

Accepted Manuscript

Structural biology and regulation of protein import into the nucleus

Mary Christie, Chiung-Wen Chang, Gergely Róna, Kate M. Smith, Alastair G. Stewart, Agnes A.S. Takeda, Marcos R.M. Fontes, Murray Stewart, Beáta G. Vértessy, Jade K. Forwood, Bostjan Kobe

PII: S0022-2836(15)00616-6
DOI: doi: [10.1016/j.jmb.2015.10.023](https://doi.org/10.1016/j.jmb.2015.10.023)
Reference: YJMBI 64897

To appear in: *Journal of Molecular Biology*

Received date: 17 July 2015
Revised date: 16 October 2015
Accepted date: 24 October 2015



Please cite this article as: Christie, M., Chang, C.-W., Róna, G., Smith, K.M., Stewart, A.G., Takeda, A.A.S., Fontes, M.R.M., Stewart, M., Vértessy, B.G., Forwood, J.K. & Kobe, B., Structural biology and regulation of protein import into the nucleus, *Journal of Molecular Biology* (2015), doi: [10.1016/j.jmb.2015.10.023](https://doi.org/10.1016/j.jmb.2015.10.023)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Structural biology and regulation of protein import into the nucleus

Mary Christie^{1,2*}, Chiung-Wen Chang^{3,4*}, Gergely Róna^{5,6*}, Kate M. Smith^{7*}, Alastair G. Stewart⁸, Agnes A. S. Takeda⁹, Marcos R. M. Fontes⁹, Murray Stewart^{3,10}, Beáta G. Vértessy^{5,6}, Jade K. Forwood⁷, Bostjan Kobe³

¹The Garvan Institute of Medical Research, 384 Victoria St, Darlinghurst 2010, New South Wales, Australia.

²The University of New South Wales, Faculty of Medicine, St Vincent's Clinical School, Darlinghurst, New South Wales 2010, Australia.

³School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Queensland 4072, Australia.

⁴Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA.

⁵Institute of Enzymology, RCNS, Hungarian Academy of Sciences, Budapest, Hungary, H-1117.

⁶Department of Applied Biotechnology and Food Sciences, Budapest University of Technology and Economics, Budapest, Hungary, H-1111.

⁷School of Biomedical Sciences, Charles Sturt University, Wagga Wagga, New South Wales 2650, Australia.

⁸School of Molecular Bioscience, The University of Sydney, Sydney, NSW 2006, Australia

⁹Department of Physics and Biophysics, Institute of Biosciences, Universidade Estadual Paulista, Botucatu, São Paulo, Brazil.

¹⁰MRC Laboratory of Molecular Biology, Francis Crick Ave., Cambridge Biomedical
Campus, Cambridge CB2 0QH, UK.

*These authors contributed equally to this work.

Correspondence to: Bostjan Kobe

Phone: +617 3365 2132

Fax: +617 3365 4699

E-mail: b.kobe@uq.edu.au

ACCEPTED MANUSCRIPT

Abstract

Proteins are translated in the cytoplasm, but many need to access the nucleus to perform their functions. Understanding how these nuclear proteins are transported through the nuclear envelope and how the import processes are regulated is therefore an important aspect of understanding cell function. Structural biology has played a key role in understanding the molecular events during the transport processes and their regulation, including the recognition of nuclear targeting signals by the corresponding receptors. Here, we review the structural basis of the principal nuclear import pathways and the molecular basis of their regulation. The pathways involve transport factors that are members of the β -karyopherin family, which can bind cargo directly (e.g. importin- β , transportin-1, transportin-3, importin-13) or through adaptor proteins (e.g. importin- α , snurportin-1, symportin-1), as well as unrelated transport factors such as Hkeshi, involved in the transport of heat-shock proteins, and NTF2, involved in the transport of RanGDP. Solenoid proteins feature prominently in these pathways. Nuclear transport factors recognize nuclear targeting signals on the cargo proteins, including the classical nuclear localization signals (cNLSs), recognized by the adaptor importin- α , and the PY-NLSs, recognized by transportin-1. Post-translational modifications, in particular phosphorylation, constitute key regulatory mechanisms operating in these pathways.

Highlights

- Structural biology has played a key role in defining the molecular basis of nuclear protein import.
- We review the available structural information on proteins involved in nuclear import pathways.
- We review the available information on the molecular basis of regulation of nuclear import.

Keywords: beta-karyopherin; crystal structure; importin-alpha; nuclear localization signal (NLS); nucleo-cytoplasmic transport.

Abbreviations: IBB, importin- β binding; Imp α , importin- α ; Imp β 1, importin- β ; Imp13, importin-13; β -Kap, β -karyopherin; NLS, nuclear localization signal; Trn1, transportin-1; Trn3, transportin-3.

Overview of nuclear import pathways

The nucleus is separated from the cytoplasm by a double membrane and houses the genetic material and the transcriptional apparatus, separating it from the translational and metabolic machinery in the cytoplasm of eukaryotic cells. The key to this separation is the corresponding ability to regulate transport across the nuclear envelope. This transport occurs through nuclear pore complexes (NPCs), huge macromolecular structures that span both nuclear membranes and which are built from multiple copies of a large number of proteins termed nucleoporins (Nups) [1, 2] (see reviews by Beck, Schwartz, Lemke and Gorlich in this issue). The NPC has a large enough channel to allow proteins smaller than ~40 kDa to passively diffuse through it; however, most, if not all, proteins with functions in the nucleus use active carrier-mediated transport. Translocation of proteins through NPCs requires additional carrier proteins or transport factors. Many of these carriers belong to the β -karyopherin (β -Kap or Imp β) superfamily. All β -Kap family members are built from tandem HEAT repeats (named after huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and TOR1 [3]), each of which contains ~40-45 amino acids that form two anti-parallel α -helices (designated A and B) linked by a loop. This repetitive structure places β -Kaps in the solenoid protein category [4], which features prominently among proteins involved in nucleo-cytoplasmic transport (Fig. 1).

An additional key component of nuclear transport pathways is the small GTPase Ran. Ran cycles between GDP- and GTP-bound states [5], and the state of the bound nucleotide is determined by Ran regulatory proteins, including RanGEF (Ran guanine nucleotide exchange factor, also called RCC1, regulator of chromosome condensation 1, or Prp20p in yeast) [6], and RanGAP (Ran GTPase-activating protein, Rna1p in yeast) [7]. Like other members of the Ras-family GTPases, Ran is comprised of a small G-domain and contains two surface loops, termed switch-I and switch-II, which change conformation depending on the nucleotide

bound state of the protein. The RanGDP/GTP gradient, generated by RanGEF and RanGAP being located in the nucleus and the cytoplasm, respectively, establishes directionality in nucleo-cytoplasmic transport pathways. As a result, import receptors that bind cargo in the cytoplasm can release it in the nucleus through binding RanGTP [5, 8-10], whereas export receptors can bind cargo in the nucleus through simultaneous binding of RanGTP and release it in the cytoplasm after GTP hydrolysis is triggered.

Whether a protein localizes to the nucleus is usually determined by nuclear targeting signals. The first nuclear targeting signal discovered, and the best characterized, is the classical nuclear localization sequence (cNLS) recognized by the protein importin- α (karyopherin- α ; Imp α). Imp α is an adaptor protein that links the cargo to a carrier protein that actually takes the cargo through the NPC; its specific carrier protein is importin- β (karyopherin- β 1; Imp β 1) [11]. Imp α is also a solenoid but is built from armadillo (ARM) repeats [12, 13]. Some proteins also have nuclear export signals (NESs [14, 15]) and can shuttle in and out of the nucleus.

The members of the β -Kap family transport cargoes by either binding the NLS directly, or through adaptor molecules such as Imp α or snurportin-1 (Fig. 1). There are 20 β -Kap family members in humans, ten of which mediate transport of macromolecules into the nucleus, seven translocate macromolecules from the nucleus to the cytoplasm, two have been shown to mediate translocation in both directions, and one member remains to be characterized. The yeast genome codes for 14 β -Kaps. The reason for the large repertoire of nuclear import receptors within cells remains to be fully elucidated; in part it can be attributed to the requirement of cells to translocate hundreds of quite disparate macromolecules across the nuclear envelope. It is emerging that different β -Kap family members recognize different classes of cargoes, and moreover, tissue-specific expression of family members may differentially localize cargoes within different cell types. However, it

also appears that there is some redundancy in this system, with several β -Kap receptors able to recognize the same cargoes [16].

Structural biology has played an important role in deciphering the molecular events required for nuclear import. Here, we review the available structural information on different nuclear import pathways and the associated determinants of specificity in these pathways, and also review the regulatory mechanisms acting on the import processes. The main pathways are summarized in Fig. 1, and the representative protein structures are listed in Table 1. The structures of nuclear export complexes are reviewed by Matsuura in this issue.

Nuclear import mediated by β -karyopherin family members

Overview of the β -karyopherin family

The members of the β -Kap family are highly conserved across eukaryotes, reflecting their critical cellular function. A higher degree of similarity is often observed between orthologues in different species than between paralogues in the same species [17], and although there is low sequence identity among family members (Supplementary Table 1), they display a high degree of structural similarity. Many HEAT repeats in the β -Kap family display conserved Asp and Arg residues at positions 19 and 25 of the repeat, respectively, which form H-bonding ladders [18]. Similarities have been described between HEAT and ARM repeats, particularly for the conserved residues that form the hydrophobic cores [18-20]. The solenoid structure of these proteins (up to 20 repeats within members of the β -Kap family) enables variation in the helicoidal curvature through cumulative subtle changes throughout the protein [4, 21].

Importin- β -mediated nuclear import

Imp β 1 (Kap95p in yeast, termed yImp β 1 here), was the first member of the β -Kap family to be characterized structurally, and has been crystallized in complex with a number of cargo molecules, Ran, and Nups (Fig. 2). The first structure corresponded to Imp β 1 bound to the IBB (importin- β binding) domain of Imp α [22]. Imp β 1 was shown to contain 19 HEAT repeats arranged in a superhelix, forming a convex face (formed by the A-helices of each repeat), and a concave face (formed by the B-helices each repeat). The majority of the Imp β 1 interactions with its binding partners occur on the concave (B-helix) face, and the competition for these sites by Ran, Imp α and various cargo proteins forms the foundation for the mechanisms of assembly, disassembly and translocation. The structural flexibility of the Imp β 1 structure is essential for its function [21].

Apo-importin- β 1

The crystal structure of apo-yImp β 1 revealed that the 19 HEAT repeats are arranged in a tightly coiled and compact conformation [21] (Fig. 2). This tightly wound structure is mediated by HEAT repeats 2 and 4 (residues S⁷⁴ and D¹⁶⁷, respectively), interacting with HEAT repeat 17 residues R⁶⁹⁶, E⁷³⁷, N⁷³⁸, and G⁷³⁹. The interaction interface has a total buried surface area of only 306 Å², which is considerably smaller than most protein:protein interactions, indicating that there is a relatively small energy requirement to distort the flexible yImp β 1. This is supported by small-angle X-ray scattering (SAXS) analysis [21], which indicates that yImp β 1 exists in multiple conformations in solution, and the vast array of functions it performs are achieved by taking advantage of cumulative small structural changes that efficiently allows the transition between various conformations as internal energy is stored by continuous flexing.

Importin- β 1: cargo interactions

A number of cargo proteins are recognized directly by Imp β 1. The structurally characterized examples include the parathyroid hormone-related protein (PTHrP) [23], sterol regulatory element-binding protein 2 (SREBP2) [24], and the zinc finger protein SNAI1 (Snail1) [25] (Fig. 2). The regions on Imp β 1 to which they bind overlap, implying that only one cargo can bind at a time.

Imp β 1 binds PTHrP using B-helices spanning HEAT repeats 2-11 at three distinct binding surfaces. The N-terminal region of the PTHrP NLS (residues 67-79) binds to HEAT repeats 2-7, the central region (residues 80-86) binds HEAT repeats 8-10, and the C-terminal residues (residues 87-93) bind HEAT repeats 8-11. Overall, the PTHrP NLS, when bound to Imp β 1, exists in an extended conformation, and buries 1,027 Å² of surface area.

The crystal structure of the SREBP2 NLS bound to Imp β 1 revealed a helix-loop-helix structure for the cargo (SREBP2 residues 343-403) [24]. To accommodate this binding, Imp β 1 adopts a more open conformation and engages more of the C-terminal HEAT repeats, compared to the PTHrP complex. SREBP2 binds through the B-helices of HEAT repeats 7-17. The two long helices present in repeats 7 and 17 bind SREBP2 rather like chopsticks. A notable difference to PTHrP is that rather than salt bridges dominating the interactions, SREBP2 involves more hydrophobic interactions. This interaction buries 1355 Å² of surface area.

To bind four zinc-finger (ZF) domains of Snail1 (residues 151-264), Imp β 1 uses the B-helices of HEAT repeats 5-14, including the acidic loop in HEAT repeat 8. Unlike other cargo complexes, Imp β 1 adopts a less curved structure to accommodate the bulky snail-like NLS conformation. The ZF1 domain (residues 151-176) acts as the “head”, where the N-terminal α -helix (residues 164-176) is bound within a cleft on Imp β 1 formed from HEAT repeats 9-11. The local conformation is stabilized by a hydrophobic interaction between L¹⁶⁶ of Snail1 and P⁴⁴⁰ of Imp β 1. The Snail1 “shell” is comprised of the three domains ZF2-ZF4,

which form a compact and tight interaction with Imp β 1; the ZF2 domain (residues 177-202) forms hydrogen (H-) bonds with HEAT repeats 13-14; the α -helix within the ZF3 domain (residues 203-230) arranges antiparallel with the inner Imp β 1 α -helix of HEAT repeat 6; and the ZF4 domain (residues 231-264) interacts with residues in HEAT repeats 7 and 9. In the 5 C-terminal residues that represent the tail, C²⁶² and R²⁶⁴ bind to HEAT repeat 10. Overall, Snail1 binds through 15 intermolecular interactions in the B-helices of HEAT repeats 5-14 that bury 2205 Å² of surface area [25].

Although the HEAT repeats used to bind the 3 cargo molecules (5-14 for Snail1, 2-11 for PTHrP, 7-17 for SREBP2) overlap, the binding mechanism in each is distinctly different. The interfaces overlap with the region binding RanGTP, suggesting that regardless of the binding mechanism, the structural requirements include a large contact area and the ability for the complex to be disassembled by RanGTP binding upon entry to the nucleus.

Importin- β 1: adaptor interactions

The Imp β 1 interactions with adaptor molecules Imp α and snurportin-1 have been characterized structurally (Fig. 2; reviewed in [26]). In both cases, Imp β 1 forms a closed conformation, wrapping tightly around the IBB domains of these adaptors, forming an array of salt-bridge interactions through HEAT repeats 7–19. The N-terminal residues of Imp α (α IBB) - residues 11-23 - and snurportin (sIBB) - residues 25-40 – IBB domains mediate interactions with the HEAT repeats 7–11, whereas the C-terminal α -helical regions of α IBB (residues 24-51) and sIBB (residues 41-65) bind through HEAT repeats 12–19. Although there are common HEAT repeats involved in binding adaptor and cargo molecules, the overall mechanism of binding is distinctly different between these two groups. Not only is the overall orientation of the cargo NLSs different (for example, in SREBP the α -helices are bound perpendicular to Imp β 1, whereas the IBB domains bind parallel to the Imp β 1

superhelix) the overall helicoidal twist of Imp β 1 is different, with IBB domain-bound Imp β 1 forming a more closed conformation.

Importin- β 1:Ran interactions

Dissociation of the nuclear import cargo from Imp β 1 following entry into the nucleus is mediated by RanGTP binding [5, 8-10]. The structures for truncated human Imp β 1 [27] and full-length yImp β 1 [28] in complex with RanGTP (Fig. 2) have been determined, revealing important allosteric mechanisms of cargo-release control. The N-terminal fragment of Imp β 1, encompassing HEAT repeats 1-10 (residues 1-462) bound to RanGTP, identified two main contact areas: (1) HEAT repeat 1 interacting with Ran switch-II region residues W⁶⁴, G⁷⁴, and Q⁸², HEAT repeat 2 interacting with Ran residues E⁷⁰, D⁷⁷G⁷⁸, N¹⁰³, and D¹⁰⁷, and HEAT repeat 3 interacting with Ran R¹¹⁰, V¹¹¹, and E¹¹³; and (2) HEAT repeat 7 and the highly conserved acidic loop in HEAT repeat 8 interacting with Ran residues R¹⁴⁰KKNLQYY, K¹⁵⁹, R¹⁶⁶, and P¹⁷²N¹⁷³. The structure of the complex with the full-length yeast orthologue (Kap95p; yImp β 1) revealed an additional interaction site within RanGTP, involving the Ran switch-I loop and the C-terminal arch of yImp β 1, producing a change in curvature that locks yImp β 1 in a conformation incompatible with cargo binding. A sequential binding mechanism has been proposed to occur for RanGTP binding at 3 distinct sites, with residues in the switch-II loop binding to the CRIME motif in N-terminal HEAT repeats 1-4 first, followed by the basic patch on Ran (K¹³⁴, H¹³⁹RK) binding to acidic residues in HEAT repeats 7 and 8, and finally the switch-I loop binding HEAT repeats 12-15 residues R²⁹, K³⁷ forming salt bridges and H-bonds, F³⁵ hydrophobic interactions, and K¹⁵², N¹⁵⁴, N¹⁵⁶, and F¹²⁷ also contributing to the interface. The third site, where the RanGTP switch-I loop binds the C-terminal arch of yImp β 1, is crucial for locking the molecule in a conformation with increased curvature that cannot bind cargo. This allosteric mechanism enables the release of

cargo, because it results in Imp β 1 becoming locked in a conformation that prevents the flexibility required to bind to different partners. The size of the yImp β 1:RanGTP interface (2159 Å², [28]) is similar to the interfaces found in cargo and adaptor complexes, and there is limited overlap between the binding sites. The data therefore strongly support an allosteric mechanism of control for cargo release.

The structure of the yImp β 1:RanGDP complex revealed further insights into Ran binding in the nucleus and cytoplasm [29]. The crystal structure showed that the Ran switch-I and -II regions are induced by yImp β 1 into a GTP-bound conformation, so that rather than these switch regions precluding binding to yImp β 1, they are forced into conformations that enable yImp β 1 binding. This is not inconsistent with other reports of binding partner-induced conformational changes within Ran and Ras switch regions [27, 30]. One region that was shown to be differentially positioned in yImp β 1:RanGDP and yImp β 1:RanGTP structures corresponds to the “basic loop” contained with Ran, comprising residues 133-144, which move 8 Å between the two structures. Significantly, this structure highlights that the yImp β 1 HEAT repeats 7 and 8 are exposed in the RanGDP complex, allowing binding partners such as Imp α to disassemble the complex [29].

Importin- β and RanGTP dissociation and recycling

After translocation to the cytoplasm, recycling of Imp β 1 and other transport factors is achieved by the conversion of Ran to the GDP-bound state. Accessory proteins located at the cytoplasmic face of the NPC facilitate the dissociation of the kinetically stable importin:RanGTP complexes and the generation of RanGDP; the intrinsic GTPase activity of Ran is low and GTP hydrolysis is prevented when Ran is in complex with β -Kap family members [31]. Although the precise mechanism of RanGTP dissociation from transport factors is unclear, one model implicates RanGAP, together with Ran-binding domains

(RanBDs) from either Ran binding protein 1 (RanBP1) or RanBP2 [31]. The structure of Ran in complex with RanBD1 (from RanBP2) and a fragment of Imp β 1 (residues 1-462) has been determined [32] (Fig. 3a), and superimposing it on the Imp β 1:RanGTP complex identifies clashes between the Imp β 1 C-terminal HEAT repeats and RanBD1. This suggests that the primary role of RanBDs is to destabilize the interaction between the transport factors and RanGTP through steric hindrance.

The structure of Ran:GppNHp (a non-hydrolysable GTP analogue) in complex with RanBD1 (of RanBP2) and RanGAP has also been determined [33] (Fig. 3b), and comparison with the Imp β 1:RanGTP complex reveals further clashes between Imp β 1 and RanGAP. *In vitro* radiolabeled-nucleotide assays have demonstrated that RanBP1 does not affect the intrinsic GTPase activity of Ran in isolation. However, RanBP1 has been shown to co-stimulate, with RanGAP, GTP hydrolysis by Ran [34], as well as increase the association rate of RanGTP and RanGAP [35]. This is consistent with the crystal structures of the ternary complexes; RanBP1 forms a molecular embrace with Ran, sequestering Ran's C-terminal region that inhibits RanGAP-mediated GTP hydrolysis. Moreover, the structure of the ternary complex indicated that, unlike other Ras-family GAPs, RanGAP does not provide catalytic residues to stimulate RanGTP hydrolysis [36, 37]. Instead, Ran contains all requisite catalytic machinery and the primary role of RanGAP appears to be the stabilization of the Ran switch-II loop and the repositioning of Ran's catalytic Q⁶⁹ residue towards the active site [33].

In contrast to its GTP-bound state, RanGDP has low affinity for transport factors (\sim 2 μ M for Imp β 1) [29], RanBDs (\sim 10 μ M) [38, 39] and RanGAP (\sim 100 μ M) [35]. GTP hydrolysis therefore precludes rebinding of cytoplasmic RanGDP to transport factors or accessory proteins, rendering dissociation essentially irreversible on the one hand, while enabling the recycling of its binding partners for further rounds of import and disassembly on the other.

