CHARACTERIZATION OF NUCLEAR LOCALIZATION SIGNALS OF KARYOPHERIN-MEDIATED NUCLEAR IMPORT

APPROVED BY SUPERVISORY COMMITTEE

Yuh Min Chook, Ph. D.

Beatriz M.A. Fontoura, Ph. D.

Joseph P.Albanesi, Ph. D.

Melanie H. Cobb, Ph. D

Nick V. Grishin, Ph. D.

DEDICATION

I would like to express my appreciation and thanks to my mentor Yuh Min Chook for her continuous guidance and support throughout my Ph.D. education. She has been a great role model for me in pursuit of my science career. She always pushes me out my comfort zone to pursuit further and her dedication to science is the great inspiration for me. She is the best mentor I could ever have. I also want to thank every member of the Chook Lab past and present for making the lab a very enjoyable environment. I would especially like to thank Garen Collett for helping me with cloning. I would like to thank the members of my committee Dr.s Beatriz Fontoura, Joseph Albanesi, Melanie Cobb, Nick Grishin for their time and invaluable guidance throughout my education as a scientist. I would especially like to thank Dr. Xuewu Zhang who taught me a lot about X-ray crystallography for his kindness and support, Dr. Jimin Pei who helped me with bioinformatics, and Dr. Kate Luby-Phelps who helped me with microscopy. I would like to thank my friends without whose support I would not have made it through the graduate school, especially Pei-Ling, Jen-Hsuang, Yuwen, Lan, Wentao and Huawei. I would like to express my great thanks to my dearest parents. Their endless love and support have been the strongest foundation for me to pursue my career. Finally, I would like to thank my boyfriend Lei for his support and patience, whose gentleness and wit made me through the ups and downs.

CHARACTERIZATION OF NUCLEAR LOCALIZATION SIGNALS OF KARYOPHERIN-MEDIATED NUCLEAR IMPORT

by

ZI CHAO ZHANG

DISSERTATION / THESIS

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

June, 2011

CHARACTERIZATION OF NUCLEAR LOCALIZATION SIGNALS OF KARYOPHERIN-MEDIATED NUCLEAR IMPORT

ZI CHAO ZHANG, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2011

Supervising Professor Yuh Min Chook, Ph.D

Nucleocytoplasmic transport is mediated by Karyopherin beta (Kap beta) proteins in a Ran-dependent manner. Ten import Kap betas recognize their cargos through the nuclear localization signals (NLSs) and carry them into the nucleus. Recent structural and biochemical work on Kap beta2 (or Transportin) and its well-characterized hnRNP A1-NLS (or M9NLS) reveal that NLSs recognized by Kap beta2 are structurally disordered, have overall positive charges and contain a loose N-terminal hydrophobic or basic motif followed by a C-terminal conserved R/H/K$X_{(2-5)}$PY motif. The newly defined PY-NLSs are further divided into two subclasses: hydrophobic or basic PY-NLSs (hPY or bPY). Bioinformatic searches using these physical characteristics predicted 81 new PY-NLSs. Of the 77 tested new PY-NLSs, 13 showed strong binding to Kap beta2, 8 showed moderate binding and 56 have very weak or no binding.

Comparison of Kap beta2 in complex with hnRNP A1 and M NLSs suggest that PY-NLSs are multivalent and each epitope has different contribution to the overall binding energy, which lead to the design of the chimeric M9M peptide. M9M as a Kap beta2-

specific inhibitor mislocalizes the Kap beta2 cargos, hnRNP A1, HuR and hnRNP M but has no effect on HDAC1, a cargo for Impα/β pathway.

