

# Localization of the Nucleolar Protein NO38 in Amphibian Oocytes

Brenda A. Peculis and Joseph G. Gall

Department of Embryology, Carnegie Institution, Baltimore, Maryland 21210; and Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218

**Abstract.** To examine the role of primary amino acid sequence in the localization of proteins within the nucleus, we studied the nucleolar protein NO38 of amphibian oocytes. We synthesized NO38 transcripts in vitro, injected them into the oocyte cytoplasm, and followed the distribution of the translation products. The injected RNA contained a short sequence encoding an epitope derived from the human *c-myc* protein. We used an mAb against this epitope to detect translation products from injected RNAs by Western blots and by immunofluorescent staining of cytological preparations.

When full-length transcripts of NO38 were injected into oocytes, the translation products accumulated efficiently in the germinal vesicle, and a major fraction was localized in the multiple nucleoli. To identify protein domains involved in this nucleolus-specific ac-

cumulation, we prepared a series of carboxy-terminal deletions of the cDNA. Oocytes injected with RNA encoding truncated forms of NO38 were examined for altered patterns of protein accumulation. We defined a domain of about 24 amino acids near the carboxy terminus that was essential for nucleolar localization of NO38. This domain is separated by more than 70 amino acids from two putative nuclear localization signals near the middle of the molecule. Hybrid constructs were made which encoded part of *Escherichia coli*  $\beta$ -galactosidase or pyruvate kinase fused to a long segment of NO38 containing the essential domain. Injection of RNA from these constructs showed that the essential domain was not sufficient to target the hybrid proteins to the nucleolus. We suggest that nucleolar accumulation of NO38 requires more than a single linear domain.

FOR years the specific accumulation of proteins within subcellular compartments has been examined at the molecular level (reviewed in Watson, 1984). It is well accepted that many proteins can cross membranes and accumulate against gradients to concentrate in the nucleus, mitochondria, ER, or Golgi complex. In the case of the nucleus, early studies showed that the nuclear envelope provides an effective barrier against free diffusion of large macromolecules (Paine and Feldherr, 1972; Paine et al., 1975). Nevertheless, many nuclear proteins were known to exceed the size defined by these studies, and it seemed unlikely that they would all gain access to the nucleus during the short period when the envelope breaks down at mitosis. Nucleophilic, or karyophilic, proteins were described as those nuclear proteins that accumulate rapidly in the germinal vesicle (GV)<sup>1</sup> when injected into the cytoplasm of amphibian oocytes

(Bonner, 1975a,b; De Robertis et al., 1978; Dabauvalle and Franke, 1982). Subsequent studies, particularly on nucleoplasmin (Dingwall et al., 1982, 1988; Bürglin and De Robertis, 1987; Robbins et al., 1991) and the SV-40 large T antigen (Kalderon et al., 1984a,b), defined short runs of amino acids that are essential and sufficient for nuclear accumulation. The SV-40 sequence, Pro Lys Lys Lys Arg Lys Val, is typical of such nuclear targeting or nuclear localization signals (NLS), most of which contain a short run of basic amino acids. Nuclear proteins containing NLS's have been found in many organisms, including yeast, *Xenopus laevis*, and mammals (reviewed by Dingwall and Laskey, 1986; Silver and Goodson, 1989).

Once inside the nucleus, some proteins may remain free in the nucleoplasm, but others become associated with specific subnuclear structures, such as the nuclear envelope, chromosomes, and nucleoli. Several investigators have examined the amino acid sequences necessary for nucleolar localization of the heat-shock protein hsp70 (Munro and Pelham, 1984; Milarski and Morimoto, 1989; Dang and Lee, 1989). Siomi et al. (1988) defined a highly basic sequence of 19 amino acids in the Rex protein of HTLV-I that is essential for nucleolar localization and is also sufficient to target other proteins to the nucleolus. A somewhat shorter, but

B. Peculis' present address is Department of Molecular Biophysics and Biochemistry, Yale University Medical School, New Haven, CT 06510.

This work is in partial fulfillment of the requirements for a doctoral degree by B. Peculis in the Department of Biology, The Johns Hopkins University, Baltimore, MD 21218.

1. *Abbreviations used in this paper:* GV, germinal vesicle; NLS, nuclear localization signal.

equally basic nucleolar targeting sequence was identified in the Tat protein of HIV (Dang and Lee, 1989). In each case, these "nucleolar signals" are comprised of, or contain, sequences resembling previously characterized NLS's.

We have studied the nucleolar localization of NO38, a non-ribosomal, nucleolar protein in the frog *Xenopus laevis*, previously described by Schmidt-Zachmann et al. (1987). We isolated a cDNA clone whose conceptual translation is 87% identical to that described by Schmidt-Zachmann et al. (1987). We used our clone and various derivatives of it to identify sequences required for the nucleolar-specific accumulation of NO38. NO38 does not contain a region resembling the "nucleolar signal" identified in the viral proteins by Siomi et al. (1988) and Dang and Lee (1989), nor are the NLS's alone sufficient to target the protein to the nucleoli. We identified a carboxy-terminal domain that is essential for localization, but is not sufficient to target hybrid proteins to the nucleolus. These results indicate that the mechanisms and/or binding sites which facilitate the nucleolar accumulation of NO38 must be different from those proposed for the viral proteins. We suggest that subnuclear localization of NO38 and other nuclear proteins may depend more on specific protein-protein and protein-nucleic acid interactions than on general linear targeting signals.

## Materials and Methods

### Antibodies

mAbs against *Xenopus* GV proteins were generated in mice (Roth et al., 1990). One stabilized cell line produced mAb 164, which stained the multiple nucleoli in oocytes of *Xenopus*, but not of *Notophthalmus*. Subsequent experiments showed that mAb 164 recognizes the NO38 protein of *Xenopus*. mAb 9E10 specifically recognizes a 10-amino acid peptide in the human *c-myc* protein (Evan et al., 1985). All immunodetections were carried out with undiluted culture supernate from hybridoma cell lines.

### Clones

mAb 164 selected two positive clones in a screen of 80,000 plaques from a lambda ZAP expression library, constructed from *Xenopus* ovary cDNA (Stratagene, La Jolla, CA). The insert from one clone, 164-1, was used as a hybridization probe to isolate an additional 12 clones from the same plates.

All clones were converted into the corresponding Bluescript SK(-) phagemid (Short et al., 1988), and were partially sequenced by the dideoxy chain termination method (Sanger et al., 1977). The clones differed primarily in the extent of 3' or 5' untranslated region. Clone 164-1 was completely sequenced and was used in all subsequent manipulations.

To construct clones that encode *myc*-tagged proteins, we inserted the following oligonucleotide into the HindIII-EcoRI sites of Bluescript KS(+): 5' T TCG AAG CCG CTG GAG TGG ATG GAG CAA AAG CTC ATT TCT GAA GAG GAC TTG AAT TC 3'; and 5' GA ATT CAA GTC CTC TTC AGA AAT GAG CTT TTG CTC CAT CCA CTC CAG CGG CTT CGA A 3'. This oligonucleotide encodes an AUG translation initiation codon followed by 10 amino acids (No. 410-419) from the human *c-myc* protein (Glu Gln Lys Leu Iso Ser Glu Glu Asp Leu). We refer to this clone as the *myc* vector. We cloned the EcoRI insert from 164-1 into the EcoRI site of the *myc* vector. When a cDNA derived from an expressing lambda ZAP clone is inserted in this way, the *myc* tag and the coding region of the cDNA should be in frame. We verified that this was the case for the *myc*-tagged 164-1 clone by sequencing.

### In Vitro RNA Synthesis

Plasmids were linearized at the unique BamHI site in the polylinker downstream of the NO38 sequence, and were used as templates for RNA synthesis in vitro (Melton et al., 1984). Synthesis was carried out with T3 polymerase in the presence of diguanosine triphosphate (Pharmacia Fine Chemicals, Piscataway, NJ) and ribonucleoside triphosphates to produce 5'

capped messages. Clone 164-1 encoded 59 adenosine residues beginning 20 nucleotides downstream of an AATAAA polyadenylation sequence. Thus, RNA transcripts from this template were polyadenylated. After extraction with phenol, the RNA was precipitated with ethanol and redissolved at a concentration of 1 mg/ml.

### Oocytes

Adult female newts, *Notophthalmus viridescens*, were purchased from Lee's Newt Farm (Oak Ridge, TN). *Xenopus laevis* females were purchased from Xenopus I (Ann Arbor, MI). Oocytes from both *Xenopus* and the newt were injected with in vitro-transcribed RNA (Bonner, 1975a,b). *Xenopus* oocytes were maintained in OR2 saline (Wallace et al., 1973). Newt oocytes were maintained in 0.6x DME, 10 mM MOPS, pH 7.2, with 0.5% antibiotics and antimycotics (Bethesda Research Laboratories, Gaithersburg, MD). After injection, the oocytes were held at 18°C for 10-12 h. Translation products were then analyzed by Western blots of extracted proteins or by immunofluorescence of cytological preparations.

### Western Blots

After incubation, the oocytes were transferred to a solution containing 83 mM KCl, 17 mM NaCl, 6.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 3.5 mM KH<sub>2</sub>PO<sub>4</sub>, 10 mM MgCl<sub>2</sub>, 1 mM EDTA, and 1 mM DTT. GVs were hand isolated, washed free of cytoplasm and yolk, and collected in a microfuge tube. The tube was sonicated for 20 s and centrifuged for 20 min at 15,000 g to generate a GV pellet fraction (membranes, chromosomes, nucleoli, and protein aggregates) and a GV supernate fraction (soluble proteins). Enucleated oocytes were sonicated and centrifuged to yield a yolk-free soluble cytoplasmic fraction. The proteins from these three fractions were precipitated with acetone and then boiled in sample buffer (Laemmli, 1970).

Proteins were resolved on 8, 10, or 12.5% polyacrylamide-SDS gels, transferred to nitrocellulose membranes, and analyzed with the appropriate primary antibody followed by biotinylated goat anti-mouse IgG. Detection was by the Vectastain HRP kit (Vector Laboratories, Inc., Burlingame, CA).

### Lampbrush Chromosome Preparations

Spread nuclear preparations were made as described by Gall et al. (1981), with the addition of Mg<sup>2+</sup> to the isolation and dispersal media. GVs were hand isolated in a medium containing 83 mM KCl, 17 mM NaCl, 6.5 mM Na<sub>2</sub>HPO<sub>4</sub>, 3.5 mM KH<sub>2</sub>PO<sub>4</sub>, and 1 mM MgCl<sub>2</sub>. The nuclear contents were dispersed in the same medium at 1/4 strength, with Mg<sup>2+</sup> maintained at 1 mM and with either 0.1 or 0.01% paraformaldehyde. After centrifugation at 2,500 g for 45 min, the preparations were fixed in 70% ethanol, hydrated through a descending series of ethanols, and washed in PBS + 10 mM MgCl<sub>2</sub>. The slides were treated with 3 M urea for 3 min (Hausen and Dreyer, 1982) followed by another wash in PBS + 10 mM MgCl<sub>2</sub>. Preparations were "blocked" with 10% horse serum in TBS for 15 min, after which the primary antibody was applied. The secondary antibody was goat anti-mouse IgG, labeled with rhodamine or fluorescein (Cappel Laboratories, Malvern, PA).