Importin-β1:nucleoporin interactions

To mediate translocation across the nuclear envelope, Impβ1 interacts directly with Nups that contain tandem repeats of motifs based on a Phe-Gly core (FG-Nups). The structures of Nup FxFG and GLFG motifs bound to yImpβ1 and Impβ1 show that the interaction sites are primarily located on the convex outer surface in pockets between the A-helices (Fig. 2). HEAT repeats 5-6 bind both the GLFG and FxFG peptides [40, 41]. The high-affinity Nup1p (residues 963-1076) binding sites on yImpβ1 are located between the A-helices of HEAT repeats 5-8. The first and second Nup1p (residues 974-988) binding sites are between HEAT repeats 7-8 and 6-7, respectively, with the third site between HEAT repeats 5-6. Because of the repetitive sequences within the Nup1p C-terminal domain, it is unclear whether the binding involves Nup1p residues 999-1011 or 1019-1031. The interactions within all three sites are predominantly hydrophobic, and involve the Phe aromatic rings, as well as contributions from the adjacent hydrophobic residues (site 1, F⁹⁷⁷ and P⁹⁷⁹; site 2, F⁹⁸⁷ and I⁹⁸⁵; site 3, F¹⁰⁰⁸ and I¹⁰⁰⁷ or F¹⁰²⁷ and I¹⁰²⁶). There is also an additional hydrophobic interaction between P⁹⁸³ and HEAT repeat 7, as well as several H-bonds to the peptide backbone of Nup1p at each site. The total buried surface area of interactions between Impβ1 and other FG Nup cores is ~1,000 Å². However, for the yImpβ1:Nup1p FxFG, the interaction of the buried surface area of the three binding sites is twice that observed for the other interactions, with 2210 Å² buried [41]. Although the higher affinity of Nup1p for yImpβ1 cannot offer a mechanism for translocation, concentrating yImpβ1 at the nuclear face can enhance the kinetics of import complex dissociation and so the overall rate of transport [41].

Bednenko et al [42] proposed that there was a second, weaker, FG binding site on human Impβ1, located between HEAT repeats 14-16 and which involved Leu⁶¹², Phe⁶⁸⁸ and

Leu⁶⁹⁵. Although mutations of this site alone did not impair the binding of a FG peptide, these mutations did impair function in conjunction with mutations in the primary FG binding site. Molecular dynamics calculations [43] indicated that there may be additional FG binding sites on Imp β 1, but this work has not been validated by mutagenesis.

Transportin-1/PY-NLS mediated nuclear import

Transportin-1 (Trn1) imports a broad spectrum of cargoes, many of which are mRNA-binding proteins or transcription factors. Like other β -Kaps, Trn1 is comprised of a series of HEAT repeats arranged as C- and N-terminal arches (reviewed in [16]). Compared to Imp β 1, it contains one additional HEAT repeat and a large 62-residue loop that connects the two helices within HEAT repeat 8 and which appears to be involved in cargo release upon RanGTP binding. Trn1 binds cargoes directly, recognizing them through a broad range of loosely related NLSs that have a characteristic C-terminal Pro-Tyr (PY) motif, which generally has an Arg preceding it by 2-5 residues, giving a consensus of RX₂₋₅PY [44]. The motif in the best-characterized cargo, the splicing factor hnRNP A1, is known as the M9-NLS [45-47]. These motifs lack defined elements of secondary structure, so that they can adopt a conformation to match the binding surface on Trn1. These NLSs generally also contain either a basic or a hydrophobic cluster N-terminal to the PY motif. The hydrophobic PY-NLSs contain two motifs separated by 8–13 amino acids, an N-terminal Φ -G/A/SU motif (Φ represents a hydrophobic amino acid) and a C-terminal sequence R/K/H-X₂₋₅-P-Y. Structures of a range of these NLSs bound to Trn1 show considerable variability in the way they bind. The region in the C-terminal arch of Trn1, to which the N-terminal clusters in the NLSs bind, is rich in negatively-charged residues and also contains a number of scattered clusters of hydrophobic residues, enabling it to accommodate a considerable range of

different NLS sequences [48, 49]. Structure-guided mutations have not always been successful in disrupting these interactions, reflecting the complexity with which Trn1 recognizes its cargoes. Like other β -Kaps, Trn1 binds its cargoes in the cytoplasm, where Ran is primarily in its GDP-bound state, and releases it in the nucleus when RanGTP binds.

Transportin-1:hydrophobic PY-NLS complexes

The available structures include the complexes with the NLSs from the heterogeneous nuclear ribonucleoproteins hnRNP A1 and D, nuclear RNA export factor TAP, hnRNP D-like protein JKTBP and fused in sarcoma protein (FUS) [44, 50, 51] (Fig. 4). In the Trn1:hnRNP A1 crystal structure, the so-called M9-NLS (residues 257-305) binds the C-terminal arch of Trn1, comprising HEAT repeats 8-20. The hydrophobic N-terminal motif residues F²⁷³ and P²⁷⁵ form hydrophobic contacts with Trn1 residues I⁷⁷³ and W⁷³⁰, respectively. Beyond P²⁸⁸Y²⁸⁹, the PY-NLS is disordered. Residues 263-266 in hnRNP A1 interact with a hydrophobic patch in HEAT repeats 18-19 on the outer convex surface, and also with HEAT 20 on the inner concave surface. Residues 267-269 bind the loop in HEAT repeat 18, while the rest of hnRNP A1 follows the inner concave C-terminal arch to contact HEAT repeats 8-17. Isothermal calorimetry analysis of site-directed mutants indicates that the most significant energy contributions to the interaction come from the N-terminal hydrophobic motif [48]. The hnRNP A1 M9-NLS is antiparallel to the Trn1 superhelix, and the interaction buries 3432 Å² of surface area.

Trn1 cargo binding involving other hydrophobic PY-NLSs follows an analogous pattern, with the hydrophobic hnRNP D PY-NLS (residues 332-355) binding to Trn1 HEAT repeats 8-18, whereas the TAP NLS (residues 53-82), binds to HEAT repeats 8-13 (but not HEAT repeats 14-18); and the JKTBP NLS (residues 396-420) binds to HEAT repeats 8-13.

Similar to other PY-NLSs, FUS PY-NLS (residues 498-526) occupies the C-terminal arch of Trn1, but unlike other PY-NLSs that are structurally disordered, the central segment of the FUS peptide (residues 514-522) forms a 2.5-turn α -helix. The FUS PY-NLS interacts with Trn1 at three major sites. The first site involves the N-terminal residues 508-511 of FUS forming hydrophobic interactions with Trn1 residues W⁷³⁰ and I⁷⁷³. Within this N-terminal hydrophobic motif, FUS K⁵¹⁰ makes hydrophobic interactions with Trn1 residues W⁷³⁰ and salt bridges with Trn1 E⁶⁵³ and D⁶⁹³. In the second “central” site, FUS residues 514-522, arranged as an α -helix, interact with Trn1 HEAT repeats 9-12 at residues D⁵⁰⁹, D⁵⁴³, D⁵⁵⁰, E⁵⁸⁸, and D⁶⁴⁶, forming H-bonds and salt-bridge interactions with all 5 of its basic residues (R⁵¹⁴, H⁵¹⁷, R⁵¹⁸, R⁵²¹ and R⁵²²). In the third site, the C-terminal residues P⁵²⁵Y⁵²⁶ interact primarily hydrophobically with Trn1 residues A³⁸¹, L⁴¹⁹, I⁴⁵⁷, and W⁶⁴⁰. Overall, the FUS PY-NLS binds Trn1 through hydrophobic interactions at both the N- and C-terminus of the peptide and through electrostatic interactions in its central α -helix. Recently, FUS binding to Trn1 was shown to also occur within the adjacent unmethylated RGG3 domain (residues 472-507), suggesting it could act independently as a Trn1-dependent NLS, or as an accessory domain capable of extending the PY-NLS [52].

Overall, the C-terminal PY motif and the N-terminal hydrophobic motif in these PY-NLSs are recognised by Trn1 HEAT repeats 8-13 and HEAT repeats 14-18, respectively.

Transportin-1:basic PY-NLS complexes

In the structure of Trn1 in complex with yNab2 (residues 205-242) [51] (Fig. 4), Nab2 residues T²³⁴RFNPL²⁴⁰ bind Trn1 in an extended conformation, occupying the same binding site observed by the RX₂₋₅PY motifs in other Trn1:PY-NLS structures. The Nab2 PY-NLS structure contains a “PL” instead of the canonical “PY”. The N-terminal R²³⁵ makes salt-bridge and H-bond interactions with D⁵⁴³, T⁵⁰⁶, E⁵⁰⁹, and T⁵⁴⁷ of Trn1. Additionally, F²³⁶

makes hydrophobic contacts with Trn1 residues A⁴⁹⁹, E⁴⁹⁸ and W⁴⁶⁰; P²³⁸ interacts predominantly through hydrophobic interactions with Trn1 residues L⁴¹⁹, I⁴⁵⁷, W⁴⁶⁰; and L²³⁹ interacts hydrophobically with Trn1 residues L⁴¹⁹, A³⁸¹, A⁴²², and W⁴⁶⁰. The Tyr aromatic ring can make more hydrophobic or polar interactions, compared to the Leu in the PL motif, resulting in a higher energetic contribution.

The structure of the Trn1:HCC1 (hepato cellular carcinoma protein 1) complex (PDB 4OO6) has been deposited, however a detailed description of the structure has not yet been published. We provide a brief comparison with other similar NLSs. HCC1 contains a basic PY-NLS and analysis of the interface with Trn1 reveal residues R⁹²GRYRSPY⁹⁹ that bind HEAT repeats 8-14, with H-bonds formed between HCC1 Y⁹⁵, R⁹⁶ and Y⁹⁹, and Trn1 residues D³⁸⁴, A⁴²³, S⁵⁰², T⁵⁰⁶.

The basic PY-NLS of hnRNP M binds Trn1 HEAT repeats 8-16 [48]. Unlike hnRNP A1 that binds the convex side of the N-terminal Trn1, the N-terminus of hnRNP M binds toward the Trn1 arch opening. In the N-terminal basic motif of hnRNP M, residues E⁵¹KNI⁵⁴ bind the same region of Trn1 as hnRNP A1 residues 274-277. However, in contrast to hnRNP A1, hnRNP M is ordered beyond the PY motif with 5 residues extending C-terminally. Hydrophobic interactions involve aliphatic portions of hnRNP M residues K⁵² and I⁵⁴ and Trn1 residues W⁷³⁰, I⁶⁴², D⁶⁴⁶, and Q⁶⁸⁵. Surprisingly, unlike hnRNP A1, the most significant energy contributions come from the C-terminal PY domain, rather than the N-terminal motif.

Overall, although Trn1 uses the same residues (HEAT repeats 8-13) to bind the C-terminal PY-NLS motifs in all PY-NLSs, the N-terminal regions of basic and hydrophobic PY-NLSs only partially overlap.

Transportin-1:Ran complex

Similar to other β -Kap family members, Ran binding to Trn1 in the nucleus displaces the imported cargo. Trn1 has a characteristic acidic H8 loop, and the proposed NLS dissociation mechanism involves RanGTP binding to the Trn1 N-terminal arch. This causes conformational changes that push the H8 loop into the principal cargo-binding site in the C-terminal arch of Trn1, causing cargo release [53]. The structure of Trn1:RanGTP complex (Fig. 4) [54] reveals two distinct binding interfaces, an N-terminal interface involving interactions between Trn1 HEAT repeats 1-4 and the switch regions of Ran (residues 64-110); and a more centrally located interface involving HEAT repeats 7-8, 14-15, and loop-8 (residues 311-373) of Trn1 binding to Ran α -helices α 4 and α 5, β -strand β 6, and the intervening loops. In the N-terminal interface, the C-terminal region of Ran switch-I (residues 44-47) interacts with Trn1 HEAT repeat 1; the switch-II region of Ran (residues 72-82) is buried at the interface by hydrophobic contacts to Trn1 HEAT repeat 2. Overall, these interactions mediate a combined buried surface area of 3900 \AA^2 and involve both polar and hydrophobic contacts, with most of the polar contacts contributed by Ran R¹⁰⁶ and R¹¹⁰ and Trn1 S¹⁶⁵, D¹⁶⁴, E¹⁶¹.

Transportin-3/RS-repeat NLS-mediated nuclear import

Transportin-3 (Trn3), comprised of 20 HEAT repeats, mediates the nuclear import of many proteins containing arginine-serine (“RS”) repeat NLSs [55-57]. These proteins are typically involved in mRNA metabolism, and include the alternative splicing factor/splicing factor 2 (ASF/SF2). Structures of Trn3 bound to ASF/SF2 and Ran, and its unbound form, are available [58] (Fig. 5). The flexibility of the HEAT repeat region is important for binding of the RS domains as well as the RNA recognition motif in ASF/SF2. The significant overlap in binding regions between RanGTP and the cargo explains the structural basis of release.

Transportin-3:ASF/SF2 cargo complex

The ASF/SF2 protein is an RNA-splicing factor that contains 2 RRM domains and an RS domain. Residues within HEAT repeat 15, particularly Arg-rich regions within B-helices, are the key binding determinants for the recognition of the phosphorylated RS domain of ASF/SF2 [58]. There are three major binding regions, which bury a total of ~4,300 Å² of surface area. The RRM domain (residues 116-191) binds HEAT repeats 4-7 and 19-20, the RS regions (residues 198-211) are bound by HEAT repeat domains 14-17, and the linker region between the RRM and RS domains is bound by HEAT repeat domains 12-13 [58]. These interactions are mediated by an extensive array of salt bridges. Although the RRM domain exhibited 60% of the buried surface area, the RS domain is actually the major contributor to binding; this is based on (i) mutagenesis studies; (ii) the fact that many RS domain-containing proteins that interact with Trn-3 do not contain RRM domains; and (iii) the RS-domain NLS being necessary and sufficient to mediate nuclear transport [58].

Transportin-3:Ran complex

Similar to other β-Kap family members bound to Ran, the structure of Trn3 in complex with RanGTP shows that Ran contacts the B-helices on the concave side of the transport receptor. Sites that mediate binding within Ran include the switch-I and switch-II regions, which interact with the Trn3 HEAT repeats 1-3 and 17-18, respectively [58]. In particular, the switch-I region of Ran inhibits the ability of HEAT repeat 15 to interact with the RS domain.

Apo-transportin-3

The 20 HEAT repeats in apo-Trn3 are arranged in a circular shape, whereby the N- and C-terminal repeats face each other. In most β-Kap family members, HEAT repeats pack

in a rather uniform manner, but in Trn3 there are several notable exceptions: HEAT repeats 1 and 2 pack perpendicular to each other; the stacking of HEAT repeats 3/4 and 9/10 displays pronounced left-handed twists; and HEAT repeat 20 contains an additional C-terminal α -helix. Interestingly, the crystal structure revealed a significant molecular interface mediating homo-dimer formation, and SAXS analysis is consistent with this observation; however, the functional role of dimerization is unclear at this stage [58].

Importin-13-mediated nuclear transport

Importin-13 (Imp13) is the closest paralogue of Trn3, but has distinctly different cargo recognition specificity [59]. Whereas Trn3 predominantly interacts with flexible RS domains, Imp13 mediates the nuclear import of several transcription factors containing histone-fold motifs (comprised of ~70 amino acids arranged as three α -helices). Similar to Trn1, Imp13 contains 20 consecutive HEAT repeats [60] and interactions with cargo occur on the inner concave surface; however, Imp13 is able to mediate transport of cargoes both into and out of the nucleus. Recent co-crystal structures of Imp13 with the exon-junction-complex components Mago and Y14, and the E2 SUMO-conjugating enzyme UBC9, show that the flexibility of Imp13 is important for cargo binding (Fig. 6).

Importin-13:Mago-Y14 cargo complex

The first structure of Imp13 corresponds to the complex with Mago-Y14, revealing that 15 HEAT repeats are involved in binding the cargo [60]. Imp13 adopts a closed ring-like conformation, whereby the N- and C-terminal arches are facing each other, and Mago-Y14 bind to the inner concave-surface helices of the C-terminal arch. HEAT repeats 8 and 9 interact with the Mago β -sheet, at the site where the N terminal region of Y14 is bound, and

HEAT repeat 15 binds at the opposite side of the Mago β -sheet. HEAT repeats 17, 18, and 20 interact with Mago α -helices, and HEAT repeats 4-7 and 19-20 surround Y14.

Importin-13:UBC9 cargo complex

The crystal structure of the Imp13:UBC9 cargo complex showed a unique cargo recognition mode, with UBC9 bound within the N-terminal arch of Imp13 to occupy the RanGTP-binding site in that region [61]. Unlike the Imp13:Mago-Y14 complex, the N- and C-terminal HEAT repeats of Imp13 are positioned away from each other. UBC9 lies between HEAT repeats 1-9, and forms interactions within the inner concave surface of Imp13. UBC9 makes interactions mainly through three of its loops. The first interacting region involves hydrophobic interactions mediated by a loop and a helix that binds both helices within HEAT repeat 1, and the B helix of HEAT repeat 2 of Imp13. I¹²⁵ of UBC9 makes contact with Imp13 residues Y³⁴, E⁷³ and Y⁷⁶, while an additional hydrophobic interaction involves UBC9, Y¹³⁴, positioned towards Imp13 L³³ and Y³⁴.

Importin-13:Ran complex

Similar to Imp β 1 and Trn1, cargo release and directionality of nuclear import of Imp13 are achieved by Ran. However, cargo release of Imp β 1 and Trn1 relies on the characteristic acidic loop within HEAT repeat 8, which is lacking in Imp13, and therefore the mechanism of cargo release is likely to be different in these transport molecules. The Imp13:RanGTP structure shows RanGTP interacting with the inner concave helices contained within the N-terminal arch of Imp13 at 3 sites, similar in position to those identified in Imp β 1, Trn1 and Trn3 [62]. The Ran switch-I loop binds Imp13 at HEAT repeats 16-19 with predominantly polar and electrostatic contacts (eg Ran K³⁹K⁴⁰ binding Imp13 D⁷⁸⁵/D⁷⁸⁸) [62]. The Ran switch-II loop binds Imp13 HEAT repeats 1-3, with

hydrophobic interactions involving Ran L⁷⁷ and electrostatic contacts between Ran D⁷⁹ and Imp13 R¹²². The helix adjacent to the switch-II loop also contacts HEAT repeats 3 and 4, in particular Ran residues R¹⁰⁸ and R¹¹² contacting Imp13 residues E¹⁷⁵ and E¹⁷⁶ [62]. The third binding site of RanGTP is through Imp13 HEAT repeats 8-9, with Ran residues R¹⁶⁸K¹⁶⁹ approaching negatively charged Imp13 residues D⁴¹⁵E⁴¹⁶ on helix 9B. Thus, unlike the mechanism of cargo release by Imp β 1, Imp13 releases cargo due to its direct steric clashes with RanGTP.

Importin- α mediated nuclear import

In this pathway, the Imp α :Imp β 1 hetero-dimer binds to cargo proteins containing cNLSs [63]. The translocation through the nuclear pore is achieved through transient interactions between Imp β 1 and Nups. This process, known as the classical nuclear import pathway, is thought to be the most extensively used nuclear import mechanism in the cell [64-66].

Monopartite and bipartite cNLSs

The first nuclear targeting motif was identified in the simian virus 40 (SV40) large T-antigen (TAg) through mutational studies. It comprises a small stretch of positively charged amino-acid residues (P¹²⁶KKKRRV¹³²). Non-conservative substitutions of residues within this motif abrogated nuclear distribution of the cargo protein [67]. Furthermore, fusion of this motif to cytoplasmic proteins such as β -galactosidase induced their nuclear accumulation [68]. A similar, but more complex signal was later defined for the *X. laevis* nucleoplasmin protein, consisting of two clusters of basic amino acids separated by a 10-12 residue linker region (K¹⁵⁵RPAATKKAGQAKKKK¹⁷⁰) [69]. Substitution of residues within either basic cluster altered the nuclear distribution of the protein, suggesting that both motifs were required for

nuclear targeting. By contrast, mutation of residues within the linker region had no effect on nuclear distribution [70].

The two sequences are now commonly described as the prototypic monopartite (SV40-TAg) and bipartite (nucleoplasmin) cNLSs and numerous cNLS-containing cargo proteins have since been identified based on sequence similarity with these two [65] (Supplementary Table 2). Using *in vitro* transport assays in digitonin-permeabilized cells, the active import of cNLS sequences was shown to be dependent on soluble cytoplasmic factors [71]. This *in vitro* system was later used to identify and characterize essential transport factors, Imp α and Imp β 1, that could reconstitute nuclear import of cNLS sequences when reintroduced to cytosol depleted cells [63, 72-76]. Consequently, cNLS cargoes are defined by the presence of one or two sequence clusters rich in Arg and Lys that are necessary and sufficient for nuclear import by the Imp α :Imp β 1 complex.

Structure of importin- α

Imp α has a modular structure comprised of a short N-terminal auto-inhibitory region that also mediates binding to Imp β 1 [77] (the IBB domain) and larger C-terminal NLS-binding domain comprised of 10 ARM repeats [78, 79]. The ARM-repeat motif, first described for the *Drosophila melanogaster* armadillo protein [80], is comprised of three α -helices (H1, H2 and H3). The continuous stacking of the tandem ARM repeats generates a superhelical solenoid, with the H3 helices forming the Imp α inner concave surface (Fig. 7a). A rotation between consecutive ARM repeats creates a groove along the superhelical axis of the protein, where the NLS-binding sites are located [11]. The ARM repeat solenoid appears to be much less flexible than the HEAT repeat solenoids in β -Kaps. The structures of Imp α proteins from different organisms have been determined (human [81-85], *S. cerevisiae* [13, 79, 86], mouse [12, 87-102], rice [100, 103], *Arabidopsis thaliana* [104], *Neurospora crassa*

[105]; Supplementary Table 3). These structures all comprise 10 ARM repeats, but their curvatures vary, particularly between proteins from different phylogenetic families [128]. The structural variations result in differences in binding NLSs, as observed for rice and mouse Imp α bound to the same NLS peptide, for example [66].

Structural basis of cNLS recognition by importin- α

X-ray crystallography has been used extensively to elucidate the molecular details of cNLS binding to Imp α . The concave surface formed by Imp α H3 helices comprises the cNLS-binding site, which displays a high degree of sequence conservation between Imp α proteins from various organisms (Fig. 7b). More specifically, conserved (^R/K)XXWXXXN motifs (where x is any residue, and ^R/K is any residue other than Arg/Lys, typically an acidic or hydrophilic residue) within the H3 helices form an array of binding pockets along the inner concave groove of the Imp α adaptor. The conserved Asn residues form H-bonds with the cNLS backbone, whereas the invariant Trp side-chains form an array of binding cavities on the adaptor surface, typically with acidic residues (^R/K) located at the end of these pockets. Thus, the aliphatic moieties of long basic side-chains, such as Lys and Arg, can interact with the stacked indole rings of the Trp array, while the positively charged portion of the side-chain can simultaneously form H-bonds and salt bridges with the hydrophilic residues that line the pockets, and also form cation- π interactions with the electron clouds of the tryptophan indoles.

