Journal of Cell Science 107, 631-638 (1994)
Printed in Great Britain © The Company of Biologists Limited 1994
JCS8392

Role of different domains in the self-association of rat nucleoporin p62

Folma Buss[1], Helen Kent[1], Murray Stewart[1,*], Susanne M. Bailer[2] and John A. Hanover[2]

[1]MRC Laboratory of Molecular Biology, Hills Rd, Cambridge CB2 2QH, UK
[2]National Institutes of Health, Building 10, Room 9B07, Bethesda, MD 20892, USA

*Author for correspondence

SUMMARY

We have expressed rat nucleoporin p62 cDNA in *Escherichia coli* to obtain material for structural and self-association studies. Electron microscopy and circular dichroism spectroscopy are consistent with a rod-shaped molecule with an α-helical coiled-coil domain at its C terminus and a cross-β structure at its N terminus, separated by a threonine-rich linker, which has a less-defined secondary structure. Electron microscopy and the solubility properties of fragments produced using thrombin and CNBr digestion indicate that p62 molecules associate to form linear chains and that a small region near the C terminus is an important determinant of assembly. This association may have important consequences for pore structure and function; for example, one way p62 could associate would be to form rings in nuclear pores that could function like barrel hoops.

Key words: nuclear pore, macromolecular assembly, structure

INTRODUCTION

Nuclear pores perforate the nuclear envelopes of eukaryotic cells and are the gateways for transport of large molecules between nucleus and cytoplasm (reviewed by Akey, 1991; Miller et al., 1991; Silver, 1991; Forbes, 1992; Stewart, 1992). The nuclear pores are huge macromolecular assemblies about 120 nm in diameter that have 8-fold rotational symmetry and a molecular mass of the order of 125,000 kDa (Reichelt et al., 1990; Hinshaw et al., 1992; Forbes, 1992; Stewart, 1992). A series of studies (reviewed by Forbes, 1992; Stewart, 1992) indicate that the pores are constructed from a central cylindrical 'plug-spoke' unit sandwiched between nucleoplasmic and cytoplasmic rings. Although it is likely that nuclear pores are constructed from as many as 100 different proteins, there is only limited information on their structure and arrangement. The detailed molecular characterisation of nuclear pores has been frustrated by the low volume fraction of cells that they represent. Nuclear pore proteins are not very abundant and consequently have proved difficult to isolate and characterise. However, significant progress has now been made using a combination of immunological, genetic and biochemical methods, and a number of nuclear pore proteins or 'nucleoporins' have now been identified (Davis and Blobel, 1986, 1987; Davis and Fink, 1990; Cordes et al., 1991; Nehrbass et al., 1990; Greber et al., 1990; Wozniak et al., 1989; Starr et al., 1990; Sukegawa and Blobel, 1993; Wente et al., 1992; Wimmer et al., 1992), many of which share characteristic repeating sequence motifs.

Many nucleoporins have been found to be glycoproteins, which have characteristic O-linked *N*-acetylglucosamine residues that can be recognised by the lectin wheat germ agglutanin (WGA) and by several monoclonal antibodies raised against nuclear envelopes (Davis and Blobel, 1986, 1987; Snow et al., 1987; Featherstone et al., 1988). Some of these glycoproteins appear to be important for nucleocytoplasmic transport, which is inhibited by WGA or the monoclonal antibodies that bind to epitopes containing common sugar residues in nucleoporins (Dabauvalle et al., 1988, 1990; Featherstone et al., 1988; Finlay et al., 1987, 1991; Snow et al., 1987). Significantly, although nuclei formed in *Xenopus* egg reassembly systems depleted of these glycoproteins using WGA-Sepharose are defective in nucleocytoplasmic transport, effective transport can be restored by adding back material eluted from the WGA-Sepharose using appropriate sugars (Finlay et al., 1991).

The best-understood nucleoporin is probably the glycoprotein p62 (Davis and Blobel, 1986), which is the component of vertebrate nuclear pores most easily recognised by WGA (Davis and Blobel, 1987). This protein appears to exist in a complex with at least two other glycoproteins (Finlay et al., 1991; Dabauvalle et al., 1990) and its cDNA has been cloned for man (Carmo-Fonseca et al., 1991), rat (Starr et al., 1990), mouse and *Xenopus* (Cordes et al., 1991). Analysis (Cordes et al., 1991; Carmo-Fonseca et al., 1991; Starr and Hanover 1991) of the predicted amino acid sequence of p62 indicates that the molecule can be divided into three segments as illustrated in Fig. 1: a C-terminal domain (residues 330-525) that contains the heptad repeat characteristic of α-helical coiled-coil proteins; an N-terminal domain (residues 1-140) that contains at least 11 copies of a degenerate repeat based on a central five-residue xFxFG core (where x is commonly a small residue such as Gly, Ser, Thr or Asn); and a central threonine-rich linker (residues 190-330). Sequence analysis also suggests that the yeast homologue of p62 may be NSP1, an essential protein that has

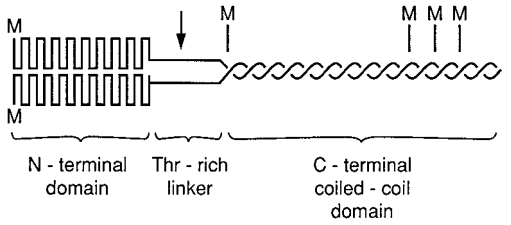

Fig. 1. Schematic illustration of the domain structure of the p62 molecule based on analysis of sequence data (Carmo-Fonseca et al., 1991; Cordes et al., 1991; Starr and Hanover, 1991). The N-terminal domain is based on a degenerate xFxFG repeat (x is usually a small residue, such as G, S or T) and probably contains alternating β-sheets and loops, whereas the C-terminal domain has the heptad repeat characteristic of an α-helical coiled-coil. A threonine-rich region joins these two domains. The position of methionine residues is marked by M and the arrow marks the site of cleavage of expressed rat p62 by thrombin.

analogous coiled-coil and xFxFG domains (Carmo-Fonseca et al., 1991; Nehrbass et al., 1991).