### Oocyte Sections

Oocytes were frozen for a few seconds in liquid isopentane at -161°C and were then dehydrated in 100% ethanol at -80°C for 3 d. The ethanol was warmed to room temperature and the oocytes were transferred through tertiary butanol to liquid paraffin at 60°C. After embedding, sections were cut with a glass knife at 4 μm and floated onto 10 mM Mg<sup>2+</sup>. Sections were placed on glass slides and deparaffinized with xylene. The slides were transferred to 100% ethanol and rehydrated through a series of ethanols into PBS + 10 mM Mg<sup>2+</sup>. The slides were treated with 3 M urea for 3 min and washed with PBS + 10 mM Mg<sup>2+</sup>. Immunofluorescent staining was the same as for chromosome preparations.

### Oligonucleotides

Oligonucleotides were purchased from Research Genetics (Huntsville, AL), or were synthesized on a PCR Mate DNA Synthesizer (model 391; Applied Biosystems, Foster City, CA); they were purified by preparative 20% PAGE (Maniatis et al., 1982). After purification, complementary DNA strands were combined in equimolar concentrations in 100 mM NaCl, heated to 85°C, and cooled slowly to allow annealing. When possible, double stranded oligonucleotides were cloned into asymmetric restriction sites in the plas-

mid. (a) Exonuclease III oligonucleotide. This oligonucleotide was inserted into the unique BsmI and SphI sites in the 3' untranslated region of NO38 cDNA. The oligonucleotide contained restriction sites for EcoRV and AatII upstream of three translation termination codons, one in each reading frame. After cutting, the EcoRV site is a substrate for exonuclease III digestion, whereas the AatII site is not. 5' CAA GAT ATC GCT GAC TGA ACG ACG TCT GAT TGA TTG AGA ATG CT 3' and 5' CAT TCT CAA TCA ATC AGA CGT CGT TCA GTC AGC GAT ATC TTG CAT 3'. (b) NLS. This oligonucleotide was inserted into the EcoRI site between the *myc* tag and the NO38 cDNA, destroying the EcoRI site at the *myc*-NLS border. 5' AAT TTG GCT CCA AAG AAG AAG CGT AAG GTT G 3' and 5' AAT TCA ACC TTA CGC TTA TTT GGA GCC AAA TT 3'. Ribosomes initiating at the AUG in the *myc* sequence read through this segment to produce a peptide corresponding to the SV-40 NLS: (Ala) Pro Lys Lys Lys Arg Lys Val. (c) Multiple cloning site. This oligonucleotide was inserted between the EcoRI site at the end of the NLS and the SphI site at the beginning of the exonuclease III oligonucleotide. It replaces all of the NO38 coding region but leaves 300 bp of untranslated spacer, the polyadenylation signal, and the poly A tail. The oligonucleotide was designed to accept partial cDNAs of  $\beta$ -galactosidase and pyruvate kinase, maintaining a single open reading frame from the *myc* AUG codon to the translation-termination sites in the exonuclease III oligonucleotide (see Fig. 11 A). 5' AAT TCG CCA TGG ACG TAC GGT ACG TAA AAG ATC TAC CGT CTA CGA AAT GCA TG 3' and 5' CAT TTC GTA GAC GGT AGA TCT TTT ACG TAC CGT ACG TCC ATG GCG 3'. (d) PK-NO38 fusion. This oligonucleotide was inserted between the EcoRI site at the 3' end of the *myc* tag and the EcoRI site at the 5' end of NO38, destroying the EcoRI site at the NO38 junction in the process. The oligonucleotide accepts the 390-amino acid fragment of pyruvate kinase and the 150-amino acid segment of NO38 containing the two NLSs and flanking region (see Fig. 11 B). 5' AAT TCC ATG GAA GAT CTA TAC GTA AAT 3' and 5' AAT TAT TTA CGT AAT GAT CTT CAT GG 3'.

### Exonuclease III Digestions

Plasmids were cut with EcoRV, whose terminus is a substrate for exonuclease III, and AatII, whose terminus is resistant to this enzyme. The digested vector was extracted with phenol, precipitated with ethanol, redissolved in the recommended buffer, and treated with exonuclease III (New England Nuclear, Boston, MA) at 27°C. Aliquots were removed every 30 s. The deleted DNAs were blunt ended with mung bean nuclease (United States Biochemicals, Cleveland, OH) and recircularized with T4 ligase (Boehringer Mannheim Biochemicals, Indianapolis, IN). Competent *Escherichia coli* cells were transformed with the ligated DNA. Several deletion constructs were analyzed, initially by restriction digests and later by nucleotide sequencing.

## Results

### NO38 cDNAs

Our studies of the nucleolar protein NO38 began with the isolation and characterization of mAb 164. This antibody was derived from a mouse that had been injected with *Xenopus* GV proteins (Roth et al., 1990). mAb 164 stained the multiple nucleoli from *Xenopus* GVs, but did not stain the chromosomes or any of several types of extrachromosomal granules (Fig. 1). On Western blots of *Xenopus* GV proteins, mAb 164 recognized a protein of 38 kD. It did not react with any protein on Western blots or spread nuclear preparations from the newt, *Notophthalmus viridescens*, or two other tailed amphibians, *Pleurodeles waltl* and *Ambystoma mexicanum*.

mAb 164 was used to screen a lambda ZAP cDNA library (Short et al., 1988) made from poly A(+) RNA isolated from *Xenopus laevis* ovaries. This screen identified two positive clones, which were converted into the Bluescript SK(-) phagemid and sequenced (Fig. 2). One of the cDNAs, 164-1, was 1,387 bp in length and encoded an open reading frame of 296 amino acids, including a putative initiating methionine. Comparison of the conceptual translation of 164-1 to

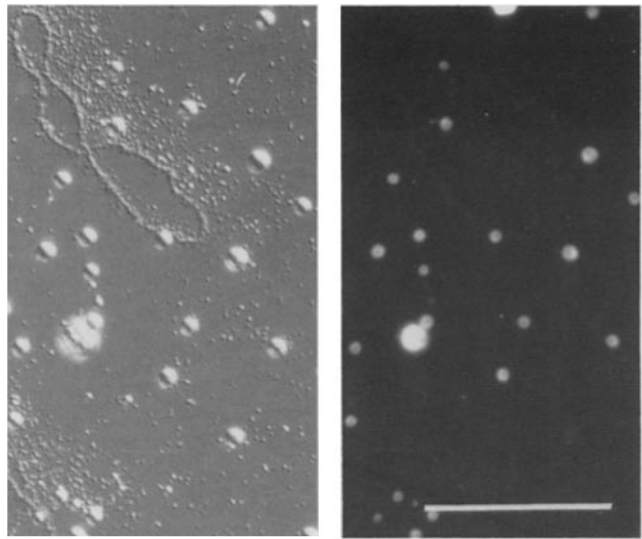


Figure 1. Localization of endogenous NO38 protein in a *Xenopus* GV. Normarski contrast image (left) and corresponding immunofluorescence image (right) showing spread nuclear contents from an uninjected *Xenopus* oocyte, stained with mAb 164 followed by rhodamine-labeled goat anti-mouse IgG. Nucleoli are intensely stained, but chromosomes and other nuclear inclusions are unstained. Bar, 50  $\mu$ m.

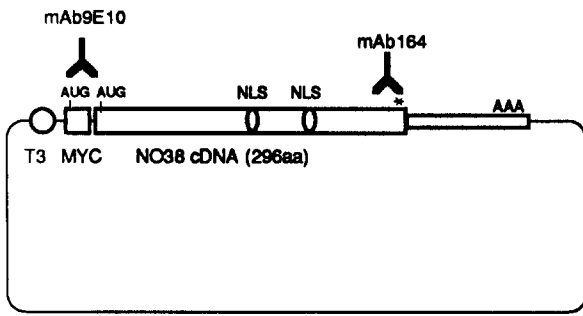
known nucleolar proteins identified the protein as NO38, a nonribosomal, nucleolus-specific protein previously described in *Xenopus* by Schmidt-Zachmann et al. (1987). The nucleotide sequences of 164-1 and the clone isolated by Schmidt-Zachmann et al. (1987) were 85% identical and their conceptual translations were 87% identical.

### *myc* Peptide Tag

In our first injection experiments, we transcribed RNA from the full-length clone 164-1, injected the transcripts into *Notophthalmus* oocytes, and detected the translation products with mAb 164. Since mAb 164 does not react with any endogenous newt protein, we could detect newly translated NO38 protein in these oocytes. These initial experiments could not be performed with *Xenopus* oocytes because mAb 164 reacts with the endogenous NO38 protein. To use *Xenopus* oocytes we needed a way to tag the translation products from injected transcripts. Furthermore, truncated proteins were to be examined for their ability to localize in nucleoli and it was possible that some of the truncations would remove the epitope recognized by mAb 164. For these reasons we constructed a vector that encoded a new epitope. The epitope consisted of 10 amino acids (No. 410-419) derived from the human *c-myc* protein (Evan et al., 1985). An oligonucleotide encoding an AUG (methionine) initiation codon followed by the codons for the 10 amino acids in the *myc* epitope was cloned upstream of, and in frame with, the NO38 cDNA (Fig. 3). Evan et al. (1985) produced an antibody, mAb 9E10, that specifically recognizes this peptide epitope. Thus the *myc* peptide provides an epitope tag that can be used to distinguish the products of injected transcripts from endogenous proteins. In addition, since the *myc* tag is external to the protein coding region of NO38, the tag remains intact through all manipulations of the NO38 cDNA. Finally, the *myc* tag is at the extreme amino terminus of the pro-

