Primary Sequence and Heterologous Expression of Nuclear Pore Glycoprotein p62

Christopher M. Starr, Mara D'Onofrio, Min K. Park, and John A. Hanover

Laboratory of Biochemistry and Metabolism, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Abstract. The major nuclear pore protein p62 is modified by O-linked *N*-acetylglucosamine and functions in nuclear transport. We have cloned, sequenced, and expressed the full-length rat p62 cDNA. The rat p62 mRNA is 2,941 nucleotides long and encodes a protein of 525 amino acids containing 30% serine and threonine residues. The amino acid sequence near the amino-terminus contains unique tetrapeptide repeats while the carboxy-terminus consists of a series of predicted alpha-helical regions with hydrophobic heptad repeats. Heterologous expression of rat p62 in African Green Monkey kidney COS-1 cells and CV-1 cells was detected using a species-specific antipeptide serum. When transiently expressed in COS-1 cells, rat p62 binds wheat germ agglutinin and concentrates at the spindle poles during mitosis. In CV-1 cells cotransfected with rat p62 cDNA and SV40 viral DNA, rat p62 associates with the nuclear membrane without interfering with the nuclear transport of SV40 large T antigen. The ability to express p62 in tissue culture cells will facilitate analysis of the role of this pore protein in nuclear transport.

THE nuclear pore complex (NPC),[1] a component of all eukaryotic cell nuclei, is a supramolecular structure with an estimated mass of 100–125 × 10^6 D and a diameter of 1,000 Å. Located in the center of the NPC is an aqueous channel through which proteins and nucleic acids are transported between the nuclear and cytoplasmic compartments. Nuclear protein import through the nuclear pore is gated and requires the recognition of specific nuclear localization signals by components of the NPC. In vitro experiments and studies in living cells suggest that O-linked *N*-acetylglucosamine (GlcNAc)–modified pore glycoproteins play a role in the regulation of transport through the NPC (Finlay et al., 1987; Wolff et al., 1988; Featherstone et al., 1988; Dabauvalle et al., 1988). Protein import has been functionally divided into two steps: a binding step that is followed by an ATP-dependent translocation step (Newmeyer and Forbes, 1988). The lectin wheat germ agglutinin (WGA) inhibits the translocation step without interfering with the binding step, suggesting that the pore glycoproteins are involved in the transport of proteins through the pore but are not involved in the binding of nucleophilic proteins to pore "receptors."

mAbs directed against the nuclear pore show a broad binding specificity for a family of glycoproteins ranging in molecular weight from 54,000–210,000 D (Davis and Blobel, 1986; Park et al., 1987; Snow et al., 1987; Davis and Blobel, 1987; Dabauvalle et al., 1988). These glycoproteins contain multiple sites of single GlcNAc moieties attached directly to the polypeptide backbone via O-glycosidic linkages to serine (and possibly threonine) residues (Hanover et al., 1987; Holt et al., 1987). The major glycoprotein of the nuclear pore has a molecular mass of 62 kD and has been referred to as p62. To define the sequence requirement for GlcNAc modification, we previously sequenced a site of O-linked GlcNAc addition in p62. From a partial rat p62 cDNA, we determined that the peptide sequence Pro-Ala-Asp-Thr-Ser-Asp-Pro contains a serine modified by O-linked GlcNAc (D'Onofrio et al., 1988). Unlike Asn-linked glycosylation and O-linked GalNAc addition, which take place in the lumen of the endomembrane system, O-linked GlcNAc addition is believed to occur in the cytosol either cotranslationally or within minutes of the completion of protein synthesis (Holt et al., 1987; Hanover et al., 1987; Davis and Blobel, 1987; Starr and Hanover, 1990).

In this report, we have cloned and sequenced a full-length cDNA-encoding rat p62. Using a rat-specific antipeptide serum we are able to detect the expression of rat p62 in transfected COS-1 and CV-1 cells and demonstrate that nuclear transport occurs normally in cultured cells overexpressing the pore protein.

Materials and Methods

Isolation and Analysis of cDNA Clones

To obtain a full-length rat p62 cDNA, 1.5 × 10^6 pfu of a rat thyroid FRTL-5 lambda gt11 expression library were screened by plaque hybridization using a partial 690 bp cDNA of rat p62 described previously (D'Onofrio

1. *Abbreviations used in this paper*: GlcNAc, *N*-acetylglucosamine; NPC, nuclear pore complex; NRK, normal rat kidney; WGA, wheat germ agglutinin.

© The Rockefeller University Press, 0021-9525/90/06/1861/11 $2.00

et al., 1988). The cDNA probe was labeled with [alpha-^{32}P]deoxycytidine-5′-triphosphate (3,000 Ci/mmol) (ICN Radiobiochemicals, Irvine, CA) to a specific activity of 2×10^9 cpm/μg using a multiprimer labeling kit (Amersham Corp., Arlington Heights, IL), and was hybridized to library plaque lifts on Hybond-N nylon filters (Amersham Corp.) in 6× SSC, 5× Denhardt's solution, and 0.5% SDS for 24 h at 65°C. Eight p62 cDNA clones were plaque purified from the FRTL-5 library. The Eco RI cDNA inserts of the phage clones were subcloned into the sequencing plasmid pGEM-3Zf(+) (Promega Biotec, Madison, WI) and used to transform DH5α-competent cells (Bethesda Research Laboratories, Gaithersburg, MD). These clones were sequenced on both strands using synthetic 18mer primers (The Midland Certified Reagent Co., Midland, TX) according to the strategy shown in Fig. 2 *a* by the dideoxy chain termination method using Sequenase (United States Biochemical Corp., Cleveland, OH).

The p62 5′ extension library was prepared from rat liver poly(A)-mRNA using a 32mer primer complementary to nucleotides 998–1,030 in the p62 mRNA sequence shown in Fig. 2 *b*. The first strand of the cDNA was synthesized using M-MLV reverse transcriptase (400 U) (Bethesda Research Laboratories) in a total reaction volume of 40 μl containing 40 μg of poly(A)-mRNA, 100 ng of 32mer primer, 0.5 mM dNTPs, 2 μg actinomycin D, 80 U RNasin (Promega Biotec), and 1× M-MLV buffer supplied with the enzyme. The second strand was synthesized using DNA polymerase I and RNase H (Promega Biotec) as described by Gubler (1987). The cDNA constructs were ligated to phosphorylated Eco RI linkers (New England Biolabs, Beverly, MA) and size fractionated on a CL-4B column. Material collected in the void volume of the column was cloned into the Eco RI site of pGEM-3Zf(+) (see Deininger, 1987, for general methods). The 5′ extension library was prepared from these p62-enriched cDNA constructs in DH5α cells and was screened by colony hybridization using the Bsp MI/Ava2 restriction fragment corresponding to nucleotides 747–933 in Fig. 2 *b*. The 186-bp probe was radiolabeled and the extension library was screened on nylon filters essentially as described above for phage library screening. Approximately 10% of the colonies screened hybridized to the 186-bp probe. Eight of the positive clones were colony purified and sequenced. The clone designated Ext3 in Fig. 2 *a* contains 690 bp of 5′ sequence including 22 bp of additional protein coding sequence and 668 bp of the 5′ untranslated region of p62 mRNA.

The 1.682 kbp cDNA sequence corresponding to nucleotides 608–2,295 of the p62 mRNA was assembled from clones cF2, cF7, and Ext 3 shown in Fig. 2 *a* and directionally cloned into the Xba I/Kpn I site of pGEM-3Zf(+). This construct, designated pGEMc62, contains the entire protein coding region of p62. The p62 expression plasmid pECEc62 was constructed by cloning the 1,682-bp Xba I/Kpn I cDNA from pGEMc62 into the Xba I/Sal I site located downstream of the SV-40 early promoter in the expression vector pECE-SK obtained from Frank McKeon (Harvard University). Nucleic acid and protein sequence analysis that is not specifically referenced in the text was performed using PC/GENE version 6.01 developed by Intelligenetics Inc. (Mountain View, CA).

RNA Analysis

Total RNA and poly(A)-mRNA from rat thyroid FRTL-5 cells and several rat tissues were prepared by guanidinium-phenol-chloroform extraction followed by oligo-dT chromatography (REDI column Type 7; Pharmacia Fine Chemicals, Piscataway, NJ). Electrophoresis of poly(A)-mRNA was performed in formaldehyde/formamide 0.9% agarose gels, and the RNA was transferred to nylon membranes as described previously (D'Onofrio et al., 1987). The 690-bp rat p62 cDNA was radiolabeled to a specific activity of 2×10^9 cpm/μg as described above and 5×10^6 cpm/ml of the labeled probe was added to a hybridization solution containing 5× SSPE, 50% formamide, 5× Denhardt's solution, and 10% dextran sulfate. The filters were hybridized for 24 h at 42°C, washed in 2× SSPE, and subjected to a final wash in 0.1× SSPE, 0.1% SDS for 20 min at 37°C.

