Published in final edited form as:
Proteins. 2012 August ; 80(8): 2110–2116. doi:10.1002/prot.24102.

Atomic Structure of the Nuclear Pore Complex targeting domain of a Nup116 homologue from the yeast, *Candida glabrata*

Parthasarathy Sampathkumar[1,2,*], **Seung Joong Kim**[3,4,5], **Danalyn Manglicmot**[2], **Kevin T. Bain**[2], **Jeremiah Gilmore**[2], **Tarun Gheyi**[2], **Jeremy Phillips**[3,4,5,6], **Ursula Pieper**[3,4,5], **Javier Fernandez-Martinez**[7], **Josef D. Franke**[7], **Tsutomu Matsui**[8], **Hiro Tsuruta**[8,⊥], **Shane Atwell**[2], **Devon A. Thompson**[2], **J. Spencer Emtage**[2], **Stephen R. Wasserman**[9], **Michael P. Rout**[7], **Andrej Sali**[3,4,5], **J. Michael Sauder**[2], **Steven C. Almo**[1], and **Stephen K. Burley**[2]

[1]Department of Biochemistry, Ullmann Building, Room 409, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461

[2]Translational Sciences and Technology (TS&T), Eli Lilly and Company, Lilly Biotechnology Center, 10300 Campus Point Drive, Suite 200, San Diego, CA 92121, USA

[3]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA 94158

[4]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158

[5]California Institute for Quantitative Biosciences, University of California, San Francisco, CA 94158

[6]Graduate Group in Biological and Medical Informatics, University of California, San Francisco, CA 94158

[7]Laboratory of Cellular and Structural Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

[8]Stanford Synchrotron Radiation Lightsource, Stanford Linear Accelerator Center, 2575 Sand Hill Road, MS 69, Menlo Park, CA 94025-7015, USA

[9]LRL-CAT, Eli Lilly and Company, Advanced Photon Source, Argonne National Laboratory, Building 401, 9700 South Cass Avenue, Argonne, IL 60439 USA

Abstract

The nuclear pore complex (NPC), embedded in the nuclear envelope, is a large, dynamic molecular assembly that facilitates exchange of macromolecules between the nucleus and cytoplasm. The yeast NPC is an eight-fold symmetric annular structure composed of ~456 polypeptide chains contributed by ~30 distinct proteins termed nucleoporins (Nups). Nup116, identified only in fungi, plays a central role in both protein import and mRNA export through the NPC. Nup116 is a modular protein with N-terminal “FG” repeats containing a Gle2p-binding sequence motif (GLEBS motif) and a NPC targeting domain at its C-terminus. We report the crystal structure of the NPC targeting domain of *Candida glabrata* Nup116, consisting of residues 882-1034 [CgNup116(882-1034)], at 1.94 Å resolution. The X-ray structure of

[*]Corresponding author’s psampath@aecom.yu.edu, Department of Biochemistry, Ullmann Building, Room 409, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461., Phone: +1-(718) 430-2745; Fax: +1-(718)-430-8565.
[⊥]Deceased August 25, 2011

Protein Data Bank Codes

Atomic coordinates and structure factors of CgNup116(882-1034) were deposited to the PDB on 09 June 2010 with accession codes 3NF5. The NYSGXRC target identifier for CgNup116 in TargetDB (http://targetdb.pdb.org) is “NYSGXRC-15100c”. Expression clone sequences and selected interim experimental results are available in PepcDB (http://pepcdb.pdb.org/).

CgNup116(882-1034) is consistent with the molecular envelope determined in solution by Small Angle X-ray Scattering (SAXS). Structural similarities of CgNup116(882-1034) with homologous domains from *Saccharomyces cerevisiae* Nup116, *S. cerevisiae*Nup145N, and human Nup98 are discussed.

Keywords

Nuclear Pore Complex; Nup116; Nup98; Nup100; Nup145; mRNA export; structural genomics

INTRODUCTION

Transport of macromolecules between nucleus and cytoplasm is an essential eukaryotic process facilitated by the nuclear pore complex (NPC). In addition to its role in normal physiology, NPC loss of function has been implicated in cancer and autoimmune disease[1,2]. In yeast (e.g., *Saccharomyces*), NPCs are large, eight-fold symmetric dynamic macromolecular assemblies composed of at least 456 polypeptide chains derived from multiple copies of ~30 distinct nucleoporins (Nups) [3,4]. Several of these components share similar structural motifs and form stable subcomplexes that contribute to the overall organization of the assembly, which includes two outer rings (the nuclear and cytoplasmic rings), two inner rings, and a membrane-associated ring[5,6].