We report here the expression of rat p62 cDNA in *Escherichia coli* and characterisation of the resultant protein. Our results support the general structural model derived from sequence analysis, while an examination of a series of fragments indicates that a small region near the C terminus of the molecule is important for its assembly. These results also suggest a possible manner in which p62 could contribute to nuclear pore structure and function.

MATERIALS AND METHODS

Expression and purification of p62

Rat liver p62 cDNA (Starr et al., 1990) was cloned into expression vector pET3a (Studier et al., 1990) and, after sequencing on both strands to confirm the coding sequence, the resultant plasmid transformed into *E. coli* strain BL21/DE3 containing plasmid pLysS (Studier et al., 1990). Single colonies of freshly transformed cells were grown in 10 ml culture overnight and then diluted to 1 litre with 2×TY medium and grown to an A_{600} of 0.5-0.6 when they were induced by the addition of IPTG to 0.4 mM. After growing for a further 3-4 hours, cells were harvested and frozen. Inclusion bodies containing p62 were isolated and washed as described (Moir et al., 1991) and then dissolved in 8 M urea, 20 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM dithiothreitol (DTT). After centrifugation to remove insoluble material, the p62 was purified by ion-exchange chromatography on a 2.5 cm × 10 cm DE-52 (Whatman) column using a 0-100 mM linear NaCl gradient. Typical yields were 50 mg/l of culture.

To renature p62, samples at approximately 0.5 mg/ml in 8 M urea, 20 mM Tris-HCl, pH 8, 1 mM EDTA, 2 mM dithiothreitol were diluted twofold with 50 mM Tris-HCl, 5 mM EDTA, 10 mM dithiothreitol, pH 8, and, after standing for 2 hours at 4°C, were dialysed exhaustively against 20 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM DTT at 4°C.

Production of fragments

Thrombin digests of p62 were obtained by incubating 50 mg of p62 in 20 mM Tris-HCl, pH 8, 1 mM DTT, 1 mM $CaCl_2$ with 250 units of thrombin (Sigma T-6634) for several days at room temperature. The resultant 35 kDa fragment was purified by ion exchange chromatography (Whatman DE52) in 8 M urea, 20 mM Tris-HCl, 1 mM DTT, 1 mM EGTA, pH 8, using a linear 0-200 mM NaCl gradient. For cyanogen bromide digestion, p62 was dialysed against 1% acetic acid, 1 mM β-mercaptoethanol and freeze-dried. This material was dissolved in the minimum amount of 70% formic acid and a 5- to 100-fold excess of CNBr added. After 24-48 hours at 37°C, the material was freeze-dried and dissolved in 8 M urea, 20 mM Tris-HCl, 1 mM DTT, 1 mM EGTA, pH 8, and purified using a 2.5 cm × 100 cm Sephadex G-75 (Pharmacia) gel filtration column.

Electron microscopy

Samples for electron microscopy were applied to freshly cleaved mica in either 20% glycerol or pure water and then centrifuged to produce a thin film as described by Nave et al. (1989). After drying at 10^{-7} Torr (1 Torr=133.3 Pa), samples were shadowed with platinum/carbon at a nominal angle of 10° followed by carbon at 90° using a Cressington CFE-50 freeze-fracture apparatus. Specimens were then examined at 80 kV in Philips EM301 or 420 electron microscopes and images were recorded on Kodak SO-163 film at nominal magnifications between ×15,000 and ×30,000. Molecular lengths were measured from ×90,000 prints, using molecules that were straight, separated from their neighbours, and aligned approximately perpendicular to the shadowing direction as described (Stewart and Edwards, 1984).

Solubility

To determine the solubility of p62 and its fragments as a function of ionic strength, the proteins in 20 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM DTT were dialysed against the same buffer containing increasing concentrations of NaCl. Samples were then centrifuged at 200,000 ***g*** in an airfuge for 30 minutes and the concentration of protein remaining in the supernatant was determined.

Protein chemical techniques

Protein concentration was determined by A_{280} using the following absorption coefficients for a concentration of 1 mg/ml calculated from amino acid composition: p62, 0.39; thrombin fragment, 0.80; CNBr fragment, 0.32; CNBr-thrombin fragment, 0.71. We found that the Bradford (1976) dye binding assay was not accurate for this material because the N-terminal domain tended to bind less dye than most proteins (perhaps because of its hydrophobic nature) and so tended to seriously underestimate the concentration of p62 and the CNBr fragment. SDS-PAGE was carried out as described by Laemmli (1971) using 7.5% to 17.5% gradient minigels. Microsequencing of protein samples was performed by Mark Skehel of the MRC Laboratory of Molecular Biology using an Applied Biosystems gas-phase sequenator. Analytical gel filtration was performed with a Pharmacia FPLC System using a Superose 6 column eluted with 2 M urea, 10 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM DTT at 20°C. The column was calibrated using rabbit cardiac tropomyosin (M_r 66,000), rabbit skeletal short S-2 (M_r 80,000), rabbit skeletal light meromyosin (M_r 180,000) and rabbit skeletal myosin rod (M_r 250,000). Circular dichroism spectra were recorded at protein concentrations of 0.2-0.5 mg/ml in 5 mM sodium phosphate buffer, pH 7.0, at room temperature using a ISA CD-6 spectrometer and 0.1 mm or 1 mm quartz cells as appropriate.