Antibodies and Protein Analysis

The p62 antipeptide serum designated AS474 was prepared against the synthetic peptide TTPSTSLPSLATQTS (Applied Biosystems, Foster City, CA) which corresponds to the underlined amino acid residues Thr_{55} through Ser_{69} in Fig. 2 *b*. Before immunization the peptide was conjugated to Keyhold limpet hemocyanin (Calbiochem-Behring Corp., San Diego, CA) at a coupling ratio of 1 mg peptide/2 mg carrier using 0.5% glutaraldehyde in 0.1 M $NaPO_4$ buffer, pH 7.5. New Zealand White Rabbits were injected intracutaneously in the lower abdomen every 2 wk with 200 μg of the peptide conjugate in Freund's complete adjuvant. The rabbits were bled every 2 wk and the serum of the sixth bleed was used in the experiments described. The mouse mAb TIB117 (ATCC TIB 117) against the SV40 large T antigen was obtained from the American Type Culture Collection (Rockville, MD). Rat liver nuclei were isolated according to the procedure described by Blobel and Potter (1966). SDS-PAGE was performed on a reducing 9% acrylamide/bisacrylamide (29.2:0.8) separating gel that was blotted to nitrocellulose. Protein immunoblots were incubated with AS474 (1:1,000) or with preimmune serum (1:1,000). Immunoreactivity was detected using biotinylated affinity-purified goat anti-rabbit IgG and Vectastain AP reagent (Vector Laboratories, Inc., Burlingame, CA).

Transfection of Cultured Cells

African Green monkey kidney COS-1 cells (ATCC CRL 1650), CV-1 cells (ATCC CCL 70), and normal rat kidney (NRK) fibroblast cells (ATCC CRL 1570) from the American Type Culture Collection were maintained in DME containing 10% FBS (Gibco Laboratories, Grand Island, NY) at 37°C in 5% CO_2. 16 h before transfection, cells were split to 10^4 cells/ml and transferred to single-chambered LABTEX glass culture slides (PGC Scientific, Gaithersburg, MD). DNA was introduced into cells by cationic liposome-mediated transfection using Lipofectin (Bethesda Research Laboratories) at a ratio of 1 μg Lipofectin/0.5 μg DNA per 10^4 cells or by $CaPO_4$-mediated transfection (1 μg DNA per 10^4 cells) as described by van der Eb and Graham (1980), except that the DNA was removed from the cells after only 4.5 h without glycerol shock. Cells transfected with the p62 expression vector pECEc62 are designated *trans*COS-1. Untransfected COS-1 cells are designated wtCOS-1. CV-1 cells cotransfected with SV40 viral DNA (Bethesda Research Laboratories) and pECEc62 are designated *trans*CV-1 cells.

Immunofluorescence and EM

For immunofluorescence, cells grown on glass slides for 18 h were fixed in 2% formaldehyde for 20 min, permeabilized for 5 min in methanol at −20°C, and incubated in primary antiserum diluted in PBS containing 0.05% Tween-20 (TPBS) for 1 h at room temperature. The cells were rinsed three times in TPBS and incubated in the appropriate TRITC or FITC-conjugated affinity-purified secondary antibody (Jackson Immunologicals, West Grove, PA) for 30 min. Cells stained with succinylated FITC-WGA (Vector Laboratories, Inc.) were incubated with 5 μg/ml of the lectin in TPBS for 30 min. Chromatin was stained using 1 μg/ml bisbenzimide (Hoescht stain; Sigma Chemical Co., St. Louis, MO) for 1 min. Slides were mounted in fluid containing 90% glycerol, 10% PBS, and 0.1% 1,4-phenylenediamine (Aldrich Chemical Co., Milwaukee, WI) and photographed using a Zeiss IM inverted fluorescence microscope equipped with a 100× planapochromatic objective and Kodak Tri-X pan ISO 400/27° print film. Immunoferritin antibody labeling was done using the ferritin bridge method after glutaraldehyde fixation as described previously (Park et al., 1987).

Results

Construction of a Full-length p62 cDNA

The 690-bp cDNA, designated cB5 in Fig. 2 *a*, encodes 155 amino acids of the carboxy-terminal region of rat p62 and 218 nucleotides of 3′ nontranslated sequence (D'Onofrio et al., 1988). To determine the size of the p62 mRNA, the cB5 clone was used to probe a Northern blot of poly(A)-mRNA isolated from FRTL-5 cells and selected rat tissues. A single transcript at 2.9 kb was detected in FRTL-5 cells and was relatively abundant in rat liver, lung, and ovary (Fig. 1). To obtain the entire p62 coding sequence, the FRTL-5 phage library was rescreened using cB5. Eight p62 clones were isolated including clones cF2 and cF7 shown in Fig. 2 *a*. The clone Ext3, which contains the remaining amino-terminal sequence of p62, was isolated from a p62 5′ extension library as described in Materials and Methods. The full-length p62 cDNA was constructed from clones Ext3, cF2, and cF7 diagrammed in Fig. 2 *a*. The complete sequence of rat p62 mRNA is 2,941 nucleotides long (Fig. 2 *b*), which is in

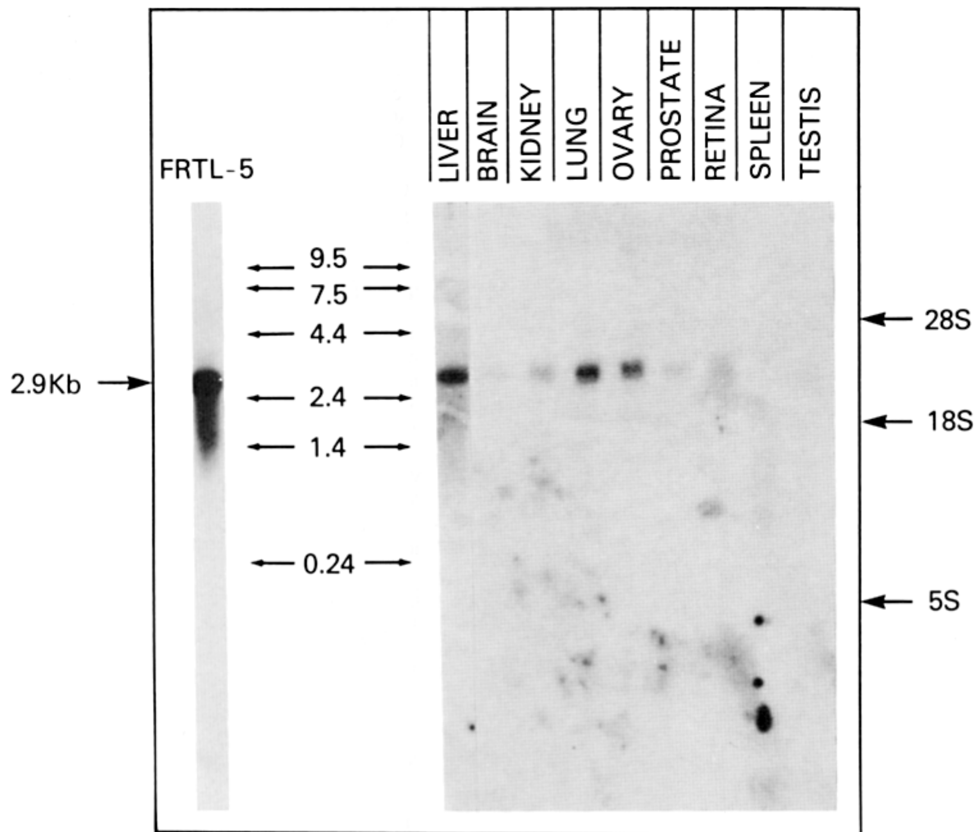


Figure 1. Expression of p62 in selected rat tissues. Poly(A)-mRNA was isolated from FRTL-5 cells and from rat tissue as described in Materials and Methods. 10 μg of each RNA was applied to a 0.9% agarose gel, separated, and blotted onto nylon membrane. The Northern blot was probed with the rat p62 cDNA clone cB5. A single band at ~2.9 kb was detected. The location of 28S, 18S, and 5S rRNA are shown on the right. The location of RNA size markers (in kilobase) are shown down the center.