Disruption of the (^R/K)XXWXXXN motif within ARM repeats 5 and 6 has been observed in all Imp α proteins with known structure. This disruption effectively creates and segregates two distinct binding regions on the Imp α surface termed the major (ARM repeats 2-4) and minor (ARM repeats 6-8) binding sites (Fig. 7). Monopartite and bipartite cNLSs bind to the Imp α binding sites in an extended conformation (Fig. 7). Structural analyses have

demonstrated that monopartite cNLSs preferentially bind to the major binding site, reflected through lower crystallographic B-factors and the presence of more extensive electron density when compared to that observed at the minor binding site [79, 87, 97]. In addition, substitution of residues within the major binding site can abrogate nuclear accumulation of monopartite cNLS cargoes, whereas minor binding-site mutations have minimal effects [106]. By contrast, mutation of residues at either minor or major binding sites can severely disrupt the interaction with bipartite cNLSs [106], which interact simultaneously with the two binding regions on the Imp α surface.

The major binding site of Imp α is comprised of four principal binding cavities that bind the side-chains of cNLS residues P2-P5 (Fig. 7). Structures of numerous cNLSs bound to Imp α have been determined (Supplementary Table 2). For all characterized monopartite and bipartite sequences, the most crucial structural determinant is a Lys residue located at position P2. Its side-chain forms a salt bridge with a highly conserved Imp α Asp side-chain (Fig. 7). Consistently, substitution of this conserved Asp residue results in an ~300-400 fold decrease for monopartite and bipartite cNLS binding in γ Imp α [106]. Although an Arg side-chain can bind at the P2 position (PDB entry 4HTV; Supplementary Table 2), mutational studies on the SV40-TAg cNLS have demonstrated that a Lys is energetically favored at this position [107, 108].

Although preference for long basic side-chains is observed at the other major binding site positions (P3, P4, P5), cNLSs can have a range of different amino-acids in these positions (Supplementary Table 2). The calculated free-energy contributions of the SV40-TAg basic side-chains at the P3-P5 positions are between 1/4 to 2/3 of that observed for the P2 Lys residue, with the P4 position contributing the least free energy to the interaction with Imp α [107]. This suggests that non-basic side-chains are not strictly necessary at the P3, P4 and P5 positions, provided that the overall affinity of the cNLS cluster is sufficient to

constitute a functional cNLS motif. This is achieved by maximizing interactions either at the other major binding site pockets, the regions directly flanking the major binding site, or binding at the minor binding site in the case of bipartite cNLSs.

The two key amino-acids in the N-terminal region of bipartite cNLS (positions P1'-P2') bind in a conserved manner to the minor NLS-binding site, but adjacent auxiliary cavities can be used differentially in a cNLS-specific manner (Supplementary Table 2). In bipartite cNLSs, a "KR" motif is observed predominantly at these positions (Supplementary Table 2), with the P2' Arg side-chain forming a salt bridge with a conserved Imp α Glu residue. The total energetic contribution of the P1' and P2' pockets has been calculated to be 3.2 kcal/mol, comparable to that observed for a basic residue at the P3 or P5 position [107]. Although this interaction is modest, the addition of a KR motif N-terminal to a non-functional SV40-TAg variant, whereby the critical P2 Lys residue was replaced with a Thr, was sufficient to direct nuclear accumulation of the protein [109]. Thus, compared to monopartite motifs, the sequence requirements at the major binding site are not as strict in bipartite cNLSs, due to the additional interactions at the minor binding site, as well as cNLS-specific linker region interactions.

Structural studies showed that a minimum of ten residues between the P2' and P2 positions is required to allow functional cNLSs to interact simultaneously with both the major and minor binding sites on the Imp α surface [56]. Furthermore, early localization assays demonstrated that the linker region could tolerate non-conservative substitutions, as well as insertions [67], and these observations are consistent with the minimal interactions observed between bipartite cNLS linker regions and the Imp α surface in crystal structures. Consistently, bipartite cNLS linker region residues are not well ordered and have higher crystallographic B-factors than residues at the major and minor binding sites. In some cNLSs, electron density is absent for most linker-region residues (Supplementary Table 2, residues in italics indicate

residues not visible in crystallographic models), suggesting that bipartite linker regions longer than 12 residues likely bulge away from the Imp α surface, but are functional, provided that sufficient contacts at the major and minor binding regions are maintained. Notably, peptide library studies have demonstrated a preference for acidic residues in bipartite linker region sequences [110], and structural studies have shown that negatively charged side-chains can form electrostatic interactions with the basic surface of Imp α ARM repeats 4-6 [90].

Taken together, structural and biochemical data have revealed the molecular determinants of cNLS binding to the Imp α adaptor. These studies have therefore enabled the elucidation of consensus sequences of both types of cNLSs. The monopartite cNLS motif is defined as **K(K/R)X(K/R)**, whereas the bipartite cNLS consensus sequences correspond to KRX₁₀₋₁₂**KRRK**, KRX₁₀₋₁₂**K(KR)(KR)** and KRX₁₀₋₁₂**K(K/R)X(K/R)** (where X corresponds to any residue, Lys residues in bold indicate the critical P2 lysine, and minor site-binding KR motifs are underlined) [65, 90].

Atypical importin- α -dependent NLSs

Several nuclear targeting signals are dissimilar to cNLSs described above, but are nevertheless recognized by Imp α . An analysis of binding of a random peptide library to Imp α variants revealed six classes of NLSs, including two types of non-classical NLSs, which were annotated ‘plant-specific’ (consensus sequence LGKR[K/R][W/F/Y]) and ‘minor site-specific’ (consensus sequences KRX[W/F/Y]XXAF and [R/P]XXKR[K/R][\wedge DE]) NLSs [111]. Both of these types of NLSs feature a short basic cluster flanked C-terminally by hydrophobic residues.

Unique features of the binding of atypical NLSs to Imp α have been identified by structural and biochemical studies. Crystal structures of mouse Imp α in complex with poorly

basic NLSs (G²⁵⁷KISKHWTGI²⁶⁶; G²⁷³SIIRKWN²⁸⁰) from human phospholipid scramblase isoform 1 and 4 (hPLSCR1/4) show binding to the major and minor NLS binding sites, respectively [95, 99]. The exclusive binding to the minor NLS-binding site is also observed in naturally-occurring NLSs (e.g. the mouse RNA helicase II (Gu α) NLS (K⁸⁴²RSFSKAF⁸⁴⁹) [101]), whereas as the mitotic regulator protein TPX2 NLS (K²⁸⁴RKH²⁸⁷) binds predominantly to the minor site, but could be considered an atypical bipartite NLS (with K³²⁷MIK³³⁰ binding to the major site) [91]. Unlike other NLSs that bind in an extended conformation, structural analysis of the Gu α NLS and four other ‘minor site-specific’ NLSs in complex with mouse Imp α revealed that the C-terminal residues of these NLSs form an α -helical turn [101]. This distinct structure of the NLS is stabilized by internal H-bond and cation- π interactions between the aromatic residues from the NLSs and the positively charged residues from Imp α . Such a conformation is prevented sterically at the Imp α major binding site, explaining the minor site preference of these motifs [101]. Although contacts between ‘minor site-specific’ NLSs are observed at the major binding site (Supplementary Table 2), the NLS peptides at the minor binding region have more extensive interactions and lower crystallographic B-factors.

Synthetic peptides corresponding to ‘plant-specific’ NLSs show preferential binding to the minor NLS-binding site of rice Imp α , although the structural determinants of their binding mode are different from the ones observed for other ‘minor site-specific’ NLSs [100]. Although putative naturally occurring ‘plant-specific’ NLSs can be found using sequence analyses [100], they have not yet been characterized experimentally. Additional atypical NLSs have been identified that have not been characterized structurally, for example in Borna disease virus (BDV) P10 protein (R⁶LTLLELVRRRLNGN¹⁹) [112]. Bioinformatic analyses of the distribution of different classes of NLSs in diverse eukaryotes indicate that the atypical NLSs are much less prevalent than the monopartite and bipartite cNLSs [66].

Structures of importin- α in complex with native proteins

Recognition of NLSs by Imp α requires that these linear sequence motifs adopt an extended conformation. Consistently, cNLSs are located in disordered (and thus flexible) regions of native proteins, and so structural characterization of the Imp α -cNLS interaction has predominantly involved the use of peptide sequences that correspond to NLS segments. Structural analyses of Imp α in complex with native cNLS-containing proteins or domains have only been described for the influenza virus PB2 C-terminal fragment (residues 628-759) [83, 84], the human cap-binding protein 80 (CBP80; in complex with CBC20) [81] and the Ebola virus VP24 protein [113]. The PB2 C-terminal fragment has been crystallized in complex with 4 different human Imp α isoforms (Imp α 5 [83] and Imp α 1, Imp α 3, Imp α 7 [84]). In all these structures, a PB2 globular domain Lys residue located outside the canonical bipartite cNLS sequence interacts *in trans* with residues that comprise the Imp α P3' pocket. In the absence of the globular domain, the PB2 bipartite cNLS is able to interact with the P3'-binding pocket, causing a register shift in the cNLS minor binding site cavities (Supplementary Table 2) [84]. Differences in binding registers have also been described for structures of the SV40-TAg cNLS peptide (Supplementary Table 2), suggesting that these linear motifs can differentially bind to the Imp α surface to maximize favorable interactions.

In contrast to the NLSs described above, the Ebola virus VP24 protein interacts with Imp α through a distinctly different mechanism. The crystal structure of truncated human Imp α (ARM repeats 7-10) in complex with VP24 shows that the virus protein primarily contacts the extreme C-terminus of the adaptor through three interspersed clusters on the surface of the folded VP24 structure [113]. Moreover, VP24 interacts with the opposite surface of Imp α compared to NLS peptide sequences, with the H2 helices of ARM repeats 9 and 10 defining the VP24 interface [113]. Although the binding surfaces of NLS-containing

cargo and VP24 do not overlap, a two-fold difference in binding is observed between the nucleoplasmin cNLS and Imp α in the presence of VP24. This suggests that minor binding site interactions of cNLS sequences are allosterically affected by VP24 interaction with the outer Imp α H2 helices [113].

Although bioinformatics analyses have suggested that a moderate proportion of yImp α -binding proteins lack a detectable linear NLS [64], nuclear localization may also occur through “piggy-back” mechanisms, whereby translocation of a non-NLS containing protein by Imp α is mediated via interaction with an NLS-containing binding partner. The prevalence of proteins that mediate Imp α binding through non-canonical means such as the VP24 protein is not known and cannot be identified by current cNLS prediction algorithms.

Auto-inhibition by the importin- α IBB domain

The N-terminal IBB domain of Imp α contains a cNLS-like sequence. Similar to cNLS motifs, the Imp α IBB domain is rich in basic amino acids and interacts with the NLS binding pockets in the absence of cargo. In the mouse Imp α structure, the IBB domain residues K⁴⁹RRN⁵² are bound to the major NLS-binding site (and correspond to cNLS positions P2-P5) [12]. This is consistent with lower affinity of cNLS binding to full-length Imp α , compared to truncated Imp α proteins that lack the IBB domain [114]. Likewise, increased affinity for cNLSs was observed when the corresponding K⁵⁴RR⁵⁶ motif in the IBB domain of yeast Imp α was substituted with alanine residues [115]. This suggests that the IBB domain inhibits cNLS binding. In rice Imp α , the K⁴⁷KRR⁵⁰ motif in the IBB domain was found to form analogous interactions with the major NLS-binding site [100]. Distinct from the mouse structure however, the G²⁵RRRR²⁹ motif in the rice Imp α IBB domain additionally interacts with the minor NLS-binding site. Similar to rice Imp α , minor and major binding site interactions are observed between the IBB domain of the yeast protein and the solenoid

domain when in complex with Cse1p and RanGTP [116]. The auto-inhibitory mechanism in plant and yeast Imp α proteins may therefore differ from the mammalian proteins.

The IBB domain of Imp α also mediates binding to Imp β 1 and so the interaction with Imp β 1 promotes cNLS binding to Imp α . The IBB domain-mediated auto-inhibitory mechanism presumably reduces futile import of empty adaptors, and hinders cNLS binding when Imp β 1 is not present for nuclear translocation. Most structural studies of Imp α binding to NLSs have therefore employed a truncated protein lacking the IBB domain.

IBB domain-like NLSs

A recent report describing the structures of the Heh1 and Heh2 inner membrane protein NLSs in complex with yeast Imp α have revealed a bipartite mode of binding distinct from that observed in cNLSs [117]. Structural and biochemical analyses of these NLSs suggest similarities to the IBB domain interaction with the Imp α adaptor: (1) the conformation of Heh2 NLS is similar to that observed for the IBB domain of full-length yImp α in the auto-inhibited state (when in complex with RanGTP and Cse1p [116]); (2) mutational analyses have identified that the key structural determinant for these NLS motifs is the P2' pocket; this is in contrast to bipartite cNLS motifs, where nuclear accumulation is only modestly abrogated by disruption of P2' interaction [70, 106]; and (3) pull-down assays demonstrate that the Heh1 and Heh2 NLS can efficiently relieve IBB domain auto-inhibition of Imp α in the absence of Imp β 1, unlike typical bipartite cNLS sequences [117]. However, in contrast to the Imp α IBB domain, direct binding to Imp β 1 was not detected for the Heh1 and Heh2 NLSs [117]. Similar to the atypical monopartite NLSs, Heh1 and Heh2 mediate extensive contacts at the Imp α minor binding site, with residues in this region having lower average B-factors compared to residues bound at the major binding pockets.

Importin- α variants

In some organisms, several Imp α variants exist as a result of duplication events. The metazoan paralogues can be divided into three clades ($\alpha 1$, $\alpha 2$ and $\alpha 3$); Imp α proteins from Viridiplantae and Fungi belong to the $\alpha 1$ -like clade [118-120]. The members of different subfamilies share ~50% sequence identity, whereas within a subfamily, the identities are >80% [118].

The structure and recognition mechanism are highly conserved among Imp α proteins from different species [90, 105]. However, Imp α variants can display preferences for specific NLSs, which may be important for development and tissue-specific roles [118]. In *D. melanogaster*, which encodes three Imp α proteins ($\alpha 1$, $\alpha 2$ and $\alpha 3$), oogenesis depends on Imp $\alpha 2$, and neither Imp $\alpha 1$ nor Imp $\alpha 3$ can substitute [121]. The mouse genome codes for six variants and Imp $\alpha 7$ plays an essential role during the early stages of embryo development [122]. Some of the seven Imp α variants in humans also display preferential interactions with specific cargoes, for example Imp $\alpha 3$ with RCC1 [123] and Imp $\alpha 5$ for STAT proteins [124, 125].

The overall structure and the key NLS-binding features of the different Imp α variants are conserved in the crystal structures determined to date (Supplementary Table 3, Fig. 7). Therefore, the reasons for specific NLS-binding preferences by certain variants are not entirely clear, but may relate to amino-acid differences in the vicinity of NLS-binding sites affecting NLS binding and auto-inhibition.

Structural analyses showed that plant-specific NLSs bind specifically to the minor NLS-binding site of rice Imp α , but preferentially to the major site of mouse Imp α [100], which has been attributed to specific amino-acid differences in the C-terminal region of Imp α . In particular, a Thr-to-Ser mutation prevents the binding of a plant-specific NLS peptide to the minor site of mouse Imp α through steric hindrance. Indeed, it has been shown that SV40-

TAg NLS-binding mode in the minor NLS site is not the same among Imp α structures [105] (Supplementary Table 2). The structures of mouse and yeast Imp α in complex with the SV40-TAg NLS peptide show “KK” residues at P1'-P2' positions, whereas rice and *N. crassa* Imp α structures show “KR” residues at these positions. The minor binding site may play a more important role in the α 1-like Imp α family, which includes the rice and *N. crassa* proteins.

Human is the only organism for which the structures of different Imp α variants are available [83-85, 113, 126] (Supplementary Table 3). The overall structures and the NLS-binding sites are conserved, consistent with the equivalent affinity *in vivo* of the influenza A PB2 NLS fragment for different Imp α variants [84]. Differences identified between variants include a higher flexibility and reduced affinity between bipartite NLSs and Imp α 3, as well as different levels of auto-inhibition [84].

Cargo release and recycling of importin- α

A consequence of the auto-inhibitory function of the Imp α IBB domain is the facilitation of cargo release in the nucleus; once the trimeric complex has traversed the nuclear pore, dissociation of Imp β 1 from Imp α in the nucleus allows the IBB domain to compete for the NLS-binding site. In addition to this mechanism, some reports have implicated Nup50 (Nup2p in yeast) in Imp α -cargo disassembly [86, 102, 116]. Solution-binding assays suggest that the addition of Nup2p can accelerate displacement of NLS cargo from the yeast protein [127]. Consistently, structures of Imp α in complex with Nup50 peptide segments reveal interactions between the Nup side-chains and the Imp α minor binding site (Fig. 7e) [86, 102, 126]. In addition, contacts are observed outside the NLS binding region between the Nup and the Imp α C-terminal region.

Imp α is recycled back to the cytoplasm by its export factor CAS (Cse1 in yeast), which binds preferentially to cargo-free Imp α [128]. Nuclear export of Imp α further depends on RanGTP [129, 130], and the formation of the trimeric CAS:Imp α :RanGTP complex has been shown to be highly cooperative [128]. Binding of CAS-RanGTP to Imp α displaces Nup50 through steric hindrance. Like other members of the β -Kap family, CAS has a superhelical HEAT repeat architecture and wraps around RanGTP and the Imp α C-terminal region [116]. Extensive interactions are observed between the outer surface of Imp α ARM repeats 8-10, in a region that overlaps with the VP24-binding site. Together, these mechanisms ensure that cargo is efficiently displaced from Imp α and that recycling of the adaptor to the cytoplasm occurs only after cargo disassembly.

Snurportin-mediated nuclear import

The nuclear import of assembled spliceosomal subunits, the uridine-rich small ribonucleoprotein particles (UsnRNPs), employs a variation of the classical nuclear import pathway that utilizes a distinct adaptor protein termed snurportin-1. This protein, first identified via UV cross-linking to a m₃G-capped oligonucleotide, binds Imp β 1 with an IBB domain similar to that found in Imp α , but lacks the canonical ARM-repeat region [131], instead adopting a double β -sheet fold to form the m₃G-cap binding pocket [132]. Snurportin-1 binds both the hyper-methylated cap and the first nucleotide of the RNA in a stacking conformation, with the specificity determined by a highly solvent-exposed tryptophan [132].

Symportin-1-mediated nuclear import

Recent reports have described a new adaptor protein termed symportin-1 (Syo1 for synchronized import) that links nuclear cargo to Trn1 for nuclear translocation. Syo1 was

identified through tandem affinity purification (TAP) analysis of the yeast ribosomal protein Rpl5. Biochemical assays demonstrated direct binding of Syo1 to Rpl5 and the related protein Rpl11, as well as stable trimeric Syo1:Rpl5:Rpl11 complexes [133]. As Rpl5 and Rpl11 form a functional cluster within the ribosome, simultaneous binding to Syo1 suggested concomitant import of the ribosomal subunits, unlike other import pathways described to date that mediate binding of individual cargoes.

Syo1 is recognized by Trn1 through an N-terminal PY-NLS motif. The structure of *Chaetomium thermophilum* Syo1 revealed an unusual combination of four N-terminal ARM repeats fused to six C-terminal HEAT repeats in its globular cargo-binding domain (Fig. 8a). The absence of the ([^]R/K)XXWXXXN motif that forms the binding pockets on the Imp α surface suggests that cNLS sequences are not able to bind to Syo1. However, structural characterization of the Syo1-Rpl5 peptide complex reveals that the inner concave surface of the Syo1 HEAT repeats mediates binding to the Rpl5 N-terminal region in a manner highly similar to the Imp α -NLS interaction [133]. The Rpl5 peptide binds to Syo1 in an extended conformation with a short helical segment [133]. Conserved basic and aromatic residues throughout the Rpl5 linear motif mediate Syo1 binding. By contrast, the structure of the trimeric complex revealed that Rpl11 binding is mediated by the outer surface of the Syo1 superhelix, with additional contacts from a helical region in the large acidic loop of HEAT repeat 1 (Fig. 8b) [134].

Other nuclear import pathways

Although most nuclear proteins depend on β -Kaps to reach their subcellular destination, some alternative pathways exist. These include the pathway operating during heat-shock stress that involves the carrier Hikeshi, and the RanGDP import pathway that involves the nuclear transport factor NTF2 (see below). The actin-capping protein CapG also

uses the interaction with NTF2 and Ran to enter the nucleus [135]. The RaDAR (RanGDP/ankyrin repeat) pathway has recently been characterized as an importin-independent nuclear import pathway for a number of ankyrin-repeat proteins, with the signal identified as a hydrophobic residue at a specific position of two consecutive repeats [136]. The calcium-binding protein calmodulin can function as an import factor independent of β -Kaps, GTP and Ran, for a range of cargoes, particularly transcription factors [137-139]. Some proteins enter the nucleus independent of carrier molecules, for example by direct binding to Nups, diffusion through the NPC and interaction with nuclear components (e.g. the ARM-repeat protein β -catenin [140]). Lectins have been described as import factors for glycosylated proteins, and viruses disrupt the nuclear envelope during infection. Transport factors remaining to be characterized are involved in light-dependent nucleo-cytoplasmic trafficking in plants [141]. Some proteins can “piggy-back” through interactions with proteins with NLSs [142-145]. Many small proteins (for example histones) are imported by active transport mechanisms, although they could freely diffuse into the nucleus [146-149]. Proteins do not always use a single nuclear import pathway, which may be important under circumstances when conventional pathways are inhibited [150].