RESULTS

Expression of p62 in *E. coli*

We obtained high levels of expression of rat nucleoporin p62 cDNA (Starr et al., 1990) using the bacteriophage T7-based pET expression system (Studier et al., 1990). After induction with IPTG, a major new band was present at approximately 55 kDa (Fig. 2B), which cross-reacted with a monoclonal antibody to p62 (Starr et al., 1990) on western blots (not shown). The expressed material was present in inclusion

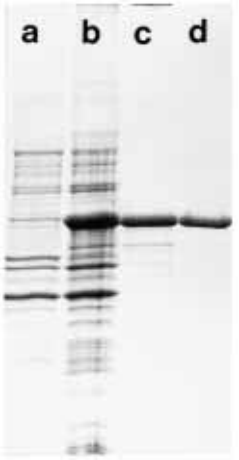

Fig. 2. SDS-PAGE showing expression and isolation of p62 in *E. coli*. (a) Cells before induction with IPTG; (b) cells 4 hours after induction; (c) isolated inclusion bodies; (d) p62 after purification by ion-exchange chromatography.

bodies (Fig. 2C) that were easily dissolved in 8 M urea and purified to homogeneity using ion-exchange chromatography (Fig. 2D). The material in 8 M urea was renatured by dialysis against 20 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM DTT after first diluting twofold to reduce the urea concentration to 4 M where coiled-coil assembly could start to occur. We explored a range of renaturation conditions, including those used by Cordes et al. (1991) for *Xenopus* p62, but found no difference in the properties of the material produced.

Electron microscopy of shadowed p62 preparations showed rod-shaped molecules 35±4 nm (mean ± s.d., n=169) long that frequently tended to aggregate to form chains (Fig. 3), probably by way of interactions near their ends. Problems were encountered due to the aggregation of p62 and its low affinity for mica, and generally better results were obtained by centrifugal spreading of solutions on mica (Nave et al., 1989) rather than by spraying. Also, the addition of glycerol tended to promote aggregation, and so the best results were obtained when material was spread from pure water. Many shadowed molecules appeared to have a slight protuberance at one end (Fig. 3), which we tentatively identify with the N-terminal domain (see Fig. 1), and in some instances there was a suggestion that this portion of the molecule bifrucated (Fig. 3B and C).

The circular dichroism spectrum of p62 in 5 mM phosphate buffer (Fig. 4) showed a substantial negative ellipticity at 220 nm and a positive ellipticity at 200 nm, consistent with the α-helix content of the coiled-coil rod domain. Overall, the shape of the spectrum resembled closely that produced by α-helical secondary structure (Greenfield and Fasman, 1969), albeit the intensity was only about a third of that for a completely α-helical structure. The close approximation to the shape of an α-helical spectrum indicated that the contributions from β and random conformations approximately cancelled in p62, which indicated that they were present in the ratio of approximately 1 to 3. Quantitation of the spectra and comparison with standards (Greenfield and Fasman, 1969) indicated that the molecule contained approximately 31% α, 22% β and 47% random coil, which was in good agreement with the values suggested from sequence analysis (Carmo-Fonseca et al., 1991; Cordes et al., 1991; Starr and Hanover, 1991). The 195 residues of the rod constitute 37% of the sequence. The heptad pattern characteristic of an α-helical coiled coil is interrupted in several places and so, if the rod contained 80% α structure, one would anticipate p62 having roughly 33% α conformation overall. If the threonine-rich linker, which constitutes roughly 25% of the sequence, was mainly random and the repeats in the N-terminal domain contained roughly equal amounts of β and random, then one would anticipate whole p62 having 16%