Nup116[7], a nucleoporin identified only in fungi, is involved in both protein import and in mRNA export[8]. Nup116 shows an asymmetric radial distribution within the NPC, with a bias towards the cytoplasmic face[9]. Nup116 is homologous to yeast Nup100; it is also homologous to yeast Nup145N and human Nup98, both of which are derived from a larger precursor by autoproteolysis[10,11]. *Candida glabrata* Nup116 is a modular protein with N-terminal “FG” repeats (residues 2-643) and a C-terminal domain (residues 890-1035) supporting NPC localization[12]. The “FG” repeats are thought to transiently interact with nuclear transport factors to ensure the transport of specific proteins and ribonucleoprotein (RNP) complexes[13]. A distinguishing feature of Nup116, when compared to both Nup100 and Nup145N, is the presence of an N-terminal Gle2p-binding[14] sequence (GLEBS, ~60 amino acid residue motif), responsible for targeting the RNA export factor Rae1/Gle2p to the NPC. A GLEBS motif is also present in human Nup98[15]. The crystal structure of a human Rae1:Nup98-GLEBS domain complex revealed that GLEBS contains a hairpin motif required for the interaction with Rae1[16].

While the N-terminal domain of Nup116 mediates interactions with nuclear transport factors, its C-terminal domain (referred to as NPC targeting domain) localizes Nup116 to the NPC and plays an essential role in NPC assembly. The NPC targeting domain of Nup116 interacts directly with the Nup82-Nsp1-Nup159 complex[12]. Herein, we report the 1.94 Å resolution crystal structure of the NPC targeting domain of *Candida glabrata* Nup116 [residues 882-1034; CgNup116(882-1034)] and the results of complementary solution studies using Small Angle X-ray Scattering (SAXS). We also present detailed structural comparisons with previously reported structures of *Saccharomyces cerevisiae* Nup116 (ScNup116; residues 967-1113; apo form determined by NMR spectroscopy[17] and the heterotrimer with Nup82:Nup159 complex determined by X-ray crystallography[18]), Nup145N (ScNup145N; residues 443-605 X-ray[19]), and human Nup98 (HsNup198; residues 716-870; X-ray[10, 11]).

MATERIALS AND METHODS

Cloning, expression, and purification of CgNup116(882-1034)

The gene encoding Nup116 from *Candida glabrata* was cloned from genomic DNA of strain 2001D-5_CBS 138 (American Type Culture Collection, USA). The desired truncation (encoding residues 882-1034) was PCR amplified using GATGGCATTGATGATCTAGAATTTG and CTAATGCATGATCAACAGTGAAGCAG as forward and reverse primers, respectively. The purified PCR product was TOPO® (Invitrogen, USA) cloned into pSGX3, a derivative of pET26b(+), yielding a protein with a non-cleavable C-terminal hexa-histidine tag. The resulting plasmid was transformed into BL21(DE3)-Condon+RIL (Invitrogen, USA) cells for expression. Production of Se-Met protein[20] was carried out in 1L of HY media at 22°C containing 50μg/ml of kanamycin and 35μg/ml of chloramphenicol. Protein expression was induced by addition of 0.4mM IPTG. Cells were harvested after 21 hours by centrifugation at 4°C.

For purification, the *E. coli* cell pellet was resuspended in 30mL of cold buffer containing 20mM Tris HCl *p*H 8.0, 500mM NaCl, 25mM imidazole, and 0.1% (v/v) Tween20 and the cells were lysed by sonication. Cell debris was removed by centrifugation at 4°C. The supernatant was applied to a 5mL HisTrapHP column (GE Health Care, USA) charged with nickel and pre-equilibrated with 20mM Tris HCl *p*H 8.0, 500mM NaCl, 10% (v/v) glycerol, and 25mM imidazole. The sample was washed with 5 column volumes (CV) of 20mM Tris HCl *p*H 8.0, 500mM NaCl, 10% (v/v) glycerol, and 40mM imidazole, and subsequently eluted with 2 CV of same buffer with an imidazole concentration of 250mM. Eluted protein was further purified over a 120ml Superdex 200 size exclusion column equilibrated with 10mM HEPES *p*H 7.5, 150mM NaCl, 10% (v/v) glycerol, and 5mM DTT (protein storage buffer). SDS-PAGE analysis demonstrated greater than 95% purity. Protein fractions corresponding to the central portion of the size exclusion chromatography profile were pooled, concentrated by AMICON spin filtration and aliquots frozen in liquid nitrogen and stored at -80°C.