Unexpected redundant import pathways for NXF1 are also discovered using M9M peptide. The N-terminal disordered region of human NXF1 contains NLSs for Imp beta, Kap beta2, Imp4, Imp11 and Imp alpha. Mutation of the NLSs in NXF1 abolished binding to the Karyopherins, mislocalized NXF1 to the cytoplasm and significantly compromised its mRNA export function. Sequence examination of NXF1 from divergent eukaryotes and the interactions of NXF1 homologs with various Karyopherins have revealed the redundancy of nuclear import pathways for NXF1 increased progressively from fungi to nematodes and insects to chordates.

TABLE OF CONTENT

CHAPTER SEVEN

CHAPTER EIGHT

PRIOR PUBLICATIONS

Zhang ZC, Satterly N, Fontroura BMA, Chook YM. *Evolutionary development of redundant nuclear localization signals in the mRNA export factor NXF1*. Mol Biol Cell 2011, under revision.

Satterly, N, Yarbrough, M, Stubbs S, Zhang ZC, Chook YM, Conrad N, BMA Fontoura BMA. *The Cleavage and Polyadenylation factor CPSF30 Inhibits NXF1 Mediated mRNA Export*. Mol Biol Cell. 2011, under revision.

Zhang ZC and Chook YM. *Structural Analysis of Karyopherin-mediated nucleocytoplasmic transport.* Nuclear Transport. 2009, edited by Ralph Kehlenbach, Landes Bioscience

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 May;14(5):452-4.

Lee BJ, Cansizoglu AE, Suel KE, Louis TH, Zhang ZC, Chook YM. *Rules for nuclear localization sequence recognition by karyopherin beta 2.* Cell. 2006 Aug 11;126(3):543-58.

LIST OF FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

ATP	adenosine triphosphate
B-ME	beta-mercaptoethanol
DTT	dithiothreitol
E.coli	*Eschericheria coli*
EDTA	ethylenediamine tetra-acetic acid
GAP	GTPase activating protein
GDP	guanosine 5‘-diphosphate
GEF	guanine nucleotide exchange factor
GST	glutathione S-transferase
GTP	guanosine 5‘-triphosphate
HEAT and TOR1	Huntington, Elongation factor 3, _A‘ subunit of protein phosphatase-2A,
HEPES	4-(2-hydroxyethyl)-1-piperizine-ethanesulfonic acid
hnRNP	heterogeneous nuclear ribonucleoprotein
Impβ	Karyopherin beta-1, importi beta
IPTG	Isopropyl β-D-thiogalactoside

ITC	isothermal titration calorimetry
Kap	karyopherin
Kapα	Karyopherin alpha
Kapβ2	Karyopherin beta-2
K_D	disociation constant
kDa	kilo Dalton
LB	Luria Bertoni
MBP	Maltose Binding Protein
NES	nuclear export signal
NLS	nuclear localization signal
NPC	nuclear pore complex
PCR	polymerase chain reaction
PY-NLS	proline-tyrosine nuclear localization signal
r.m.s.d.	Root Mean Square Deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis

CHAPTER ONE

INTRODUCTION

Overview of Nucleocytoplasmic Transport

Eukaryotic cells are characterized by physical separation of their genomic material from the rest of the cell by the nuclear envelope (NE), a double membrane system that is contiguous with the endoplasmic reticulum (ER)(D'Angelo and Hetzer 2006). DNA replication and RNA processing are restrained in the nucleus, while protein synthesis and other cellular processes take place in the cytoplasm. This compartmentalization evidently benefits the eukaryotes in evolution: it may restrict accessibility to large genomic material thus stabilizing it, and it may also provide additional regulatory strategies to refine responses to complex environments. On the other hand, compartmentalization leads to a requirement for the exchange of huge volumes of materials cross the NE(Chook, Cingolani et al. 1999; Kuersten, Ohno et al. 2001; Damelin, Silver et al. 2002; Fried and Kutay 2003). Not only must the nuclear proteins synthesized in the cytoplasm be imported into the nucleus where they execute their functions, but many RNA and ribonucleoproteins (RNPs) also need to be exported into the cytoplasm to function in translation. In interphase cells, this nucleocytoplasmic exchange is restricted through the nuclear pore complex (NPC), a huge protein complex that penetrates the NE to form a channel for material exchange. Ions, metabolites and other small molecules can freely diffuse through the NPC (Paine, Moore et al. 1975). However, the NPC is

impermeable to macromolecules, such as proteins and RNPs larger than 30KDa (Gorlich and Kutay 1999).