1 31  
AAT TCC GGA AGG AGA ATA GCA GGA GTG TAT TTG GTT AAA ATG GAA GAC TCT **ATG** GAT ATG  
MET asp met  
61/4 91/14  
GAC ACC ATC GCT CCC CTC CGC CCG CAG AAT TTC CTC TTT GGA TGT GAA CTT AAA GGA GAC  
asp thr ile ala pro leu arg pro gln asn phe leu phe gly cys glu leu lys gly asp  
121/24 151/34  
AAA AAA GAG TAT TCA TTT AAA GTT GAT GAT GAA AAT GAG CAC CAG TTG TCA TTG AGA ACG  
lys lys glu tyr ser phe lys val asp asp glu asn glu his gln leu ser leu arg thr  
181/44 211/54  
GTT AGT TTG GGT GCT TCT GCT AAA GAT GAA CTT CAT GTT GTT GAG GCT GAA GGA ATA AAT  
val ser leu gly ala ser ala lys asp glu leu his val val glu ala glu gly ile asn  
241/64 271/74  
TAT GAG GGT AAA ACC ATT AAA ATT GCA CTT GCC TCA CTG AAG CCA TCT GTG CAA CCA ACG  
tyr glu gly lys thr ile lys ile ala leu ala ser leu lys pro ser val gln pro thr  
301/84 331/94  
GTA TCG CTA GGT GGG TTT GAG ATA ACA CCA CCA GTA ATT TTG CGG TTG AAA TCG GGT TCA  
val ser leu gly gly phe glu ile thr pro pro val ile leu arg leu lys ser gly ser  
361/104 **C186** 391/114  
GGA CCA GTA TAT GTC AGT GGG CAG CAT CTC GTG GCT TTA GAA GAT CTT GAA TCC TCC GAT  
gly pro val tyr val ser gly gln his leu val ala leu glu asp leu glu ser ser asp  
421/124 451/134 **C153**  
GAA GAC GAT GAT GAA CAA ATA CCC AGT CCA AAA AAT GCA AAA AGG GCA GCA CCA AAT GCA  
glu asp asp asp glu gln ile pro ser pro lys asn ala lys arg ala ala pro asn ala  
481/144 511/154  
GCA TCA AAG GTT CCT CTG AAA AAA ACA CGT CTG GAG GAG GAA GCA GAT TCA GAT GAG GAG  
ala ser lys val pro leu lys lys thr arg leu glu glu glu ala asp ser asp glu glu  
541/164 571/174  
GAC AAT GAT GAT GAT GTG GAT GAT GAA GTA GAT GAG GAA GAA GAA ACA CCA GTT AAA  
asp asn asp asp asp val asp asp glu val asp glu glu glu glu glu thr pro val lys  
601 **C108** 631/194  
AAG ACT GAC TTG TCT ACA AAA TCA AAG GTG GCT CAA AAG TTA AAT CAC AAT GGG AAG GCT  
lys thr asp leu ser thr lys ser lys val ala gln lys leu asn his asn gly lys ala  
661/204 691/214  
TCG GAA TTG TCT GCA ACA CCA AAG ACT CCA AAA ACT CCA GAG CAG AAA GGA AAA CAA GAT  
ser glu leu ser ala thr pro lys thr pro lys thr pro glu gln lys gly lys gln asp  
721/224 751/234  
ACA AAG TCT CAG ACA CCA AAG ACA CCA AAA ACT CCT TTG TCT TTT GAG GAA ATG AAA GCA  
thr lys ser gln thr pro lys thr pro lys thr pro leu ser phe glu glu met lys ala  
781/244 811/254 **C35**  
AAG ATG CAG ACT ATA TTA GAA AAG GGA ACT GTT CTT CCT AAA GTG GAA GTA AAG TTT GCA  
lys met gln thr ile leu glu lys gly thr val leu pro lys val glu val lys phe ala  
841/264 871/274 **C11**  
AAC TAT GTG AAG AAT TGC TTC AGG ACT GAC AGT CAG AAG GTA ATT CAA GAT CTT TGG AAA  
asn tyr val lys asn cys phe arg thr asp ser gln lys val ile gln asp leu trp lys  
901/384 **C5** 931  
TGG AGA CAG TCA CTG AAA GAT GGA AAA **TAA** GCC TGC CTC TCT GTA AAT TTT AAT TTT TTT  
trp arg gln ser leu lys asp gly lys **OCH**  
961 991  
TTT TTT TTT TTA AAT CTA TCA TGG GCC AAG TGT TTT AAA ATG TTA AGA CTA AAA CGC ATG  
1021 1051  
CTT TAT CAA AAT TAA TTT TGT TCT TAC ATT GAC GTT GAA ATT GTG CTC AGG CAA ATG GTT  
1081 1111  
AGA ATG CTC TGA TGT ACA TTA TTT TAC ATT AAA AAA AAA AAA GTT AAT TAC AGG TAG CCC  
1141 1171  
GTA AAA TTG TTC ACT GTA AAC GCA AAT AAT ATT TTT AAA TGG ATT TTA TTG GTG AGA CGA  
1201 1231  
CAA GGG AAC ATT TCA ATA AAT GTG TAC CTT CTC CTT TAA GCA GAC TAC ATG GAG ACA TGG  
1261 1291  
AAC ATG ATT GAT GTT TAT GTA TTC TCA ATT TCC CTA CAA TAA AGC ACA CAG TCA TAT TTA  
1321 1351  
ACC AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA  
1381/461  
AAG AAT TCC

Figure 2. Nucleotide sequence and conceptual translation of NO38 cDNA 164-1. The putative translation initiation and termination codons are indicated in bold type. The two putative NLS's are double underlined. The single underlined residues constitute the "NLS region" used in the PKNLS construct. The extent of the six deletion constructs (C5-C186) are indicated in bold type above the appropriate amino acids. These sequence data are available from EMBL/GenBank/DBJ under accession No. X56039 *X. laevis* mRNA NO38.



**Figure 3.** Diagram of *myc*-tagged 164-1 cDNA showing the relative positions of epitopes. The oligonucleotide marked *MYC* encodes a translation initiation codon (*AUG*) followed by 10 amino acids from the human *c-myc* protein. This peptide forms an epitope recognized by mAb 9E10. Clone 164-1 (NO38 cDNA) is in frame with the *AUG* in the *myc* peptide. The NO38 protein is 296 amino acids in length and contains two putative NLS's. 164-1 encodes a translation termination codon (\*) followed by ~400 bp of 3' untranslated spacer ending in a polyadenylation signal and a poly A tail (*AAA*). mAb 164 recognizes an epitope located within the carboxy terminal 35 amino acids.

tein, where it may be less likely to be buried within the folded protein.

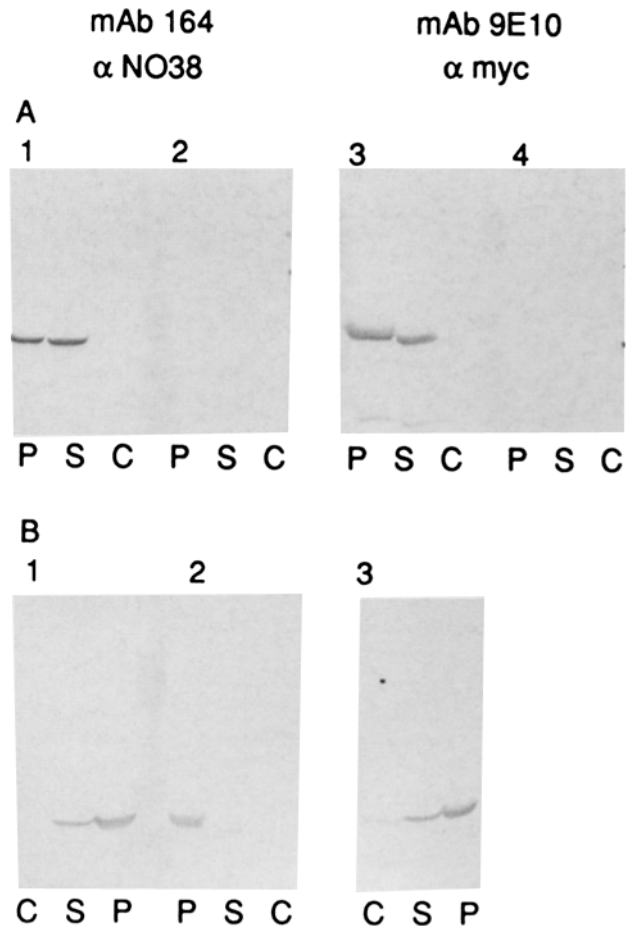
mAb 9E10 detects the human *c-myc* protein, but does not cross react with other mammalian *myc* variants (Evan et al., 1985). In our experiments we saw no immunoreactive proteins on Western blots of cytoplasm, GV supernate, or GV pellet fractions from uninjected oocytes (Fig. 4). However, mAb 9E10 consistently stained a few small loops on several *Notophthalmus* lampbrush chromosomes (not shown). Since there was no detectable stain elsewhere in the nucleus of control oocytes, this small amount of endogenous cross reaction did not detract from the usefulness of the *myc* tag.

### Full-Length NO38 RNA Injections

As described above, the first experiments involved untagged NO38 RNA injected into *Notophthalmus* oocytes. All subsequent experiments were carried out with *myc*-tagged constructs injected into both *Xenopus* and *Notophthalmus* oocytes. Because there were no differences in the behavior of tagged and untagged translation products, only the experiments with *myc*-tagged constructs are detailed here.

