Title: **Structural mechanism of nuclear transport mediated by importin β and flexible amphiphilic proteins**

Authors: Shige H. Yoshimura*, Masahiro Kumeta, and Kunio Takeyasu

Affiliations: Graduate School of Biostudies, Kyoto University

Yoshida-konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan

Contact Information (corresponding author):

*Shige H. Yoshimura, Ph.D.

Graduate School of Biostudies, Kyoto University

Yoshida-konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan

Tel&fax: +81-75-753-7906,

E-mail: yoshimura@lif.kyoto-u.ac.jp

Additional Footnotes: none

Running Title: Nuclear transport mediated by amphiphilic proteins

Summary

Karyopherin β family proteins mediate the nuclear/cytoplasmic transport of various proteins through the nuclear pore complex (NPC), although they are substantially larger than the size limit of the NPC. To elucidate the molecular mechanism underlying this paradoxical function, we focused on their unique structures called HEAT repeats, which consist of repetitive amphiphilic α-helices. An *In vitro* transport assay and FRAP analyses demonstrated that not only karyopherin β family proteins but also other proteins with HEAT repeats could pass through the NPC by themselves, and serve as transport mediators for their binding partners. Biochemical and spectroscopic analyses and molecular dynamics simulations of purified HEAT-rich proteins revealed that they interact with hydrophobic groups, including phenyl and alkyl groups, and undergo reversible conformational changes in tertiary structures, but not in secondary structures. These results show that conformational changes in the flexible amphiphilic motifs play a critical role in translocation through the NPC.

Introduction

Macromolecular trafficking between the cytoplasm and nucleoplasm is crucial for establishing proper intracellular protein distributions. The nuclear pore complex (NPC), which penetrates the double membrane of the nuclear envelope, controls the flow of inbound and outbound traffic. The relationships between the structure and function of the NPC have been clarified by recent proteomic and bioinformatics analyses (Cronshaw et al., 2002, Ori et al., 2013, Rout et al., 2000, Tamura et al., 2010). The NPC is composed of more than 30 different subunits called nucleoporins (Nups). The central channel is especially rich in Nups bearing Phe-Gly (FG) motifs (FG-Nups), and forms a hydrophobic environment. These Nups form a hydrogel structure *in vitro* via hydrophobic interactions between phenylalanine residues (Frey et al., 2006, Mohr et al., 2009).

Because of the macromolecular crowding, the NPC has been regarded as a molecular sieve for macromolecules in the cytoplasm and nucleoplasm. Molecules smaller than 40 kDa can pass through the NPC by passive diffusion, whereas proteins larger than 40 kDa cannot pass through by themselves (for review, (Gorlich et al., 1999)). Numerous larger proteins are actively transported by so-called transport receptors, including karyopherin β family proteins. These proteins can pass through the NPC individually and can also mediate the passage of other proteins, including those that are larger than 40 kDa (for review, (Chook et al., 2001)).

Important questions regarding NPC transport that remain unanswered are how transport receptors mediate the translocation of the cargo in spite of their large molecular sizes, and how nuclear proteins that do not interact with any known transport mediators are transported across the

NPC (Lange et al., 2007). A number of previous studies have demonstrated the importance of the hydrophobic environment in the NPC for receptor-mediated nuclear transport. Karyopherin β proteins strongly bind to hydrophobic phenyl sepharose columns (Ribbeck et al., 2002). The crystal structure of an importin β-FG motif complex showed that importin β has several hydrophobic pockets in its convex surface that interact with the hydrophobic residues in FG-Nups (Bayliss et al., 2000, Bayliss et al., 2002, Liu et al., 2005, Otsuka et al., 2008). Other studies have demonstrated that a protein's hydrophobicity is closely related to its ability to pass through the NPC (Naim et al., 2009, Ribbeck et al., 2002). These lines of evidence imply that the size of a protein and its hydrophobicity are critical determinants of permeability. However, the strong surface hydrophobicity and its strong interaction with Nups cannot fully explain its rapid translocation through the NPC, because hydrophobic proteins tend to become trapped in the NPC and cannot easily escape into the nucleoplasm or the cytoplasm.

The crystal structures of karyopherin β family proteins exhibit significant similarity in their overall molecular shape (Figure S1), though the primary sequence similarity is very low (15–20% sequence identity) (O'Reilly et al., 2011, Xu et al., 2010). They are composed of a number of HEAT motifs (19–21 repeats), each comprising two amphiphilic α-helices (A-helix and B-helix) connected by a short linker region (Figures 1A and 1B) (Chook et al., 1999, Cingolani et al., 1999). A helical wheel representation of the α-helices in HEAT repeats shows that the α-helices in importin β are amphiphilic, with the hydrophobic sides facing each other toward the inside of the molecule and the hydrophilic sides facing the outside (solvent) (Figures 1C and 1D). Due to such repetitive helices,

karyopherin β has large structural flexibility, which has been demonstrated to play important roles in binding to cargo proteins and RanGTP (Chook et al., 2001, Conti et al., 2006, Forwood et al., 2010, Lee et al., 2000, Stewart, 2007).

In this report, we focused on the structural flexibility of karyopherin β and other HEAT motif-rich proteins, and examined their involvements in passage through the NPC. The flexible structure with amphiphilic α-helices seems suitable for passing through the hydrophobic crowding environment of the NPC, and travel between the cytoplasm and the nucleoplasm. Therefore, conformational changes that occur in such a flexible amphiphilic structure play a critical role not only in interaction with cargo but also in fast passage through the NPC. Our extensive functional and structural analyses revealed that importin β and other HEAT motif-rich proteins undergo conformational changes induced by various hydrophobic groups.

Results

The HEAT motif is a suitable structure for passage through the NPC

Non-karyopherin β family member proteins which contain multiple HEAT motifs are listed in Figure 2A. The number of HEAT motifs and their positions within the polypeptide vary from protein to protein. Since the primary sequence similarity of HEAT motif is very low, we also examined their secondary structures (amphiphilic α-helices, Figure 1C) and found additional possible HEAT motifs within the molecule (see Supplemental Experimental Procedures for detail). The results are summarized in Figure 2A. We expressed these HEAT-rich proteins in bacterial cells either as full-length proteins or as partial fragments containing HEAT-rich regions, purified them by affinity chromatography, and subjected them to circular dichroism (CD) spectra analysis. All the HEAT-rich fragments showed typical α-helix-rich spectra (Figure S2A) and high α-helical content comparable to that of importin β (Figure S2B).

We then tested whether these HEAT-rich fragments could travel through the NPC. EGFP-fused HEAT-rich proteins, and non-HEAT proteins as controls, were expressed in bacteria, affinity purified, and incubated with digitonin-treated HeLa cells. As shown in Figure 2B, HEAT-rich proteins were able to enter the nucleus through the NPC without the assistance of other proteins. This influx is not due to the EGFP tag, as we observed similar results with HA-tagged proteins that were immunostained with an anti-HA antibody (Figure S2C). It should be noted that all of these HEAT-rich proteins are far larger than the size limit of the NPC (~40 kDa). The flux rate constant was plotted against the size of the protein (Figure 2C) (The Stokes radius was obtained by gel

filtration chromatography as described in Figure S2D). The flux rates of non-HEAT proteins were drastically reduced above a Stokes radius of 4 nm. In contrast, all of the HEAT-rich proteins showed high flux rates, indicating that they can overcome the size barrier of the NPC regardless of molecular size. The influx of HEAT-rich proteins was almost completely blocked by wheat germ agglutinin (Figure 2B, +WGA), which indicated that they indeed traveled through the NPC. The pull-down assay also demonstrated direct interactions between HEAT-rich proteins and nucleoporins (Figure 2D), confirming that HEAT-rich proteins interact with Nups when they go through the pore.

The intracellular distribution and dynamics of the HEAT-rich proteins were also examined by expressing EGFP-fusion proteins in HeLa cells. The steady-state distributions varied from protein to protein; some proteins strongly accumulated in the nucleus, whereas others were less in the nucleus and mainly located in the cytoplasm (Figure 3, and summarized in Figure 2A). The proteins that showed strong nuclear signals were then subjected to FRAP analysis to examine their steady-state shuttling across the nuclear envelope. The analysis of fluorescence recovery demonstrated that these proteins shuttle between the nucleoplasm and the cytoplasm (Figure 3). The flux rates obtained *in vivo* (FRAP analysis, Figure 3) were in general smaller than those *in vitro* (transport assay, Figure 2B), suggesting that these HEAT-rich proteins may form a complex *in vivo* and may regulate shuttling of the protein complex through the NPC.

HEAT-rich proteins mediate the translocation of other proteins through the NPC

We then examined whether HEAT-rich proteins assist the nuclear transport of other proteins and act as potential transport mediators. CAND1 (~120 kDa) contains 27 HEAT repeats and is known to form a complex with cullin1 and cullin4B (Fischer et al., 2011, Goldenberg et al., 2004) (Figures 4A and 4B). Purified cullin4B was excluded from the nucleus of permeabilized HeLa cells, but was able to enter the nucleus when it formed a complex with CAND1 (Figure 4C), demonstrating that CAND1 stimulates the influx of cullin4B despite the increase in total molecular mass. We confirmed this function *in vivo* by measuring the mobility of EGFP-fused cullin across the nuclear envelope in CAND1-knockdown cells (Figure 4D). As demonstrated by the FRAP analysis, fluorescence recovery in the nucleus was drastically reduced by the depletion of CAND1 (Figure 4E). In particular, the recovery of cullin4B was significantly reduced in CAND1-knockdown cells. The shuttling of cullin1 in control cells was slower than that of cullin4B, but was also significantly reduced by knockdown of CAND1.

The same effect was observed for the protein phosphatase 2A (PP2A) complex, in which the HEAT-rich subunit (A subunit), the catalytic subunit (C subunit), and the regulatory subunit (B subunit) form a trimeric complex (Figure 5A) (Cho et al., 2007). HeLa cells express two isoforms of the A subunit (PPP2R1A (Aα) and PPP2R1B (Aβ)). When the EGFP-fused C subunit was expressed in HeLa cells, it localized to both the cytoplasm and nucleoplasm. Bleaching of the nuclear fluorescent signal resulted in fast recovery of the signal (Figures 5B and 5C). Knockdown of the HEAT subunit (Figure 5D) reduced the flux rate of the EGFP-fused C subunit across the nuclear envelope (PPP2R1A KD and PPP2R1B KD, Figures 5B and 5C). These results

demonstrated that some HEAT-rich proteins can not only pass through the pore, but also could assist the passage of other proteins. Since a previous report demonstrated direct interaction between PPP2R1A and importin 9 (Lubert et al., 2003), we also tested the involvement of importin 9 in the nuclear shuttling of PP2A. Knockdown of importin 9 only slightly decreased the mobility of the PP2A C subunit, while double-knockdown of importin 9 and PPP2R1A resulted in a further decrease of the recovery (Figure S3). This result indicates that the PP2A A subunit is a major transport mediator of the PP2A complex and importin 9 may have additional regulatory roles in the intracellular distribution of the enzyme (see Discussion).

HEAT-rich proteins bind to hydrophobic groups and undergo conformational changes

The central channel of the NPC is protein-rich environment with high content of hydrophobic residues; the pore-forming Nups contain a number of phenylalanine and leucine, as well as high numbers of non-bulky side chains such as serine, threonine, and glycine (Figure S4A). In addition, significant portions of these Nups are assigned as unstructured (Figure S4B) and do not show any significant secondary structures based on the CD spectra (Figure S4C) (Denning et al., 2003). The interaction between HEAT-rich proteins and various hydrophobic groups were investigated by hydrophobic interaction chromatography. As shown in Figure 6A, purified HEAT-rich proteins firmly bind not only to phenyl sepharose, as was demonstrated for importin β in previous studies (Ribbeck et al., 2002), but also to other alkyl groups such as butyl (C4) and octyl (C8). This result well matches to the pull-down assay between HEAT-rich proteins and FG-rich domain of

nucleoporin (Figure 2D). Increasing the salt concentration strengthened the binding (Figure 6A, NaCl+), demonstrating that the interaction is mainly governed by hydrophobic interactions.

Since HEAT motifs are composed of amphiphilic helices (Figure 1), we tested whether hydrophobic groups can induce conformational changes of HEAT motifs. The tertiary structure of purified HEAT-rich proteins in various organic solvents was investigated by examining the fluorescence spectra of internal tryptophan residues. As summarized in Figure 6B, the center of the fluorescence peak was more or less red-shifted as the alcohol concentration increased, implying that the residue is exposed to the solvent. The size of the shift largely depended on the alkyl species; as the alkyl chain became longer, the effect became stronger (Figure 6B), indicating that the conformational change was induced by an alkyl group, and not by a hydroxyl group. The details of conformational flexibility was further investigated by mutating tryptophan residues in importin β. As summarized in Figure S4D, structural changes were more prominent in the middle regions (HEAT #7-11) than those in both termini (HEAT #1-4 and #19). It should also be noted that these structural changes are reversible, since HEAT-rich proteins, which was once exposed to 50 % alcohol and then placed back to alcohol-free solution exhibited the tertiary structure similar to the non-treated protein (Figure S4E).

In contrast to the tertiary structures, the secondary structures were not affected by the hydrophobic groups. The CD spectra of HEAT-rich proteins showed a typical α-helix-rich shape in the absence and presence of hydrophobic groups (Figure S4F). The molar ellipticity at 222 nm, which corresponds to an α-helix-specific negative peak, was summarized in Figure 6C. The

α-helical contents of HEAT-rich proteins were not reduced in the presence of high concentrations of hydrophobic groups and rather slightly increased due to the stabilizing effect of alcohol.

Structural flexibility is required for translocation through the NPC

The importance of structural flexibility was examined in an *in vitro* transport assay. Purified HEAT-rich proteins were fixed by a crosslinker (BS^3) to restrict conformational flexibility. The fixed proteins showed almost the same elution profile as non-treated proteins in gel filtration chromatography, but less structural dynamics upon exposure to the hydrophobic groups (Figures S5A and S5B). These prefixed HEAT-rich proteins showed a slower influx rate in the *in vitro* transport assay (Figure 7, crosslinked), indicating that structural flexibility plays a critical role in traveling through the NPC. This effect was not due to the modification of side chains by the crosslinker, since the same modification without crosslinking did not significantly affect the flux rate (Figure 7, non-crosslinked). These results indicate that structural flexibility plays a critical role in translocation through the NPC.

Molecular dynamics simulation of conformational changes

The molecular dynamics (MD) simulation of the structural changes occurring in importin β correlated well with the experimental results (Figure 6) and provided more detailed structural information. The MD simulation was initiated from the previously reported crystal structure for mouse importin β (Protein Data Bank ID: 1UKL) (Figure S1), and was executed in two different

solvents (water or 50% (v/v) TFE/water mixture) until the structure became stable (Figure S6A). As shown in Figures 8A and 8B, and as quantified in Figures 8C and 8D, most of the α-helices in the HEAT motifs remained unchanged, although occasional and partial elongations were detected in 50% TFE (Figure 8D). The structural fluctuation is even smaller in 50% TFE than in water (the root mean square fluctuation was 0.714 in 50% TFE and 0.948 in water). This result is in accordance with the results from the CD spectra (Figure 6), and it indicates that most of the secondary structure is retained (in fact, it is even more stabilized) in the water/alcohol mixture.

In contrast to the secondary structure, the interactions among adjacent helices are significantly affected by the environment. As shown in Figures 8A and 8B, and quantified in Figures 8E and8F, the distances between adjacent α-helices were increased. Changes in the carboxyl terminal half were particularly prominent (Figure 8F). The collapse in tertiary structure was also observed in other trajectories in TFE, although the position and timing varied from trajectory to trajectory (Figure S5B). The dissociation of adjacent helices is as expected, since hydrophobic interaction generally becomes weaker in a hydrophobic environment. This separation of α-helices resulted in the exposure of hydrophobic side chains to the molecular surface (Figure 8G–I). The solvent-excluded surface area (SESA) assigned to the hydrophobic residues was 1.7 times larger in 50% TFE than in water (Figure 8I). In 50% TFE, the first hydration shell was severely collapsed and the TFE molecules were in close contact with hydrophobic residues at the protein surface (Figure S5F), indicating that hydration energy is closely related to the conformational change of importin β. TFE attacks hydrophobic surface regions including the previously identified binding

pockets for the FG motif (Bayliss et al., 2000, Bayliss et al., 2002, Bednenko et al., 2003) and induces separation of the HEAT helices.