The transport of macromolecules is facilitated by specific transport receptors that recognize designated signals in their cargos (Gorlich and Kutay 1999; Chook and Blobel 2001; Fried and Kutay 2003; Cook, Bono et al. 2007). For nuclear import, the import receptors (named importins) recognize nuclear localization signals (NLSs) of the cargos in the cytoplasm and after translocation, they release the cargos in the nucleus with the help of RanGTP (Chook and Blobel 2001; Stewart 2007) (Figure 1-1). For nuclear export, the export receptors (named exportins) bind nuclear export signals (NESs) in the cargos and RanGTP cooperatively in the nucleus and the export complex is disassembled in the cytoplasm to release the cargos (Cook, Bono et al. 2007; Cook and Conti 2010) (Figure 1-1). Even small proteins or RNAs like histones and tRNAs use the facilitated transport process (Zasloff 1983; Breeuwer and Goldfarb 1990; Arts, Fornerod et al. 1998; Kutay, Lipowsky et al. 1998; Jäkel, Albig et al. 1999). Bidirectionary nucleoctyoplasimc transport is highly selective, controlled and quite different from protein transport into the ER, mitochondria, and chloroplasts, where proteins pass through the membranes only once in unfolded form. Proteins and complexes preserve their native folds during the nucleoctyoplasimc transport process (Gorlich and Kutay 1999).

Central to facilitated nuclear transport are the transport receptors, most of which belong to a family of proteins called Karyoherins (Kaps) (Mosammaparast and Pemberton 2004). (Chook and Blobel 2001). There are also a few non-Kap transport receptors, such as nuclear transport factor 2 (NTF2) for Ran import (Ribbeck, Lipowsky et al. 1998; Smith,

Brownawell et al. 1998) and nuclear export factor 1 (NXF1/Mex67) for mRNA export (Segref, Sharma et al. 1997; Grüter, Tabernero et al. 1998; Herold, Klymenko et al. 2001). A limited number of proteins besides the transport factors, such as β-catenin, can mediate their own transport via direct interaction with the NPC (Fagotto, Gluck et al. 1998). Research in this thesis focused mainly on Karyopherin-mediated nuclear import.

