Multiple Conformations in the Ligand-binding Site of the Yeast Nuclear Pore-targeting Domain of Nup116p*

Received for publication, May 9, 2005, and in revised form, August 10, 2005
Published, JBC Papers in Press, August 16, 2005, DOI 10.1074/jbc.M505068200

Michael A. Robinson‡§1,2, Sunghyouk Park‡1,3, Zhen-Yu J. Sun†, Pamela A. Silver‡*‡, Gerhard Wagner§‡‡3, and James M. Hogle‡‡4

From the Departments of Biological Chemistry and Molecular Pharmacology and Institute of Systems Biology, Harvard Medical School, and the Dana-Farber Cancer Institute, Boston, Massachusetts 02115, the Committee on Higher Degrees in Biophysics, Harvard University, Cambridge, Massachusetts 02138, and the Department of Biochemistry, College of Medicine, Inha University, Shinheung-dong, Chung-gu, Incheon 400-712, Korea

The yeast nucleoporin Nup116p plays an important role in mRNA export and protein transport. We have determined the solution structure of the C-terminal 147 residues of this protein, the region responsible for targeting the protein to the nuclear pore complex (NPC). The structure of Nup116p-C consists of a large β-sheet sandwiched against a smaller one, flanked on both sides by α-helical stretches, similar to the structure of its human homolog, NUP98. In unliganded form, Nup116p-C exhibits evidence of exchange among multiple conformations, raising the intriguing possibility that it may adopt distinct conformations when bound to different partners in the NPC. We have additionally shown that a peptide from the N terminus of the nucleoporin Nup145p-C binds Nup116p-C. This previously unknown interaction may explain the unusual asymmetric localization pattern of Nup116p in the NPC. Strikingly, the exchange phenomenon observed in the unbound state is greatly reduced in the corresponding spectra of peptide-bound Nup116p-C, suggesting that the binding interaction stabilizes the domain conformation. This study offers a high resolution view of a yeast nucleoporin structural domain and may provide insights into NPC architecture and function.

Carefully controlled nucleocytoplasmic transport is critical for the eukaryotic cell, playing a role in key functions such as gene expression and cell division. This transport is mediated by nuclear pore complexes (NPCs). Large proteinaceous assemblies embedded in the nuclear envelope. Each NPC consists of ~30 distinct proteins termed nucleoporins, each present in at least eight copies, reflecting the octagonal symmetry of the complex (1, 2). The core region of the NPC additionally possesses bilateral symmetry in the plane of the membrane, whereas peripheral structures (the cytoplasmic filaments and the nuclear basket) distinguish the two faces.

Although NPCs have been studied extensively via genetics, biochemistry, and electron microscopy, there currently exist only limited atomic resolution data on the structures of NPC components. Thus, an important goal in improving our understanding of NPC function is the structural characterization of its constituent parts. Here, we focus on a structural domain from the yeast nucleoporin Nup116p.

Nup116p is involved in mRNA export and protein transport (3–5). It is a member of the GLFG class of nucleoporins, containing a large number of Gly-Leu-Phe-Gly sequence motifs that interact with soluble transport receptors (5). The Nup116p deletion mutant is temperature-sensitive, forming a double membrane seal over the cytoplasmic face of NPCs and shutting down nucleocytoplasmic transport when shifted to the non-permissive temperature of 37 °C (6).

Nup116p has an unusual localization pattern within the NPC: it is found on both faces of the pore, but the majority is localized to the cytoplasmic face (1). This suggests that Nup116p possesses at least two distinct binding partners within the pore. One known binding partner, Nup82p, is found exclusively on the cytoplasmic face of the pore (7, 8), whereas the nuclear binding partner is not known. Nup82p is required for poly(A)⁺ RNA export (9); and thus, Nup116p may contribute to the formation of a subcomplex at the cytoplasmic face of the pore that is responsible for a terminal step in mRNA export (10).

Nup116p consists of three functional domains (Fig. 1A): an N-terminal Gle2p-binding site, which allows it to form a stable complex with Gle2p, an important nucleoporin component of the mRNA export machinery; a large GLFG repeat region spanning the N-terminal two-thirds of the sequence; and a ~150-residue C-terminal domain responsible for NPC targeting. This latter domain is predicted to have a high degree of secondary structure, in contrast to much of the rest of the sequence, and may act as a tether to attach the long flexible N-terminal GLFG repeat regions and the Gle2p-binding site to the NPC (4, 7, 8).

Given the economy of composition of the yeast NPC (1), it is intriguing that Nup116p is homologous to two other nucleoporins, Nup100p and Nup145p-N (Fig. 1A), most likely due to gene duplication events (11). This apparent redundancy could play a biological role in ensuring continued NPC function in the event that one of the components is mutated or absent (12); however, in the higher eukaryotic NPC, a single nucleoporin (NUP98) fulfills the role of the three yeast proteins (13).

The Nup116p homolog Nup145p possesses autoproteolytic activity that is unique among yeast nucleoporins (14). This protein is expressed as a single polypeptide chain, cleaving itself post-translationally near the...
middle of the sequence to produce two species, Nup145p-N and Nup145p-C (15, 16). The human homolog, which originates as an H11011 190-kDa polyprotein, has a similar activity, cleaving itself into the N-terminal NUP98 and the C-terminal NUP96 (17). Although the biological purpose of this cleavage is not clear, it is essential in each case for proper localization of the cleaved species, suggesting an important role for this activity (14).