agreement with the size determined by Northern analysis (Fig. 1). The 651 nucleotides of the 3′ nontranslated region contain a consensus polyadenylation signal located between nucleotide positions 2,931 through 2,936.

p62 Shows Little Similarity to Known Proteins

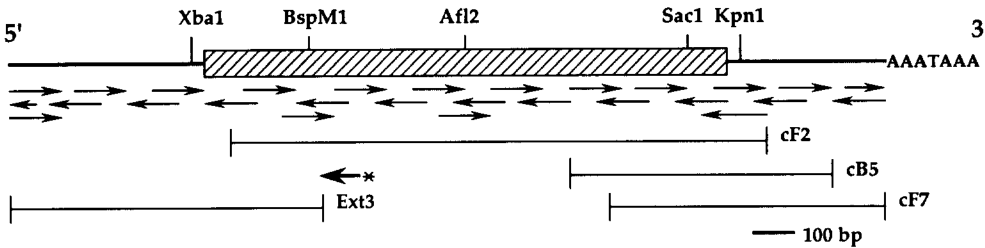
The amino-terminal region of p62 contains a single methionine codon, designated Met_I in Fig. 2 *b*. The nucleotide sequence surrounding Met_I (CAGCCATGAG) has a high degree of homology to the reported consensus sequence for eukaryotic translation initiation (CCGCCATGG) (Kozak, 1986). The 2,941 kb p62 mRNA encodes a protein with a calculated molecular mass of 54.3 kD (Fig. 2 *b*). No signal sequence for ER or nuclear localization was identified. A search of the EMBL/GenBank Genetic Sequence Database (Bilofsky and Burks, 1988) revealed no significant nucleic acid sequence homology and only limited regions of sequence similarities at the protein level (see Discussion).

Detection of p62 Expression in COS-1 Cells

Expression of rat p62 in transfected COS-1 and CV-1 cells was assayed by immunofluorescence 18 h after transfection with pECEc62 DNA. Two previously described pore-specific mAbs CHON-211 (Park et al., 1987) and RL-1 (Snow et al., 1987) detected rat p62 expression in transfected COS-1 (*trans*COS-1) cells, but both antibodies also stained the nuclear envelope of untransfected monkey cells (wtCOS-1) (Starr and Hanover, 1990). We sought, therefore, to develop an immunofluorescent detection system for rat p62 that was species specific. Our strategy was to prepare rabbit antipeptide sera against several peptide sequences of rat p62. The resultant antipeptide sera were screened for their ability to bind rat p62 without recognizing monkey nuclear pore proteins. The rabbit antipeptide serum designated AS474, generated against the 15-residue amino-terminal peptide underlined in Fig. 2 *b*, satisfied both criteria. AS474 recognized p62 on protein immunoblots of isolated rat liver nuclei (Fig. 3, lane *1*) as well as protein bands at ~120–130 kD which may represent dimers of p62. AS474 stained the nuclear envelope of NRK cells by immunofluorescence (Fig. 4 *a*), but did not stain the nuclear envelope of wtCOS-1 cells (Fig. 4 *c*).

Immunofluorescent staining of *trans*COS-1 cells with AS474 revealed diverse patterns of expression. AS474 stained the nuclear envelope of *trans*COS-1 cells expressing rat p62 with a fine punctate pattern consistent with pore binding (Fig. 4, *e* and *g*, Fig. 5, *a* and *c*, and Fig. 7, *a*, *d*, and *g*). In most *trans*COS-1 cells, AS474 also recognized phase-dense cytoplasmic aggregates and showed diffuse cytoplasmic staining (Fig. 4, *e* and *f*). The cytoplasmic aggregates were often located in a perinuclear position (Fig. 4, *g* and *h*, and Fig. 5, *a* and *c*). Although the perinuclear staining pattern was reminiscent of the immunofluorescent staining of proteins localized to the Golgi complex, staining these cells with FITC-ricin (a galactose-binding Golgi membrane marker) showed that the perinuclear structures recognized by AS474 and the Golgi complex were often but not always in the same region of the cell (data not shown). We often observed all three patterns of expression (i.e., staining of the nuclear envelope, cytoplasmic aggregates, and diffuse cytoplasmic staining) in the same *trans*COS-1 cell. We have examined transfected cells by EM using immuno-ferritin labeling with AS474 to

a



b

```
CATTCTGATATATTCCCATTCAACATGGATATGTGTCTATGGTACTATGTGGGCTATCCAACTTTGCATC
TTGGTCTCATTGTCCCACACCAGTGTGGGGGGAATTGCCCTGGGAAGATGATGCAATGTTCCCTGTCATC
ATCATCGCCATATAAGCTTTGTATATGGATGTACCTGGATGACTCTGCCTGGAGCGGGCGTAAGTCAAAA
CTTTGAGAGTCTAGTGCTTTTGCATGGGATGGACTGTGGCTGCTCTTGGCTTATGGAGTGCTTGGTCTCT
TTTCCCCTTGTCCCTCCCAACCACTACTGGTTTCCACCTAGAGAAAACCCCAAGTAACATCTGATTAGCT
GACCTGGGTCTCCACTTATCCCTAAGCACTGGGCTCTCATACCCACCCTTGTGTCCTGGTTCCAACTCAC
TAGAGTGCCTTCTTGTGAACTGCCTCAGACTTCCACCTTGGTAGTGGGACGTGCGCCTGCCCAGGAGCAC
TCAACTCCCATTGGTTTACTGGAGTTGGTCATGTTAGCAAAGGAAAGAGCAGGGGTGCTGCCAGCTCTGC
ATCCAGCCCACTACCCTCCACACCACTGAGCTTGCCTGTCCCAAACTGCCATCTGACCATGTCACTTCCT
CTGCTCTAGACATCTTCCCAAGGTTGCAGACGCTGACGGATTTGCCCCAGAGTCGAAGTAACTGCCACCT

                                                                    770
GGAAATTCCACAGCCATGAGTGGGTTTAACTTTGGAGGCACCGGGGCTCCTGCTGGCGGCTTTACATTTG
                M  S  G  F  N  F  G  G  T  G  A  P  A  G  G  F  T  F
1

                                                                    840
GGACCGCAAAGACTGCGACCACCACACCCGCCACTGGCTTTTCCTTCTCTGCTTCTGGCACCGGCACCGG
G  T  A  K  T  A  T  T  T  P  A  T  G  F  S  F  S  A  S  G  T  G  T  G
19
                                                                    910
AGGGTTTAATTTGGGACTCCCTCCCAGCCAGCTGCGACCACCCCTTCCACCAGCCTCTTCTCACTTGCC
  G  F  N  F  G  T  P  S  Q  P  A  A  T  T  P  S  T  S  L  F  S  L  A
43
                                                                    980
ACACAGACCTCAACCACACAGACCCCAGGATTCAACTTTGGAACAACACCTGCTTCTGGGGGAACAGGCT
 T  Q  T  S  T  T  Q  T  P  G  F  N  F  G  T  T  P  A  S  G  G  T  G
66
                                                                   1050
TCTCCCTGGGGATCAGCACCCCAAAGCTCAGCCTAAGTAGCACAGCCGCCACACCAGCCACAGCCAACAC
F  S  L  G  I  S  T  P  K  L  S  L  S  S  T  A  A  T  P  A  T  A  N  T
89
                                                                   1120
TGGCAGCTTTGGGCTTGGCAGCAGTACTCTTACCAATGCCATCTCAGGTGCCAGCACCTCCAGCCAGGGG
  G  S  F  G  L  G  S  S  T  L  T  N  A  I  S  G  A  S  T  S  S  Q  G
113                                                                1190
ACAGCCCCCACTGGCTTTGTCTTTGGCTCTTCTACCACCTCTGCTCCGTCCACCGGCACCACGGGATTCT
 T  A  P  T  G  F  V  F  G  S  S  T  T  S  A  P  S  T  G  T  T  G  F
136
                                                                   1260
CATTCACCAGTGGCAGTGCATCCCAGCCTGGAGCCTCCGGCTTCAACATTGGCTCTGTGGGTAGTTTGGC
S  F  T  S  G  S  A  S  Q  P  G  A  S  G  F  N  I  G  S  V  G  S  L  A
159
                                                                   1330
CCAGCCCACAGCACTGTCTGGCTCTCCCTTCACTCCAGCCACTCTGGCGACAACTACGGCAGGAGCAACA
  Q  P  T  A  L  S  G  S  P  F  T  P  A  T  L  A  T  T  T  A  G  A  T
183
                                                                   1400
CAGCCAGCTGCTGCTACACCCACTGCTGCCACCACCAGTGCAGGGTCTACACTGTTTGCTTCAATAGCTG
 Q  P  A  A  A  T  P  T  A  A  T  T  S  A  G  S  T  L  F  A  S  I  A
206
                                                                   1470
CTGCTCCTGCCTCATCCAGTACTACAGTGCTTTCCCTCTCAGCTCCAGCGACAACTGCAGCCACTCCTAC
A  A  P  A  S  S  S  T  T  V  L  S  L  S  A  P  A  T  T  A  A  T  P  T
229
                                                                   1540
TGCTGGGACTTTGGGCTTCAGCCTCAAGGCCCCTGGAGCAGCTCCTGGTGCCTCCACCACCAGCACCACC
  A  G  T  L  G  F  S  L  K  A  P  G  A  A  P  G  A  S  T  T  S  T  T
253
```