Hikeshi-mediated nuclear import of Hsp70 proteins

Nuclear import of heat-shock proteins from the Hsp70 family has been shown to be mediated by the nuclear transport factor Hikeshi [151] (see the review by Imamoto in this issue). The crystal structure of Hikeshi reveals a dimeric two-domain protein, with the N-terminal domains (NTDs) responsible for the interaction with Nups [152]. The asymmetric nature of the dimer has been suggested to be important for the recognition of the ATP-bound form of Hsp70.

NTF2-mediated nuclear import of RanGDP

The conformational changes generated by nuclear RanGTP binding to β -Kaps lead to the release of their macromolecular cargo and adaptors (such as Imp α), but the karyopherins can only participate in another import cycle after Ran is released, following stimulation of its GTPase activity in the cytoplasm by RanGAP. The RanGDP generated in this way is then returned to the nucleus for recharging with GTP by the chromatin-bound RanGEF. Although Ran is a 25-kDa protein, its rate of nuclear import using simple diffusion appears to be insufficiently rapid to maintain adequate levels of karyopherin-based nuclear transport and is augmented by NTF2 [153, 154] (reviewed in [155]).

The structure of NTF2 features an extensive β -sheet flanked by three helices, yielding a cone-shaped molecule that dimerizes in solution (Fig. 9). The β -sheets of the protomers form an extensive interface in the NTF2 dimer, in which a considerable number of hydrophobic residues are buried [156, 157]. The arrangement of the helices that flank the β -sheet generates an extensive cavity that is lined by hydrophobic residues and which forms the principal interaction interface with RanGDP [158]. NTF2 recognizes the GDP-bound state of Ran through binding to the switch-II loop (Fig. 9). In the RanGDP conformation, F⁷² in the switch-II loop inserts into the NTF2 cavity and this essentially hydrophobic interaction is complemented by salt bridges formed between K⁷¹ and R⁷⁶ of Ran and D⁹²/D⁹⁴ and E⁴² of NTF2, respectively [158]. To mediate movement through the nuclear pore, NTF2 also binds to FxFG motifs present in many Nups, with the Phe residues of these motifs becoming buried in a hydrophobic cavity formed between the two chains in the dimer (Fig. 9) at a position opposite from that to which RanGDP binds [40, 159]. The affinity of NTF2 for RanGDP is of the order of 100 nM [157], which ensures that the dissociation rate of the NTF2:RanGDP complex is sufficiently slow for it to remain intact during nuclear transport, whereas the affinity of NTF2 for the nucleoporin FxFG repeats is weaker ($\sim 5 \mu\text{M}$), consistent with its

forming much more rapidly dissociating complexes that enable the NTF2:RanGDP complex to move through the nuclear pore transport channel rapidly, using transient binding to Nups [159].

Regulation of nuclear import pathways

One of the key features of limiting movement in and out of the nuclear compartment is the opportunity to regulate these transport processes. Regulation is essential for fine-tuning transport activities according to the actual cellular needs. Nucleo-cytoplasmic trafficking is regulated on several levels (see [160-162] for reviews), with new mechanisms continuing to be discovered (Fig. 10). The nuclear accumulation rate of cargoes is directly related to their binding affinities for their import receptors [163, 164] and so transport processes can be regulated by modulating these binding affinities either by direct changes to the NLS, or by physically blocking importin:cargo interactions through inter-molecular or intra-molecular NLS masking. Post-translational modification-induced changes are perhaps the best-described means of modulating transport processes. Although phosphorylation plays a central role, there is also a growing number of examples based on methylation or acetylation [165-169]. Post-translational modifications link nucleo-cytoplasmic transport to a variety of signaling pathways including the cell cycle, gene transcription, RNA metabolism, immune responses, apoptosis and the DNA damage response. Several of the case studies examined in the literature involve proteins that constantly shuttle in and out of the nucleus and the balance between nuclear import and export establishes the specific localization pattern. However, many of the mechanistic studies fail to investigate precisely how a post-translational modification perturbs the dynamics. For instance, if the post-translational modifications in the nucleus lead to increased cytoplasmic accumulation, this could result from changes in nuclear export (whether it is enhanced) or nuclear import (whether it is inhibited) or both.

Regulation by phosphorylation*Modulation of importin: cargo binding affinity*

There are several examples in the literature of the introduction of a negative charge by phosphorylation inhibiting NLS binding [160, 161, 170-172]. Well-established examples include the phosphorylation of the yeast transcription factor Pho4, which disrupts the interaction with its dedicated carrier Pse1, an β -Kap family member [173, 174], and the inhibitory phosphorylation by Cdk1 (cyclin-dependent kinase 1), which introduces negative charges that interfere with importin binding in a number of proteins [93, 175, 176]. In the case of the human dUTPase, structural work suggests that phosphorylation in the vicinity of the NLS leads to altered intra-NLS contacts that prevent favorable interactions with Imp α , resulting in the cytoplasmic accumulation of the phosphorylated form [93]. The NLS:importin dissociation constants fall into a rather wide range [177, 178], and the effect of phosphorylation will depend on whether it is capable of moving the affinity over the threshold, so it falls outside the functional NLS range [65, 90, 107, 177]. A high-affinity cargo complex would require a more substantial alteration to make the NLS non-functional.

Phosphorylation can also enhance nuclear accumulation through increasing NLS:Imp α affinity. A well-established example is protein kinase CK2-mediated phosphorylation of the SV40-Tag NLS at positions S^{111/112}, which enhances affinity for Imp α 2-fold, leading to a considerable increase in the nuclear import rate [179]. The nuclear-transport efficiency also increases for the Epstein-Barr virus nuclear antigen 1 (EBNA-1) if its NLS is phosphorylated at S³⁸⁵, which increases its affinity for Imp α 5. However, phosphorylation in two other neighboring positions, S³⁸³ and S³⁸⁶, decreases the nuclear import rate [180, 181]. The precise structural reasons behind the enhanced affinity due to phosphorylation remain unclear [89]. Negative charges in the linker region of bipartite cNLSs

were shown to have a positive effect on Imp α binding. Generation of peptide inhibitors against the classical nuclear transport pathway led to bipartite NLSs that have several Glu or Asp residues in their linker regions. These help in maximizing the possible interactions in the cargo:carrier complex [90, 110]. Clearly, phosphorylation has an effect specific to the position of the phosphorylated residue relative to the positive cluster of the NLS [175].

Intra-molecular NLS masking

Intra-molecular NLS masking can also inhibit cargo:carrier complex formation, through induced structural changes in the cargo making the NLS inaccessible to Imp α . In the case of the *X. laevis* b-Myb protein, the C-terminal domain simultaneously inhibits DNA binding and NLS function. During embryo development, b-Myb is subjected to several modifications, which result in the NLSs that facilitate nuclear accumulation of the protein becoming accessible [182]. STAT1 activation through the phosphorylation of a tyrosine residue (Y⁷⁰¹) is one of the central events in cytokine signaling and the regulation of immune responses. Phosphorylation induces a structural rearrangement that shifts STAT1 dimers from an antiparallel to a parallel conformation, exposing a non-classical NLS that is recognized by Imp α 5. Dimerization and phosphorylation are essential for efficient nuclear accumulation during STAT1 activation, even though the phospho-tyrosine residue is not a binding determinant for Imp α 5 [124, 181, 183-186].

Inter-molecular NLS masking and organelle-specific retention

Inter-molecular masking can occur if the binding of a heterologous protein prevents the interaction of the cargo and its carrier. One of the best-described examples is the NF- κ B p50/p65 heterodimer, a transcription factor regulating immune and stress responses, apoptosis and differentiation. NF- κ B is kept inactive by its inhibitor, I κ B α , which impedes

NF- κ B from being recognized by the nuclear import machinery. Crystal structures show how the NLSs of both NF- κ B p50 and p65 subunits are covered by the ankyrin repeat region of I κ B α [187, 188]. I κ B α binds the NLS of NF- κ B until phosphorylation licenses its ubiquitin-mediated degradation, which enables Imp α 3 and Imp α 4 to access the unmasked NF- κ B NLSs [189-192].

A recently described E3 ubiquitin ligase that binds phospho-NLSs, the BRCA1-binding protein 2 (BRAP2), was shown to reduce the nuclear accumulation of several viral and endogenous proteins, depending on their phosphorylation state. Although it does not completely sequester its targets in the cytoplasm, it fine-tunes their localization pattern [193, 194].

DNA or RNA can also be responsible for intermolecular masking. For example, the DNA binding region and the Imp β 1-recognized NLS of the human sex-determining factor SRY overlap. DNA binding inhibits Imp β 1 binding, and *vice versa*. This mechanism may also facilitate the release of the SRY:Imp β 1 complex once it enters the nucleus [195]. Interestingly, acetylation of SRY is necessary for proper Imp β 1:SRY complex formation, showing the interplay among different modes of regulation [196]. The NLS of HDAC4, a class-IIa histone deacetylase, is masked by phosphorylation-induced 14-3-3 protein binding, making both NLSs inaccessible to nuclear import factors and causing cytoplasmic retention [197]. A similar mechanism seems to apply for HDAC5, 7 and 9, where phosphorylation near the NLSs, along with 14-3-3 binding, inhibits nuclear translocation [198].

Regulation by methylation and acetylation

Post-translational modifications of histones, including acetylation and methylation, are crucial in epigenetics. A growing number of examples show that acetylation and methylation of Lys or Arg residues, in addition to those in histones, regulates a variety of

cellular functions [199-201]. Interestingly, Imp α itself is targeted for acetylation within its IBB domain by p300/CBP, increasing its ability to bind Imp β 1 *in vitro* [202]. p300/CBP interacts with several components of the nuclear transport machinery, possibly fine-tuning their functions by affecting their intracellular distribution [203]. Methylation and acetylation of Lys and Arg residues can directly modulate NLS/NES function through altering the interaction with the transport machinery by modulating the residue charge. These modifications may also have indirect effects on localization through alteration of binding between interaction partners or by promoting conformational changes. Although the precise mechanisms are mostly unclear [204], several well-documented examples are summarized below.

Modulation of importin: cargo binding affinity

In c-Abl, lysine acetylation occurs in the NLS. The protein is involved in apoptosis when nuclear, whereas it responds to proliferative signals in the cytoplasm. It can be acetylated in the nucleus by P/CAF (an acetyltransferase) within one of its NLSs, leading to cytoplasmic accumulation. It is hypothesized that when the acetylated form is exported to the cytoplasm, it cannot re-enter the nucleus because its NLS is no longer recognized by the import machinery [205]. A similar mechanism was proposed for RECQL4 (a DNA helicase important for genomic integrity maintenance), which is acetylated by p300 [206], and for HMGB1 (a protein involved in transcriptional control) [207]. Acetylation within the NLS of poly(A) polymerase (PAP) leads to its cytoplasmic accumulation. Acetylation directly interferes with PAP binding to Imp α /Imp β 1 [208]. In the case of another P/CAF substrate, E1A, a clear negative effect of NLS acetylation on Imp α 3 binding has been demonstrated [209]. Interestingly, for P/CAF itself, intramolecular acetylation is required for its nuclear localization. However, a mutant incapable of autoacetylation is strictly cytoplasmic, despite

the fact that the acetylated form of P/CAF shows decreased binding to both Imp α 1 and Imp β in *in vitro* pull-down assays, compared to non-acetylated P/CAF. Acetylation possibly affects the P/CAF localization pattern through regulation of proteasome-dependent degradation processes [210]. Besides acetylation, phosphorylation also influences the localization of P/CAF. Phosphorylation may promote the dissociation of P/CAF-PP1/PP2a cytoplasmic complexes, and allow the nuclear import of P/CAF [211]. Although the precise mechanism is not yet known, acetylation near NLSs of Net1A (a RhoA GEF protein) alters the dynamics of its nucleo-cytoplasmic shuttling, probably by slowing its nuclear re-import rate, leading to cytoplasmic accumulation [212]. Acetylation can also directly enhance importin: cargo interactions. It has been shown that p300-mediated acetylation in the proximity of the NLS of SRY (K¹³⁶) is needed for nuclear localization, facilitating Imp β 1 binding [196].

Intra-molecular NLS masking

Conformational changes induced by acetylation can alter nucleo-cytoplasmic transport processes in a fashion similar to phosphorylation. The transcription factor HNF-4 shuttles between the nuclear and cytoplasmic compartments and CBP-mediated acetylation is hypothesized to induce conformational changes that make the NES inaccessible to CRM1. Acetylation also enhances the ability of HNF-4 to bind DNA and CBP, both of which help its nuclear anchoring [213]. One of the key aspects of the many modes of regulation of p53 localization is its phosphorylation and acetylation-dependent tetramerization, which influences the accessibility of its NES and NLSs (reviewed in [214, 215]). Acetylation of lysine residues in the C-terminal region of p53 inhibits its oligomerization, resulting in an exposed NES and effective nuclear export [216]. By contrast, CBP-dependent acetylation of survivin at K¹²⁹ enhances its oligomerization, making its NES sequence inaccessible to CRM1 and leading to nuclear accumulation of the protein [217, 218]. The acetylation of K⁴³³

in the protein kinase PKM2 by p300 is speculated to influence its nuclear transport through stabilizing its dimeric state, because the tetrameric state may bury its NLS [219], or by modulating a “piggy-back” nuclear entry mechanism of PKM2.

Inter-molecular NLS masking and organelle-specific retention

The localization of Yap, one of the key components of the Hippo signaling pathway, is regulated through lysine methylation. Yap is monomethylated at K⁴⁹⁴ by Set7, which leads to its cytoplasmic retention through an unknown mechanism [220]. Hsp70, a protein with roles in folding and degradation, mainly localizes to the nucleus when dimethylated on K⁵⁶¹, whereas the unmethylated form is predominantly cytoplasmic. A small fraction of the overall Hsp70 pool was reported to be dimethylated in cancer cells and this form is thought to interact specifically with Aurora kinase B, which might be responsible for the chromatin association of methylated Hsp70 [221]. RNA helicase A shuttles between the nuclear and cytoplasmic compartments, and PRMT1-mediated arginine methylation is hypothesized to inhibit its interaction with a putative cytoplasmic retention factor binding its NLS containing C-terminal region [222]. CtBP2 is actively exported from the nucleus in a CRM1-dependent manner, but this is prevented by K¹⁰ acetylation-dependent nuclear sequestration, mediated by p300 [223]. Interestingly, the localization of its closely related protein, CtBP1, is regulated by sumoylation (enhanced nuclear entry or retention) [224], cytoplasmic retention [225] and phosphorylation [226]. Finally, acetylation of retinoblastoma protein (Rb) by P/CAF was shown to be important for Rb to remain nuclear during keratinocyte differentiation. Because acetylation presumably happens in the nucleus, it does not have an effect on nuclear import, although the modification is within the NLS [227].

Regulation by ubiquitination and sumoylation

Other post-translational covalent modifications, including ubiquitination or sumoylation, can also regulate protein nuclear import. These modifications can directly target the Lys residues in the NLS. A well-documented example involves cytidylyltransferase (CCT), a protein involved in phosphatidylcholine biosynthesis. Monoubiquitination on K⁵⁷ masks the NLS of CCT, resulting in cytoplasmic accumulation [228]. By a similar mechanism, ubiquitinated K³¹⁹⁻³²¹ in p53 block Imp α 3 binding, inhibiting nuclear entry [229]. Nuclear entry of PTEN, the regulator of phosphatidylinositol 3-kinase signaling, is essential for its tumor suppressor function and this translocation depends on NEDD4-1 mediated monoubiquitination of two lysine residues (K¹³ and K²⁸⁹) [230-232]. Sumoylation of PAP within its NLS is essential for its nuclear import, providing regulation additional to acetylation [233]. Sumoylation can also enhance nuclear accumulation through masking NESs as in the case of Klf5, a transcription factor regulating cell proliferation [234], or through supporting nuclear retention without altering nuclear import dynamics as in the case of SAE1 (SUMO activating enzyme 1) [235]. Nuclear translocation of the enzymes in the *de novo* thymidylate biosynthesis pathway is sumoylation-dependent and takes place at the beginning of S-phase [236-238].

Other factors regulating nucleocytoplasmic trafficking

Among several other possible ways to regulate nuclear translocation, cytoplasmic anchoring and the contribution of the microtubular network are important for a number of proteins. To overcome barriers to diffusion, the microtubular system and its associated molecular motors are essential for some viral proteins to reach the nucleus (reviewed in [239, 240]), but several non-viral proteins also use them, presumably to enhance the rate and extent of their nuclear import. Nocodazole (a tubulin polymerization inhibitor) treatment of cells significantly reduces the nuclear accumulation of p53 [241], Rb [242] and PTHrP [243, 244].

These results show that NLS-containing cargoes can be transported actively to sites close to the NPCs, enhancing nuclear entry rates by moving cargoes to intracellular regions where the Imp β 1 concentrations are high [245, 246]. This may enhance the response time of cells to extra- or intracellular stimuli when rapid transport processes are needed for proper function (for example, in cell signaling and DNA damage responses).

The nuclear localization of several steroid receptors, including the glucocorticoid and estrogen receptors, is regulated through cytoplasmic retention. Without their ligands, these receptors are sequestered in the cytoplasm by Hsp90 through their ligand-binding domains. Upon ligand binding, they release Hsp90 and are imported into the nucleus using their NLSs [247]. The androgen receptor is similarly bound to Imp7, which in the absence of ligand inhibits binding of Imp α to the NLS and causes the receptor to remain in the cytoplasm. Androgen binding induces conformational changes in the receptor that result in the release of Imp7, exposing the NLS for binding to Imp α and leading to translocation into the nucleus [248].

In addition to the affinity for their receptors, the NLS copy number also has an important effect on nuclear accumulation efficiency. Higher numbers of NLSs provide advantage in competing for importins in the cellular environment, making oligomerization-induced NLS-copy number variations another way of regulating nucleo-cytoplasmic transport processes [249, 250].

Conclusions

Although the first components of nuclear import pathways were only identified in 1993 [251, 252], there is now a substantial molecular understanding of how many of the nuclear protein import pathways function. However, several key challenges remain, including determining the precise mechanism by which the cargo-carrier complex is able to overcome

the barrier function generated by the nucleoporins to facilitate translocation through the NPC; accurate prediction of nuclear targeting signals from protein sequences; defining the catalytic role of nucleoporins on the cytoplasmic and nuclear faces of NPCs in transport complex assembly and disassembly; characterizing the molecular mechanisms or regulation of nucleocytoplasmic transport; determining the structures of complexes of karyopherins with full-length cargo proteins; and targeting the components of nuclear transport machinery therapeutically. Examples of β -Kap family proteins involved in nuclear import with no structural information available currently include importin-7 and importin-8 [253, 254]. Structural biology has played a key role in characterizing the molecular events in nuclear import pathways, and will no doubt continue to do so in the future. Recent advances in cryo-electron microscopy [255] and X-ray free electron lasers [256], to name just two, should help characterize the structures of large and dynamic multi-protein complexes that form during nuclear import.

Acknowledgements

The research in the authors' laboratories has been supported by the National Health and Medical Research Council (NHMRC) Program (1000512, 565526), Project (1003326) and Research Fellowship (1003325), and Australian Research Council (ARC) Discovery Project (DP120100685) to BK; NHMRC Fellowship to AGS (1090408); UK Medical Research Council Grant U105178939 to MS; Hungarian Scientific Research Fund OTKA [NK 84008, K109486], Baross Program of the New Hungary Development Plan [3DSTRUCT, OMFB-00266/2010 REG-KM-09-1-2009-0050], Hungarian Academy of Sciences ([TTK IF-28/2012] and ICGEB CRP/HUN14-01 to BGV; Young Researcher Fellowships from the Hungarian Academy of Sciences to GR; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to MRMF; and ARC Future Fellowship to JKF (100095). We acknowledge the use of the University of Queensland Remote Operation Crystallization and X-ray Diffraction Facility (UQ ROCX) and the Australian Synchrotron.

References

- [1] Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, et al. The molecular architecture of the nuclear pore complex. *Nature*. 2007;450:695-701.
- [2] Aitchison JD, Wozniak RW. Cell biology: pore puzzle. *Nature*. 2007;450:621-2.
- [3] Andrade MA, Bork P. HEAT repeats in the Huntington's disease protein. *Nature Genet*. 1995;11:115-6.
- [4] Kobe B, Kajava AV. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends Biochem Sci*. 2000;25:509-15.
- [5] Gorlich D, Pante N, Kutay U, Aebi U, Bischoff FR. Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J*. 1996;15:5584-94.
- [6] Bischoff FR, Ponstingl H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature*. 1991;354:80-2.
- [7] Bischoff FR, Klebe C, Kretschmer J, Wittinghofer A, Ponstingl H. RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proc Natl Acad Sci U S A*. 1994;91:2587-91.
- [8] Moroianu J, Blobel G, Radu A. Nuclear protein import: Ran-GTP dissociates the karyopherin alpha-beta heterodimer by displacing alpha from an overlapping binding site on beta. *Proc Natl Acad Sci U S A*. 1996;93:7059-62.
- [9] Rexach M, Blobel G. Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell*. 1995;83:683-92.
- [10] Chi NC, Adam EJ, Visser GD, Adam SA. RanBP1 stabilizes the interaction of Ran with p97 nuclear protein import. *J Cell Biol*. 1996;135:559-69.
- [11] Gorlich D, Vogel F, Mills AD, Hartmann E, Laskey RA. Distinct functions for the two importin subunits in nuclear protein import. *Nature*. 1995;377:246-8.
- [12] Kobe B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nature Struct Biol*. 1999;6:388-97.
- [13] Conti E, Kuriyan J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure*. 2000;8:329-38.
- [14] Kutay U, Guttinger S. Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol*. 2005;15:121-4.