The autoproteolytic sites of Nup145p and NUP98/NUP96 each occur immediately C-terminal to the pore-targeting domain. Thus, although this domain occurs in the middle of each precursor polypeptide, it winds up at the C-terminal end of the N-terminal cleavage product. The resulting N-terminal products (Nup145p-N in yeast and NUP98 in humans) each bear a strong similarity to the non-cleavable Nup116p and Nup100p. In a multiple sequence alignment of the pore-targeting domains (Fig. 1B), the overall homology of these regions is evident, as well as the fact that Nup116p and Nup100p terminate only 3–4 residues downstream of the residues corresponding to the autoproteolytic sites of Nup145p and NUP98/NUP96.

The structure of the pore-targeting domain of NUP98 has been recently determined by x-ray crystallography (18). This structure, the first pore-interacting domain of any nucleoporin to be characterized in atomic detail, offers insight into the mechanism of autoproteolysis. The peptide immediately downstream of the autoproteolytic site binds along a groove in the structure; the binding interactions between the pore-targeting domain and the covalently linked partner induce sufficient strain in the scissile peptide bond to facilitate cleavage (18). The authors of this previous work were unable to determine a high quality structure of free NUP98, however; and thus, a comparison of bound and unbound structures could not be carried out.

To better understand the structure-function relationship for this class of pore-targeting domains, we have used NMR to determine the solution structure of the free 147-residue C-terminal domain of Nup116p (Nup116p-C). In addition, we show that Nup116p-C is able to bind a peptide from the N terminus of Nup145p-C in a manner analogous to the protein-peptide interaction seen in the NUP98 crystal structure. This binding interaction appears to stabilize the conformation of Nup116p-C. We discuss the implications of our structure and binding studies for NPC function.

MATERIALS AND METHODS

Cloning and Expression—Cloning was performed using the Gateway® recombination system (Invitrogen). For the GST-Nup116p-C construct, two rounds of PCR were conducted. In the first round, primers MAR116 (5'-CTGGAAGTTCTGTTCAGGGCCCAATGAGAACTACTATATCTCACC-3') and MAR117 (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGGTCTGCTCTGCAGCG-3') were used with yeast genomic DNA as a template; and in the second round, primers MAR114 (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCTGGAAGTTCTGTTCCAGGG-3') and MAR117 were used with the first-round product as a template. The resulting PCR product was recombined into the destination vector pDEST15 (Invitrogen), which encodes an N-terminal GST tag.

For the GST-TEV-Nup145p-C peptide construct, oligonucleotides MAR314 (5'-GGGGACAAGTTTGTTCAGGGCCCAATGAGAACTACTATATCTCACC-3') and MAR315 (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTTATTCATCAATTTCCGCATC-3') were used with yeast genomic DNA as a template; and in the second round, primers MAR114 (5'-GGGGACAAGTTTGTTCAGGGCCCAATGAGAACTACTATATCTCACC-3') and MAR117 were used with the first-round product as a template. The resulting PCR product was recombined into the destination vector pDEST15 (Invitrogen), which encodes an N-terminal GST tag.

For the GST-TEV-Nup145p-C peptide construct, oligonucleotides MAR314 (5'-GGGGACAAGTTTGTTCAGGGCCCAATGAGAACTACTATATCTCACC-3') and MAR315 (5'-GGGGACAAGTTTGTTCAGGGCCCAATGAGAACTACTATATCTCACC-3') were used with yeast genomic DNA as a template; and in the second round, primers MAR114 (5'-GGGGACAAGTTTGTTCAGGGCCCAATGAGAACTACTATATCTCACC-3') and MAR117 were used with the first-round product as a template. The resulting PCR product was recombined into the destination vector pDEST15 (Invitrogen), which encodes an N-terminal GST tag.
TTCTTCGTTCACCAGGCCCCAAATGCTCTGAAAATACAG-GTTTTCGAGCTGCTTTTTGTACAACCTTGCTCCC-3') were annealed and recombined in the Gateway system as described above. This expression construct encodes an N-terminal GST tag, a TEV protease cleavage site, and the N-terminal 15 amino acids of Nup116p-C.

Recombinant fusion proteins were expressed in *Escherichia coli* BL21-CodonPlus -RIL (Stratagene). Cells were grown at 37 °C to mid-log phase in media containing 16 g of Tryptone, 10 g of yeast extract, 5 g of NaCl, and 100 mg of ampicillin per liter (Qiogene, Inc.), chilled on ice for 5 min, induced with 0.5 mM isopropyl β-d-thiogalactopyranoside, and grown at 25 °C for an additional 6–8 h.

To express uniformly isotopically labeled protein, cells were grown to mid-log phase at 37 °C in M9 minimal medium with the appropriate labeled reagent substituted into the medium. The following reagents were used to achieve the corresponding isotope labeling: 98+% [15N]NH₄Cl (1 g/liter) and 99+% [13C]glucose (2 g/liter). Cells were induced as described above and grown at 25 °C for 24 h post-induction. Selective labeling of a particular amino acid was achieved by adding the following for each liter of M9 minimal medium: 125 mg of labeled amino acid; 200 mg each of the other unlabeled 19 amino acids; and 10 mg each of adenine, guanine, cytosine, and thymine.

**Nup116p-C Purification**—All purification steps were conducted at 4 °C. Frozen cell pellets containing overexpressed GST-Nup116p-C were resuspended in 50 ml of lysis buffer (50 mM Tris (pH 7.0), 500 mM NaCl, 10 mM dithiothreitol (DTT), and 1 mM EDTA). Glutathione resin was washed with 15 column volumes of wash buffer. The purified samples were analyzed by mass spectrometry (Molecular Biology Core Facilities, Dana-Farber Cancer Institute) to confirm protein mass and sample purity.

The peptide, and the pH was neutralized by adding an equal amount of HCl.

