3088–3094 *Nucleic Acids Research, 1997, Vol. 25, No. 15*

Ribonucleoprotein formation by the ORF1 protein of the non-LTR retrotransposon Tx1L in *Xenopus* oocytes

Geneviève Pont-Kingdon*, Enxi Chi+, Shawn Christensen and Dana Carroll

Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84132, USA

Received April 8, 1997; Revised and Accepted June 6, 1997

ABSTRACT

The Tx1L elements constitute a family of site-specific non-LTR retrotransposons found in the genome of the frog *Xenopus laevis*. The elements have two open reading frames (ORFs) with homology to proteins of retroviruses and other retroelements. This study demonstrates an expected activity of one of the element-encoded proteins. The RNA binding properties of ORF1p, the product of the first ORF of Tx1L, were examined after expression from RNA injected into *Xenopus* oocytes. Using sucrose gradient sedimentation and non-denaturing gel electrophoresis, we show that ORF1p associates with RNA in cytoplasmic ribonucleoprotein (RNP) particles. Discrete RNPs are formed with well-defined mobilities. The ORF1p RNPs are distinct from endogenous RNPs that contain stored oocyte mRNAs and two specific endogenous mRNAs do not become associated with ORF1p. ORF1p appears to be capable of associating with its own mRNA and with other injected RNAs, independent of specific recognition sequences. Although nuclear localization of ORF1p was anticipated, based both on the supposed mechanism of transposition and on the presence of a potential nuclear localization signal, no significant fraction of the protein was found in the oocyte nucleus. Nonetheless, the RNA binding capability of ORF1p is consistent with the proposed model for transposition of non-LTR retrotransposons.

INTRODUCTION

Transposable elements are found in the genomes of all organisms. These mobile DNAs can be classified according to their sequence organization and their modes of transposition. Two classes of elements, called retrotransposons, move via RNA intermediates that are reverse transcribed into DNA in the process of integrating new copies into the genome. One class closely resembles the retroviruses in having long terminal repeats (LTRs) and genes that encode homologs of the gag and pol proteins. They generate an extrachromosomal DNA copy of element RNA that is inserted into chromosomal sites through the integrase function of the *pol* gene.

Elements of the second class of retrotransposons (1,2) also carry sequences related to retroviral genes, but lack LTRs and apparently never produce a free DNA copy. Recent evidence suggests that these non-LTR retrotransposons (NLRs) use the 3′-hydroxyl at a single-strand nick in target DNA to prime reverse transcription, thereby generating a DNA copy that is linked to the target from its inception (3). The *pol*-related element-encoded protein apparently makes the nick that will prime synthesis (3,4). Yeast mitochondrial mobile group II introns move by a very similar mechanism (5).

Most NLR elements have two open reading frames (ORFs). The predicted ORF2 product (ORF2p) has a putative nucleic acid binding domain and sequences homologous to reverse transcriptase. The latter activity has been demonstrated in a number of cases (6, and reference therein). The ORF1 product is predicted to be an RNA binding structural protein, by analogy to retroviruses. In non-mammalian NLRs, ORF1 encodes one of two types of cysteine–histidine motifs: the majority are of the $Cx_2Cx_4Hx_4C$ type (CCHC box), analogous to the nucleocapsid motif found in retroviruses and LTR retrotransposons, while some site-specific NLRs contain a $Cx_2Cx_{12}Hx_4H$ motif, similar to that found in the transcription factor TFIIIA. Mammalian NLRs (also called LINEs) do not contain cysteine–histidine motifs in ORF1, but encode potential helical coiled coils that could mediate protein–protein interactions (7,8; Pont Kingdon *et al.*, in preparation). The mammalian ORF1 proteins have been found in a cytoplasmic macromolecular complex in association with element RNA (7,9,10) and reverse transcriptase activity (11). The ORF1 protein encoded by L1*Hs* is required for transposition in cell culture (12).

The Tx1L elements of *Xenopus laevis* belong to the family of site-specific NLRs (13). Their structure is diagramed in Figure 1A. There are ~100 such elements in the genome and all are inserted into another family of repetitive sequences, called Tx1D, that are probably cut-and-paste transposons (14). The two ORFs of Tx1L are related by a frameshift, but evidence suggests that, as in other NLRs (15–17), ORF2 is translated by ribosomal reinitiation, rather than by generation of a polyprotein, as seen in LTR elements. The predicted ORF1 product (ORF1p) has a single $Cx_2Cx_4Hx_4C$ motif. In addition, it carries a very basic motif, KKKRKFK, similar to the nuclear localization signals (NLSs) of SV40 large T antigen and the HIV-1 matrix polypeptide (18). Based on these sequence characteristics alone, ORF1p could

*To whom correspondence should be addressed at present address: Huntsman Cancer Institute, University of Utah, Building 533, Room 3260, Salt Lake City, UT 84112, USA. Tel: +1 801 585 5005; Fax: +1 801 585 3501; Email: genevieve.kingdon@genetics.utah.edu

+Present address: University of Texas Southwestern Medical Center. Dallas, TX 75235, USA

plausibly bind Tx1L transcripts and direct them to the nucleus for copying at the target.

In this paper we present the results of an investigation of the properties of the Tx1L ORF1 protein. An epitope-tagged version of ORF1p was expressed in *Xenopus* oocytes and found to bind to RNA. The formation of discrete ribonucleoprotein complexes (RNPs) is an expected characteristic of this protein and is consistent with its proposed role in element transposition.