Crystallization, data collection, and structure determination

Initial crystals of CgNup116(882-1034) were obtained in several PEG-containing conditions *via* sitting drop vapor diffusion at 21°C (~10.4 mg/ml; 0.3 μL protein + 0.3 μL reservoir solution). Subsequent optimization was carried out with an additive screen (Hampton Research, USA) and macro-seeding. Diffraction quality crystals were obtained with 100 mM MES pH 6.2, 25% (w/v) PEG MME 2K and 200 mM sodium potassium tartrate. The final sitting-drops contained 1.0 μL of CgNup116(882-1034) at 10.85 mg/mL, 0.6 μL of reservoir solution, and 0.4 μL of 5% (v/v) ethyl acetate from the additive screen. Crystals were cryo-protected by addition of glycerol [final concentration ~30% (v/v)] and flash-cooled by immersion in liquid nitrogen. Diffraction data were recorded at the LRL-CAT 31-ID beamline (Advanced Photon Source (APS)) and processed with MOSFLM[21] and SCALA (CCP4) [22]. Structures were determined by molecular replacement using PHASER[23] with a poly-alanine model of ScNup145N (PDB Code 3KEP)[19]. Initial model building was carried out with ARP/wARP[24], followed by manual rebuilding with COOT[25]. The atomic model of CgNup116(882-1034) was refined to convergence using REFMAC5[26] and exhibited excellent stereochemistry (Table 1). Illustrations were prepared with PyMol[27].

Small Angle X-ray Scattering (SAXS)

SAXS measurements of CgNup116(882-1034) were carried out at Beamline 4-2 of the Stanford Synchrotron Radiation Lightsource (SSRL). The beam energy and current were 11 keV and 200mA, respectively. A silver behenate sample was used to calibrate the ***q***-range and detector distance. Data collection was controlled with Blu-Ice[28]. We used an automatic

sample delivery system equipped with a 1.5 mm-diameter thin-wall quartz capillary within which a sample aliquot was oscillated in the X-ray beam to minimize radiation damage. The sample was placed at 1.7m from a Rayonix225 (MAR-USA, USA) CCD detector with a binned pixel size of 293 μm × 293 μm. Ten 3 sec exposures were made for each of four protein samples maintained at 15 °C. Each of the 10 diffraction images was scaled by the transmitted beam intensity, using SASTool (http://ssrl.slac.stanford.edu/~saxs/analysis/sastool.htm, formerly MarParse), and averaged to obtain fully processed data in the form of intensity *versus* $\boldsymbol{q}$ [$\boldsymbol{q}=4\pi\sin(\theta)/\lambda$, where θ is one-half of the scattering angle and λ is the X-ray wavelength]. The buffer SAXS profile was obtained in the same manner and subtracted from a protein profile. SAXS profiles of CgNup116(882-1034) were recorded at protein concentrations of 0.5, 1.0, 2.0, and 5.0 mg/ml in the protein storage buffer. Mild concentration dependence of the profiles was eliminated by extrapolating to zero concentration. The average of the lower scattering angle parts ($\boldsymbol{q}<0.15\text{Å}^{-1}$) of the lower concentration profiles (0.5-1.0 mg/ml) and the average of the higher scattering angle parts ($\boldsymbol{q}>0.12\text{Å}^{-1}$) of the higher concentration (1.5-5.0 mg/ml) profiles were merged to obtain the final experimental SAXS profile. The merged experimental SAXS profile was compared with SAXS profiles calculated for the monomer (Chain A) and for the crystallographic asymmetric unit (Chains A and B) of CgNup116(882-1034) with IMP FoXS (http://salilab.org/foxs)[29,30]. A complete monomer model of CgNup116 (882-1034), which included a C-terminal hexa-histidine tag (Gly-His-His-His-His-His-His), eight side chains not modeled in the crystal structure, and two Se-Met residues, was generated using the crystal structure with the automodel function of MODELLER[31] and customized scripts in IMP.[32] Inclusion of the missing atoms further improved the fit of the calculated and experimental profiles (χ value improved from 1.33 to 1.11). The shape of CgNup116(882-1034) was calculated from the merged experimental SAXS profile by running DAMMIF[33] and GASBOR[34] 20 times individually, followed by superposition and averaging with DAMAVER.[35] The shape of CgNup116(882-1034) was also computed from the merged experimental SAXS profile by SASTBX (http://sastbx.als.lbl.gov/wiki/) and compared with DAMMIF / GASBOR shapes.

RESULTS AND DISCUSSION

Structure of CgNup116(882-1034)

The crystal structure of CgNup116(882-1034) was determined at 1.94Å resolution [Fig. 1(A), Table 1]. The monoclinic crystals (space group $P2_1$) contain two molecules per asymmetric unit. Chain A could be traced continuously from Asp882 to Leu1034, while in chain B residues 960-963 appear disordered. Otherwise, the A and B chains are essentially identical, with a root-mean-square deviation (r.m.s.d.) of ~0.49 Å for 152 C_α atomic pairs, calculated using the SSM[36] routine as implemented in COOT. The N-terminal segment (residues 882 to 893) of CgNup116 (882-1034) is well defined in the electron density maps and adopts non-canonical secondary structure (i.e., random coil). The overall fold of CgNup116(882-1034) contains two central anti-parallel β-sheets flanked by α-helices [Fig. 1(A)]. A six stranded β-sheet is formed by β1-β2-β3-β6-β8-β7 and a two stranded β-sheet is formed by β4-β5. Helices α1, α2, and α3 (α3 is a short 3_{10} helix within loop L1) form a cap near the N-terminus and helix α4 caps the six stranded β-sheet near the C-terminus. CgNup116(882-1034) possesses three long loops including, L1 (residues 930-939 between β3 and β4), L2 (residues 958-974 between β5 and β6), and L3 (residues 980-999 between β6 and α4).