Figure 2. Cloning strategy and primary sequence of p62. (*a*) Cloning and sequencing of rat p62 mRNA. Arrows indicate the location of synthetic primers used to sequence the full-length clone. The partial p62 cDNA clones cF2, cB5, and cF7 were isolated from an FRTL-5 library. Ext3 was isolated from a 5' extension library constructed using a 32mer synthetic primer (<-*). The protein coding region of p62 is crosshatched. The Xba I/Kpn I fragment of the p62 cDNA was subcloned into the eukaryotic expression vectór pECE-SK for transient expression of rat p62 in COS-1 cells. (*b*) Primary sequence of p62. The 2,941 nucleotides of p62 mRNA are numbered on the right above each line. The polyadenylation site at the 3' end of the sequence is in bold type. The deduced amino acid sequence of p62 starting at Met_1 is numbered on the left below each line. The peptide sequence used to generate the antipeptide serum AS474 is underlined. These sequence data are available from EMBL/GenBank/DDBJ under accession number X52583.

```
                                                                    1610
ACTACTACTACCACCACCACTACTACTGCCTCCACCTCCTCCTCCACCACGACCACTGGCTTTGCCTTAA
 T  T  T  T  T  T  T  T  T  A  S  T  S  S  S  T  T  T  T  G  F  A  L
276
                                                                    1680
GCCTGAAACCCCTGGTGCCAGCTGGCCCCAGCAGCGTGGCAGCTACTGCTCTGCCTGCCTCCAGCACAGC
S  L  K  P  L  V  P  A  G  P  S  S  V  A  A  T  A  L  P  A  S  S  T  A
299
                                                                    1750
AGTTGGGACTACGACGGGTCCAGCAATGACCTACGCGCAGCTGGAAAGCCTGATCAACAAGTGGAGTCTG
  V  G  T  T  T  G  P  A  M  T  Y  A  Q  L  E  S  L  I  N  K  W  S  L
323
                                                                    1820
GAGCTGGAGGACCAGGAGCGGCACTTCCTGCAGCAGGCCACGCAGGTCAATGCCTGGGACCGCACACTGA
 E  L  E  D  Q  E  R  H  F  L  Q  Q  A  T  Q  V  N  A  W  D  R  T  L
346
                                                                    1890
TTGAGAATGGGGAGAAGATCACCAGCCTGCACCGAGAGGTGGAAAAGGTGAAGCTGGATCAGAAGCGTCT
I  E  N  G  E  K  I  T  S  L  H  R  E  V  E  K  V  K  L  D  Q  K  R  L
369
                                                                    1960
GGACCAGGAGCTGGACTTTATCCTGTCACAGCAGAAGGAGCTGGAAGACCTTCTGAGCCCGTTAGAGGAG
  D  Q  E  L  D  F  I  L  S  Q  Q  K  E  L  E  D  L  L  S  P  L  E  E
393
                                                                    2030
TCAGTGAAGGAGCAAAGTGGCACCATCTACCTTCAGCATGCTGATGAGGAGCGTGAGAAGACCTACAAGC
 S  V  K  E  Q  S  G  T  I  Y  L  Q  H  A  D  E  E  R  E  K  T  Y  K
416
                                                                    2100
TGGCTGAGAACATCGATGCTCAGCTCAAGCGCATGGCCCAGGACCTCAAGGACATTATTGAGCACCTGAA
L  A  E  N  I  D  A  Q  L  K  R  M  A  Q  D  L  K  D  I  I  E  H  L  N
439
                                                                    2170
CATGGCTGGTGGCCCTGCAGACACCAGTGACCCACTGCAGCAGATCTGCAAGATCCTCAATGCACACATG
  M  A  G  G  P  A  D  T  S  D  P  L  Q  Q  I  C  K  I  L  N  A  H  M
463
                                                                    2240
GACTCCCTTCAGTGGGTGGACCAGAGCTCTGCCCTGCTGCAGAGGAGGGTGGAGGAGGCCAGCCGTGTGT
 D  S  L  Q  W  V  D  Q  S  S  A  L  L  Q  R  R  V  E  E  A  S  R  V
486
                                                                    2310
GTGAGAGCCGGCGCAAGGAACAGGAACGCAGCCTGCGCATTGCCTTTGACTAGCCATACATGAGTTGGGG
C  E  S  R  R  K  E  Q  E  R  S  L  R  I  A  F  D
509

GTGGGGGTACCTCTTTTGTTGTAGGGCTGCTGGCTTTGGGGTGGTTGTATTGTTGAGAAAAATGCTCTTT
GGTTTGGTTCCCTGTTATCAGAAGTGGTTTGCGAAGGCATGATGTCATGGATGTAATAACCTGTGTGCTC
TGGAGACAGGTGTGAAGGTCACTGAACTCTAGGCCAGCATTTAAAGAACAAGGGCAAACTCATCCTTTAG
CCACTCCAATACAACTGTGACAGTGGGATGGCATGCTCCAGAGATGTTGGCCAAGAAACTCACCTGGAAT
GGAGACATGGCTTCTGTGCTTCCCTGCTTCTCTAACTCCCATCTAACCCATGTCCTCCAGAGCCTGTGCT
GCACAAAGCAGAGTCACCAAGGGAGACTGAGCTTAACACCGGTTTGCACTGGCTACCCCTCTCACCCCTC
TCTTGTGAGCTTCTGTGGCAGGGACTGACTGCTGATAGTTCTTTGAGAACAGAGTGGCTGCGGATTTGCT
TCTGAATGTCCCAGTGTCAAACCTTAGATTCATAGATAGTGTTGAAGGCTGGCAGTTATGTCTGGTCTAT
TTTATTTCTCCACACTCTACTGTAGCCTCTATGGTCCCTGATCAAAACATTGCTTTTCTAAATAAAGAA
AA                                                                  2941
```

(Figure 2 b continued)

determine the subcellular location of p62 and the nature of the protein aggregates (see Materials and Methods). This proved difficult since only 10–15% of transfected cells express p62. Of the cells examined, the bulk of the ferritin labeling was found on perinuclear aggregates with occasional labeling of an NPC (data not shown). The cytoplasmic aggregates showed no defined structure. The ferritin-labeled cytoplasmic aggregates were never observed in nontransfected cells.

To determine whether the cytoplasmic aggregates recognized by AS474 in *trans*COS-1 cells contain GlcNAc, we double labeled *trans*COS-1 cells with AS474 and fluorescein-conjugated succinyl-WGA, a lectin that specifically binds terminal GlcNAc residues on glycoproteins (Monsigny et al., 1980). Fig. 5 shows examples of two *trans*COS-1 cells double labeled with AS474 (Fig. 5, *a* and *c*) and FITC succinyl-WGA (Fig. 5, *b* and *d*). The panels in Fig. 5 also contain COS-1 cells located in the upper left corner of each field, which are not expressing rat p62 and therefore do not react with AS474. Succinyl-WGA stained both the perinuclear material (Fig. 5 *b*) and the cytoplasmic aggregates (Fig. 5 *d*) recognized by AS474. The binding of succinyl-WGA to the cytoplasmic aggregates was abolished when *trans*COS-1 cells were incubated with 0.5 M GlcNAc during lectin staining. Detection of GlcNAc in the perinuclear aggregates of rat p62 was difficult because this material is often colocalized with the Golgi complex of the cell and succinyl-WGA binds to Golgi membranes in both wtCOS-1 and *trans*COS-1 cells

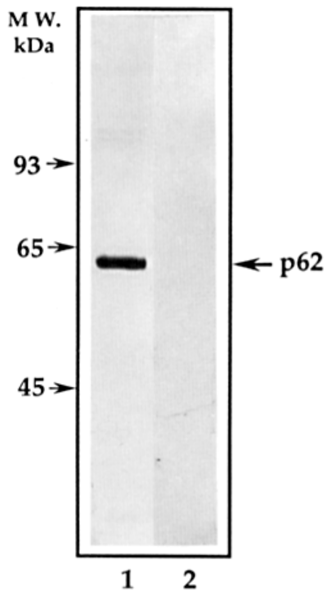


Figure 3. Detection of p62 in isolated rat nuclei by AS474. Shown is a protein immunoblot of isolated rat liver nuclei stained with the antipeptide serum AS474 (1:2,000) (lane *1*) and rabbit preimmune sera (1:2,000) (lane *2*). AS474 stains a band at 62 kD in rat liver nuclei. The migration pattern of molecular mass standards is shown on the left.