MATERIALS AND METHODS

Constructs

To allow production of RNA for injection into oocytes, Tx1L ORF1 from the original genomic isolate, λB10 (13), was cloned into the vector pSP64T (19) and endowed with a myc epitope tag. The construct pS1-1mt (Fig. 1B) was obtained as follows. A chimeric primer was synthesized that contained, 5′→3′, a *Bgl*II site, an optimal translation initiation sequence (GCCGCCACCATG; 20) and 15 nt corresponding to the first five codons of ORF1. This was used in conjunction with a primer complementary to nt 1684–1660 of ORF1 for PCR with a Tx1L template. The resulting product was cleaved with *Bgl*II and *Xba*I and ligated to the appropriate *Xba*I–*Xmn*I fragment of Tx1L. This ligation reconstructed the complete ORF1, followed by the first 46 nt of ORF2. The ligated fragment was then cloned at the *Bgl*II site of pSP64T. The sequence of the new junction was confirmed by sequencing. A double-stranded oligonucleotide encoding the myc epitope EQKLISEEDL (21) was cloned at the unique *Hin*cII site of ORF1, 94 amino acids upstream of the ORF1 stop codon. The resulting modified ORF1 has a total length of 785 amino acids.

To obtain pS1-1mt-inv, pS1-1mt was digested with the restriction enzymes *Bal*I and *Eco*RV, which both generate blunt ends. The *Bal*I–*Eco*RV fragment was reintroduced at the same position and we screened for clones in which the fragment was inverted compared with its original orientation.

The construct pS2-5mt was obtained in several steps. First, two consecutive *Hin*dIII fragments that contain the complete ORF2 of Tx1L, the 3′-untranslated region (UTR) and 200 bp of the PTR-1 target site were subcloned at the *Hin*dIII site of pBluescript KS+ (Stratagene). Serial deletions from the 5′-end with exonuclease III followed by S1 nuclease brought the first ATG of ORF2 close to the 5′-end of the insert and created an *Nco*I site between the plasmid and the insert sequence. The *Nco*I–*Cla*I fragment of this clone was ligated with the *Cla*I–*Nco*I fragment of pKSMT6 (22) at the *Nco*I site; this linked ORF2 in-frame with the first five of the six myc tags contained in pKSMT6. The resulting fragment was cloned into the *Bgl*II site of pSP64T after S1 nuclease treatment of the *Bgl*II and *Cla*I ends.

To synthesize RNA *in vitro*, pS1-1mt, pS1-1mt-inv or pS2-5mt was first cleaved with *Sal*I, which cuts 3′ of the poly(A) tail present in the vector. The linear constructs were then transcribed with SP6 RNA polymerase (Gibco BRL) in the presence of a cap analog, following a modified protocol from the mCAP™ Capping Kit from Stratagene.

Oocyte injection

Xenopus laevis females used in these studies were obtained from Xenopus I (Ann Arbor, MI). Stage V and VI oocytes were prepared as described (23). A volume of 40 nl of the desired RNA solution was microinjected into the cytoplasm of each oocyte. Oocytes were incubated at 19°C in OR-2 (24) for 3 days.

RNP extraction and sucrose gradient fractionation

Healthy oocytes were homogenized in 10 µl/oocyte 20 mM Tris, pH 7.4, 100 mM KCl containing either 10 mM EDTA or 2 mM $MgCl_2$ as indicated. The extraction buffer contained a cocktail of proteinase inhibitors (0.2 mM PMSF, 0.5 µg/ml leupeptin and 0.7 µg/ml pepstatin) and, except in the experiment in which the extract was treated with RNase A, 20 U/ml RNasin (Promega). The homogenate was centrifuged for 20 min at 9000 *g* (10 000 r.p.m. in an Eppendorf bench microcentrifuge). Glycerol was added to the clear interstitial layer to a final concentration of 5% and it was frozen at –80°C.

A volume of 500 µl oocyte extract was loaded on top of a 10–50% linear sucrose gradient prepared with homogenization buffer containing EDTA or $MgCl_2$, as above. A linear gradient of 10 ml was established on top of a 1 ml cushion of 60% sucrose. Centrifugation was performed in a Beckman SW41 rotor for times and velocities that varied among experiments and are indicated in the figure legends. Eleven to 12 fractions of 1 ml were collected from the bottom of the tube using a peristaltic pump.

RNA and protein analysis

RNA from each sucrose gradient fraction was purified by extraction with phenol/chloroform/isoamyl alcohol (25/24/1 v/v), followed by precipitation from the aqueous phase with 2.5 vol ethanol in the presence of 0.3 M sodium acetate and 20 µg glycogen (Boehringer Mannheim). Pellets were resuspended in 1 µl/oocyte H_2O. The resulting RNA was analyzed by electrophoresis in 1% agarose gels containing 0.22 M formaldehyde (25). Usually ethidium bromide (EtBr) was added to the sample to 10 µg/ml before loading (26), allowing rapid visualization of rRNAs with UV light. Northern blots were obtained by capillary transfer of the denaturing agarose gels in 20× SSC to positively charged nylon membranes (Boehringer Mannheim). The ORF1, ORF2 and amylase probes were made from agarose gel-purified (27) DNA fragments obtained from restriction digests of the relevant clones. The fragments were: a 1.4 kb *Eco*RI–*Sac*I fragment of ORF1; a 2.5 kb *Xba*I fragment of ORF2; a 0.7 kb *Sty*I–*Cla*I fragment of amylase. The β-actin probe was prepared from the human β-actin cDNA control probe from Clontech. Radioactive probes were made by random primed synthesis (28) and hybridizations were usually performed overnight at 42°C in a buffer containing 20 mM PIPES, pH 6.5, 0.8 M NaCl, 50% formamide, 1% SDS and 100 µg/ml denatured salmon sperm DNA. In the case of the β-actin experiment hybridization was at 37°C.