The reversibility of the solvent-induced conformational changes could also be seen in the MD simulation. Transferring importin β from 50% TFE to water resulted in an immediate decrease in the hydrophobic surface area; the hydrophobic SESA of importin β increased when exposed to TFE, but rapidly returned to its initial level when transferred to water (at 100 or 200 ns; Figure 8J). These results are consistent with the experimental observation that the solvent-induced conformational changes of importin β are reversible (Figure S4E).

Importin β-cargo complex also undergoes conformational change

Structural changes in importin β loaded with a cargo (importin β-binding (IBB) domain of importin α) were also characterized (Figure 9). A pull-down assay demonstrated that the IBB domain did not dissociate from importin β even in the presence of 50% TFE (Figure 9A), which is reasonable since the interaction between importin β and IBB is mainly governed by electrostatic interactions (Cingolani et al., 1999). The effect of TFE on the CD spectrum of the entire complex was similar to that of free importin β: low concentrations (10–30%) of TFE did not affect the α-helix content, whereas higher concentrations slightly increased the α-helix content (Figure 9B, red). The fluorescence spectrum suggested that the tertiary structure of the complex was significantly affected by low concentrations of TFE (Figure 9B, blue), although the effect was slightly smaller than that observed with free importin β (compare with Figure 6B), suggesting that the IBB domain helps

stabilize the tertiary structure of importin β (Cingolani et al., 2000).

MD simulation of the importin β-IBB complex was performed using the previously reported crystal structure (PDB ID: 1QGK; Figure S1) as the starting structure. Interactions between importin β and the IBB domain in 50% TFE were retained throughout our simulation time (400 ns; Figures 9C and 9D; see also Figure S5D), which is in agreement with the pull-down experimental results (Figure 9A). As observed with free importin β, the HEAT-HEAT distances (Figures 9E and 9F) and hydrophobic SESA (Figures 8I, 9G and 9H) significantly increased in the presence of TFE, whereas the length of each α-helix remained unaffected in both cases (Figure S6E). The pattern of the increase in distance (Figure 9F) slightly differed from that of free importin β (Figure 8F), in part due to the interaction of importin β with IBB via its carboxyl terminal HEAT domains (Cingolani et al., 1999). These results demonstrate that the importin β-IBB complex undergoes conformational changes similar to those of free importin β, although the changes in the entire conformation are slightly smaller due to the interaction with the IBB domain.

Discussion

It has long been known that the NPC restricts the flow of molecules based mainly on their size and shape. Smaller molecules can pass through the NPC, whereas larger molecules cannot. However, such larger proteins can be translocated through the NPC with the aid of transport receptors. The unresolved questions about NPC and transport mediators are i) how importin β family proteins pass through the NPC despite their large molecular size, and ii) whether more than 1,000 nuclear

proteins are imported by only dozens of known transport receptors. In this study, we focused on the flexible amphiphilic structure of importin β and other HEAT motif-rich proteins, and clearly demonstrated that not only importin β but also other HEAT motif-rich proteins can pass through the NPC by themselves and, in some cases, together with cargo. Our findings provide an answer to the size-barrier paradox of receptor-mediated nuclear transport, propose one general structural property of "transport receptors", and suggest the existence of numerous cargo-specific transport receptors and "transport-regulating subunits".

Amphiphilic motifs are suitable for passage through the NPC

Our results shown in Figures 2 and 3 demonstrated that not only importin β but also many other proteins that contain a number of HEAT motifs can pass through the NPC. All of the proteins we investigated are larger than 40 kDa and, therefore, exceed the "size barrier" of the NPC. There have been several previous reports that demonstrated importin β-independent nuclear translocation of over-sized proteins. These include albumin (66 kDa) with hydrophobic modifications on the surface (Naim et al., 2009, Ribbeck et al., 2002), amphiphilic triple-helix of spectrin repeat (Kumeta et al., 2012), and importin α and β-catenin, both of which contain amphiphilic helical repeats similar to HEAT repeats (ARM repeats) (Fagotto et al., 1998, Koike et al., 2004, Wiechens et al., 2001). These lines of evidence strongly suggest that hydrophobic/hydrophilic property of the protein is closely related to its passage through the NPC. The protein structure database contains a number of proteins with repetitive amphiphilic α-helical motifs in addition of the HEAT motif (D'Andrea et al., 2003,

Groves et al., 1999, Marsella et al., 2009, Neuwald et al., 2000). Although the detailed arrangements of helices (and sheets) vary from motif to motif, they are structurally flexible and involved in interaction with other proteins (Bella et al., 2008, Neuwald et al., 2000, Shi, 2009). Therefore, it might be the case that not only HEAT-rich proteins, but also proteins with repetitive amphiphilic motifs might translocate through the NPC.

We further demonstrated that non-karyopherin HEAT-rich proteins, CAND1 and PPP2R1, function as specific nuclear transport receptors for their binding partners (Figures 4 and 5). These proteins have been identified as a component of multi-subunit protein complexes, but not as a transport receptor. PPP2R1 is one of the three subunits of protein phosphatase 2A (PP2A) complex and has been regarded as a scaffolding subunit (A subunit) which directly interacts with both of the other two subunits (B and C subunits) (Figure 5A). PP2A is highly conserved from yeast to human and is known to dephophorylate a number of cellular proteins (Janssens et al., 2008, Shi, 2009). The regulatory mechanism of PP2A is extraordinarily complex because of the existence of multiple isoforms in each subunit; 2 isoforms (Aα and Aβ) for A subunit, 2 isoforms (Cα and Cβ) for C subunit, and at least 19 isoforms for B subunit (Janssens et al., 2008). Especially, the B subunit affects the substrate specificity and the catalytic activity of the holoenzyme. On the other hand, the function of the HEAT-rich A subunit has not been well understood, except that it functions as a flexible scaffold. Our result presented here demonstrated that the A subunit plays a role in the translocation of the holocomplex through the NPC. Considering the fact that the B subunit also has

a HEAT-like structure (Figure 5A, blue) (Cho et al., 2007), both A and B subunits may regulate the intracellular localization of the enzyme.

Double-knockdown of importin 9 and the A subunit resulted in a further decrease of the flux rate of the C subunit (Figure S3). This implies that both importin 9 and the HEAT-rich subunit are involved in nuclear shuttling. It might be the case that PPP2R1A confers a constitutive permeability to the holoenzyme, while importin 9 plays more regulatory roles in the intracellular distribution of the complex. It is intriguing that PP2A directly interacts with HDAC4 and dephosphorylates it, which results in the nuclear import of HDAC4 (Paroni et al., 2008). Since PP2A is shuttling between cytoplasm and nucleoplasm, it might be the case that PP2A regulates the intracellular localization of various substrates by converting phosphorylated and dephosphorylated states. Our result on PPP4R1 (Figures 2 and 3), which is a regulatory subunit of protein phosphatase 4 (PP4), suggests that the intracellular distribution of PP4 is also regulated by this HEAT-rich subunit. Therefore, one of the important functions of HEAT-rich and other amphiphilic proteins in a protein complex might be regulating the intracellular distribution of the complex in a karyopherin β-independent manner.

Conformational changes in HEAT-rich proteins in hydrophobic environment

Our structural analysis on importin β and other HEAT-rich proteins demonstrated that reversible conformational change is critical in traveling through hydrophobic crowding of the NPC (Figures 6 and 7). Although a direct interaction between Nups and transport receptors has been reported in

previous studies, a simple binding event cannot explain the passage mechanism because strong binding to Nups results in a longer residence time in the NPC and poor escape to the opposite side. Repetitive amphiphilic structures and their flexible and reversible conformational changes could clearly explain the mechanism of such fast passage. The structural flexibility of karyopherins has been reported and discussed in previous studies, mainly based on the crystal structures of karyopherins bound with cargo, RanGTP or FG-Nups (Chook et al., 2001, Conti et al., 2006, Forwood et al., 2010, Lee et al., 2000, Stewart, 2007). The interaction with RanGTP and some of the cargos mainly occurs at the concave surface (B-helices) of importin, whereas interaction with FG-motifs (FG-Nups) and some of the cargo occur at the convex surface. Our results (Figure 6B, S4D, 8 and 9) suggest that Nup-induced conformational changes of importin β occur over a wide area of the molecule, especially in the middle HEAT repeats (#5-17). Repetitive alignment of A- and B-helices on the opposite sides of the molecule (A-helix on the convex and B-helix on the concave) and their flexible conformation enables HEAT proteins to interact both with the cargo and nucleoporins without affecting each other, which plays a critical role in efficient passage through the NPC as a protein complex.

Our result shown in Figure 6 indicates that conformational changes in HEAT motifs are induced by interaction with various kinds of hydrophobic groups (ethyl, butyl, propyl, octyl, and phenyl). This type of conformational change is in clear contrast to ligand-dependent conformational changes, which are mediated by specific interactions between the ligand and specific residues in the target enzyme. In the case of importin β, our MD simulation revealed a number of weak interactions

between hydrophobic groups in the solvent and many hydrophobic side chains (Figure S6F). Such weak interaction between importin β and hydrophobic groups of Nups plays an important role in the fast kinetics of the translocation process, which had been observed in previous reports using bulk import assay and single-molecule observation (several tens of milliseconds) (Sun et al., 2008, Tu et al., 2011). In good agreement with this is the transport ability of importin β which lacks FG-binding pocket(s). Importin β contains multiple binding pockets for FG-motifs (Bayliss et al., 2000, Bednenko et al., 2003). However, mutating these critical amino acids could not completely abolish the transport ability (Figure S7), suggesting that the interaction between importin β and nucleoporins is not of a site-specific manner. Dissociation rate constant (k_{off}) between importin β and FG-Nups obtained in *in vitro* binding assay (10^{-3}–10^{-4}·sec^{-1}) does not explain such fast kinetics, since hydrophobic interaction *in vitro* condition is extremely stronger than that in the hydrophobic environment of the NPC.

Molecular Events within the NPC

Although HEAT-rich proteins can mediate translocation of the cargo, karyopherin β family proteins are distinct from others because they can catch or release cargo depending on the RanGTP/GDP cycle. Importin β releases the cargo when bound to RanGTP, which is rich in the nucleoplasm (Figure 10A). Due to this mechanism, importin and exportin can produce a cargo gradient across

the nuclear envelope (Gorlich et al., 2003), even though their passage through the NPC itself is bi-directional and energy-independent. In contrast to karyopherin β proteins, other HEAT-rich proteins do not actively produce a cargo gradient across the nuclear envelope, because their passage through the NPC is mainly diffusion-based (Figure 2B), and their interaction with the cargo is Ran-independent (Figure 4B). In good agreement with this is the previous reports that ARM repeat proteins can mediate the nuclear import of other proteins (CaMKIV by importin α (Kotera et al., 2005) and lef-1 by β-catenin (Asally et al., 2005)) but do not produce a strong accumulation of the cargo in the nucleus (Yokoya et al., 1999). Although non-karyopherin β HEAT-rich proteins cannot produce the cargo gradient in Ran-dependent manner, they might have other regulatory mechanisms on cargo-release.

The possible molecular events that occur during the passage of an amphiphilic protein through the NPC are depicted in Figure 10B. When an amphiphilic protein or protein complex enters the NPC from the cytoplasmic side, it may initially bind to part of the NPC with some structural and positional fluctuations. Nup358 (RanBP2) may help capture and anchor karyopherin β to the cytoplasmic side of the NPC. The protein eventually moves into the central channel by repeated interaction with hydrophobic groups, or in some cases, it escapes back into the cytoplasm. Once inside the central pore of the NPC, the protein travels in the pore by diffusion. As the interactions between the hydrophobic surface of the protein and FG-Nups are very weak in a hydrophobic environment, the diffusion rate becomes much higher than that obtained *in vitro*. At a

certain point during diffusion, the protein reaches the nucleoplasmic border of the NPC, and occasionally dissociates from the NPC.

The detailed structural information of the central channel of the NPC will lead further understanding of the molecular event of protein passage. Long-term MD simulation (as described by (Isgro et al., 2005)) including such structural information will also provide useful information on the dynamics of molecular crowding (water, polar groups, and hydrophobic groups) within the NPC, and on the contribution of amphiphilic motifs in migrating through such a crowding environment.

Experimental Procedures

DNA construction and protein purification

GST-tagged mouse importin β was expressed in bacterial cells and purified as described in a previous study (Bayliss et al., 2000, Bayliss et al., 2002, Liu et al., 2005, Otsuka et al., 2008). The cDNAs encoding human HEAT-rich proteins and non-HEAT proteins were either cloned by PCR from a cDNA pool of HeLa cells (PPP2R1A, PPP2R1B, PPP4R1, CAPG), or purchased from the Kazusa DNA library. cDNA fragments of HEAT-rich regions and the fragments of cullin were amplified by PCR and subcloned into the pGEX vector (GE Healthcare). The proteins were expressed as GST-fusion proteins and were subjected to specific protease digestion (PreScision; GE Healthcare) or thrombin (Nacalai Tesque) if necessary. The cDNA fragment encoding the importin β-binding domain of rat importin α (IBB; a.a. 1–69) was amplified by PCR from full-length cDNA and cloned into a pET29 vector (Novagen). The protein was expressed in *E. coli* and purified by ion-exchange chromatography (Hi-Trap SP; GE Healthcare) followed by gel filtration chromatography (Superdex 75; GE Healthcare).

Protein fixation

Purified importin β was diluted into different solvents (0 or 50% TFE with 50 mM KPO_4) to a final concentration of 2 mg/mL. A bifunctional crosslinker (BS^3; Thermo Science) or sulfo-NHS (Thermo Science) was added at a final concentration of 1 mM, and the mixture was incubated at 25 °C for 30 min. The reaction mixture was then subjected to a gel filtration column to replace the buffer with 100 mM KPO_4.

CD and fluorescence spectra

The CD spectra of purified karyopherins and other HEAT proteins were measured using a J-805 (JASCO) with a 0.1-cm cuvette. Data was acquired every 0.1 nm between 200 and 250 nm. Fluorescence spectra of purified importin β (1 μM) and free N-acetyl tryptophan (8 μM) were measured using a fluorometer FP-8200 (JASCO) with a 3 mm × 3 mm cuvette. Samples were dissolved in 50 mM of KPO_4 (pH 7.4) with various concentrations of organic solvents (0–50%). The difference of the peak center wavelength between importin β and free N-acetyl tryptophan was calculated at each solvent condition.

Molecular dynamics simulations

The MD simulations were conducted using the Amber11 package with an ff99SB force field. TFE parameters were set by a general AMBER force field. A GPU version of the particle mesh Ewald method was used with an NVIDIA Tesla M2090. The time step was 2 fsec, a constant 310 K temperature was maintained using Langevin dynamics with a collision frequency of 1.0 $psec^{-1}$, and the pressure was kept at 1 atm with a relaxation time of 2.0 ps. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm, and the long-range interaction cut-off was set to 15.0 Å.

FRAP analysis

EGFP-fused HEAT-rich proteins were expressed in HeLa cells cultured in DMEM supplemented with 10% FCS. The microscopic observations were performed using confocal laser scanning microscope system (FV1200, Olympus), equipped with a stage chamber (Tokai Hit). The nucleus

was bleached by a 488 nm laser at maximum output for 5 sec. After bleaching, the time lapse observation was continued every 15 sec. Signal intensities of the nucleus and cytoplasm were quantified using the MetaMorph software (Molecular Imaging). For RNAi analysis, HeLa cells were also transfected with siRNA purchased from Invitrogen (for CAND1, PPP2R1A and PPP2R1B) by Lipfectamin RNAiMAX (Invitrogen). The antibodies against PP2A A subunit and CAND1 (TIP120A) were purchased from Cell Signaling and Abcam, respectively.

In vitro transport assay

In vitro transport assay was performed as described in the previous report (Yoshimura et al., 2013). The detailed procedures of measurement and data analysis are described also in Supplemental Experimental Procedures.

Author Contributions

S.H.Y. performed all the experiments and analyses with assistance from technicians. S.H.Y. designed the study and analyzed the data. S.H.Y., M.K., and K.T. discussed the results and wrote the paper.

Acknowledgements

This study was financially supported by a Funding Program for Next Generation World-leading Researchers (S.H.Y.), a Grant-in-Aid for Scientific Research (B) (S.H.Y.), a Grant-in-Aid for Young Scientists (A) (S.H.Y.), and a Grant-in-Aid for Scientific Research on Priority Areas (S.H.Y. and

K.T.) from the Japan Society for the Promotion of Science (JSPS), and a Grant-in-Aid for Scientific Research on Innovative Areas (M.K. and K.T.) from the Ministry of Education, Culture, Sports, Science and Technology, Japan. We thank S. Iwasaka, Y. Takashima, M. Minobe and K. Ogawa for technical assistances. The authors declare no competing financial interest. Correspondence and requests for materials should be addressed to S.H.Y. (yoshimura@lif.kyoto-u.ac.jp).

