACE) analysis

For an analysis of the 5′-ends of the La mRNAs we chose the 5′-RACE system supplied by Gibco BRL Life Technologies. It includes a first-strand cDNA synthesis and a PCR amplification step. During reverse transcription 40 units of RNase inhibitor were added to each 1 $\mu$g RNA sample. The RNA samples were isolated from (adult) liver of the tumour patient, human fetal spleen (20th week), the PBL of a patient (Ma) with pSS, the PBL of an (adult) control person, endothelial cells obtained from primary cultures (second passage) of the umbilical vein, and mouse LTA cells, which were untransfected or transfected with the human either La gene or La exon 1 La cDNA [18]. The liver tissue of the tumour patient was taken from a part of the liver tissue that did not obviously show metastasis. In spite of this the tissue displayed an abundant expression of the c-*myc* oncogene (results not shown).

The RNA was isolated as described earlier [14]. The reactions and the following 5′-RACE steps were performed according to the instructions of the supplier. For the first-strand synthesis, the primer P1 located within exon 4 was used (P1, CACTGATTTCCATGAGTTCTGCCTTGG). This primer was extracted with phenol/chloroform and diluted in diethyl pyrocarbonate-treated double-distilled water. The first amplification was performed using the Anchor primer of the supplier as upstream primer and the primer P2 located in the exon 3 (P2, TGTCCCGTGGCAAATTGAAGTCGCC) as downstream primer.

For detection of exons 1′ and 1″ PCR fragments the 5′-RACE products were amplified using the nested primer pair P3/P4 consisting of the upstream primer P3 (P3, ACCGCCTTCTAGTCTCACCGAA) and the downstream primer P4 (P4, GATGACAGATTTTGGCCTCCAG).

For analysis of the 5′-start of the exon 1″ La mRNA isoform PCR amplification was performed using as downstream primer P5 (P5, TTTCAGCCATTGCGGCTATCCA) in combination with the slightly modified universal amplification primer (UAP) according to the supplier P6 (P6, TAGGCCACGCGTCGACTAGTAC). The primer P5 started at the 5′-end with three of the 4 nts of the exon 1″-specific insert between exon 1′ and exon 2. The primer P5 allowed the specific amplification of an exon 1″ fragment from a plasmid DNA template containing an exon 1″ La cDNA derivative, while it failed to amplify a fragment when using a plasmid DNA template containing an exon 1′ La cDNA derivative (results not shown). The detection of the fourth nt of the exon 1″-specific insert in the sequence of a PCR fragment obtained with P5 as downstream primer allowed the clear-cut conclusion that the PCR fragment was indeed derived from an exon 1″ mRNA.

The same exon 1″ PCR fragments obtained by amplification of the 5′-RACE products with the primer pair P5/P6 were also used for construction of the exon 1″ pCI-neo transfection construct (see below).

### Preparation of exon 1, exon 1′ and exon 1″ constructs for transfection studies

Exons 1, 1′ and 1″ constructs were cloned in the pCI-neo transfection vector according to the following strategy. Cloning started from the La cDNA La23 in pBluescript SK(−). The La-insert in La23 started at the 5′-site with an oligo(dT)-tail and represented a 5′-shortened exon 1′ La mRNA derivative. Therefore in the first step an exon 1′ full-length La cDNA was constructed. In order to construct the exon 1″ full-length La

cDNA, the portion of exon 1′ La cDNA between the *Eco*O109 [in exon 1′ downstream of the oligo(dT)-tail] and the *Kpn*I site (in exon 3) was replaced by the respective exon 1″ sequence. In the case of the exon 1 full-length La cDNA, a PCR fragment starting with exon 1 and ending at the *Kpn*I site was cloned into the same *Kpn*I site. Then the respective exon 1, 1′, and 1″ construct was cloned into the transfection vector pCI-neo. During construction we learned that cDNA La23 contained a frame shift mutation in exon 7. As all constructs were derivatives of La23, in a further cloning step the mutated reading frame was replaced with the correct reading frame, which was obtained from La cDNA La19. Based on this strategy, an exons 1, 1′, and 1″ La construct was obtained.

For this cloning procedure 5′-exon 1′, 5′-exon 1, and 5′-exon 1″ fragments were required and prepared by PCR using the proof reading pfu polymerase.

For the exon 1′ fragment PCR was performed using DNA of a genomic subclone as substrate. The subclone was prepared from the charon phage Lambda 2.1 [17]. Restriction of the Lambda DNA with *Eco*RI resulted in a 4.4 and a 4.6 kb fragment. The 4.6 kb *Eco*RI fragment was isolated and subcloned into pBluescript SK(−). This subclone contained, besides exons 1 and 2, an intron between exons 1 and 2 including the exon 1′. PCR was performed using as upstream primer P7 (P7, CGCTTT ACTAGT GCGCGACTGCGCGTTTCC; the artificial *Spe*I site is underlined) and as downstream primer P8 (P8, CTGAAA-CCTGATGTGAGCGATG). The resulting exon 1′ fragment started at the predicted 5′-start of exon 1′ [14] and ended downstream of an *Eco*O109 site, which located downstream of the oligo(dT)-tail in the exon 1′ La sequence. The exon 1′ fragments were cleaved with *Spe*I and *Eco*O109 and cloned into the respective sites of La23. For this purpose La23 was linearized with *Spe*I and the isolated DNA was partially digested with *Eco*O109. In spite of using pfu polymerase, the length of the oligo(dT)-stretch was incorrect in the vast majority of isolated clones.