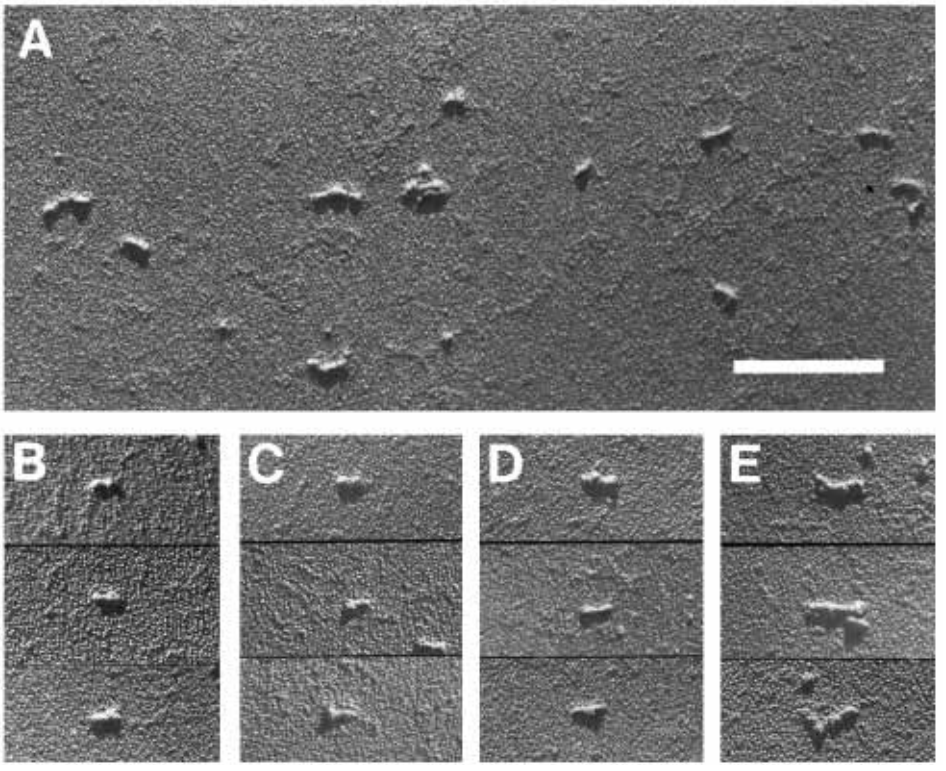

Fig. 3. Electron microscopy of p62 spread on mica from pure water using centrifugation and shadowed unidirectionally with platinum-carbon at a nominal angle of 10°. (A) A field containing rod-shaped molecules; there is a mixture of lengths corresponding to one, two and three molecules joined end-to-end. (B and C) Galleries of isolated 35 nm long molecules, which show a slight protuberance at one end (at the right-hand end of the molecules in (B) and the left-hand end in (C)). In some instances the protuberance appeared to bifurcate slightly. (D) A gallery of dimers linked end-to-end whereas (E) shows higher aggregates. In most cases p62 polymers were formed by molecules joining end-to-end to produce chains. Bar, 200 nm.

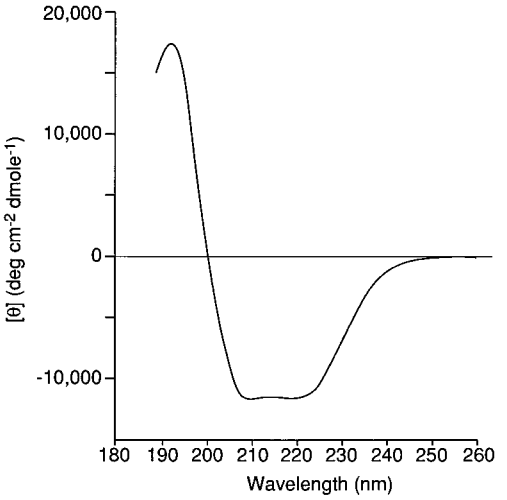

Fig. 4. Circular dichroism spectrum of p62 showing a marked negative ellipticity at 210-220 nm and positive ellipticity below 200 nm characteristic of the α-conformation. Comparison with standards indicated a secondary structure content of 31%, α; 22%, β; and 47%, random.

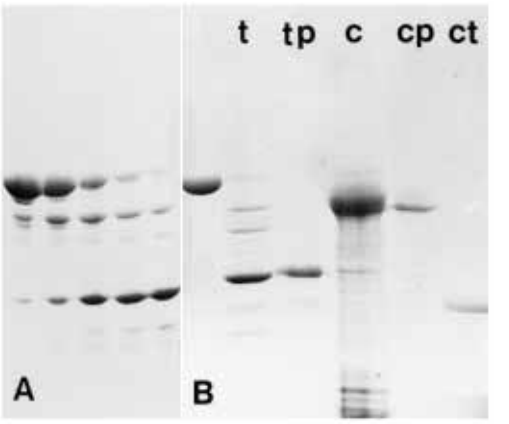

Fig. 5. SDS-PAGE of digestion of p62. (A) Thrombin produced a limiting digest that gave a fragment of 35 kDa (B) Thrombin digest (t), purified fragments produced by thrombin (tp), cyanogen bromide digest (c), purified CNBr fragments (cp) and a thrombin-CNBr double digest (ct). Cyanogen bromide cleavage produced a triplet of fragments corresponding to cleavage at Met450; Met463 and Met485.

β and 41% random coil. Because of its tendency to aggregate, it was only possible to use gel filtration to determine the molecular mass of p62 in the presence of 2 M urea, which was able to dissociate the molecules from one another without dissociating the chains within a molecule. In these conditions, p62 eluted with an apparent M_r of 95,000. Since a single chain has a M_r of 55,000, this result was consistent with the presence of two-chain p62 molecules.

Generation of fragments

We investigated a range of methods to generate fragments of p62 suitable for assessing the roles of the different molecular domains. Although most common proteases (trypsin, chymotrypsin, V8) gave a series of terminal digest products, thrombin gave a particularly stable fragment (Fig. 5A) that was easily isolated and purified by ion-exchange chromatography. Solid-phase sequence analysis of this fragment gave its N-terminal sequence as APGAAPGASTTSTTTTTTT, which indicated that proteolysis had taken place roughly in the middle of the p62 polypeptide chain at Lys261 and that the proteolytic fragment produced therefore corresponded to the rod domain of the molecule plus a small segment of the threonine-rich linker (see Fig. 6).