For protein analysis SDS loading buffer (29) was added to aliquots of the gradient fractions to a final concentration of 1×; samples were boiled and directly loaded onto an SDS–8% polyacrylamide gel and run at 30 mA for 4.5 h. The proteins were then electrotransfered to positively charged nylon membranes in Tris–glycine transfer buffer (29) for 14 h at 20 V. The myc-tagged protein products of the ORFs were detected using an anti-myc antibody prepared from hybridoma cell line MYC1-9E10 (21) (American Type Culture Collection). Bound antibody was revealed with an HRP-conjugated goat anti-mouse IgG (Cappel). The antibody used to control the protein content of the nuclear extract was anti-MBP-4F (30). Polyclonal antibodies that reveal the major *Xenopus* oocyte mRNP proteins FRGY2a, FRGY2b,

p50 and p54 were obtained from Dr John Sommerville (31). The MBP-4F and mRNP proteins were visualized with a peroxidase-conjugated goat anti-rabbit IgG (Sigma). In all cases the chemiluminescence kit Renaissance from Dupont/NEN was used for detection.

Native gel analysis

RNPs were analyzed after migration in 1.3% agarose gels in 1× TBE buffer. Gels were usually run for 4 h at 50 V. Positions of ribosomal subunits or rRNAs were visualized by EtBr staining. The gels were transferred by capillary action in 20× SSC or electrotransferred to positively charged nylon membranes in the Tris–glycine transfer buffer used for protein analysis. The same membrane was successively probed with a DNA hybridization probe, as described for RNA detection, then with antibodies, as described for protein detection.

RESULTS

Expression of the Tx1L ORF1 protein (ORF1p) in *Xenopus* oocytes

Tx1L ORF1p was expressed in *Xenopus* oocytes following injection of RNA transcribed *in vitro* from plasmid pS1-1mt (Fig. 1B). To facilitate detection of the protein, we introduced a myc epitope tag near its C-terminus, far from any conserved sequences. An optimal translation initiation sequence was engineered around the initiator ATG (13). In addition, the pSP64T vector provides a 3′ poly(A) tail and 5′- and 3′-UTRs from *Xenopus* β-globin cDNA that contribute to translational efficiency (32). Transcription of linearized pS1-1mt with SP6 RNA polymerase in the presence of a cap analog led to synthesis of a functional mRNA.

Injection of this RNA into the oocyte cytoplasm caused production of ORF1p that was readily detectable in a Western blot with an anti-myc antibody (Fig. 2). A single band was observed with a mobility corresponding to 105 kDa, which is somewhat larger than the expected 83 kDa. In contrast to observations for mammalian NLR ORF1 products (7,33,34), no evidence for ORFlp phosphorylation was obtained by Western analysis of phosphatase-treated protein (data not shown). Like other NLR ORF1 proteins, but unlike those of LTR elements, no evidence of proteolytic processing was observed, although small amounts of probable degradation products were occasionally seen (for example in Fig. 3). Other possible post-translational modifications were not tested.

After 3 days incubation we estimate that ~100 ng ORF1p are produced in a single oocyte (judged by comparison of the intensity of the Western signal with that of the same protein expressed in *Escherichia coli* and partially purified; data not shown). Despite the presence of a putative nuclear localization signal in the ORF1 sequence, no ORF1p was detectable in nuclei of manually dissected oocytes (Fig. 2, lanes 2 and 3). An antibody directed to the nuclear protein MBP-4F (30) revealed, on the same blot, the expected band in the nuclear lane, confirming the success of the dissection (data not shown).

Association of ORFlp in a macromolecular complex

Extracts of oocytes expressing ORF1p were fractionated by sucrose gradient sedimentation. Western blotting of individual fractions showed that the protein migrated in a discrete complex that was somewhat slower than monomeric 80S ribosomes (Fig. 3, top). Fractions containing ribosomes or ribosomal subunits were determined by EtBr staining of denaturing agarose gels of RNA from each fraction. The sedimentation behavior of ORF1p was not affected by the presence of EDTA at a concentration sufficient to dissociate ribosomal subunits (Fig. 3, bottom). The apparent sedimentation coefficient of ~60 S is much greater than that predicted for a monomeric protein of ~100 kDa. For comparison,

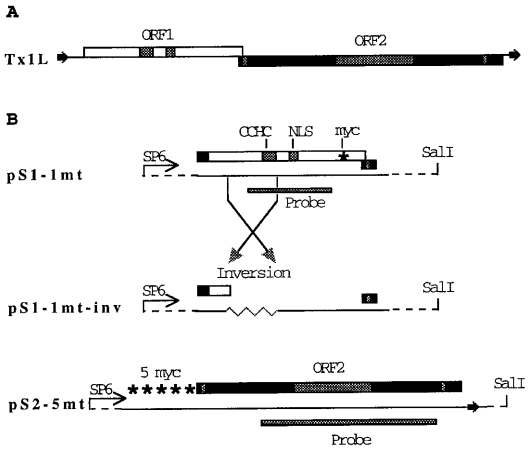

Figure 1. (**A**) Structure of Tx1L. Black arrows represent the 23 bp duplication that flanks the element (13). Within the ORFs shaded boxes indicate, from left to right, the CCHC motif and potential nuclear localization signal of ORF1 and the endonuclease homology, the reverse transcriptase homology and the putative metal binding site of ORF2. (**B**) Diagram of the constructs pS1-1mt, pS1-1mt-inv and pS2-5mt. Sequences encoding ORF1 (pS1 constructs) or ORF2 (pS2–5mt) were cloned into the pSP64T vector, where they are surrounded by the 5′- and 3′-UTRs of β-globin (dotted lines). Boxes placed above the constructs indicate their predicted coding capability. The black box at the 5′-end of ORF1 indicates the introduced Kozak consensus sequence. Stars indicate the positions of the myc tags. The zigzag line in pS1-1mt-inv indicates the pS1-1mt DNA segment that was inverted. After digestion of the constructs with *Sal*I, *in vitro* transcription was performed using SP6 RNA polymerase. Segments of the constructs used as probes in native gel electrophoresis analysis are underlined with thin shaded boxes.