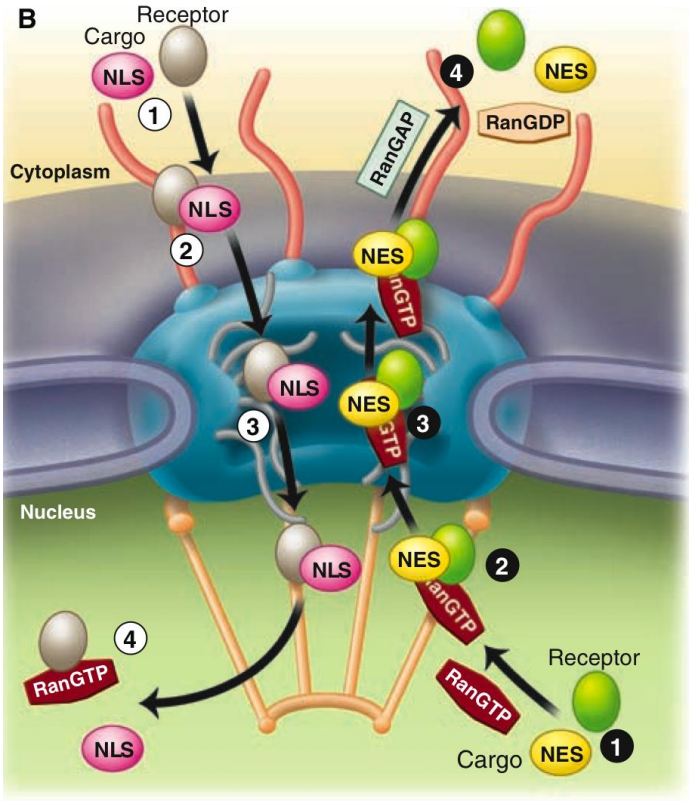

Figure 1-1 Nucleocytoplasmic transport across the nuclear envelope. [Adapted from (Terry, Shows et al. 2007)]

Nuclear Pore Complex

In order to accommodate nucleocytoplasmic transport, the outer and inner nuclear membranes of the NE fuse at specific sites to form aqueous pores, where proteinaceous structures of the NPCs are embedded(D'Angelo and Hetzer 2006). The NPC is probably one of the largest protein complexes in eukaryotic cells. It has a molecular weight of ~60-125 MDa in mammals (Reichelt, Holzenburg et al. 1990) and ~40-60 MDa in yeasts (Rout and Blobel 1993; Yang, Rout et al. 1998). The NPC measures about 100-150 nm in diameter and 50-70 nm in thickness under the electronic microscope (EM) (Ryan and Wente 2000; Lim, Ullman et al. 2008). The NPC is a cylindrical structure with eight-fold rotational symmetry, and its overall structure is evolutionarily conserved from yeasts to mammals. A core scaffold surrounds a central channel in the NE-embedded portion of the NPC, with eight filaments emanating from the scaffold to the cytoplasm and nucleoplasm respectively. The cytoplasmic filaments have loose ends, but the nuclear ones are connected in a distal ring, forming a structure called nuclear basket (Reviewed in(Lim and Fahrenkrog 2006; D'Angelo and Hetzer 2008; Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010)).

Recent advance of new electron microscopy technology have generated higher resolution 3-dimensional views of the NPC (Alber, Dokudovskaya et al. 2007). This giant complex is modular and composed of spokes and rings. There are only about 30 different proteins known as Nucleoporins (Nups) in the NPC, each of which is present in multiples of eight copies due to the structural symmetry. These Nups are associated with each other to form relatively stable subcomplexes, which are considered the “building blocks” for the

NPC(Alber, Dokudovskaya et al. 2007),(Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010). Nups can be divided according to their locations and functions into four classes: transmembrane, core scaffold, linker and so-called FG Nups that contain distinct phenylalanine-glycine (FG), GLFG (L, leucine), or FxFG (x, any) repeats (Alber, Dokudovskaya et al. 2007),(Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010). Three transmembrane Nups (Ndc1, Pom152 and Pom34 in yeast, Gp210, Ndc1 and Pom121 in vertebrate) span the pore membrane, the specified NE region where the outer and inner membranes fuse together. They form an outer luminal ring that interact with the core scaffold to anchor the NPC (orange ring in Figure 1-2). (Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010). The core scaffold is comprised of two inner rings associated with two outer rings, one on the cytoplasmic side and the other on the nucleorplasm side. The inner rings mainly contain the Nup170 complex (yeast) or the Nup155 complex (vertebrate) (purple rings in Figure 1-2), whereas the outer rings contain the Nup84 complex (yeast) or Nup107 complex (vertebrate) (yellow rings in Figure 1-2) (Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010). The core scaffold Nups comprise about half the NPC mass and cover the highly curved portion of the pore membrane, giving the NPC its shape and stabilizing the NE. Linker Nups (Nup82 and Nic 96 in yeast, Nup88 and Nup93 in vertebrate) connect the inner and outer rings, and provide the major attachment sites for the FG Nups (cyan rings in Figure 1-2) (Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010). These rings and linkers are aligned to form eight perpendicular spokes that surround the central pore, which is filled up with mostly symmetrically distributed FG Nups (green and red fibers in Figure 1-2) (Strambio-De-Castillia, Niepel et al. 2010; Wente and Rout 2010). This last