Cyanogen bromide cleavage produced a triplet of fragments corresponding to cleavage at Met450, Met463 and Met485 (Fig. 5), which we confirmed by N-terminal sequence determinations of the resultant three small peptides produced (see Fig. 5). Methionine 331 at the junction of the rod and threonine-rich linker was resistant to CNBr cleavage, probably because it was followed by a threonine residue. The three large CNBr fragments were easily separated from whole p62, but could not be resolved from one another by either ion-exchange or gel-filtration chromatography. However, they were useful as fragments that corresponded to most of the p62 molecule except for a small segment at the C terminus. We also produced a smaller fragment by a double digest of p62 with thrombin

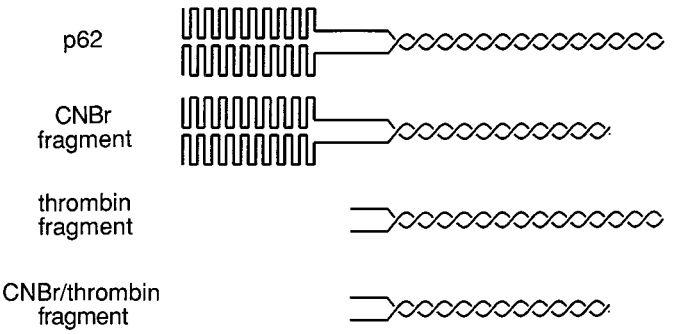

Fig. 6. Schematic illustration of the fragments of p62 produced.

Table 1. Secondary structure predicted from circular dichroism for p62 and its fragments

Sample	α (%)	β (%)	Random (%)
Whole p62	31	22	47
CNBr fragment	20	20	60
Thrombin fragment	52	0	48
CNBr/thrombin	33	0	67

followed by CNBr. Fig. 6 summarises the various fragments produced. The circular dichroism spectra of these fragments (Table 1) were generally those expected from their position in the molecule. For example, the thrombin fragment that corresponded to the rod domain and part of the linker had the highest content of α-helix, whereas the CNBr fragments that had lost part of the rod had the lowest content of α-helix. Electron microscopy of shadowed preparations of these fragments showed rod-shaped molecules (Fig. 7), although those which had been digested with thrombin were significantly shorter than native p62 or its CNBr fragment (Table 2).

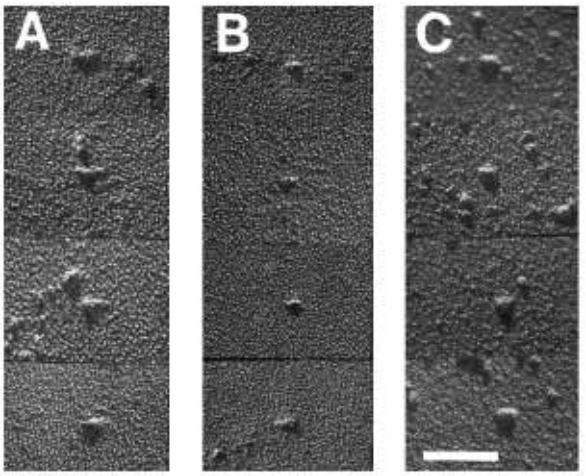

Fig. 7. Electron micrographs of shadowed preparations of p62 fragments. (A) CNBr large fragment; (B) thrombin fragment; (C) CNBr/thrombin double digest. Bar, 100 nm.

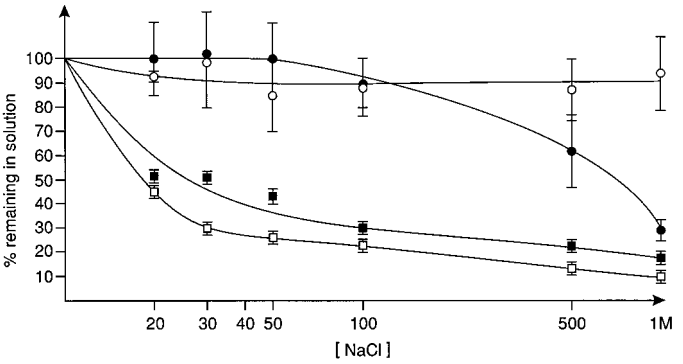

Fig. 8. Solubility of p62 and its fragments as a function of ionic strength. (■) p62; (□) thrombin fragment; (●) large CNBr fragment; (○) CNBr/thrombin double digest. Although the solubility of whole p62 and the thrombin fragment decreased rapidly with increasing ionic strength, fragments from with the extreme C terminus had been removed by CNBr remained soluble up to quite high ionic strengths.

Table 2. Length of p62 and its fragments in electron micrographs of shadowed material

Sample	Mean length (nm)	Standard deviation	Number of observations
p62	35	4	169
CNBr fragment	36	6	59
Thrombin fragment	29	4	59
CNBr/thrombin	27	3	66

Solubility

The solubility and aggregation behaviour of the p62 fragments lacking a small segment at the C terminus as a result of CNBr digestion was strikingly different from that of the intact molecule or the thrombin fragment lacking the N-terminal domain (Fig. 8). Whereas the solubility of whole p62 and the thrombin fragment was generally low and decreased rapidly with increasing ionic strength, fragments lacking the C terminus of the rod domain were very much more soluble and were only precipitated at higher salt concentrations. The similarity in behaviour between the thrombin fragment and whole p62 indicated that a principal determinant for molecular aggregation resided in the rod domain of the molecule and was not influenced markedly by the presence of the N-terminal domain.