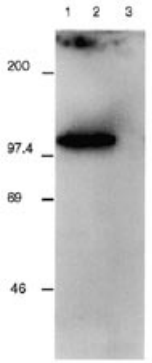

Figure 2. Expression of Tx1L ORF1p in *Xenopus* oocytes. An aliquot of 40 ng RNA transcribed *in vitro* from pS1-1mt was injected into the cytoplasm of stage VI oocytes. After 72 h incubation at 19°C, cytoplasm and nucleus were separated by manual dissection. Each lane contains protein from four oocytes separated by electrophoresis in an 8% SDS–polyacrylamide gel. After electrotransfer, the membrane was probed with an anti-myc antibody. Lane 1, total oocytes; lane 2, cytoplasm; lane 3, nuclei. Molecular weight markers are indicated in kDa.

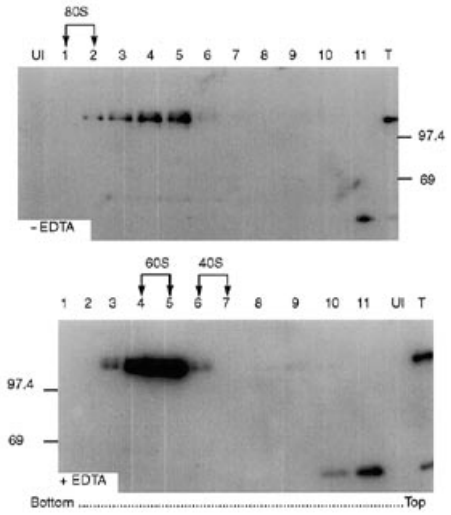

Figure 3. Sedimentation of ORF1p in the presence or absence of EDTA. Injections were as described in the legend to Figure 2. Oocyte extracts were prepared in buffers with Mg^{2+} (top) or with EDTA (bottom). Protein extracts from 50 oocytes were loaded on top of a linear 10–50% sucrose gradient in a SW41 rotor (Beckman). Centrifugation was for 20 h at 26 000 r.p.m. Fractions of 1 ml were collected from the bottom of the tube. Around 15% of each fraction was denatured by addition of SDS and boiling and analyzed by 8% SDS–PAGE. The Western blot was developed with an anti-myc antibody. UI indicates lanes where protein extracts from uninjected oocytes were loaded. T indicates lanes where total, unfractionated protein extracts from injected oocytes were loaded. Molecular weights are indicated in kDa.

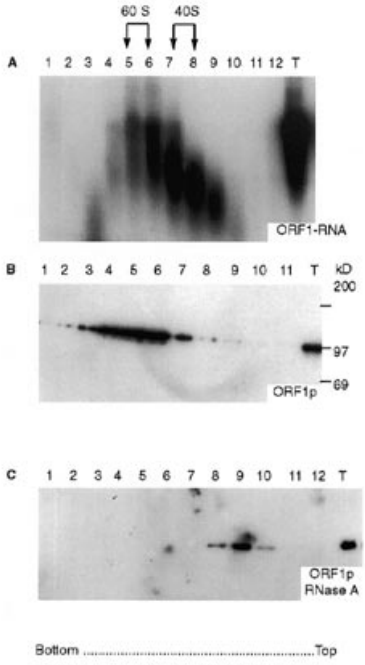

Figure 4. Sedimentation analysis of ORF1p RNPs. Oocyte injections were as described in Figure 2. Extracts from half the oocytes were prepared in a buffer containing EDTA and RNase inhibitor (**A** and **B**). No RNase inhibitor was added to the buffer used to extract RNP from the other half (**C**). RNase A was added to this second half and incubation was for 20 min at room temperature. Both extracts were loaded on top of sucrose gradients and were centrifuged for 5.5 h at 41 000 r.p.m. RNA was phenol extracted from fractions of the extract not treated with RNase A and analyzed by Northern blotting with an ORF1 DNA probe (A). Proteins from 15% of each fraction were analyzed by Western blotting as in Figures 2 and 3 (B and C). T indicates lanes where total extracts from whole injected oocytes were loaded. Positions of the 60S and 40S rRNA were visualized by EtBr staining of the RNA gel and are indicated.

the 60S ribosomal subunit has a mass of ~3000 kDa. The data also show that ORF1p is not associated with ribosomes or polysomes.

Association of ORF1p with RNA

Two analyses were performed to determine if ORF1p was incorporated into RNP particles. First, the location of the ORF1 RNA was determined in sucrose gradient fractions from injected oocytes. As shown in the Northern blot in Figure 4A, ORF1 RNA is broadly distributed in the gradient, in a pattern that has considerable overlap with that of ORF1p in the same experiment (Fig. 4B). There were problems with degradation of the RNA from these native gradient fractions, but repeated experiments confirmed the basic observations. The partial overlap suggests that ORF1 RNA and protein may be in the same complex.