The 5′-exon 1 fragment was obtained as follows. As substrate we used the La cDNA M13-3, which was a gift from Dr. E. K. L. Chan (Scripps Clinic, La Jolla, CA, U.S.A.). The primer P9 (P9, CGCTTT ACTAGT CGGTCCCCATCTTCTTGG; the artificial *Spe*I site is underlined) served as the upstream exon 1 primer. P1 (see above) served as downstream primer, which was located downstream of the *Kpn*I site in exon 4 of the La sequence. The exon 1 PCR fragments were cleaved with *Spe*I and *Kpn*I and cloned into the respective sites of La23. For this purpose La23 was linearized with *Spe*I and the isolated, linearized DNA was partially digested with *Kpn*I.

In order to conserve the correct oligo(dT)-tail of the exon 1′ La cDNA construct the following strategy was developed for cloning of the exon 1″ construct. First a 5′-shortened exon 1″ clone starting downstream of the oligo(dT) with an 5′-artificial *Spe*I site was prepared. This construct contained, upstream of the 4 nts, the *Eco*O109 site in exon 1′. After cleavage with *Spe*I/*Eco*O109 the full-length exon 1″ was obtained by cloning the 5′-*Spe*I/*Eco*O109 fragment obtained from the exon 1′ pBluescript construct (see above) into the respective sites. For cloning of the 5′-shortened exon 1″ construct the required 5′-terminal fragment was obtained as follows. First a 5′-RACE/PCR exon 1″ product was prepared using as downstream primer P5, in combination with the UAP primer P6 (see above). In order to remove the 5′-oligo(dG)-tail of the 5′-RACE product, nested PCR was performed with the primer P10 as upstream primer (= P3 with a 5′-artificial *Spe*I site; P10, GAGAGA ACTAGT ACCGCCTTCTAGTCTCACCGAA; the 5′-artificial *Spe*I site is underlined) and P5 as downstream primer. The resulting PCR fragment started downstream of the oligo(dT)-tail with the artificial *Spe*I site, contained the 4 nt insert between the 3′-end of the exon 1′ and the 5′-start of exon 2 and ended about 60 nts downstream of the *Eco*O109 site. This fragment served as a megaprimer in the following PCR step. PCR was performed using in addition to the megaprimer an overlapping La cDNA fragment starting at the 5′-end with the *Eco*O109 site and the primer pair P10/P1. P10 served as upstream primer and P1 as downstream primer. As described above the downstream primer P1 located downstream of the *Kpn*I site in exon 4 of the La sequence. The resulting PCR fragment was purified, digested with *Spe*I and *Kpn*I and cloned into the respective sites of the exon 1′ pBluescript SK(−) construct. The resulting construct started downstream of the oligo(dT)-tail and contained the 4 nts between exon 1′ and exon 2. This 5′-shortened exon 1″ construct was cleaved with *Spe*I and partially digested with *Eco*O109. Then the *Spe*I/*Eco*O109 fragment of the full-length exon 1′-pBluescript SK(−) construct was cloned into the respective sites resulting in the full-length exon 1″ construct.

The exon 1, 1′ and 1″ La inserts were isolated from the pBluescript SK(−) constructs by cleavage with *Spe*I and *Xho*I and cloned into pCI-neo, which was restricted with *Neh*I and *Sal*I.

Finally, the reading frame of the exons 1, 1′ and 1″ construct was corrected as follows. La19 cDNA containing the correct La coding sequence was restricted with *Bst*EII, which cleaved at exon 9 of the La sequence, and *Bgl*II, which cleaved at exon 5 of the La sequence. The exons 1, 1′ and 1″ pCI-neo constructs were linearized with *Bst*EII and after isolation of the linearized DNA partially digested with *Bgl*II. Then the *Bgl*II/*Bst*EII fragment of La19 was cloned at the respective sites of the exon 1, 1′ and 1″ pCI-neo construct.

### Transfection, cell culture and staining procedures

Mouse LTA cells were grown in DMEM containing 10% (w/v) fetal calf serum (FCS) in a humidified $CO_2$ atmosphere either in culture flasks for preparation of extracts or on coverslips for epifluorescence microscopy. The mouse LTA cells were transfected transiently according to the following protocol. Transfection was performed in six-well tissue culture plates (35 mm) containing cover slips. Cells were grown to a confluency of 70 to 80% in 4 ml of DMEM containing 10% FCS. Before transfection the serum/medium was removed and the cells were washed for 30 min with 2 ml of DMEM without FCS and antibiotics. In parallel, 1.5 $\mu$g of plasmid DNA (see below) was dissolved in 100 $\mu$l of Opti-MEM medium and combined with 100 $\mu$l of DMEM (without antibiotics and FCS) containing 6 $\mu$l of Lipofectamine. After removal of 2 ml of DMEM, 0.8 ml of DMEM lacking FCS and antibiotics was added to the cells and the DNA mixture followed. After a 5 h incubation, 1 ml of DMEM containing 20% FCS and antibiotics was added. At 24 h after the start of transfection the medium was replaced by 2 ml of DMEM containing 10% FCS and antibiotics. At 44 h after transfection the cells were either harvested for preparation of total extracts or fixed with methanol/EGTA for immunofluorescence microscopy.

In addition to transiently transfected cells, permanently transfected LTA cell lines were used. These lines were kindly provided by Dr. K. Keech of the group of Professor Dr. J. McCluskey and Professor Dr. T. Gordon (Flinders Medical Center, Bedford, South Australia). The cells were transfected with either the human La gene or the human exon 1 La cDNA [22]. The permanently transfected LTA cells were grown in the presence of 0.02% (w/v) geneticin.