(Fig. 5, *a* and *b*). However, succinyl-WGA binding to the cytoplasmic aggregates recognized by AS474 (Fig. 5, *c* and *d*) strongly suggests that rat p62 expressed in COS-1 cells is modified by O-linked GlcNAc. This conclusion is supported by our findings that a significant fraction of rat p62 expressed in *trans*COS-1 cells binds to succinyl-WGA affinity resin and comigrates at 62 kD on SDS-PAGE with glycosylated p62 isolated from rat liver nuclei (Starr and Hanover, 1990).

Rat p62 Associates with the Mitotic Spindle during Mitosis

In *trans*COS-1 cells undergoing mitosis, AS474 shows intense staining in the region of the mitotic spindle poles suggesting that p62 may become associated with these structures during cell division. Fig. 6 shows *trans*COS-1 cells at three different stages of mitosis double labeled with AS474 and Hoechst stain to visualize the chromatin of the dividing cell. During metaphase, p62 concentrates at the poles of the mitotic spindle and is observed migrating with the spindle aster during anaphase (Fig. 6, *a–c*). During late anaphase, p62 also appears to associate with the outer sheath surrounding the spindle fibers (Fig. 6, *d–f*). In early telophase, aggregates of p62 are found in both daughter cells in the region of condensed chromatin (Fig. 6, *g–i*). The aggregates become excluded from the daughter nuclei in late telophase.

Nuclear Accumulation of SV40 Large T Antigen Is Normal in CV-1 Cells Expressing p62

Because GlcNAc-modified pore glycoproteins have been implicated in mediating transport through the pore, we investigated whether the observed association of rat p62 with the nuclear envelope of transfected cells significantly interferes with nuclear protein transport. COS-1 cells synthesize SV40 large T antigen which contains a nuclear localization signal sequence and is transported into the nucleus. The mAb TIB117, which binds the SV40 large T antigen, was used to determine the distribution of T antigen in cultured cells. TIB117 staining of the nuclei of *trans*COS-1 cells was consistent with the nuclear localization of SV40 large T antigen (Fig. 7, *a–c*), however, endogenous T antigen was probably present in the nucleus of these cells before transfection. To insure that the synthesis and nuclear accumulation of SV40 T antigen occurred during a period of rat p62 synthesis and association with the nuclear envelope, we cotransfected CV-1 cells, which do not express any endogenous T antigen with pECEc62 and SV40 viral DNA (*trans*CV-1). In *trans*CV-1 cells, both SV40 T antigen and rat p62 are synthesized during the same period. *Trans*CV-1 cells show nuclear rim staining with AS474 consistent with the association of rat p62 with the nuclear envelope (Fig. 7, *d* and *g*) and TIB117 staining of the nucleus resulting from the nuclear import of SV40 T antigen (Fig. 7, *e* and *h*). The CV-1 cell in the upper right corner of *d–f* is not expressing SV40 T antigen or rat p62. Taken together, these data demonstrate that the overexpression of rat p62 does not significantly interfere with protein transport through the pore.

Discussion

Rat p62 has a calculated molecular mass of 54.3 kD, but native p62 from isolated rat liver nuclei migrates at 62 kD on SDS-PAGE (Snow et al., 1987; Davis and Blobel, 1987; Hanover et al., 1987). At least two factors may explain the reduced mobility of p62 on SDS-PAGE. First, rat p62 has been estimated to contain 8–10 residues of O-linked GlcNAc (Holt et al., 1987), which would contribute ∼2,200 D to the molecular mass of the protein. This estimate of the contribution of O-linked GlcNAc to the molecular mass of p62 is consistent with our finding that nonglycosylated p62 synthesized in vitro using wheat germ extract migrates at 59 kD on SDS-PAGE (Starr and Hanover, 1990). To date, no additional posttranslational modifications have been described for p62. Second, the derived amino acid sequence of p62 has an acidic pK of 4.8 with clusters of glutamate and aspartate residues in the carboxy-terminal region of the protein. Acidic nuclear proteins that contain tracts of polyglutamic acid have also been found to exhibit reduced mobility on SDS-PAGE (Kleinschmidt et al., 1985). The reduced mobility of p62 on SDS-PAGE may be partially due to the decreased binding of the cationic detergent to these negatively charged regions of the protein.

Differential extraction of the nuclear envelope with detergents and salt suggests that the NPC is composed of both peripheral and integral membrane proteins (Davis and Blobel, 1986; Hanover et al., 1987; Snow et al., 1987). The hydrophobicity profile of p62 did not suggest the presence of membrane spanning domains. Moreover, using the algorithm of Klein et al. (1985), p62 was classified as a peripheral membrane protein which is consistent with reports that p62 can be extracted from the nuclear envelope with 0.5 M NaCl (Davis and Blobel, 1986; Hanover et al., 1987; Snow et al., 1987). It seems likely, therefore, that p62 is anchored to the pore complex through association with other pore proteins.

p62 Has Three Distinct Sequence Regions

Based on the primary sequence, p62 can be roughly divided into three regions. The amino-terminal region of p62 is

