Crystal structure of the Xpo1p nuclear export complex bound to the SxFG/PxFG repeats of the nucleoporin Nup42p

Masako Koyama[1a], Hidemi Hirano[1], Natsuki Shirai[1] and Yoshiyuki Matsuura[1, 2*]

[1]Division of Biological Science, and [2]Structural Biology Research Center, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan.

[*]*Correspondence*: matsuura.yoshiyuki@d.mbox.nagoya-u.ac.jp

[a]*Present address*: Laboratory of Structural Biology, Research Institute of Advanced Science and Engineering, Waseda University, Tokyo 162-8480, Japan.

Running title: Xpo1p-Nup42p interactions

Abstract

Xpo1p (yeast CRM1) is the major nuclear export receptor that carries a plethora of proteins and ribonucleoproteins from the nucleus to cytoplasm. The passage of the Xpo1p nuclear export complex through nuclear pore complexes (NPCs) is facilitated by interactions with nucleoporins (Nups) containing extensive repeats of phenylalanine-glycine (so called FG repeats), although the precise role of each Nup in the nuclear export reaction remains incompletely understood. Here we report structural and biochemical characterization of the interactions between the Xpo1p nuclear export complex and the FG repeats of Nup42p, a nucleoporin localized at the cytoplasmic face of yeast NPCs and has characteristic SxFG/PxFG sequence repeat motif. The crystal structure of Xpo1p-PKI-Nup42p-Gsp1p-GTP complex identified three binding sites for the SxFG/PxFG repeats on HEAT repeats 14-20 of Xpo1p. Mutational analyses of Nup42p demonstrated that the conserved serines and prolines in the SxFG/PxFG repeats contribute to Xpo1p-Nup42p binding. Our structural and biochemical data suggest that SxFG/PxFG-Nups such as Nup42p and Nup159p at the cytoplasmic face of NPCs provide high-affinity docking sites for the Xpo1p nuclear export complex in the terminal stage of NPC passage and that subsequent disassembly of the nuclear export complex facilitates recycling of free Xpo1p back to the nucleus.

Introduction

In eukaryotes, nuclear pore complexes (NPCs) perforate the nuclear envelope and mediate nucleocytoplasmic exchange of macromolecules (reviewed in Gorlich & Kutay 1999; Wente & Rout 2010). NPCs are large (60 MDa in yeast, 120 MDa in vertebrates) macromolecular assemblies that are constructed from multiple copies of ~30 different protein subunits collectively termed nucleoporins (Nups) in both yeast (Rout *et al*. 2000) and vertebrates (Cronshaw *et al*. 2002). Throughout eukaryotes, NPCs have a modular architecture (reviewed in Knockenhauer & Schwartz 2016; Hoelz *et al*. 2016; Beck & Hurt 2017). About half of ~30 Nups are structured and generate a scaffolding core of NPCs that has a three-ringed, sandwiched architecture (the cytoplasmic ring, the inner ring, and the nuclear ring) with eight-fold rotational symmetry, the axis of which lies perpendicular to the plane of the nuclear envelope. The symmetric core recruits transport-channel Nups containing intrinsically disordered non-globular domains rich in phenylalanine-glycine (FG) repeat sequences (the so-called FG domains) that fill a ~40 nm central channel of NPCs through which macromolecules are transported. In addition, the symmetric core of NPCs is decorated on the cytoplasmic and nuclear faces with peripheral components such as the nuclear basket.

Approximately one third of all Nups contain a conserved FG repeat sequence motif. The FG repeat sequences can be classified based on conserved amino acids in the neighboring residues of the core FG sequence. The four major types of FG repeat motifs are FxFG, GLFG, SxFG, and PxFG motifs, where x is any amino acid. These motifs are not evenly distributed within NPCs. In general, SxFG/PxFG repeats are mainly found in Nups located at the cytoplasmic face of NPCs (such as Nup42p and Nup159p in yeast; Fig. 1A); GLFG repeats are present in Nups forming the central channel of NPCs; and FxFG repeats are enriched in nuclear Nups. The FG repeat motifs are separated by

spacers of variable length, with a typical spacing of ~20 amino acids. The sequences of the spacers between FG repeats do not appear to be particularly conserved, but are rich in hydrophilic residues. A subset of FG-Nups have some charged amino acids in the spacers between FG repeats, while other FG-Nups have relatively uncharged spacers. The FG-Nups with uncharged spacers tend to be cohesive via hydrophobic attractions, while the FG-Nups with charged spacers are noncohesive, and thus the FG-Nups can be classified into two categories depending on cohesiveness (Patel *et al*. 2007; Yamada *et al*. 2010).

NPCs are equipped with a permeability barrier that restricts passive diffusion of macromolecules. Although water, ions, metabolites, and small proteins diffuse freely through NPCs, the rate of passive diffusion of macromolecules through NPCs decreases with increasing molecular mass of the diffusing molecule (Mohr et al. 2009; Timney *et al*. 2016). However, large macromolecules can be transported rapidly through NPCs if the macromolecules are recognized and escorted by appropriate nuclear transport receptors (NTRs) that are able to penetrate the permeability barrier of NPCs (reviewed in Gorlich & Kutay 1999; Wente & Rout 2010). The FG domains of various Nups are involved in forming the selective permeability barrier of NPCs (Patel *et al*. 2007; Timney *et al*. 2016; Strawn *et al*. 2004; Frey & Gorlich 2007; Hulsmann *et al*. 2012), and, although the precise mechanism of how the FG domains form the selective barrier to transport remains incompletely understood, all known NTRs interact with the FG domains in such a way that allows facilitated diffusion of the NTR-cargo complexes through the selective permeability barrier of NPCs (reviewed in Wente & Rout 2010; Schmidt & Gorlich 2016; Stewart 2007; Rout *et al*. 2003).

The majority of NTRs belong to the karyopherin-β (Kap-β) family of proteins that can be classified into importins (nuclear import receptors), exportins (nuclear

export receptors), and bidirectional Kap-βs that can mediate both nuclear import and export. The active nuclear transport mediated by Kap-βs requires input of energy, which is provided by the Ran GTPase system. Ran (Gsp1p in yeast) cycles between its GTP- and GDP-bound states and functions as a molecular switch (reviewed in Vetter & Wittinghofer 2001). In the nucleus, Ran exists primarily in its GTP-bound state, because the chromatin-bound protein RCC1 (Prp20p in yeast), the guanine nucleotide exchange factor for Ran, is localized exclusively in the nucleus. In the cytoplasm, GTP bound to Ran is rapidly hydrolyzed, due to the actions of RanGAP (Rna1p in yeast) and its co-activators RanBP1 (Yrb1p in yeast) and RanBP2. Consequently, Ran returns to the resting GDP-bound state in the cytoplasm.

CRM1 (for chromosome region maintenance 1 (Adachi & Yanagida 1989); Xpo1p in yeast) was one of the first exportins discovered (Fornerod *et al*. 1997; Fukuda *et al*. 1997; Stade *et al*. 1997; Ossareh-Nazari *et al*. 1997). Among exportins, CRM1 is exceedingly promiscuous and carries hundreds of proteins and RNA-protein complexes from the nucleus to cytoplasm (Kirli *et al*. 2015). In most cases, CRM1 recognizes cargoes through short linear nuclear export signals (NESs), which harbor 4-5 critical hydrophobic residues with various spacings (Wen *et al*. 1995; Fischer *et al*. 1995; Dong *et al*. 2009; Monecke *et al*. 2009; Guttler *et al*. 2010; Fung *et al*. 2015; Fung *et al*. 2017). CRM1 binds cargoes cooperatively with Ran-GTP in the nucleus. The assembly of the CRM1-cargo-Ran-GTP complex (the CRM1 nuclear export complex) can be accelerated by a predominantly nuclear Ran-binding protein RanBP3 (Yrb2p in yeast) (Koyama *et al*. 2014), which functions as a CRM1 (Xpo1p)-specific nuclear export cofactor (Taura *et al*. 1998; Noguchi *et al*. 1999; Englmeier *et al*. 2001; Lindsay *et al*. 2001). After passing through NPCs, the CRM1 nuclear export complex is disassembled in the cytoplasm. The RanBD of RanBP1 (Yrb1p in yeast) or RanBP2 accelerates the release of NES-containing cargoes from CRM1 (Xpo1p) (Koyama & Matsuura 2010),

and then RanGAP (Rna1p) accelerates the GTP hydrolysis by Ran (Gsp1p), rendering the disassembly reaction virtually irreversible. The free CRM1 returns to the nucleus. Ran-GDP is imported back to the nucleus by NTF2 (nuclear transport factor 2) (Ribbeck *et al*. 1998; Smith *et al*. 1998), and recharged with GTP due to the action of RCC1 in the nucleus.

Crystal structures of unliganded CRM1 (Saito & Matsuura 2013; Monecke *et al*. 2013), of CRM1-cargo-Ran-GTP complexes and their assembly- and disassembly intermediates (Dong *et al*. 2009; Monecke *et al*. 2009; Guttler *et al*. 2010; Fung *et al*. 2015; Fung *et al*. 2017; Koyama *et al*. 2014; Koyama & Matsuura 2010) have been determined. These structures have provided rich insights into the molecular mechanism of how cargo is loaded onto CRM1 in the nucleus and how cargo is unloaded in the cytoplasm (reviewed in Matsuura 2016). CRM1 (Xpo1p) is a toroid-shaped molecule constructed from a tandem array of 21 HEAT repeats, each of which consists of two antiparallel α-helices (designated A- and B-helices) connected by loops of varying length. Ran-GTP binds to the interior surface of CRM1, whereas the NES binds to a hydrophobic cleft on the outer surface of CRM1 formed between HEAT repeats 11 and 12. A long β-hairpin loop of CRM1 (termed "HEAT9 loop" (Koyama & Matsuura 2010) or "acidic loop" (Monecke *et al*. 2009)) plays an important role in the regulation of cargo loading and unloading.

The crystal structure of the Xpo1p-Yrb2p-Gsp1p-GTP complex, an assembly intermediate of the Xpo1p nuclear export complex in yeast, revealed for the first time how FG repeats bind the exportin Xpo1p (Koyama *et al*. 2014). Similar mode of FG-repeat recognition was subsequently observed in the structure of CRM1-Spn1-Nup214-Ran-GTP complex (Port *et al*. 2015), which revealed two additional FG-repeat binding sites. In the present study we report crystal structure of the

Xpo1p-PKI-Nup42p-Gsp1p-GTP complex that provides new insights into previously uncharacterized interactions between the SxFG/PxFG repeats and Xpo1p. Together with our biochemical characterization of Nup42p, the structure reported here extends our understanding of the role of the SxFG/PxFG-Nups at the cytoplasmic face of NPCs in the terminal stage of nuclear export. The structure presented here also provides insights into evolutionary conservation of the FG repeat binding sites on CRM1 across species.