To produce isotopically labeled peptide, the GST-TEV-Nup145p-C peptide fusion protein was purified as described above for GST-Nup116p-C. Following glutathione elution, the fusion protein was concentrated to <2 ml. To this sample were added 20 μl of 0.1 M DTT, 40 μl of 0.5 mM EDTA, and 1 μl of AcTEV™ protease (Invitrogen) per mg of fusion protein and gel filtration buffer sufficient to bring the sample to 2 ml. The sample was incubated overnight at room temperature to allow protease cleavage and then filtered and applied to a HiPrep 16/60 Sephacryl™ S-100 HR gel filtration column (Amersham Biosciences) equilibrated with wash buffer. Peptide desalting was accomplished using a 3-ml Oasis hydrophilic-lipophilic balance cartridge (Waters Corp.). The purified peptide was analyzed by mass spectrometry to confirm peptide mass and purity.

**NMR Samples**—Purified isotopically labeled Nup116p-C was dia- lyzed into NMR buffer (100 mM sodium/potassium phosphate (pH 6.5), 50 mM NaCl, and 1 mM DTT) and concentrated to 0.5–1.0 mM using Vivaspin-20 concentrators with a 5-kDa molecular mass cutoff. In some cases, the peptide was mixed with the concentrated protein sample to form a complex. 25 μl of D₂O (99.8+%; Cambridge Isotope Laboratories, Inc., and Aldrich) was added to 250 μl of the protein or complex mixture to provide a lock signal.

**NMR Experiments**—NMR spectra were acquired at 293 K (unless indicated otherwise) using Varian Unity Plus 400-MHz and Unity Inova 500-, 600-, and 750-MHz and Bruker Avance 500- and 600-MHz spec- trometers. All triple resonance experiments utilized cryoprobe- equipped spectrometers. Spectra were processed using NMRPipe (19) and visualized and analyzed using NMRView (20).

The following spectra of Nup116p-C were acquired for use in backbone assignment: 15N heteronuclear single quantum correlation (HSQC), HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB, HN(CO)CABC, C(CO)NH, 15N total correlation spectroscopy (TOCSY)-HSQC, and 15N nuclear Overhauser effect (NOE) correlation spectroscopy (NOESY)-HSQC (21–23). IBIS (24), an automated assignment software package, was used to generate an initial assignment. Additionally, assignment was aided using 15N HSQC spectra of samples selectively 15N-labeled at the following residues: alanine, arginine, iso- leucine, leucine, lysine, tyrosine, and valine. The 15N NOE-HSQC spectrum was used to confirm assignments.

**H(CO)NH, HCCCH-TOCSY, D₂O-TOCSY, and 13C HSQC methyl spectra** were acquired for use in side chain assignment. Side chain resonances were assigned using the above spectra as well as 13CO TOSCY-HSQC and 15N NOE- HSQC spectra. 13C NOE-HSQC and D₂O-NOE(13C HSQC) spectra (mixing times of 100 ms) spectra were acquired for distance constraints. Based upon these spectra as well as the 13N NOE-HSQC spectrum (mixing time of 80 ms), NOE resonances were assigned and tabulated. A constant time 13C HSQC spectrum acquired from a 10% 13C-labeled Nup116p-C sample was used for stereospecific methyl assignments.

Distance and angle constraints were used to determine the structure of Nup116p-C. The intensity of each assigned NOE cross-peak was categorized as strong, medium, weak, or very weak and converted into a distance upper limit of 3.0, 4.2, 5.2, or 6.0 Å, respectively, using NMRView. For non-stereospecifically assigned diastereotopic protons and aromatic and methyl homotopic protons, the ambiguous distance con- straint method was used (25). In each case, the lower distance bound was set to 1.8 Å. TALOS (torsion angle likelihood obtained from shift and sequence similarity) (26) was used to generate ϕ and ψ angle constraints based on comparing observed chemical shifts with data base values.

Structures were calculated using the CYANA (27) and CNS (28) soft- ware packages. For the reported structure, 100 structure calculations were performed using CNS, and the 15 structures with the lowest ener- gies were chosen for the structure ensemble. MOLMOL (29) and Mol- Script (30) were used to visualize the structures and to make figures. The quality of the final structure ensemble was evaluated using PRO- CHECK-NMR (31).
Isothermal Titration Calorimetry (ITC)—Purified Nup116p-C was dialyzed extensively into ITC buffer (25 mM sodium/potassium phosphate (pH 6.5)) and brought to a concentration of 10 \( \mu \text{M} \). Dissolved peptide stock was diluted 20-fold with ITC buffer. The final peptide concentration was adjusted to 200 \( \mu \text{M} \) with ITC buffer.

A VP-ITC calorimeter (MicroCal, LLC), under the control of Origin 5.0 software, was used to perform ITC on protein-peptide complexes. The sample chamber was filled to capacity with protein and equilibrated at 25 °C. The peptide was then titrated into the sample chamber in 10-\( \mu \text{l} \) increments (30 in total), with 210 s between injections and a mixing speed of 270 rpm.

For base-line determination, identical experiments were run with peptide titrated into blank ITC buffer as well as with blank ITC buffer titrated into protein sample. The heats of dilution determined from these experiments were subtracted from the main experimental data using the Origin software. The data were fit to a single binding site model, from which thermodynamic parameters were determined.

RESULTS

Structure Determination of Nup116p-C—The solution structure of Nup116p-C was determined using standard methods of NMR spectroscopy. This domain consists of the final 147 residues of Nup116p as well as two extraneous residues (Gly-Pro) from the protease cleavage site at the N terminus of the domain. An ensemble of 15 calculated structures (Fig. 2A and TABLE ONE) was chosen based on the criteria described under “Materials and Methods.”