DISCUSSION

A major difficulty in studying nuclear pore proteins (nucleoporins) is the very small quantities that can usually be obtained because of the low number of copies of these proteins in cells. Even in yeast, where the volume fraction of nuclear envelope is probably greatest, isolation of nucleoporins in any quantity has required great effort and ingenuity (Wente et al., 1992; Wimmer et al., 1992). Thus the production of large quantities of nucleoporins from their corresponding cDNAs offers an attractive route for obtaining sufficient material for structural studies and for investigating their assembly and interaction with other nuclear pore components, and indeed Cordes et al. (1991) have already expressed murine p62 for use in assessing glycosylation sites. In the present study we expressed rat p62 cDNA in *E. coli* to obtain sufficient material to establish that the molecular structure is indeed similar to that predicted by analysis of sequence data, while the generation of fragments has given some insight into the role of the extreme C terminus of the coiled-coil rod domain in self-assembly.

Both electron microscopy and circular dichroism spectroscopy support a molecular structural model in which the C-terminal third of the sequence, from residue 330 to 525, has an α-helical coiled-coil conformation. Furthermore, the circular dichroism spectrum also suggests that the repetitive N-terminal domain contains roughly equal proportions of β-sheet and random coil secondary structure, which would be consistent with a domain based on a cross-β conformation in which short segments of the polypeptide chain (particularly those with the xFxFG repeat) alternate with surface loops as suggested by Carmo-Fonseca et al. (1991), Cordes et al. (1991) and Starr and Hanover (1991). The variability in length of the segments between successive xFxFG motifs would also be consistent with such a model, because surface loops are usually the most variable parts of protein sequences. The location of the thrombin digestion site to the threonine-rich linker would be consistent with suggestions that this segment of the sequence may adopt a relatively open structure. Thrombin usually cleaves at arginine residues, especially those followed by glycine, isoleucine, valine, asparagine, cysteine or arginine, and so the actual cleavage at a lysine is somewhat unusual and probably is a consequence of this portion of the molecule having a very open conformation. The coiled-coil rod domain contains 13 arginine residues and so the failure of thrombin to cleave at any of these sites in the rod would be consistent with their being located in a coiled-coil conformation, which, because of its compact nature, usually shows considerable resistance to proteolysis. Coiled-coil molecules that are located within cells usually are two-stranded and the gel filtration data indicate that p62 indeed contains two polypeptide chains.

The axial spacing between successive residues along an α-helical coiled-coil is 0.15 nm, and so the putative 195-residue rod domain of p62 would be expected to have a length of about

29 nm. The length of the molecules was certainly consistent with this model, with the excess length of 6 nm probably being accounted for by the bulk of the N-terminal domains. Support for this observation came from the lengths determined for the fragments of p62 (Table 2). Although the change in length produced by removing a small fragment from the molecule's C terminus was too small to be detected in shadowed molecules, there was a significant decrease in molecular length when the N-terminal domain was removed by thrombin digestion. Thus, intact p62 and its large CNBr fragment were of the order of 35 nm long whereas the thrombin fragment and the thrombin/CNBr double-digest fragment were both of the order of 28 nm long.

The generally low solubility of p62 indicated that there was considerable interaction between molecules and indeed the decrease in solubility with increasing salt concentration is very similar to that seen with cytoplasmic intermediate filament proteins (Steinert and Roop, 1988). Moreover, p62 expressed in COS cells (Starr et al., 1990) is also insoluble in vivo, indicating that the low solubility we observed with the material expressed in bacteria was a genuine property of the molecule and not an artefact due, for example, to improper renaturation. Electron microscopy indicated that the self-association of p62 molecules was probably a specific interaction in which molecules aggregated to form chains. This formation of linear chains of molecules appeared analogous to that seen with nuclear lamins (Heitlinger et al., 1991), although clearly the precise nature of the interaction would not be expected to be the same. The substantial increase in solubility observed in fragments in which the last 40-75 residues had been removed by CNBr cleavage (Fig. 8) was consistent with this observation and further suggested that a significant determinant of p62 aggregation was located at the extreme C terminus of the molecule. It is perhaps interesting to note in this context that other aggregating coiled-coil molecules, such as myosin (Atkinson and Stewart, 1990), tropomyosin (Johnson and Smillie, 1977) and intermediate filaments (reviewed by Stewart, 1993), also have small regions at their C termini that are important determinants of molecular assembly. The aggregation and solubility properties of the thrombin digest indicated that the other site important for self-association was probably near the N terminus of the coiled-coil rod domain. Although it was not possible to establish unequivocally whether molecules aggregated parallel (so that the interaction was between N and C termini of the rod) or antiparallel (so that the interaction was between C termini only), the observation that linear aggregates containing more than two molecules would suggest strongly that the interaction was primarily between opposite ends of the molecules. In nuclear pores p62 is associated with at least two other proteins to form a complex (Dabauvalle et al., 1990; Findlay et al., 1991) that appears to contain more than one copy of each polypeptide chain, which would be consistent with self-association interactions analogous to those we have observed in vitro.