Results

Nup42p forms a stable complex with the Xpo1p-NES-Gsp1p-GTP complexes

Previously, solution binding assays and pull-down experiments from whole yeast cell extract showed that Xpo1p has a strong preference for the SxFG/PxFG-Nups on the cytoplasmic face of NPCs (Nup159p and Nup42p) versus those localized to the central channel or the nuclear face of NPCs (Allen *et al*. 2001; Allen *et al*. 2002). However, it has been controversial whether or not the binding of NES-cargo and Nup42p to Xpo1p are mutually exclusive. Yeast two hybrid assays indicated that Xpo1p bridges the interaction between Nup42p and HIV-1 Rev (an NES-cargo) (Neville *et al*. 1997), whereas solution binding assays suggested that the binding of Xpo1p-HIV-1 Rev-Ran-GTP complex to Nup42p results in release of HIV-1 Rev and formation of a ternary Nup42p-Xpo1p-Ran-GTP complex (Floer & Blobel 1999). We revisited this issue using GST pull-down assays. Purified GST-Nup42p (full-length) was immobilized on glutathione-Sepharose beads and incubated with Xpo1p in the presence or absence of various proteins (Gsp1p, NES-cargo, and Yrb1p) (Fig. 1B). Although Xpo1p alone bound only weakly to Nup42p (Fig. 1B, lane 1), Gsp1p-GTP substantially enhanced the binding of Xpo1p to Nup42p (Fig. 1B, lane 3). Likewise, the binding of Xpo1p to Nup42p was quite strong in the presence of both Gsp1p-GTP and NES-cargo (Fig. 1B, lanes 7, 12). Because NES-cargo alone did not bind Nup42p directly (Fig. 1B, lanes 5 and 10), the clearly visible bands of NES-cargo (PKI or HIV-1 Rev) in lanes 7 and 12 of Fig. 1B indicate that NES-cargo bind to Nup42p via Xpo1p. Thus, although it is unclear why our results disagree with the observation reported by Floer and Blobel (1999), our results strongly support the previously suggested notion that Xpo1p bridges the interaction between Nup42p and NES-cargo (Neville *et al*. 1997).

Strong binding of Xpo1p to Nup42p in the presence of Gsp1p-GTP (but not in the absence of Gsp1p) was also observed for Nup42p residues 31-213, which is a fragment of the FG domain of Nup42p (Fig. 1C, lanes 1 and 3). The binding of Xpo1p to this Nup42p fragment was also strong in the presence of both Gsp1p-GTP and NES-cargo (PKI or HIV-1 Rev) (Fig. 1C, lanes 7 and 12) and, the binding of NES-cargo (PKI or HIV-1 Rev) to the immobilized Nup42p fragment was observed only in the presence of both Xpo1p and Gsp1p-GTP (Fig. 1C, lanes 5-7, 10-12).

Using GST pull-down assay, we also found that Yrb1p (yeast RanBP1) weakens the binding of the Xpo1p nuclear export complex to Nup42p (Fig. 1B, lanes 9 and 14; Fig. 1C, lanes 9 and 14). Yrb1p was not detected in the bound fractions (Fig. 1B, lanes 9 and 14; Fig. 1C, lanes 9 and 14), indicating that the binding of Nup42p to Xpo1p-Yrb1p-Gsp1p-GTP complex, which is formed concomitant with Yrb1p-induced release of NES-cargo from Xpo1p and Gsp1p-GTP (Koyama & Matsuura 2010), is undetectably weak. These data suggest that Yrb1p prevents Xpo1p from being tightly trapped on the cytoplasmic face of NPCs in the terminal stage of nuclear export.

Crystal structure of the Xpo1p nuclear export complex bound to the SxFG/PxFG repeats of Nup42p

To elucidate the structural basis for the strong binding of the SxFG/PxFG repeats of Nup42p to the Xpo1p nuclear export complex, we attempted to cocrystallize Xpo1p-PKI-Gsp1p-GTP complex with Nup42p, but failed. As an alternative approach, we soaked crystals of Xpo1p-PKI-Gsp1p-GTP complex with a synthetic peptide that corresponds to Nup42p residues 88-122. We obtained an X-ray diffraction data set of the peptide-soaked crystals to 2.20 Å resolution using synchrotron radiation at SPring-8 and determined the structure by molecular replacement (Table 1; Fig. 2). The structure

was refined to free and working *R*-factor values of 22.0% and 19.1%, respectively, and the final model had excellent stereochemistry (Table 1).

Nup42p had unambiguously clear electron density at three sites (termed sites 1, 2 and 3) on the outer surface of HEAT repeats 14-20 of Xpo1p (Figs. 2A-2C), suggesting multivalent binding of the SxFG/PxFG repeats to Xpo1p. Site 1 was located on HEAT repeats 17-20 and coincides with one of the two FG-repeat binding sites on Xpo1p previously identified in the crystal structure of Xpo1p-Yrb2p-Gsp1p-GTP complex (Koyama *et al*. 2014). Sites 2 and 3 were located on HEAT repeats 16-18 and HEAT repeats 14-15, respectively, and were not previously identified as FG-repeat binding sites on Xpo1p in the Yrb2p complex (Koyama *et al*. 2014).

The overall structure of the Xpo1p-PKI-Nup42p-Gsp1p-GTP complex (Fig. 2D) was very similar to the previously determined structure of the Xpo1p-PKI-Gsp1p-GTP complex (PDB code, 3WYG; Koyama *et al*. 2014) with Cα r.m.s.d. (root mean square deviation) of 0.62 Å over 1188 equivalent residues of Xpo1p, PKI, and Gsp1p. Thus the binding of Nup42p did not induce major conformational change of Xpo1p. However, minor conformational changes of Xpo1p were observed at sites 2 and 3, as described below.

The SxFG/PxFG repeat binding sites on Xpo1p

Details of Nup42p-Xpo1p interactions at the three binding sites are shown in Fig. 3, Fig. 4, and Fig. 5, and the interactions are illustrated schematically in Fig. 6. At site 1, eight residues of Nup42p (90-SAFGAPAF-97) bound along HEAT repeats 17-20 of Xpo1p, with the two phenylalanine aromatic rings occupying nonpolar depressions between adjacent HEAT repeats (Fig. 3A). The side chain of $F92^{Nup42p}$ fitted into a nonpolar pocket

(pocket P1; Fig. 3B) formed by Xpo1p residues L842, N846, D887, A888, and W891. The side chain of F97^{Nup42p} fitted into another nonpolar pocket (pocket P2; Fig. 3B) formed by Xpo1p residues S937, F941, T944, M998, A1002, and F1003. These hydrophobic interactions were supplemented by a H-bond between E938^{Xpo1p} and Nup42p main chain. The binding of Nup42p to this site was not associated with conformational changes in Xpo1p (Fig. 3C).

At site 2, nine residues of Nup42p (90-SAFGAPAFG-98) bound along HEAT repeats 16-18 of Xpo1p (Fig. 4A), occupying two adjacent nonpolar pockets (pockets P3 and P4; Fig. 4B). The side chain of F92^{Nup42p} fitted into the nonpolar pocket P3 formed by Xpo1p residues L834, L877, and L884. The binding of Nup42p to this pocket was associated with a slight movement (indicated by an arrow in Fig. 4C) of a loop of Xpo1p containing L877 to open up this pocket to accommodate the phenylalanine side chain. The aliphatic side chain of P95^{Nup42p} packed against the aromatic ring of F97^{Nup42p}, and these two hydrophobic side chains formed intimate contacts with the nonpolar pocket P4 formed by Xpo1p residues L796, E797, M800, L832, and Q835. The binding of Nup42p to this site was further stabilized by H-bonds formed between Nup42p main chain and Xpo1p residues N793, Q835, and E839.

At site 3, eight residues of Nup42p (110-TTSAFGAP-117) bound along the cleft formed between HEAT repeats 14 and 15 (Fig. 5A). The side chain of F114^{Nup42p} fitted into a deep nonpolar pocket (pocket P5; Fig. 5B) and made intimate contacts with Xpo1p residues W676, I679, V680, L690, L731, and A734. The side chain of S112^{Nup42p} packed against the aromatic ring of F114^{Nup42p} and formed a H-bond with Q730^{Xpo1p} at the entrance of the pocket P5. Nup42p residues G115, A116, and P117 packed against a shallow nonpolar pocket (pocket P6; Fig. 5B) formed by Xpo1p residues T688, L691, A734, S737, and M738. In addition, there were H-bonds made between Nup42p main

chain and Xpo1p residues S683, T684, N727, and Q730. The binding of Nup42p to this site was associated with a conformational change of Xpo1p (indicated by an arrow in Fig. 5C) to open the cleft between HEAT repeats 14 and 15 to accommodate Nup42p, indicating that this binding site has conformational plasticity.

Mutational analyses of Nup42p-Xpo1p interactions

Although the serine and proline residues of the SxFG/PxFG repeat motifs of Nup42p did not make specific interactions with Xpo1p at site 1, the side chain of serine of SxFG was involved in specific recognition by Xpo1p at site3, and the nonpolar side chain of proline of PxFG made direct contacts with Xpo1p at sites 2 and 3. We used Nup42p mutants to evaluate the contribution made by the serines and prolines in the SxFG/PxFG repeat motifs for the binding to Xpo1p in solution (Fig.7). The design of the Nup42p mutants is depicted in Fig. 7A. In the "SA" mutant, all of the serines in the SxFG sequences are changed to alanines so that the serine side chain cannot form a H-bond with Xpo1p; in the "PD" mutant, all of the prolines in the PxFG sequences are changed to aspartates to disrupt hydrophobic interaction between the proline side chain and Xpo1p; the "SA/PD" mutant has mutations that are combinations of the "SA" and "PD" mutations.

Two Nup42p fragments (residues 31-213 or 31-122) were used in GST pull-down assay to analyze the binding of Xpo1p to Nup42p in the presence of Gsp1p-GTP (Fig. 7B). Although the SA mutations of Nup42p residues 31-213 hardly affected the binding of Xpo1p and Gsp1p-GTP to this immobilized Nup42p fragment (Fig. 7B, lanes 1 and 2), the PD mutations strikingly weakened the binding of Xpo1p and Gsp1p-GTP (Fig. 7B, lane 3) and, likewise, the binding to the SA/PD mutant was undetectably weak (Fig. 7B, lane 4). The binding of Xpo1p (in the presence of

Gsp1p-GTP) to Nup42p residues 31-122 was weaker than the binding to Nup42p residues 31-213 (Fig. 7B, lanes 1 and 5). In the case of the shorter fragment (Nup42p residues 31-122), the SA mutations weakened Xpo1p binding in the presence of Gsp1p-GTP (Fig. 7B, lane 6) and, the binding of Xpo1p to the PD mutant or the SA/PD mutant (in the presence of Gsp1p-GTP) was undetectably weak (Fig. 7B, lanes 7 and 8).