Consistent with the crystal structure of the human NUP98 homolog (18), Nup116p-C adopts a predominantly \( \beta \)-strand structure (Fig. 2B). The molecule consists of a six-stranded \( \beta \)-sheet sandwiched against a two-stranded \( \beta \)-sheet and flanked by \( \alpha \)-helical regions. The N-terminal helical region consists of two short helices (residues 12–16 and 21–24), whereas the stretch on the opposite side of the molecule consists of a single, longer helix.

Nup116p-C also contains three long, less well structured loop regions. Loop 1 (residues 42–59) occurs between strands \( \beta 3 \) and \( \beta 4 \) and links the large and small \( \beta \)-sheets; loop 2 (residues 69–87) occurs between strands \( \beta 5 \) and \( \beta 6 \) and also links the \( \beta \)-sheets; and loop 3 (residues 93–111) occurs between strand \( \beta 6 \) and helix \( \alpha 3 \).

Comparison with the Crystal Structure of NUP98—Superposition of a representative solution structure of Nup116p-C and the crystal structure of NUP98 (Fig. 2C) makes it clear that the structures share a common architecture. However, the domains differ functionally in that NUP98 is an autoproteolytic protein, cleaving itself into N- and C-terminal molecules (32), whereas Nup116p does not possess this activity. The NUP98 crystal structure includes this autoproteolytic juncture, and the first several residues of the C-terminal cleavage product (NUP96) are visible in the electron density. In contrast, the Nup116p-C structure does not include a bound peptide. Thus, it is not surprising that the NUP98 binding site and the corresponding region of Nup116p-C exhibit notable differences.

In particular, helix \( \alpha 3 \) in the Nup116p-C structure is shifted up to 5Å away from the corresponding helix and binding site in NUP98, creating an expanded cleft between helix \( \alpha 3 \) and strand \( \beta 5 \). In the crystal structure, this cleft is occupied by the NUP96 peptide, with which the long helix forms key hydrogen bonds and hydrophobic contacts. These contacts are clearly absent in Nup116p-C.
In the ensemble of calculated structures (Fig. 2A), helix α3 exhibits variation mainly along the direction of the helical axis from structure to structure, with less variation toward or away from the small β-sheet. Thus, the substantial observed shift of the long helix in this latter dimension, compared with NUP98, is presumed to be a real difference between the two structures, rather than an artifact of the uncertainty in the NMR structure determination.

As mentioned above, the conformations of the loop regions of Nup116p-C, particularly loop 3, vary substantially among calculated models, presumably due in part to flexibility and consequently fewer NOE constraints per residue in these regions. It is therefore difficult to make direct comparisons with the loop conformations in NUP98. Nevertheless, it is worth noting that the sequences in loop 3 and, to a lesser extent, in loop 2 are fairly well conserved (Fig. 1B). In NUP98, loop 3, which has several basic residues, may contribute to peptide binding by contacting the acidic portion of the peptide that is not visible in the electron density, whereas loop 2 plays an important role in autoproteolysis (18). The sequence conservation suggests that, although Nup116p-C does not autoproteolyze, it may be capable of binding a partner peptide in a functionally similar manner to NUP98.

**Nup116p-C Interacts with the Nup145p-C Peptide**—Of the three yeast homologs of NUP98, only Nup145p possesses autoproteolytic activity, post-translationally cleaving itself into Nup145p-N and Nup145p-C. It has been shown that only a minimal region of 10 residues or less, C-terminal to the cleavage site, is required to maintain autoproteolytic activity in Nup145p (14) and NUP98/NUP96 (32), and autoproteolytic activity is thought to depend strongly on correct peptide orientation in the binding pocket (18). Additionally, only the first 7 peptide residues have readily interpretable electron density in the x-ray structure of NUP98, suggesting that only a short peptide is required for the interaction. Because the amino acids that make up the NUP98 binding site are surprisingly well conserved across the human and yeast homologs (18), we hypothesized that Nup116p-C may bind a peptide in a similar fashion to NUP98. A plausible binding partner could be Nup145p-C, the yeast homolog of NUP96. To test this idea, we synthesized a short peptide corresponding to the first 15 residues of Nup145p-C. Residues 3–5 of this peptide (Trp-Gly-Leu) are very similar to NUP96 residues 3–5 (Tyr-Gly-Leu), which are the most highly ordered of the peptide residues in the crystal structure and which participate in the bulk of the observed protein-peptide interactions. The Nup145p-C peptide also includes a substantial stretch of acidic residues in its C-terminal half, a conserved feature that may be important in binding.

ITC was used to assess whether the Nup145p-C peptide binds to Nup116p-C. The observed injection trace of rate of heat evolved versus time (Fig. 3A) demonstrated that protein-peptide binding was occurring. By fitting the integrated values of enthalpic change for each injection point to a best fit curve, thermodynamic parameters could be determined. The dissociation constant was found to be 2.32 ± 0.05 μM, with a stoichiometry of one molecule of peptide/molecule of Nup116p-C. Values for ΔH and ΔTS were found to be −5.94 ± 0.04 and +1.75 ± 0.06 kcal/mol, respectively. Both the negative change in enthalpy and the positive change in entropy contributed favorably to a spontaneous binding reaction. The quality of the ITC data and the close match to the best fit curve, especially in comparison with ITC binding data for Nup116p-C and another likely ligand (see supplemental data), provide a strong argument in favor of the significance of the interaction between Nup116p-C and the Nup145p-C peptide.

With peptide binding established by ITC, the protein-peptide interaction was examined by NMR to determine whether peptide binding occurs at the expected site. Excess unlabeled peptide was added to an 15N-labeled sample of Nup116p-C, and an 15N HSQC spectrum was acquired. The resulting peptide-bound spectrum clearly differs from the unbound spectrum, with a strikingly large percentage of resonances noticeably shifted, although the overall pattern is similar (Fig. 3B). This result not only confirmed that Nup116p-C binds the Nup145p-C peptide, but because of the large number of shifted resonances, it also suggested that a significant conformational change in Nup116p-C may occur upon binding.