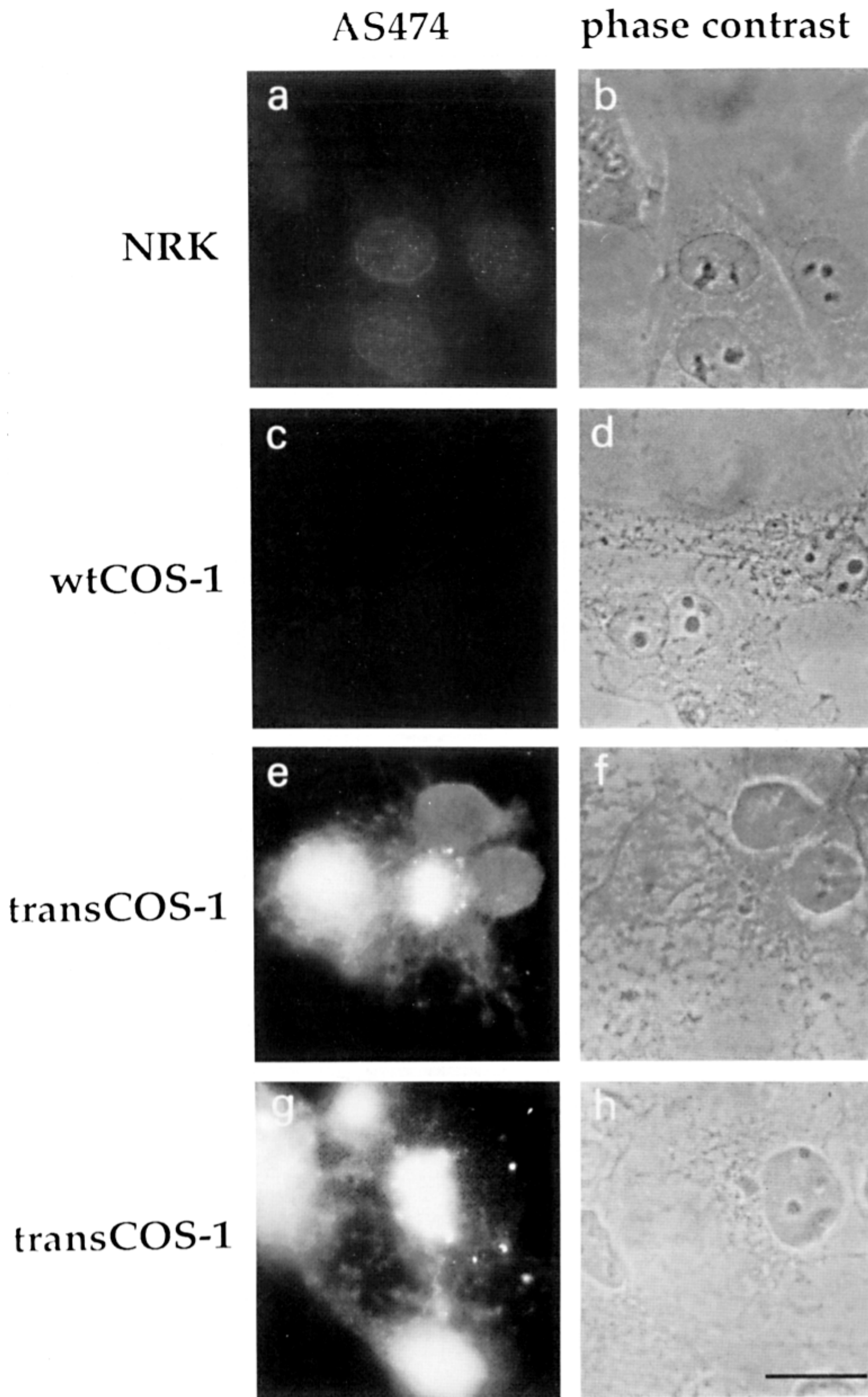


Figure 4. Indirect immunofluorescent detection of rat p62 expression in transfected cells. Panels show the paired immunofluorescent and corresponding phase images of NRK cells (*a* and *b*); wtCOS-1 cells (*c* and *d*); and *trans*COS-1 cells (*e–h*) stained with the antipeptide serum AS474 (1:1,000) followed by TRITC-conjugated AP goat anti-rabbit IgG (1: 20). AS474 stains the nuclei of NRK cells with a finely punctate pattern (*a*). AS474 does not stain wtCOS-1 cells (*c*). In addition to staining the nuclei of *trans*COS-1 cells, AS474 stains peripheral regions of the cytoplasm as well as perinuclear and cytoplasmic aggregates (*e–h*). Bar, 10 μm.

highly repetitive; the tetrapeptide Thr-Pro-Ala-Thr is repeated four times with eight additional repeats containing conservative amino acid substitutions (Thr-Pro is always conserved). The peptide repeats of p62 are somewhat reminiscent of the repeating Gly-Pro-X sequence in the alpha chain of collagen. A central region containing an almost uninterrupted stretch of 25 serine and threonine residues (between Ser_{270} and Thr_{294} in Fig. 2 *b*) may have important im-

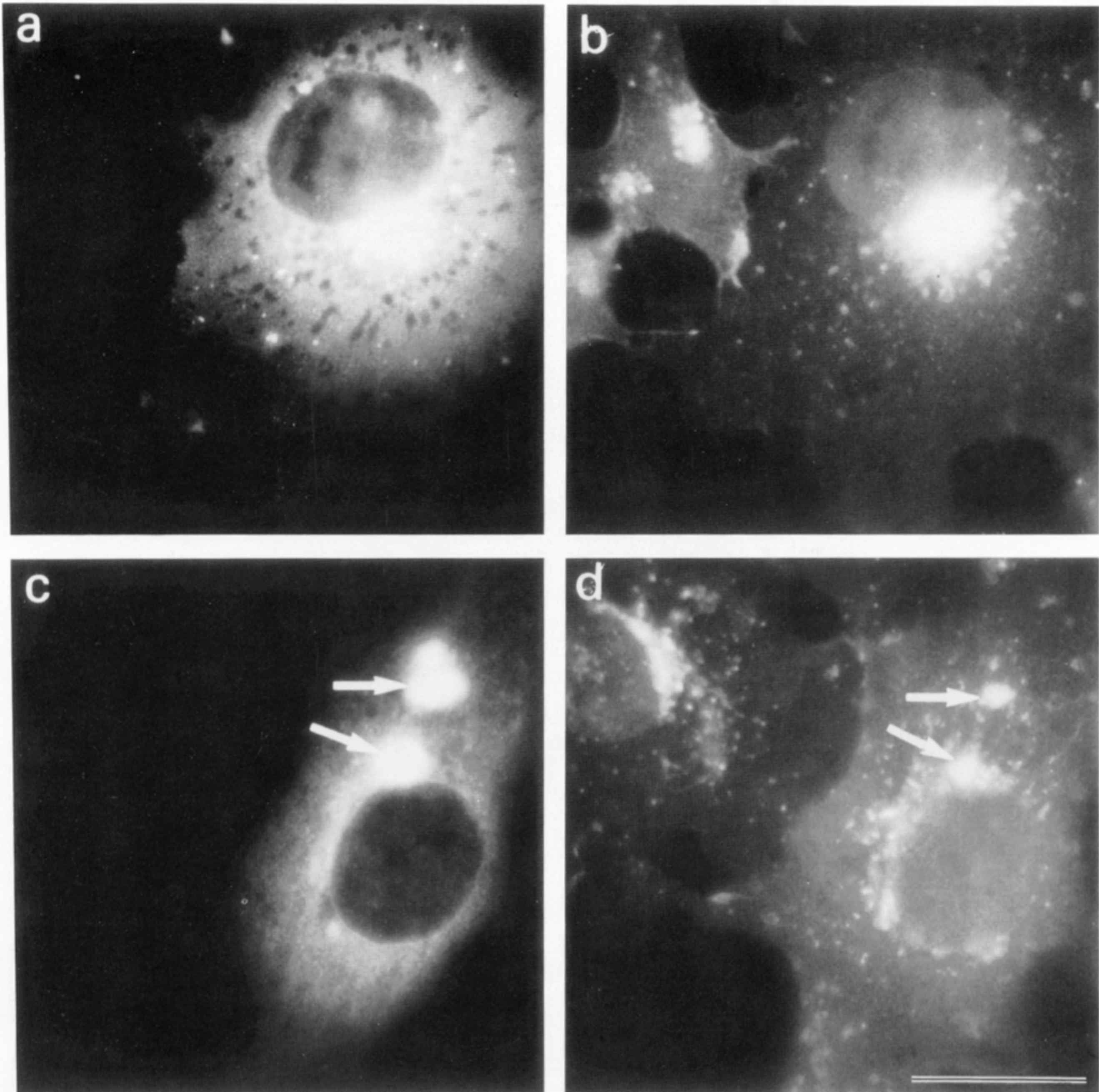


Figure 5. WGA binding to rat p62 in transfected cells. *Trans*COS-1 cells were stained with both AS474 and WGA. Panels show indirect immunofluorescence of AS474 (1:1,000) using TRITC-conjugated AP goat anti-rabbit IgG (1:20) (*a* and *c*) and the corresponding fluorescein images from FITC-conjugated succinyl-WGA (1:400) staining (*b* and *d*). Transfected cells often present a perinuclear staining pattern (*a*) that is coincident with the succinyl-WGA staining (*b*). Corresponding arrows in the cell shown in *c* and *d* indicate that cytoplasmic aggregates react with both AS474 (*c*) and WGA (*d*). Bar, 10 μm.

plications regarding the glycosylation of p62. Regions rich in serine and threonine residues have also been found in the transcription factor Spl which contains O-linked GlcNAc (Jackson and Tjian, 1988). A high percentage of hydroxy amino acids may be a common feature of O-linked GlcNAc-modified proteins. The third sequence motif, located within the carboxy-terminal region of p62, is a series of hydrophobic heptad repeats (between amino acid residues 350 and 525 in Fig. 2 *b*) that are similar to those of the k-m-e-f class of coiled-coil filamentous proteins (e.g., myosin, lamins, and intermediate filaments). Secondary structure predictions based on the algorithm of Garnier et al. (1978) suggest that