- [15] Cook A, Bono F, Jinek M, Conti E. Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem.* 2007;76:647-71.
- [16] Chook YM, Suel KE. Nuclear import by karyopherin-betas: recognition and inhibition. *Biochim Biophys Acta.* 2011;1813:1593-606.
- [17] Quan Y, Ji ZL, Wang X, Tartakoff AM, Tao T. Evolutionary and transcriptional analysis of karyopherin beta superfamily proteins. *Mol Cell Proteomics.* 2008;7:1254-69.
- [18] Andrade MA, Petosa C, O'Donoghue SI, Muller CW, Bork P. Comparison of ARM and HEAT protein repeats. *J Mol Biol.* 2001;309:1-18.
- [19] Kobe B, Gleichmann T, Horne J, Jennings IG, Scotney PD, Teh T. Turn up the HEAT. *Structure.* 1999;7:R91-R7.
- [20] Malik HS, Eickbush TH, Goldfarb DS. Evolutionary specialization of the nuclear targeting apparatus. *Proc Natl Acad Sci U S A.* 1997;94:13738-42.
- [21] Forwood JK, Lange A, Zachariae U, Marfori M, Prest C, Grubmuller H, et al. Quantitative structural analysis of importin-beta flexibility: paradigm for solenoid protein structures. *Structure.* 2010;18:1171-83.
- [22] Cingolani G, Petosa C, Weis K, Muller CW. Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature.* 1999;399:221-9.
- [23] Cingolani G, Bednenko J, Gillespie MT, Gerace L. Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta. *Mol Cell.* 2002;10:1345-53.
- [24] Lee SJ, Sekimoto T, Yamashita E, Nagoshi E, Nakagawa A, Imamoto N, et al. The structure of importin-beta bound to SREBP-2: nuclear import of a transcription factor. *Science.* 2003;302:1571-5.
- [25] Choi S, Yamashita E, Yasuhara N, Song J, Son SY, Won YH, et al. Structural basis for the selective nuclear import of the C2H2 zinc-finger protein Snail by importin beta. *Acta Crystallogr D Biol Crystallogr.* 2014;70:1050-60.
- [26] Lott K, Cingolani G. The importin beta binding domain as a master regulator of nucleocytoplasmic transport. *Biochim Biophys Acta.* 2011;1813:1578-92.
- [27] Vetter IR, Arndt A, Kutay U, Gorlich D, Wittinghofer A. Structural view of Ran-importin beta interaction at 2.3 Å resolution. *Cell.* 1999;97:635-46.
- [28] Lee SJ, Matsuura Y, Liu SM, Stewart M. Structural basis for nuclear import complex dissociation by RanGTP. *Nature.* 2005;435:693-6.
- [29] Forwood JK, Lonhienne TG, Marfori M, Robin G, Meng W, Guncar G, et al. Kap95p binding induces the switch loops of RanGDP to adopt the GTP-bound conformation:

- implications for nuclear import complex assembly dynamics. *J Mol Biol.* 2008;383:772-82.
- [30] Vetter IR, Nowak C, Nishimoto T, Kuhlmann J, Wittinghofer A. Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. *Nature.* 1999;398:39-46.
- [31] Bischoff FR, Gorlich D. RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors. *FEBS Lett.* 1997;419:249-54.
- [32] Saric M, Zhao X, Korner C, Nowak C, Kuhlmann J, Vetter IR. Structural and biochemical characterization of the Importin-beta.Ran.GTP.RanBD1 complex. *FEBS Lett.* 2007;581:1369-76.
- [33] Seewald MJ, Korner C, Wittinghofer A, Vetter IR. RanGAP mediates GTP hydrolysis without an arginine finger. *Nature.* 2002;415:662-6.
- [34] Bischoff FR, Krebber H, Smirnova E, Dong W, Ponstingl H. Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J.* 1995;14:705-15.
- [35] Seewald MJ, Kraemer A, Farkasovsky M, Korner C, Wittinghofer A, Vetter IR. Biochemical characterization of the Ran-RanBP1-RanGAP system: are RanBP proteins and the acidic tail of RanGAP required for the Ran-RanGAP GTPase reaction? *Mol Cell Biol.* 2003;23:8124-36.
- [36] Scheffzek K, Ahmadian MR, Kabsch W, Wiesmuller L, Lautwein A, Schmitz F, et al. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science.* 1997;277:333-8.
- [37] Rittinger K, Walker PA, Eccleston JF, Nurmahomed K, Owen D, Laue E, et al. Crystal structure of a small G protein in complex with the GTPase-activating protein rhoGAP. *Nature.* 1997;388:693-7.
- [38] Villa Braslavsky CI, Nowak C, Gorlich D, Wittinghofer A, Kuhlmann J. Different structural and kinetic requirements for the interaction of Ran with the Ran-binding domains from RanBP2 and importin-beta. *Biochemistry.* 2000;39:11629-39.
- [39] Kuhlmann J, Macara I, Wittinghofer A. Dynamic and equilibrium studies on the interaction of Ran with its effector, RanBP1. *Biochemistry.* 1997;36:12027-35.
- [40] Bayliss R, Littlewood T, Strawn LA, Wentz SR, Stewart M. GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. *J Biol Chem.* 2002;277:50597-606.

- [41] Liu SM, Stewart M. Structural basis for the high-affinity binding of nucleoporin Nup1p to the *Saccharomyces cerevisiae* importin-beta homologue, Kap95p. *J Mol Biol.* 2005;349:515-25.
- [42] Bednenko J, Cingolani G, Gerace L. Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. *J Cell Biol.* 2003;162:391-401.
- [43] Isgro TA, Schulten K. Binding dynamics of isolated nucleoporin repeat regions to importin-beta. *Structure.* 2005;13:1869-79.
- [44] Lee BJ, Cansizoglu AE, Suel KE, Louis TH, Zhang Z, Chook YM. Rules for nuclear localization sequence recognition by karyopherin beta 2. *Cell.* 2006;126:543-58.
- [45] Bonifaci N, Moroianu J, Radu A, Blobel G. Karyopherin beta2 mediates nuclear import of a mRNA binding protein. *Proc Natl Acad Sci U S A.* 1997;94:5055-60.
- [46] Pollard VW, Michael WM, Nakielny S, Siomi MC, Wang F, Dreyfuss G. A novel receptor-mediated nuclear protein import pathway. *Cell.* 1996;86:985-94.
- [47] Fridell RA, Truant R, Thorne L, Benson RE, Cullen BR. Nuclear import of hnRNP A1 is mediated by a novel cellular cofactor related to karyopherin-beta. *J Cell Sci.* 1997;110 (Pt 11):1325-31.
- [48] Cansizoglu AE, Lee BJ, Zhang ZC, Fontoura BM, Chook YM. Structure-based design of a pathway-specific nuclear import inhibitor. *Nat Struct Mol Biol.* 2007;14:452-4.
- [49] Imasaki T, Shimizu T, Hashimoto H, Hidaka Y, Kose S, Imamoto N, et al. Structural basis for substrate recognition and dissociation by human transportin 1. *Mol Cell.* 2007;28:57-67.
- [50] Zhang ZC, Chook YM. Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc Natl Acad Sci U S A.* 2012;109:12017-21.
- [51] Soniat M, Sampathkumar P, Collett G, Gizzi AS, Banu RN, Bhosle RC, et al. Crystal structure of human karyopherin beta2 bound to the PY-NLS of *Saccharomyces cerevisiae* Nab2. *J Struct Funct Genomics.* 2013;14:31-5.
- [52] Dormann D, Madl T, Valori CF, Bentmann E, Tahirovic S, Abou-Ajram C, et al. Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. *EMBO J.* 2012;31:4258-75.
- [53] Chook YM, Jung A, Rosen MK, Blobel G. Uncoupling Kapbeta2 substrate dissociation and ran binding. *Biochemistry.* 2002;41:6955-66.

- [54] Chook YM, Blobel G. Structure of the nuclear transport complex karyopherin-beta2-RanGppNHp. *Nature*. 1999;399:230-7.
- [55] Kataoka N, Bachorik JL, Dreyfuss G. Transportin-SR, a nuclear import receptor for SR proteins. *J Cell Biol*. 1999;145:1145-52.
- [56] Lai MC, Lin RI, Huang SY, Tsai CW, Tarn WY. A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *J Biol Chem*. 2000;275:7950-7.
- [57] Lai MC, Lin RI, Tarn WY. Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proc Natl Acad Sci U S A*. 2001;98:10154-9.
- [58] Maertens GN, Cook NJ, Wang W, Hare S, Gupta SS, Oztop I, et al. Structural basis for nuclear import of splicing factors by human Transportin 3. *Proc Natl Acad Sci U S A*. 2014;111:2728-33.
- [59] Mingot JM, Kostka S, Kraft R, Hartmann E, Gorlich D. Importin 13: a novel mediator of nuclear import and export. *EMBO J*. 2001;20:3685-94.
- [60] Bono F, Cook AG, Grunwald M, Ebert J, Conti E. Nuclear import mechanism of the EJC component Mago-Y14 revealed by structural studies of importin 13. *Mol Cell*. 2010;37:211-22.
- [61] Grunwald M, Bono F. Structure of Importin13-Ubc9 complex: nuclear import and release of a key regulator of sumoylation. *EMBO J*. 2011;30:427-38.
- [62] Grunwald M, Lazzaretti D, Bono F. Structural basis for the nuclear export activity of Importin13. *EMBO J*. 2013;32:899-913.
- [63] Gorlich D, Kostka S, Kraft R, Dingwall C, Laskey RA, Hartmann E, et al. Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Curr Biol*. 1995;5:383-92.
- [64] Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem*. 2007;282:5101-5.
- [65] Marfori M, Mynott A, Ellis JJ, Mehdi AM, Saunders NF, Curmi PM, et al. Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim Biophys Acta*. 2011;1813:1562-77.
- [66] Chang CW, Counago RM, Williams SJ, Boden M, Kobe B. The distribution of different classes of nuclear localization signals (NLSs) in diverse organisms and the utilization of

the minor NLS-binding site in plant nuclear import factor importin- α . *Plant Signal Behav.* 2013;8.

[67] Kalderon D, Richardson WD, Markham AF, Smith AE. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature.* 1984;311:33-8.

[68] Kalderon D, Roberts BL, Richardson WD, Smith AE. A short amino acid sequence able to specify nuclear location. *Cell.* 1984;39:499-509.

[69] Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J Cell Biol.* 1988;107:841-9.

[70] Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell.* 1991;64:615-23.

[71] Adam SA, Marr RS, Gerace L. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J Cell Biol.* 1990;111:807-16.

[72] Adam EJ, Adam SA. Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope. *J Cell Biol.* 1994;125:547-55.

[73] Gorlich D, Prehn S, Laskey RA, Hartmann E. Isolation of a protein that is essential for the first step of nuclear protein import. *Cell.* 1994;79:767-78.

[74] Chi NC, Adam EJ, Adam SA. Sequence and characterization of cytoplasmic nuclear protein import factor p97. *J Cell Biol.* 1995;130:265-74.

[75] Imamoto N, Shimamoto T, Kose S, Takao T, Tachibana T, Matsubae M, et al. The nuclear pore targeting complex binds to nuclear pores after association with a karyophile. *FEBS Letters.* 1995;368:415-9.

[76] Radu A, Blobel G, Moore MS. Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins. *Proc Natl Acad Sci U S A.* 1995;92:1769-73.

[77] Weis K, Ryder U, Lamond AI. The conserved amino-terminal domain of hSRP1 α is essential for nuclear protein import. *EMBO J.* 1996;15:1818-25.

[78] Kobe B, Heierhorst J, Feil SC, Parker MW, Benian GM, Weiss KR, et al. Giant protein kinases: domain interactions and structural basis of autoregulation. *EMBO J.* 1996;15:6810-21.

- [79] Conti E, Uy M, Leighton L, Blobel G, Kuriyan J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*. 1998;94:193-204.
- [80] Riggleman B, Wieschaus E, Schedl P. Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a *Drosophila* segment polarity gene. *Genes Dev*. 1989;3:96-113.
- [81] Dias SM, Wilson KF, Rojas KS, Ambrosio AL, Cerione RA. The molecular basis for the regulation of the cap-binding complex by the importins. *Nat Struct Mol Biol*. 2009;16:930-7.
- [82] Jeong S, Kim K, Lee JH, Cha JS, Khadka P, Cho HS, et al. Akt-mediated phosphorylation increases the binding affinity of hTERT for importin α to promote nuclear translocation. *J Cell Sci*. 2015.
- [83] Tarendeau F, Boudet J, Guilligay D, Mas PJ, Bougault CM, Boulo S, et al. Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat Struct Mol Biol*. 2007;14:229-33.
- [84] Pumroy RA, Ke S, Hart DJ, Zachariae U, Cingolani G. Molecular determinants for nuclear import of influenza A PB2 by importin alpha isoforms 3 and 7. *Structure*. 2015;23:374-84.
- [85] Trowitzsch S, Viola C, Scheer E, Conic S, Chavant V, Fournier M, et al. Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nat Commun*. 2015;6.
- [86] Matsuura Y, Lange A, Harreman MT, Corbett AH, Stewart M. Structural basis for Nup2p function in cargo release and karyopherin recycling in nuclear import. *EMBO J*. 2003;22:5358-69.
- [87] Fontes MR, Teh T, Kobe B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J Mol Biol*. 2000;297:1183-94.
- [88] Fontes MR, Teh T, Jans D, Brinkworth RI, Kobe B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *J Biol Chem*. 2003;278:27981-7.
- [89] Fontes MR, Teh T, Toth G, John A, Pavo I, Jans DA, et al. Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-alpha. *Biochem J*. 2003;375:339-49.

- [90] Marfori M, Lonhienne TG, Forwood JK, Kobe B. Structural basis of high-affinity nuclear localization signal interactions with importin- α . *Traffic*. 2012;13:532-48.
- [91] Giesecke A, Stewart M. Novel binding of the mitotic regulator TPX2 (target protein for *Xenopus* kinesin-like protein 2) to importin- α . *J Biol Chem*. 2010;285:17628-35.
- [92] Roman N, Christie M, Swarbrick CM, Kobe B, Forwood JK. Structural characterisation of the nuclear import receptor importin α in complex with the bipartite NLS of Prp20. *PLoS One*. 2013;8:e82038.
- [93] Rona G, Marfori M, Borsos M, Scheer I, Takacs E, Toth J, et al. Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. *Acta Crystallogr D Biol Crystallogr*. 2013;69:2495-505.
- [94] Cutress ML, Whitaker HC, Mills IG, Stewart M, Neal DE. Structural basis for the nuclear import of the human androgen receptor. *J Cell Sci*. 2008;121:957-68.
- [95] Chen MH, Ben-Efraim I, Mitrousis G, Walker-Kopp N, Sims PJ, Cingolani G. Phospholipid scramblase 1 contains a nonclassical nuclear localization signal with unique binding site in importin α . *J Biol Chem*. 2005;280:10599-606.
- [96] Takeda AA, de Barros AC, Chang CW, Kobe B, Fontes MR. Structural basis of importin- α -mediated nuclear transport for Ku70 and Ku80. *J Mol Biol*. 2011;412:226-34.
- [97] Yang SN, Takeda AA, Fontes MR, Harris JM, Jans DA, Kobe B. Probing the specificity of binding to the major nuclear localization sequence-binding site of importin- α using oriented peptide library screening. *J Biol Chem*. 2010;285:19935-46.
- [98] Mynott AV, Harrop SJ, Brown LJ, Breit SN, Kobe B, Curmi PM. Crystal structure of importin- α bound to a peptide bearing the nuclear localisation signal from chloride intracellular channel protein 4. *FEBS J*. 2011;278:1662-75.
- [99] Lott K, Bhardwaj A, Sims PJ, Cingolani G. A minimal nuclear localization signal (NLS) in human phospholipid scramblase 4 that binds only the minor NLS-binding site of importin α 1. *J Biol Chem*. 2011;286:28160-9.
- [100] Chang CW, Counago RL, Williams SJ, Boden M, Kobe B. Crystal structure of rice importin- α and structural basis of its interaction with plant-specific nuclear localization signals. *Plant Cell*. 2012;24:5074-88.

- [101] Chang CW, Counago RM, Williams SJ, Boden M, Kobe B. Distinctive conformation of minor site-specific nuclear localization signals bound to importin- α . *Traffic*. 2013;14:1144-54.
- [102] Matsuura Y, Stewart M. Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. *EMBO J*. 2005;24:3681-9.
- [103] Chang CW, Williams SJ, Counago RM, Kobe B. Structural basis of interaction of bipartite nuclear localization signal from *Agrobacterium* VirD2 with rice importin- α . *Mol Plant*. 2014;7:1061-4.
- [104] Wirthmueller L, Roth C, Fabro G, Caillaud MC, Rallapalli G, Asai S, et al. Probing formation of cargo/importin- α transport complexes in plant cells using a pathogen effector. *Plant J*. 2015;81:40-52.
- [105] Bernardes NE, Takeda AA, Dreyer TR, Freitas FZ, Bertolini MC, Fontes MR. Structure of importin- α from a filamentous fungus in complex with a classical nuclear localization signal. *PLoS One*. 2015;10:e0128687.
- [106] Leung SW, Harreman MT, Hodel MR, Hodel AE, Corbett AH. Dissection of the karyopherin α nuclear localization signal (NLS)-binding groove: functional requirements for NLS binding. *J Biol Chem*. 2003;278:41947-53.
- [107] Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *J Biol Chem*. 2001;276:1317-25.
- [108] Colledge WH, Richardson WD, Edge MD, Smith AE. Extensive mutagenesis of the nuclear location signal of simian virus 40 large-T antigen. *Mol Cell Biol*. 1986;6:4136-9.
- [109] Makkerh JPS, Dingwall C, Laskey RA. Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr Biol*. 1996;6:1025-7.
- [110] Kosugi S, Hasebe M, Entani T, Takayama S, Tomita M, Yanagawa H. Design of peptide inhibitors for the importin α /beta nuclear import pathway by activity-based profiling. *Chem Biol*. 2008;15:940-9.
- [111] Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, et al. Six classes of nuclear localization signals specific to different binding grooves of importin α . *J Biol Chem*. 2009;284:478-85.
- [112] Wolff T, Unterstab G, Heins G, Richt JA, Kann M. Characterization of an unusual importin α binding motif in the Borna disease virus p10 protein that directs nuclear import. *J Biol Chem*. 2002;277:12151-7.

- [113] Xu W, Edwards MR, Borek DM, Feagins AR, Mittal A, Alinger JB, et al. Ebola virus VP24 targets a unique NLS binding site on karyopherin alpha 5 to selectively compete with nuclear import of phosphorylated STAT1. *Cell Host Microbe*. 2014;16:187-200.
- [114] Fanara P, Hodel MR, Corbett AH, Hodel AE. Quantitative analysis of nuclear localization signal (NLS)-importin alpha interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding. *J Biol Chem*. 2000;275:21218-23.
- [115] Harreman MT, Cohen PE, Hodel MR, Truscott GJ, Corbett AH, Hodel AE. Characterization of the auto-inhibitory sequence within the N-terminal domain of importin alpha. *J Biol Chem*. 2003;278:21361-9.
- [116] Matsuura Y, Stewart M. Structural basis for the assembly of a nuclear export complex. *Nature*. 2004;432:872-7.
- [117] Lokareddy RK, Hapsari RA, van Rheenen M, Pumroy RA, Bhardwaj A, Steen A, et al. Distinctive properties of the nuclear localization signals of inner nuclear membrane proteins Heh1 and Heh2. *Structure*. 2015.
- [118] Kohler M, Speck C, Christiansen M, Bischoff FR, Prehn S, Haller H, et al. Evidence for distinct substrate specificities of importin alpha members in nuclear protein import. *Mol Cell Biol*. 1999;19:7782-91.
- [119] Goldfarb DS, Corbett AH, Mason DA, Harreman MT, Adam SA. Importin alpha: a multipurpose nuclear-transport receptor. *Trends Cell Biol*. 2004;14:505-14.
- [120] Mason DA, Stage DE, Goldfarb DS. Evolution of the metazoan-specific importin alpha gene family. *J Mol Evol*. 2009;68:351-65.
- [121] Mason DA, Fleming RJ, Goldfarb DS. *Drosophila melanogaster* importin alpha1 and alpha3 can replace importin alpha2 during spermatogenesis but not oogenesis. *Genetics*. 2002;161:157-70.
- [122] Rother F, Schmidt T, Popova E, Krivokharchenko A, Hugel S, Vilianovich L, et al. Importin alpha7 is essential for zygotic genome activation and early mouse development. *PLoS One*. 2011;6:e18310.
- [123] Talcott B, Moore MS. The nuclear import of RCC1 requires a specific nuclear localization sequence receptor, karyopherin alpha3/Qip. *J Biol Chem*. 2000;275:10099-104.

- [124] Sekimoto T, Imamoto N, Nakajima K, Hirano T, Yoneda Y. Extracellular signal-dependent nuclear import of Stat1 is mediated by nuclear pore-targeting complex formation with NPI-1, but not Rch1. *EMBO J.* 1997;16:7067-77.
- [125] Ma J, Cao X. Regulation of Stat3 nuclear import by importin alpha5 and importin alpha7 via two different functional sequence elements. *Cell Signal.* 2006;18:1117-26.
- [126] Pumroy RA, Nardozi JD, Hart DJ, Root MJ, Cingolani G. Nucleoporin Nup50 stabilizes closed conformation of armadillo repeat 10 in importin alpha5. *J Biol Chem.* 2012;287:2022-31.
- [127] Gilchrist D, Mykytka B, Rexach M. Accelerating the rate of disassembly of karyopherin-cargo complexes. *J Biol Chem.* 2002;277:18161-72.
- [128] Kutay U, Bischoff FR, Kostka S, Kraft R, Gorlich D. Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell.* 1997;90:1061-71.
- [129] Gorlich D, Dabrowski M, Bischoff FR, Kutay U, Bork P, Hartmann E, et al. A novel class of RanGTP binding proteins. *J Cell Biol.* 1997;138:65-80.
- [130] Koepp DM, Wong DH, Corbett AH, Silver PA. Dynamic localization of the nuclear import receptor and its interactions with transport factors. *J Cell Biol.* 1996;133:1163-76.
- [131] Huber J, Cronshagen J, Kadokura M, Marshallsay C, Wada T, Sekine M, et al. Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J.* 1998;17:4114-26.
- [132] Strasser A, Dickmanns A, Luhrmann R, Ficner R. Structural basis for m3G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1. *EMBO J.* 2005;24:2235-43.
- [133] Kressler D, Bange G, Ogawa Y, Stjepanovic G, Bradatsch B, Pratte D, et al. Synchronizing nuclear import of ribosomal proteins with ribosome assembly. *Science.* 2012;338:666-71.
- [134] Calvino FR, Kharde S, Ori A, Hendricks A, Wild K, Kressler D, et al. Symportin 1 chaperones 5S RNP assembly during ribosome biogenesis by occupying an essential rRNA-binding site. *Nat Commun.* 2015;6.
- [135] Van Impe K, Hubert T, De Corte V, Vanloo B, Boucherie C, Vandekerckhove J, et al. A new role for nuclear transport factor 2 and Ran: nuclear import of CapG. *Traffic.* 2008;9:695-707.