The aggregation of p62 to form linear chains may be important for its function in nuclear pores. One possible product of such end-to-end polymerisation would be very fine filaments. Although there have indeed been reports of fine filaments projecting from nuclear pores (see Franke, 1974; Maul, 1977), ultrastructural studies have indicated that p62 is located in the body of the nuclear pore complex (reviewed by Stewart, 1992) and so we think it is unlikely that p62 could constitute either nucleoplasmic or cytoplasmic filaments radiating from pores. However, an alternative form of end-to-end interaction might be to link a number of p62 molecules into a circle to form a ring such as that illustrated in Fig. 9 (which contains eight molecules because nuclear pores have prominent 8-fold rotational symmetry; see Stewart, 1992). Eight molecules linked in this way would would be consistent with the formation of a ring of diameter of the order of 80 nm, which is roughly that of the central disc domain and of the inner diameter of the nucleoplasmic and cytoplasmic rings of nuclear pores (see Stewart, 1992). Such a ring of coiled-coil proteins would be strong in tension, and so could function like barrel hoops and provide strength to the nuclear pore, particularly against radial stresses such as those that might arise as the pore distends to accommodate the transport of large particles such as ribosome subunits and snRNPs. Moreover, the repetitive N-terminal domains and threonine-rich linker, to which most of the *N*-acetylglucosamine residues are attached (Cordes and Krohne, 1993), could face inwards to provide a generally hydrophilic lining of the pore and/or potential attachment sites for other pore proteins, especially those associated with the central 'transporter' region. Electron microscopic location of glycoprotein epitopes, using either WGA or antibodies indicates a likely radius of about 10-12 nm, which would be consistent with such a model (Akey and Goldfarb, 1989). However, clearly more precise location of p62 and its different domains will be required to establish such a structural model more firmly.

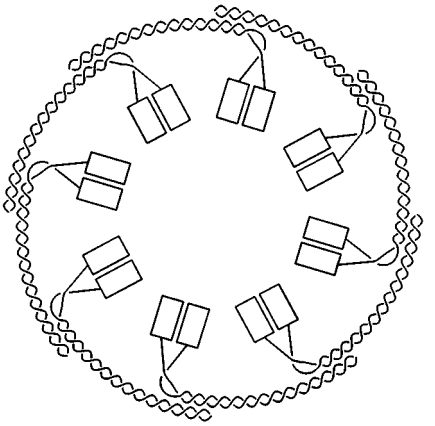

Fig. 9. Hypothetical model of how p62 molecules could join to produce rings that could act like barrel hoops in nuclear pores. In this highly schematic model, eight p62 molecules are linked through their rod domains to form a ring of diameter approximately 80 nm, with the N-terminal domain and linkers facing inward where they might provide a hydrophilic lining to the central pore channel or provide support for a centrally located transporter unit in the pore complex.

Because the p62 expressed in *E. coli* has not been glycosylated, its properties may differ slightly from native material. However, glycosylation generally takes place on surface residues of proteins, and so results in minimal conformational change. However, one possible result of glycosylation could be to increase the rigidity of the threonine-rich linker domain as suggested by Cordes et al. (1991) and so our circular dichroism studies may have overestimated the content of

random coil structure in this region of the native molecule. However, the observation that the glycosylated p62 overexpressed in COS cells (Starr et al., 1990) forms insoluble aggregates indicates that the solubility properties we observed were probably not due to a lack of glycosylation. Moreover, our studies with fragments indicated that the solubility of the molecule could be increased greatly by removing a small number of residues from the C terminus of the rod, whereas glycosylation occurs primarily outside the rod domain (Cordes and Krohne, 1993; D'Onofrio et al., 1988).

In summary, our electron microscopy and circular dichroism results support the general structural model proposed for p62 on the basis of its sequence (Carmo-Fonseca et al., 1991; Cordes et al., 1991; Starr and Hanover, 1991). Although precise information on the detailed conformation of the domains of p62 will require higher-resolution studies using either X-ray crystallography or NMR spectroscopy, our studies establish the broad outline of the structure of the molecule and are consistent with a 30 nm α-helical coiled-coil rod domain linked to an N-terminal domain formed from alternating β-sheets and loops. Moreover, the solubility data obtained from fragments indicate that an important contribution to the aggregation of p62 to form linear chains observed by electron microscopy resides at the C terminus of the rod domain.

We thank our colleagues in Cambridge and Washington for their many helpful comments, criticisms and suggestions, Dr M. Suzuki for assistance with circular dichroism, Mark Skehel for amino acid sequence determination, Annette Lenton and Patrick Sadler for artwork and Sue Whytock for photographic assistance. F.B. was supported by an EMBO long-term postdoctoral fellowship.

REFERENCES

Akey, C. W. and Goldfarb, D. S. (1989). Protein import through the nuclear pore is a multi-stage process. *J. Cell Biol.* **109**, 971-982.

Akey, C. W. (1991). Probing the structure and function of the nuclear pore complex. *Semin. Cell Biol.* **2**, 167-177.

Atkinson, S. J. and Stewart, M. (1991). Expression in *Escherichia coli* of fragments of the rod domain of rabbit myosin: influence of different regions of the molecule on aggregation and paracrystal formation. *J. Cell Sci.* **99**, 823-836.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.

Carmo-Fonseca, M., Kern, H. and Hurt, E. (1991). Human nucleoporin p62 and the essential yeast nuclear pore protein NSP1 show sequence homology and a similar domain organisation. *Eur. J. Cell Biol.* **55**, 17-30.

Cordes, V. and Krohne, G. (1993). Sequential O-glycosylation of nuclear pore complex protein p62 in vitro. *Eur. J. Cell Biol.* **60**, 185-195.

Cordes, V., Waizenegger, I. and Krohne, G. (1991). Nuclear pore complex glycoprotein p62 of *Xenopus laevis* and mouse: cDNA cloning and identification of its glycosylation region. *Eur. J. Cell Biol.* **55**, 31-47.