We also analyzed how the mutations of Nup42p SxFG/PxFG repeats affect the binding of Xpo1p to Nup42p in the presence of both Gsp1p-GTP and a NES-cargo (HIV-1 Rev) (Fig. 7C). The binding of Xpo1p to Nup42p residues 31-213 in the presence of both Gsp1p-GTP and HIV-1 Rev was quite stable and the SA mutations or the PD mutations alone did not affect Xpo1p binding appreciably (Fig. 7C, lanes 1-3). However, when combined, the SA/PD mutations weakened Xpo1p binding to Nup42p residues 31-213 (in the presence of both Gsp1p-GTP and HIV-1 Rev; Fig. 7C, lane 4). In the case of the shorter fragment of Nup42p (residues 31-122), although the SA mutations hardly affected Xpo1p binding (in the presence of both Gsp1p-GTP and HIV-1 Rev; Fig. 7C, lanes 5 and 6), the PD mutations or the SA/PD mutations abolished Xpo1p binding (in the presence of both Gsp1p-GTP and HIV-1 Rev; Fig. 7C, lanes 7 and 8).

Taken together, the results of the mutational analyses suggest that both serines and prolines in the SxFG/PxFG repeats of Nup42p contribute to Xpo1p binding.

Conservation of the FG-repeat binding sites on CRM1 (Xpo1p) across species

Our structural data also highlight the evolutionary conservation of FG-repeat binding sites on CRM1 (Xpo1p). The structure of yeast Xpo1p nuclear export complex bound to Nup42p peptide identified five binding pockets for the phenylalanine side chains of

Nup42p on the C-terminal region of Xpo1p (Fig. 8A). Similarly, the crystal structure of human CRM1-Spn1-Nup214-Ran-GTP complex (Port *et al*. 2015) identified five FG binding sites on HEAT repeats 14-20 of CRM1 (Fig. 8B). Interestingly, four FG-binding sites (pockets P1, P2, P3, and P5 in the yeast structure) are common in the two structures (Fig. 8C). Thus, as previously pointed out (Port *et al*. 2015), the FG-repeat binding sites of CRM1 (Xpo1p) appear to be highly conserved from yeast to humans.

Discussion

Structural characterization of the interactions between the Xpo1p nuclear export complex and various FG-Nups is a prerequisite to understand the molecular mechanism of nuclear export. However, our attempts to co-crystallize the Xpo1p nuclear export complex with various FG-Nups have been unsuccessful. As an alternative approach, we soaked FG repeat peptides to preformed crystals of the Xpo1p-PKI-Gsp1p-GTP complex. This soaking strategy worked for a Nup42p peptide as reported here. Although previous cell biological studies have shown that the FG domains of Nup42p and Nup159p are dispensable for NES export from the nucleus (Zeitler and Weis, 2004), biochemical studies have suggested that the characteristic SxFG/PxFG repeats on Nup42p and Nup159p provide high affinity binding site for Xpo1p (Allen *et al*. 2001; Allen *et al*. 2002). In agreement with our biochemical data that Nup42p can bind to the Xpo1p nuclear export complex without displacing NES-cargo, the structure of Xpo1p-PKI-Nup42p-Gsp1p-GTP complex showed that the Nup42p peptide can bind to Xpo1p nuclear export complex at three sites on the outer surface of HEAT repeats 14-20 of Xpo1p without disturbing the binding of PKI (an NES-cargo) to Xpo1p. An obvious limitation of the soaking strategy used in this study is that the crystal packing interactions could potentially hinder access of the Nup42p peptide to its binding sites on Xpo1p. It is also likely that the crystallographic data can reveal only strong binding sites for the Nup42p peptide. Therefore, there may be additional FG repeat binding sites on Xpo1p in addition to the three binding sites of Nup42p (sites 1, 2, and 3) observed in the current crystal structure. Nevertheless, our structure revealed multivalent binding of the SxFG/PxFG repeats of Nup42p to the Xpo1p nuclear export complex, analogous to the multivalent binding of the FG repeats of human Nup214 to the CRM1 nuclear export complex (Port *et al*. 2015). Our structure showed atomic detail of how the serines and prolines of the SxFG/PxFG sequence repeat motif, together with the key

phenylalanines, interact directly with Xpo1p. Consistently, mutational analyses of Nup42p provided strong evidence that the serines and prolines of the SxFG/PxFG repeats contribute to Xpo1p binding. Thus, our biochemical and structural data suggest that the SxFG/PxFG-Nups such as Nup42p and Nup159p provide high-affinity binding sites for the Xpo1p nuclear export complex on the cytoplasmic face of NPCs.

The cytoplasmic face of NPCs would be optimal locale for cargo dissociation in a terminal step of nuclear export. Importantly, our biochemical data showed that Yrb1p weakens the binding of the Nup42p FG domain to the Xpo1p nuclear export complex. Thus, Yrb1p would not only displace NES-cargo from Xpo1p and Gsp1p-GTP (Koyama & Matsuura 2010) but also prevent the Xpo1p nuclear export complex from being tightly trapped at the cytoplasmic face of NPCs. Our biochemical data also showed that the binding of unliganded Xpo1p to the Nup42p FG domain is quite weak. The weak binding would have high off-rate. Therefore, after complete disassembly of the Xpo1p nuclear export complex in the cytoplasm, the weak and transient binding of the SxFG/PxFG-Nups to free Xpo1p would facilitate rapid recycling of free Xpo1p back to the nucleus without trapping free Xpo1p on the cytoplasmic face of NPCs.

Although we can only speculate the mechanism by which Yrb1p destabilizes the binding of Nup42p to the Xpo1p nuclear export complex, one possibility would be that the affinity of FG domain to Xpo1p is sensitive to Yrb1p-induced alterations in the conformational flexibility of the Xpo1p HEAT repeats. In the crystal structure of Xpo1p-PKI-Gsp1p-GTP complex, Gsp1p-GTP and the HEAT9 loop bind to the inner surface of HEAT repeats 15-19 and, these interactions are abolished in the Xpo1p-Yrb1p-Gsp1p-GTP complex (Koyama & Matsuura 2010; Koyama *et al*. 2014). This raises a possibility that Yrb1p increases the conformational flexibility of HEAT repeats 14-20 and, as a consequence, weakens the binding of the Nup42p FG domain to

the outer surface of these HEAT repeats. Dissociation of Gsp1p-GTP and Yrb1p from Xpo1p would further increase the conformational flexibility of Xpo1p and thereby further weaken the Xpo1p-FG domain interactions. Experimental analysis of conformational flexibility of Xpo1p will be required to test this hypothesis.

Experimental procedures

Protein expression and purification

The recombinant proteins for GST pull-down assays were expressed in the *E. coli* host strain BL21-CodonPlus(DE3)RIL (Stratagene). His/S-Xpo1p, His/S-PKI, His-Gsp1p, and His/S-Yrb1p were expressed as described (Matsuura & Stewart 2004; Koyama & Matsuura 2010). DNAs encoding Nup42p mutants and HIV-1 Rev were synthesized by GenScript. His/S-HIV-1 Rev was expressed from pET30a (Novagen). His-tagged proteins were purified over Ni-NTA (Novagen) and by gel filtration over Superdex200 or Superdex75 (GE Healthcare). GST-Nup42p (full-length or residues 31-213 or residues 31-122) was expressed from pGEX-TEV (Matsuura & Stewart 2004) and purified over glutathione-Sepharose 4B (GE Healthcare) and gel filtration over Superdex200 (GE Healthcare). All DNA constructs were verified by DNA sequencing.

Xpo1p (*S. cerevisiae*, a functional deletion mutant in which residues 377-413 are deleted; Koyama & Matsuura 2010), PKI (human, full-length S35L mutant that has increased affinity to CRM1; Guttler *et al*. 2010), and Gsp1p (*S. cerevisiae*, residues 1-182, Q71L mutant) used for crystallization were expressed in *E. coli* and Xpo1p-PKI-Gsp1p-GTP complex for crystallization was purified as described (Koyama *et al*. 2014).

GST pull-down assays

Pull-down assays were performed in binding buffer (10 mM Tris-HCl (pH7.5), 150 mM NaCl, 5 mM $Mg(OAc)_2$, 0.05% Tween-20, 2 mM β-mercaptoethanol). GST-fusion proteins were immobilized on 10 μl of packed glutathione-Sepharose 4B (GE

Healthcare) beads and each binding reaction was performed by incubating the beads with reaction mixtures in a total volume of 50 µl for 1 h at 4 °C. The amounts of proteins used are indicated in the figure legends. Beads were then spun down and washed twice with 1 ml of binding buffer, and bound proteins were analyzed by SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) and Coomassie staining.

X-ray crystallography

Nup42p peptide (88-KPSAFGAPAFGSSAPINVNPPSTTSAFGAPSFGST-122) was synthesized by Sigma-Aldrich Japan. Crystals of Xpo1p-PKI-Gsp1p-GTP complex were grown as described (Koyama *et al*. 2014). The Xpo1p-PKI-Gsp1p-GTP-Nup42p complex was formed by soaking the crystals of Xpo1p-PKI-Gsp1p-GTP complex for 3 h in a stabilizing solution (0.1 M Tris-HCl pH 7.8 and 13% PEG20000) containing 2 mg/ml Nup42p peptide (residues 88-122). Crystals were cryoprotected using mother liquor containing 16% PEG20000 and 15% glycerol and flash-cooled in liquid nitrogen. Preliminary X-ray diffraction experiments were performed at Photon Factory beamline BL-1A (Tsukuba, Japan), and a 2.2 Å resolution data set used for final structure determination was collected at 100 K at SPring-8 beamline BL41XU. The crystals had $P2_12_12_1$ symmetry (a = 101.6 Å, b = 107.6 Å, c = 149.1 Å) with one complex in the asymmetric unit. Diffraction data were processed using MOSFLM and CCP4 programs (Winn *et al*. 2011). The structure of Xpo1p-PKI-Gsp1p-GTP-Nup42p complex was solved by molecular replacement using MOLREP (Vagin & Teplyakov, 2010) using the structure of Xpo1p-PKI-Gsp1p-GTP complex (Koyama *et al*. 2014; PDB code, 3WYG) as a search model. Iterative cycles of model building using COOT (Emsley & Cowtan 2004) and refinement using PHENIX (Adams *et al*. 2010) yielded a final model with an R_{free} of 22.0% (R_{cryst} 19.1%). A TLSMD analysis (Painter & Merritt 2006) was used to

define TLS groups for the final cycles of refinement. MolProbity (Chen *et al*. 2010) was used to validate the final structure. Atomic coordinates and structure factors of the Xpo1p-PKI-Gsp1p-GTP complex bound to the Nup42p peptide have been deposited in the Protein Data Bank with accession code 5XOJ. Structural figures were produced using CCP4MG (McNicholas *et al*. 2011).