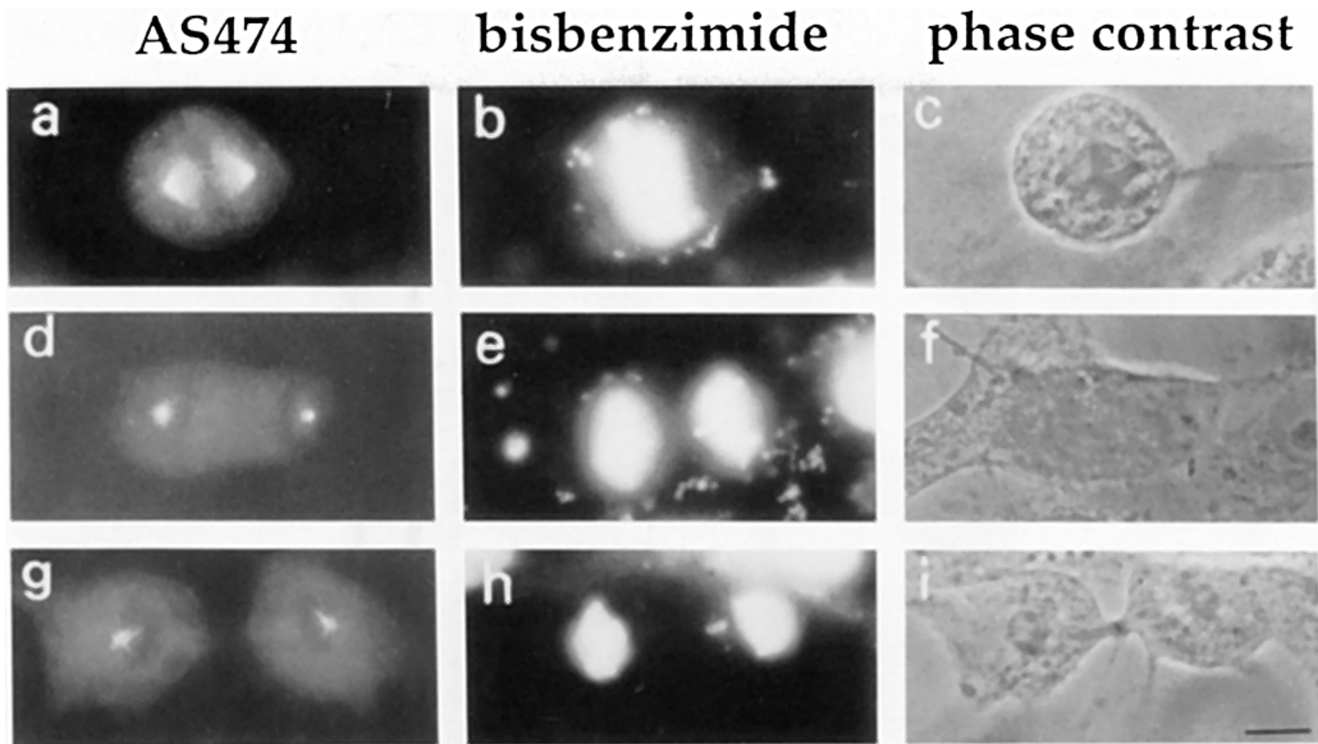


Figure 6. Distribution of rat p62 during mitosis in COS-1 cells. Corresponding panels of *trans*COS-1 cells at various stages of mitosis stained with AS474 (1:1,000) (*a*, *d*, and *g*), Hoechst stained to visualize the chromosomes (*b*, *e*, and *h*) and the phase images (*c*, *f*, and *i*). Shown are cells in early anaphase (*a–c*); late anaphase (*d–f*); and telophase (*g–i*). Rat p62 concentrates at the astral centers of the mitotic spindle at metaphase and early anaphase (*a*), migrates with the poles of the mitotic spindle during anaphase (*d*), and remains associated with the residual aster in each daughter cell during telophase (*g*). Bar, 10 μm.

this carboxy-terminal region contains four alpha-helical segments each ~30–40 residues long beginning around residue 375 and extending to the carboxy-terminus of the protein. As with myosin, the arrangement of the heptad repeats of p62 is consistent with the formation of alpha helices with apposing charged and hydrophobic surfaces. These surfaces may interact in vivo to form interchain or intrachain helical bundles. Alternatively, this region of p62 may associate with complementary regions on other pore proteins to form a multiprotein complex.

Heterologous Expression of p62

Immunofluorescent staining of transfected cells expressing rat p62 suggests that p62 associates with the nuclear envelope. In addition to associating with the nuclear envelope, rat p62 also forms large phase dense protein aggregates in the cytoplasm of transfected cells. These aggregates show a vesicle-like morphology when stained with AS474 that is clearly visible when the focal plane is moved through the cell. The aggregates may represent filamentous bundles of p62 which may exclude the antibody from the interior of the bundle and cause a vesicular staining pattern.

In *trans*COS-1 cells undergoing cell division, we found that while p62 is detected in the cytoplasm, a significant fraction of the rat p62 concentrates at the poles of the mitotic spindle. This finding is in contrast to reports that during mitosis, endogenous pore proteins become dispersed in the cytoplasm after nuclear envelope breakdown (Park et al., 1987; Davis and Blobel, 1986; Snow et al., 1987). Indeed, the NRK cell located in the upper left corner of Fig. 4, *a* and *b*, which is in early anaphase, does not show staining of the mitotic spindle pole by immunofluorescence using AS474. Clearly, there are differences in the immunofluorescent staining pattern between endogenous p62 and rat p62 expressed in *trans*COS-1 cells. The rat p62 in *trans*COS-1 cells and endogenous p62 may behave differently during mitosis. It is possible that the localization of rat p62 observed in *trans*COS-1 cells results from the physical displacement of the cytoplasmic aggregates of overexpressed rat p62 brought about by the expanding mitotic apparatus. In that case, the pattern of p62 expression observed in dividing *trans*COS-1 cells would have little relationship to the actual in vivo behavior of endogenous p62. Alternatively, the pattern of rat p62 localization in dividing *trans*COS-1 cells may accurately reflect the normal behavior of endogenous p62; the immunofluorescent staining being greatly enhanced in *trans*COS-1 cells due to the large amounts of p62 present in these cells. Furthermore, the perinuclear localization of rat p62 in nondividing cells (Fig. 4, *e–h* and Fig. 5 *a*) is consistent with the position of the centrosomes that are capable of nucleating microtubules in interphase and mitotic cells (Vorbjev and Nadehzdina, 1987). Components of the nuclear envelope are known to associate with the mitotic apparatus (see Hepler and Wolniak, 1984 for review). For example, during the "closed" intranuclear mitosis of the budding yeast *Saccharomyces cerevisiae*, the spindle is anchored at each end to a dense plate on the inner surface of the nuclear envelope.