The interface area and the gap volume index[37] between the A and B chains of CgNup116(882-1034), calculated using the NOXclass (http://noxclass.bioinf.mpi-sb.mpg.de/index.php)[38], are ~630 Å^2 and 7.5, respectively. The two copies of CgNup116(882-1034) observed in the crystal asymmetric unit are not likely to represent a

physiological dimer, as the merged experimental SAXS profile [Fig. 1(B)] is well matched (χ = 1.11) to the SAXS profile calculated from the complete monomer model of CgNup116(882-1034). The SAXS profile calculated from the complete dimer model resulted in an unacceptably high χ value of 7.67. The measured radius of gyration (R_g) of 18.38±0.24 Å, determined with AutoRg,[39] is almost identical to the value of 18.1 Å calculated from the complete monomer model of CgNup116(882-1034) (the calculated value of R_g for the A and B chain complex, representing the crystallographic asymmetric unit, is 20.7 Å). Moreover, the "*ab initio*" shape computed from the merged experimental SAXS profile with DAMMIF[33] (not shown), GASBOR [Fig. 1(C)],[34] and SASTBX [Fig. 1(D)] shows very considerable similarity to our X-ray structure of the CgNup116(882-1034) monomer. Finally, based on the merged experimental SAXS profile, OLIGOMER[40] estimates 100% monomer composition. Thus, our SAXS analyses of the solution behavior of CgNup116(882-1034) and the X-ray crystallographic structure of the monomer are fully consistent with each other.

Comparison of CgNup116(882-1034) with the structures of ScNup116

A pairwise local alignment of CgNup116(882-1034) and ScNup116, computed using LALIGN (http://www.ch.embnet.org/software/LALIGN_form.html), shows sequence identity of 60.1%, while sequence identities of CgNup116(882-1034) drop to 34.5% and 30.6% for the autoproteolytic domains of ScNup145N and HsNup98, respectively. A multiple structural alignment was obtained with the Multiprot[41] (http://bioinfo3d.cs.tau.ac.il/MultiProt/) and STACCATO programs to enable identification of structurally conserved residues across CgNup116(882-1034), ScNup116 bound to the Nup82-Nup159 complex,[18] ScNup145N[19], and HsNup98[11]. The alignment demonstrates conservation of the overall fold, despite varying pairwise sequence identities, with an average r.m.s.d. of 1.28 Å over 102 alignment positions with C_α atoms from all four structures. The alignment also reveals positions with a conserved residue type as well as gaps in loop regions [Fig. 2(A)]. The alignment of CgNup116(882-1034) with ScNup116 bound to the Nup82-Nup159 complex suggests that ScNup116 undergoes only a minimal conformational change upon binding to the N-terminal seven bladed β-propeller domain of Nup82, with only α4-helix (structurally equivalent to helix αB in ScNup116)[18] and loop L3 showing significant structural differences [Fig. 2(B)]. ScNup116 contributes (i) a hydrophobic groove on its surface between the β5-strand and the αB-helix, which forms a binding pocket for the "FGL" motif from the 3D4A loop of ScNup82, and (ii) loop L3, referred to as the "K-loop", situated between the β6-strand and αB-helix (in particular, the conserved Lys1063 of ScNup116 interacts with Asp204 of ScNup82).[16] 7 out of 10 ScNup116 residues involved in its interaction with ScNup82 are identical in CgNup116 [Fig. 2(A) and Supplementary Figure S1]. ScNup116 residues Lys1029, Cys1031, and Ile1033 (all from β5-strand) are replaced by Met953, Val955, and Leu957, respectively, at structurally equivalent positions in CgNup116, suggesting possible species specific differences in Nup116:Nup82 interactions.

A structural comparison (not shown) of CgNup116(882-1034) with the solution NMR structure of ScNup116 (PDB Code 2AIV)[17] also revealed a similar overall structure. The N-terminal α-helices and the β-strands of both central β-sheets are arranged similarly, with the largest difference between the two structures occurring in the L2 loop connecting the β5 and β6 strands. Residues comprising the β5-strand and the N-terminus of the L2 loop have been implicated in the binding of ScNup145C-peptide to ScNup116[17]. In addition, loop L3 and the polypeptide chain segment following α4-helix exhibit significant conformational differences, which is consistent with the conformational flexibility revealed by the solution NMR structures in this region[17].