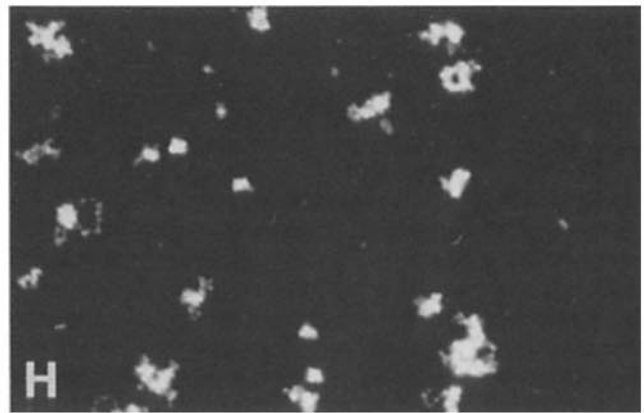
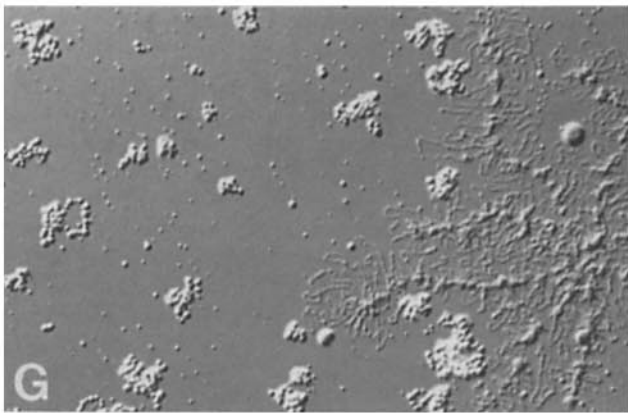
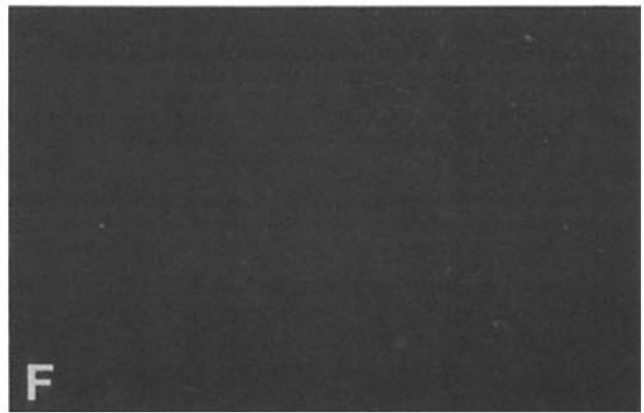
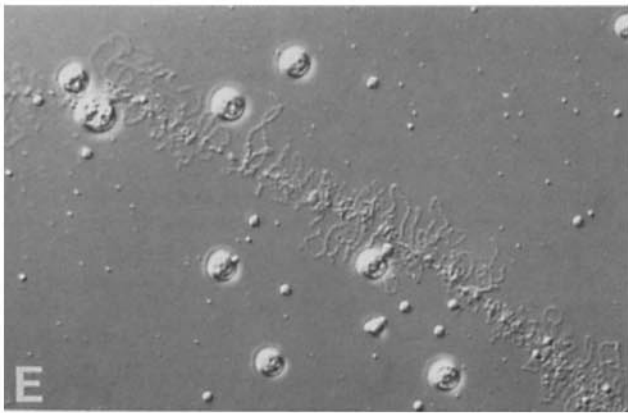
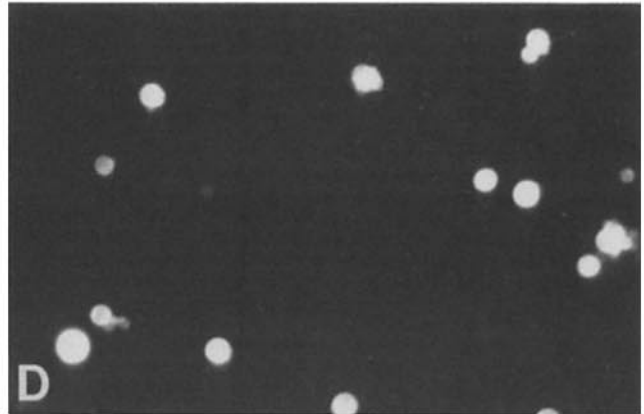
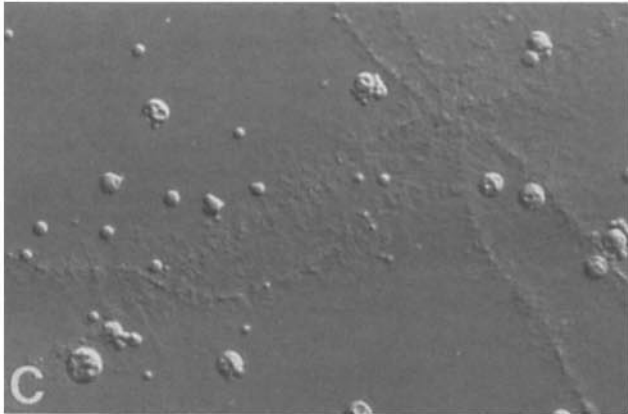
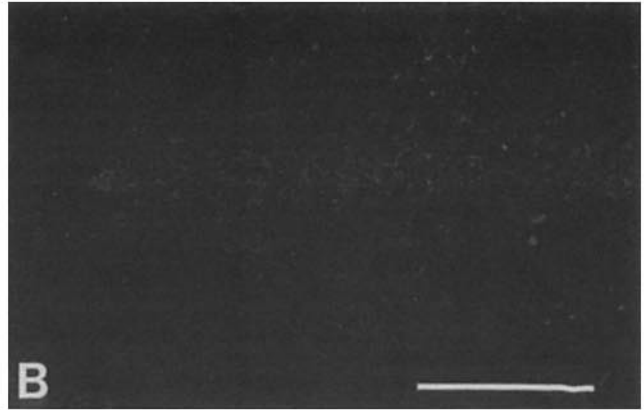
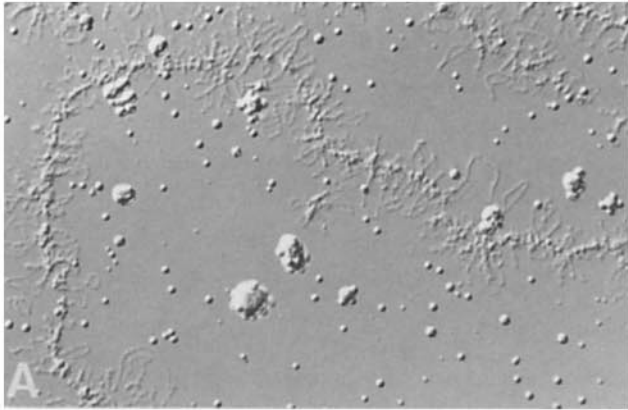
Capped, polyadenylated, sense strand RNA was synthesized *in vitro*. Approximately 25 nl (25 ng) of RNA was injected into the cytoplasm of *Xenopus* or *Notophthalmus* oocytes. The oocytes were incubated for 10–12 h at 18°C in OR2 saline (frog) or 0.6× DME (newt). Proteins from hand-isolated GVs and enucleated oocytes were assayed on Western blots to determine the relative nuclear versus cytoplasmic distribution of the NO38 protein. In addition, the GV contents were centrifuged in a Mg<sup>2+</sup> buffer to yield supernate and pellet fractions. Because nucleoli pellet under the centrifugation conditions used, these fractions provided a rough estimate of “soluble” versus nucleolus-associated NO38. The precise localization of the protein within the GV was studied by indirect immunofluorescence of spread GV contents.

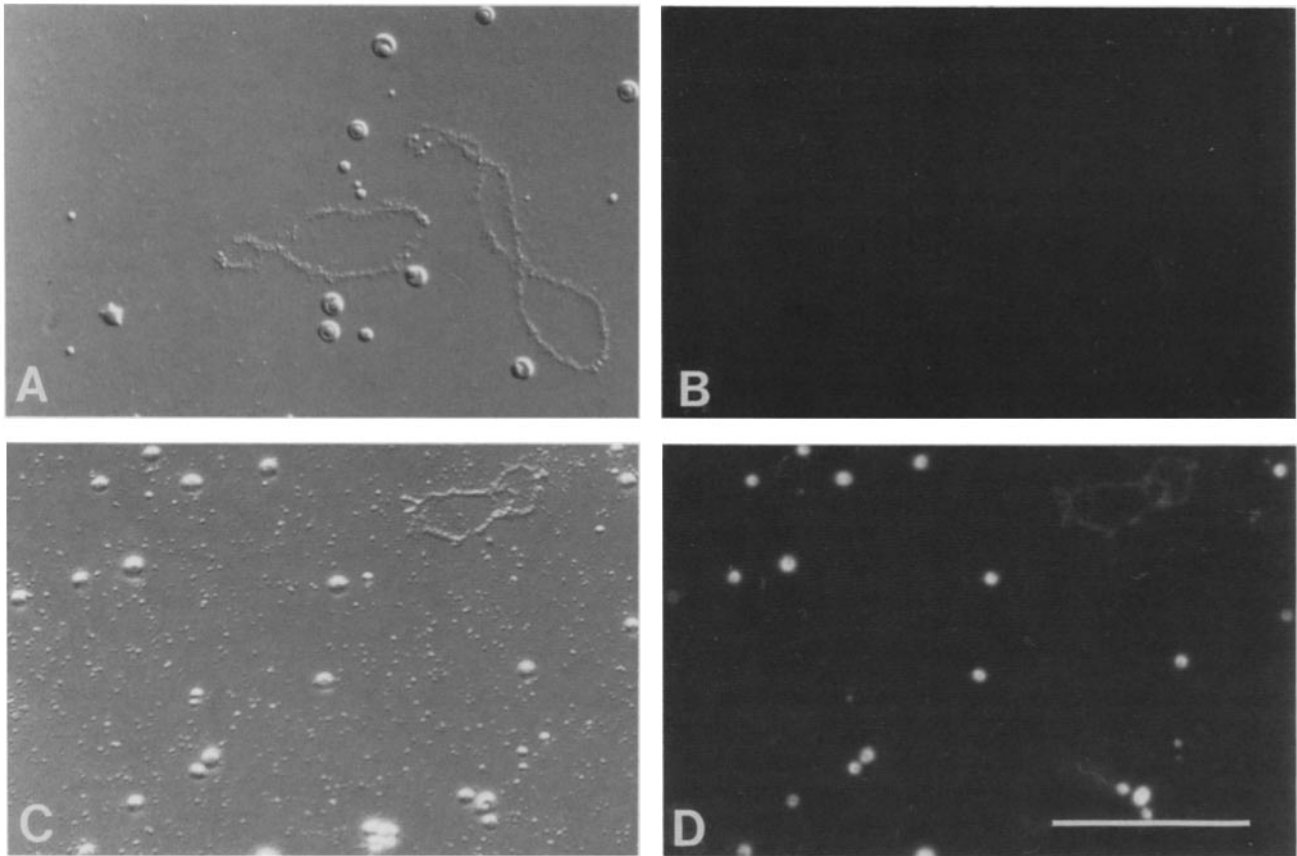
When transcripts encoding full-length NO38 protein were injected into the cytoplasm of oocytes, essentially all translation products detected by mAbs 9E10 and 164 on Western



**Figure 4.** Western blots of proteins from uninjected oocytes and oocytes injected with RNA coding for full-length, *myc*-tagged NO38. GVs were centrifuged to generate GV supernate (S) and GV pellet (P) fractions; yolk was removed from enucleated oocytes by centrifugation to yield soluble cytoplasmic proteins (C). (A) *Notophthalmus* oocytes injected with *myc*-tagged NO38 RNA produced a 39-kD protein, detected with mAb 164, which recognizes NO38 (A1), or with mAb 9E10, which recognizes the *myc* epitope tag (A3). Proteins from uninjected newt oocytes (A2 and A4) did not react with either antibody. (B) Endogenous NO38 protein was detected by mAb 164 in control *Xenopus* oocytes, where it partitions almost completely into the GV pellet fraction (B2). When proteins from injected *Xenopus* oocytes are probed with mAb 164, some newly synthesized protein is seen in the soluble fraction; new and endogenous proteins overlap in the pellet fraction (B1). Only the new *myc*-tagged protein is detected by mAb 9E10 in injected oocytes (B3). In every case, protein translated from injected RNA was efficiently transported into the nucleus, where it distributed about evenly between the GV supernate and GV pellet fractions.

blots localized to the nucleus. Within the nucleus, the immunoreactive protein was about equally distributed between the soluble and pellet fractions (Fig. 4). Control *Notophthalmus* oocytes showed no immunoreactive protein with either antibody (Fig. 4, A2 and A4). The absence of a protein detectable with mAb 164 in control newt GVs is almost certainly due to the epitope specificity of the antibody and not to the absence of an NO38 homologue. Schmidt-Zachmann et al. (1987) clearly demonstrated the presence of NO38 in oocytes of the salamander *Pleurodeles* using an antibody with a broader species cross reactivity.





**Figure 6.** Representative areas from spread nuclear preparations of an uninjected *Xenopus* oocyte (A and B) and one injected with full-length, *myc*-tagged NO38 RNA (C and D). Nomarski contrast image (A and C) and corresponding immunofluorescence image (B and D) after staining with mAb 9E10 and rhodamine-labeled goat anti-mouse IgG. The nucleoli stain intensely in the injected (D), but not in the uninjected (B) sample. Bar, 50  $\mu$ m.

When mAb 9E10 was used for detection of *myc*-tagged protein in *Notophthalmus* or *Xenopus*, there was a single immunoreactive band of 39 kD (Fig. 4, A3 and B3). In *Xenopus* GV, mAb 164 identified a somewhat broader band, which could sometimes be resolved into a doublet, presumably due to the detection of both the endogenous NO38 and the slightly larger translation product of the tagged RNA (Fig. 4 B1). Uninjected or mock-injected *Xenopus* oocytes showed no immunoreactive material with mAb 9E10 (not shown). It should be noted that there is no "soluble" NO38 protein in control *Xenopus* oocytes: that is, the endogenous NO38 is found only in the pellet fraction after GV contents are sonicated and centrifuged in  $Mg^{++}$  buffer (Fig. 4 B2).