For immunofluorescence microscopy the cells were fixed with methanol containing 0.02% (w/v) EGTA at −20 °C for 1 h. Indirect immunofluorescence of cells with anti-La mAb SW5 was performed by incubating the fixed cells, which had been rehydrated for 5 min, with PBS and with cell culture supernatant of hybridomas secreting the respective anti-La mAb for at least 15 min. The cells were washed with PBS (5 min) and the bound anti-La mAb was detected using Cy3-conjugated anti-mouse antibody developed in goat. The incubation time for the secondary antibody was 15 min. The staining occurred at room temperature and the unbound secondary antibodies were removed by washing with PBS (twice, 5 min). The stained cells were mounted using PBS/glycerol (1:1, v/v). Confocal laser scanning microscopy (cLSM) was performed using a Zeiss LSM 10. The stained specimens were cut automatically into horizontal sections (512 × 512 pixels/8 bit, objective lenses Plan-Neofluar 40 × /1.3 oil). Evaluation of the stored stacks of the horizontal optical sections was performed with the LSM 10 image processing unit.

### Preparation of cell extracts

Total extracts from transiently transfected cell lines were prepared by incubation with 350 $\mu$l of SDS/PAGE sample solution (95 °C) [17]. Total extracts from permanently transfected LTA cells were harvested by adding 100 $\mu$l of hot cell lysis buffer [100 mM $Na_2HPO_4$, pH 8.3/200 mM dithiothreitol/1% (v/v) SDS/10% (v/v) glycerol] per cm². The lysed cells were heated for 5 min to 95 °C and centrifuged at 14000 ***g*** for 5 min at 4 °C. Aliquots (5 $\mu$l) were mixed with SDS/PAGE sample solution and used for SDS/PAGE.

### SDS/PAGE and immunoblotting

SDS/PAGE and immunoblotting were performed as described [23,24]. The blocked membrane was incubated with monospecific anti-La antibodies isolated from the serum of the patient Ma [12,17]. The formed immune complexes were visualized using the alkaline phosphatase system and BCIP/NBT as substrate.

## RESULTS

Recently three La pseudogenes (PS1 to PS3) were isolated, characterized and sequenced [13]. All pseudogenes represented retrogenes. As schematically summarized in Figure 1, PS1 and PS3 were derived from an exon 1 La mRNA. PS3 lacked the complete exon 3. PS2 was derived from an exon 1′ La mRNA but lacked a portion of 172 nts of the exon 1′ downstream of the oligo(dT)-tail. The deletion in PS2 and PS3 happened at splice sites. Thus the most likely explanation (see also the Discussion section) for the deletions is that these exon sequences had already been removed by alternative splicing of those La mRNAs, which served as templates during reverse transcription and development of the retrogenes.

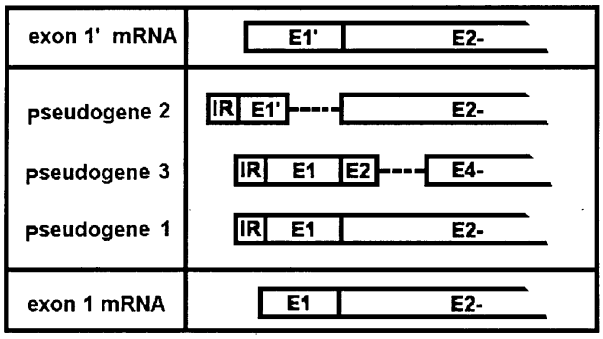



**Figure 1 Comparison of the three La pseudogenes with exon 1 and exon 1′ La mRNA**

E1, exon 1; E1′, exon 1′; E2, exon 2; IR, 5′-direct repeats (insert sites of the La pseudogenes); (---) absent in the pseudogene.