References

Asally, M., and Yoneda, Y. (2005). Beta-catenin can act as a nuclear import receptor for its partner transcription factor, lymphocyte enhancer factor-1 (lef-1). Exp. Cell Res. *308*, 357-363.

Bayliss, R., Littlewood, T., and Stewart, M. (2000). Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell *102*, 99-108.

Bayliss, R., Littlewood, T., Strawn, L.A., Wente, S.R., and Stewart, M. (2002). GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. J. Biol. Chem. *277*, 50597-50606.

Bednenko, J., Cingolani, G., and Gerace, L. (2003). Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. J. Cell Biol. *162*, 391-401.

Bednenko, J., Cingolani, G., and Gerace, L. (2003). Nucleocytoplasmic transport: navigating the channel. Traffic *4*, 127-135.

Bella, J., Hindle, K.L., McEwan, P.A., and Lovell, S.C. (2008). The leucine-rich repeat structure. Cell Mol. Life Sci. *65*, 2307-2333.

Cho, U.S., and Xu, W. (2007). Crystal structure of a protein phosphatase 2A heterotrimeric holoenzyme. Nature *445*, 53-57.

Chook, Y.M., and Blobel, G. (1999). Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. Nature *399*, 230-237.

Chook, Y.M., and Blobel, G. (2001). Karyopherins and nuclear import. Curr. Opin. Struct. Biol. *11*, 703-715.

Cingolani, G., Lashuel, H.A., Gerace, L., and Muller, C.W. (2000). Nuclear import factors importin alpha and importin beta undergo mutually induced conformational changes upon association. FEBS Lett. *484*, 291-298.

Cingolani, G., Petosa, C., Weis, K., and Muller, C.W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. Nature *399*, 221-229.

Conti, E., Muller, C.W., and Stewart, M. (2006). Karyopherin flexibility in nucleocytoplasmic transport. Curr. Opin. Struct. Biol. *16*, 237-244.

Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., and Matunis, M.J. (2002). Proteomic analysis of the mammalian nuclear pore complex. J. Cell Biol. *158*, 915-927.

D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. Trends Biochem. Sci. *28*, 655-662.

Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., and Rexach, M. (2003). Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. Proc. Natl. Acad. Sci. U S A *100*, 2450-2455.

Erickson, H.P. (2009). Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biol. Proced. Online *11*, 32-51.

Fagotto, F., Gluck, U., and Gumbiner, B.M. (1998). Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. Curr. Biol. *8*, 181-190.

Fischer, E.S., Scrima, A., Bohm, K., Matsumoto, S., Lingaraju, G.M., Faty, M., Yasuda, T., Cavadini, S., Wakasugi, M., Hanaoka, F., et al. (2011). The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. Cell *147*, 1024-1039.

Forwood, J.K., Lange, A., Zachariae, U., Marfori, M., Preast, C., Grubmuller, H., Stewart, M., Corbett, A.H., and Kobe, B. (2010). Quantitative structural analysis of importin-beta flexibility: paradigm for solenoid protein structures. Structure *18*, 1171-1183.

Frey, S., Richter, R.P., and Gorlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science *314*, 815-817.

Goldenberg, S.J., Cascio, T.C., Shumway, S.D., Garbutt, K.C., Liu, J., Xiong, Y., and Zheng, N. (2004). Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. Cell *119*, 517-528.

Gorlich, D., and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. Annu. Rev. Cell Dev. Biol. *15*, 607-660.

Gorlich, D., Seewald, M.J., and Ribbeck, K. (2003). Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. EMBO J. *22*, 1088-1100.

Groves, M.R., and Barford, D. (1999). Topological characteristics of helical repeat proteins. Curr. Opin. Struct. Biol. *9*, 383-389.

Isgro, T.A., and Schulten, K. (2005). Binding dynamics of isolated nucleoporin repeat regions to importin-beta. Structure *13*, 1869-1879.

Janssens, V., Longin, S., and Goris, J. (2008). PP2A holoenzyme assembly: in cauda venenum (the sting is in the tail). Trends Biochem. Sci. *33*, 113-121.

Koike, M., Kose, S., Furuta, M., Taniguchi, N., Yokoya, F., Yoneda, Y., and Imamoto, N. (2004). beta-Catenin shows an overlapping sequence requirement but distinct molecular interactions for its bidirectional passage through nuclear pores. J. Biol. Chem. *279*, 34038-34047.

Kotera, I., Sekimoto, T., Miyamoto, Y., Saiwaki, T., Nagoshi, E., Sakagami, H., Kondo, H., and Yoneda, Y. (2005). Importin alpha transports CaMKIV to the nucleus without utilizing importin beta. EMBO J. *24*, 942-951.

Kumeta, M., Yamaguchi, H., Yoshimura, S.H., and Takeyasu, K. (2012). Karyopherin-independent spontaneous transport of amphiphilic proteins through the nuclear pore. J. Cell Sci. *125*, 4979-4984.

Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E., and Corbett, A.H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. J. Biol. Chem. *282*, 5101-5105.

Lee, S.J., Imamoto, N., Sakai, H., Nakagawa, A., Kose, S., Koike, M., Yamamoto, M., Kumasaka, T., Yoneda, Y., and Tsukihara, T. (2000). The adoption of a twisted structure of importin-beta is essential for the protein-protein interaction required for nuclear transport. J. Mol. Biol. *302*, 251-264.

Liu, S.M., and Stewart, M. (2005). Structural basis for the high-affinity binding of nucleoporin Nup1p to the Saccharomyces cerevisiae importin-beta homologue, Kap95p. J. Mol. Biol. *349*, 515-525.

Lubert, E.J., and Sarge, K.D. (2003). Interaction between protein phosphatase 2A and members of the importin superfamily. Biochem. Biophys. Res. Commun. *303*, 908-913.

Marsella, L., Sirocco, F., Trovato, A., Seno, F., and Tosatto, S.C. (2009). REPETITA: detection and discrimination of the periodicity of protein solenoid repeats by discrete Fourier transform.

Bioinformatics *25*, i289-295.

Mohr, D., Frey, S., Fischer, T., Guttler, T., and Gorlich, D. (2009). Characterisation of the passive permeability barrier of nuclear pore complexes. EMBO J. *28*, 2541-2553.

Naim, B., Zbaida, D., Dagan, S., Kapon, R., and Reich, Z. (2009). Cargo surface hydrophobicity is sufficient to overcome the nuclear pore complex selectivity barrier. EMBO J. *28*, 2697-2705.

Neuwald, A.F., and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. Genome Res. *10*, 1445-1452.

O'Reilly, A.J., Dacks, J.B., and Field, M.C. (2011). Evolution of the karyopherin-beta family of nucleocytoplasmic transport factors; ancient origins and continued specialization. PLoS One *6*, e19308.

Ori, A., Banterle, N., Iskar, M., Andres-Pons, A., Escher, C., Khanh Bui, H., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., Lemke, E.M. and Beck, M. (2013). Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. Mol. Syst. Biol. *9*, 648.

Otsuka, S., Iwasaka, S., Yoneda, Y., Takeyasu, K., and Yoshimura, S.H. (2008). Individual binding pockets of importin-beta for FG-nucleoporins have different binding properties and different sensitivities to RanGTP. Proc. Natl. Acad. Sci. U S A *105*, 16101-16106.

Paroni, G., Cernotta, N., Dello Russo, C., Gallinari, P., Pallaoro, M., Foti, C., Talamo, F., Orsatti, L., Steinkuhler, C., and Brancolini, C. (2008). PP2A regulates HDAC4 nuclear import. Mol. Biol. Cell *19*, 655-667.

Ribbeck, K., and Gorlich, D. (2002). The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. EMBO J. *21*, 2664-2671.

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. J. Cell Biol. *148*, 635-651.

Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. Cell *139*, 468-484.

Stewart, M. (2007). Molecular mechanism of the nuclear protein import cycle. Nat. Rev. Mol. Cell

Biol. *8*, 195-208.

Sun, C., Yang, W., Tu, L.C., and Musser, S.M. (2008). Single-molecule measurements of importin alpha/cargo complex dissociation at the nuclear pore. Proc. Natl. Acad. Sci. U S A *105*, 8613-8618.

Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., and Hara-Nishimura, I. (2010). Identification and characterization of nuclear pore complex components in Arabidopsis thaliana. Plant Cell *22*, 4084-4097.

Tu, L.C., and Musser, S.M. (2011). Single molecule studies of nucleocytoplasmic transport. Biochim. Biophys. Acta *1813*, 1607-1618.

Wiechens, N., and Fagotto, F. (2001). CRM1- and Ran-independent nuclear export of beta-catenin. Curr. Biol. *11*, 18-27.

Xu, D., Farmer, A., and Chook, Y.M. (2010). Recognition of nuclear targeting signals by Karyopherin-beta proteins. Curr. Opin. Struct. Biol. *20*, 782-790.

Yokoya, F., Imamoto, N., Tachibana, T., and Yoneda, Y. (1999). beta-catenin can be transported into the nucleus in a Ran-unassisted manner. Mol. Biol. Cell *10*, 1119-1131.

Yoshimura, S.H., Otsuka, S., Kumeta, M., Taga, M. and Takeyasu, K. (2013) Intermolecular disulfide bonds among nucleoporins regulate karyopherin-dependent nuclear transport. J. Cell Sci. *126*, 3141-3150.

Figure Legends

Figure 1 - Amphiphilic structure of importin β and its hydrophobic interactions (A) Crystal structure of yeast importin β (Kap95) modified from PDB file 3ND2. The structures of other karyopherin β family proteins and HEAT-rich proteins are also shown in Figure S1. (B) Each HEAT motif contains 2 α-helices: the A-helix (light green) and the B-helix (dark green). (C) Wheel-helix models of the α-helices in HEAT motifs #14–16 of mouse importin β1. Hydrophobic amino acids are represented by shaded circles. (D) Amino acid composition at the inner and outer surfaces of the HEAT A-helices and B-helices in mouse importin β1. The distributions of hydrophobic, polar, and charged residues within the helix are summarized.

Figure 2 - HEAT motif-rich proteins pass through the NPC (A) HEAT motif-containing proteins in the human protein database. HEAT motif-rich proteins, which do not belong to karyopherin β family, are listed. The name of the protein, total amino acids, and number of HEAT motifs assigned by the National Center for Biotechnology Information (NCBI) and by our own strategy (Supplemental Experimental Procedures) are shown. The intracellular distribution examined by EGFP-tagged proteins expressed in HeLa cells, $t_{1/2}$ value obtained from FRAP analysis (Figure 3), and the influx rate constant (k_{in}) obtained from the *in vitro* assay shown in (B) are summarized. n.d.: not determined. n.e.: not examined. Asterisk (*) represents the value obtained from HEAT-rich domain fragment and not from the full-length protein. For nuclear accumulation, (-): not detectable, (+): weakly detectable, (++): equivalent to the cytoplasm and (+++): strongly accumulated. (B) *In*

vitro import assay of EGFP-fused HEAT-rich proteins. Left panels, images captured at 30 min; Right panels, fluorescence intensity ratio (nucleus [nuc.] over external [ext.] media) over time, with or without wheat germ agglutinin (WGA). k_{in}, influx rate constant without WGA; bar = standard deviation from >5 different measurements. The rate constant obtained here are also summarized in (A). The import assay results using HA-tagged HEAT proteins are also shown in Figure S2C. (C) The summary of *in vitro* import assay using HEAT-rich proteins (●) and non-HEAT proteins (◆). The list of non-HEAT proteins is in Figure S2D. The Stokes radius of the protein was obtained by gel filtration chromatography using reference proteins (Erickson, 2009). (D) Pull-down assay between HEAT-rich proteins and nucleoporin. GST tagged Nup62(FG) was incubated with purified HEAT-rich proteins and analyzed by SDS-PAGE.

Figure 3 - Shuttling of HEAT-rich proteins across the nuclear envelope in living cells EGFP-fused HEAT-rich proteins were expressed in HeLa cells. The fluorescence signal in the nucleus was bleached by irradiation with a strong laser (488 nm) at t = 0, and then time-lapse observation continued for 15 min. Fluorescence images at t = –0.5, +0.5, and 15 min are shown (left panels). The fluorescence signals from the nucleus (black) and the cytoplasm (gray) were quantified and plotted against time (right panel). The bars represent the standard deviation of at least three independent measurements.

Figure 4 - HEAT motif-rich subunit can assist the nuclear transport of CAND1-cullin complex

(A) Crystal structures of a CAND1-cullin1 complex (modified from PDB: 1U6G). CAND1 (HEAT-rich subunit, dark grey) forms a heterotrimeric complex with cullin1 (light grey) and Roc1 (not shown). **(**B) Pull-down assay between CAND1 and GST-fused cullin (cullin1 [1–497] and cullin4B [192–638]). Bound (b) and unbound (u) fractions are shown. (C) GST-tagged cullin4B incubated with digitonin-treated HeLa cells with or without CAND1 was detected by indirect immunofluorescence with an anti-GST antibody. (D) Control- and CAND1-knockdown HeLa cells were subjected to immunoblot analysis using anti-CAND1 antibody. (E) FRAP analysis of EGFP-fused cullin1 (upper panels) or cullin4B (lower panels) in CAND1-knockdown and control-knockdown cells. Left panels: snap shots prior to bleaching (pre-bleach), after bleaching (post-bleach, t = 0), and 30 min after bleaching (t = 30). Right panel: fluorescence intensities in the nucleus (black) and cytoplasm (grey) of the bleached cell (indicated by an arrow) were plotted against time.

Figure 5 - HEAT motif-rich subunit can assist the nuclear transport of PP2A complex (A) Crystal structures of protein phosphatase 2A complex (PDB: 2IAE) (Cho et al., 2007). PP2A is a heterotrimeric complex with an A-subunit (HEAT-rich, black), B-subunit (regulatory, light grey), and C-subunit (catalytic, dark grey). (B, C) FRAP analysis of PP2A. The HEAT-rich A-subunit of human PP2A (PPP2R1A or PPP2R1B) was knocked-down in HeLa cells by RNAi as shown in D. The EGFP-fused C-subunit (PPP2Cα) was expressed, and flux across the nuclear envelope was analyzed by FRAP as described in Figure 4E. Snapshot images prior to bleaching (pre-bleach), after

bleaching (post-bleach, t = 0), and 15 min after bleaching (t = 15) are shown (B). Bleached cells are indicated by arrows. Fluorescence intensity in the nucleus was quantified and plotted against time (C). (D) RNAi of A subunit. Control- and PPP2R1A-knockdown HeLa cells were subjected to immunoblot analysis using anti-A-subunit antibody. The position of PPP2R1A (~66 kDa) is indicated by an arrow. The knockdown effect of importin 9, which is known to interact with A subunit, is shown in Figure S3.

Figure 6 - Conformational changes of HEAT-rich proteins (A) HEAT-rich proteins firmly interact with hydrophobic groups. Purified HEAT-rich proteins (importin β, CAND1 and PPP2R1A) were incubated with octyl-, phenyl-, or butyl-sepharose in the absence (-) or presence (+) of 500 mM NaCl. The bound proteins were analyzed by SDS-PAGE and CBB staining (left panel). The band intensities from four independent experiments on importin β are representatively quantified and summarized in the right panel. (B) The fluorescence spectra of purified HEAT-rich proteins in various solvents indicated were measured at an excitation wavelength of 290 nm. The solvent-induced shifts of the peak wavelength were plotted against the concentration of alcohol (mean value from four independent experiments, relative to that in the absence of alcohol). The direct effect of the solvent on tryptophan fluorescence peak was examined by measuring free N-acetyl tryptophan in the same solvent, and was subtracted from the results. The results from the mutants (Trp is replaced by Phe) are summarized in Figure S4D. (C) CD spectra of HEAT-rich proteins in various organic solvents indicated. The molar ellipticity at 222 nm (the α-helix-specific

peak) is plotted against the concentration of organic solvent. In the case of importin β, ~60% of the molecule is in α-helical conformation, which is in good agreement to the crystal structure (Figure 1). Raw spectra data are shown in Figure S4F. For simplicity, only mean values from three independent experiments are plotted.