Precise localization of *myc*-tagged translation products was determined by indirect immunofluorescence. Ethanol fixed preparations of spread nuclear contents from injected newt oocytes were stained with mAb 164 (Fig. 5, A–D) or mAb 9E10 (Fig. 5, E–H) followed by a rhodamine- or fluo-

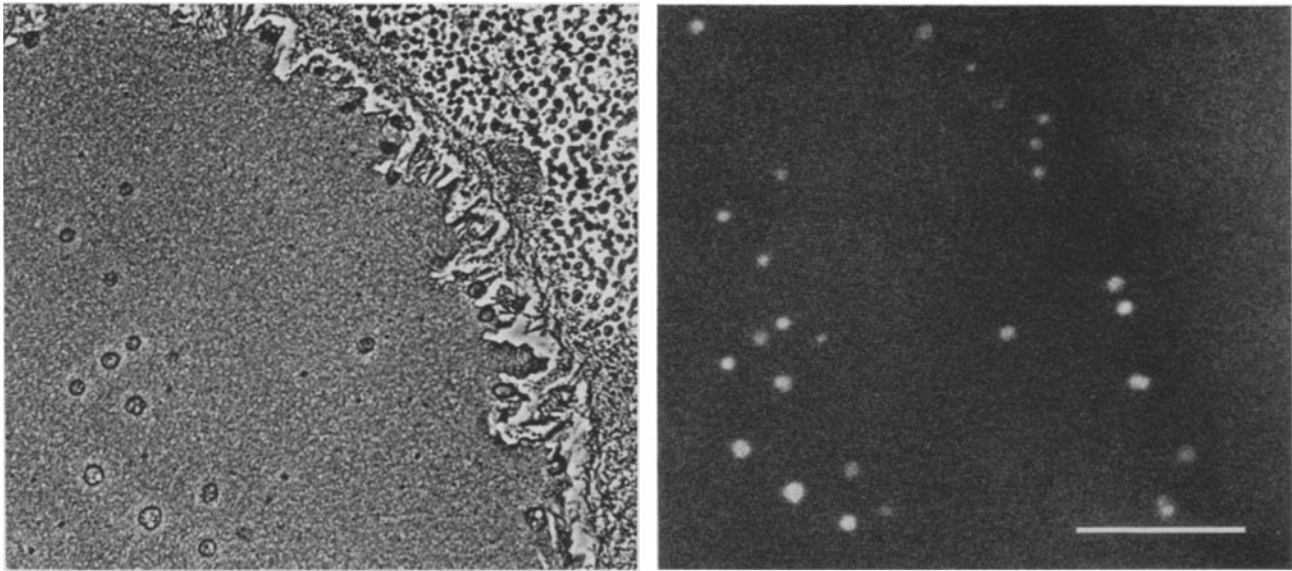
rescein-conjugated secondary antibody. Staining was limited to the nucleoli of injected oocytes (Fig. 5, D and H). Preparations from uninjected and mock-injected controls showed no specific labeling of any nuclear structure (Fig. 5, B and F). Essentially identical observations were made on spread nuclear contents of *Xenopus* oocytes (Fig. 6).

When paraffin sections of injected oocytes were stained with mAb 9E10, immunofluorescence was seen only in the multiple nucleoli of the GV (Fig. 7). The specific labeling of nucleoli in sections and in spread GV contents demonstrates that newly synthesized NO38 is more highly concentrated in the nucleoli than in any other cytological structure. However, the detection of proteins by immunofluorescence is quite sensitive to their concentration. A soluble fraction of newly synthesized NO38 could be distributed throughout the nucleoplasm, yet remain undetected in the stained sections and spreads.

In some cases we followed the accumulation of newly syn-

**Figure 5.** Representative areas from spread nuclear preparations of injected and uninjected *Notophthalmus* oocytes. Nomarski contrast and corresponding immunofluorescence images from uninjected oocytes (A and B, and E and F) and oocytes injected with full-length, *myc*-tagged NO38 RNA (C and D, and G and H). Preparations were stained with mAb 164 (B and D) or mAb 9E10 (F and H), followed by rhodamine-labeled goat anti-mouse IgG. Nucleoli from injected oocytes stain strongly with both antibodies, but those from uninjected oocytes are negative. Bar, 50  $\mu$ m.





**Figure 7.** Paraffin section through a newt oocyte injected with full-length NO38 RNA. Phase contrast image (*left*) and corresponding immunofluorescence image (*right*) after staining with mAb 9E10. The nuclear envelope runs diagonally across the field, separating the GV on the left from the cytoplasm on the right. Several nucleoli are tucked into folds of the nuclear envelope, whereas others are scattered in the nucleoplasm. The nucleoli appear uniformly stained, showing that NO38 protein is not limited to their surfaces. Bar, 50  $\mu\text{m}$ .

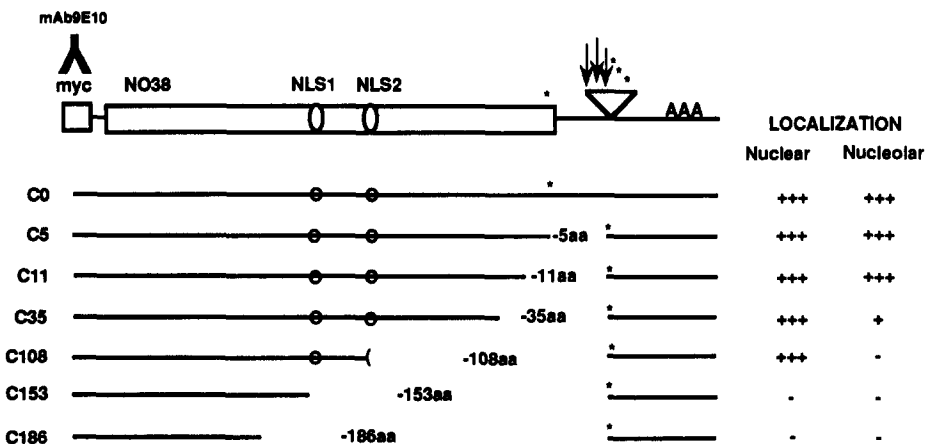
thesized NO38 over time. By 1 h after injection, Western blots showed a detectable amount of protein in the GV, distributed about equally between the supernate and pellet fractions. The total signal increased for the duration of the experiment (16 h), but the ratio between supernate and pellet did not change (data not shown). The earliest that we could detect NO38 in nucleoli by immunofluorescence was 2 h after injection, and the intensity of staining increased with time.

Taken together, the results from the Western blots and the immunofluorescent staining demonstrated that we could mimic most aspects of the endogenous localization process. Microinjected RNA was stable enough to support protein synthesis in the oocyte cytoplasm for at least 16 h, and the

newly translated protein was efficiently transported to the nucleus, where a readily detectable fraction localized in the multiple nucleoli. Furthermore, the *myc* tag had no adverse effect on the localization or detection of newly translated protein.

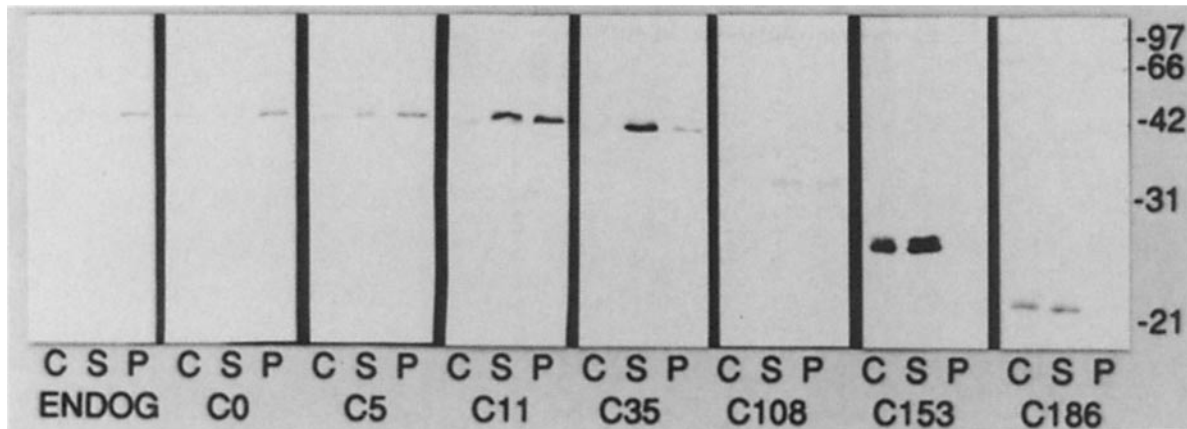
### Exonuclease III Deletions

Having established that full-length transcripts, both tagged and untagged, could be used in localization studies, we next synthesized RNAs that encoded truncated forms of the NO38 protein. Our aim was to identify a protein domain(s) involved in nucleolar accumulation. Since the amino terminus of NO38 resembles nucleoplasmin, a soluble protein not



**Figure 8.** Diagram of deletion constructs. An oligonucleotide was inserted into the 3' untranslated region of the NO38 clone. From a restriction site within this oligonucleotide, exonuclease III generated a set of 3' deletions of the cDNA template. CO denotes the full length construct. C5, C11, C35, C108, C153, and C186 designate clones coding for proteins with 5–186 amino acids missing from the carboxy terminus. The translation termination codon (\*) is provided by the oligonucleotide. C153 and C186 remove both NLS's, whereas C108 removes only half of the second one. The extent of nuclear and nucleolar accumulation, determined by Western blots and immunofluorescence respectively, is indicated for each construct.





**Figure 9.** Translation products from truncated NO38 constructs, detected with mAb 9E10 in *Xenopus* oocytes. Hand-isolated GV's were centrifuged to give GV supernate (S) and GV pellet (P) fractions; enucleated oocytes were the source of soluble cytoplasmic proteins (C). Proteins from constructs with two NLS's (C0, C5, C11, and C35) were efficiently transported into the nucleus, where they distributed about equally between the supernate and pellet fractions. Proteins from C153 and C186, which lack both NLS's, are present in both nucleus and cytoplasm (note that the cytoplasmic lane is derived from 2 oocytes, whereas the nuclear lane contains 10 GV's). In this gel, protein from C108 does not show up well; in most cases it was predominantly in the GV supernate fraction, with some in the GV pellet and cytoplasm.

present in the nucleoli, we began our search for such a domain by generating cDNAs that encoded proteins with carboxy terminal deletions. An oligonucleotide that contained two unique restriction sites (EcoRV and AatII) was inserted into the 3' untranslated region of the *myc*-tagged 164-1 clone (Fig. 8). When the clone was cleaved at these sites, exonuclease III digestion could proceed only toward the 5' end of the construct. In this way we produced a set of nested deletions proceeding toward the 5' end of the coding region. Just downstream of the restriction sites, the oligonucleotide encoded three translation termination codons, one in each reading frame. Thus the open reading frame of each deletion ended within the NO38 sequence and all of the RNAs had an untranslated spacer region of the same length (306 nucleotides) between the translation termination codon and the polyadenylation site.

Several exonuclease III deletion clones were produced, of which six were examined in detail (Fig. 8). RNA was synthesized *in vitro* from these six constructs and injected into the cytoplasm of *Xenopus* and *Notophthalmus* oocytes. After incubation at 18°C for 10–12 h the oocytes were assayed for the presence and distribution of translation products.