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

Dabauvalle, M-C., Loos, K. and Scheer, U. (1990). Identification of a soluble precursor complex essential for nuclear pore assembly in vitro. *Chromosoma* **100**, 56-66.

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

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

Davis, L. I. and Fink, G. R. (1990). The NUP1 gene encodes an essential component of the yeast nuclear pore complex. *Cell* **61**, 965-978.

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

Featherstone, C., Darby, M. K. and Gerace, L. (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., Newmeyer, D. D., Price, T. M. and Forbes, D. J. (1987). Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores. *J. Cell Biol.* **104**, 189-200.

Finlay, D. R., Meier, E., Bradley, P., Horecka, J. and Forbes, D. J. (1991). A complex of nuclear pore proteins required for pore function. *J. Cell Biol.* **114**, 169-183.

Forbes, D. J. (1992). Structure and function of the nuclear pore. *Annu. Rev. Cell Biol.* **8**, 495-527.

Franke, W. W. (1974). Structure, biochemistry and function of the nuclear envelope. *Int. Rev. Cytol. Suppl.* **4**, 71-236.

Greber, U. F., Senior, A. and Gerace, L. (1990). A major glycoprotein of the nuclear pore complex is a membrane-spanning polypeptide with a large lumenal domain and a small cytoplasmic tail. *EMBO J.* **9**, 1495-1502.

Greenfield, N. and Fasman, G. D. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **68**, 4108-4116.

Heitlinger, E., Peter, M., Haener, M., Lustig, A., Aebi, U. and Nigg, E. A. (1991). Expression of chicken lamin B2 in *Escherichia coli:* characterization of its structure, assembly and molecular interaction. *J. Cell Biol.* **113**, 485-495.

Hinshaw, J. E., Carragher, B. O. and Milligan, R. A. (1992). Architecture and design of the nuclear pore complex. *Cell* **69**, 1133-1141.

Johnson, P. and Smillie, L. B. (1977). Polymerizability of rabbit skeletal tropomyosin: effects of enzymatic and chemical modification. *Biochemistry* **16**, 2264-2269.

Laemmli, U. K. (1971). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Maul, G. G. (1977). The nuclear and cytoplasmic pore complexes: structure, dynamics, distribution and evolution. *Int. Rev. Cytol. Suppl.* **6**, 75-186.

Miller M., Park, M. K. and Hanover, J. A. (1991). Nuclear pore complex: structure, function and regulation. *Physiol. Rev.* **71**, 909-949.

Moir R. M., Donaldson, A. D and Stewart, M. (1991). Expression in *Escherichia coli* of human lamins A and C: influence of head and tail domains on assembly properties and paracrystal formation. *J. Cell Sci.* **99**, 363-372.

Nave, R., Furst, D. O. and Weber, K. (1989). Visualization of the polarity of isolated titin molecules: a single globular head on a long rod as the M band anchoring domain? *J. Cell Biol.* **109**, 2177-2187.

Nehrbass, U., Kern, H., Mutvei, A., Horstmann, H., Marshallsay, B. and Hurt, E. (1990). NSP1: a yeast nuclear envelope protein localized at the nuclear pores exerts its essential function by its carboxy-terminal domain. *EMBO J.* **61**, 979-989.

Reichelt, A., Holzenberg, A., Buhle, L., Jarnick, M. and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J. Cell Biol.* **110**, 883-894.

Silver, P. A. (1991). How proteins enter the nucleus. *Cell* **64**, 489-497.

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

Starr, C. M. and Hanover, J. A. (1991). A common structural motif in nuclear pore proteins (nucleoporins). *BioEssays* **13**, 145-146.

Starr, C. M., D'Onofrio, M., Park, M. K. and Hanover, J. A. (1990). Primary sequence and heterologous expression of nuclear pore glycoprotein p62. *J. Cell Biol.* **110**, 1861-1871.

Steinert, P. M. and Roop, D. R. (1988). Molecular and cellular biology of intermediate filaments. *Annu. Rev. Biochem.* **57**, 593-625.

Stewart, M. (1992). Nuclear pore structure and function. *Semin. Cell Biol.* **3**, 267-277.

Stewart, M. (1993). Intermediate filament structure and assembly. *Curr. Opin. Cell Biol.* **5**, 3-11.

Stewart, M. and Edwards, P. (1984). Length of myosin rod and its fragments determined by electron microscopy. *FEBS Lett.* **168**, 75-78.

Studier, F. W., Rosenberg A. R. and Dunn, J. J. and Dubendorff, J. W.

(1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* **185**, 60-69.

Sukegawa, J. and Blobel, G. (1993). A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and faces the nucleoplasm. *Cell* **72**, 29-38.

Wente, S., Rout, M. D. and Blobel, G. (1992). A new family of yeast nuclear pore complex proteins. *J. Cell Biol*. **119**, 705-723.

Wimmer, C., Doye, V., Graudi, P., Nehrbass, U. and Hurt, E. C. (1992). A new subclass of nucleoporins that functionally interact with nuclear pore protein NSP1. *EMBO J.* **11**, 5051-5061.

Wozniak, R. W., Bartnik, E. and Blobel, G. (1989). Primary structure analysis of an integral membrane glycoprotein of the nuclear pore. *J. Cell Biol*. **108**, 2083-2092.

(Received 6 July 1993 - Accepted, in revised form, 28 October 1993)