Figure 7 - Conformational flexibility is necessary for NPC passage. Purified HEAT-rich proteins were treated with either BS^3 (bifunctional crosslinker) or sulfo-NHS (mono-functional group, negative control), and then were subjected to the *in vitro* transport assay as described in Figure 2B. The microscopic images from sulfo-NHS-treated (non-crosslinked) and BS^3-treated (crosslinked) proteins during the time-lapse observation are shown in left panels. The nuclear signal intensity was plotted against time (right panel). See also Figure S5 for the structural properties of crosslinked proteins.

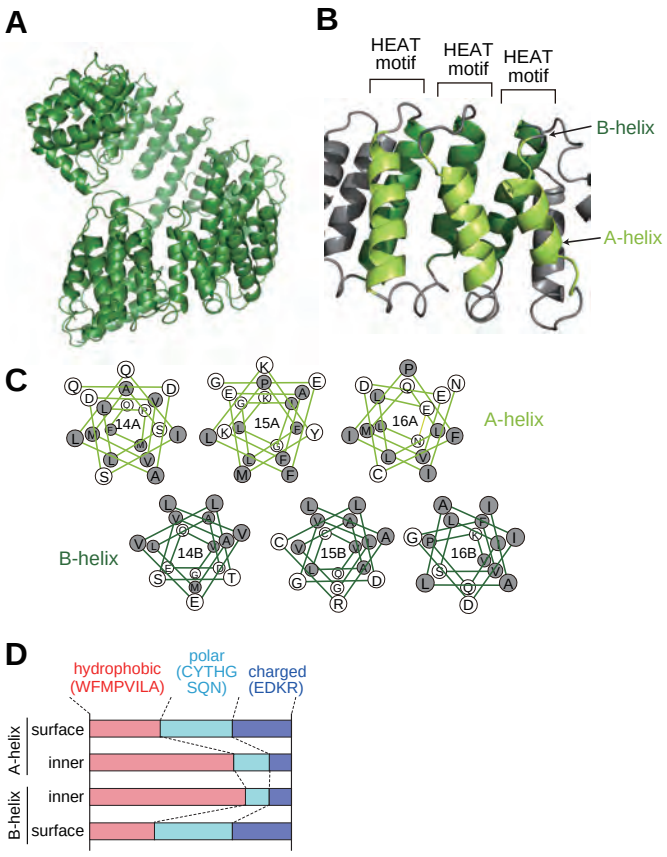
Figure 8 - Molecular dynamics simulation of importin β in different solvents The initial structure of importin β (modified from PDB: 1UKL) was mixed with solvent molecules (water or 50% (v/v) TFE/water), and the molecular dynamics simulation was performed. The RMSd plot during the simulation time is shown in Figure S6A. One representative result from the 3 independent trajectories is summarized here (A, C, E, and G in water and B, D, F, and H in 50% TFE). (A, B) Representative structures at 200 ns. (C, D) Changes in the length of each α-helix relative to the initial state. The bottom panel: color scale (±10 Å). (E, F) Changes in the distance

between adjacent HEAT repeats relative to the initial state. The bottom panel: color scale (±7 Å). The results from other two trajectories are in Figure S6B. (G, H) Molecular surface representation with hydrophobic residues in red. (I) Quantification of surface hydrophobicity. The SESA of hydrophobic residues were quantified by using water molecule as a probe (r_s = 1.4 Å) at 200 ns; bar = standard deviation from 3 trajectories. The detailed evaluation of the surface area is described in Figure S6C. (J) The hydrophobic SESA was plotted against time. Magenta, importin β in 50% TFE; Cyan, in water. Importin β was transferred from TFE to water at 100 (light blue) or 200 (dark blue) ns.

Figure 9 - Structural analysis of cargo-loaded importin β (A) Pull-down assay between GST-fused importin β-binding (IBB) domain of importin α and importin β in the presence of different concentrations of TFE. (B) The effect of TFE on the tryptophan fluorescence peak (blue) and CD value at 220 nm (red) of IBB-importin β complex was examined as described in Figures 6B and 6C and summarized. The IBB fragment does not contain any tryptophan residues. (C–H) The crystal structure of the importin β-IBB complex (PDB: 1QGK) was used as an initial structure, and was subjected to molecular dynamics simulation in water (C, E and G) and 50% TFE (D, F and H) as described in Figure 8. Structures at 200 ns (C, D) with IBB in yellow. Distance between adjacent HEAT motifs (0–200 ns) (E, F). The right panel: color scale (±7 Å). Molecular surface representations at 200 ns (G, H). Red, hydrophobic surface; yellow, IBB. The RMSd value and length of each α-helix during the simulation time are plotted in Figures S6D and S6E, respectively.

Figure 10 - Molecular mechanism of protein translocation across the NPC (A) Fundamental mechanism for molecular trafficking through the NPC. Proteins with amphiphilic HEAT motifs pass through the NPC, and in some cases transport bound proteins together. Karyopherins uniquely catch and release cargo proteins in response to nucleoplasm-enriched RanGTP, thus producing a concentration gradient of the cargo across the nuclear envelope. (B) Flexible amphiphilic proteins undergo conformational changes during passage through the NPC. The details are described in the text.

Figure1



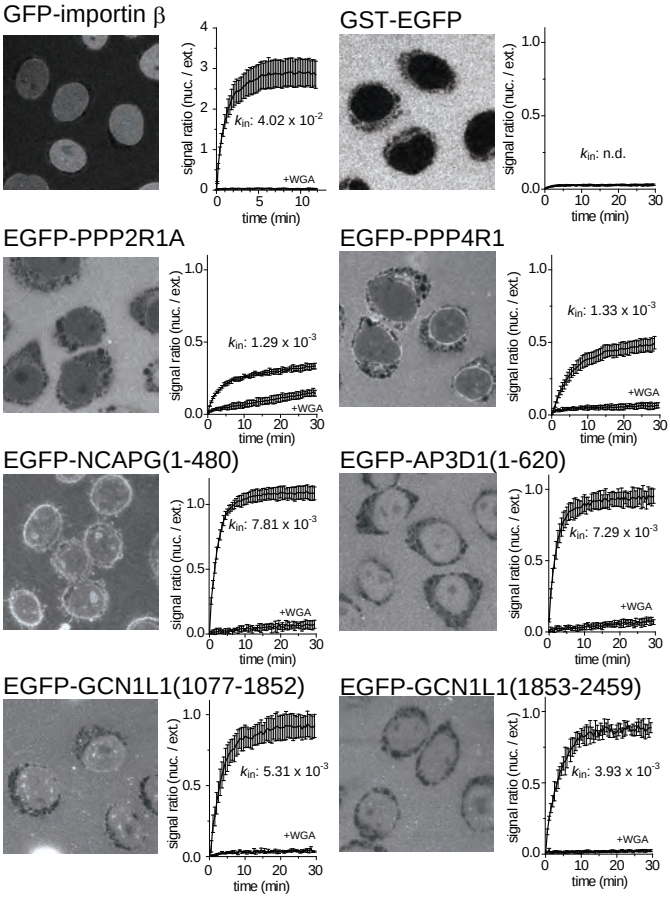
Yoshimura et al. Figure 1

Figure2

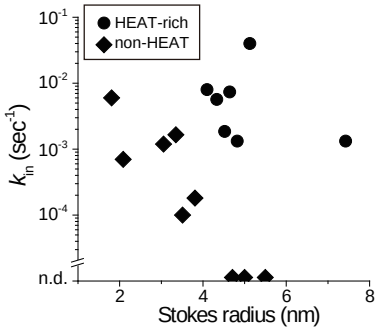
A

protein name	total amino acid	# HEAT (by NCBI)	# HEAT (by eyes)	nuclear localization	$t_{1/2}$ (min) by FRAP	k_{in} by *in vitro*
GCN1L1	2671	24	61-64	-	n.d.	5.31 x 10^{-3}*
HEATR1	2144	2	41-43	-	n.d.	n.e.
CKAP5	2032	10	32-35	-	n.d.	n.e.
BTAF1/TAF172	1849	8	35-37	+	3.01	n.e.
HEATR7B2	1585	15	31-33	+	3.62	n.e.
NCAPD3	1498	4	24-26	+	3.11	5.87 x 10^{-3}*
SAP155	1304	11	19-22	n.e.	n.e.	n.e.
CAND1/TIP120	1230	27	27-28	+++	7.52	n.e.
AP3D1	1153	11	19-20	+	5.31	7.29 x 10^{-3}*
NCAPG	1015	10	22-23	+	2.81	7.81 x 10^{-3}*
HEATR4	979	4	10-12	-	n.d.	n.e.
PPP4R1A	950	14	16	++	4.51	1.33 x 10^{-3}
HEATR2	855	10	16-18	-	n.d.	n.e.
C7orf27	821	2	11-13	+++	3.36	n.e.
HEATR3	680	3	6-10	+++	5.28	2.67 x 10^{-3}
PPP2R1B	601	15	15	++	4.31	1.83 x 10^{-3}
PPP2R1A	589	15	15	++	2.46	1.27 x 10^{-3}