class of Nups contains distinct phenylalanine-glycine (FG), GLFG (L, leucine), or FxFG (x, any) repeats that are interspersed with charged or polar spacer sequences (Rout and Wente 1994; Lim, Huang et al. 2006). These FG regions have been shown to be structurally disordered regions (Denning, Patel et al. 2003). FG Nups are the docking sites for transport complexes and directly mediate translocation of macromolecules through the NPC (Bayliss, Leung et al. 2002; Grant, Neuhaus et al. 2003; Isgro and Schulten 2005; Liu and Stewart 2005). Removal of the FG regions or blocking its binding to transport receptors leads to disruption of nucleocytoplasmic transport (Strawn, Shen et al. 2004; Terry, Shows et al. 2007).

The composition of the NPC is quite dynamic. Fluorescence recovery after photobleaching (FRAP) experiments, using GFP-tagged Nups have shown that residence times of individual Nups varies greatly (Rabut, Doye et al. 2004). Scaffold Nups are relatively stable during interphase with residence times that are slightly longer than the average cell cycle. In contrast, periphery FG Nups turn over faster with residence times of seconds to minutes. Linker Nups that connect the scaffold and FG Nups have intermediate residence times. It has been suggested that the mobility of Nups may help deliver transport complexes to the NPCs (Griffis, Craige et al. 2004). Alternatively, such mobility may reflect changes in NPC composition in response to different transport requirements. The discovery of tissue or developmental-specific Nups (Fan, Liu et al. 1997; Cai, Gao et al. 2002; Olsson, Scheele et al. 2004) provided support for this hypothesis. It still remains an interesting question for further investigation.

In recent years, numerous high-resolution structures of Nup domains and Nup complexes have become available(Brohawn, Partridge et al. 2009; Strambio-De-Castillia, Niepel et al. 2010). Domain analysis and three-dimensional structures show that scaffold Nups are exclusively formed from domains containing β-propeller or α-solenoid motifs, or a specific combination of both (*e.g.* β-propeller at the amino-terminus followed by a α-solenoid at the carboxyl-terminus) (Devos, Dokudovskaya et al. 2004; Devos, Dokudovskaya et al. 2006; Brohawn, Partridge et al. 2009; DeGrasse, DuBois et al. 2009). Such architectural organization resembles other membrane-associated complexes such as the clathrin coat in endocytosis as well as the COPI and COPII coats in vesicular transport (Devos, Dokudovskaya et al. 2004; Devos, Dokudovskaya et al. 2006). This finding suggests that the NPC and vesicle coats may have originated from a common ancestral membrane-coating module that may have allowed early eukaryotes to form intracellular membrane systems to distinguish them from the prokaryotes (Devos, Dokudovskaya et al. 2004; Devos, Dokudovskaya et al. 2006; DeGrasse, DuBois et al. 2009).

Although numerous atomic resolution Nup structures are now available, the fundamental problems of how the NPC maintain selective permeability and the mechanism of translocation remain unresolved. The idea that periphery FG Nups in the central channel play a major role is widely accepted. Based on observations of Nup properties, several models of translocation have been proposed. The "virtual gate model" suggests the existence of an energetic barrier rather than a physical barrier (Rout, Aitchison et al. 2003). Larger molecules would lose more of their entropy when they enter the narrow