We next sought to identify the residues whose peaks had shifted most dramatically. An overall value for the chemical shift change was calculated for each residue, and residues with the largest changes were plotted on a ribbon representation of the protein (Fig. 3C). Because this is a map of backbone amide chemical shift changes, rather than a map of residues physically contacted by the peptide, residues outside the binding pocket were expected to be affected through indirect perturbations of their chemical microenvironments. Still, the immediate vicinity of the hypothesized binding site is the region most affected by chemical shift changes. In particular, residues along strand β5 of the small β-sheet are shifted substantially, as high as 0.6 ppm in the case of Thr66. In the NUP98 crystal structure, Val59, which corresponds to Cys57 in Nup116p-C, forms two backbone–backbone hydrogen bonds with the peptide residue Tyr1. In that structure, the peptide effectively functions as a third strand in the small β-sheet, interacting extensively with its neighboring strand. The results described here support the idea that the Nup145p-C peptide interacts in a similar manner with strand β5 of Nup116p-C.

The other region exhibiting the largest chemical shift changes is the C-terminal tail of Nup116p-C, residues 146–149. Although it is unclear what role this short stretch may play in binding, it is not surprising to see these residues shifted to such a large extent. In NUP98 and its autoproteolytic homologs, this corresponds to the region that is cleaved, forming the bound peptide visible in the structure. Upon Nup116p binding of
the Nup145p-C peptide, the environment of the tail region is likely to be substantially altered, as the chemical shift data show.

Notably, although a few residues in the long α3 helix exhibit significant backbone amide chemical shift changes, most do not. In NUP98, this helix forms hydrophobic side chain interactions with Leu5 of the peptide, and the final residue of the helix, Gln642, makes two side chain hydrogen bonds with the peptide backbone at Gly4. However, the backbone amide groups are principally involved in helical backbone hydrogen bonds; thus, although the overall position and environment of the helix likely change significantly upon peptide binding (Fig. 2), this would not necessarily be reflected in the 15N HSQC spectrum.

Peptide Binding Stabilizes the Conformation of Nup116p-C—In addition to the chemical shift changes described above, peptide addition also caused new resonances to appear in the 15N HSQC spectrum. Intriguingly, during the process of backbone assignment of peptide-bound Nup116p-C, we observed that the peaks from Ile69 and Tyr70 were easily assigned, unlike in the case of the unbound protein, for which these residues had no identifiable backbone amide resonances. These residues immediately follow strand β5 of the small β-sheet and, in fact, may form a continuation of this strand in the presence of the peptide. Indeed, in the NUP98 structure, the homologous residues Val784 and Tyr785 form the last 2 residues of the corresponding β-strand. Although these residues do not interact directly with the peptide in that structure, they do form multiple important interactions that stabilize the small β-sheet and its position relative to the rest of the structure.

The lack of resonances for Ile69 and Tyr70 in the unbound form of Nup116p-C suggested that conformational exchange occurs in the intermediate exchange regime. This exchange hypothesis is consistent with the observation that the adjacent residues, Ile68 and Ala71, exhibit the shortest T2 15N relaxation times of all measured residues, 23.2 and 21.3 ms, respectively. These are each less than half of the average T2 value of 49.0 ms (data not shown). In contrast, the presence of discrete backbone amide resonances for Ile69 and Tyr70 in the new spectra indicated that these residues may lock into a fixed conformation in the context of peptide binding. To confirm the appearance of resonances for Ile69 and Tyr70 upon peptide addition, the spectra of selectively labeled Nup116p-C samples were acquired in the absence and presence of the peptide. In the peptide-bound [15N]Ile HSQC spectrum (Fig. 4A), a new resonance is clearly visible, the location of which agrees with our sequential assignment of peptide-bound Ile69. Significant amide chemical shift changes in isoleucines 59, 68, and 89 are also evident. Ile69 corresponds to Ile774 in the NUP98 structure, with which Tyr785 forms a hydrogen bond; Ile68 is a neighboring residue on strand β5, and Ile69 lies on strand β6, with the side chain facing into the peptide-binding site. Each of these residues could play a role in stabilizing the small β-sheet or interacting with the peptide.

A new resonance can also be seen in the peptide-bound [15N]Tyr HSQC spectrum (Fig. 4B). Because there are fewer resonances and no degeneracy in these spectra, the appearance of a new resonance upon peptide binding is even clearer than in the isoleucine spectra. As expected, the new resonance is consistent with the previous assignment of Tyr70.

In a similar fashion, the tail residues 146–148 exhibit evidence of intermediate exchange in the unbound state, with the Glu147 amide peak unassigned and the Ala146 and Gln148 amide peaks almost too weak to detect. This raises the suggestion that the tail alternates among discrete conformations, possibly flipping in and out of the binding pocket. Upon peptide binding, a clear Glu147 peak appears, and the Ala146 and Gln148 peaks are strengthened significantly, suggesting a stabilized conformation.