Western blot analysis using the *myc*-specific antibody indicated that each construct produced a protein of approximately the predicted size. Translation products from deletion constructs C5, C11, C35, and C108, like the full-length protein from C0, efficiently accumulated within the nucleus (Fig. 9). The deletions in constructs C153 and C186 removed both of the putative NLS's. Translation products from these two constructs, not unexpectedly, distributed between nucleus and cytoplasm roughly according to volume; all of the nuclear signal was in the GV supernate fraction (Fig. 9). The polypeptides from C153 and C186 have expected molecular masses of <29 kD and may enter the nucleus by passive diffusion (Bonner, 1975a,b).

Precise localization of truncated proteins was determined by indirect immunofluorescence. The truncated forms of NO38 encoded by C5 and C11, like the full-length NO38

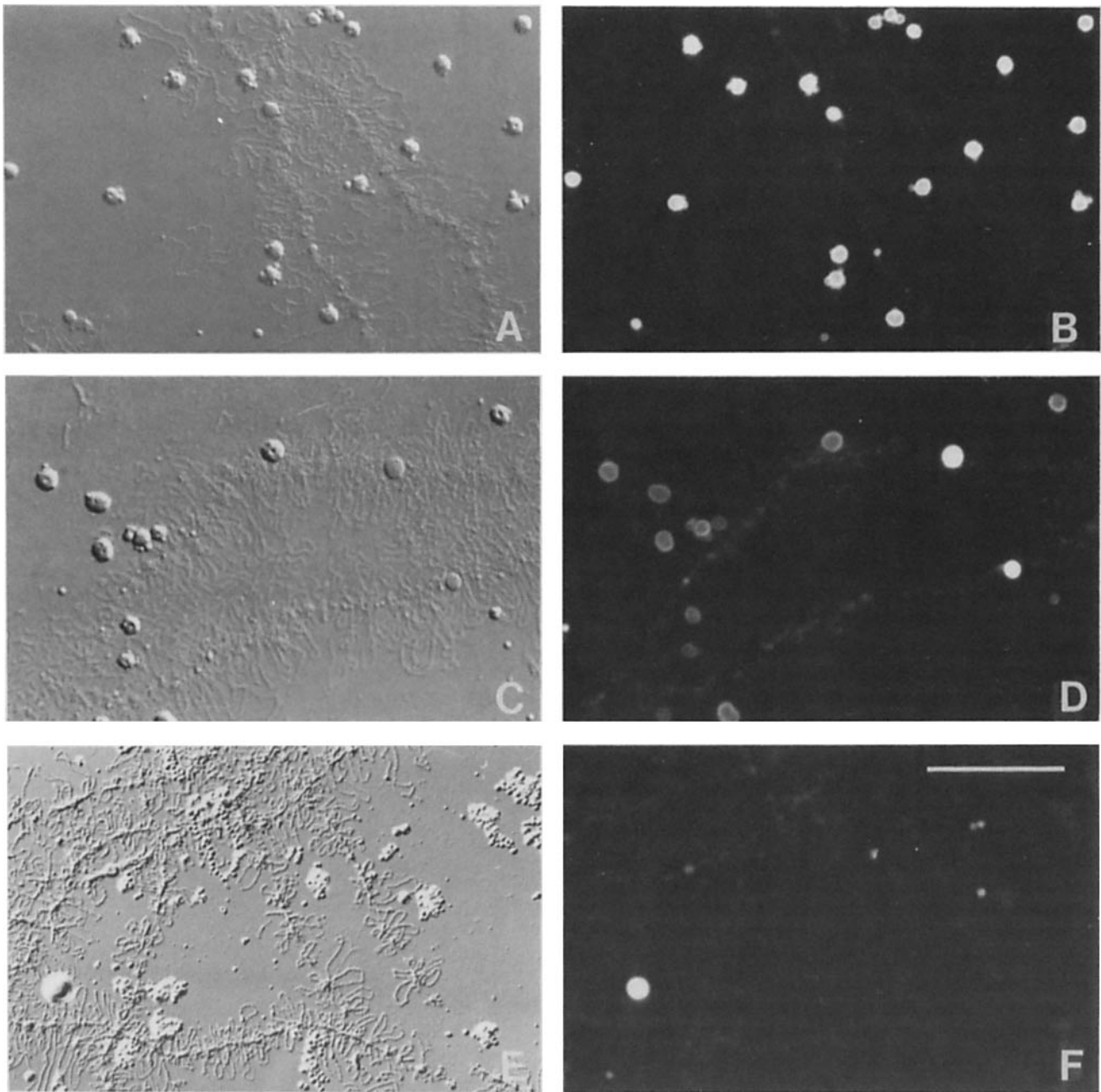
from C0, were concentrated in the nucleoli, as shown by intense staining with the *myc*-specific antibody (Fig. 10, A and B). Thus, deletion of up to 11 amino acids from the carboxy terminus did not detectably alter the ability of NO38 to accumulate in nucleoli. Deletion of an additional 24 amino acids (35 in total) in construct C35 resulted in a dramatic decrease in the intensity of nucleolar label (Fig. 10, C and D). Oocytes injected with construct C108, which deletes the carboxy terminal 108 amino acids of NO38, failed to show any nucleolar staining (Fig. 10, E and F).

Protein from the C108 construct was, however, readily detectable in a small number of subnuclear structures called spheres (Fig. 10, E and F). Weaker labeling of the spheres was also visible with C11 and C35, but not with C0. Spheres are distinct from nucleoli in both morphology and composition (Callan, 1986). They contain snRNPs (Gall and Callan, 1989; Wu et al., 1991), but they do not contain rRNA, nor are they known to be connected in any way with ribosome biogenesis. The accumulation in spheres of the truncated NO38 proteins, particularly the C108 polypeptide, was unanticipated and remains unexplained.

The deletion experiments demonstrate that a domain near the carboxy terminus of the NO38 protein is essential for nucleolar localization. When amino acids 262–286 are eliminated (C35), the protein does not accumulate efficiently in the nucleoli. When additional upstream amino acids are removed (C108), there is no detectable label in nucleoli. It should be noted that the truncated protein from C35 is almost exclusively nuclear in localization, and the protein from C108 is predominantly so (Fig. 9). Thus, the failure of the truncated proteins to localize in the nucleoli is not simply due to an inability to enter the nucleus.

#### *Is the Essential Domain Sufficient for Nucleolar Localization?*

The exonuclease III deletions identified a protein domain that is essential for accumulation of NO38 in nucleoli. To de-

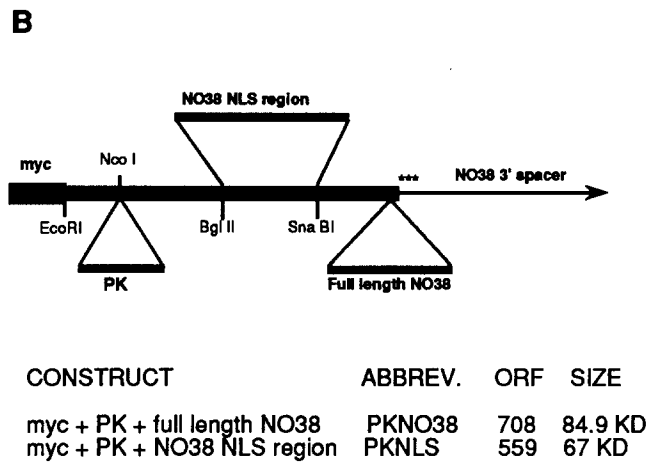
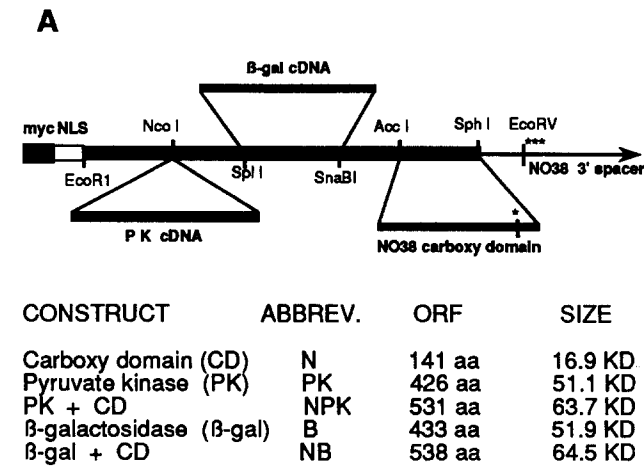


**Figure 10.** Detection of truncated NO38 proteins in spread GV preparations from *Notophthalmus*. Corresponding Nomarski (A, C, and E) and immunofluorescence (B, D, and F) images after staining with the *myc*-specific mAb 9E10 and rhodamine-labeled goat anti-mouse IgG. Oocytes were injected with RNA from constructs C11 (A and B), C35 (C and D), or C108 (E and F). Nucleoli in B are stained as strongly as in oocytes injected with a full-length construct. Nucleoli in D are only weakly stained, and in F they are not detectably stained. The brightly stained structures in D and F are spheres, which differ from nucleoli in both morphology and composition. Bar, 50  $\mu$ m.

termine if this domain is sufficient to target other proteins to nucleoli, we designed hybrid constructs with the characteristics shown in Fig. 11 A. Each construct contained a *myc* tag at the 5' end followed by a nuclear localization sequence to assure that the translation products would enter the nucleus. Two hybrid constructs (*NB* and *NPK*) contained a segment encoding 105 amino acids from the carboxy terminus of NO38, including the domain required for nucleolar accumulation. This segment of NO38 was downstream from the coding region for either *E. coli*  $\beta$ -galactosidase or the cytoplasmic protein, pyruvate kinase. Three control constructs

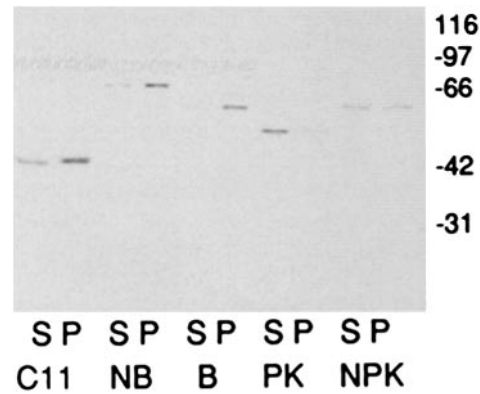
(*N*, *B*, and *PK*) contained the NO38 segment,  $\beta$ -galactosidase, or pyruvate kinase alone. As an additional control, *PKNO38* contained the entire NO38 coding region cloned downstream of the pyruvate kinase fragment in a vector containing a *myc* tag but no SV-40 NLS (Fig. 11 B). This construct tests whether the pyruvate kinase moiety prevents nucleolar accumulation of an otherwise intact NO38 protein.

RNA was transcribed *in vitro* from each of the constructs and injected into the cytoplasm of newt and *Xenopus* oocytes. Protein production and nuclear accumulation were assayed by Western blots, using the *myc*-specific antibody.