- [136] Lu M, Zak J, Chen S, Sanchez-Pulido L, Severson DT, Endicott J, et al. A code for RanGDP binding in ankyrin repeats defines a nuclear import pathway. *Cell*. 2014;157:1130-45.
- [137] Sweitzer TD, Hanover JA. Calmodulin activates nuclear protein import: a link between signal transduction and nuclear transport. *Proc Natl Acad Sci U S A*. 1996;93:14574-9.
- [138] Hanover JA, Love DC, DeAngelis N, O'Kane ME, Lima-Miranda R, Schulz T, et al. The high mobility group box transcription factor Nhp6Ap enters the nucleus by a calmodulin-dependent, Ran-independent pathway. *J Biol Chem*. 2007;282:33743-51.
- [139] Argentaro A, Sim H, Kelly S, Preiss S, Clayton A, Jans DA, et al. A SOX9 defect of calmodulin-dependent nuclear import in campomelic dysplasia/autosomal sex reversal. *J Biol Chem*. 2003;278:33839-47.
- [140] Sharma M, Johnson M, Brocardo M, Jamieson C, Henderson BR. Wnt signaling proteins associate with the nuclear pore complex: implications for cancer. *Adv Exp Med Biol*. 2014;773:353-72.
- [141] Klose C, Viczian A, Kircher S, Schafer E, Nagy F. Molecular mechanisms for mediating light-dependent nucleocytoplasmic partitioning of phytochrome photoreceptors. *New Phytol*. 2015;206:965-71.
- [142] Jans DA, Briggs LJ, Gustin SE, Jans P, Ford S, Young IG. The cytokine interleukin-5 (IL-5) effects cotransport of its receptor subunits to the nucleus in vitro. *FEBS Lett*. 1997;410:368-72.
- [143] Liku ME, Nguyen VQ, Rosales AW, Irie K, Li JJ. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. *Mol Biol Cell*. 2005;16:5026-39.
- [144] Steidl S, Tuncher A, Goda H, Guder C, Papadopoulou N, Kobayashi T, et al. A single subunit of a heterotrimeric CCAAT-binding complex carries a nuclear localization signal: piggy back transport of the pre-assembled complex to the nucleus. *J Mol Biol*. 2004;342:515-24.
- [145] Zhang N, Wei P, Gong A, Chiu WT, Lee HT, Colman H, et al. FoxM1 promotes beta-catenin nuclear localization and controls Wnt target-gene expression and glioma tumorigenesis. *Cancer Cell*. 2011;20:427-42.
- [146] Baake M, Bauerle M, Doenecke D, Albig W. Core histones and linker histones are imported into the nucleus by different pathways. *Eur J Cell Biol*. 2001;80:669-77.

- [147] Baake M, Doenecke D, Albig W. Characterisation of nuclear localisation signals of the four human core histones. *J Cell Biochem.* 2001;81:333-46.
- [148] Mosammaparast N, Guo Y, Shabanowitz J, Hunt DF, Pemberton LF. Pathways mediating the nuclear import of histones H3 and H4 in yeast. *J Biol Chem.* 2002;277:862-8.
- [149] Mosammaparast N, Pemberton LF. Karyopherins: from nuclear-transport mediators to nuclear-function regulators. *Trends Cell Biol.* 2004;14:547-56.
- [150] Wagstaff KM, Jans DA. Importins and beyond: non-conventional nuclear transport mechanisms. *Traffic.* 2009;10:1188-98.
- [151] Kose S, Furuta M, Imamoto N. Hikeshi, a nuclear import carrier for Hsp70s, protects cells from heat shock-induced nuclear damage. *Cell.* 2012;149:578-89.
- [152] Song J, Kose S, Watanabe A, Son SY, Choi S, Hong H, et al. Structural and functional analysis of Hikeshi, a new nuclear transport receptor of Hsp70s. *Acta Crystallogr D Biol Crystallogr.* 2015;71:473-83.
- [153] Ribbeck K, Lipowsky G, Kent HM, Stewart M, Gorlich D. NTF2 mediates nuclear import of Ran. *EMBO J.* 1998;17:6587-98.
- [154] Wong DH, Corbett AH, Kent HM, Stewart M, Silver PA. Interaction between the small GTPase Ran/Gsp1p and Ntf2p is required for nuclear transport. *Mol Cell Biol.* 1997;17:3755-67.
- [155] Stewart M. Insights into the molecular mechanism of nuclear trafficking using nuclear transport factor 2 (NTF2). *Cell Struct Funct.* 2000;25:217-25.
- [156] Bullock TL, Clarkson WD, Kent HM, Stewart M. The 1.6 angstroms resolution crystal structure of nuclear transport factor 2 (NTF2). *J Mol Biol.* 1996;260:422-31.
- [157] Chaillan-Huntington C, Braslavsky CV, Kuhlmann J, Stewart M. Dissecting the interactions between NTF2, RanGDP, and the nucleoporin XFXFG repeats. *J Biol Chem.* 2000;275:5874-9.
- [158] Stewart M, Kent HM, McCoy AJ. Structural basis for molecular recognition between nuclear transport factor 2 (NTF2) and the GDP-bound form of the Ras-family GTPase Ran. *J Mol Biol.* 1998;277:635-46.
- [159] Bayliss R, Ribbeck K, Akin D, Kent HM, Feldherr CM, Gorlich D, et al. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J Mol Biol.* 1999;293:579-93.

- [160] Poon IK, Jans DA. Regulation of nuclear transport: central role in development and transformation? *Traffic*. 2005;6:173-86.
- [161] Pouton CW, Wagstaff KM, Roth DM, Moseley GW, Jans DA. Targeted delivery to the nucleus. *Adv Drug Deliv Rev*. 2007;59:698-717.
- [162] McLane LM, Corbett AH. Nuclear localization signals and human disease. *IUBMB Life*. 2009;61:697-706.
- [163] Hodel AE, Harreman MT, Pulliam KF, Harben ME, Holmes JS, Hodel MR, et al. Nuclear localization signal receptor affinity correlates with in vivo localization in *Saccharomyces cerevisiae*. *J Biol Chem*. 2006;281:23545-56.
- [164] Timney BL, Tetenbaum-Novatt J, Agate DS, Williams R, Zhang W, Chait BT, et al. Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo. *J Cell Biol*. 2006;175:579-93.
- [165] Biggar KK, Li SS. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol*. 2015;16:5-17.
- [166] Sadoul K, Wang J, Diagouraga B, Khochbin S. The tale of protein lysine acetylation in the cytoplasm. *J Biomed Biotechnol*. 2011;2011:970382.
- [167] Verdin E, Ott M. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol*. 2015;16:258-64.
- [168] Wei H, Mundade R, Lange KC, Lu T. Protein arginine methylation of non-histone proteins and its role in diseases. *Cell Cycle*. 2014;13:32-41.
- [169] Zhang X, Wen H, Shi X. Lysine methylation: beyond histones. *Acta Biochim Biophys Sin (Shanghai)*. 2012;44:14-27.
- [170] Jans DA. The regulation of protein transport to the nucleus by phosphorylation. *Biochem J*. 1995;311:705-16.
- [171] Jans DA, Hubner S. Regulation of protein transport to the nucleus: central role of phosphorylation. *Physiol Rev*. 1996;76:651-85.
- [172] Nardoizzi JD, Lott K, Cingolani G. Phosphorylation meets nuclear import: a review. *Cell Commun Signal*. 2010;8:32.
- [173] Kaffman A, Rank NM, O'Shea EK. Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes Dev*. 1998;12:2673-83.
- [174] Komeili A, O'Shea EK. Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science*. 1999;284:977-80.

- [175] Rona G, Borsos M, Ellis JJ, Mehdi AM, Christie M, Kornyei Z, et al. Dynamics of re-constitution of the human nuclear proteome after cell division is regulated by NLS-adjacent phosphorylation. *Cell Cycle*. 2014;13:3551-64.
- [176] Kosugi S, Hasebe M, Tomita M, Yanagawa H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A*. 2009;106:10171-6.
- [177] Harreman MT, Kline TM, Milford HG, Harben MB, Hodel AE, Corbett AH. Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. *J Biol Chem*. 2004;279:20613-21.
- [178] Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *J Biol Chem*. 2001;276:1317-25.
- [179] Hübner S, Xiao CY, Jans DA. The protein kinase CK2 site (Ser111/112) enhances recognition of the simian virus 40 large T-antigen nuclear localization sequence by importin. *J Biol Chem*. 1997;272:17191-5.
- [180] Kitamura R, Sekimoto T, Ito S, Harada S, Yamagata H, Masai H, et al. Nuclear import of Epstein-Barr virus nuclear antigen 1 mediated by NPI-1 (Importin alpha5) is up- and down-regulated by phosphorylation of the nuclear localization signal for which Lys379 and Arg380 are essential. *J Virol*. 2006;80:1979-91.
- [181] Nardozi J, Wenta N, Yasuhara N, Vinkemeier U, Cingolani G. Molecular basis for the recognition of phosphorylated STAT1 by importin alpha5. *J Mol Biol*. 2010;402:83-100.
- [182] Humbert-Lan G, Pieler T. Regulation of DNA binding activity and nuclear transport of B-Myb in *Xenopus* oocytes. *J Biol Chem*. 1999;274:10293-300.
- [183] Fagerlund R, Mélen K, Kinnunen L, Julkunen I. Arginine/lysine-rich nuclear localization signals mediate interactions between dimeric STATs and importin alpha 5. *J Biol Chem*. 2002;277:30072-8.
- [184] McBride KM, Banninger G, McDonald C, Reich NC. Regulated nuclear import of the STAT1 transcription factor by direct binding of importin-alpha. *EMBO J*. 2002;21:1754-63.
- [185] Wenta N, Strauss H, Meyer S, Vinkemeier U. Tyrosine phosphorylation regulates the partitioning of STAT1 between different dimer conformations. *Proc Natl Acad Sci U S A*. 2008;105:9238-43.

- [186] Meyer T, Begitt A, Lodige I, van Rossum M, Vinkemeier U. Constitutive and IFN-gamma-induced nuclear import of STAT1 proceed through independent pathways. *EMBO J.* 2002;21:344-54.
- [187] Huxford T, Huang D-B, Malek S, Ghosh G. The crystal structure of the I κ B α /NF- κ B complex reveals mechanisms of NF- κ B inactivation. *Cell.* 1998;95:759-70.
- [188] Jacobs MD, Harrison SC. Structure of an I κ B α /NF- κ B complex. *Cell.* 1998;95:749-58.
- [189] Fagerlund R, Kinnunen L, Kohler M, Julkunen I, Melen K. NF- κ B is transported into the nucleus by importin α 3 and importin α 4. *J Biol Chem.* 2005;280:15942-51.
- [190] Beg AA, Ruben SM, Scheinman RI, Haskill S, Rosen CA, Baldwin AS, Jr. I κ B interacts with the nuclear localization sequences of the subunits of NF- κ B: a mechanism for cytoplasmic retention. *Genes Dev.* 1992;6:1899-913.
- [191] Henkel T, Machleidt T, Alkalay I, Kronke M, Ben-Neriah Y, Baeuerle PA. Rapid proteolysis of I κ B α is necessary for activation of transcription factor NF- κ B. *Nature.* 1993;365:182-5.
- [192] Oeckinghaus A, Hayden MS, Ghosh S. Crosstalk in NF- κ B signaling pathways. *Nat Immunol.* 2011;12:695-708.
- [193] Asada M, Ohmi K, Delia D, Enosawa S, Suzuki S, Yuo A, et al. Brap2 functions as a cytoplasmic retention protein for p21 during monocyte differentiation. *Mol Cell Biol.* 2004;24:8236-43.
- [194] Fulcher AJ, Roth DM, Fatima S, Alvisi G, Jans DA. The BRCA-1 binding protein BRAP2 is a novel, negative regulator of nuclear import of viral proteins, dependent on phosphorylation flanking the nuclear localization signal. *FASEB J.* 2010;24:1454-66.
- [195] Forwood JK, Harley V, Jans DA. The C-terminal nuclear localization signal of the sex-determining region Y (SRY) high mobility group domain mediates nuclear import through importin beta 1. *J Biol Chem.* 2001;276:46575-82.
- [196] Thevenet L, Mejean C, Moniot B, Bonneaud N, Galeotti N, Aldrian-Herrada G, et al. Regulation of human SRY subcellular distribution by its acetylation/deacetylation. *EMBO J.* 2004;23:3336-45.

- [197] Nishino TG, Miyazaki M, Hoshino H, Miwa Y, Horinouchi S, Yoshida M. 14-3-3 regulates the nuclear import of class IIa histone deacetylases. *Biochem Biophys Res Commun.* 2008;377:852-6.
- [198] Harrison BC, Huynh K, Lundgaard GL, Helmke SM, Perryman MB, McKinsey TA. Protein kinase C-related kinase targets nuclear localization signals in a subset of class IIa histone deacetylases. *FEBS Lett.* 2010;584:1103-10.
- [199] Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science.* 2009;325:834-40.
- [200] Guo A, Gu H, Zhou J, Mulhern D, Wang Y, Lee KA, et al. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol Cell Proteomics.* 2014;13:372-87.
- [201] Moore KE, Carlson SM, Camp ND, Cheung P, James RG, Chua KF, et al. A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Mol Cell.* 2013;50:444-56.
- [202] Bannister AJ, Miska EA, Gorlich D, Kouzarides T. Acetylation of importin-alpha nuclear import factors by CBP/p300. *Curr Biol.* 2000;10:467-70.
- [203] Ryan CM, Harries JC, Kindle KB, Collins HM, Heery DM. Functional interaction of CREB binding protein (CBP) with nuclear transport proteins and modulation by HDAC inhibitors. *Cell Cycle.* 2006;5:2146-52.
- [204] Yang XJ. Lysine acetylation and the bromodomain: a new partnership for signaling. *Bioessays.* 2004;26:1076-87.
- [205] di Bari MG, Ciuffini L, Mingardi M, Testi R, Soddu S, Barila D. c-Abl acetylation by histone acetyltransferases regulates its nuclear-cytoplasmic localization. *EMBO Rep.* 2006;7:727-33.
- [206] Dietschy T, Shevelev I, Pena-Diaz J, Huhn D, Kuenzle S, Mak R, et al. p300-mediated acetylation of the Rothmund-Thomson-syndrome gene product RECQL4 regulates its subcellular localization. *J Cell Sci.* 2009;122:1258-67.
- [207] Bonaldi T, Talamo F, Scaffidi P, Ferrera D, Porto A, Bachi A, et al. Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *EMBO J.* 2003;22:5551-60.
- [208] Shimazu T, Horinouchi S, Yoshida M. Multiple histone deacetylases and the CREB-binding protein regulate pre-mRNA 3'-end processing. *J Biol Chem.* 2007;282:4470-8.

- [209] Madison DL, Yaciuk P, Kwok RP, Lundblad JR. Acetylation of the adenovirus-transforming protein E1A determines nuclear localization by disrupting association with importin- α . *J Biol Chem*. 2002;277:38755-63.
- [210] Blanco-Garcia N, Asensio-Juan E, de la Cruz X, Martinez-Balbas MA. Autoacetylation regulates P/CAF nuclear localization. *J Biol Chem*. 2009;284:1343-52.
- [211] Wong K, Zhang J, Awasthi S, Sharma A, Rogers L, Matlock EF, et al. Nerve growth factor receptor signaling induces histone acetyltransferase domain-dependent nuclear translocation of p300/CREB-binding protein-associated factor and hGCN5 acetyltransferases. *J Biol Chem*. 2004;279:55667-74.
- [212] Song EH, Oh W, Ulu A, Carr HS, Zuo Y, Frost JA. Acetylation of the RhoA GEF Net1A controls its subcellular localization and activity. *J Cell Sci*. 2015;128:913-22.
- [213] Soutoglou E, Katrakili N, Talianidis I. Acetylation regulates transcription factor activity at multiple levels. *Mol Cell*. 2000;5:745-51.
- [214] Liang SH, Clarke MF. Regulation of p53 localization. *Eur J Biochem*. 2001;268:2779-83.
- [215] O'Brate A, Giannakakou P. The importance of p53 location: nuclear or cytoplasmic zip code? *Drug Resist Updat*. 2003;6:313-22.
- [216] Kawaguchi Y, Ito A, Appella E, Yao TP. Charge modification at multiple C-terminal lysine residues regulates p53 oligomerization and its nucleus-cytoplasm trafficking. *J Biol Chem*. 2006;281:1394-400.
- [217] Rodriguez JA, Span SW, Ferreira CG, Kruyt FA, Giaccone G. CRM1-mediated nuclear export determines the cytoplasmic localization of the antiapoptotic protein Survivin. *Exp Cell Res*. 2002;275:44-53.
- [218] Wang H, Holloway MP, Ma L, Cooper ZA, Riolo M, Samkari A, et al. Acetylation directs survivin nuclear localization to repress STAT3 oncogenic activity. *J Biol Chem*. 2010;285:36129-37.
- [219] Lv L, Xu YP, Zhao D, Li FL, Wang W, Sasaki N, et al. Mitogenic and oncogenic stimulation of K433 acetylation promotes PKM2 protein kinase activity and nuclear localization. *Mol Cell*. 2013;52:340-52.
- [220] Oudhoff M, Freeman SA, Couzens AL, Antignano F, Kuznetsova E, Min PH, et al. Control of the hippo pathway by Set7-dependent methylation of Yap. *Dev Cell*. 2013;26:188-94.

- [221] Cho HS, Shimazu T, Toyokawa G, Daigo Y, Maehara Y, Hayami S, et al. Enhanced HSP70 lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B. *Nat Commun.* 2012;3:1072.
- [222] Smith WA, Schurter BT, Wong-Staal F, David M. Arginine methylation of RNA helicase a determines its subcellular localization. *J Biol Chem.* 2004;279:22795-8.
- [223] Zhao LJ, Subramanian T, Zhou Y, Chinnadurai G. Acetylation by p300 regulates nuclear localization and function of the transcriptional corepressor CtBP2. *J Biol Chem.* 2006;281:4183-9.
- [224] Lin X, Sun B, Liang M, Liang YY, Gast A, Hildebrand J, et al. Opposed regulation of corepressor CtBP by SUMOylation and PDZ binding. *Mol Cell.* 2003;11:1389-96.
- [225] Riefler GM, Firestein BL. Binding of neuronal nitric-oxide synthase (nNOS) to carboxyl-terminal-binding protein (CtBP) changes the localization of CtBP from the nucleus to the cytosol: a novel function for targeting by the PDZ domain of nNOS. *J Biol Chem.* 2001;276:48262-8.
- [226] Barnes CJ, Vadlamudi RK, Mishra SK, Jacobson RH, Li F, Kumar R. Functional inactivation of a transcriptional corepressor by a signaling kinase. *Nat Struct Biol.* 2003;10:622-8.
- [227] Pickard A, Wong PP, McCance DJ. Acetylation of Rb by PCAF is required for nuclear localization and keratinocyte differentiation. *J Cell Sci.* 2010;123:3718-26.
- [228] Chen BB, Mallampalli RK. Masking of a nuclear signal motif by monoubiquitination leads to mislocalization and degradation of the regulatory enzyme cytidyltransferase. *Mol Cell Biol.* 2009;29:3062-75.
- [229] Marchenko ND, Hanel W, Li D, Becker K, Reich N, Moll UM. Stress-mediated nuclear stabilization of p53 is regulated by ubiquitination and importin-alpha3 binding. *Cell Death Differ.* 2010;17:255-67.
- [230] Howitt J, Lackovic J, Low LH, Naguib A, Macintyre A, Goh CP, et al. Ndfip1 regulates nuclear Pten import in vivo to promote neuronal survival following cerebral ischemia. *J Cell Biol.* 2012;196:29-36.
- [231] Trotman LC, Wang X, Alimonti A, Chen Z, Teruya-Feldstein J, Yang H, et al. Ubiquitination regulates PTEN nuclear import and tumor suppression. *Cell.* 2007;128:141-56.
- [232] Wang X, Trotman LC, Koppie T, Alimonti A, Chen Z, Gao Z, et al. NEDD4-1 is a proto-oncogenic ubiquitin ligase for PTEN. *Cell.* 2007;128:129-39.

- [233] Vethantham V, Rao N, Manley JL. Sumoylation regulates multiple aspects of mammalian poly(A) polymerase function. *Genes Dev.* 2008;22:499-511.
- [234] Du JX, Bialkowska AB, McConnell BB, Yang VW. SUMOylation regulates nuclear localization of Kruppel-like factor 5. *J Biol Chem.* 2008;283:31991-2002.
- [235] Truong K, Lee TD, Li B, Chen Y. Sumoylation of SAE2 C terminus regulates SAE nuclear localization. *J Biol Chem.* 2012;287:42611-9.
- [236] Anderson DD, Eom JY, Stover PJ. Competition between sumoylation and ubiquitination of serine hydroxymethyltransferase 1 determines its nuclear localization and its accumulation in the nucleus. *J Biol Chem.* 2012;287:4790-9.
- [237] MacFarlane AJ, Anderson DD, Flodby P, Perry CA, Allen RH, Stabler SP, et al. Nuclear localization of de novo thymidylate biosynthesis pathway is required to prevent uracil accumulation in DNA. *J Biol Chem.* 2011;286:44015-22.
- [238] Woeller CF, Anderson DD, Szebenyi DM, Stover PJ. Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. *J Biol Chem.* 2007;282:17623-31.
- [239] Dohner K, Nagel CH, Sodeik B. Viral stop-and-go along microtubules: taking a ride with dynein and kinesins. *Trends Microbiol.* 2005;13:320-7.
- [240] Leopold PL, Pfister KK. Viral strategies for intracellular trafficking: motors and microtubules. *Traffic.* 2006;7:516-23.
- [241] Giannakakou P, Sackett DL, Ward Y, Webster KR, Blagosklonny MV, Fojo T. p53 is associated with cellular microtubules and is transported to the nucleus by dynein. *Nat Cell Biol.* 2000;2:709-17.
- [242] Roth DM, Moseley GW, Glover D, Pouton CW, Jans DA. A microtubule-facilitated nuclear import pathway for cancer regulatory proteins. *Traffic.* 2007;8:673-86.
- [243] Lam MH, Thomas RJ, Loveland KL, Schilders S, Gu M, Martin TJ, et al. Nuclear transport of parathyroid hormone (PTH)-related protein is dependent on microtubules. *Mol Endocrinol.* 2002;16:390-401.
- [244] Roth DM, Moseley GW, Pouton CW, Jans DA. Mechanism of microtubule-facilitated "fast track" nuclear import. *J Biol Chem.* 2011;286:14335-51.
- [245] Lowe AR, Tang JH, Yassif J, Graf M, Huang WY, Groves JT, et al. Importin-beta modulates the permeability of the nuclear pore complex in a Ran-dependent manner. *ELife.* 2015;4.