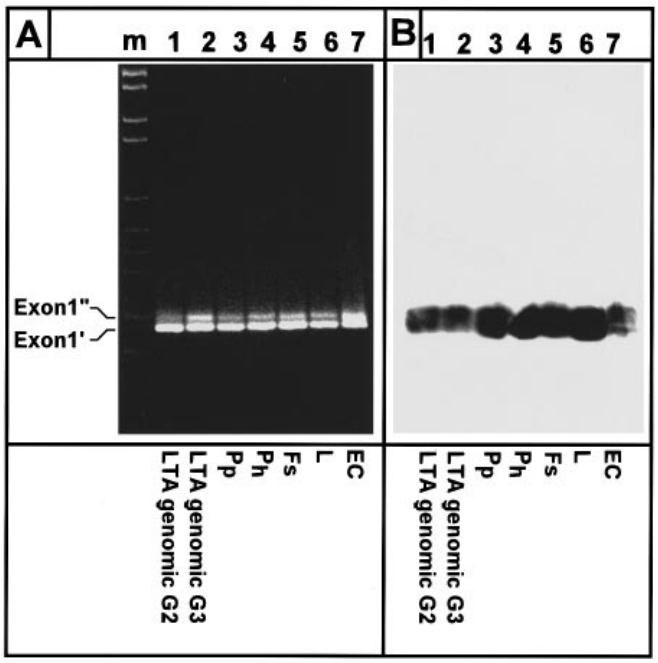



**Figure 2 Identification of exon 1″ La mRNA**

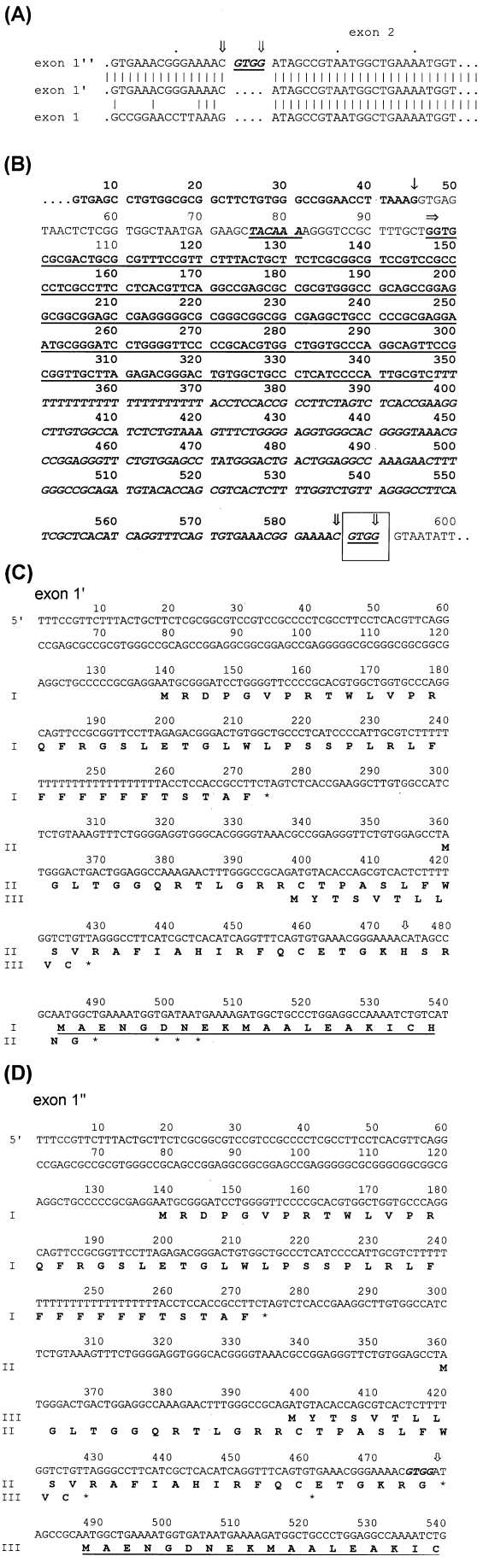
(**A**) 5′-RACE/PCR products obtained with the primer pair P3/P4 separated on a NuSieve agarose gel. (**B**) Southern blot of (**A**). The PCR products were from either (adult) liver of the tumour patient (L, lanes 6), human fetal spleen (20th week) (Fs, lanes 5), PBL of the patient (Ma) with pSS ($P_p$, lanes 3) or PBL of an (adult) control person ($P_h$, lanes 4), endothelial cells (EC, lanes 7) obtained from primary cultures (second passage) of the umbilical vein, and the mouse LTA cell lines (LTA genomic G2,G3, lanes 1 and 2), which were permanently transfected with the human La gene; m, marker lane.

### 5′-RACE analysis to detect further alternative La mRNAs

Therefore we looked for the existence of such alternatively spliced La mRNA isoforms. For this purpose 5′-RACE products were prepared from a series of tissues (see the Materials and methods section) starting from exon 4. The resulting PCR fragments were analysed by Southern blot analysis, and the detected fragments were eluted, subcloned and sequenced (results not shown). In spite of all this effort, we were not able to detect such alternative La mRNA isoforms.

### Detection of a further alternative La mRNA form: exon 1″

We found, however, that when the 5′-RACE/PCR products were separated on a NuSieve agarose gel (Figure 2A) the exon 1′ band was separated into two bands. Both bands hybridized with the exon 1′ probe (Figure 2B). The upper band represented a further alternatively spliced exon 1′ derivative (see below), which was termed ‘exon 1″ La mRNA form’. The exon 1″ La mRNA form was detected in all tissues analysed including human liver [Figures 2A and 2B (L) lanes 6], the PBL of a healthy donor




[Figures 2A and 2B ($P_h$) lanes 4], the PBL of a patient with pSS ($P_p$; lanes 3), the two mouse LTA cell lines (G2,G3) permanently transfected with the human La gene (LTA genomic G2, G3; lanes 1 and 2), cultured primary endothelial cells (EC; lanes 7) and also in fetal spleen (Fs; lanes 5). The ratio exon 1′/exon 1″ varied between 1:10 and 1:1.

The exon 1″ fragments were excised from the agarose gel and either directly sequenced or subcloned in pGEM-T and then sequenced. Thereby it became obvious that the exon 1″ La mRNA form differed from exon 1′ La mRNA due to an insert of 4 nts between exon 1′ and exon 2 (Figure 3A). When looking to the genomic La sequence (Figure 3B) it became obvious that the 4 nt insert was due to the use of a further alternative donor splice site. This donor splice site was located just the 4 nts downstream of the exon 1′ donor splice site. When exon 1″-specific 5′-RACE/PCR products were analysed, we found that the longest 5′-exon 1″ products started in the same region as was determined for exon 1′ (Figure 3B, nt position 96, position 14). In consequence, like the exon 1′ La mRNA, the exon 1″ La mRNA was the result of a combination of a promoter switching and an alternative splicing pathway.

In consequence, like the exon 1′ La mRNA, the exon 1″ La mRNA had a GC-rich 5′-terminus and contained an oligo(U)-tail of 23 uridine residues. Moreover, it encoded for the three upstream open reading frames (ORFs) 1 to 3 (Figures 3D and 4C). As shown in Figures 3(C) and 4(B) the reading frame of ORF1 in exon 1′ La mRNA was in the reading frame of the La protein but interrupted by two stop codons. ORF2 was in the (+1) reading frame and overlapped the La protein reading frame. The ORF3 was in the (−1) reading frame. As shown in Figures 3(D) and 4(C), the insert of the 4 nts between exons 1′ and 2 in the exon 1″ La mRNA, caused ORF1 to now be in the (−1) reading frame of the La protein, ORF2 to be in the (+1) reading frame and ORF3 to be in the correct reading frame of the La protein. Moreover, the 4 nt insert caused ORF2 to now end upstream of the La protein reading frame. Thus in the case of the exon 1″ La mRNA form, all upstream ORFs are interrupted by stop codons from the La protein reading frame.