Comparison of CgNup116(882-1034) with yeast Nup145 and human Nup98

Both the ScNup145N and the human Nup98 are generated from larger precursors *via* post-translational autoproteolysis at a conserved Phe-Ser peptide bond[10,11] [Fig. 2(A)]. Nup116, Nup100, and Nup145N are paralogues, and they share an orthologous relationship with human Nup98. CgNup116(882-1034) and ScNup145N's autoproteolytic domain share moderate sequence identity (34.5%). However, overall structures of CgNup116(882-1034) and ScNup145N (PDB Code 3KEP)[19] are virtually identical [Fig. 2(C)]. The only notable conformational difference between these two structures is in loop L1, which includes the 3_{10} helix α3. This difference could result from an insertion within loop L1 in ScNup145N [Fig. 2(A)].

CgNup116(882-1034) is also similar to human Nup98 (PDB Code 2Q5X)[11] [Fig. 2(A) and Fig. 2(D)]. The main structural differences are the consequence of deletions in loops L1 and L2 as well as an insertion in loop L3. In addition, the structures of CgNup116(882-1034) and human Nup98 differ in the random-coil segment that precedes strand β1 [Fig. 2(D)]. In the human Nup98 autoproteolytic domain, these residues (712-723) fold towards the core. In contrast, the equivalent residues (882-893) of CgNup116(882-1034) project away from the core [Fig. 2(D)]. Moreover, the conformational plasticity of residues 882-893 in CgNup116(882-1034) is revealed by the absence of any features corresponding to these residues in the solution shapes computed from the SAXS profiles [Fig. 1(C) and (D)] of CgNup116(882-1034). These results suggest that the residues preceding the β1 strand may adopt different conformations in the NPC targeting domain of CgNup116 and the autoproteolytic domain of human Nup98.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the Rout and Sali laboratories for their help and advice. Funding for the NYSGXRC and NYSGRC were provided by NIH Grants U54 GM074945 (PI: S.K. Burley) and U54 GM094662 (PI: S.C. Almo), respectively. Additional funding for this work was provided by NIH R01 GM062427 (MPR), NIH R01 GM083960 (A. Sali), and NIH U54 RR022220 (A. Sali and M.P. Rout). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Basic Energy Sciences. Access to the LRL-CAT beam line facilities at Sector 31 of the APS was provided by Eli Lilly, which operates the facility. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program (P41RR001209). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

References

1. Xu S, Powers MA. Nuclear pore proteins and cancer. Seminars in Cell and Develop Biology. 2009; 20:620–630.
2. Cronshaw JM, Matunis JM. The nuclear pore complex: disease associations and functional correlations. Trends in Endocrinology and Metabolism. 2004; 15:34–39. [PubMed: 14693424]
3. Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao YM, Chait BT. The yeast nuclear pore complex: Composition, architecture, and transport mechanism. J Cell Biol. 2000; 148:635–651. [PubMed: 10684247]
4. D'Angelo MA, Hetzer MW. Structure, dynamics and function of nuclear pore complexes. Trends Cell Biol. 2008; 18:456–466. [PubMed: 18786826]