Second, extracts from injected oocytes were treated with RNase A before sedimentation (Fig. 4C). The ORF1 RNA was no longer detectable (not shown) and ORF1p was shifted to much more slowly sedimenting material. ORF1p still migrated more rapidly than expected for free monomeric protein, suggesting residual association with RNase-resistant cores or with proteins in the extracts. The RNase-induced shift clearly indicates incorporation of ORF1p into RNA-containing particles.

Analysis of ORF1p RNPs by native gel electrophoresis

Because the resolving power of sucrose gradient sedimentation is limited, we turned to non-denaturing gel electrophoresis to further characterize the ORF1p RNPs. Both composite (2.25% acrylamide, 0.5% agarose) (35) and 1.3% agarose gels proved effective. After transferring the fractionated species to a membrane, it was important to first hybridize with a nucleic acid probe to detect the desired RNA species, then to detect proteins with their respective antibodies.

As shown in Figure 5A (lane 3), ORF1 RNA migrates in such gels as a discrete species that is dramatically retarded with respect to deproteinized RNA from the same sample (lane 1). The ORF1p signal (Fig. 5B, lane 3) overlaps that of the RNA completely and is somewhat broader.

To determine the mobility of ORF1 RNA in the absence of ORF1 protein, we constructed pS1-1mt-inv (Fig. 1B). Transcription of this plasmid produces a mRNA which has the same size and many of the same sequences as intact ORF1 mRNA. The inversion made in the ORF1 coding sequence introduces a stop codon upstream of the CCHC box and the myc tag, at codon 77; it seems unlikely that the truncated protein would have retained any binding activity. In oocytes injected with RNA made from pS1-1mt-inv no ORF1p is made (Fig. 5B, lane 2). The RNA is apparently associated with endogenous oocyte proteins, since its mobility is shifted (Fig. 5A, lane 2) relative to free RNA (lane 1), but its migration is clearly different from that seen in oocytes expressing ORF1p (lane 3). The supershift induced by ORF1p and its co-migration with ORF1 RNA strongly suggest that the two are associated in discrete RNPs.

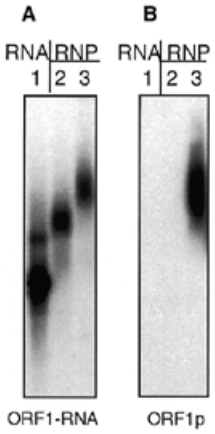

Figure 5. Native gel electrophoresis of ORF1p RNPs. Oocyte injections of pS1-1mt and pS1-1mt-inv RNA were as described in Figure 2. RNA from pS1-1mt-injected oocytes (lane 1) was extracted and one oocyte equivalent was loaded on a 1.3% agarose gel. RNPs from oocytes injected with pS1-1mt (lane 3) or pS1-1mt-inv (lane 2) were extracted and one oocyte equivalent was loaded on the same gel. After migration, the material was transferred to a nylon membrane without prior denaturation. (**A**) The membrane was probed with a labeled ORF1-encoding DNA fragment to reveal the injected RNA. (**B**) The membrane was then probed with a myc antibody to detect epitope-tagged ORF1p.

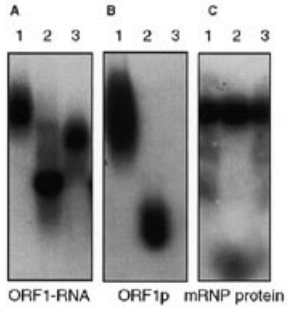

Figure 6. Comparison of ORF1 RNP and endogenous mRNPs in native gel electrophoresis. An agarose gel was loaded with RNP extracted from oocytes injected with pS1-1mt (lanes 1 and 2) or pS1-1mt-inv (lane 3) RNA. In lane 2, SDS was added to a final concentration of 0.1% and the sample was boiled for 10 min prior to loading. (**A**) The nylon membrane obtained after transfer of the gel was hybridized with a radioactive DNA probe that reveals the injected ORF1 RNA. (**B**) The probe used was an anti-myc antibody, which detects myc-tagged ORF1p. (**C**) The same blot was probed with a mix of antisera against the major mRNP proteins FRGY2a and FRGY2b.

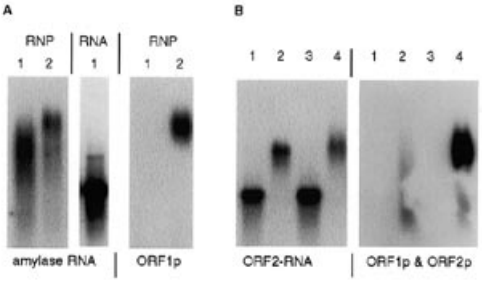

Figure 7. (**A**) Binding of ORF1p to injected amylase mRNA. Lane 1, only amylase mRNA was injected into oocytes. Lane 2, amylase mRNA and pS1-1mt RNA were co-injected. The nylon membrane resulting from transfer of a native agarose gel loaded with RNP (RNP lanes) or RNA (RNA lane) extracted from injected oocytes was successively probed with an amylase DNA probe (amylase RNA) and the anti-myc antibody (ORF1p). (**B**) Binding of ORF1p to ORF2 RNA. Lanes 1 and 2, ORF2 RNA was injected into oocytes; lanes 3 and 4, ORF2 and pS1-1mt RNA were co-injected. RNA (lanes 1 and 3) and RNP (lanes 2 and 4) extracted from oocytes were analyzed by hybridization with a DNA probe containing a fragment of ORF2 of Tx1L (left) and with the anti-myc antibody (right).