Because the experiments described above give information only on backbone amide changes, we also examined the effect of peptide binding on Nup116p-C side chains. In particular, unbound and bound 13C HSQC methyl spectra were compared (Fig. 4C). Strikingly, the number of observed peaks dropped dramatically, from >140 in the unbound spectrum to ~85 in the bound spectrum. The number of expected resonances would ordinarily equal the number of distinct methyl groups in the protein side chains, which add to 84 (one methyl group each from 6 alanines and 8 threonines and two methyl groups each from 12 isoleucines, 13 leucines, and 10 valines). Thus, peptide addition brought the

---

**FIGURE 3.** Results of ITC experiment examining the interaction between Nup116p-C and a synthetic peptide from the N terminus of Nup145p-C. A: upper panel, the raw data show the titration of the peptide into a sample cell containing the Nup116p-C protein. Lower panel, data points were integrated to compute overall enthalpy change for each peptide injection. The heats of dilution of both the protein and peptide were each subtracted, and the results are plotted against the molar ratio of the peptide to the Nup116p-C protein. The best fit curve is shown as a solid line. B: 15N-labeled Nup116p-C was analyzed by 13N HSQC in the absence (filled black) and presence (open red) of the unlabeled Nup145p-C peptide. Dramatic chemical shift changes are evident for many of the amide resonances. C: residues whose amide resonances are shifted 0.12 ppm or more are plotted in red against a ribbon diagram of Nup116p-C. The “smooth” feature of MOLMOL was used to create this figure.
number of peaks in the spectrum into almost exact agreement with the number of methyl groups. As a key example, Leu$^{120}$, which is located toward the C-terminal end of helix $\alpha_3$, exhibits multiple resonances in the unbound state for each methyl group. H-$\alpha^3$ gives rise to multiple weak resonances, whereas H-$\delta^2$ is split into two clear, distinct peaks. Upon peptide addition, each set of peaks collapses into a single, discrete peak (Fig. 4C). These observations suggest that, in the unbound state, the Leu$^{120}$ side chain exists in multiple conformational states undergoing slow exchange, whereas in the peptide-bound state, a single conformation is stabilized, likely involving hydrophobic contacts between the leucine side chain and the peptide. Similarly, two unbound Ala$^{145}$ methyl peaks merge into a single peak upon peptide addition, suggesting that the conformation of this tail residue is stabilized by the presence of the peptide.

**Binding Studies with Isotopically Labeled Nup145p-C Peptide**—To better characterize the protein-peptide complex, isotopically labeled Nup145p-C peptide was produced and analyzed by acquisition of $^{15}$N HSQC spectra in the absence and presence of excess unlabeled Nup116p-C (Fig. 4D). The spectrum of the free peptide reveals 14 strong peaks, corresponding to backbone amides and the side chain of Trp$^3$. In the spectrum of the complexed peptide, nine strong peaks corresponding to Asn$^7$–Glu$^{15}$ are visible; a weaker Val$^7$ peak is evident; and four more extremely weak peaks can be seen corresponding to Trp$^3$–Leu$^6$ and the Trp$^3$ side chain. By comparing free and bound spec-
Solution Structure of Nup116p-C

In this work, we have used NMR to determine the solution structure of the pore-targeting domain of Nup116p. As expected, Nup116p-C has a fold similar to that of its human homolog, NUP98, whose structure was determined while these studies were in progress. HSQC spectra indicate that numerous residues of Nup116p-C in the unbound state appear to exist in an ensemble of discrete conformations undergoing slow to intermediate conformational exchange. By analogy with the human homolog, we reasoned that Nup116p-C may bind the N-terminal portion of Nup145p-C, and we were able to demonstrate an interaction using ITC and NMR. Intriguingly, the conformational exchange observed in spectra of unliganded Nup116p-C disappears or is greatly reduced upon peptide binding, suggesting that the presence of a binding partner stabilizes the domain conformation.

Conformational Diversity of Unliganded Nup116p-C—The strongest evidence of conformational exchange comes from the $^{15}$N HSQC and $^{13}$C HSQC spectra (Fig. 4). The $^{15}$N HSQC spectra reveal that no backbone amide peaks are present for several residues, including Ile$^{69}$ and Tyr$^{70}$, the residues immediately following strand $\beta$5. The absence of peaks for these residues, demonstrated unambiguously with selectively labeled spectra, as well as the very weak intensities of resonances for nearby residues, suggests that these amide groups exist in two or more discrete conformations, alternating among them on an intermediate time scale. $T_2$ relaxation data support this idea, as the adjacent residues, Ile$^{68}$ and Ala$^{71}$, have extremely short transverse relaxation times, ruling out the possibility that this region of the protein is highly flexible in the manner of a disordered loop region.

The $^{13}$C HSQC spectrum of unbound Nup116p-C also provides evidence of conformational exchange, with many more peaks than side chain methyl groups in the spectrum. Interestingly, the time regimes of the backbone and side chain dynamics differ: the backbone amides of Ile$^{69}$ and Tyr$^{70}$ exhibit intermediate exchange, as evidenced by the absence of expected peaks, whereas the side chain methyl groups of Leu$^{120}$ and Ala$^{145}$ exhibit slow exchange, as shown by the presence of additional peaks.

Based on the ensemble of calculated structures, loop 3 in particular exhibits a dramatic degree of conformational diversity, and the lateral position of helix $\alpha$3 varies appreciably across structures relative to the highly fixed six-stranded $\beta$-sheet. Consistent with the lack of backbone amide resonances for Ile$^{69}$ and Tyr$^{70}$, the conformations of these residues vary substantially from one calculated structure to the next. These conformationally variable regions all map to the putative ligand-binding site.

The presence of multiple conformations may explain why extensive efforts to crystallize unliganded Nup116p-C were unsuccessful. Conformational changes from one molecule to the next or changes within a single molecule over time could easily disrupt the regular network of subunit-subunit interactions that must form for a crystal to grow. In the case of the crystal structure of the human homolog, NUP98, only the liganded form of the pore-targeting domain could be solved. Although crystals of unliganded NUP98 were able to be grown, the crude structural model derived from these crystals could not be properly refined, suggesting substantial disorder in the structure (18). Thus, unliganded NUP98 may exhibit multiple conformations in a similar fashion to Nup116p-C.