5. Alber F, Dokudovskaya S, Veenhoff LM, Zhang WH, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, Sali A, Rout MP. The molecular architecture of the nuclear pore complex. Nature. 2007; 450:695–701. [PubMed: 18046406]
6. Alber F, Dokudovskaya S, Veenhoff LM, Zhang WZ, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, Rout MP, Sali A. Determining the architectures of macromolecular assemblies. Nature. 2007; 450:683–694. [PubMed: 18046405]
7. Wente SR, Rout MP, Bobel G. A new family of yeast nuclear pore complex proteins. J Cell Biol. 1992; 119:705–723. [PubMed: 1385442]
8. Strawn L, Shen T, Wente SR. The GLFG regions of Nup116p and Nup100p serve as binding sites of both Kap95p and Mex67p at the nuclear pore complex. J Biol Chem. 2001; 276:6445–6452. [PubMed: 11104765]
9. Ho AK, Shen T, Ryan KJ, Kiseleva E, Levy MA, Allen TD, Wente SR. Assembly and preferential localization of Nup116p on the cytoplasmic face of the nuclear pore complex by interaction with Nup82p. Mole Cell Biol. 2000; 20:5736–5748.
10. Hodel EA, Hodel MR, Griffis ER, Hennig KA, Ratner GA, Xu S, Powers MA. The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. Mol Cell. 2002; 10:347–358. [PubMed: 12191480]
11. Sun Y, Guo H-C. Structural constraints on autoprocessing of the human nucleoporin Nup98. Protein Sci. 2008; 17:494–505. [PubMed: 18287282]
12. Bailer SM, Balduf C, Katahira J, Podtelejnikov A, Rollenhangen C, Mann M, Pante N, Hurt E. Nup116p associates with the Nup82p-Nsp1p-Nup159p nucleoporin complex. J Biol Chem. 2000; 275:23540–23548. [PubMed: 10801828]
13. Suntharalingam M, Wente SR. Peering through the pore: nuclear pore complex structure, assembly, and function. Develop Cell. 2003; 4:775–789.
14. Bailer SM, Siniossoglou S, Podtelejnikov A, Hellwig A, Mann M, Hurt E. Nup116p and nup100p are interchangeable through a conserved motif which constitutes a docking site for the mRNA transport factor gle2p. The EMBO Journal. 1998; 17(4):1107–1119. [PubMed: 9463388]
15. Blevin MB, Smith AM, Phillips EM, Powers MA. Complex formation among the RNA export proteins Nup98, Rae1/Gle2 and TAP. J Biol Chem. 2003; 278:20979–20988. [PubMed: 12637516]
16. Ren Y, Seo H-S, Blobel G, Hoelz A. Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. Proc Natl Acad Sci. 2010; 107:10406–10411. [PubMed: 20498086]
17. Robinson MA, Park S, Sun ZJ, Silver PA, Wagner G, Hogle JM. Multiple conformations in the ligand-binding site of yeast nuclear pore-targeting domain of Nup116p. J Biol Chem. 2005; 280:35723–35732. [PubMed: 16105837]
18. Yoshida K, Seo H-S, Debler EW, Blobel G, Hoelz A. Structural and functional analysis of an essential nucleoporin heterotrimer on the cytoplasmic face of the nuclear pore complex. Proc Natl Aca Sci USA. 2011; 108:16571–16576.
19. Sampathkumar P, Ozyurt SA, Do J, Bain KT, Dickey M, Rodgers LA, Gheyi T, Sali A, Kim SJ, Phillips J, Pieper U, Fernandez-Martinez J, Franke JD, Martel A, Tsuruta H, Atwell S, Thompson DA, Emtage JS, Wasserman SR, Rout MP, Sauder MJ, Burley SK. Structure of the autoproteolytic domain from the *Saccharomyces Cerevisiae* nuclear pore complex component, Nup145. Proteins: Structure Function and Bioinformatics. 2010; 78:1992–1998.
20. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. Atomic structures of human immunophilin FKBP-12 complexes with FK506 and Rapamycin. J Mol Biol. 1993; 229:105–124. [PubMed: 7678431]
21. Leslie AGW, Brick P, Wonacott AJ. An improved program package for the measurement of oscillation photographs. CCP4 News. 1986; 18:33–39.
22. Collaborative Computing Project Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994; 50:760–763. [PubMed: 15299374]
23. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. PHASER crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

24. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nature Struct Biol. 1999; 6:458–463. [PubMed: 10331874]
25. Emsley P, Cowtan K. COOT: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]
26. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the Maximum-Likelihood Method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
27. DeLano, WL. The PyMOL Molecular Graphics System. 2002. http://pymol.sourceforge.net
28. McPhillips TM, McPhillips SE, Chiu HJ, Cohen AE, Deacon AM, Ellis PJ, Garman E, Gonzalez A, Sauter NK, Phizackerly RP. Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. J Synchrotron Radiat. 2002; 9:401–406. [PubMed: 12409628]
29. Forster F, Webb B, Krukenberg KA, Tsuruta H, Agard DA, Sali A. Integration of small-angle X-ray scattering data into structural modeling of proteins and their assemblies. J Mol Biol. 2008; 382:1089–1106. [PubMed: 18694757]
30. Schneidman-Duhovny D, Hammel M, Sali A. FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res. 2010 Jul 1; 38(Web Server issue):W540–W544. [PubMed: 20507903]
31. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993; 234:779–815. [PubMed: 8254673]
32. Russel D, Lasker K, Webb B, Velazquez-Muriel J, Tjioe E, Schneidman-Duhovny D, Peterson B, Sali A. Putting the pieces together: integrative structure determination of macromolecular assemblies. PLoS Biol. 2012 in press.
33. Franke D, Svergun DI. DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. J Appl Cryst. 2009; 42:342–346.
34. Svergun DI, Petoukhov MV, Koch MHJ. Determination of domain structure of proteins from X-ray solution scattering. Biophys J. 2001; 80:2946–2953. [PubMed: 11371467]
35. Volkov VV, Svergun DI. Uniqueness of *ab initio* shape determination in small-angle scattering. J Appl Cryst. 2003; 36:860–864.
36. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr. 2004; 60:2256–2268. [PubMed: 15572779]
37. Bahadur R, Chakrabarti P, Rodier F, Janin J. A dissection of specific and non-specific protein-protein interfaces. J Mol Biol. 2004; 336:943–955. [PubMed: 15095871]
38. Zhu H, Domingues FS, Sommer I, Lengauer T. NOXclass: prediction of protein-protein interaction types. BMC Bioinformatics. 2006; 7:27. http://noxclass.bioinf.mpi-sb.mpg.de/index.php. [PubMed: 16423290]
39. Konarev PV, Kikhney AG, Franke D, Petoukhov MV, Svergun DI. Automated SAXS Data Processing on the X33 Beamline. HASY LAB Annual Report. 2007; Part II:339–340.
40. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Cryst. 2003; 36:1277–1282.
41. Shatsky M, Nussinov R, Wolfson HJ. A method for simultaneous alignment of multiple protein structures. PROTIENS: Struc Func Bioinformatics. 2004; 56:143–156.
42. Ramakrishnan C, Ramachandran GN. Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. Biophys J. 1965; 5:909–933. [PubMed: 5884016]
43. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB III, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucl Acids Res. 2007; 35:W375–W383. http://molprobity.biochem.duke.edu/. [PubMed: 17452350]