- [246] Tanaka M, Nishi M, Morimoto M, Sugimoto T, Kawata M. Yellow fluorescent protein-tagged and cyan fluorescent protein-tagged imaging analysis of glucocorticoid receptor and importins in single living cells. *Endocrinology*. 2003;144:4070-9.
- [247] Echeverria PC, Picard D. Molecular chaperones, essential partners of steroid hormone receptors for activity and mobility. *Biochim Biophys Acta*. 2010;1803:641-9.
- [248] Ni L, Llewellyn R, Kesler CT, Kelley JB, Spencer A, Snow CJ, et al. Androgen induces a switch from cytoplasmic retention to nuclear import of the androgen receptor. *Mol Cell Biol*. 2013;33:4766-78.
- [249] Falces J, Arregi I, Konarev PV, Urbaneja MA, Svergun DI, Taneva SG, et al. Recognition of nucleoplasmin by its nuclear transport receptor importin alpha/beta: insights into a complete import complex. *Biochemistry*. 2010;49:9756-69.
- [250] Rona G, Palinkas HL, Borsos M, Horvath A, Scheer I, Benedek A, et al. NLS copy-number variation governs efficiency of nuclear import--case study on dUTPases. *FEBS J*. 2014;281:5463-78.
- [251] Melchior F, Paschal B, Evans J, Gerace L. Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. *J Cell Biol*. 1993;123:1649-59.
- [252] Moore MS, Blobel G. The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature*. 1993;365:661-3.
- [253] Jakel S, Albig W, Kutay U, Bischoff FR, Schwamborn K, Doenecke D, et al. The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *EMBO J*. 1999;18:2411-23.
- [254] Dean KA, von Ahsen O, Gorlich D, Fried HM. Signal recognition particle protein 19 is imported into the nucleus by importin 8 (RanBP8) and transportin. *J Cell Sci*. 2001;114:3479-85.
- [255] Kuhlbrandt W. Biochemistry. The resolution revolution. *Science*. 2014;343:1443-4.
- [256] Schlichting I. Serial femtosecond crystallography: the first five years. *IUCr*. 2015;2:246-55.

Tables

Table 1. Structurally characterized transport receptors involved in protein import into the nucleus

| Transport receptor (abbreviation) | Other Names | Transport direction | Representative available structures (PDB entry) |
|------------------------------------|--|---------------------|--|
| Importin- β (Imp β 1) | Importin-90, karyopherin β -1, nuclear factor p97, pore-targeting complex 97 kDa subunit, (PTAC97), Kap95p (yeast; yImp β 1) | Import | <i>Apo</i> : yImp β 1 (3ND2) <i>Cargo complex</i> : Imp β 1:PTHRP (1M5N); Imp β 1:SREBP2 (1UKL); Imp β 1:SNAIL1 (3W5K) <i>Adaptor complex</i> : Imp β 1:Imp α -IBB domain (1QGK); Imp β 1:snurportin-IBB domain (3LWW) <i>Ran complex</i> : yImp β 1:RanGTP (2BKU); yImp β 1:RanGDP (3EA5) <i>Nup complex</i> : Imp β 1:GLFG (1O6P); Imp β 1:FxFG (1F59); yImp β 1:Nup1p (2BPT) |
| Transportin-1 (Trn1) | Importin- β 2, karyopherin- β 2, M9 region interaction protein (MIP); Kap104p (yeast; yTrn1) | Import | <i>Apo</i> : Trn1 (2QMR, 2Z5J) <i>Cargo complex</i> : Trn1:hnRNPA1 M9 NLS (2H4M); Trn1:Tap NLS (2Z5K); Trn1:hnRNP D NLS (2Z5N); Trn1:hnRNP M NLS (2OT8); Trn1:JKTBP NLS (2Z5O); Trn1:FUS NLS (4FQ3); Trn1:Nab2 NLS (4JLQ); Trn1:HCC1 (4OO6) <i>Ran complex</i> : Trn1:Ran (1QBK) |
| Transportin-3 (Trn3) | Importin-12, Transportin-SR; Kap111p (yeast, yTrn3) | Import | <i>Apo</i> : Trn3 (4COP) <i>Cargo complex</i> : Trn3:ASF/SF2 (4C0O) <i>Ran complex</i> : Trn3:Ran (4C0Q) |
| Importin-13 (Imp13) | Ran-binding protein 13 (RanBP13), karyopherin-13 (Kap13) | Bi-directional | <i>Apo</i> : Imp13 (3ZKV) <i>Cargo complex</i> : Imp13:Mago-Y14 (2X1G); Imp13:UBC9 (2XWU) <i>Ran complex</i> : Imp13:Ran (2X19) |
| Importin- α (Imp α) | Karyopherin- α (Kap α); Kap60p (yeast; yImp α) | Import | <i>Apo</i> : Imp α (1IAL); Supplementary Table 3 <i>Cargo complex</i> : Imp α :SV40-TAg NLS (1EJL, 1BK6); Imp α :nucleoplasmin NLS (1EE5, 3UL1); Imp α :PB2 (2JDQ); Imp α :CBP80 (3FEY); Imp α :VP24 (4U2X); Supplementary Table 2 <i>Nup complex</i> : yImp α :Nup2p (2C1T); Imp α :Nup50 (2C1M); Supplementary Table 2 |
| Snurportin-1 | | Import | <i>Cargo complex</i> : snurportin1 m3G-cap-binding domain:m3GpppG-cap dinucleotide (1XK5) |
| Symportin-1 | Syo1 (synchronized import) | Import | <i>Apo</i> : Syo1 (4GMO) <i>Cargo complex</i> : Syo1:RpL5:RpL11 (5AFF) |
| Hikeshi | | Import | <i>Apo</i> : Hikeshi (3WVZ) |

Figure legends

Fig. 1. Overview of the main nuclear import pathways. Schematic illustration of three different nuclear import pathways. (i)-(ii) classic import pathway; cargo (nucleoplasmin, shown in red – PDB entry 1K5J), Imp α (shown in green – PDB entry 1IAL and 1EE5) and Imp β 1 (shown in yellow - PDB entry 1QGK) form a ternary complex before translocating across the membrane via the nuclear pore (shown in grey); RanGTP (shown in blue – PDB entry 2BKU) and Cse1p [or CAS] (shown in orange – PDB entry 1WA5) dissociate the complex and release the cargo. (iii)-(iv) Snurportin-1 import pathway; cargo (U1A-UTR, shown in red and wheat – PDB entry 1AUD), snurportin-1 (shown in green – PDB entry 1UKL) and Imp β 1 translocate across the membrane; RanGTP and CRM1 (shown in orange – PDB entry 3GJX) release the cargo. (v)-(vi) Import pathway involving direct cargo binding to Imp β 1; cargo (SREBP-2, shown in red – PDB entry 1UKL) and Imp β 1 form a binary complex before translocating across the membrane; RanGTP dissociates the complex.

Fig. 2. Structures of Imp β 1. PDB entries: PTHrP complex (1M5N), SREBP2 complex (1UKL), Snail1 complex (3W5K), Imp α -IBB complex (1QGK), RanGTP complex (2BKU), GLFG complex (1O6P), FxFG complex (1F59), Nup1p complex (2BPT), apo (3ND2). The HEAT repeats involved in cargo binding are highlighted in dark yellow. A representative Imp β 1 from Imp β 1:Ran complex has all cargoes overlaid.

Fig. 3. Structures of (A) Ran:RanBD1:Imp β 1(1-462) [32] and (B) RanGppNHp:RanBD1:RanGAP complexes (PDB entry 1K5D). RanBD1 forms a molecular embrace with Ran, sequestering Ran's C-terminus. Superimposition of the full-length

RanGTP-Imp β structure onto the ternary complexes reveals steric clashes between Imp β and RanBD1 or RanGAP.

Fig. 4. Structures of Trn1. PDB entries: hnRNP A1 complex (2H4M), hnRNP D complex (2Z5N), TAP complex (2Z5K), JKTBP complex (2Z5O), FUS complex (4FQ3), Nab2 complex (4JLQ), HCC1 complex (4OO6), hnRNP M complex (2OT8), RanGppNHP complex (1QBK), apo (2Z5J). The HEAT repeats involved in cargo binding are highlighted in dark yellow. A representative Trn1 in apo-form has all cargoes overlaid.

Fig. 5. Structures of Trn3. PDB entries: ASF/SF2 complex (4C0O), Ran complex (4C0Q), apo (4C0P). The HEAT repeats involved in cargo binding are highlighted in dark yellow. A representative Trn3 in apo-form has all cargoes overlaid.

Fig. 6. Structures of Imp13. PDB entries: Mago-Y14 complex (2X1G), UBC9 complex (2XWU), Ran complex (2X19), apo (3ZKV). The HEAT repeats involved in cargo binding are highlighted in dark yellow. A representative Imp13 from Imp13:Ran complex has all cargoes overlaid.

Fig. 7. NLS binding by importin- α . (A) Structure of rice Imp α (PDB entry 4BQK) with H3 helices coloured in green. (B) Structure of mouse Imp α (PDB entry 3UL1) in complex with nucleoplasmin (Npl) cNLS (shown as black sticks). The mouse adaptor is coloured by sequence conservation based on known Imp α structures (human Imp α 1, human Imp α 3, human Imp α 5, human Imp α 7, *M. musculus* Imp α 1, *S. cerevisiae* Imp α , *O. sativa* Imp α , *A. thaliana* Imp α 3, *N. crassa* Imp α 1). (C) Schematic representation of a monopartite NLS binding at the Imp α major and minor site pockets. Conserved Asn and Trp residues of Imp α

shown in green and yellow, respectively. Monopartite NLS main-chain and side-chains are shown as black and blue lines, respectively. Dashed lines indicate common salt-bridge interactions at the P2 and P2'-binding cavities. (D) Structure of mouse Imp α with atypical minor site-binding Gu α NLS shown in blue. Imp α residues comprising (^R/K)XXWXXXN motif in ARM repeats 7 and 8 are shown in stick representation. Gu α residues that bind to minor site cavities are indicated. (E) Structure of full-length yeast Imp α (PDB entry 1WA5; IBB domain shown in green) superimposed onto yeast Imp α :Nup2p (PDB entry 2C1T; Nup2p shown in magenta) and yeast Imp α :nucleoplasmin cNLS (PDB entry 1EE5; NLS shown in orange) complexes. For clarity, only one Imp α ARM repeat domain is shown in grey surface representation.

Fig. 8. The structure of symportin-1. (A) Structure of the *C. thermophilum* Syo1 adaptor (PDB entry 4GMO). The protein has an extended superhelical conformation comprised of a unique chimera of 4 ARM (residues 65-260) and 6 HEAT (residues 274-675) repeats. (B) Structure of the Syo1:Rpl5:Rpl11 complex (PDB entry 5AFF). The Rpl5 peptide binds along the inner solenoid surface, while Rpl11 interacts with the outer surface of the Syo1 superhelix as well as a helical region from the HEAT repeat 1 acidic loop.

Fig. 9. NTF2 dimer (blue) bound to two chains of RanGDP (green) and two FxFG nucleoporin motif cores (yellow) (based on PDB entries 5BXQ and 1GYB). (A) The two chains of the NTF2 dimer interact through an extensive β -sheet, whereas the remainder of the molecule generates a cavity into which Phe⁷² of the RanGDP switch-II loop binds. The FxFG motif cores bind in a hydrophobic cavity generated between the two NTF2 chains. (B) Binding of the RanGDP switch-II loop to NTF2. Ran Phe⁷² inserts into the hydrophobic

cavity and is supplemented by salt bridges formed between Lys⁷¹ and Arg⁷⁶ of Ran and Asp^{92/94} and Glu⁴² of NTF2, respectively.

Fig. 10. Regulatory mechanisms in nucleo-cytoplasmic trafficking. The control of trafficking in and out of the nucleus at the protein level operates by several types of mechanisms. First, different chemical moieties can be attached to the cargo or transport factors, as shown in the first row. These post-translational modifications may alter the thermodynamics and kinetics of the interactions between the cargo and the karyopherin (A) or may lead to masking of the interacting groups (B and C). Other mechanisms involve the microtubular system, which can influence the concentration gradient of cargo proteins such that these may accumulate around the nuclear pores where importins are readily available (D). “Piggy-backing” is the indirect coupling between cargo and karyopherins (E). NLS-copy number variation for oligomeric cargo proteins constitutes a fine-tuning control that provides advantage to homo-oligomers or hetero-oligomers with an increased number of NLS segments (F).