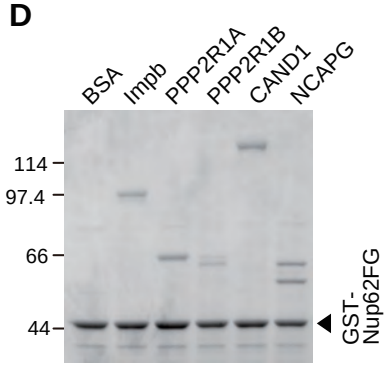
B



C

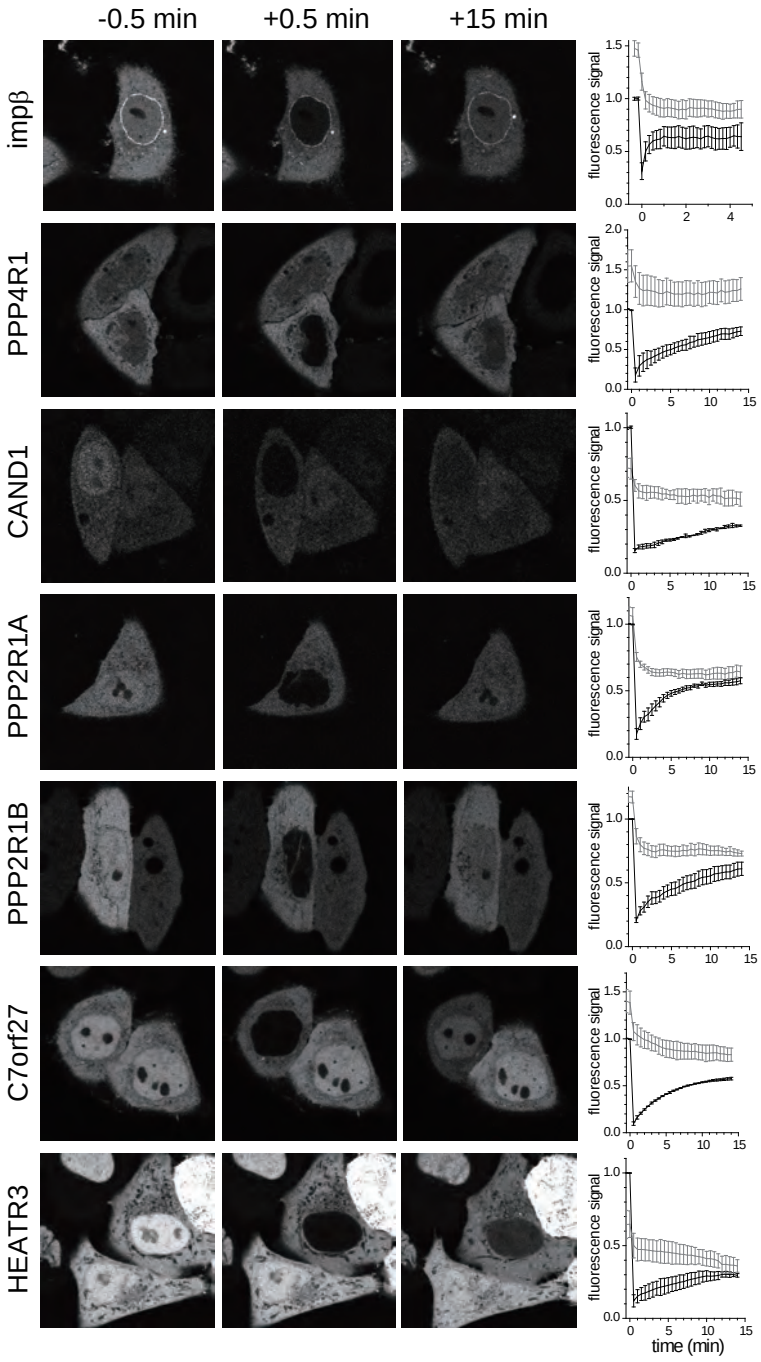


D



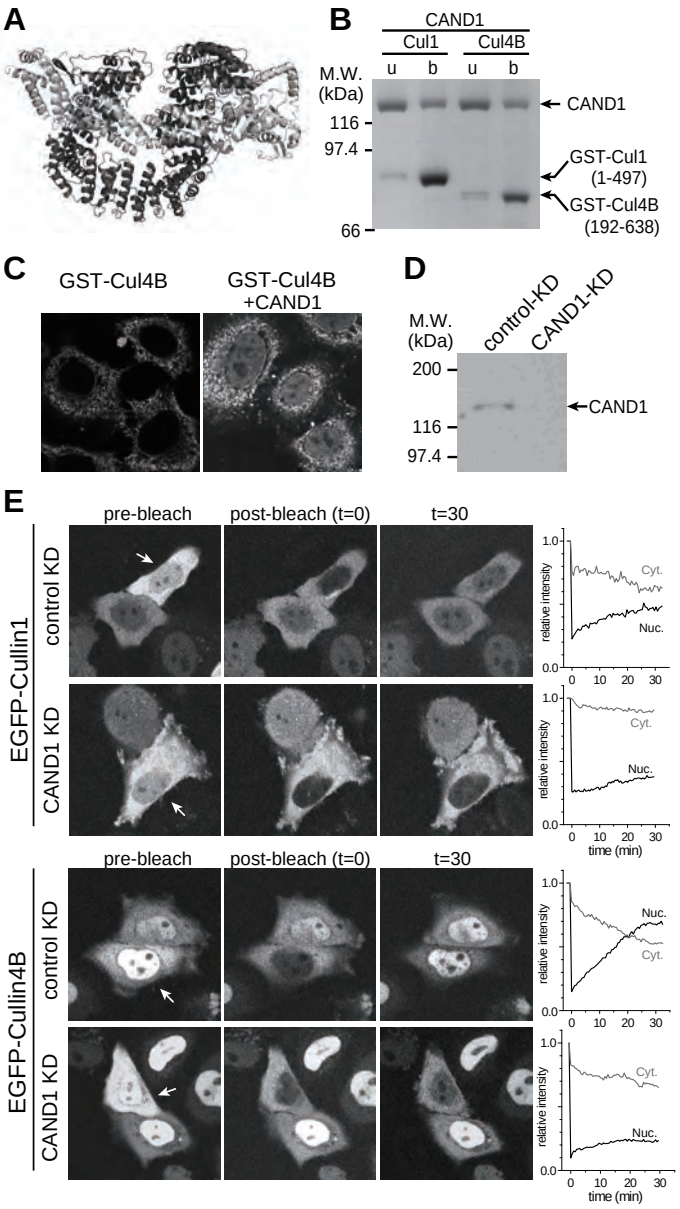
Yoshimura et al. Figure 2

Figure3



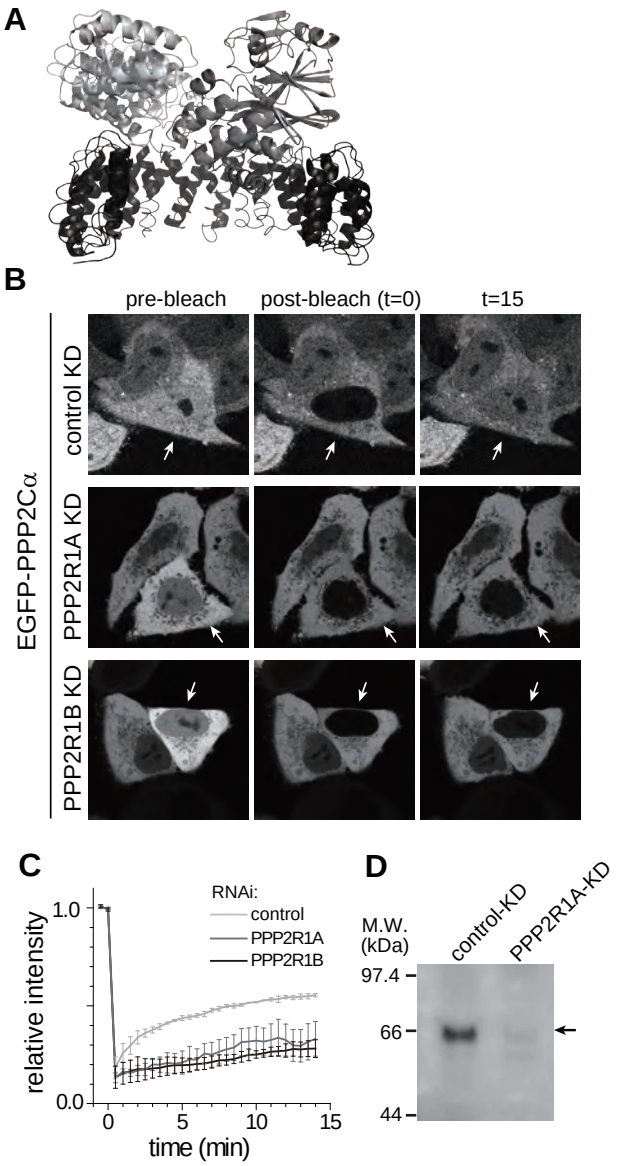
Yoshimura et al. Figure 3

Figure4



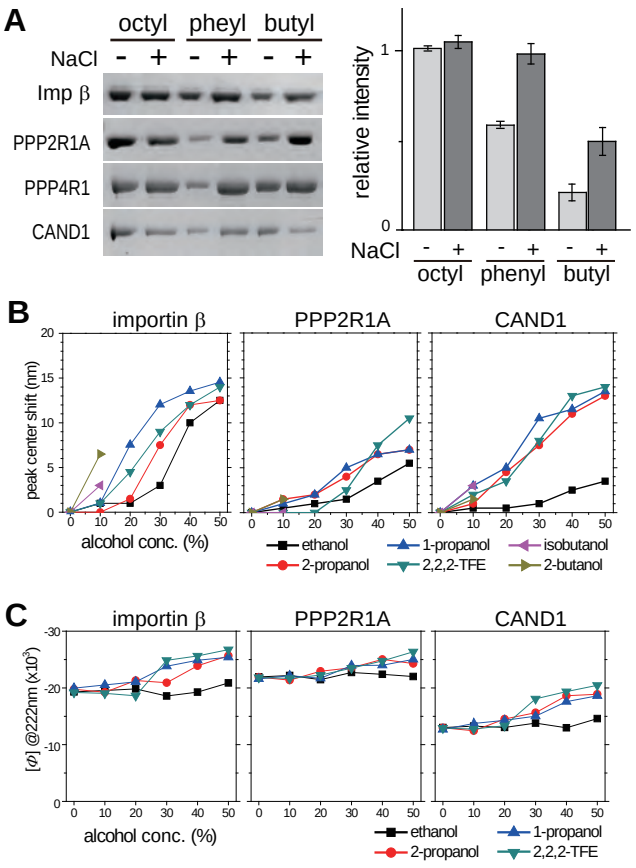
Yoshimura et al. Figure 4

Figure5



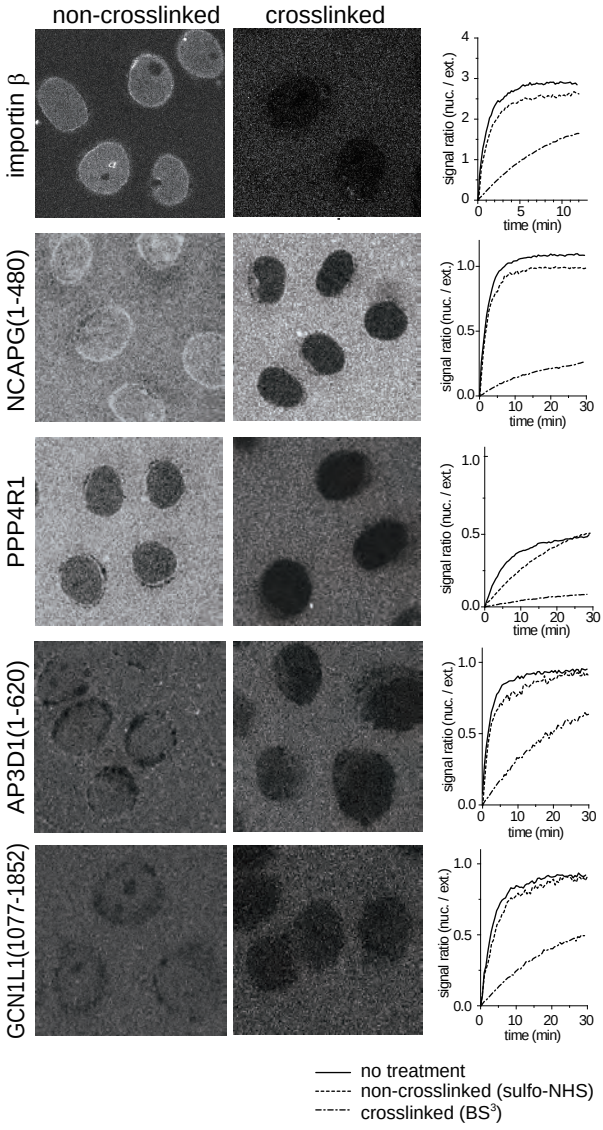
Yoshimura et al. Figure 5

Figure6



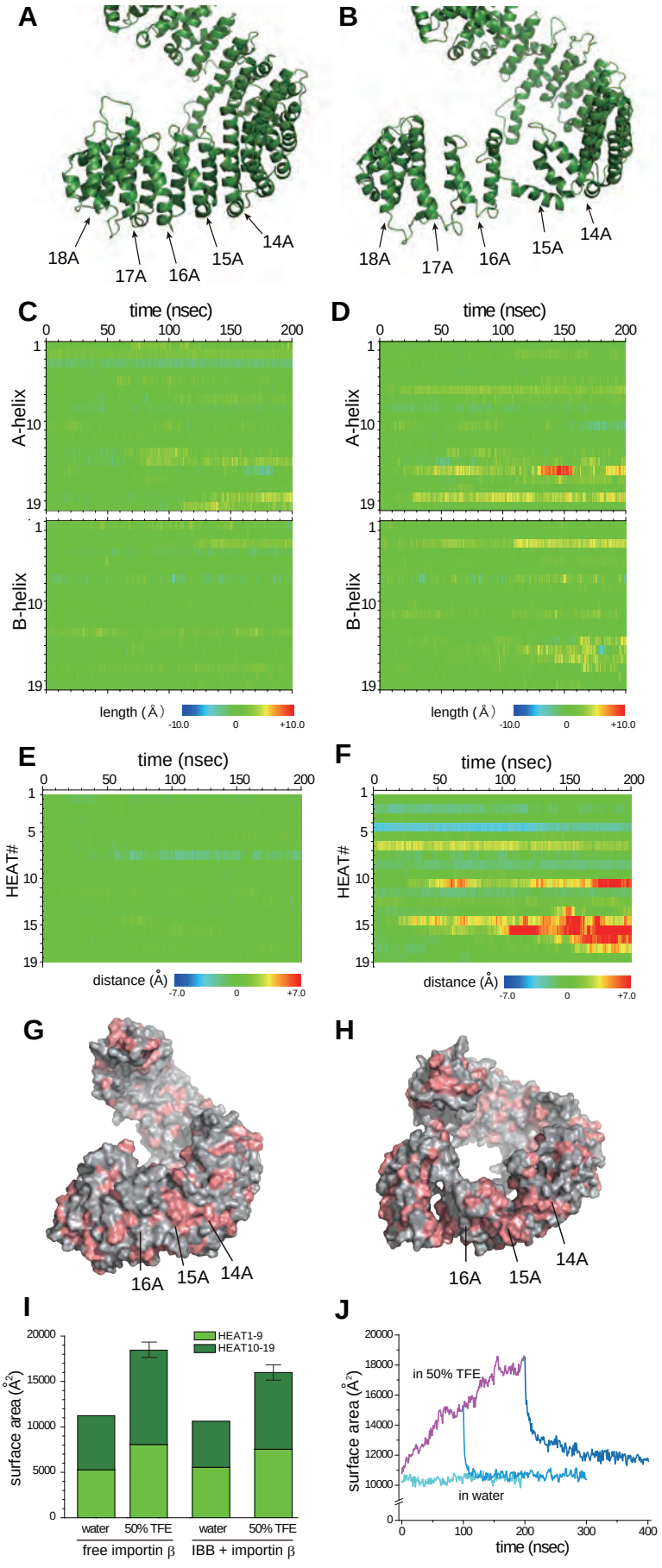
Yoshimura et al. Figure 6

Figure7



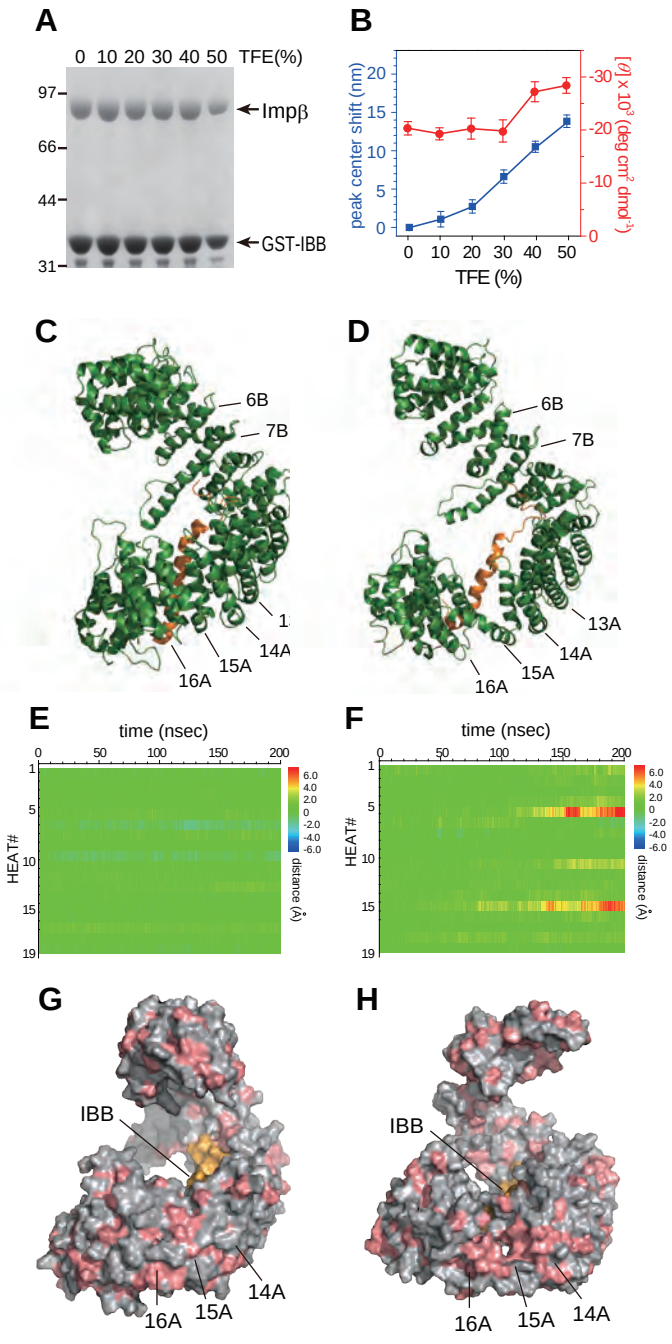
Yoshimura et al. Figure 7

Figure8



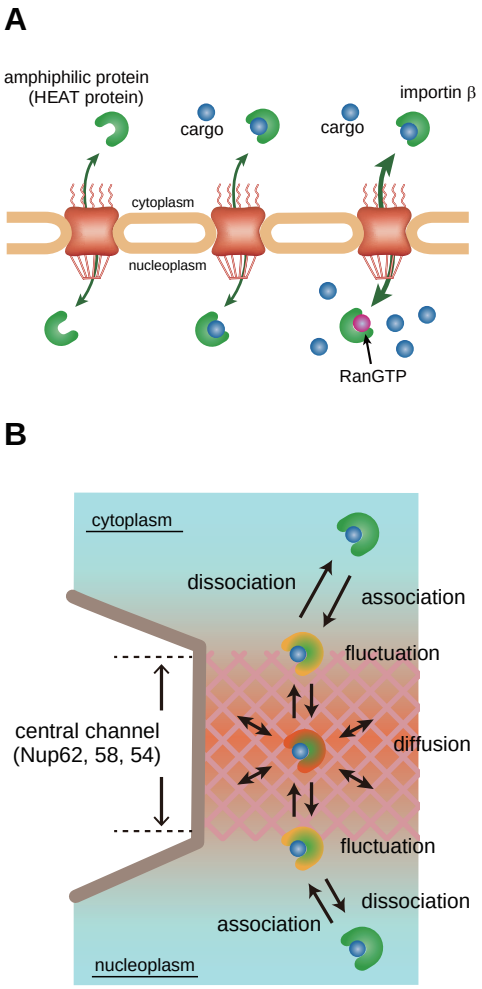
Yoshimura et al. Figure 8

Figure9



Yoshimura et al. Figure 9

Figure10



Yoshimura et al Figure 10

Supplemental Information

Title: Structural mechanism of nuclear transport mediated by importin β and flexible amphiphilic proteins

Authors: Shige H. Yoshimura, Masahiro Kumeta and Kunio Takeyasu

Graduate School of Biostudies, Kyoto University

List of Contents

Figure S1, Related to Figure 1; Structure of karyopherin β family proteins and other HEAT motif-containing proteins.

Figure S2, Related to Figure 2; HEAT motif-rich proteins pass through the NPC.

Figure S3, Related to Figure 5; Involvement of importin 9 in nuclear shuttling of PP2A complex.

Figure S4, Related to Figure 6; Structural properties of pore-forming FG-nucleoporins (FG-Nups) and alcohol-induced conformational change of HEAT-rich proteins.

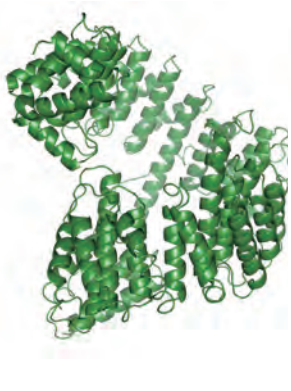
Figure S5, Related to Figure 7; Characterization of crosslinked importin β.

Figure S6, Related to Figure 8 and 9; Additional analyses from the MD simulation.

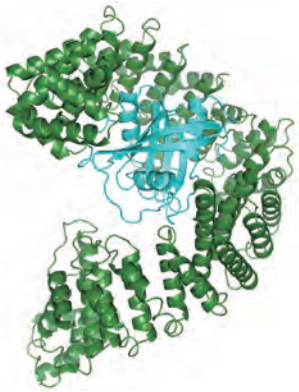
Figure S7, Related to Figure 10; importin b contains multiple binding pockets for FG-Nups.