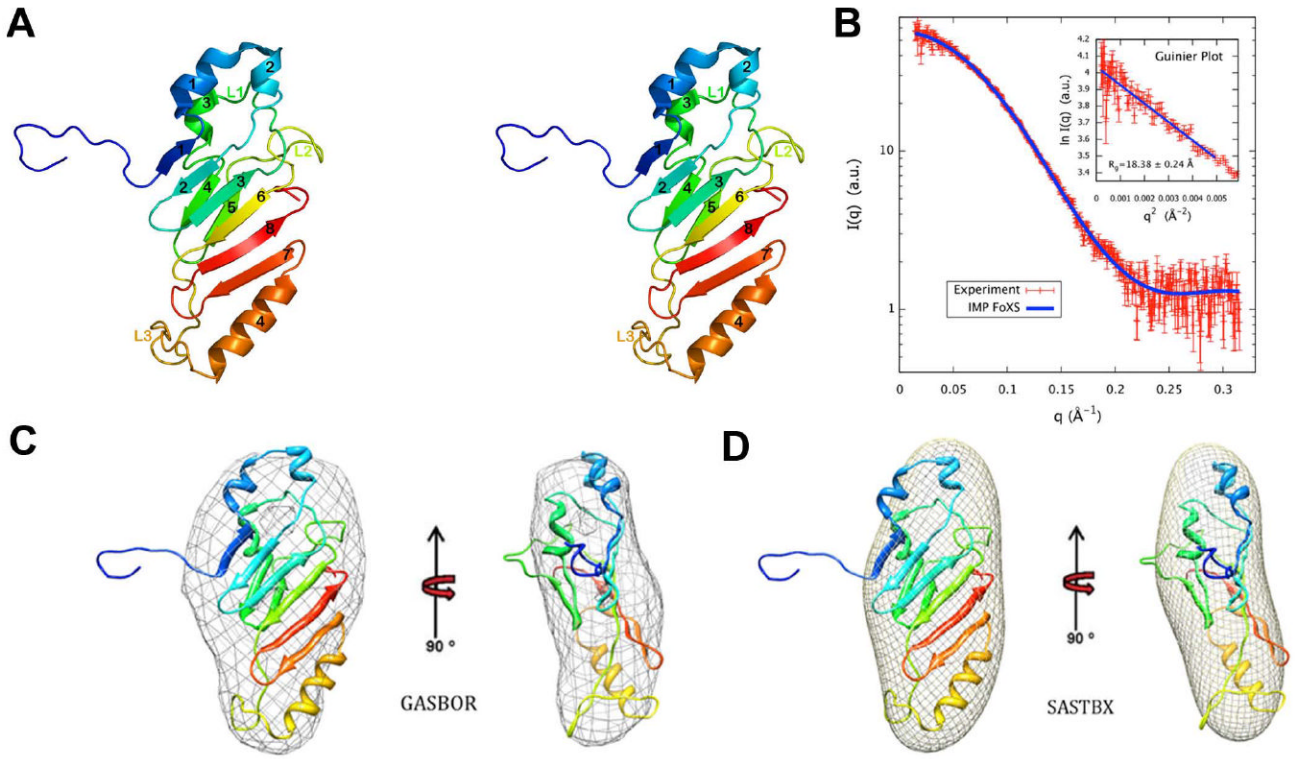

Figure 1.

A: Stereoview of the CgNup116(882-1034) monomer. Cartoon of Chain A is shown as a rainbow from blue to red from N- to C-terminus.