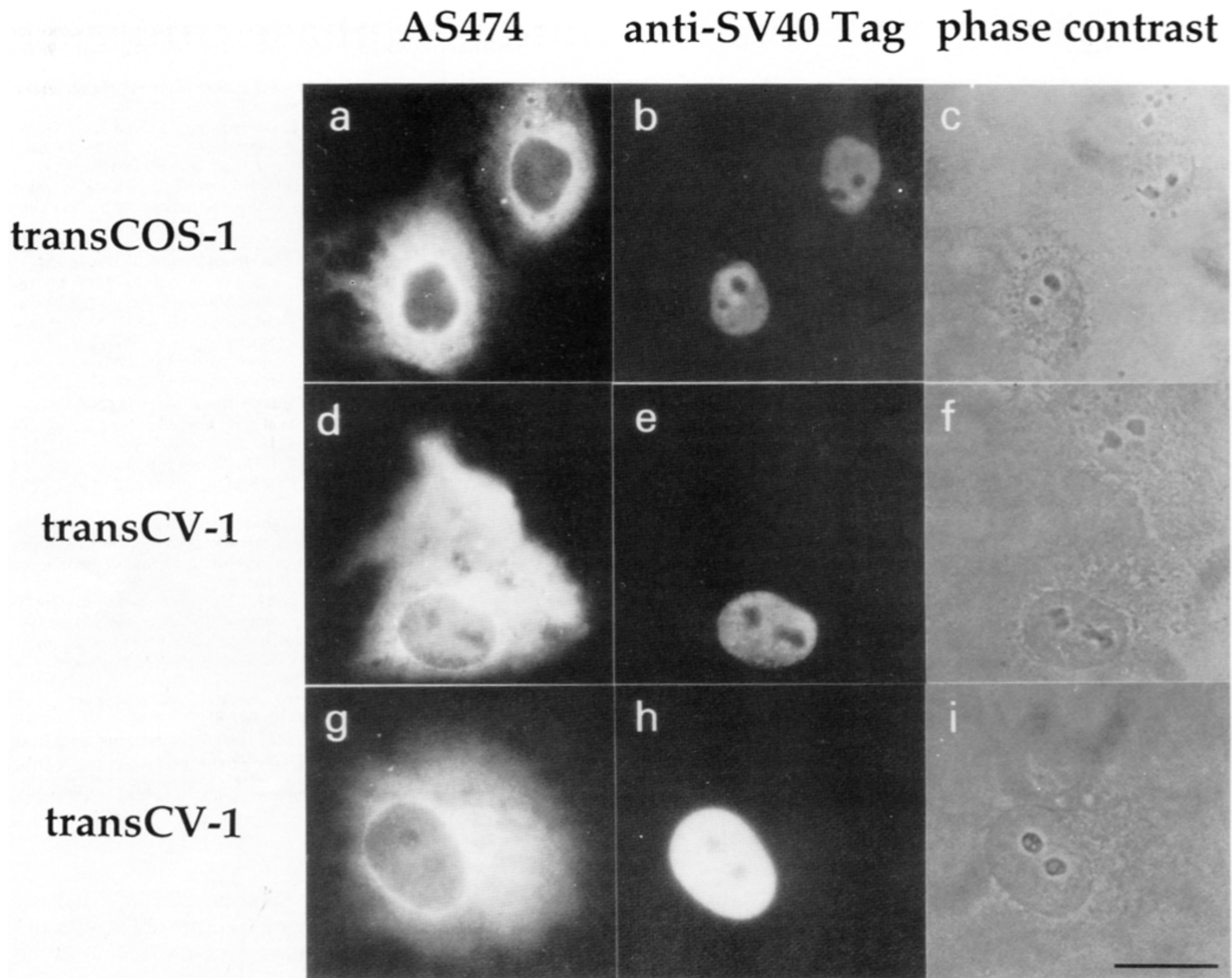


Figure 7. Nuclear accumulation of SV40 large T antigen in cells overexpressing rat p62. Rat p62 expression was detected using AS474 and SV40 large T antigen was detected using the mouse mAb TIB117. Shown in *a–c* are *trans*COS-1 cells expressing rat p62 that were double labeled with AS474 (1:1,000) and TIB117 (1:1,000) followed by TRITC-conjugated AP goat anti-rabbit IgG and FITC-conjugated AP goat anti-mouse IgG. Shown in *d–f* and *g–i* are CV-1 cells cotransfected with pECEc62 and SV40 DNA double labeled with AS474 to show rat p62 expression (*d* and *g*), and with TIB117 to show nuclear accumulation of SV40 T antigen (*e* and *h*). The CV-1 cell in the upper right corner of *d–f* is not expressing SV-40 T antigen or rat p62. The nuclear accumulation of SV40 T antigen is normal in cells expressing rat p62. Bar, 10 μm.

During "open" mitosis, the extranuclear spindle is surrounded by a sheath composed of overlapping membrane vesicles derived from the ER and nuclear envelope. It is possible that a subset of nuclear pore components remains associated with the nuclear envelope during mitosis. The association of nuclear pore proteins with the mitotic apparatus would assure that each daughter cell receives pore proteins after mitosis.

Role of p62 in Nuclear Pore Function and Assembly

Studies of lectin inhibition (Wolf et al., 1988; Finlay et al., 1987) and of antibody binding (Featherstone et al., 1987; Dabauvalle et al., 1987) suggest that O-linked GlcNAc containing nuclear pore glycoproteins are involved in transport through the pore. We have documented in both COS-1 and CV-1 cells that rat p62 becomes associated with the nuclear envelope without interfering with the nuclear accumulation of SV40 large T antigen. After submission of this paper, Finlay and Forbes (1990) reported that nuclear transport could be restored to transport defective reconstituted *Xenopus* nuclear pores by adding back rat pore glycoproteins. These in vitro results are consistent with our in vivo findings that the cross-species association of a heterologous pore glycoprotein, rat p62, with the nuclear envelope of monkey cells does not interfere with protein import. These results further suggest that pore glycoproteins from diverse species may be structurally and functionally interchangeable in the assembly of the NPC and in mediating protein import. Our findings also suggest that some incorporation or exchange of pore glycoproteins may occur in interphase cells that have an intact nuclear envelope.

We wish to thank Gilbert Ashwell, William Jakoby, and Rachel Myerowitz for critical reading of the manuscript. We would like to acknowledge Mark Willingham for the immuno-ferritin labeling and EM.

C. M. Starr is supported by a research associateship from the National Research Council.

Received for publication 11 October 1989 and in revised form 2 March 1990.

References

Blobel, G., and V. Potter. 1966. Nuclei from rat liver: isolation method that combines purity with high yield. *Science (Wash. DC).* 154:1662-1665.

Bilofsky, H. S., and C. Burks. 1988. The GenBank (R) genetic sequence data bank. *Nucleic Acid Res.* 16:1861-1864.

Dabauvalle, M.-C., R. Benavente, and N. Chaly. 1988. Monoclonal antibodies to a Mr 68000 pore complex glycoprotein interfere with nuclear uptake in *Xenopus* oocytes. *Chromosoma (Berl.).* 97:193-197.

Davis, L. I., and G. Blobel. 1986. Identification and characterization of a nuclear pore complex protein. *Cell.* 45:699-709.

Davis, L. I., and G. Blobel. 1987. Nuclear pore complex contains a family of glycoproteins that includes p62: glycosylation through a previously unidentified cellular pathway. *Proc. Natl. Acad. Sci. USA.* 84:7552-7556.

Deininger, P. L. 1987. Full-length cDNA clones: vector-primed cDNA synthesis. *Methods Enzymol.* 152:371-389.

D'Onofrio, M., C. M. Starr, M. K. Park, G. D. Holt, R. S. Haltiwanger, G. W. Hart, and J. A. Hanover. 1988. Partial cDNA sequence encoding a nuclear pore protein modified by O-linked N-acetylglucosamine. *Proc. Natl. Acad. Sci. USA.* 85:9595-9599.

Featherstone, C., M. K. Darby, and L. Gerace. 1988. A monoclonal antibody against the nuclear pore complex inhibits nucleocytoplasmic transport of protein and RNA in vivo. *J. Cell Biol.* 107:1289-1297.

Finlay, D. R., and D. J. Forbes. 1990. Reconstitution of biochemically altered nuclear pores: transport can be eliminated and restored. *Cell.* 60:17-29.

Finlay, D. R., D. D. Newmeyer, T. M. Price, and D. J. Forbes. 1987. Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores. *J. Cell Biol.* 104:189-200.

Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of a simple method for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97-120.

Gubler, U. 1987. Second strand cDNA synthesis: mRNA fragments as primers. *Methods Enzymol.* 152:330-335.

Hanover, J. A., C. Cohen, M. C. Willingham, and M. K. Park. 1987. O-linked N-acetylglucosamine is attached to proteins of the nuclear pore: evidence for cytoplasmic and nucleoplasmic glycoproteins. *J. Biol. Chem.* 262:9887-9894.

Hepler, P. K., and S. M. Wolniak. 1984. Membranes in the mitotic apparatus: their structure and function. *Intern. Rev. Cytol.* 90:169-238.

Holt, G. D., C. M. Snow, A. Senior, R. S. Haltiwanger, L. Gerace, and G. W. Hart. 1987. Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. *J. Cell Biol.* 104:1157-1164.

Jackson, S. P., and R. Tjian. 1988. O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell.* 55:125-133.

Klein, P., M. Kanehisa, and C. Delisi. 1985. The detection and classification of membrane spanning proteins. *Biochim. Biophys. Acta.* 815:468-476.

Kleinschmidt, J. A., C. Dingwall, G. Maier, and W. W. Franke. 1985. Molecular characterization of a karyophilic histone-binding protein: cDNA cloning, amino acid sequence and expression of nuclear protein N1/N2 of *Xenopus laevis. EMBO (Eur. Mol. Biol. Organ.) J.* 5:3547-3552.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiation codon that modulates translation by eukaryotic ribosome. *Cell.* 44:285-292.

Monsigny, M., A.-C. Roche, C. Sene, R. Maget-Dana, and F. Delmotte. 1980. Sugar-lectin interactions: how does wheat germ agglutinin bind sialoglycoconjugates. *Eur. J. Biochem.* 104:147-153.

Newmeyer, D. D., and D. J. Forbes. 1988. Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation. *Cell.* 52:641-653.

Park, M. K., M. D'Onofrio, M. C. Willingham, and J. A. Hanover. 1987. A monoclonal antibody against a family of nuclear pore proteins (nucleoporins): O-linked N-acetylglucosamine is part of the immunodeterminant. *Proc. Natl. Acad. Sci. USA.* 84:6462-6466.

Snow, C. M., A. Senior, and L. Gerace. 1987. Monoclonal antibodies identify a group of nuclear pore complex glycoproteins. *J. Cell Biol.* 104:1143-1156.

Starr, C. M., and J. A. Hanover. 1990. Glycosylation of nuclear pore protein p62: reticulocyte lysate catalyses O-linked N-acetylglucosamine addition *in vitro. J. Biol. Chem.* In press.

van der Eb, A. J., and F. L. Graham. 1980. Assay of transforming activity of tumor virus DNA. *Methods Enzymol.* 65:826-839.

Vorobjev, I. A., and E. S. Nadehzdina. 1987. The centrosome and its role in organizing microtubules. *Int. Rev. Cytol.* 106:227-284.

Wolff, B., M. C. Willingham, and J. A. Hanover. 1988. Nuclear protein import: specificity for transport across the nuclear pore. *Exp. Cell Res.* 178:318-334.