Acknowledgements

We thank our colleagues in Nagoya, especially Junya Kobayashi, for assistance and discussion. We thank the staff of Photon Factory and SPring-8 for assistance during data collection. The X-ray diffraction data collection experiments at the beamline BL-1A of Photon Factory were performed with the approval of the Photon Factory Program Advisory Committee (Proposal No. 2011G585). The X-ray diffraction data collection experiments at the beamline BL41XU of SPring-8 were performed with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2012B1291). This work was supported by the Sumitomo Foundation (090488) and JSPS/MEXT KAKENHI (18687010, 21770109, 23770110, 25121713, 10J05907, and 24030003). MK was supported by JSPS Research Fellowship.

References

Adachi, Y. & Yanagida, M. (1989) Higher order chromosome structure is affected by cold-sensitive mutations in a Schizosaccharomyces pombe gene crm1+ which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. *J Cell Biol* **108**, 1195-1207.

Adams, P.D., Afonine, P.V., Bunkoczi, G. *et al*. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.

Allen, N.P., Huang, L., Burlingame, A. & Rexach, M. (2001) Proteomic analysis of nucleoporin interacting proteins. *J Biol Chem* **276**, 29268-29274.

Allen, N.P., Patel, S.S., Huang, L., Chalkley, R.J., Burlingame, A., Lutzmann, M., Hurt, E.C. & Rexach, M. (2002) Deciphering networks of protein interactions at the nuclear pore complex. *Mol Cell Proteomics* **1**, 930-946.

Beck, M. & Hurt, E. (2017) The nuclear pore complex: understanding its function through structural insight. *Nat Rev Mol Cell Biol* **18**, 73-89.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. & Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21.

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

Dong, X., Biswas, A., Suel, K.E., Jackson, L.K., Martinez, R., Gu, H. & Chook, Y.M. (2009) Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* **458**, 1136-1141.

Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.

Englmeier, L., Fornerod, M., Bischoff, F.R., Petosa, C., Mattaj, I.W. & Kutay, U. (2001) RanBP3 influences interactions between CRM1 and its nuclear protein export substrates. *EMBO Rep* **2**, 926-932.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W. & Luhrmann, R. (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475-483.

Floer, M. & Blobel, G. (1999) Putative reaction intermediates in Crm1-mediated nuclear protein export. *J Biol Chem* **274**, 16279-16286.

Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I.W. (1997) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051-1060.

Frey, S. & Gorlich, D. (2007) A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* **130**, 512-523.

Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M. & Nishida, E. (1997) CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390**, 308-311.

Fung, H.Y., Fu, S.C., Brautigam, C.A. & Chook, Y.M. (2015) Structural determinants of nuclear export signal orientation in binding to exportin CRM1. *eLife* **4**.

Fung, H.Y., Fu, S.C. & Chook, Y.M. (2017) Nuclear export receptor CRM1 recognizes diverse conformations in nuclear export signals. *eLife* **6**.

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

Guttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M. & Gorlich, D. (2010) NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* **17**, 1367-1376.

Hoelz, A., Glavy, J.S. & Beck, M. (2016) Toward the atomic structure of the nuclear pore complex: when top down meets bottom up. *Nat Struct Mol Biol* **23**, 624-630.

Hulsmann, B.B., Labokha, A.A. & Gorlich, D. (2012) The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* **150**, 738-751.

Kirli, K., Karaca, S., Dehne, H.J., Samwer, M., Pan, K.T., Lenz, C., Urlaub, H. & Gorlich, D. (2015) A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *eLife* **4**.

Knockenhauer, K.E. & Schwartz, T.U. (2016) The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell* **164**, 1162-1171.

Koyama, M. & Matsuura, Y. (2010) An allosteric mechanism to displace nuclear export cargo from CRM1 and RanGTP by RanBP1. *EMBO J* **29**, 2002-2013.

Koyama, M., Shirai, N. & Matsuura, Y. (2014) Structural insights into how Yrb2p accelerates the assembly of the Xpo1p nuclear export complex. *Cell Rep* **9**, 983-995.

Lindsay, M.E., Holaska, J.M., Welch, K., Paschal, B.M. & Macara, I.G. (2001) Ran-binding protein 3 is a cofactor for Crm1-mediated nuclear protein export. *J Cell Biol* **153**, 1391-1402.

Matsuura, Y. & Stewart, M. (2004) Structural basis for the assembly of a nuclear export complex. *Nature* **432**, 872-877.

Matsuura, Y. (2016) Mechanistic Insights from Structural Analyses of Ran-GTPase-Driven Nuclear Export of Proteins and RNAs. *J Mol Biol* **428**, 2025-2039.

McNicholas, S., Potterton, E., Wilson, K.S. & Noble, M.E. (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67**, 386-394.

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

Monecke, T., Guttler, T., Neumann, P., Dickmanns, A., Gorlich, D. & Ficner, R. (2009) Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* **324**, 1087-1091.

Monecke, T., Haselbach, D., Voss, B., Russek, A., Neumann, P., Thomson, E., Hurt, E., Zachariae, U., Stark, H., Grubmuller, H., Dickmanns, A. & Ficner, R. (2013) Structural basis for cooperativity of CRM1 export complex formation. *Proc Natl Acad Sci U S A* **110**, 960-965.

Neville, M., Stutz, F., Lee, L., Davis, L.I. & Rosbash, M. (1997) The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr Biol* **7**, 767-775.

Noguchi, E., Saitoh, Y., Sazer, S. & Nishimoto, T. (1999) Disruption of the YRB2 gene retards nuclear protein export, causing a profound mitotic delay, and can be rescued by overexpression of XPO1/CRM1. *J Biochem* **125**, 574-585.

Ossareh-Nazari, B., Bachelerie, F. & Dargemont, C. (1997) Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* **278**, 141-144.

Painter, J. & Merritt, E.A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-450.

Patel, S.S., Belmont, B.J., Sante, J.M. & Rexach, M.F. (2007) Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* **129**, 83-96.

Port, S.A., Monecke, T., Dickmanns, A., Spillner, C., Hofele, R., Urlaub, H., Ficner, R. & Kehlenbach, R.H. (2015) Structural and Functional Characterization of CRM1-Nup214 Interactions Reveals Multiple FG-Binding Sites Involved in Nuclear Export. *Cell Rep* **13**, 690-702.

Ribbeck, K., Lipowsky, G., Kent, H.M., Stewart, M. & Gorlich, D. (1998) NTF2 mediates nuclear import of Ran. *EMBO J* **17**, 6587-6598.

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

Rout, M.P., Aitchison, J.D., Magnasco, M.O. & Chait, B.T. (2003) Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol* **13**, 622-628.

Saito, N. & Matsuura, Y. (2013) A 2.1-A-resolution crystal structure of unliganded CRM1 reveals the mechanism of autoinhibition. *J Mol Biol* **425**, 350-364.

Schmidt, H.B. & Gorlich, D. (2016) Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends Biochem Sci* **41**, 46-61.

Smith, A., Brownawell, A. & Macara, I.G. (1998) Nuclear import of Ran is mediated by the transport factor NTF2. *Curr Biol* **8**, 1403-1406.

Stade, K., Ford, C.S., Guthrie, C. & Weis, K. (1997) Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* **90**, 1041-1050.

Stewart, M. (2007) Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol* **8**, 195-208.

Strawn, L.A., Shen, T., Shulga, N., Goldfarb, D.S. & Wente, S.R. (2004) Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat Cell Biol* **6**, 197-206.

Taura, T., Krebber, H. & Silver, P.A. (1998) A member of the Ran-binding protein family, Yrb2p, is involved in nuclear protein export. *Proc Natl Acad Sci U S A* **95**, 7427-7432.

Timney, B.L., Raveh, B., Mironska, R., Trivedi, J.M., Kim, S.J., Russel, D., Wente, S.R., Sali, A. & Rout, M.P. (2016) Simple rules for passive diffusion through the nuclear pore complex. *J Cell Biol* **215**, 57-76.

Vagin, A. & Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22-25.

Vetter, I.R. & Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304.

Wen, W., Meinkoth, J.L., Tsien, R.Y. & Taylor, S.S. (1995) Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463-473.

Wente, S.R. & Rout, M.P. (2010) The nuclear pore complex and nuclear transport. *Cold Spring Harbor Perspect Biol* **2**, a000562.

Winn, M.D., Ballard, C.C., Cowtan, K.D. *et al*. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242.

Yamada, J., Phillips, J.L., Patel, S., Goldfien, G., Calestagne-Morelli, A., Huang, H., Reza, R., Acheson, J., Krishnan, V.V., Newsam, S., Gopinathan, A., Lau, E.Y., Colvin, M.E., Uversky, V.N. & Rexach, M.F. (2010) A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol Cell Proteomics* **9**, 2205-2224.

Zeitler, B. & Weis, K. (2004) The FG-repeat asymmetry of the nuclear pore complex is dispensable for bulk nucleocytoplasmic transport in vivo. *J Cell Biol* **167**, 583-590.

Figure legends

Figure 1 Biochemical characterization of Nup42p that has characteristic SxFG/PxFG sequence repeat motif. (A) Amino acid sequence of full-length Nup42p in single letter code. (B)-(C) GST pull-down assay of the assembly and disassembly of the Xpo1p-cargo-Nup42p-Gsp1p-GTP complex. (B) Immobilized 44 μg GST-Nup42p (full-length) was incubated with 10 μg Xpo1p alone (lane 1) or 30 μg Gsp1p-GTP alone (lane 2) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP (lane 3) or 10 μg Xpo1p together with 30 μg Gsp1p-GDP (lane 4) or 30 μg PKI alone (lane 5) or 10 μg Xpo1p together with 30 μg PKI (lane 6) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP and 30 μg PKI (lane 7) or 10 μg Xpo1p together with 30 μg Gsp1p-GDP and 30 μg PKI (lane 8) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP, 30 μg Yrb1p and 30 μg PKI (lane 9) or 30 μg HIV-1 Rev alone (lane 10) or 10 μg Xpo1p together with 30 μg HIV-1 Rev (lane 11) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP and 30 μg HIV-1 Rev (lane 12) or 10 μg Xpo1p together with 30 μg Gsp1p-GDP and 30 μg HIV-1 Rev (lane 13) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP, 30 μg Yrb1p and 30 μg HIV-1 Rev (lane 14). (C) Immobilized 26 μg GST-Nup42p (residues 31-213) was incubated with 10 μg Xpo1p alone (lane 1) or 30 μg Gsp1p-GTP alone (lane 2) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP (lane 3) or 10 μg Xpo1p together with 30 μg Gsp1p-GDP (lane 4) or 30 μg PKI alone (lane 5) or 10 μg Xpo1p together with 30 μg PKI (lane 6) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP and 30 μg PKI (lane 7) or 10 μg Xpo1p together with 30 μg Gsp1p-GDP and 30 μg PKI (lane 8) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP, 30 μg Yrb1p and 30 μg PKI (lane 9) or 30 μg HIV-1 Rev alone (lane 10) or 10 μg Xpo1p together with 30 μg HIV-1 Rev (lane 11) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP and 30 μg HIV-1 Rev (lane 12) or 10 μg Xpo1p together with 30 μg Gsp1p-GDP and 30 μg HIV-1 Rev (lane 13) or 10 μg Xpo1p together with 30 μg Gsp1p-GTP, 30 μg Yrb1p and 30 μg HIV-1 Rev (lane 14).