channel crowded by extended FG repeats, which means a higher entropic barrier. This entropic penalty can be paid off through the binding of transport receptors to the Nups thus providing kinetic advantages for cargos that are bound to receptors. The "oily spaghetti model" stems from a similar idea that extended FG repeats are constantly moving in the central channel and transport complexes can push the FG spaghetti to one side and pass through by a binding-release mechanism (Macara 2001). In contrast, the "selective phase model" proposes that the FG Nups form weak hydrophobic interactions with each other to form a sieve-like meshwork that mechanically restricts the passage of molecules larger than the pore of the meshwork (Ribbeck and Görlich 2001). The binding of transport receptors to the FG repeats is proposed to dissolve the meshwork, allowing selective partitioning of transport receptors into this FG Nups phase. Finally, a "reduction of dimensionality model" proposes the existence of selective filter formed by FG Nups in the central channel, and only transport complexes that bind the continuous FG surface could enter the filter and slide through like ferries (Peters 2005). There is substantial disagreement and controversy with regard to these models of translocation. No single model is sufficient to explain all the observed NPC properties. It is likely that a combination of these models would be required to explain the NPC gating mechanism.

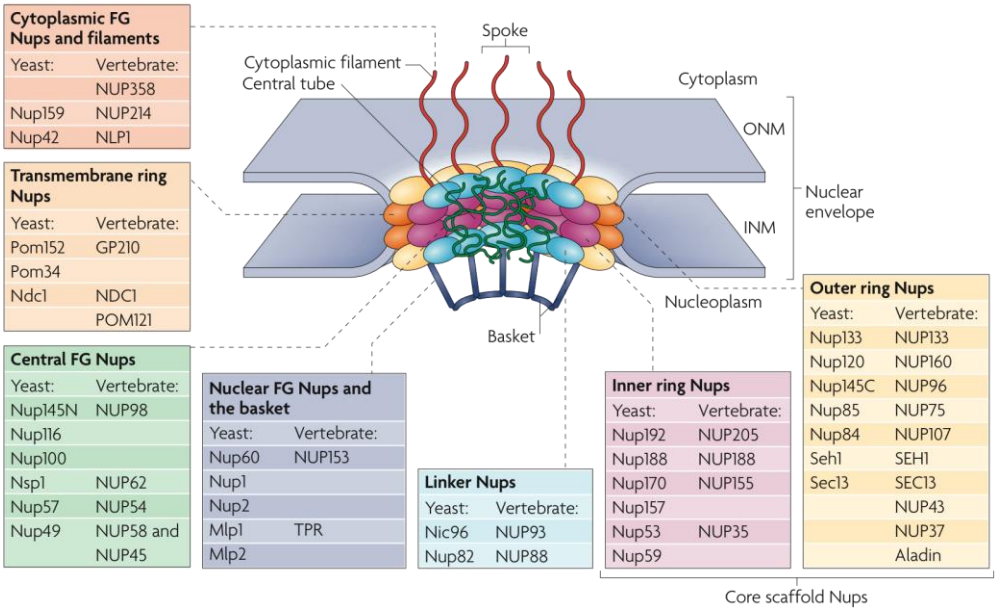

Figure 1-2 Structural model of nuclear pore complex (NPC). The double layer nuclear membrane (light grey sheets) fuses at the nuclear pore. ONM, outer nuclear membrane; INM, inner nuclear membrane. NPC is composed of transmembrane ring, core scafford (outer and inner rings), linker ring and FG Nups filling the central pore. The protein components of each ring are listed**.** [Adapted from (Strambio-De-Castillia, Niepel et al. 2010)]

The Ran GTPase System

The Ran GTPase system provides directionality and energy for nuclear transport. The system includes Ran itself (Gsp1p and Gsp2p in yeast)(Drivas, Shih et al. 1990; Bischoff and Ponstingl 1991; Belhumeur, Lee et al. 1993; Kadowaki, Goldfarb et al. 1993), the guanine nucleotide exchange factor RCC1 (Prp20p in yeast)(Ohtsubo, Kai et al. 1987; Aebi, Clark et al. 1990; Bischoff and Ponstingl 1991), the RanGTPase-activating protein RanGAP1 (Ran1p in yeast)(Atkinson, Dunst et al. 1985; Bischoff, Klebe et al. 1994;