Supplemental Experimental Procedures

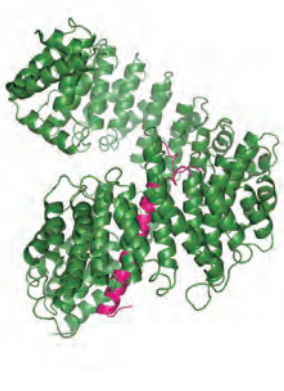
Supplemental references



yeast Importin β
PDB: 3ND2



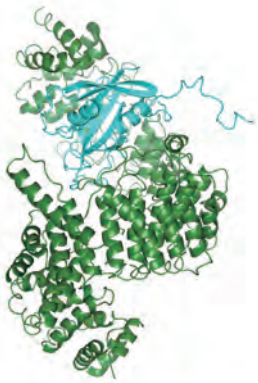
yeast Importin β
+ RanGTP
PDB: 2BKU



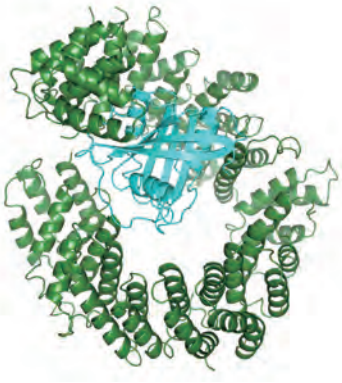
human Importin β
+ importin α IBB
PDB: 1QGK

mouse Importin β
+ SREBP2
PDB: 1UKL

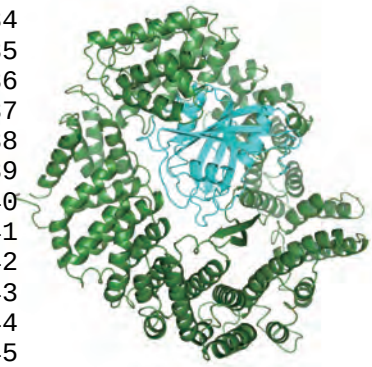
human transportin
PDB: 2QMR



human transportin
+ RanGTP
PDB: 2QMR



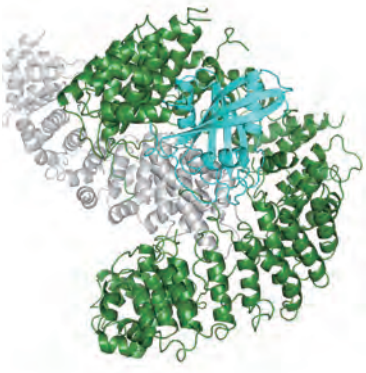
human Importin 13
+ RanGTP
PDB: 2X19



mouse Crm1
+ RanGTP
PDB: 3N1C

yeast Crm1
PDB: 3VYC

yeast Cse1
PDB: 1Z3H

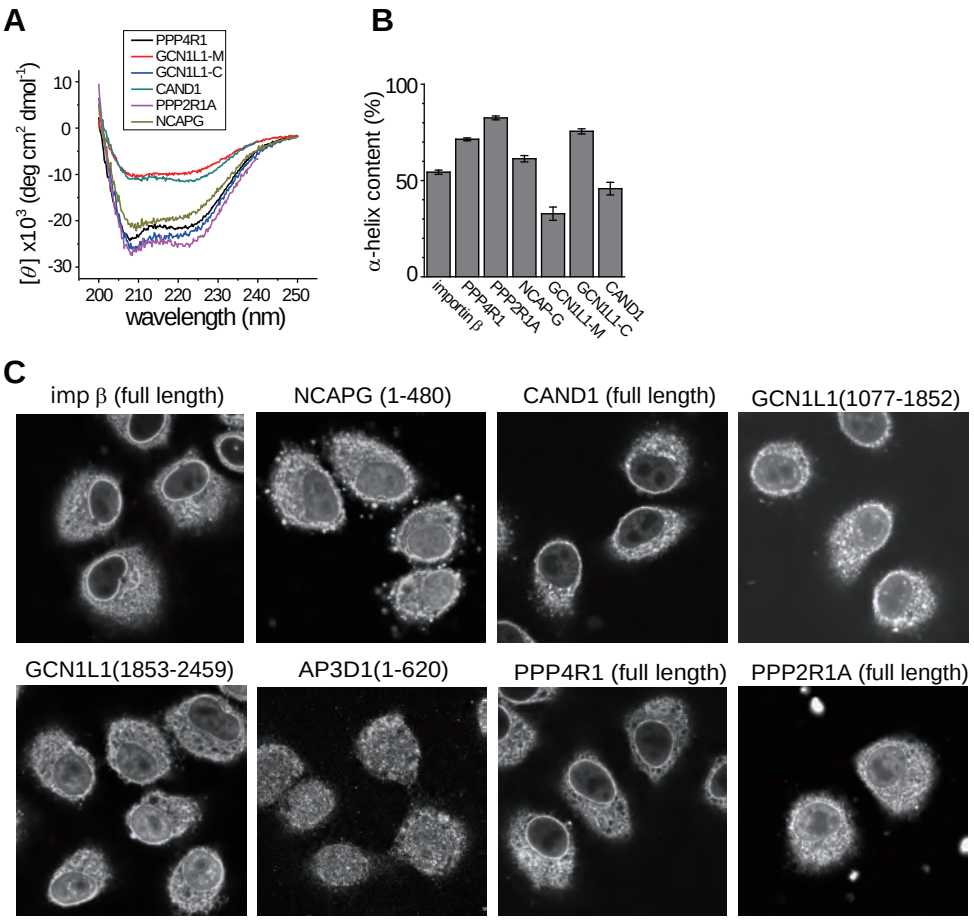


yeast Cse1
+ RanGTP
PDB: 1WA5

human PPP2R1A
from PDB: 2NYL

human CAND1
from PDB: 1U6G

Figure S1, Related to Figure 1; Structure of karyopherin β family proteins and other HEAT motif-containing proteins. Crystal structures of karyopherin β family proteins (green), and those complexed with RanGTP (cyan) or a cargo (magenta) are shown. In yeast CseI:RanGTP complex, the cargo (Kap60p) is shown in gray. Non-karyopherin β proteins which contain HEAT repeats (PPP2R1A and CAND1) are also shown.



D

	Stokes rad. (nm)	k_{in} (sec^{-1})
EGFP	1.88	6.25 x 10^{-3}
mCherry	2.09	9.89 x 10^{-4}
IBB-mCherry	3.36	1.62 x 10^{-3}
MBP-EGFP	3.83	1.74 x 10^{-4}
MBPx2-EGFP	4.70	n.d.
EGFP-CKB	4.98	n.d.
GST-EGFP-CKB	5.52	n.d.
Alexa-ovalbumin	3.05	1.22 x 10^{-3}
TxRed-BSA	3.55	9.70 x 10^{-5}

Figure S2, Related to Figure 2; HEAT motif-rich proteins pass through the NPC. (A, B) Circular dichroism analysis of recombinant HEAT-rich proteins. The raw spectra are shown in (A). The α-helix content was calculated based on the molar ellipticity at 222 nm using the following equation: fa = −([θ]$_{222}$ + 2340)/30300), and summarized in (B). (C) GST-HA-tagged HEAT-rich proteins (or protein fragments) described in Figure 2A were expressed as GST-fusion proteins in bacteria, purified by glutathione beads, and separated from the GST moiety by proteolytic cleavage. HeLa cells were treated with digitonin (40 μg/mL), incubated with the purified proteins at 37 ºC for 30 min, and immediately fixed with 4% paraformaldehyde for 10 min. The cells were then immunostained with an anti-HA antibody (16B12) and observed by confocal laser scanning microscope. Fluorescence was detected in the nucleoplasm. Fluorescence was also detected in the cytoplasmic region, due to non-specific crosslinking of free proteins in the external solution. (D) The stokes radius and influx rate constant (k_{in}) of non-HEAT proteins used in Figure 2C are summarized. The Stokes radii were obtained by gel filtration chromatography as described in Erickson et al., 2009.

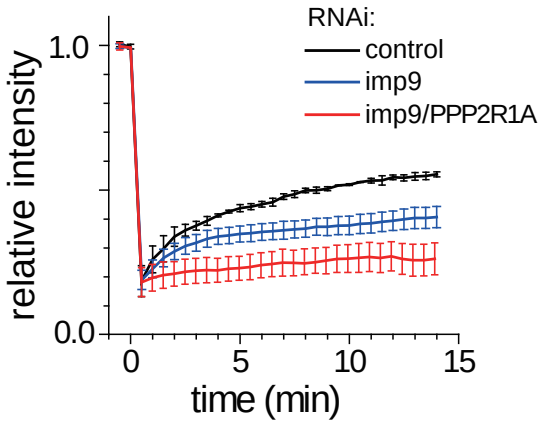


Figure S3, Related to Figure 5; involvement of importin 9 in nuclear shuttling of PP2A complex. FRAP analysis described in Figures 5B and 5C was performed under importin 9-knockdown condition. Knockdown of importin 9 slightly reduced nuclear shutling of PP2A c subunit. The double knowckdown of impotin 9 and PP2A a subunit (PPP2R1A) further reduced the rate.

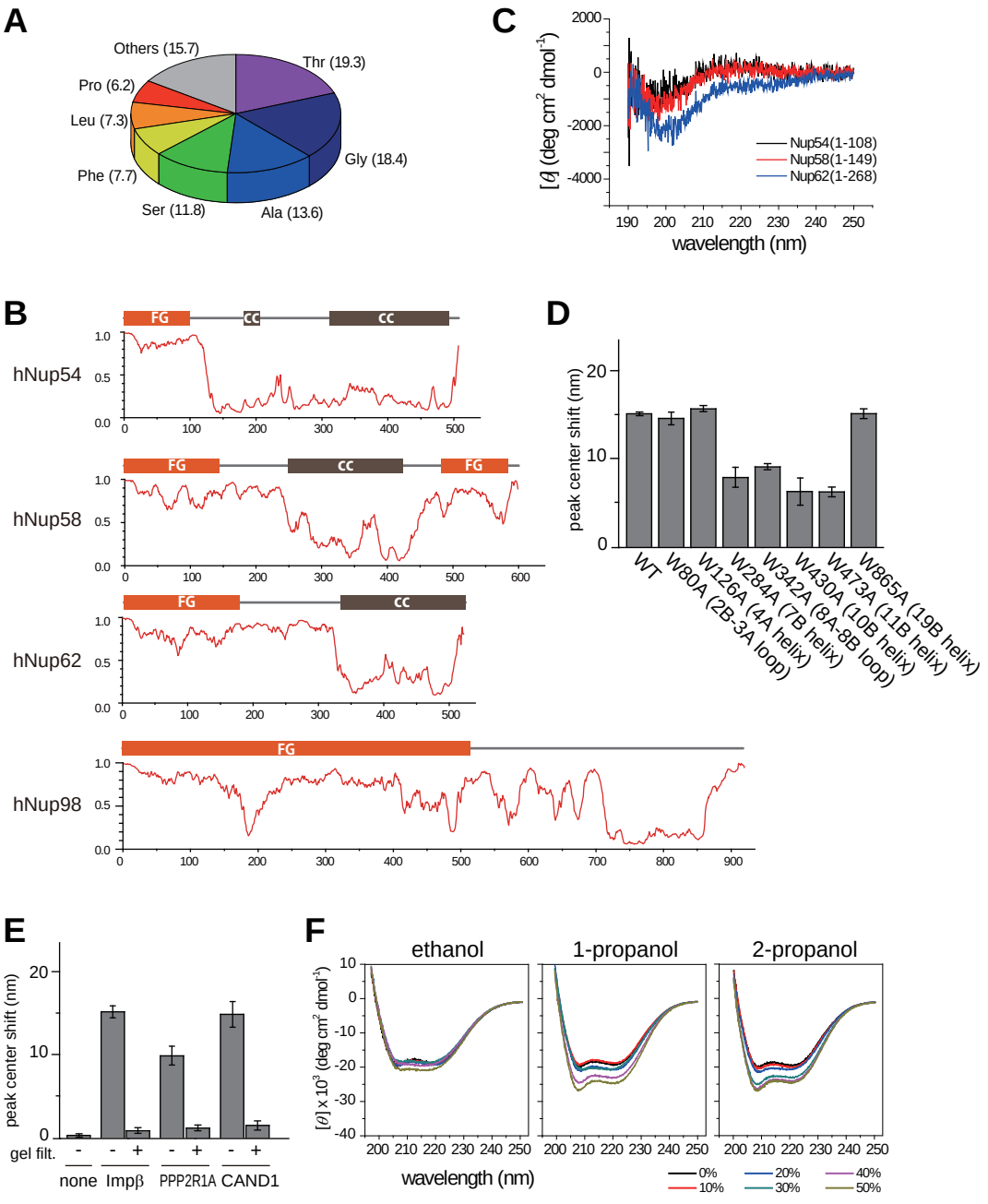


Figure S4, Related to Figure 6; Structural properties of pore-forming FG-nucleoporins (FG-Nups) and alcohol-induced conformational change of HEAT-rich proteins. (A) Amino acid compositions of pore-forming Nups (Nup54, Nup58, Nup62 and Nup98). (B) Structural predictions of pore-forming nucleoporins. The intrinsically disordered regions (IDRs) predicted by the PONDR-fit algorithm are plotted. The coiled-coil regions (cc) and FG-rich regions (FG) are also depicted above the plot. (C) CD spectra of purified IDRs. Human Nup54 (1-108), Nup58 (1-149) and Nup62 (1-268) were expressed in bacterial cells and purified as hexahistidine-tagged proteins, and subjected to CD measurement in 50 mM phosphate buffer (pH 7.4). (D) The analysis of tertiary structures by tryptophan fluorescence. Tryptophan residue in importin β was substituted by phenylalanine, and subjected to fluorescence spectra analysis as described in Figure 6B. The position of the mutated residue in the HEAT repeat is indicated below. (E) Reversible conformational change of HEAT proteins. Purified HEAT-rich proteins were exposed to organic solvent (2-propanol), and subjected to the fluorescence spectrum analysis as described in Figure 6B (-). The exposed proteins were then passed through a gel filtration column to remove organic solvent, and subjected to the same spectrum analysis (+). (F) CD spectra of purified importin β in hydrophobic solvent (ethanol, 1-propanol and 2-propanol). One representative spectrum from three independent experiments is shown.

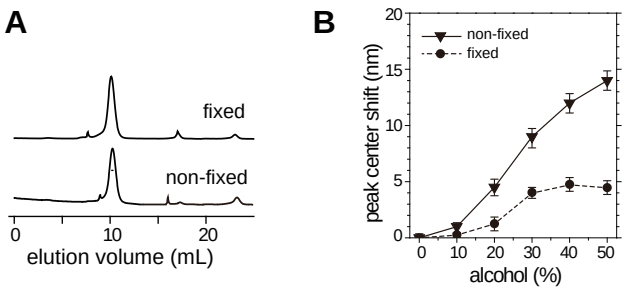
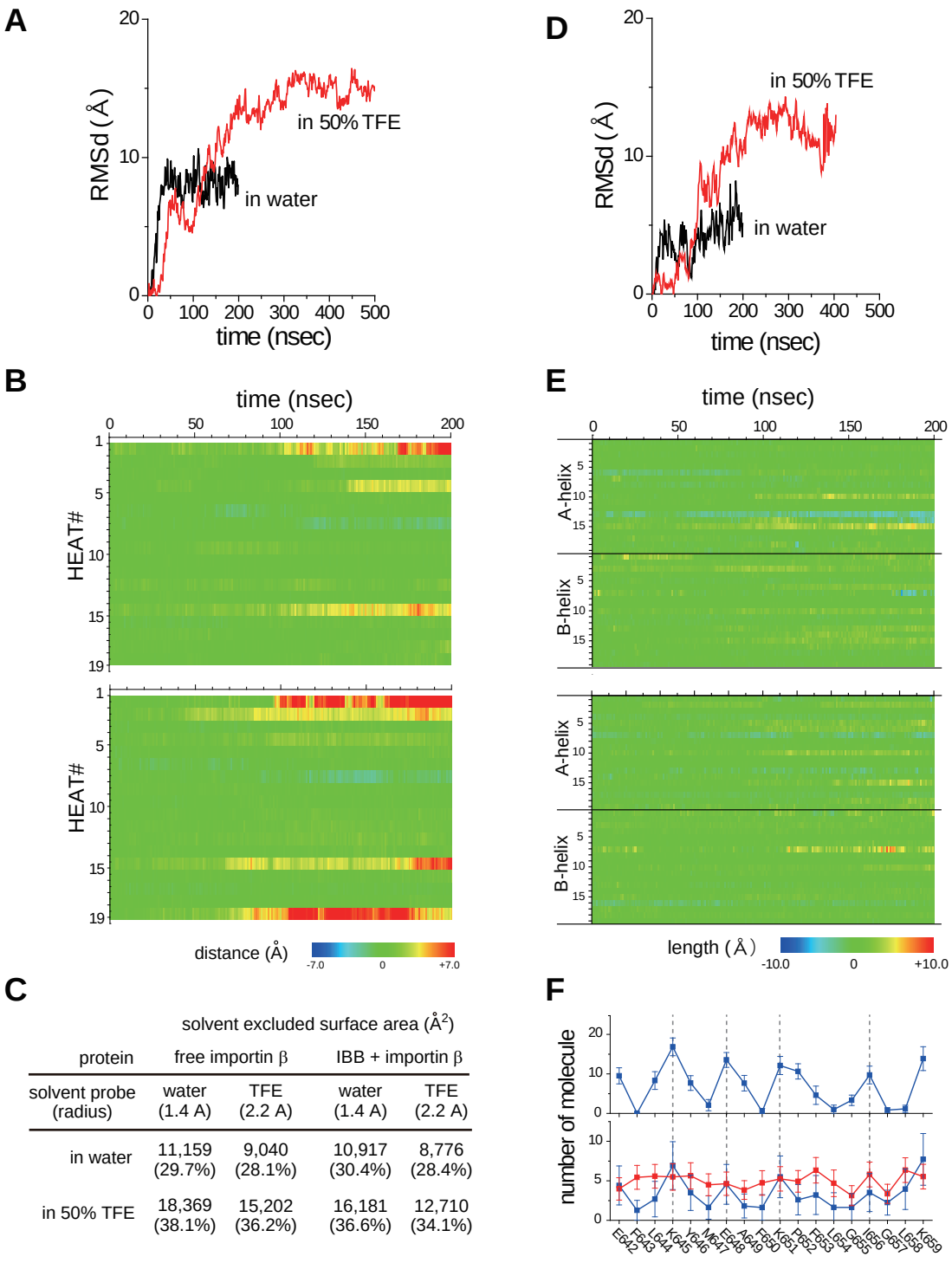


Figure S5, Related to Figure 7; Characterization of fixed importin β. Purified importin β was fixed by a crosslinker (BS^3) to restrict conformational changes. Importin β contains 27 lysine residues, 18 pairs of which could be candidates for the crosslinker (spacer length, 11.4 angstroms). The crosslinked protein was subjected to gel filtration chromatography (Superdex 200HR) to examine inter-molecular crosslinking(A), and to the analysis of fluorescence spectrum as described in Figure 6B to examine the structural flexibility (B). The peak center shift was plotted against the concentration of TFE (%).