**Figure 11.** (A) Diagram of hybrid constructs involving the carboxy terminus of the NO38 protein. A *myc* vector was constructed that encodes an SV-40 NLS followed by an oligonucleotide that can accept 1,170 bp of the pyruvate kinase (*PK*) gene, 1,191 bp of  $\beta$ -galactosidase ( $\beta$ -gal), or 492 bp of NO38 (encoding the carboxy terminal 105 amino acids, the translation stop codon, and 100 bp of untranslated 3' spacer). Five separate constructs were made from these fragments. (B) Diagram of hybrid constructs using NO38 NLS's. A *myc* vector was constructed that contained an oligonucleotide that could accept the 1,170-bp fragment of the pyruvate kinase (*PK*) gene. Downstream were sites that could accept either the 450-bp NLS region of NO38 (Fig. 2, *underlined*) or the full length NO38 coding region. The lengths of the open reading frame (ORF) and predicted molecular masses (SIZE) of the hybrid proteins are indicated.

Each construct produced a protein whose mobility corresponded roughly to that predicted from its molecular mass, and in each case the majority of the protein localized in the nucleus (Fig. 12). The distribution between GV supernate and GV pellet fractions was variable. To assay for specific localization of the hybrid proteins, spread nuclear preparations were stained with the *myc*-specific antibody. The PKN038 fusion protein localized specifically to nucleoli, and the staining intensity was equal to that of the NO38 protein alone. The presence of the pyruvate kinase fragment did not, therefore, inhibit nucleolar accumulation. However, there was no detectable labeling of the nucleoli or any other subnuclear structure with the two hybrid constructs (*NPK*



**Figure 12.** Western blot of translation products from hybrid constructs, detected with mAb 9E10 in *Notophthalmus* GV. In each case the immunoreactive protein is distributed between the GV supernate (S) and GV pellet (P) fractions (other experiments showed that the hybrid proteins are exclusively nuclear). The gel mobilities differ slightly from those predicted from the molecular masses (Fig. 11). *C11*, positive control, nearly full-length NO38; *NB*,  $\beta$ -galactosidase fragment plus NO38 domain; *B*,  $\beta$ -galactosidase fragment alone; *PK*, pyruvate kinase fragment alone; *NPK*, pyruvate kinase fragment plus NO38 domain.

and *NB*) that contained the carboxy terminus of NO38 fused to pyruvate kinase or  $\beta$ -galactosidase. As expected, the constructs containing only pyruvate kinase (*PK*) or  $\beta$ -galactosidase (*B*) did not accumulate in the nucleoli (data not shown). These results indicate that although the carboxy domain is needed for nucleolar accumulation, it alone is not sufficient to target hybrid proteins to the nucleolus.

Because soluble nuclear proteins might be lost from spread nuclear preparations, paraffin sections of injected oocytes were used in an attempt to visualize the hybrid proteins containing the carboxy terminus of NO38. Injected oocytes were incubated for 10–12 h, fixed, embedded in paraffin, and sectioned. The sections were treated with the *myc*-specific antibody and examined by epifluorescence. The nucleoli were not detectably stained, nor was the level of nucleoplasmic staining different from that in control, uninjected oocytes. It should be noted that nucleoplasmic staining in sections was not above background level in any experiment, even when the nucleoli were strongly stained.

### The Central Domain of NO38

The more extensive deletion constructs C153 and C186 (Fig. 2) removed the putative NLS's of NO38, and proteins translated from them failed to concentrate in the nucleus. To test whether the central domain removed by these constructs can target other sequences to the nucleus, we generated construct PKNLS (Fig. 11 B). PKNLS contained a *myc* tag, the pyruvate kinase fragment, and a 150-amino acid region containing the two putative NLS's and the surrounding acidic domain (Fig. 2, *underlined*). RNA was transcribed from PKNLS and injected into *Xenopus* oocytes. Western blots demonstrated that a protein of the appropriate size was produced, which localized entirely within the nucleus. Spread nuclear preparations stained with the *myc*-specific antibody failed to show detectable label in the nucleoli (data not shown).

In summary, hybrid proteins of the expected molecular

weights were synthesized in the oocyte cytoplasm and were targeted to the nucleus by endogenous NO38 signals (PKN-O38 and PKNLS) or a synthetic SV-40 NLS (cloned into the remaining hybrid constructs). Only PKNO38, which contained the entire NO38 sequence, produced a protein that localized in the nucleoli. We conclude, therefore, that the essential carboxy terminal domain is not sufficient to target hybrid proteins to the nucleolus. Furthermore, construct PKNLS suggests that the nucleolar accumulation of NO38 is not a result of an "extended" NLS region, similar to those reported by Siomi et al. (1988) and Dang and Lee (1989).

## Discussion

We isolated a *Xenopus* cDNA clone that encodes the nucleolus-specific protein NO38. This clone is 85% identical in nucleotide sequence to one originally described by Schmidt-Zachmann et al. (1987). By injecting RNA synthesized in vitro from our NO38 clone into the cytoplasm of *Xenopus* and *Notophthalmus* oocytes, we examined the ability of newly synthesized NO38 protein to localize within the GV. RNA transcribed from full-length cDNA templates was efficiently translated by the oocyte for at least 18 h, and the new NO38 protein was found almost exclusively in the GV. Newly translated protein was distinguished from endogenous NO38 by two mAbs. In some cases we injected untagged transcripts into *Notophthalmus* oocytes and detected the new protein with an mAb specific for *Xenopus* NO38. In most experiments we used a template that encoded, in addition to NO38, a short amino-terminal sequence derived from the human *c-myc* gene. This epitope tag was detected with a *myc*-specific mAb, 9E10.

In control *Xenopus* oocytes NO38 protein is easily detectable in the multiple nucleoli by immunofluorescent staining. That NO38 is limited to nucleoli is suggested by the observation that it is found exclusively in the pellet fraction when GVs are sonicated and centrifuged at 15,000 g for 20 min. Under these fractionation conditions, the chromosomes and nucleoli pellet, but ~95% of the GV proteins remain in the supernate (Benavente et al., 1984). In our injection experiments, we found that newly synthesized, full-length NO38 was about equally distributed between the pellet and supernate fractions; the ratio between these fractions did not change appreciably over a period of 16 h. Because the nucleoli show strong immunofluorescence in these experiments, we presume that the pelletable NO38 corresponds to protein rather firmly associated with nucleoli. The cytological location of the nonpelletable NO38 is unknown. It is possible that all newly synthesized NO38 is localized in the nucleoli, but a loosely bound fraction becomes soluble after the GVs are isolated and sonicated in buffer. It is equally possible that the soluble NO38 is dispersed in the nucleoplasm. In this case newly synthesized NO38 would be too dilute to detect by immunofluorescent staining of sections (Fig. 6) and might be washed away during preparation of GV spreads.

Since we were able to inject RNA encoding NO38 and mimic the localization process of the endogenous protein, we next attempted to identify regions in NO38 which were essential for the specific accumulation. There are three general types of experiments that have been used by others to define functional nuclear localization sequences. Deletion analyses are used to identify regions of the protein involved in the lo-

calization, gene fusions test the sufficiency of a protein domain to target hybrid proteins, and mutation experiments systematically determine the sequence requirements of a defined set of amino acids. Two of these three types of experiments were used to identify regions in NO38 involved in both nuclear and nucleolar accumulation.

The injection experiments with truncated constructs identified a domain at the carboxy terminus of NO38 that is essential for nucleolar accumulation. A protein lacking the carboxy-terminal 35 amino acids (C35) showed only weak nucleolar association, and one lacking 108 amino acids (C108) was not detectable in nucleoli by immunofluorescence. C35 was exclusively nuclear and C108 was predominantly so. Thus, the failure of the deleted constructs to accumulate in the nucleoli was not due to their absence from the nucleus.

When gene fusions were constructed to join the carboxy-terminal 108 amino acids of NO38 with either of two non-nuclear proteins ( $\beta$ -gal or pyruvate kinase), and an SV-40 NLS was provided to ensure nuclear accumulation, the hybrid proteins were found in the nucleus as expected. However, these hybrid proteins did not accumulate in nucleoli. The control hybrid protein, a pyruvate kinase fragment joined to the full-length NO38 protein, did enter the nucleus and accumulate in nucleoli, demonstrating that the pyruvate kinase fragment does not inhibit nucleolar accumulation. These results indicate that the carboxy domain is needed for proper accumulation, but is not sufficient to target hybrid proteins to nucleoli.

The domain necessary for nucleolar accumulation is distinct from two sequences near the middle of NO38 that are probably involved in nuclear localization. The first of these (amino acids 137-153) contains consensus elements that define the nucleoplasmin-like nuclear localization motif (Robbins et al., 1991). The second (amino acids 181-187) resembles the SV-40 NLS (Kalderon et al., 1984a,b). Translation products from deletion constructs C0, C5, C11, and C35, which contain both putative NLS's, localize exclusively in the nucleus. However, proteins from constructs C153 and C186, which lack these sequences, distribute between nucleus and cytoplasm more or less in proportion to volume. The proteins encoded by C153 and C186 are <29 kD in size and, like other small proteins, can probably enter the nucleus by diffusion (Bonner, 1975a,b; Paine et al., 1975). C108 includes one complete and one partially deleted sequence (Figs. 2 and 8). In most experiments C108 translation products showed predominantly nuclear localization, although some cytoplasmic protein was detected. This observation suggests that NO38 may localize more efficiently when both signals are intact. More direct evidence that these sequences are functional NLS's is provided by the hybrid construct PKNLS. This construct contains the region encoding the two putative NLS's, the long acidic region which separates them, and some additional flanking sequence (see Fig. 2, *underlined*), fused to pyruvate kinase. Translation products from this construct localize exclusively in the nucleus, but do not accumulate within nucleoli.

An unexpected finding of the deletion studies was the accumulation of the protein encoded by C108 in spheres. This construct leaves intact, and perhaps in an exposed configuration, the long glutamic and aspartic acid-rich domain of the NO38 protein. The next shorter deletion construct, C153,

lacks this domain, and its products fail to accumulate in the spheres. It is possible, therefore, that the product of C108 interacts with the spheres via this acid domain. We have no reason to believe that the accumulation in spheres, whatever its mechanism, is more than an accidental curiosity.

Several studies have been made to identify the sequences responsible for the nucleolar localization of the heat-shock protein, hsp70. In unstressed cells, hsp70 is found in both cytoplasm and nucleus, but it shows a striking association with the nucleolus after heat-shock (Welch and Feromisco, 1984). Munro and Pelham (1984) expressed deletion constructs of the *Drosophila* hsp70 gene in monkey COS cells and showed that nucleolar localization in heat-shocked cells required an evolutionarily conserved region in the amino half of the molecule. Dang and Lee (1989) made a construct that encoded an 18-amino acid sequence from within this region fused to pyruvate kinase. Some of the chimeric protein expressed by their construct localized in the nucleolus, while the remainder was found in the nucleoplasm and cytoplasm. Milarski and Morimoto (1989) examined the effect of internal deletions of the human hsp70 protein on nucleolar localization and identified an essential region in the carboxy half of the molecule. It is not clear whether the *Drosophila* and human hsp70 proteins differ in the sequences responsible for nucleolar localization. It is likely that sequences in both the amino and carboxy regions of the molecule are involved in localization.