Comparison of ORFlp RNPs with endogenous mRNPs

Previous work has characterized an abundant class of mRNPs in *Xenopus* oocytes (36,37). These particles sediment in the range 40–80 S and contain mRNAs that are masked from the oocyte translational machinery (38). To test if ORFlp was found associated with these mRNPs, we probed our blots with a mixture of antibodies against some of the major mRNP proteins: FRG Y2a (pp60), FRG Y2b (pp56), mRNP3 (50–54 kD) (39,40 and references therein).

There was some, but not complete, overlap of the distributions of ORF1p and these mRNP proteins on sucrose gradient sedimentation (data not shown). Electrophoretic analysis, however, showed clear distinctions between the RNPs.

As seen previously (Fig. 5), ORF1 RNA associates with oocyte proteins and is supershifted by expression of ORF1p. These complexes are dissociated with 0.1% SDS (Fig. 6A and B, lanes 2); both the ORF1 RNA and protein migrate faster under these conditions and in non-overlapping positions. In contrast, migration of the endogenous mRNPs is not affected either by expression of ORF1p (Fig. 6C, lane 3) or by 0.1% SDS (Fig. 6C, lane 2). The mRNPs run as a discrete complex, but this complex does not appear to bind either injected ORF1 RNA or expressed ORF1 protein.

Association of ORFlp with other injected RNAs

In order to test the specificity of ORFlp binding, we analyzed the electrophoretic migration pattern of two other RNAs isolated from oocytes that were or were not expressing ORFlp. Human α-amylase mRNA synthesized *in vitro* (41) and injected into oocytes behaved very similarly to ORF1 RNA. Like pS1-1mt-inv RNA, its migration was slowed by association with oocyte proteins (Fig. 7A, lanes 1) and it was further retarded by association with ORF1p (lanes 2). The amylase mRNA and ORF1p were shown to co-migrate.

The same behavior was observed for an injected mRNA encoding the ORF2 protein of Tx1L (Fig. 7B). This case was somewhat more complicated because the ORF2p was also myc tagged (five copies of the epitope at its N-terminus) in the expression construct pS2-5mt (Fig. 1B). Injection of the ORF2 mRNA alone into oocytes led to its association with endogenous proteins (Fig. 7B, lanes 1 and 2). As seen in the Western blot of this gel (right panel), the myc-tagged ORF2p migrated as a broad smear of which only the tail overlapped the position of ORF2 RNA. Thus little if any ORF2p was associated with the RNA that encodes it. When ORF1p was co-expressed there was a modest shift in migration of the ORF2 RNP (lane 4) and its position corresponded to that of ORF1p. The reason for the smaller shift in this case may be that ORF2 RNA is larger (4.8 kb) than those

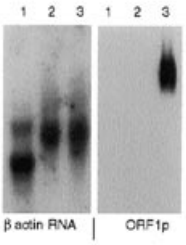

Figure 8. Lack of association of β-actin mRNA with ORF1p. Lanes 1 and 2, extracts were from uninjected oocytes (RNA in lane 1 and RNP in lane 2). An RNP extract from oocytes injected with pS1-1mt RNA was loaded in lane 3. The left panel shows the result of hybridization of the blot with a human β-actin probe while the right panel shows the result of developing the blot with the anti-myc antibody.

of ORF1 (2.8 kb) or amylase (1.7 kb) and its complexes run in a less well-resolved region of the gel.

ORF1p does not associate with endogenous RNA

The above results indicate that ORF1p associates with any injected RNA molecule, independent of its sequence, and possibly in co-association with endogenous oocyte proteins. To test whether this promiscuity extended to RNAs present naturally in the oocytes, we examined the effect of ORF1p expression on the electrophoretic mobility of endogenous RNPs. Both β-actin and Vg1 (42) were tested; the results for β-actin are shown in Figure 8. In an extract from uninjected oocytes, β-actin mRNA is apparently associated with proteins, since its mobility (lane 2) is retarded relative to free RNA (lane 1). No change in mobility was observed when ORF1p was expressed in oocytes for 3 days and there was no overlap in the migration of ORF1p and β-actin mRNA (lanes 3). Comparable results were obtained for Vg1 (not shown). Evidently ORF1p does not displace or add to the proteins of endogenous RNPs. The mobilities of β-actin and Vg1 RNPs were also distinct from that of the masking proteins examined earlier (Fig. 6).

DISCUSSION

The first ORF of NLR elements encodes a protein that is believed to play a largely structural role: binding to element RNA, perhaps directing it to the cell nucleus and possibly enforcing a conformation that is favorable for reverse transcription. An obvious requirement for such a function is the ability to bind element RNA. Two types of motif with potential nucleic acid binding properties are found in different classes of these ORFs. Mammalian LINEs contain sequences that can potentially form a helical coiled coil, which is a feature of some DNA binding proteins (43; Pont Kingdon *et al.*, in preparation). Cytoplasmic RNPs containing the human L1 RNA and L1 ORF1p have been described (7,9,10). Non-mammalian NLRs, including *Xenopus* Tx1L, have a zinc finger motif in ORF1. This motif is analogous to the CCHC sequences found in the nucleocapsid segment of the gag proteins of retroviruses. In this study we expressed ORF1p of Tx1L in *Xenopus* oocytes and employed several techniques to show that it associates with RNA to form discrete RNPs.

ORF1p RNPs

The combination of sucrose gradient sedimentation and non-denaturing gel electrophoresis shows that ORF1p expressed in oocytes is associated with other components to form discrete, high molecular weight complexes. That there is RNA in these complexes was demonstrated in sedimentation experiments by showing that they are largely destroyed by RNase treatment. The electrophoretic analysis showed co-migration of ORF1p with various injected RNAs at mobilities that were different from those observed in the absence of ORF1p. Thus, ORF1p is assembled into RNP complexes. These RNPs are not sensitive to treatment with EDTA, so they are not dependent on ribosome structure. They are disrupted by 0.1% SDS, which distinguishes them from endogenous, masked oocyte mRNPs.