## Characterization of exons 1′ and 1″ La mRNAs as functional mRNAs

The properties of exons 1, 1′ and 1″ La mRNAs are schematically compared in Figure 4. In order to determine whether the exons 1′ and 1″ La mRNAs represent translatable La mRNAs, mouse LTA cells were transiently transfected with either an exon 1, an exon 1′ or an exon 1″ construct and stained with the anti-La mAb SW5. Stained cells were analysed using the cLSM technique. As

**Figure 3 Characterization of three La mRNA isoforms**

(**A**) Difference in the sequence of the two 5′-RACE/PCR fragments separated in Figure 2 on NuSieve agarose. Subcloning and sequencing of the 5′-RACE/PCR products showed that the upper band (Figure 2A) differed from the exon 1′ band with respect to an insert (box in **B**) of the 4 nts (5′-*GTGG*) between the exon 1′ donor splice site (↓) and the acceptor splice site of the exon 2 (↓). (**B**) Genomic Sequence. The exons 1′ and 1″ were part of the intron between exons 1 and 2 (EMBL Acc. No.: Z35127). The longest 5′-RACE/PCR exons 1′ and 1″ fragment started (→) approx. 20 nts downstream of the putative TFIID binding site 5′-*TACAAA*. (↓) Exon 1′ donor splice site. (↓) Exon 1″ donor splice site. (**C**) and (**D**) Properties of the exons 1′ and 1″ La mRNA. The 5′-terminus of the exons 1′ and 1″ is GC-rich, and contains an oligo(dT)-tail and the three ATGs (I to III). The amino acid sequences of these upstream ORFs are indicated by bold amino acid symbols. The N-terminal portion of the La protein reading frame is shown by bold and underlined amino acid symbols. Stop codon (*). (↓) Start of exon 2. (**C**) In exon 1′ ATG (I) is in the (0) reading frame of the La protein, ATG (II) is in the (+1) reading frame, and ATG (III) is in the (−1) reading frame. (**D**) In exon 1″ ATG (I) is in the (+1) reading frame, ATG (II) is in the (−1) reading frame, and ATG (III) is in the (0) reading frame. (*GTGG*) represents the 4 nts inserted between 3′-exon 1′ and 5′-exon 2 in exon 1″.

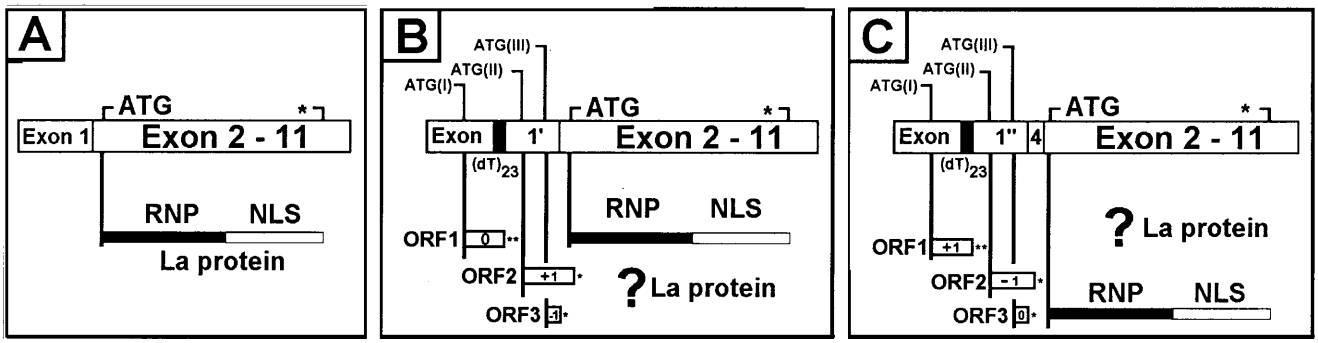



**Figure 4 Schematic summary of the properties of exon 1 (A) , exon 1′ (B) and 1″ (C) La mRNAs**

The human 50 kDa La protein consists of two domains including a 29 kDa N-terminal domain (box filled with black in **A** to **C**) and 25 kDa C-terminal domain (white box in **A** to **C**). The N-terminal domain contains a ribonuclear protein particle (RNP)-consensus motif, while the nuclear location signal (NLS) locates in the C-terminal domain. The La protein reading frame starts in exon 2 (ATG in **A** to **C**). (*) indicate stop codons. Due to the upstream ORFs (1 to 3) in exon 1′ (**B**) and 1″ (**C**) La mRNA, it remained unclear (indicated by ? in **B** and **C**) whether these La mRNAs are translated to La protein. ATG (I) to (III) in (**B**) and (**C**) indicate the start of the upstream ORFs (1 to 3). The black box within the exon 1′ box in (**B**) and also in the exon 1″ box in (**C**) indicates the oligo(dT)$_{23}$-stretch in exon 1′ and 1″. (**B**) In the exon 1′ La mRNA the ORF1 is in the same (0) reading frame as the La protein, while the ORF2 is in the (+1) and the ORF3 is in the (−1) reading frame. The stop codons of the ORF1 and ORF3 are located upstream of the La protein reading frame. ORF2 overlaps with the La protein frame. (**C**) In the exon 1″ La mRNA ORF3 is in the same (0) reading frame as the La protein, while ORF1 is now in the (+1) and ORF2 is in the (−1) reading frame. Due to the insert of 4 nts (indicated by the 4 in the exon 1″ box), ORF2 is now also upstream of the La protein frame.

shown in Figure 5, LTA cells transfected with the pCI-neo vector control lacking a La insert did not result in immunofluorescence staining with the anti-La mAb SW5 (Figure 5c), indicating that this anti-La mAb did not cross-react with the endogenous mouse La protein under the conditions used. In contrast, mouse LTA cells transiently transfected with either the exon 1 (Figure 5e), the exon 1′ (Figure 5f) or the exon 1″ (Figure 5g) pCI-neo construct gave a nuclear staining pattern similar to the pattern of mouse LTA cells permanently transfected with either the human La gene (Figure 5a) or the exon 1 La cDNA (Figure 5b).