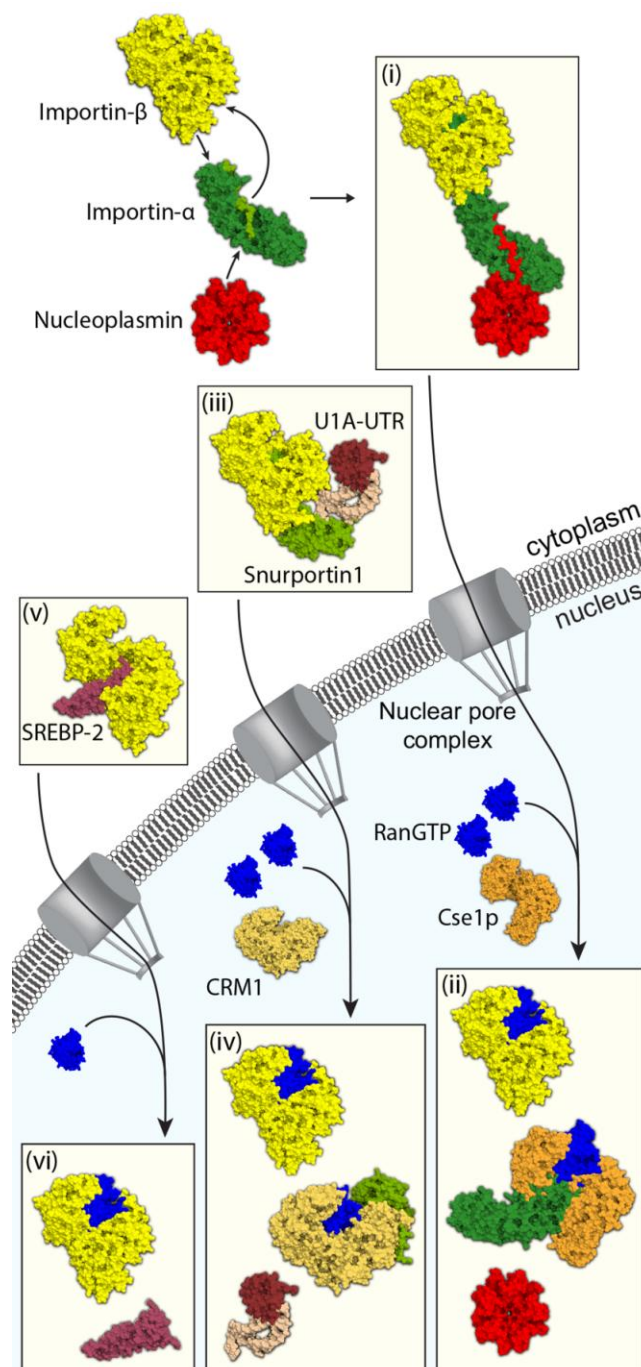


Figure 1

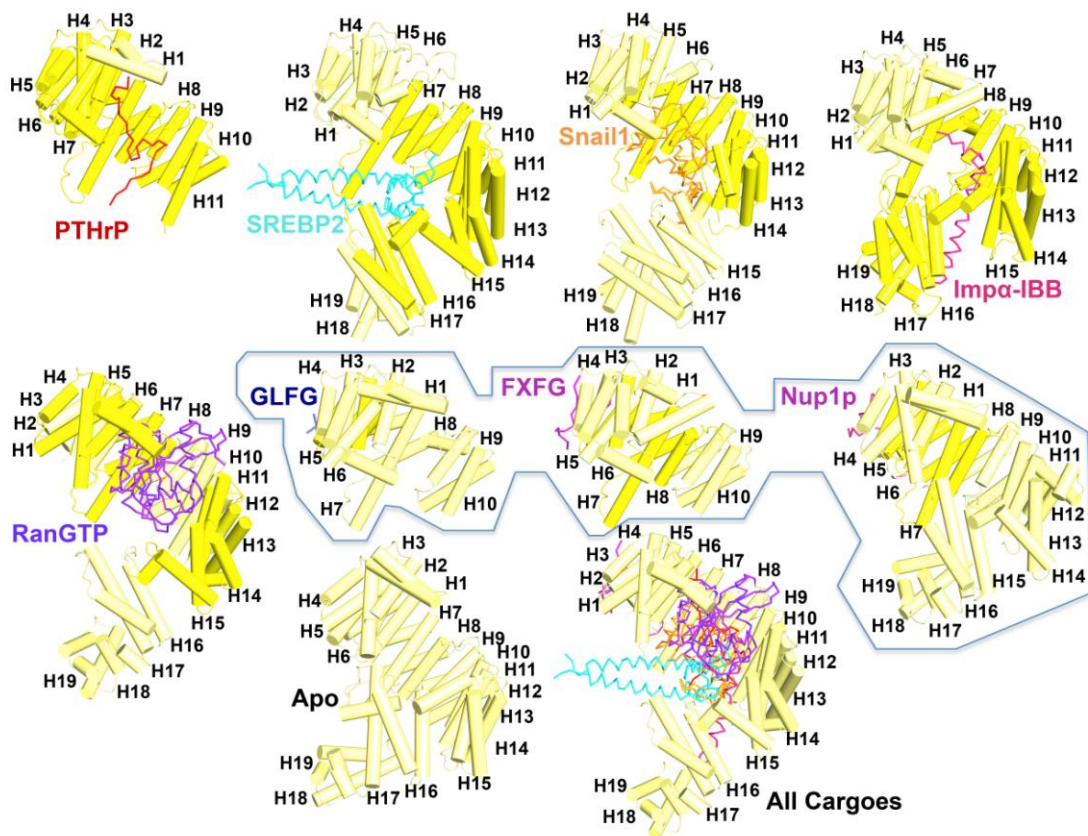


Figure 2

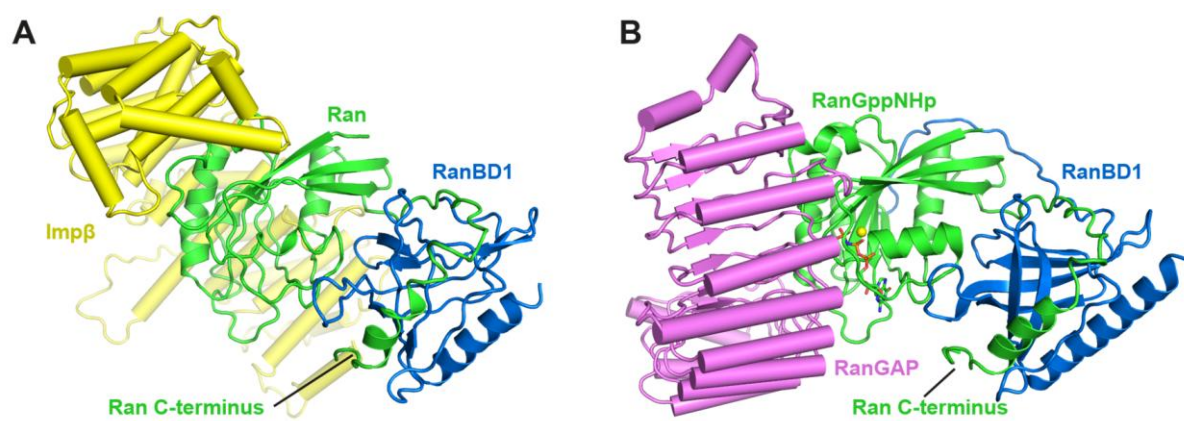


Figure 3

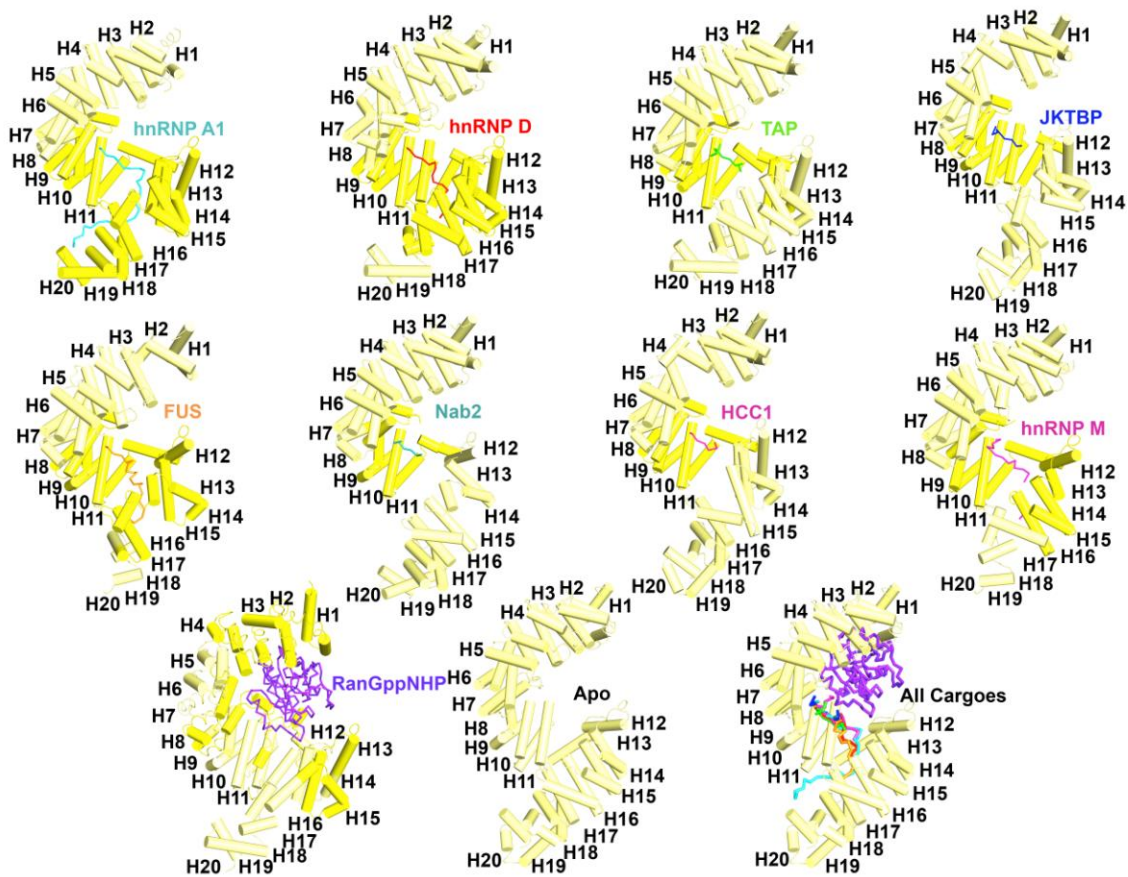


Figure 4

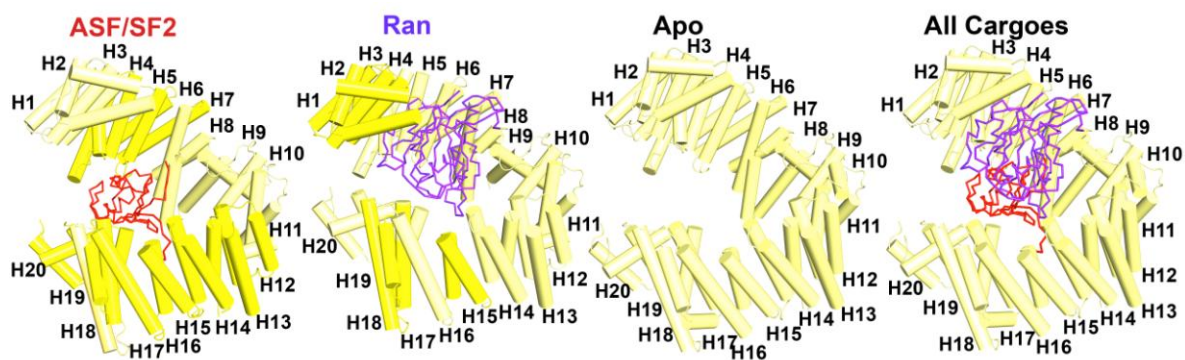


Figure 5

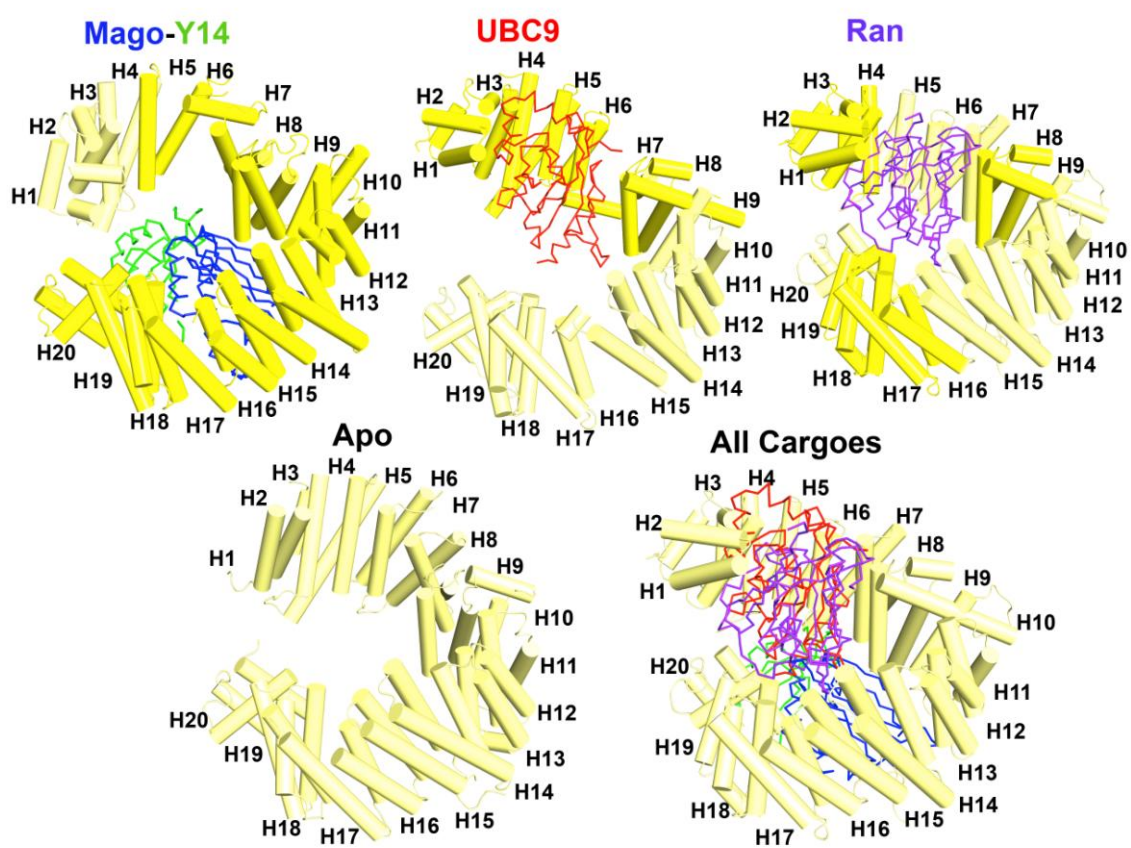


Figure 6

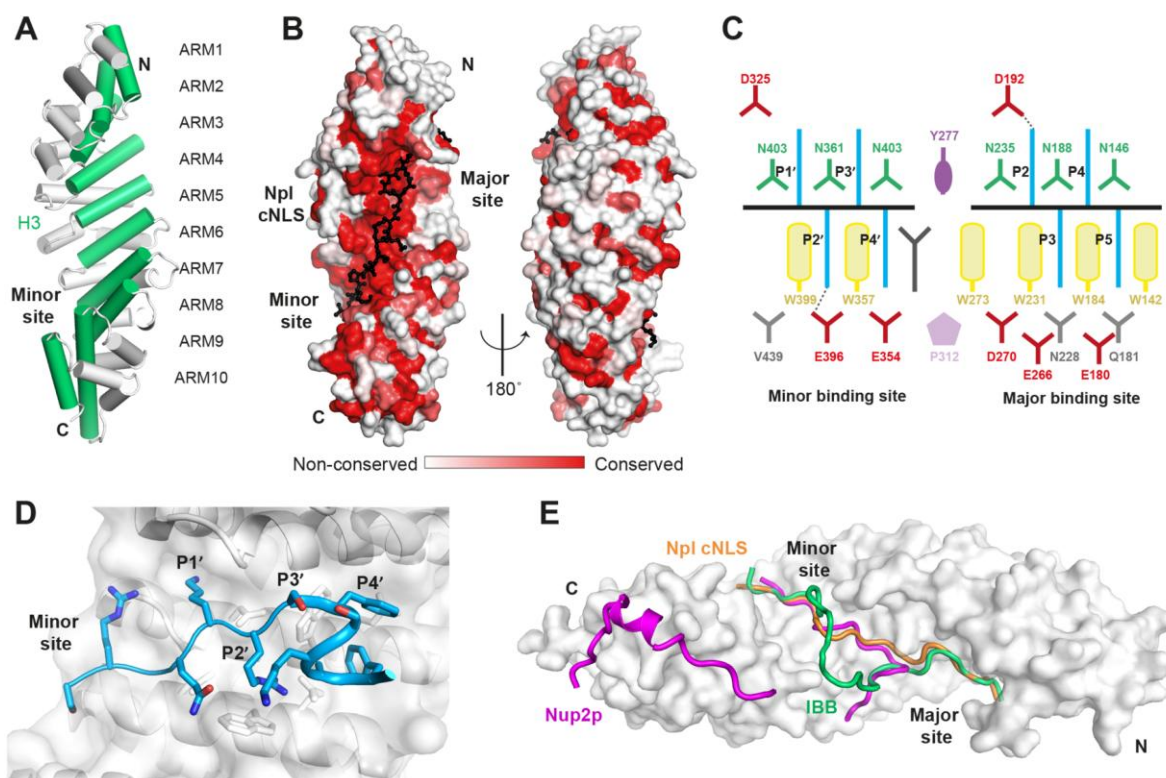


Figure 7

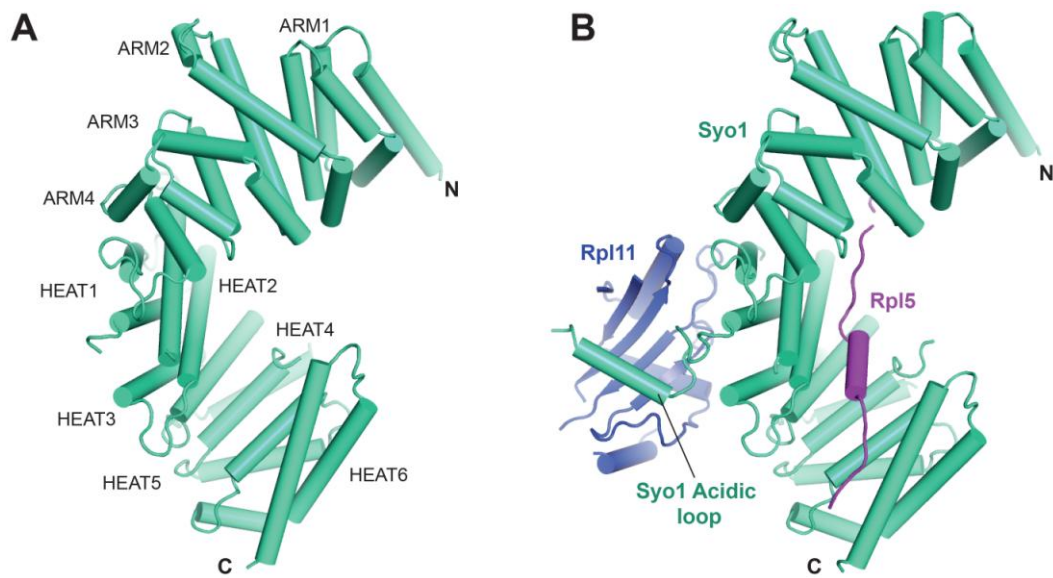


Figure 8

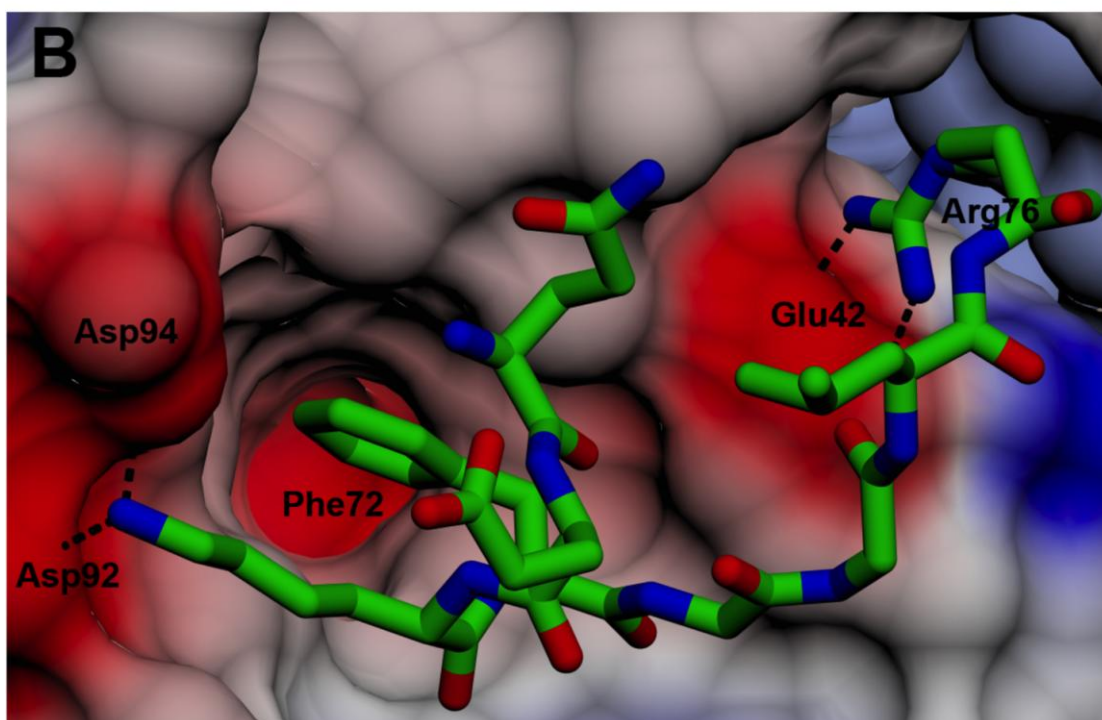
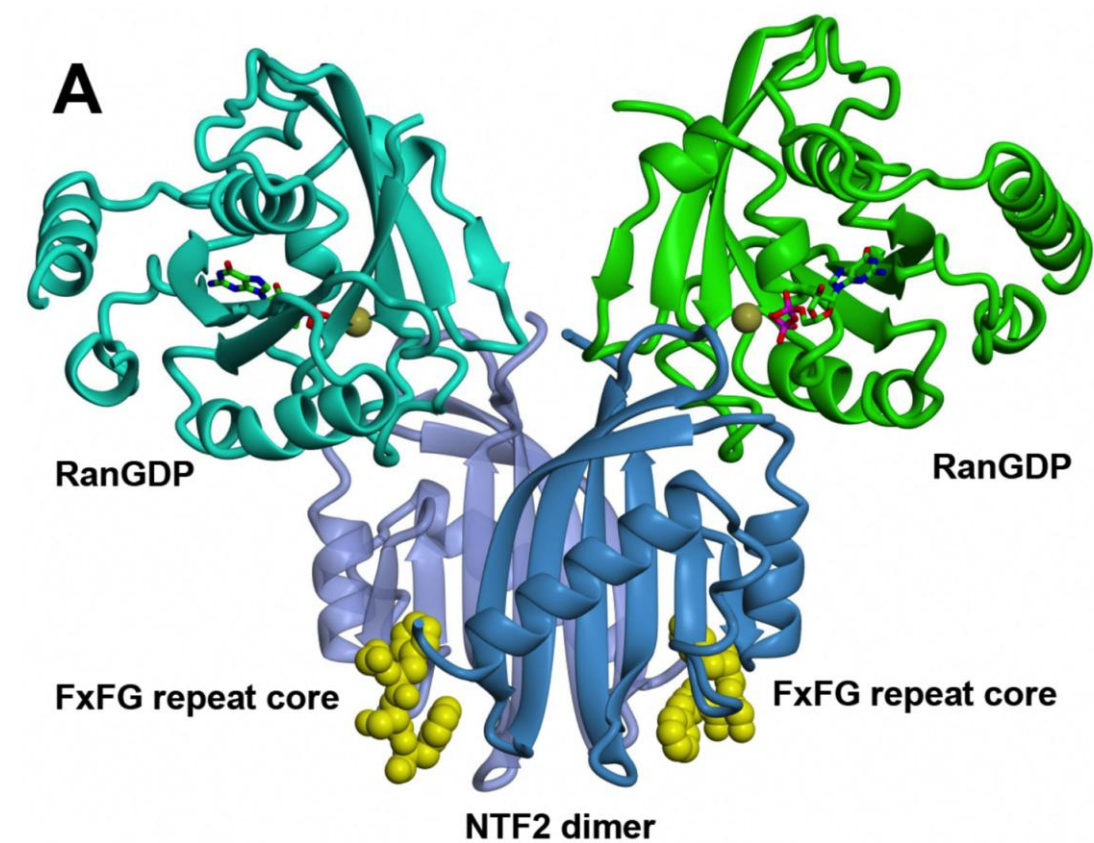


Figure 9

Post-translational modification-regulated transport

P phosphorylation
 Ac acetylation
 Me methylation
 Ub ubiquitination
 Su sumoylation

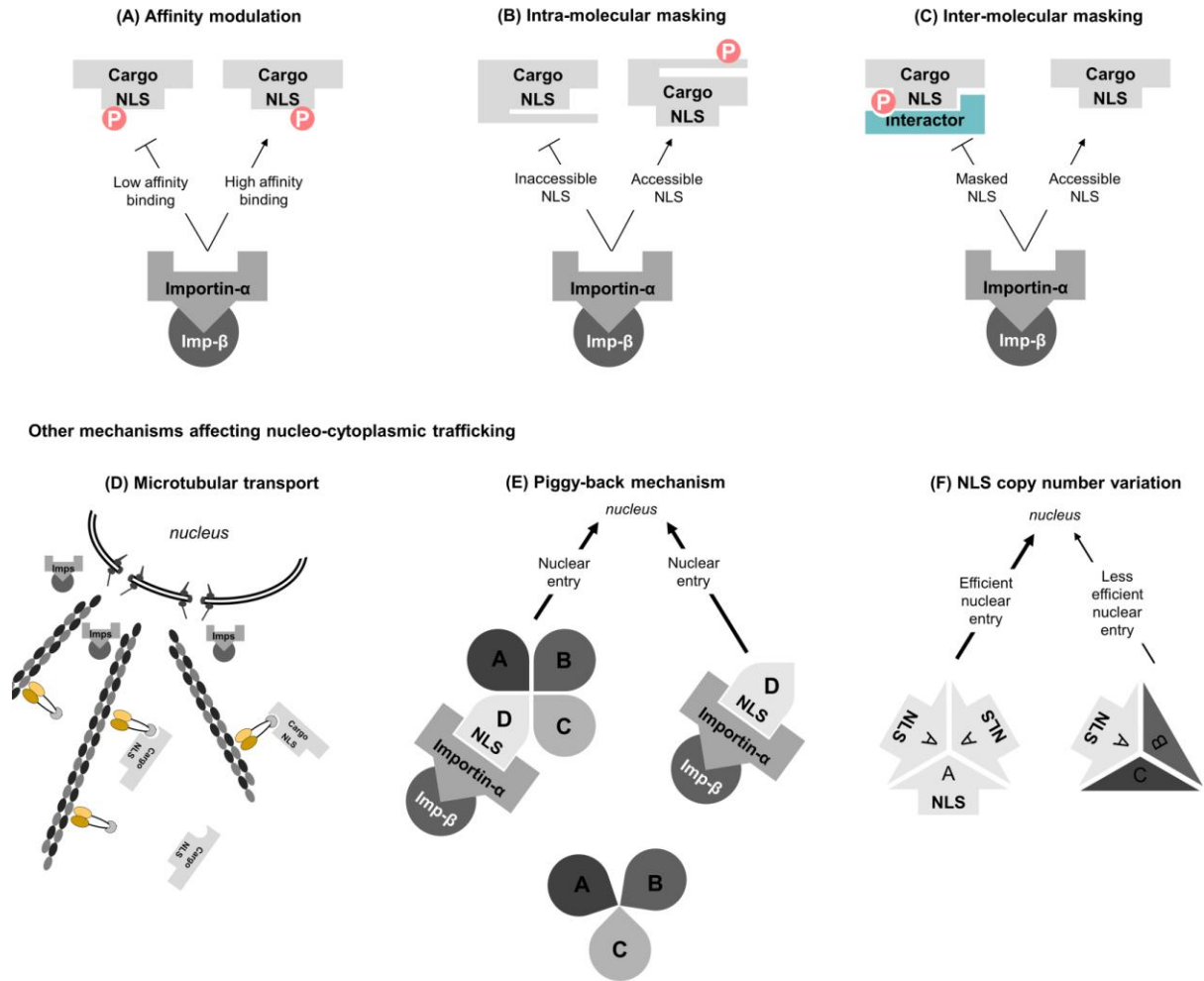
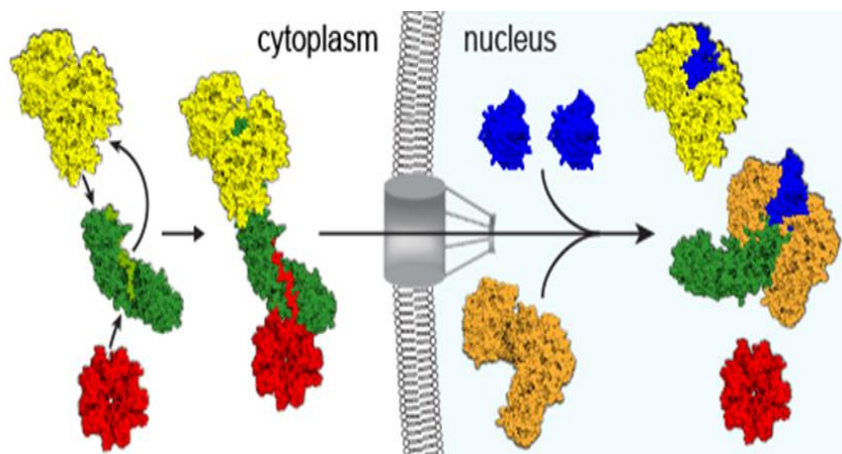


Figure 10



Graphical abstract