Targeting experiments have been carried out with several other proteins found in the nucleolus. Siomi et al. (1988) and Nosaka et al. (1989) studied the Rex protein of HTLV-I (also known as p27x-III), one of three proteins transcribed from the pX region. They identified a highly basic sequence of 19 amino acids at the amino terminus of Rex (Met Pro Lys Thr Arg Arg Arg Pro Arg Arg Ser Gln Arg Lys Arg Pro Pro Thr Pro), the first part of which resembles the NLS in SV-40 and the second part the NLS in polyoma virus. Hybrid proteins consisting of the Rex sequence fused to  $\beta$ -galactosidase or to protein p40x of HTLV-I localized in the nucleolus, as did the intact Rex protein. Several hybrid proteins with small deletions or amino acid substitutions in this region accumulated in the nucleus but lost their specific nucleolar localization.

Dang and Lee (1989) identified a 12-amino acid basic sequence (Glu Arg Lys Lys Arg Arg Gln Arg Arg Ala Pro) that they believe is responsible for nucleolar localization of the Tat protein of HIV. When a fusion was made between this sequence and pyruvate kinase, the hybrid protein was found in the nucleolus, as well as in the nucleoplasm and cytoplasm. They also examined a sequence consisting of the NLS of *c-myc* (Pro Ala Ala Lys Arg Val Lys Leu Asp) adjacent to the carboxy-terminal 6 amino acids of the HIV Tat sequence (Gln Arg Arg Arg Ala Pro). Pyruvate kinase containing this chimeric signal localized predominantly in the nucleolus. Dang and Lee (1989) suggest that a short array of basic amino acids is sufficient for nuclear localization, but a long array may be necessary for nucleolar accumulation.

The underlying question raised by the targeting experiments is whether there is a common "nucleolar localization sequence" for most or all nucleolar proteins. A similar issue, of course, underlies nuclear targeting experiments. To reach the interior of the nucleus, all proteins larger than some minimal size must presumably pass through the same barrier, the nuclear pore. In this case, it is easy to imagine that there may

be a limited number of mechanisms for protein import, such as a single binding site at the pore or a few carrier proteins. Consistent with this assumption, several nuclear proteins contain well-defined NLS's. The best studied of these sequences consist of runs of basic amino acids (reviewed in Dingwall and Laskey, 1986; Silver and Goodson, 1989), but even here no universal consensus sequence exists.

The situation with respect to subnuclear localization is probably more complex. Inside the nucleus, there are no obvious physical barriers to relatively free movement of macromolecules, and consequently no obvious need for transport systems to bring molecules to the chromosomes, nucleoli, or other subnuclear components. It seems more likely to us that most cases of intranuclear localization will be explicable in terms of protein-protein and protein-nucleic acid binding. Chromosomal DNA and nascent RNA transcripts may provide the primary binding sites; proteins bound to these components could then serve as additional sites for a wide variety of protein-protein interactions.

If this view is correct, nucleolar targeting experiments will define only those sequences used by a particular macromolecule in its binding to the nucleolus. Recent studies suggest that the Tat protein binds to an RNA stem-loop structure in the HIV long terminal repeat (Weeks et al., 1990; Cordingley et al., 1990). The RNA-binding region is included in the localization domain defined by nucleolar targeting experiments. Very possibly, the Tat protein binds to a similar stem-loop in ribosomal RNA in the nucleolus. If this is the case, one would expect only a limited number of nucleolar proteins to bind in this way. NO38 clearly differs from the viral proteins Rex and Tat in the amino acid sequences required for nucleolar localization. The central region of NO38, which contains two putative NLS's, targets pyruvate kinase to the nucleus but not to the nucleolus (construct PKNLS). Nucleolar accumulation, therefore, is not conferred by an "extended" NLS, as in the viral proteins. Moreover, the essential domain defined by our experiments is located near the carboxy terminus some distance from the putative NLS's. This domain contains few basic residues and is not sufficient to target hybrid proteins to the nucleolus. One possibility is that nucleolar accumulation of NO38 requires the cooperative action of sequences in both the amino and carboxy halves of the molecules; that is, the protein associates with the nucleolus only when noncontiguous sequences are in a properly folded configuration. In such a case, one could not define a nucleolar localization sequence by targeting experiments that involve short runs of amino acids, or even larger domains, attached to a reporter protein.

We thank Mark Roth for producing, stabilizing, and initially characterizing mAb 164. We also thank Christine Murphy for expert technical assistance, Sean Munro for the cell line that produces mAb 9E10, Bruce Roberts for the pyruvate kinase cDNA, and Andrew Fire for the  $\beta$ -galactosidase cDNA. Douglas Fambrough kindly helped with the synthesis of oligonucleotides.

Supported by grant GM 33397 from the National Institute of General Medical Sciences. J. G. Gall is American Cancer Society Professor of Developmental Genetics.

Received for publication 30 July 1991 and in revised form 25 September 1991.

#### References

Benavente, R., G. Krohne, M. S. Schmidt-Zachmann, B. Hügler, and W. W.

- Franke. 1984. Karyoskeletal proteins and the organization of the amphibian oocyte nucleus. *J. Cell Sci.* 1(Suppl.):161-186.
- Bonner, W. 1975a. Protein migration into nuclei I. Frog oocyte nuclei in vitro accumulate microinjected histones, allow entry to small proteins, and exclude large proteins. *J. Cell Biol.* 64:421-430.
- Bonner, W. 1975b. Protein migration into nuclei II. Frog oocyte nuclei accumulate a class of microinjected oocyte nuclear proteins and exclude a class of oocyte cytoplasmic proteins. *J. Cell Biol.* 64:431-437.
- Bürglin, T. R., and E. M. De Robertis. 1987. The nuclear migration signal of *Xenopus laevis* nucleoplasmin. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2617-2625.
- Callan, H. G. 1986. Lampbrush Chromosomes. Springer-Verlag New York Inc., New York. 1-254.
- Cordingley, M. G., R. L. LaFemina, P. L. Callahan, J. H. Condra, V. V. Sardana, D. J. Graham, T. M. Nguyen, K. LeGrow, L. Gotlib, A. J. Schlabach, and R. J. Colonna. 1990. Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive element of human immunodeficiency virus type 1 in vitro. *Biochemistry.* 87:8985-8989.
- Dabauvalle, M.-C., and W. W. Franke. 1982. Karyophilic proteins: polypeptides synthesized in vitro accumulate in the nucleus on microinjection into the cytoplasm of amphibian oocytes. *Proc. Natl. Acad. Sci. USA.* 79:5302-5306.
- Dang, C., and W. M. F. Lee. 1989. Nuclear and nucleolar targeting sequences of c-erb-A, c-myc, N-myc, p53, HSP70, and HIV Tat proteins. *J. Biol. Chem.* 264:18019-18023.
- De Robertis, E. M., R. F. Longthorne, and J. B. Gurdon. 1978. Intracellular migration of nuclear proteins in *Xenopus* oocytes. *Nature (Lond.)*. 272:254-256.
- Dingwall, C., and R. A. Laskey. 1986. Protein import into the cell nucleus. *Annu. Rev. Cell Biol.* 2:367-390.
- Dingwall, C., S. V. Sharnick, and R. A. Laskey. 1982. A polypeptide domain that specifies migration of nucleoplasmin into the nucleus. *Cell.* 30:449-458.
- Dingwall, C., J. Robbins, S. M. Dilworth, B. Roberts, and W. D. Richardson. 1988. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J. Cell Biol.* 107:841-849.
- Evan, G. I., G. K. Lewis, G. Ramsay, and J. M. Bishop. 1985. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell Biol.* 5:3610-3616.
- Gall, J. G., and H. G. Callan. 1989. The sphere organelle contains small nuclear ribonucleoproteins. *Proc. Natl. Acad. Sci. USA.* 86:6635-6639.
- Gall, J. G., E. C. Stephenson, H. P. Erba, M. O. Diaz, and G. Barsacchi-Pilone. 1981. Histone genes are located at the sphere loci of newt lampbrush chromosomes. *Chromosoma (Berl.)*. 84:159-171.
- Hausen, P., and C. Dreyer. 1982. Urea reactivates antigens in paraffin sections for immunofluorescent staining. *Stain Technol.* 57:321-324.
- Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith. 1984a. A short amino acid sequence able to specify nuclear location. *Cell.* 39:499-509.
- Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith. 1984b. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature (Lond.)*. 311:33-38.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of a bacteriophage T4. *Nature (Lond.)*. 227:680-685.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning—a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zimm, and M. R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12:7035-7056.
- Milarski, K. L., and R. I. Morimoto. 1989. Mutational analysis on the human HSP70 protein: distinct domains for nucleolar localization and adenosine triphosphate binding. *J. Cell Biol.* 109:1947-1962.
- Munro, S., and H. R. B. Pelham. 1984. Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of *Drosophila* hsp70. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:3087-3093.
- Nosaka, T., H. Siomi, Y. Adachi, M. Ishibashi, S. Kubota, M. Maki, and M. Hatanaka. 1989. Nucleolar targeting signal of human T-cell leukemia virus type I Rex-encoded protein is essential for cytoplasmic accumulation of unspliced viral mRNA. *Proc. Natl. Acad. Sci. USA.* 86:9798-9802.
- Paine, P. L., and C. M. Feldherr. 1972. Nucleocytoplasmic exchange of macromolecules. *Exp. Cell Res.* 74:81-98.
- Paine, P. L., L. C. Moore, and S. B. Horowitz. 1975. Nuclear envelope permeability. *Nature (Lond.)*. 254:109-114.
- Robbins, J., S. M. Dilworth, R. A. Laskey, and C. Dingwall. 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell.* 64:615-623.
- Roth, M. B., C. Murphy, and J. G. Gall. 1990. A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian oocyte nucleus. *J. Cell Biol.* 111:2217-2223.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.
- Schmidt-Zachmann, M. S., B. Hügler-Dörr, and W. W. Franke. 1987. A constitutive nucleolar protein identified as a member of the nucleoplasmin family. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:1881-1890.
- Short, J. M., J. M. Fernandez, J. A. Sorge, and W. D. Huse. 1988. Lambda Zap: a bacteriophage lambda expression vector with in vitro excision properties. *Nucleic Acids Res.* 16:7583-7600.
- Silver, P., and H. Goodson. 1989. Nuclear protein transport. *Crit. Rev. Biochem. Mol. Biol.* 24:419-436.
- Siomi, H., H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka. 1988. Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. *Cell.* 55:197-209.
- Wallace, R. A., D. W. Jared, J. N. Dumont, and M. W. Sega. 1973. Protein incorporation by isolated amphibian oocytes: III. Optimum incubation conditions. *J. Exp. Zool.* 184:321-333.
- Watson, M. E. E. 1984. Compilation of published sequences. *Nucleic Acids Res.* 12:5145-5164.
- Weeks, K. M., C. Ampe, S. C. Schultz, T. A. Steitz, and D. A. Crothers. 1990. Fragments of the HIV-1 protein specifically bind TAR RNA. *Science (Wash. DC)*. 249:1281-1285.
- Welch, W. J., and J. R. Feramisco. 1984. Nuclear and nucleolar localization of the 72,000-dalton heat shock protein in heat shocked mammalian cells. *J. Biol. Chem.* 259:4501-4513.
- Wu, Z., C. Murphy, H. G. Callan, and J. G. Gall. 1991. Small nuclear ribonucleoproteins and heterogeneous nuclear ribonucleoproteins in the amphibian germinal vesicle: loops, spheres, and snurposomes. *J. Cell Biol.* 113:465-483.