The results obtained by immunofluorescence microscopy were confirmed by SDS/PAGE/immunoblotting. For this purpose extracts were prepared from untransfected cells or cells that were transfected with equivalent amounts of either the exon 1, the exon 1′ or the exon 1″ construct. The transfection experiments were repeated five times. No differences between the individual experiments were observed. The results obtained for one of these experiments are shown in Figure 6. The extracts of the transfected cells were analysed by SDS/PAGE/immunoblotting using the monospecific patient's anti-La antibody. The patient's anti-La antibody cross-reacted with the endogenous mouse La protein in all cell extracts analysed (Figure 6, lanes 1 to 5). In cells transiently transfected with either the exon 1′ (Figure 6, lane 3), the exon 1″ (Figure 6, lane 4), or the exon 1 (Figure 6, lane 2), the patient's anti-La antibody detected, in addition to the endogenous mouse La protein, the human La protein. The human La protein was also detected in the extract of the LTA cell lines permanently transfected with the human La gene (Figure 6, lane 5), while it was not detectable in LTA cells transfected with the pCI-neo vector control lacking a human La-specific insert (Figure 6, lane 1). In all five experiments the respective extract of cells transfected with the exon 1 construct appeared to contain a higher concentration of La protein.

## DISCUSSION

Frequently sera of patients with pSS or SLE contain self-reacting antibodies directed to nuclear antigens. One of the targets of anti-nuclear antibodies is the nuclear autoantigen La/SS-associated antigen B (SS-B) [1]. The La protein was described as a house keeping protein [18]. It was proposed to be involved in transcription/termination of RNA polymerase III and in internal initiation of translation, especially of (polio)virus mRNAs [9–13].

In a recent study we searched for alternative La mRNAs by screening of cDNA library made from PBL of a patient with pSS using her own anti-La serum. Thereby five La cDNAs were isolated. Among them two La cDNAs (La19, La23; EMBL Acc. No. X69804) were found, in which exon 1 was replaced with the alternative exon 1′. Analysis of the La gene showed that the transcription of the alternative La mRNA started in the intron between exons 1 and 2 using a promoter switch. The resulting transcript was processed using an alternative splicing pathway [14]. Both La mRNA forms were found to represent finally processed, abundant, cytoplasmic mRNAs, that were expressed in parallel at ratios between 1:1 and 1:5 of exon 1′ to exon 1 [20,21]. The expression ratio varied in dependence on the analysed tissue. Both mRNAs were highly expressed in salivary gland tissue of patients with pSS, including the infiltrating lymphocytes. During genomic analysis we characterized and sequenced the three La retropseudogenes PS1 to PS3 (EMBL Acc. Nos. X91336; X91337; X91338) [13]. PS3 lacked the complete exon 3. PS2 lacked a portion of 172 nts of exon 1′ between the acceptor splice site of exon 2 and a putative alternative donor splice site in exon 1′. The lack of these sequences in PS2 and PS3 can be explained by several mechanisms. One explanation is that these sequences did not exist in the La gene at the time when the respective retrogenes were developed. Another possibility is that they are due to a deletion after development of the respective retrogene. Alternatively, the La mRNAs that served as substrate during retrotransposition represented alternatively spliced La mRNAs that already lacked the respective sequences. Based on the amount of mutations, we estimated that the retrogenes developed approx. 4 (PS1), 4.5 (PS2) and 5 (PS3) million years ago [25]. Having in mind that the mouse, rat and bovine La sequences already contain exon 3 [22,26,27], it appears rather unlikely that the human La gene lacked exon 3 at the time of development of PS3. Moreover, as the 5′- and 3′-sites flanking the deletions in PS2 and PS3 represent donor and acceptor splice sites, it appeared most likely that the lack of exon 3 in PS3 and the lack of the 172 nts in PS2, was due to alternative splicing.