The composition of the ORF1p RNPs is not known. Because injected RNAs evidently associate with endogenous oocyte proteins when ORF1p is not present, it is possible that the ORF1p complexes contain both the expressed and endogenous proteins in some ratio. The sedimentation behavior of the ORF1p RNPs varied slightly among experiments. This may reflect differences in levels of protein expression and resulting variations in protein:RNA ratio. In a typical experiment, 40 ng ORF1 mRNA (8×10^{11} molecules) were injected into a single oocyte and led to the accumulation, after 3 days incubation, of 100 ng (3×10^{13} molecules) ORF1 protein. The resulting ratio is 40 molecules of protein for each molecule of RNA, or one protein for each 70 bases of RNA. This may not be sufficient to saturate the RNA with ORF1p and the observed sedimentation behavior would reflect different levels of ORF1p bound.

The ORF1p RNPs described here are less polydisperse and slower sedimenting than the RNPs described for other NLR elements (7,9,11). Those studies did not employ non-denaturing gel electrophoresis, but that technique may prove very useful in further characterization of such RNPs. We have demonstrated reliable sequential detection of RNA and proteins on a single blot from these gels.

Specificity of ORF1p binding

It might be expected *a priori* that ORF1p would have both sequence-specific and non-specific modes of binding to RNA. The first would underlie recognition of Tx1L element RNA for packaging and the second would permit association with the entire length of the RNA. In our experiments we found that ORF1p could associate with the three injected RNAs that were tested: ORF1 mRNA, ORF2 mRNA and amylase mRNA. It is clear that non-specific binding occurs, but our observations do not rule out preferences for particular sequences. Although the ORF1 and ORF2 messages were both derived from Tx1L, it is arguable that none of the RNAs carried the necessary recognition sequences, since the 555 nt 5′-UTR of the element was missing in all cases. Non-specific RNA binding properties have been demonstrated *in vitro* for ORF1p of the mouse element L1*Md*, which does not carry a CCHC motif (44).

While ORF1p exhibited non-specific RNA binding capability in oocytes, it did not associate with the endogenous mRNAs or mRNPs tested. Neither β-actin nor Vg1 mRNP was affected by ORF1p expression and the protein was apparently not incorporated into RNPs that include masked messages. However, there is reason to think that some endogenous RNPs bind ORF1p, because the distributions of the protein and injected RNA do not

completely overlap, either in the sedimentation (Fig. 4) or electrophoretic (Figs 5 and 7) analyses. We cannot predict what would happen if transcription of an oocyte RNA and translation of ORF1p occurred simultaneously. In such a situation there would be competition for assembly of various types of RNP and ORF1p would not have to disrupt pre-assembled structures.

Cytoplasmic versus nuclear localization of ORF1p

The current model for NLR transposition (1) predicts that full-length element RNA must associate with chromosomal DNA to permit reverse transcription at the target site. This might require nuclear disruption at mitosis in dividing cells, as is seen for many retroviruses (45). Alternatively, element RNA may enter the intact nucleus. Cellular nuclear RNAs rely on association with proteins to provide nuclear localization signals and the same is true of HIV-1, which is among the minority of retroviruses that propagate in quiescent cells (46,47).

ORF1 of Tx1L contains one plausible nuclear localization signal. In a direct experiment, however, we obtained no evidence for nuclear accumulation of ORF1p. This situation is similar to mammalian LINEs, where element RNA and ORF1 protein have been found in cytoplasmic fractions (7,11,34), and to the I factor of *Drosophila melanogaster*, where ORF1 has been detected only in the cytoplasm of germline cells, where transposition occurs (A.Bucheton, personal communication). Conceivably, a necessary component of the complete RNP complex was missing in our experiments; the full-length element RNA and the ORF2 protein are candidates. Alternatively, it may be that only a small fraction of the RNP translocates to the nucleus and that this was below our level of detection. ORF1 of NLR elements other than Tx1L do not encode a nuclear localization signal and the mechanism by which NLR RNA gains access to genomic DNA is still an open question.

RNPs are believed to be an important intermediate in the retrotransposition cycle of NLRs. Their structures and their roles are expected to differ from those of retroviruses and retrovirus-like elements because of the apparent differences in transposition mechanism. Expression of ORF1p and other Tx1L products in *Xenopus* oocytes offers the potential for detailed functional characterization of the various components.

ACKNOWLEDGEMENTS

We are grateful to Jon Trautman for oocyte preparation and help with injection experiments. We thank Drs B.Bass and J.Sommerville for providing us with antibodies and Drs S.Martin and A.Bucheton for allowing us to cite results prior to publication. We also thank members, past and present, of the Carroll laboratory for discussions. This work was supported by a grant from the American Cancer Society (NP-803C).