Figure 2 Crystal structure of the Xpo1p nuclear export complex bound to the SxFG/PxFG repeat peptide of Nup42p. (A)-(C) The electron density of the Nup42p peptide bound to Xpo1p on the outer surface of (A) HEAT repeats 17-20 (site 1), (B) HEAT repeats 16-18 (site 2), and (C) HEAT repeats 14-15 (site 3). The omit F_o–F_c electron density map (contoured at 2.5σ) covering the Nup42p peptide at sites 1-3 is shown in blue mesh with the refined model of the Nup42p peptide (stick representation). Xpo1p is shown in ribbon representation. (C) Overview of the Xpo1p-PKI-Gsp1p-GTP complex bound to the Nup42p peptide. Xpo1p (gold), PKI (purple), and Gsp1p (cyan) are shown in ribbon representation. The HEAT repeats 9-20 of Xpo1p are labeled H9-H20. GTP and Nup42p are shown in space-filling representation.

Figure 3 Xpo1p-Nup42p interactions at site 1. (A) A stereo view of the interactions between Xpo1p (ribbon representation) and Nup42p peptide (stick representation). A dashed line indicates a H-bond. (B) A stereo view of the molecular surface of Xpo1p colored according to electrostatic potential, shaded from blue (potential +250 mV) through white (potential 0 mV) to red (potential –250 mV). (C) A stereo view of overlay of Xpo1p (ribbon representation in yellow) in the Xpo1p-PKI-Gsp1p-GTP-Nup42p complex and Xpo1p (ribbon representation in brown) in the Xpo1p-PKI-Gsp1p-GTP complex (PDB code, 3WYG).

Figure 4 Xpo1p-Nup42p interactions at site 2. (A) A stereo view of the interactions between Xpo1p (ribbon representation) and Nup42p peptide (stick representation). Dashed lines indicate H-bonds. (B) A stereo view of the molecular surface of Xpo1p colored according to electrostatic potential, shaded from blue (potential +250 mV) through white (potential 0 mV) to red (potential –250 mV). (C) A stereo view of overlay of Xpo1p (ribbon representation in yellow) in the Xpo1p-PKI-Gsp1p-GTP-Nup42p

complex and Xpo1p (ribbon representation in brown) in the Xpo1p-PKI-Gsp1p-GTP complex (PDB code, 3WYG).

Figure 5 Xpo1p-Nup42p interactions at site 3. (A) A stereo view of the interactions between Xpo1p (ribbon representation) and Nup42p peptide (stick representation). Dashed lines indicate H-bonds. (B) A stereo view of the molecular surface of Xpo1p colored according to electrostatic potential, shaded from blue (potential +250 mV) through white (potential 0 mV) to red (potential –250 mV). (C) A stereo view of overlay of Xpo1p (ribbon representation in yellow) in the Xpo1p-PKI-Gsp1p-GTP-Nup42p complex and Xpo1p (ribbon representation in brown) in the Xpo1p-PKI-Gsp1p-GTP complex (PDB code, 3WYG).

Figure 6 Schematic illustration of Xpo1p-Nup42p interactions on the outer surface of HEAT repeats 14-20 of Xpo1p. Hydrophobic interactions or van der Waals interactions are shown as continuous lines. H-bonds are shown as dashed lines.

Figure 7 Mutational analyses of the SxFG/PxFG repeats of Nup42p. (A) Amino acid substitutions in the Nup42p mutants. (B) GST pull-down assay. Immobilized 25 μg GST-Nup42p (residues 31-213) (lane 1, wild-type; lane 2, SA mutant; lane 3, PD mutant; lane 4, SA/PD mutant) or 25 μg GST-Nup42p (residues 31-122) (lane 5, wild-type; lane 6, SA mutant; lane 7, PD mutant; lane 8, SAPD mutant) was incubated with 10 μg Xpo1p and 30 μg Gsp1p-GTP. (C) GST pull-down assay. Immobilized 25 μg GST-Nup42p (residues 31-213) (lane 1, wild-type; lane 2, SA mutant; lane 3, PD mutant; lane 4, SA/PD mutant) or 25 μg GST-Nup42p (residues 31-122) (lane 5, wild-type; lane 6, SA mutant; lane 7, PD mutant; lane 8, SA/PD mutant) was incubated with 10 μg Xpo1p, 30 μg HIV-1 Rev and 30 μg Gsp1p-GTP.

Figure 8 Comparison with human Nup214. (A) Molecular surface of Xpo1p in the Xpo1p-PKI-Nup42p-Gsp1p-GTP complex colored according to electrostatic potential, shaded from blue (potential +250 mV) through white (potential 0 mV) to red (potential −250 mV). The main chain of Nup42p is shown in worm representation (green), with the phenylalanine side chains shown in stick representation. (B) Molecular surface of CRM1 in the CRM1-Spn1-Nup214-Ran-GTP complex (PDB code, 5DIS) colored according to electrostatic potential, shaded from blue (potential +250 mV) through white (potential 0 mV) to red (potential −250 mV). The main chain of Nup214 is shown in worm representation (cyan), with the side chains of F1922, F1938, F1947, F1982 and F1988 shown in stick representation. (C) Overlay of Xpo1p (in the Xpo1p-PKI-Nup42p-Gsp1p-GTP complex; ribbon representation in yellow) and CRM1 (in the CRM1-Spn1-Nup214-Ran-GTP complex; ribbon representation in pink; PDB code, 5DIS). Only HEAT repeats 14-20 are shown for clarity.

Table 1 Data collection and refinement statistics

Data collection	
Space group	$P2_12_12_1$
Unit cell dimensions	
a, *b*, *c* (Å)	101.6, 107.6, 149.1
α, β, γ (degree)	90, 90, 90
Wavelength (Å)	1.0
X-ray source	BL41XU, SPring-8
Resolution range (Å)[a]	31.58-2.20 (2.24-2.20)
No. of measured reflections[a]	427150 (20591)
No. of unique reflections[a]	81567 (4380)
Completeness (%)[a]	97.8 (96.4)
R_{merge} (%)[a]	11.9 (79.8)
Mean $I/\sigma(I)$[a]	9.4 (2.0)
Mean *I* half-set correlation CC(1/2)[a]	0.996 (0.560)
Multiplicity[a]	5.2 (4.7)
Wilson *B* factor (Å^2)	25.5
Refinement	
Resolution range (Å)[a]	31.58-2.20 (2.23-2.20)
R_{work} (%)[a]	19.1 (30.2)
R_{free} (%)[a]	22.0 (33.1)
No. of atoms	
Protein	9723
Water	494
GTP	32
Mg^{2+}	1
No. of amino acids	1212
Mean *B* factor (Å^2)	
Xpo1p	41.1
PKI	61.4
Gsp1p	32.8
Nup42p	58.3
Water	44.3
MgGTP	26.5
RMSD from ideality	
Bond lengths (Å)	0.002
Bond angles (degree)	0.656
Protein geometry[b]	
Rotamer outliers (%)	0.2
Ramachandran favored (%)	98.1
Ramachandran outliers (%)	0
Cβ deviations > 0.25 Å (%)	0
PDB code	5XOJ

[a] Values in parentheses are for the highest-resolution shell.

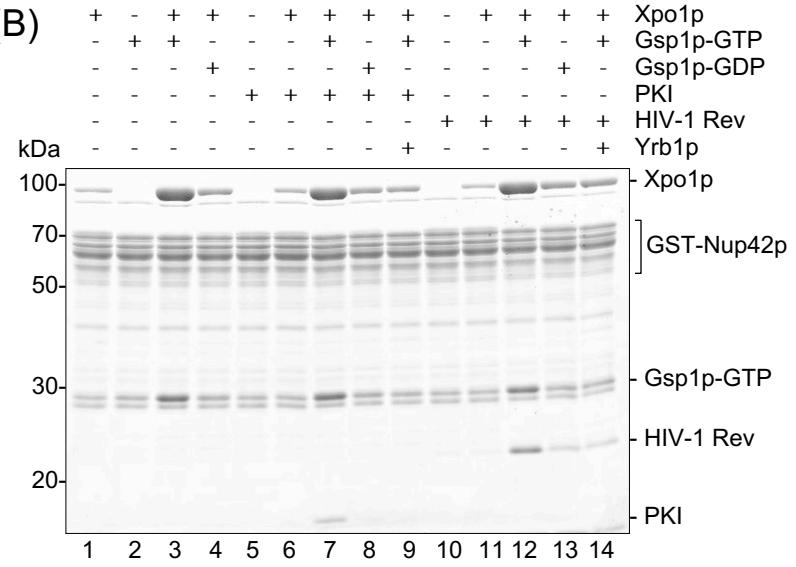
[b] As defined by MolProbity (Chen *et al*. 2010).