C

solvent excluded surface area (Å^2)

protein	free importin β		IBB + importin β	
solvent probe (radius)	water (1.4 A)	TFE (2.2 A)	water (1.4 A)	TFE (2.2 A)
in water	11,159 (29.7%)	9,040 (28.1%)	10,917 (30.4%)	8,776 (28.4%)
in 50% TFE	18,369 (38.1%)	15,202 (36.2%)	16,181 (36.6%)	12,710 (34.1%)

Figure S6, Related to Figures 8 and 9; Additional analyses from the MD simulation. The root mean square deviation (RMSd) during the MD simulation of free importin β (A) and the IBB-importin β complex (D) in water (black) and 50% TFE (red) were plotted against time. (B) HEAT-HEAT distances of importin β simulated in 50% TFE. The simulation and analysis were performed as described in Figure 8. Distances between adjacent HEAT motifs were plotted over simulation time (0-200 ns) for two independent trajectories. Bottom panel: color scale (±7 Å). (C) Summary of hydrophobic solvent excluded surface area (SESA). The SESA values of hydrophobic residues in free importin β and IBB-importin β in water or 50% TFE at 200 ns were calculated by using two different solvent probes (1.4 Å for water and 2.2 Å for TFE). In both cases, the hydrophobic surface areas were higher in 50% TFE than in water. (E) Helical length of each HEAT motif in IBB-importin β complex in water (top) and 50% TFE (bottom). Changes in the length of each α-helix (A-helix and B-helix, No. 1-19) relative to the initial state were plotted against simulation time. Color scale (±10Å). (F) The number of solvent molecules on the protein surface. The number of TFE (red) and water (blue) molecules that are within 5 Å of each amino acid residue in HEAT 15A (E642 - K659) in the simulation in water (upper panels) and in 50% TFE (lower panels) were plotted. The residues that are embedded in a hydrophobic core in water interact with TFE molecules in 50% TFE.

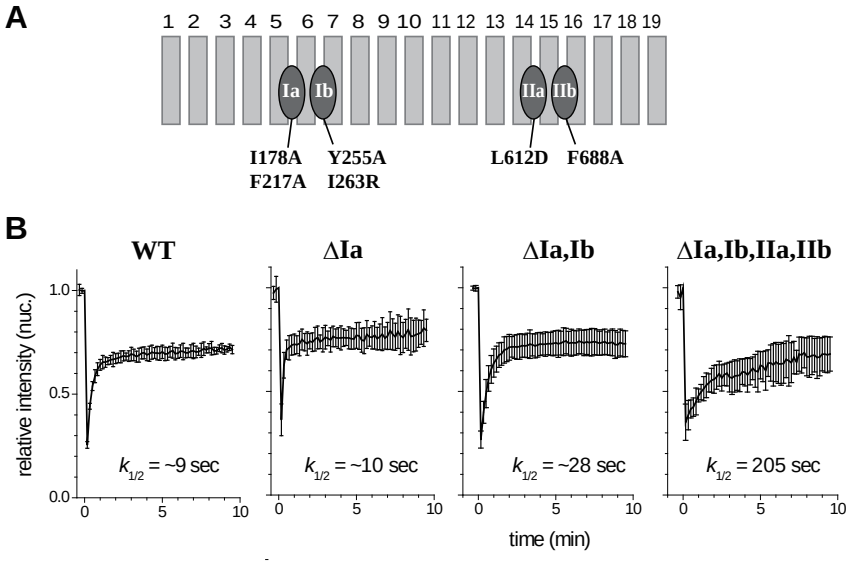


Figure S7, Related to Figure 10; importin β contains multiple binding pockets for FG-Nups. (A) The positions of the binding pockets. The HEAT repeat structure (1-19) of importin β and binding pockets for FG Nups (Ia, Ib, IIa, IIb) are depicted. The amino acid substitutions to abolish the pockets (I178A, F217A, Y256A, I263R, L612D, F688A, described in Bednenko et al., 2003) are also indicated. (B) FRAP analysis of EGFP-Importin β mutants. Importin β which carry mutations in FG-binding pockets were expressed as a fusion protein with EGFP in HeLa cells, and were subjected to FRAP analysis as described in Figure 3.

Supplemental Experimental Procedures

In vitro transport assay

HeLa cells were triple-washed with Transport Buffer (20 mM HEPES-KOH (pH 7.3), 110 mM CH_3COOK, 2 mM $(CH_3COO)_2Mg$, 5 mM CH_3COONa, 0.5 mM EGTA, and 1 mM DTT) and incubated with 0.02% digitonin at 0°C for 10 min to remove the cytoplasm and plasma membrane. The cells were again triple-washed with Transport Buffer, and incubated for 30 min at room temperature to wash away the soluble nucleoplasmic proteins. The nuclei were incubated with purified protein (1–5 μM) and observed every 15 sec for 15–30 min using a confocal laser-scanning microscope (FV1200, Olympus). To check the integrity of the nuclear envelope, 10 μg/mL Alexa568-labeled IgG was added to the sample. All image analyses were performed using MetaMorph software (Molecular Imaging). Curve-fitting and other kinetic analyses of the obtained data were performed using Origin software (Light Stone).

Calculation of flux rate constant in the in vitro transport assay

To analyze the rate constants (k_{in} and k_{out}) of passage, the movement of proteins from the cytoplasm to the nucleoplasm was approximated as a single-step event, although it occurs in multiple steps in reality: binding from the cytoplasm to the NPC, movement within the NPC, the release from the NPC to the nucleoplasm, and so forth. Therefore, k_{in} and k_{out} represent the rate constants of a net flow across the nuclear envelope, even though a significant part of the imported cargo that binds to the

NPC returns to the cytoplasm. The model system consists of 2 compartments, namely the nucleus (Nuc) and the cytoplasm (Cyt), which are connected by a pore through which macromolecules can move from one compartment to the other. The rate of macromolecule passage through the pore depends on the rate constant k and the concentration in the original compartment. Therefore, the inward and outward rate (v) of importin β movement can be defined as follows:

$$v_{in} = k_{in}[\text{Imp}\beta]_{\text{cyto}}$$

$$v_{out} = k_{out}[\text{Imp}\beta]_{\text{nuc}}$$

The rate of difference in importin β concentration in the nucleus is given by the following equation:

$$d[\text{Imp}\beta]_{\text{nuc}}/dt = k_{in}[\text{Imp}\beta]_{\text{cyto}} - k_{out}[\text{Imp}\beta]_{\text{nuc}}$$

$$d[\text{Imp}\beta]_{\text{nuc}}/dt + k_{out}[\text{Imp}\beta]_{\text{nuc}} - k_{in}[\text{Imp}\beta]_{\text{cyto}} = 0 \qquad \text{(Eq. 1)}$$

In the *in vitro* assay system, the concentration of importin β in the cytoplasm ($[\text{Imp}\beta]_{\text{cyto}}$) is regarded as constant. Therefore, Equation 1 can be solved as

$$[\text{Imp}\beta]_{\text{nuc}} = k_{in}[\text{Imp}\beta]_{\text{cyto}}/k_{out}\,(1-\exp(-k_{out}\,t))$$

Finding HEAT motifs in human protein database based on secondary structures

Amino acid sequences of the HEAT-rich proteins listed in Figure 2A are shown. HEAT motif positions are highlighted in light (A-helix) and dark (B-helix) green. For PPP2R1A and CAND1, HEAT motif positions were defined based on their crystal structures. Putative HEAT motifs were predicted in other proteins using the following procedure: i) avoid naturally disordered regions that can be predicted by DisProt (underlined regions), ii) predict secondary structures, and iii) check the amphiphilicity of each α-helix (20-25 a.a. in length) by drawing a helical-wheel presentation. The A- and B-helices were distinguished based on the existence of proline residues in the A-helix. Red: acidic residues; blue: basic residues; bold: proline; magenta: proline residues in the A-helix. The number of proline residues in the A- and B-helices is indicated.

NCAPG (1015aa) 22-23 HEAT repeats

MGAERRLLSIKEAFRLAQQPHQNQAKLVVALSRTYRTMDDKTVFHEEFIHYLKYVMVVYKREPAVERVIEFAAKFVTSFHQSDMEDDEEEEDGGLLNYLFTFLLKSHE
ANSNAVRFRVCLLINKLLGSMPENAQIDDDVFDKINKAMLIRLKDKIPNVRIQAVLALSRLQDPKDDECPVVNAYATLIENDSNPEVRRAVLSCIAPSAKTLPKIVGR
TKDVKEAVRKLAYQVLAEKVHMRAMSIAQRVMLLQQGLNDRSDAVKQAMQKHLLQGWLRFSEGNILELLHRLDVENSSEVAVSVLNALFSITPLSELVGLCKNNDGRK
LIPVETLTPEIALYWCALCEYLKSKGDEGEEFLEQILPEPVVYADYLLSYIQSIPVVNEEHRGDFSYIGNLMTKEFIGQQLILIIKSLDTSEEGGRKKLLAVLQEILI
LPTIPISLVSFLVERLLHIIIDDNKRTQIVTEIISEIRAPIVTVGVNNDPADVRKKELKMAEIKVKLIEAKEALENCITLQDFNRASELKEEIKALEDARINLLKETE
QLEIKEVHIEKNDAETLQKCLILCYELLKQMSISTGLSATMNGIIESLILPGIISIHPVVRNLAVLCLGCCGLQNQDFARKHFVLLLQVLQIDDVTIKISALKAIFDQ
LMTFGIEPFKTKKIKTLHCEGTEINSDDEQESKEVEETATAKNVLKLLSDFLDSEVSELRTGAAEGLAKLMFSGLLVSSRILSRLILLWYNPVTEEDVQLRHCLGVFF
PVFAYASRTNQECFEEAFLPTLQTLANAPASSPLAEIDITNVAELLVDLTRPSGLNPQAKTSQDYQALTVHDNLAMKICNEILTSPCSPEIRVYTKALSSLELSSHLA
KDLLVLLNEILEQVKDRTCLRALEKIKIQLEKGNKEFGDQAEAAQDATLTTTTFQNEDEKNKEVYMTPLRGVKATQASKSTQLKTNRGQRKVTVSARTNRRCQTAEAD
SESDHEVPEPESEMKMRLPRRAKTAALEKSKLNLAQFLNEDLS

A-helix with P: 9/22
B-helix with P: 10/22

NCAPD2 (1401 aa), 24-26 HEAT repeats

MAPQMYEFHLPLSPEELLKSGGVNQYVVQEVLSIKHLPPQLRAFQAAFRAQGPLAMLQHFDTIYSILHHFRSIDPGLKEDTLQFLIKVVSRHSQELPAILDDTTLSGS
DRNAHLNALKMNCYALIRLLESFETMASQTNLVDLDLGGKGKKARTKAAHGFDWEEERQPILQLLTQLLQLDIRHLWNHSIIEEEFVSLVTGCCYRLLENPTINHQKN
RPTREAITHLLGVALTRYNHMLSATVKIIQMLQHFEHLAPVLVAAVSLWATDYGMKSIVGEIVREIGQKCPQELSRDPSGTKGFAAFLTELAERVPAILMSSMCILLD
HLDGENYMMRNAVLAAMAEMVLQVLSGDQLEAAARDTRDQFLDTLQAHGHDVNSFVRSRVLQLFTRIVQQKALPLTRFQAVVALAVGRLADKSVLVCKNAIQLLASFL
ANNPFSCKLSDADLAGPLQKETQKLQEMRAQRRTAAASAVLDPEEEWEAMLPELKSTLQQLLQLPQGEEEIPEQIANTETTEDVKGRIYQLLAKASYKKAIILTREAT
GHFQESEPFSHIDPEESEETRLLNILGLIFKGPAASTQEKNPRESTGNMVTGQTVCKNKPNMSDPEESRGNDELVKQEMLVQYLQDAYSFSRKITEAIGIISKMMYEN
ITTVVQEVIEFFVMVFQFGVFQALFGVRRMLPLIWSKEPGVFEAVLNAYRQLYLNPKGDSARAKAQALIQNLSLLLVDASVGTIQCLEEILCEFVQKDELKPAVTQLL

WERATEKVACCPLERCSSVMLLGMMARGKPEIVGSNLDTLVSIGLDEKFPQDYRLAQQVCHAIANISDRRKPSLGKRHPPFRLPQEHRLFERLRETVTKGFVHPDPLW
IPFKEVAVTLIYQLAEGPEVICAQILQGGAKQALEKLEEKRTSQEDPKESPAMLPTFLLMNLLSLAGDVALQQLVHLEQAVSGELCRRRVLREEQEHKTKDPKEKNTS
SETTMEEELGLVGATADDTEAELIRGICEMELLDGKQTLAAFVPLLLKVCNNPGLYSNPDLSAAASLALGKFCMISATFCDSQLRLLFTMLEKSPLPIVRSNLMVATG
DLAIRFPNLVDPWTPHLYARLRDPAQQVRKTAGLVMTHLILKDMVKVKGQVSEMAVLLIDPEPQIAALAKNFFNELSHKGNAIYNLLPDIISRLSDPELGVEEEPFHT
IMKQLLSYITKDKQTESLVEKLCQRFRTSRTERQQRDLAYCVSQLPLTERGLRKMLDNFDCFGDKLSDESIFSAFLSVVGKLRRGAKPEGKAIIDEFEQKLRACHTRG
LDGIKELEIGQAGSQRAPSAKKPSTGSRYQPLASTASDNDFVTPEPRRTTRRHPNTQQRASKKKPKVVFSSDESSEEDLSAEMTEDETPKKTTPILRASARRHRS

A-helix with P: 9/24
B-helix with P: 7/24

SAP155 (1304 aa) 19-22 HEAT repeats
MAKIAKTHEDIEAQIREIQGKKAALDEAQGVGLDSTGYYDQEIYGGSDSRFAGYVTSIAATELEDDDDDYSSSTSLLGQKKPGYHAPVALLNDIPQSTEQYDPFAEHR
PPKIADREDEYKKHRRTMIISPERLDPFADGGKTPDPKMNARTYMDVMREQHLTKEEREIRQQLAEKAKAGELKVVNGAAASQPPSKRKRRWDQTADQTPGATPKKLS
SWDQAETPGHTPSLRWDETPGRAKGSETPGATPGSKIWDPTPSHTPAGAATPGRGDTPGHATPGHGGATSSARKNRWDETPKTERDTPGHGSGWAETPRTDRGGDSIG
ETPTPGASKRKSRWDETPASQMGGSTPVLTPGKTPIGTPAMNMATPTPGHIMSMTPEQLQAWRWEREIDERNRPLSDEELDAMFPEGYKVLPPPAGYVPIRTPARKLT
ATPTPLGGMTGFHMQTEDRTMKSVNDQPSGNLPFLKPDDIQYFDKLLVDVDESTLSPEEQKERKIMKLLLKIKNGTPPMRKAALRQITDKAREFGAGPLFNQILPLLM
SPTLEDQERHLLVKVIDRILYKLDDLVRPYVHKILVVIEPLLIDEDYYARVEGREIISNLAKAAGLATMISTMRPDIDNMDEYVRNTTARAFAVVASALGIPSLLPFL
KAVCKSKKSWQAQHTGIKIVQQIAILMGCAILPHLRSLVEIIEHGLVDEQQKVRTISALAIAALAEAATPYGIESFDSVLKPLWKGIRQHPGKGLAAFLKAIGYLIPL
MDAEYANYYTREVMLILIREFQSPDEEMKKIVLKVVKQCCGTDGVEANYIKTEILPFFKHFWQHRMALDERNYRQLVDTTVELANKVGAAEIISRIVDDLKDEAEQY
RKMVMETIEKIMGNLGAADIDHKLEEQLIDGILYAFQEQTTEDSVMLNGFGTVVNALGKRVKPYLPQICGTVLWRLNNKSAKVRQQAADLISRTAVVMKTCQEEKLMG
HLGVVLYEYLGEEYPEVLGSILGALKAIVNVIGMPKMTPPIKDLLPRLTPILKNRHEKVQENCIDLVGRIADRGAEYVSAREWMRICFELLELLKAHKKAIRRATVNT
FGYIAKAIGPHDVLATLLNNLKVQERQNRVCTTVAIAIVAETCSPFTVLPALMNEYRVPELNVQNGVLKSLSFLFEYIGEMGKDYIYAVTPLLEDALMDRDLVHRQTA
SAVVQHMSLGVYGFGCEDSLNHLLNYVWPNVFETSPHVIQAVMGALEGLRVAIGPQRMLQYCLQGLFHPARKVRDVYWKIYNSIYIGSQDALIAHYPRIYNDDKNTYI
RYELDYIL