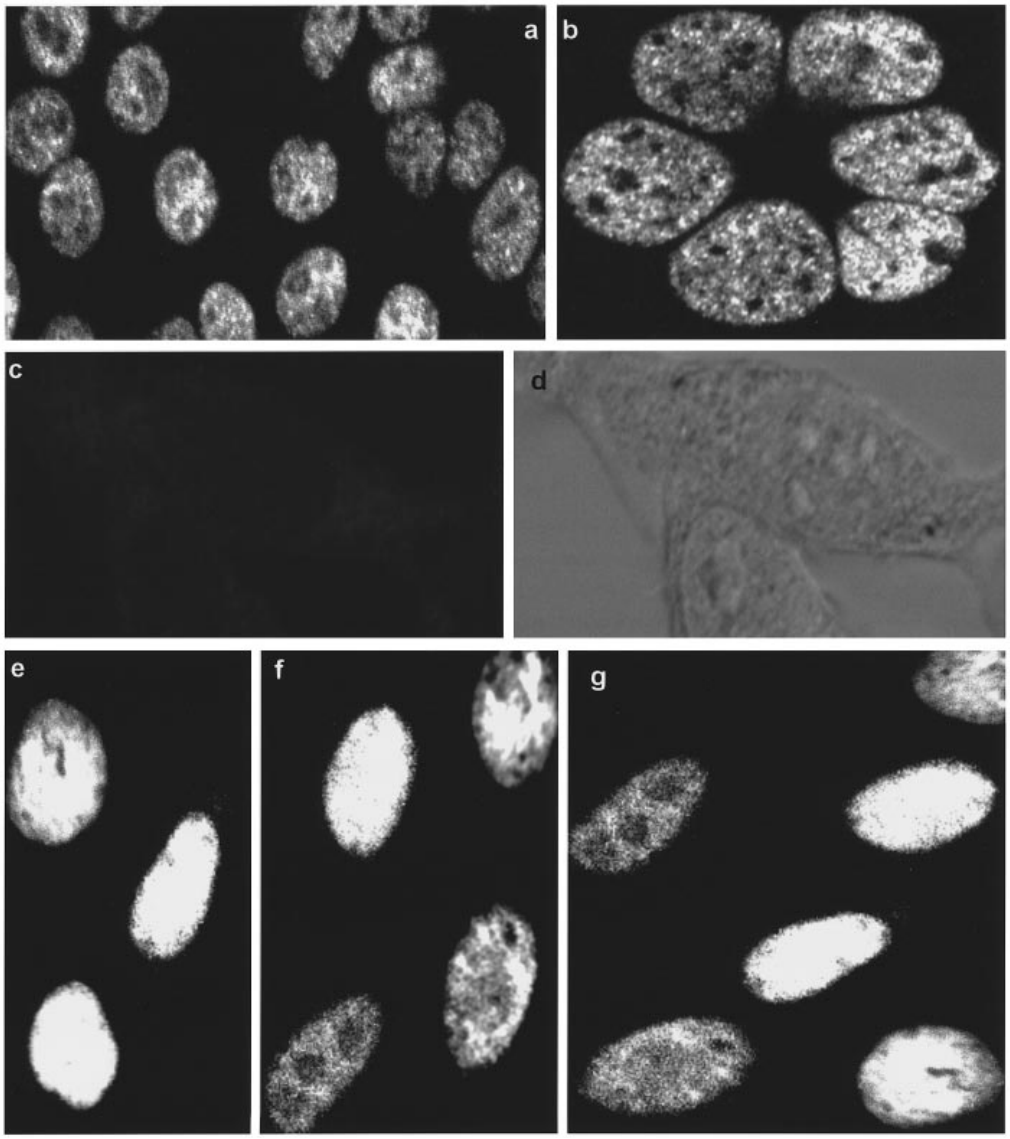



**Figure 5 cLSM analysis of transfected mouse LTA cells**

(**a**,**b**) Permanently transfected cell lines; (**c**, **d**, **e**, **f**, **g**) transiently transfected cells. (**a**) LTA cell line (G3) transfected with the human La gene. (**b**) LTA cell line transfected with the exon 1 La cDNA. (**c**, **d**) LTA cells transfected with pCI-neo lacking a La-specific insert. (**d**) Phase contrast with (**c**). LTA cells transfected with the: (**e**) exon 1 pCI-neo construct; (**f**) exon 1′ pCI-neo construct; (**g**) exon 1″ pCI-neo construct.

Up to now more than 30 La cDNAs have been described [18,19,27–30], but alternative La mRNA forms similar to the forms found in the La pseudogenes had never been described. However, in these studies La mRNAs similar to exon 1′ La mRNAs had also not been detected. Therefore we looked for further alternative La mRNA forms.

In spite of all efforts to find alternative forms as they were used as templates for the retropseudogenes, we were not able to detect such La mRNA isoforms in any of the analysed tissues including during embryonic development. Although these data do not exclude that such alternative forms can be made under certain physiological or pathophysiological conditions, it appears likely that they are only formed under rare conditions.

During these studies, in all tissues analysed, a further alternative La mRNA form was detected. This exon 1″ La mRNA form differed from the exon 1′ only with respect to a 4 nt insert and represented an alternatively spliced exon 1′ La mRNA form. In recent *in situ* hybridization studies we showed that both the exon 1 and the exon 1′ La mRNAs represent finally processed cytoplasmic mRNAs, which are expressed and regulated in parallel [20,21]. Both the exon 1 and the exon 1′ La mRNAs were shown to belong to the abundant class of mRNAs.

Due to the high homology of exons 1′ and 1″ it is now obvious that the exon 1′-specific probe used in these studies did not allow a differentiation between an exon 1′ and exon 1″ La mRNA. Therefore the *in situ* hybridization data obtained with the exon 1′-specific probe represented the sum of exons 1′ and 1″ La mRNA expression. As the exon 1′ probe localized the exon 1′ La mRNA to the cytoplasm, we concluded that the exon 1′ La mRNA is a finally processed cytoplasmic mRNA. In consequence the exon 1″ La mRNA should also represent a finally processed cytoplasmic mRNA. One prerequisite for this conclusion is that

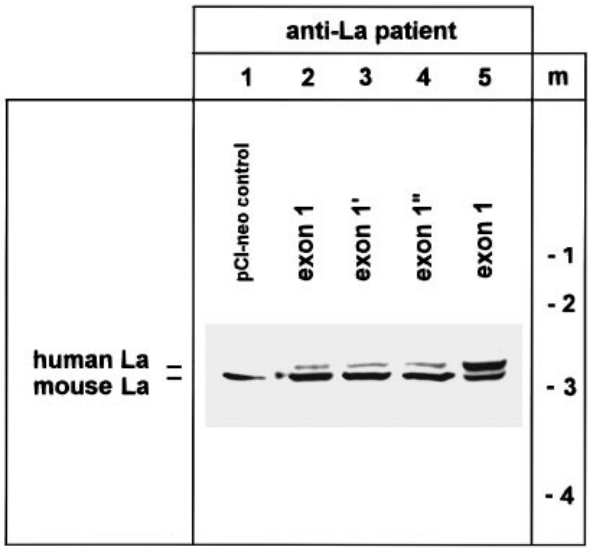



**Figure 6 SDS/PAGE/immunoblotting analysis of extracts obtained from transfected mouse cells**

The immunoblot was analysed with the patient's (Ma) anti-La antibody. Extract of mouse LTA cells permanently transfected with the exon 1 construct (lane 5). Extracts of LTA cells transfected with the pCI-neo control construct lacking a La-specific insert (lane 1), or the exon 1 (lane 2), the exon 1′ (lane 3) or the exon 1″ (lane 4) pCI-neo construct. m = marker lane, 1 = 97.5 kDa, 2 = 66 kDa, 3 = 45 kDa, 4 = 31 kDa.

the exon 1″ La mRNA is made at a relevant level compared with the exon 1′ La mRNA. As mentioned above, the exon 1″ La mRNA differs from the exon 1′ La mRNA only with respect to the 4 nt insert. Thus the primer pairs used allowed the amplification of the exon 1′ and 1″ fragments in parallel. Based on the intensities of the PCR fragments we estimated a ratio of exon 1″ to exon 1′ of between 1:10 and 1:1. Modifications of the PCR conditions did not alter the estimated ratios (results not shown). Moreover, similar ratios were also estimated in an independent laboratory for a series of embryonic tissues (Dr. Jil Byon, personal communication).