(A)

Nup42p (full-length)

```
  1 MSAFGNPFTSGAKPNLSNTSGI
 19  NPFTNNAASTNNMGG
 38  SAFGRPSFGTANTMTGGTTT
 58  SAFGMPQFGTNTGNTGNTSI
 78  SAFGNTSNAAKP
 90  SAFGAPAFGSSAPINVNPPSTT
112  SAFGAPSFGSTGFGAMAATS
132  NPFGKSPGSMG
143  SAFGQPAFGANKTAIPSSSVSNSNN
168  SAFGAASNTPLTTT
182  SPFGSLQQNASQNASSTS
200  SAFGKPTFGAATNTQ
215  SPFGTIQNTSTSSGTGV
232  SPFGTFGTNSNNK
245  SPFSNLQSGAGAGS
259  SPFGTTTSKANNNNNVGS
277  SAFGTTNNQ
286  SPFSGGSG
294  GTFGSASNLNKNTNGNFQ
312  SSFGNKG
319  FSFGITPQNDANK
332  VSQSNPSFGQTMPNTDPNISLKSNGNA
359  TSFGFGQQQMNATNVNANTATGKIRFVQ
387  GLSSEKDGILELADLAEETLKIFRANKFE
416  LGLVPDIPPPPALVA
```

(B)



(C)

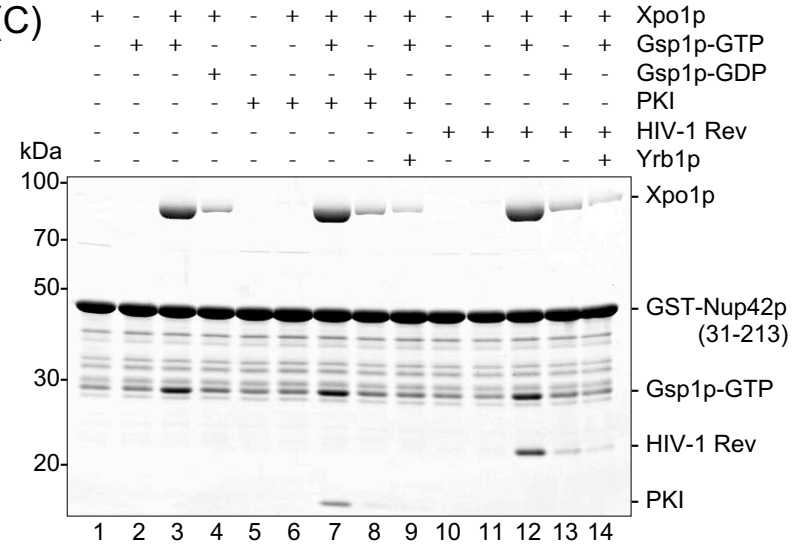


Fig. 1

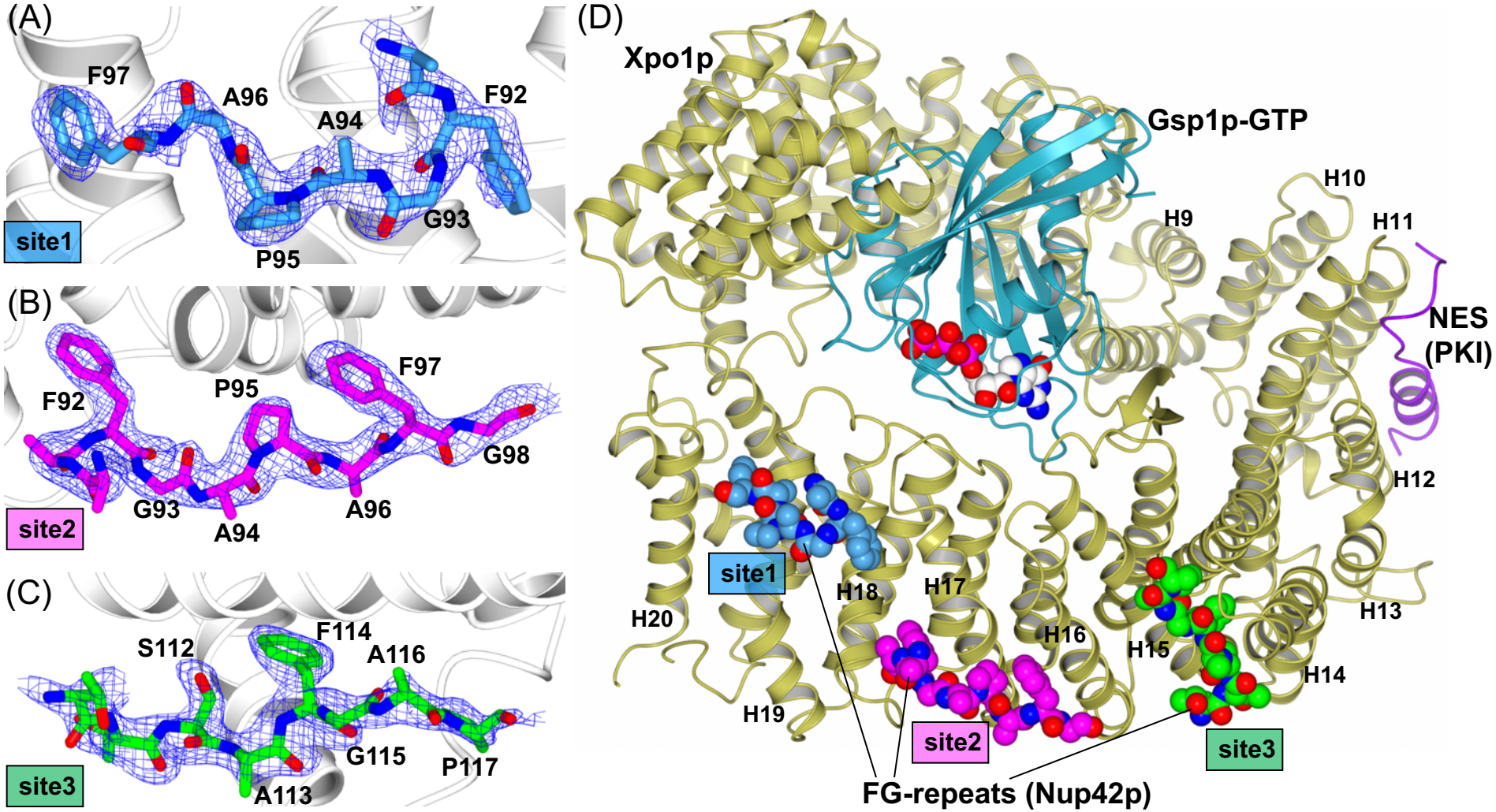

(A)
F97
A96
A94
F92
G93
P95
site1
(B)
F97
P95
F92
G98
G93
A96
A94
site2
(C)
S112
F114
A116
G115
P117
A113
site3
(D)
Xpo1p
Gsp1p-GTP
H10
H9
H11
NES
(PKI)
H12
site1
H18
H17
H20
H16
H15
H13
H14
H19
site2
site3
FG-repeats (Nup42p)

Fig. 2

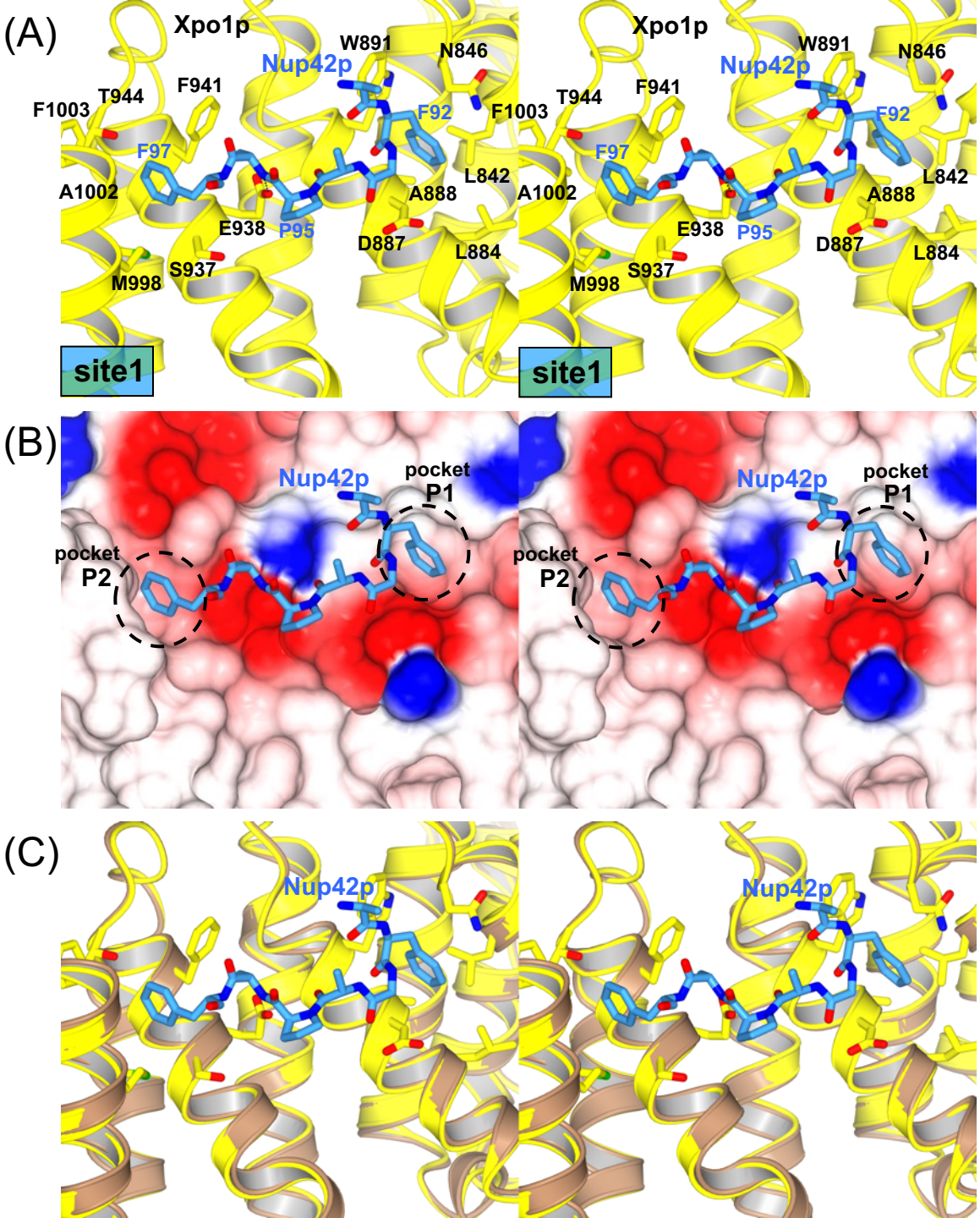

(A)
Xpo1p
Nup42p
W891
N846
F941
T944
F1003
F92
F97
L842
A1002
A888
E938
P95
D887
L884
S937
M998
site1
Xpo1p
Nup42p
W891
N846
F941
T944
F1003
F92
F97
L842
A1002
A888
E938
P95
D887
L884
S937
M998
site1
(B)
Nup42p
pocket
P1
pocket
P2
Nup42p
pocket
P1
pocket
P2
(C)
Nup42p
Nup42p