REFERENCES

1 Eickbush,T.H. (1992) *New Biologist*, **4**, 430–440.
2 Martin,S. (1991) *Curr. Opin. Genet. Dev.*, **1**, 505–508.
3 Luan,D., Korman,M.H., Jakubczak,J.L. and Eickbush,T.H. (1993) *Cell*, **72**, 595–605.
4 Feng,G., Moran,J.V., Kazazian,H.H. and Boeke,J.D. (1996) *Cell*, **87**, 905–916.
5 Zimmerly,S., Guo,H., Perlman,P.S. and Lambowitz,A.M. (1995) *Cell*, **82**, 545–554.
6 Mathias,S.L., Scott,A.F., Kazazian,H.H.J., Boeke,J.D. and Gabriel,A. (1991) *Science*, **254**, 1808–1810.
7 Hohjoh,H. and Singer,M. F. (1996) *EMBO J.*, **15**, 630–639.
8 Holmes,S.E., Singer,M.F. and Swergold,G.D. (1992) *J. Biol. Chem.*, **267**, 19765–19768.
9 Martin,S.L. (1991) *Mol. Cell. Biol.*, **11**, 4804–4807.
10 Martin,S.L. and Branciforte,D. (1993) *Mol. Cell. Biol.*, **13**, 5383–5392.
11 Deragon,J.M., Sinnett,D. and Labuda,D. (1990) *EMBO J.*, **9**, 3363–3368.
12 Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H. (1996) *Cell*, **87**, 917–927.
13 Garrett,J.E., Knutzon,D.S. and Carroll,D. (1989) *Mol. Cell. Biol.*, **9**, 3018–3027.
14 Garrett,J.E. and Carroll,D. (1986) *Mol. Cell. Biol.*, **6**, 933–941.
15 Ilves,H., Kahre,O. and Speek,M. (1992) *Mol. Cell Biol.*, **12**, 4242–4248.
16 McMillan,J.P. and Singer,M.F. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11533–11537.
17 Bouhidel,K., Terzian,C. and Pinon,H. (1994) *Nucleic Acids Res.*, **22**, 2370–2374.
18 Myers,G., Berzofsky,J.A., Pavlakis,G.N., Korber,B. and Smith,R.F. (eds) (1992) *Human Retroviruses and AIDS 1992: A Compilation and Analysis of Nucleic Acid and Amino Acids Sequences*. Los Alamos National Laboratory, Los Alamos, NM.
19 Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
20 Kozak,M. (1991) *J. Biol. Chem.*, **266**, 19867–19870.
21 Evan,G.I., Lewis,G.K., Ramsay,G. and Bishop,J.M. (1985) *Mol. Cell. Biol.*, **5**, 3610–3616.
22 Roth,M.B., Zahler,A.M. and Stolk,J.A. (1991) *J. Cell Biol.*, **115**, 587–596.
23 Maryon,E. and Carroll,D. (1989) *Mol. Cell. Biol.*, **9**, 4862–4871.
24 Wallace,R.A. (1973) *J. Exp. Zool.*, **184**, 321–324.
25 Tsang,S.S., Yin,X., Guzzo-Arkuran,C., Jones,V.S. and Davidson,A.J. (1993) *BioTechniques*, **14**, 380–381.
26 Ogretmen,B., Ratajczark,H., Kats,A. and Stark,B.C. (1993) *BioTechniques*, **14**, 932–933.
27 Lu,Z., Templer,M. and Nielsen,B.L. (1994) *BioTechniques*, **16**, 400–402.
28 Feinberg,P.A. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
29 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1987) In Janssen,K. (ed.), *Current Protocols in Molecular Biology*. John Wiley & Sons, New York, NY.
30 Bass,B., Hurst,S.R. and Singer,J.D. (1994) *Curr. Biol.*, **4**, 301–314.
31 Braddock,M., Muckenthaler,M., White,M.R.H., Thorburn,A.M., Sommerville,J., Kingsman,A.J. and Kingsman,S.M. (1994) *Nucleic Acids Res.*, **22**, 5255–5264.
32 Krieg,P., Strachan,R., Wallis,E., Tabe,L. and Colman,A. (1984) *J. Mol. Biol.*, **180**, 615–643.
33 Kolosha,V.L. and Martin,S.L. (1995) *J. Biol. Chem.*, **270**, 2868–2873.
34 Branciforte,D. and Martin,S.L. (1994) *Mol. Cell. Biol.*, **14**, 2584–2592.
35 Dahlberg,A.E. and Grabowski,P.J. (1990) *Gel Electrophoresis of Nucleic Acids*, 2nd Edn. IRL Press, Oxford, UK, pp. 101–122.
36 Darnbrough,C.H. and Ford,P.J. (1981) *Eur. J. Biochem.*, **113**, 415–424.
37 Dearsly,A.L., Johnson,R.M., Barret,P. and Sommerville,J. (1985) *Eur. J. Biochem.*, **150**, 95–103.
38 Bouvet,P. and Wolffe,A.P. (1994) *Cell*, **77**, 931–941.
39 Sommerville,J. and Ladomery,M. (1996) *FASEB J.*, **10**, 435–443.
40 Deschamps,S., Viel,A., Garrigos,M., Denis,H. and Le Maire,M. (1992) *J. Biol. Chem.*, **267**, 13799–13802.
41 Urnes,M.S. and Carroll,D. (1990) *Gene*, **95**, 267–274.
42 Weeks,D.L. and Melton,D.A. (1987) *Cell*, **51**, 861–867.
43 Alber,T. (1992) *Curr. Opin. Genet. Dev.*, **2**, 205–210.
44 Kolosha,V.O. and Martin,S.L. (1997) *Proc. Natl. Acad. Sci. USA*, in press.
45 Roe,T., Reynolds,T., Yu,G. and Brown,P. (1993) *EMBO J.*, **12**, 2099–2108.
46 von Schwedler,U., Kornbluth,R.S. and Trono,D. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 6992–6996.
47 Bukrinsky,M.I., Haggerty,S., Dempsey,M.P., Sharova,N., Adzhubei,A., Spitz,L., Lewis,P., Goldfarb,D., Emerman,M. and Stevenson,M. (1993) *Nature*, **365**, 666–669.