A-helix with P: 13/19
B-helix with P: 2/19

PPP2R1A (589 aa) 15 HEAT repeats
MAAADGDDSLYPIAVLIDELRNEDVQLRLNSIKKLSTIALALGVERTRSELLPFLTDTIYDEDEVLLALAEQLGTFTTLVGGPEYVHCLLPPLESLATVEETVVFDKA
VESLRAISHEHSPSDLEAHFVPLVKRLAGGDWFTSRTSACGLFSVCYPRVSSAVKAELRQYFRNLCSDDTPMVRRAAASKLGEFAKVLELDNVKSEIIPMFSNLASDE
QDSVRLLAVEACVNIAQLLPQEDLEALVMPTLRQAAEDKSWRVRYMVADKFTELQKAVGPEITKTDLVPAFQNLMKDCEAEVRAAASHKVKEFCENLSADCRENVIMS
QILPCIKELVSDANQHVKSALASVIMGLSPILGKDNTIEHLLPLFLAQLKDECPEVRLNIISNLDCVNEVIGIRQLSQSLLPAIVELAEDAKWRVRLAIIEYMPLLAG
QLGVEFFDEKLNSLCMAWLVDHVYAIREAATSNLKKLVEKFGKEWAHATIIPKVLAMSGDPNYLHRMTTLFCINVLSEVCGQDITTKHMLPTVLRMAGDPVANVRFNV
AKSLQKIGPILDNSTLQSEVKPILEKLTQDQDVDVKYFAQEALTVLSLA

A-helix with P: 13/15
B-helix with P: 3/15

PPP2R1B (601 aa) 15 HEAT repeats
MAGASELGTGPGAAGGDGDDSLYPIAVLIDELRNEDVQLRLNSIKKLSTIALALGVERTRSELLPFLTDTIYDEDEVLLALAEQLGNFTGLVGGPDFAHCLLPPLENL
ATVEETVVRDKAVESLRQISQEHTPVALEAYFVPLVKRLASGDWFTSRTSACGLFSVCYPRASNAVKAEIRQQFRSLCSDDTPMVRRAAASKLGEFAKVLELDSVKSE
IVPLFTSLASDEQDSVRLLAVEACVSIAQLLSQDDLETLVMPTLRQAAEDKSWRVRYMVADRFSELQKAMGPKITLNDLIPAFQNLLKDCEAEVRAAAAHKVKELGEN
LPIEDRETIIMNQILPYIKELVSDTNQHVKSALASVIMGLSTILGKENTIEHLLPLFLAQLKDECPDVRLNIISNLDCVNEVIGIRQLSQSLLPAIVELAEDAKWRVR
LAIIEYMPLLAGQLGVEFFDEKLNSLCMAWLVDHVYAIREAATNNLMKLVQKFGTEWAQNTIVPKVLVMANDPNYLHRMTTLFCINALSEACGQEITTKQMLPIVLKM
AGDQVANVRFNVAKSLQKIGPILDTNALQGEVKPVLQKLGQDEDMDVKYFAQEAISVLALA

A-helix with P: 13/15
B-helix with P: 2/15

PPP4R1 (950 aa) 16 HEAT repeats
MADLSLLQEDLQEDADGFGVDDYSSESDVIIIPSALDFVSQDEMLTPLGRLDKYAASENIFNRGMVARSLLDTLREVCDDERDCIAVLERISRLADDSEPTVRAELME
QVPHIALFCQENRPSIPYAFSKFLLPIVVRYLADQNNQVRKTSQAALLALLEQELIERFDVETKVCPVLIELTAPDSNDDVKTEAVAIMCKMAPMVGKDITERLILPR
FCEMCCDCRMFHVRKVCAANFGDICSVVGQQATEEMLLPRFFQLCSDNVWGVRKACAECFMAVSCATCQEIRRTKLSALFINLISDPSRWVRQAAFQSLGPFISTFAN
PSSSGGQYFKEESKSSEEMSVENKNRTRDQEAPEDVQVRPEDTPSDLSVSNSSVILENTMEDHAAEASGKPLGEISVPLDSSLLCTLSSESHQEAASNENDKKPGNYKS
MLRPEVGTTSQDSALLDQELYNSFHFWRTPLPEIDLDIELEQNSGGKPSPEGPEEESEGPVPSSPNITMATRKELEEMIENLEPHIDDPDVKAQVEVLSAALRASSLD
AHEETISIEKRSDLQDELDINELPNCKINQEDSVPLISDAVENMDSTLHYIHSDSDLSNNSSFSPDEERRTKVQDVVPQALLDQYLSMTDPSRAQTVDIEIAKHCAYS
LPGVALTLGRQNWHCLRETYETLASDMQWKVRRTLAFSIHELAVILGDQLTAADLVPIFNGFLKDLDEVRTGVLKHLHDFLKLLHIDKRREYLYQLQEFLVTDNSRNW

RFRAELAEQLILLLELYSPRDVYDYLRPIALNLCADKVSSVRWISYKLVSEMVKKLHAATPPTFGVDLINELVENFGRCPKWSGRQAFVFVCQTVIEDDGLPMDQFAV
HLMPHLLTLANDRVPNVRVLLAKTLRQTLLEKDYFLASASCHQEAVEQTIMALQMDRDSDVKYFASIHPASTKISEDAMSTASSTY

A-helix with P: 11/16
B-helix with P: 7/16

CAND1 (Cullin-associated NEDD-dissociating enzyme) (1230 aa) 27-28 HEAT repeats

MASASYHISNLLEKMTSSDKDFRFMATNDLMTELQKDSIKLDDDSERKVVKMILKLLEDKNGEVQNLAVRCLGPLVSKVKEYQVETIVDTLCTNMLSDKEQLRDISSI
GLKTVIGELPPASSGSALAANVCKKITGRLTSAIAKQEDVSVQLEALDIMADMLSFQGGLLVNFHPSILTCLLPQLTSPRLAVRKRTIIALGHLVMSCGNIVFVDLIE
HLLSELSKNDSMSTTRTYIQCIAAISRQAGHRIGEYLEKIIPLVVKFCNVDDDELPEYCIQAFESFVRRCPKEVYPHVSTIINICLKYLTYDPNYNYDDEDEDENAMD
ADGGDDDDQGSDDEYSDDDDMSWKVRRAAAKCLDAVVSTRHEMLPEFYKTVSPALISRFKEREENVKADVFHAYLSLLKQTRPVQSWLCDPDAMEQGETPLTMLQSQV
PNIVKALHKQMKEKSVKTRQCCFNMLTELVNVLPGALTQHIPVLVPGIIFSLNDKSSSNLKIDALSCLYVILCNHSPQVFHPHVQALVPPVVACVGDPFYKITSEAL
LVTQQLVKVIRPLDQPSSFDATPYIKDLFTCTIKRLKAADIDQEVKERAISCMGQIICNLGDNLGSDLPNTLQIFLERLKNEITRLTTVKALTLIAGSPLKIDLRPVL
GEGVPILASFLRKNQRALKLGTLSALDILIKNYSDSLTAAMIDAVLDELPPLISESDMHVSQMAISFLTTLAKVYPSSLSKISGSILNELIGLVRSPLLQGGALSAML
DFFQALVVTGTNNLGYMDLLRMLTGPVYSQSTALTHKQSYYSIAKCVAALTRAGPKEGPAVVGQFIQDVKNSRSTDSIRLLALLSLGEVGHHIDLSGQLELKSVILEA
FSSPSEEVKSAASYALGSISVGNLPEYLPFVLQEITSQPKRQYLLLHSLKEIISSASVVGLKPYVENIWALLLKHCECAEEGTRNVVAECLGKLTLIDPETLLPRLKG
YLISGSSYARSSVVTAVKFTISDHPQPIDPLLKNCIGDFLKTLEDPDLNVRRVALVTFNSAAHNKPSLIRDLLDTVLPHLYNETKVRKELIREVEMGPFKHTVDDGLD
IRKAAFECMYTLLDSCLDRLDIFEFLNHVEDGLKDEYDIKMLTFLMLVRLSTLCPSAVLQRLDRLVEPLRATCTTKVKANSVKQEFEKQDELKRSAMRAVAALLTIPE
AEKSPLMSEFQSQISSNPELAAIFESIQKDSSSTNLESMDTS

A-helix with P: 20/27
B-helix with P: 1/27

AP3D1 (AP3 complex subunit delta) (1153 aa) 19-20 HEAT repeats

MALKMVKGSIDRMFDKNLQDLVRGIRNHKEDEAKYISQCIDEIKQELKQDNIAVKANAVCKLTYLQMLGYDISWAAFNIIEVMSASKFTFKRIGYLAASQSFHEGTDV
IMLTTNQIRKDLSSPSQYDTGVALTGLSCFVTPDLARDLANDIMTLMSHTKPYIRKKAVLIMYKVFLKYPESLRPAFPRLKEKLEDPDPGVQSAAVNVICELARRNPK
NYLSLAPLFFKLMTSSTNNVVLIKIIKLFGALTPLEPRLGKKLIEPLTNLIHSTSAMSLLYECVNTVIAVLISLSSGMPNHSASIQLCVQKLRILIEDSDQNLKYLGL
LAMSKILKTHPKSVQSHKDLILQCLDDKDESIRLRALDLLYGMVSKKNLMEIVKKLMTHVDKAEGTTYRDELLTKIIDICSQSNYQYITNFEWYISILVELTRLEGTR
HGHLIAAQMLDVAIRVKAIRKFAVSQMSALLDSAHLLASSTQRNGICEVLYAAAWICGEFSEHLQEPHHTLEAMLRPRVTTLPGHIQAVYVQNVVKLYASILQQKEQA
GEAEGAQAVTQLMVDRLPQFVQSADLEVQERASCILQLVKHIQKLQAKDVPVAEEVSALFAGELNPVAPKAQKKVPVPEGLDLDAWINEPLSDSESEDERPRAVFHEE
EQRRPKHRPSEADEEELARRREARKQEQANNPFYIKSSPSPQKRYQDTPGVEHIPVVQIDLSVPLKVPGLPMSDQYVKLEEERRHRQKLEKDKRRKKRKEKEKKGKRR
HSSLPTESDEDIAPAQQVDIVTEEMPENALPSDEDDKDPNDPYRALDIDLDKPLADSEKLPIQKHRNTETSKSPEKDVPMVEKKSKKPKKKEKKHKEKERDKEKKKEK
EKKKSPKPKKKKHRKEKEERTKGKKKSKKQPPGSEEAAGEPVQNGAPEEEQLPPESSYSLLAENSYVKMTCDIRGSLQEDSQVTVAIVLENRSSSILKGMELSVLDSL
NARMARPQGSSVHDGVPVPFQLPPGVSNEAQYVFTIQSIVMAQKLKGTLSFIAKNDEGATHEKLDFRLHFSCSSYLITTPCYSDAFAKLLESGDLSMSSIKVDGIRMS
FQNLLAKICFHHHFSVVERVDSCASMYSRSIQGHHVCLLVKKGENSVSVDGKCSDSTLLSNLLEEMKATLAKC

A-helix with P: 9/20
B- helix with P: 3/20

HEATR2 (855 aa) 16-18 HEAT repeats

MAALGVAEAVAAPHPAEGAETAEAVELSRALSRLLPGLEADSKPGRRRALEALRRALEEPGPAADPTAFQGPWARLLLPRLLRCLSDPAEGCRALAVHLLDLGLRRAA
RPRDALPRLLPALAARLAGPVPARRPPEACEELRLALVQLLGLAVDLCGAALAPHLDDALRALRCSLLDPFAAVRRESCSCAAALAQATPDHFHMQSESLIGPLMQTI
SHQHWKVRVAAIEATGAVIHFGNGKSVDDVLSHFAQRLFDDVPQVRRAVASVVGGWLLCLRDRYSFFHKLIPLLLSSLNDEVPEVRQLAASLWEDVGLQWQKENEEDL
KDKLDFAPPTPPHYPPHERRPVLGCRELVFRNLSKILPALCHDITDWVVGTRVKSAQLLPVLLLHAEDHATQHLEVVLRTLFQACTDEEAAVVQSCTRSAELVGTFVS
PEVFLKLILSTLKKTPSASGLLVLASAMRGCPREALQPHLAAIATELAQAHICQASENDLYLERLLLCVQALVSVCHEDCGVASLQLLDVLLTIVALAGATGLRDKAQ
ETMDSLAMVEGVSSCQDLYRKHIGPLLERVTASHLDWTAHSPELLQFSVIVAQSGPALGEALPHVVPTLRACLQPSQDPQMRLKLFSILSTVLLRATDTINSQGQFPS
YLETVTKDILAPNLQWHAGRTAAAIRTAAVSCLWALTSSEVLSAEQIRDVQETLMPQVLTTLEEDSKMTRLISCRIINTFLKTSGGMTDPEKLIRIYPELLKRLDDVS
NDVRMAAASTLVTWLQCVKGANAKSYYQSSVQYLYRELLVHLDDPEPAIQDAILEVLKEGSGLFPDLLVRETEAVIHKHRSATYCEQLLQHVQAVPATQ

A-helix with P: 10/16
B-helix with P: 2/16

CKAP5 (2032 aa) 32-35 HEAT repeats

MGDDSEWLKLPVDQKCEHKLWKARLSGYEEALKIFQKIKDEKSPEWSKFLGLIKKFVTDSNAVVQLKGLEAALVYVENAHVAGKTTGEVVSGVVSKVFNQPKAKAKEL
GIEICLMYIEIEKGEAVQEFLLKGLDNKNPKIIVACIETLRKALSEFGSKIILLKPIIKVLPKLFESREKAVRDEAKLIAVEIYRWIRDALRPPLQNINSVQLKELEE
EWVKLPTSAPRPTRFLRSQQELEAKLEQQQSAGGDAEGGGDDGDEVPQIDAYELLEAVEILSKLPKDFYDKIEAKKWQERKEALESVEVLIKNPKLEAGDYADLVKAL
KKVVGKDTNVMLVALAAKCLTGLAVGLRKKFGQYAGHVVPTILEKFKEKKPQVVQALQEAIDAIFLTTTLQNISEDVLAVMDNKNPTIKQQTSLFIARSFRHCTASTL
PKSLLKPFCAALLKHINDSAPEVRDAAFEALGTALKVVGEKAVNPFLADVDKLKLDKIKECSEKVELIHGKKAGLAADKKEFKPLPGRTAASGAAGDKDTKDISAPKP
GPLKKAPAAKAGGPPKKGKPAAPGGAGNTGTKNKKGLETKEIVEPELSIEVCEEKASAVLPPTCIQLLDSSNWKERLACMEEFQKAVELMDRTEMPCQALVRMLAKKP

A-helix with P: 20/32
B-helix with P: 5/27

GCN1L1 (2671 aa) 61-64 HEAT repeats

A-helix with P: 39/61
B-helix with P: 18/61

Supplemental references

Bednenko, J., Cingolani, G., and Gerace, L. (2003). Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. J. Cell Biol. *162*, 391-401.

Erickson, H.P. (2009). Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biol. Proced. Online *11*, 32-51.