Fig. 3

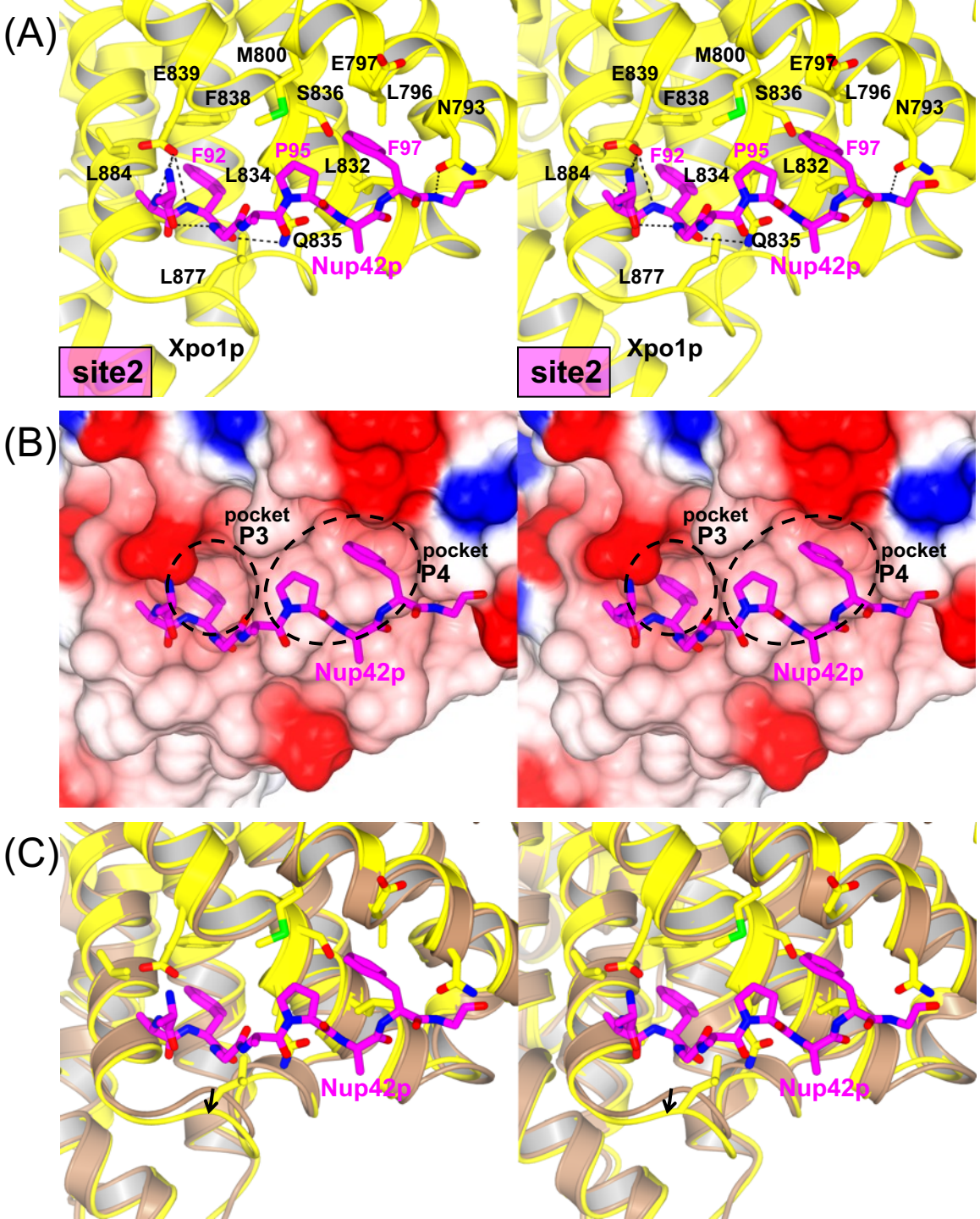

(A)
M800
E797
E839
F838
S836
L796
N793
F92
P95
F97
L884
L834
L832
Q835
Nup42p
L877
Xpo1p
site2
(B)
pocket
P3
pocket
P4
Nup42p
(C)
Nup42p

Fig. 4

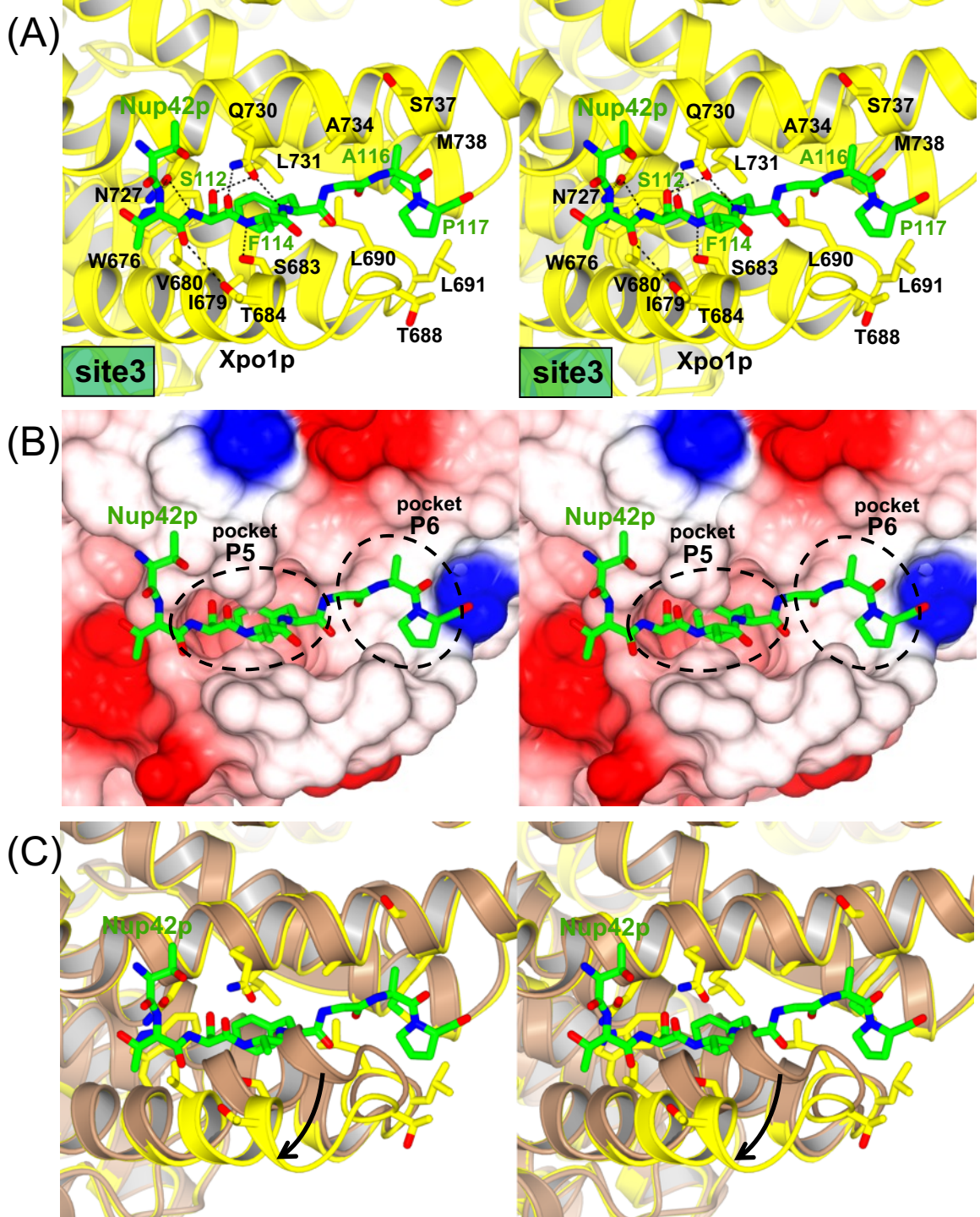

(A)
Nup42p
Q730
A734
S737
M738
L731
A116
S112
N727
P117
F114
W676
S683
L690
V680
I679
L691
T684
T688
site3
Xpo1p
Nup42p
Q730
A734
S737
M738
L731
A116
S112
N727
P117
F114
W676
S683
L690
V680
I679
L691
T684
T688
site3
Xpo1p
(B)
Nup42p
pocket
P5
pocket
P6
Nup42p
pocket
P5
pocket
P6
(C)
Nup42p
Nup42p

Fig. 5

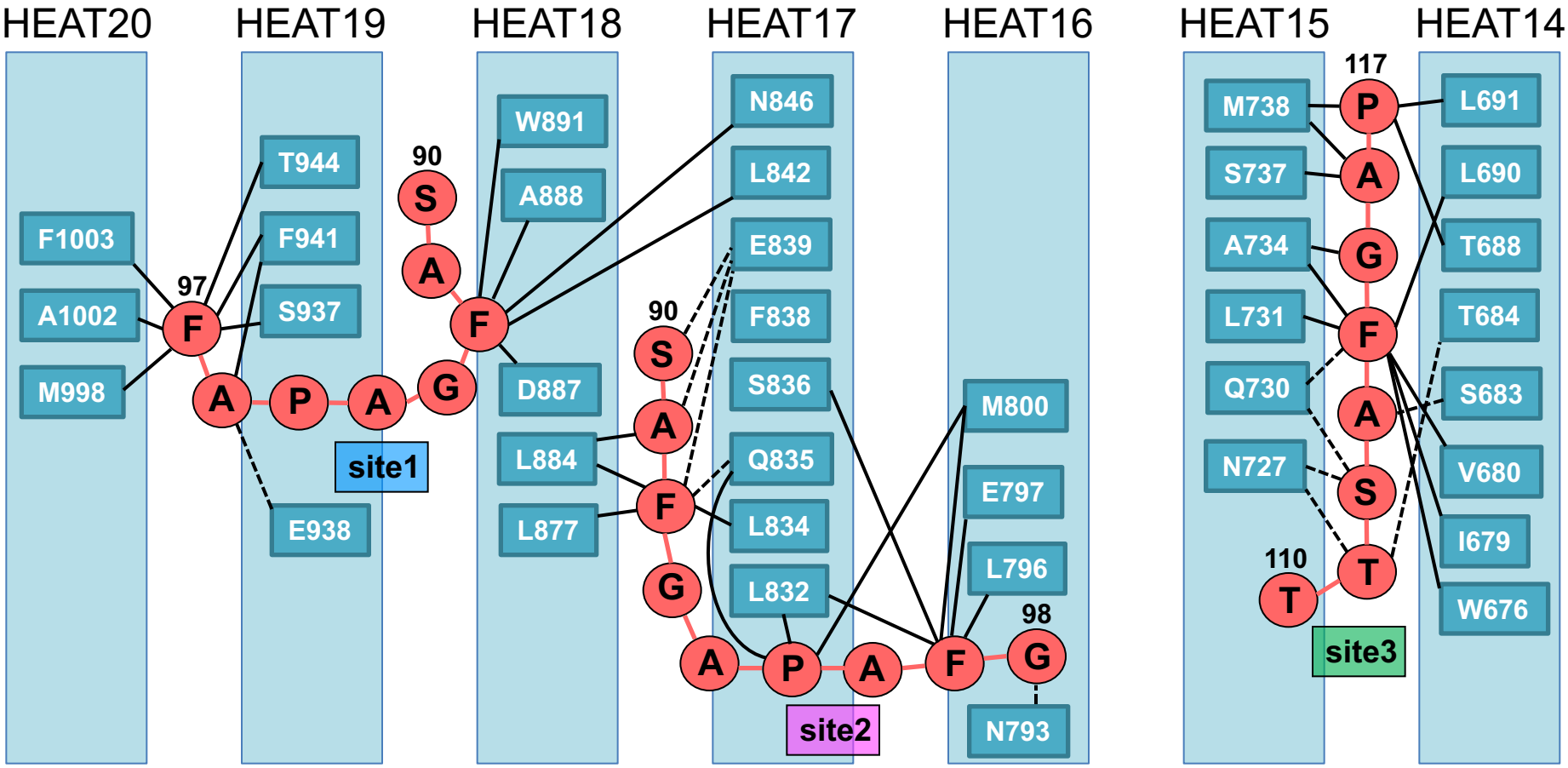

HEAT20
HEAT19
HEAT18
HEAT17
HEAT16
HEAT15
HEAT14
F1003
A1002
M998
97
F
A
P
A
G
T944
F941
S937
E938
site1
90
S
A
F
W891
A888
D887
L884
L877
90
S
A
F
G
A
P
A
F
G
N846
L842
E839
F838
S836
Q835
L834
L832
M800
E797
L796
98
N793
site2
117
M738
S737
A734
L731
Q730
N727
P
A
G
F
A
S
T
110
T
site3
L691
L690
T688
T684
S683
V680
I679
W676