Conformational diversity in Nup116p-C could exist for numerous reasons. One possibility is that the observed binding site plasticity enables Nup116p-C to bind to multiple targets within the NPC. Previous work has shown that Nup116p-C interacts with the nucleoporin Nup82p; furthermore, studies with human homologs show that NUP98 utilizes the same binding site to interact with NUP96 (its C-terminal autoprolytic partner) and NUP88 (the homolog of Nup82p) (13). Thus, the Nup116p-C binding site described in this work is likely involved in interactions with at least two nucleoporins (see supplemental data) and possibly others as well. To accommodate distinct ligands, Nup116p-C may need to exist in an ensemble of conformations in the unbound state.

Another possibility is that Nup116p is a shutting protein, much like its human homolog, NUP98 (33, 34). Although immunogold electron microscopy localization data show that Nup116p is found on both faces of the NPC, with an asymmetrical distribution skewed toward the cytoplasmic face (1), it is unclear whether this is a static or dynamic distribution. In the case of NUP98, the mobility of the protein has been hypothesized to link RNA transcription and export (33). If Nup116p behaves in a similar fashion, the conformational diversity of its C-terminal domain may result in lower affinity binding and faster dissociation rates. It is possible that the “tail” of Nup116p-C (residues 146–149) contributes to this dissociation by flipping in and out of the binding site in the absence of ligand and acting as a covalently attached competitive inhibitor in the presence of ligand. This idea is supported by the substantial chemical shift changes between unbound and Nup145p-C peptide–bound states seen for these tail residues as well as the strengthened amide resonances in the bound state. Further work, including in vivo experiments, is necessary to definitively establish whether Nup116p is in fact a mobile nucleoporin as well as the role of the tail residues in binding and dissociation.
Solution Structure of Nup116p-C

Interaction with the Nup145p-C Peptide—The crystal structure of NUP98, which reveals a binding interaction with the first several N-terminal residues of NUP98, led to the hypothesis that NUP116p-C may bind NUP145p-C in a similar fashion. Using ITC and NMR, we demonstrated an interaction using a 15-mer peptide from Nup145p-C. It should be noted that this interaction provides a convenient explanation for the unusual cytoplasmically biased localization pattern of Nup116p in the NPC. Nup145p-C is a symmetric nucleoporin, whereas Nup82p, the other known binding partner of Nup116p, is exclusively found on the cytoplasmic face of the NPC. Combining these localizations results in a cytoplasmically biased distribution, consistent with the observed pattern (Fig. 5). The relatively weak binding interaction observed between Nup116p-C and the Nup145p-C peptide gives further support to the notion discussed above that Nup116p may be a dynamic component of the pore.

In the presence of the Nup145p-C peptide, the effects of conformational exchange observed in the unbound spectra of Nup116p-C are greatly reduced, suggesting that a single bound conformation may be present. This is underscored in the selectively labeled \([15N]Ile\) and \([15N]Tyr\) HSQC spectra, in which peaks that are entirely absent in the unbound case appear as strong, single peaks in the complex. Furthermore, in the \([13C]HSQC\) methyl spectra, many more peaks than expected are present in the unbound case, whereas the number of bound peaks corresponds almost precisely to the number of side chain methyl groups. The two methyl groups of Leu\(^{120}\) provide a particularly clear example of this, with several unbound peaks coalescing into a single bound peak for each methyl group. Likewise, the two unbound peaks corresponding to the methyl group of Ala\(^{145}\) coalesce into one bound peak.

Interestingly, our studies of isotopically labeled peptide revealed evidence of peptide conformational exchange upon binding Nup116p-C. The resonances for the first several N-terminal peptide residues are considerably weaker than expected for a complex consisting of a 16.8-kDa protein and a 1.7-kDa peptide. As an extreme example, the Gly\(^{6}\) resonance is fully 25 times weaker in normalized intensity than the corresponding unbound peptide resonance (Fig. 4D, inset). Indeed, due to the weakness of the peptide resonances directly involved in binding, further structural analysis of the protein-peptide complex could not be performed.

The observed weak peaks likely reflect exchange between a single bound state and a large ensemble of disordered conformations in the unbound state. This emergence of exchange upon binding is an initially unexpected result given the reduction of exchange phenomena in the spectra of labeled Nup116p-C bound to the unlabeled peptide. However, we attribute the peptide exchange to rapid association with and dissociation from Nup116p-C, with a dissociation rate in approximately the same millisecond time regime as the NMR experiments, consistent with the relatively weak micromolar scale binding constant of the interaction. The much larger Nup116p-C protein appears to remain in its observed bound conformation even as the peptide associates and dissociates rapidly, likely because the protein responds more slowly than the peptide to changes in its environment.

Pre-existing Equilibrium Model—In general terms, Nup116p-C appears to be an instance of a growing class of protein structures that support the "pre-existing equilibrium" model (35–37) as a mechanism of protein-protein interaction. In this model, proteins exist in an ensemble of conformations in an unbound state, and the presence of a particular ligand alters this distribution to favor one or more bound states. This model challenges the dogma that sequence uniquely specifies structure and that structure in turn uniquely specifies function. Instead, a given sequence may fold into a variety of related structures, each of which may have distinct binding or other activities (38).

The notion of pre-existing equilibrium, at least as applied to Nup116p-C, goes beyond the induced fit principle of plasticity in the binding site or in the structure as a whole, suggesting that multiple discrete conformations, which in some cases may differ markedly from one another, exist for the same protein sequence. An equilibrium exists among these conformations in the unbound state, all typically at similar free energies, but only one of which may possess a particular activity (39). A ligand "selects" for its binding-competent conformation by binding only to that particular unbound state when it exists, thus biasing the conformational distribution (35). The NMR results for Nup116p-C support this model, as clear evidence is observed for multiple discrete conformations in slow or intermediate exchange in the unbound state and a single conformation (at least in the case of individual residues or chemical groups) in the bound state.