B: Comparison of the merged experimental SAXS profile (red) of CgNup116(882-1034) with the SAXS profile computed by IMP FoXS (blue) from the complete monomer model of CgNup116(882-1034). Inset shows the SAXS profiles in the Guinier plot, with a R_g fit of 18.38±0.24 Å. D_{max} of radial distribution function, P(r), is 60.65 Å.

C and D: Comparison of the shapes of CgNup116(882-1034) (represented as a mesh) calculated from the experimental SAXS profile by GASBOR (C) and SASTBX (D).

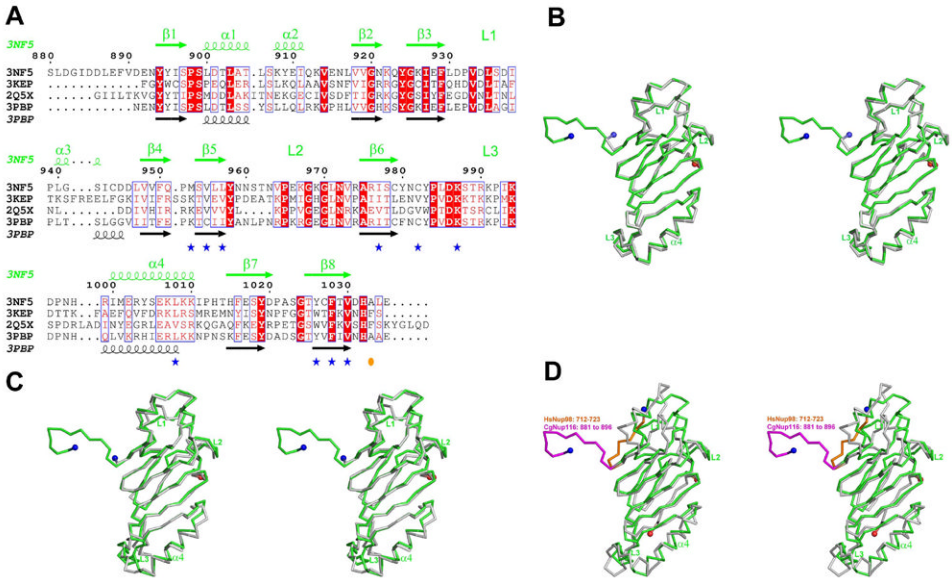

Figure 2.

A: Structure based sequence alignment of the structures of CgNup116(882-1034, PDB Code 3NF5), ScNup116 bound to ScNup82:ScNup159 complex (PDB Code 3PBP), autoproteolytic domain of ScNup145 (PDB Code 3KEP), and autoproteolytic domain of HsNup98 (PDB Code 2Q5X). Secondary structural elements of CgNup116(882-1034) and ScNup116 bound to ScNup82:ScNup159 complex are displayed in green and black, respectively. Residues of ScNup116 contributing to its binary interaction with ScNup82 β-propeller domain are marked with blue star. Sites of autoproteolysis of ScNup145N and HsNup98 are indicated with an orange circle. The Ser residue at the autoproteolytic site of ScNup145N and HsNup98 are replaced by Glu and Ala in proteins used for structure determination, respectively.

B: Stereoview of CgNup116(882-1034; green) superposed on the structure of ScNup116 bound to ScNup82:ScNup159 complex (grey). Helix α4 of CgNup116 is structurally equivalent to the helix αB of ScNup116.[18]

C: Stereoview of the CgNup116(882-1034; green) superposed on the autoproteolytic domain of ScNup145 (grey).

D: Stereoview of the CgNup116(882-1034; green) superposed on the autoproteolytic domain of HsNup98 (grey). Residues preceding the β1-strands are highlighted in magenta and orange for CgNup116(882-1034) and HsNup98, respectively. The N- and C-terminal residues are shown on the cartoons as blue and red spheres, respectively.

Table 1

Crystallographic Statistics.

Data Collection	**CgNup116(883-1034)**
PDB code:	3NF5
Space group:	$P2_1$
Unit-cell dimensions (Å):	a=48.7, b=67.7, c=55.1, β=101.4°
Matthew's coefficient (Å^3/Da):	2.38
Solvent content (%):	48
Resolution (Å):	34.37-1.94 (2.04-1.94) *
Number of unique reflections:	25957 (3756)
Completeness (%):	99.1 (98.8)
R_{symm} (%):	14.0 (47.1)
Multiplicity:	7.0 (6.9)
< I/σ(I) >:	8.3 (3.6)
Refinement	
R-factor (%):	20.7
R_{free} (%):	25.6
Root mean square deviations from ideal values	
Bond length (Å):	0.020
Bond angles (°) :	1.782
Ramachandran Plot[41]	
MolProbity[42] residues in	
Favored region (%):	97.0
Allowed region (%):	100.0

* Values in parenthesis correspond to the highest-resolution shell