Fig. 6

(A)

Nup42p residues 31-213 (wild-type)

```
 31 STNNMGG
 38 SAFGRPSFGTANTMTGGTTT
 58 SAFGMPQFGTNTGNTGNTSI
 78 SAFGNTSNAAKP
 90 SAFGAPAFGSSAPINVNPPSTT
112 SAFGAPSFGSTGFGAMAATSNPFGKSPGSMG
143 SAFGQPAFGANKTAIPSSSVSNSNN
168 SAFGAASNTPLTTT
182 SPFGSLQQNASQNASSTS
200 SAFGKPTFGAATNT
```

Nup42p residues 31-213 (SA mutant)

```
 31 STNNMGG
 38 AAFGRPSFGTANTMTGGTTT
 58 AAFGMPQFGTNTGNTGNTSI
 78 AAFGNTSNAAKP
 90 AAFGAPAFGSSAPINVNPPSTT
112 AAFGAPSFGSTGFGAMAATSNPFGKSPGSMG
143 AAFGQPAFGANKTAIPSSSVSNSNN
168 AAFGAASNTPLTTT
182 APFGSLQQNASQNASSTS
200 AAFGKPTFGAATNT
```

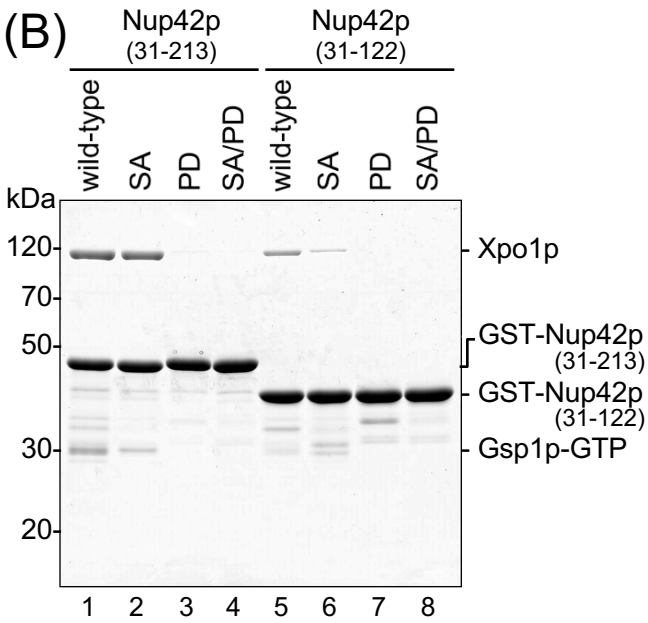
Nup42p residues 31-213 (PD mutant)

```
 31 STNNMGG
 38 SAFGRDSFGTANTMTGGTTT
 58 SAFGMDQFGTNTGNTGNTSI
 78 SAFGNTSNAAKP
 90 SAFGADAFGSSAPINVNPPSTT
112 SAFGADSFGSTGFGAMAATSNPFGKSPGSMG
143 SAFGQDAFGANKTAIPSSSVSNSNN
168 SAFGAASNTPLTTT
182 SPFGSLQQNASQNASSTS
200 SAFGKDTFGAATNT
```

Nup42p residues 31-213 (SA/PD mutant)

```
 31 STNNMGG
 38 AAFGRDSFGTANTMTGGTTT
 58 AAFGMDQFGTNTGNTGNTSI
 78 AAFGNTSNAAKP
 90 AAFGADAFGSSAPINVNPPSTT
112 AAFGADSFGSTGFGAMAATSNPFGKSPGSMG
143 AAFGQDAFGANKTAIPSSSVSNSNN
168 AAFGAASNTPLTTT
182 APFGSLQQNASQNASSTS
200 AAFGKDTFGAATNT
```

(B)



(C)

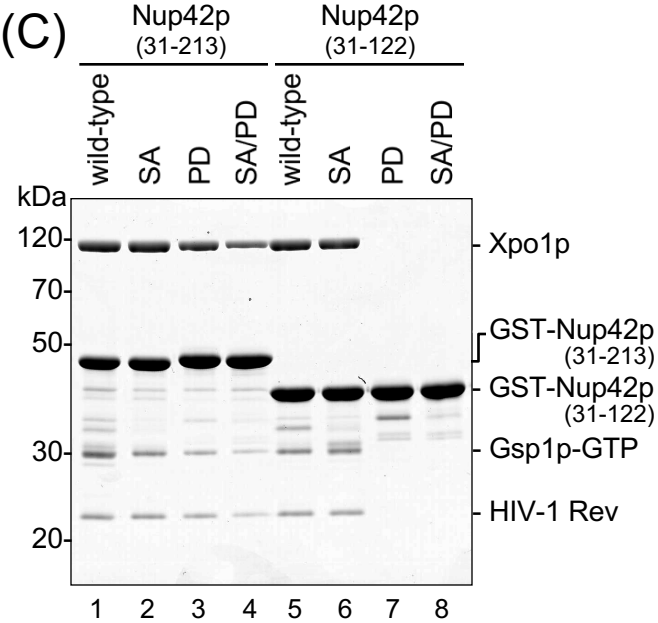


Fig. 7

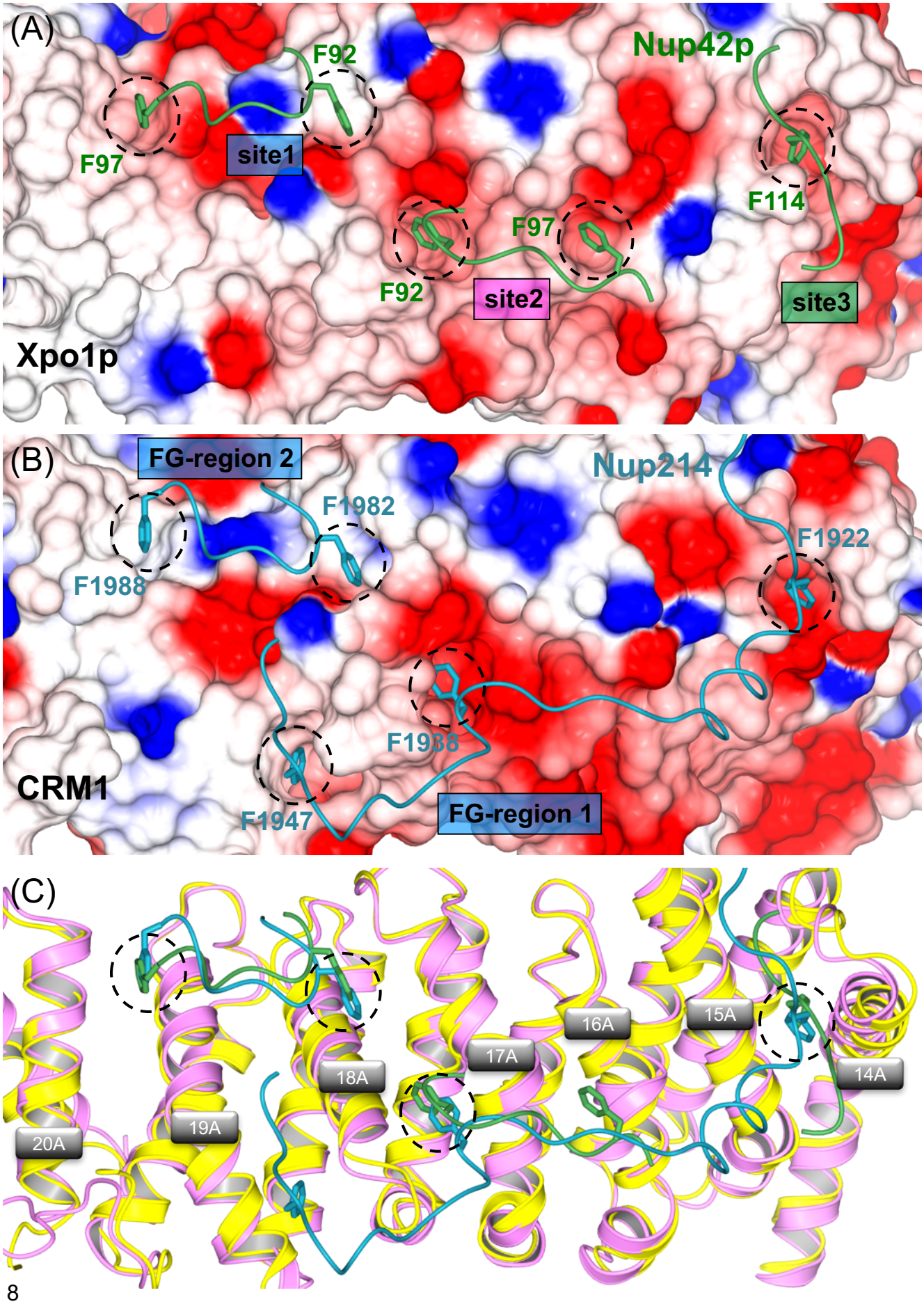

(A)
F92
Nup42p
F97
site1
F114
F97
F92
site2
site3
Xpo1p
(B)
FG-region 2
Nup214
F1982
F1922
F1988
F1938
CRM1
F1947
FG-region 1
(C)
16A
15A
17A
18A
14A
19A
20A

Fig. 8