Conclusion—We have presented structural and binding studies of a nucleoporin domain that may reveal important principles of NPC architecture. The NPC as a whole is a remarkably dynamic structure, with pronounced variations in overall shape and diameter having been observed (40); thus, it is fascinating to observe conformational diversity at the level of its constituent parts. Because so little high resolution structural information about nucleoporins is currently available, further studies of other nucleoporins will be required to determine whether these nucleoporins exhibit the type of conformational diversity seen for Nup116p-C.

The NPC has provided a daunting challenge to structural biologists, given its enormous size, large-scale conformational changes, and lack of ordered structure in many regions of its constituent nucleoporins. It is anticipated that the approach taken here could be successfully extended to other nucleoporin domains and nucleoporin-nucleoporin interfaces, with the goal of deepening our understanding of the structural underpinnings of nucleocytoplasmic transport.

Acknowledgments—We thank Marc Damelin for helpful discussions, Sven Hyberts for assistance with IBIS assignment software, and David Filman for critical review of the manuscript.

REFERENCES
1. Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000) J. Cell Biol. 148, 635–651
Solution Structure of Nup116p-C

2. Allen, T. D., Cronshaw, J. M., Bagley, S., Kiseleva, E., and Goldberg, M. W. (2000) J. Cell Sci. 113, 1651–1659
3. Strawn, L. A., Shen, T., and Wente, S. R. (2001) J. Biol. Chem. 276, 6445–6452
4. Bailer, S. M., Siniossoglou, S., Podtelejnikov, A., Hellwig, A., Mann, M., and Hurt, E. (1998) EMBO J. 17, 1107–1119
5. Iovine, M. K., Watkins, J. L., and Wente, S. R. (1995) J. Cell Biol. 131, 1699–1713
6. Wente, S. R., and Blobel, G. (1993) J. Biol. Chem. 278, 275–284
7. Bailer, S. M., Balduf, C., Katahira, J., Podtelejnikov, A., Rollenhagen, C., Mann, M., Pante, N., and Hurt, E. (2000) J. Biol. Chem. 275, 23540–23548
8. Ho, A. K., Shen, T. X., Ryan, K. J., Kiseleva, E., Levy, M. A., Allen, T. D., and Wente, S. R. (2000) Mol. Cell. Biol. 20, 5736–5748
9. Hurwitz, M. E., and Blobel, G. (1995) J. Cell Biol. 130, 1275–1281
10. Stoffler, D., Fahrenkrog, B., and Aebi, U. (1999) Curr. Opin. Cell Biol. 11, 391–401
11. Wente, S. R., Rout, M. P., and Blobel, G. (1992) J. Cell Biol. 119, 705–723
12. Fontoura, B. M., Blobel, G., and Matunis, M. J. (1999) J. Cell Biol. 144, 1097–1112
13. Emtage, J. L., Bucc, M., Watkins, J. L., and Wente, S. R. (1997) J. Cell Sci. 110, 911–925
14. Fontoura, B. M., Blobel, G., and Matunis, M. J. (1997) J. Cell Sci. 110, 911–925
15. Hodel, A. E., Hodel, M. R., Griffiths, E. R., Henning, K. A., Ratner, G. A., Xu, S., and Powers, M. A. (2002) Mol. Cell 10, 347–358
16. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
17. Johnson, B. A. (2004) Methods Mol. Biol. 278, 313–352
18. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1995) Protein NMR Spectroscopy: Principles and Practice, Academic Press, Inc., San Diego, CA
19. Bax, A., and Grzesiek, S. (1993) Acc. Chem. Res. 26, 131–138
20. Ferentz, A. E., and Wagner, G. (2000) Q. Rev. Biophys. 33, 29–65
21. Hyberts, S. G., and Wagner, G. (2003) J. Biomol. NMR 26, 355–344
22. Nilges, M. (1993) Proteins 17, 297–309
23. Cornilcussc, G., Delaglio, F., and Bax, A. (1999) J. Biomol. NMR 13, 289–302
24. Guntert, P., Munzenthaler, C., and Wuthrich, K. (1997) J. Mol. Biol. 273, 283–298
25. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
26. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 51–55, 29–32
27. Krualis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
28. Laskowski, R. A., Rullmannann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
29. Rosenblum, J., and Blobel, G. (1999) Proc. Natl. Acad. Sci. U.S.A. 96, 11370–11375
30. Griffis, E. R., Altan, N., Lippincott-Schwartz, J., and Powers, M. A. (2002) Mol. Biol. Cell 13, 1282–1297
31. Fontoura, B. M., Dales, S., Blobel, G., and Zhong, H. (2001) Proc. Natl. Acad. Sci. U.S.A. 98, 3208–3213
32. Goh, C. S., Milburn, D., and Gerstein, M. (2004) Curr. Opin. Struct. Biol. 14, 104–109
33. Ma, B., Shatsky, M., Wolfson, H. J., and Nussinov, R. (2002) Protein Sci. 11, 184–197
34. Tsai, C. J., Kumar, S., Ma, B., and Nussinov, R. (1999) Protein Sci. 8, 1181–1190
35. James, L. C., Roversi, P., and Tawfik, D. S. (2003) Science 299, 1362–1367
36. James, L. C., and Tawfik, D. S. (2003) Trends Biochem. Sci. 28, 361–368
37. Kiseleva, E., Goldberg, M. W., Cronshaw, J., and Allen, T. D. (2000) Crit. Rev. Eukaryotic Gene Expression 10, 101–112