Taken together, our results allow the conclusion that the exon 1″ La mRNA form was not the result of a rare splice artifact but represents a finally processed, cytoplasmic, abundant, alternative La mRNA form.

Finally, the transfection studies showed that, in spite of their unusual 5′-terminus, both exon 1′ and 1″ La mRNAs represented translatable mRNAs. The translation efficiency appeared to be slightly less for the exon 1′ and 1″ La mRNAs if compared with the exon 1 La mRNA in all the five transfection assays performed. However, future studies using exon 1 and exon 1′ reporter gene fusion constructs are now required to finally compare the translational efficiencies of the different La mRNA forms.

This investigation was supported by grants from the Deutsche Forschungsgemeinschaft (Ba 1145/1-2,4-1). M. B. is the recipient of a professorship granted by the 'Hermann-and-Lilly-Schilling Stiftung' in the 'Stifterverband für die Deutsche Wissenschaft'.

## REFERENCES

1 Tan, E. M. (1989) Adv. Immunol. **44**, 93–151
2 Hendrick, J. P., Wolin, S., Rinke, J., Lerner, M. and Steitz, J. A. (1981) Mol. Cell. Biol. **12**, 1138–1149
3 Madore, S. J., Wieben, E. D. and Pederson, T. (1984) J. Biol. Chem. **259**, 1929–1933
4 Rinke, J. and Steitz, J. A. (1985) Nucleic Acids Res. **13**, 2617–2629
5 Stefano, J. E. (1984) Cell **36**, 145–154
6 Pruijn, G. J. M., Slobbe, R. L. and van Venrooij, W. J. (1990) Mol. Biol. Rep. **14**, 43–48
7 Wilusz, J., Kurilla, M. G. and Keene, J. D. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 5827–5831
8 Kurilla, M. G., Cabradilla, C. D., Holloway, B. P. and Keene, J. D. (1984) J. Virol. **50**, 773–778
9 Gottlieb, E. and Steitz, J. A. (1989) EMBO J. **8**, 841–850
10 Meerovitch, K., Svitkin, Y. V., Lee, H. S., Leibkowicz, F., Kenan, D., Chan, E. K. L., Agol, V. I., Keene, J. D. and Sonenberg, N. J. (1993) J. Virol. **67**, 3798–3807
11 Meerovitch, K., Pelletier, J. and Sonenberg, N. (1989) Genes Dev. **3**, 1026–1034
12 Bachmann, M., Pfeifer, K., Schröder, H. C. and Müller, W. E. G. (1990) Cell **60**, 85–93
13 Grölz, D., Bartsch, H., Tröster, H. and Bachmann, M. (1997) Gene, in the press
14 Tröster, H., Metzger, T. E., Semsei, I., Schwemmle, M., Winterpacht, A., Zabel, B. and Bachmann, M. (1994) J. Exp. Med. **180**, 2059–2067
15 Pruijn, G. J. M., Thijssen, J. P. H., Smith, P. R., Williams, D. G. and van Venrooij, W. J. (1995) Eur. J. Biochem. **232**, 611–619
16 Smith, P. R., Williams, D. G., Venables, P. J. W. and Maini, R. N. (1985) J. Immunol. Methods **77**, 63–69
17 Tröster, H., Bartsch, H., Klein, R. R., Metzger, T. E., Pollak, G., Semsei, I., Schwemmle, M., Pruijn, G. J. M., van Venrooij, W. J. and Bachmann, M. (1995) J. Autoimmunity **8**, 825–842
18 Chambers, J. C., Kenan, D., Martin, B. J. and Keene, J. D. (1988) J. Biol. Chem. **263**, 18043–18051
19 Pruijn, G. J. M., Slobbe, R. L. and van Venrooij, W. J. (1991) Nucleic Acids Res. **19**, 5173–5180
20 Hilker, M., Tröster, H., Grölz, D., Hake, U. and Bachmann, M. (1996) Cell Tissue Res. **284**, 383–389
21 Bachmann, M., Hilker, M., Grölz, D., Tellmann, G., Hake, U., Kater, L., deWilde, P. and Tröster, H. (1996) J. Autoimmunity, **9**, 757–766
22 Keech, C. L., Gordon, T., Reynolds, P. and McCluskey, J. (1993) J. Autoimmunity **6**, 543–555
23 Laemmli, U.K. (1970) Nature (London) **227**, 680–685
24 Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038
25 Bachmann, M., Tröster, H., Bartsch, H. and Grölz, D. (1996) J. Autoimmunity, **9**, 747–756
26 Semsei, I., Tröster, H., Schwemmle, M. and Bachmann, M. (1993) Gene **126**, 265–268
27 Chan, E. K. L., Sullivan, K. F. and Tan, E. M. (1989) Nucleic Acids Res. **17**, 2233–2244
28 Kohsaka, H., Yamamoto, K., Fujii, H., Miura, H. and Miyasaka, N. (1990) J. Clin. Invest. **6**, 2233–2244
29 Rauh, A. J. G., Hornig, H. and Lührmann, R. (1988) Eur. J. Immunol. **18**, 2049–2057
30 Sturgess, A. D., Peterson, M. G., McNeilage, L. J., Wittingham, S. and Coppel, R. L. (1988) J. Immunol. **140**, 3212–3218

Received 24 June 1996/12 November 1996; accepted 20 November 1996