Bischoff, Krebber et al. 1995; Corbett, Koepp et al. 1995), the Ran-binding protein RanBP1 (Yrb1p in yeast) (Coutavas, Ren et al. 1993; Butler and Wolfe 1994; Beddow, Richards et al. 1995; Bischoff, Krebber et al. 1995; Schlenstedt, Wong et al. 1995)and homologous RanBD domains in nucleoporin RanBP2 (also known as Nup358) (Wu, Matunis et al. 1995; Yokoyama, Hayashi et al. 1995) and the nuclear transport factor 2 (NTF2) (Moore and Blobel 1994; Paschal and Gerace 1995; Corbett and Silver 1996; Nehrbass and Blobel 1996).

Ran is a member of the evolutionarily conserved Ras superfamily of small GTPases. Like other small GTPases, Ran has a core catalytic or G-domain composed of five alpha helices (A1-A5), six beta-strands (B1-B6) and five polypeptide loops (G1-G5) (Bourne, Sanders et al. 1991; Scheffzek, Klebe et al. 1995; Chook and Blobel 1999; Vetter, Arndt et al. 1999; Vetter, Nowak et al. 1999). In addition, Ran has a C-terminal extension that consists of an unstructured linker and a 16-residue α-helix (Nilsson, Weis et al. 2002). Ran exists in two nucleotide bound states: RanGDP and RanGTP (Bourne, Sanders et al. 1990). Nucleotide-free Ran is thermodynamically unstable (Klebe, Prinz et al. 1995; Klebe, Ralf Bischoff et al. 1995). Structural comparison between RanGDP and RanGTP has revealed that three regions that undergo nucleotide-dependent conformation changes: the Switch I and II regions, which interact with the bound nucleotide, and the C-terminal extension (Scheffzek, Klebe et al. 1995; Chook and Blobel 1999; Vetter, Arndt et al. 1999; Vetter, Nowak et al. 1999). In RanGDP, the C-terminal extension packs against the G-domain(Scheffzek, Klebe et al. 1995). In RanGTP, the extension is moved away from the core(Chook and Blobel 1999; Vetter, Arndt et al. 1999; Vetter, Nowak et al. 1999).

Ran has very low intrinsic rates of GTPase hydrolysis and nucleotide exchange, and thus requires regulators to obtain full of GTPase activity (Bischoff and Ponstingl 1991; Klebe, Prinz et al. 1995). Nucleotide exchange from RanGDP to RanGTP is accelerated ~ 10^5–fold by the exchange factor RCC1 as RCC1 stabilizes the intermediate nucleotide-free Ran (Bischoff and Ponstingl 1991; Klebe, Prinz et al. 1995) (Bischoff and Ponstingl 1995). Given the high GTP: GDP ratio in cells, removal of GDP will result in the production of RanGTP. RanGAP1 catalyzes the hydrolysis of RanGTP to RanGDP by enhancing the intrinsic GTPase activity of Ran 10^5–fold (Bischoff, Klebe et al. 1994; Becker, Melchior et al. 1995; Bischoff and Ponstingl 1995). RanGAP stimulated GTPase activity can be further upregulated about 10-fold by RanBP1 (Bischoff, Krebber et al. 1995; Richards, Lounsbury et al. 1995). RanBP2 is a large 358 kD nucleoporin in higher eukaryotes that contains four RanBP1-like domains that behaves like RanBP1 (Wu, Matunis et al. 1995; Yokoyama, Hayashi et al. 1995). RanBP1 is located in the cytoplasm and RanBP2 is located at the cytoplasmic filaments of the NPC (Schlenstedt, Wong et al. 1995; Wu, Matunis et al. 1995; Yokoyama, Hayashi et al. 1995; Matunis, Coutavas et al. 1996; Richards, Lounsbury et al. 1996; Mahajan, Delphin et al. 1997). In contrast, RCC1 is associated the chromosomes and resides exclusively in the nucleus (Ohtsubo, Okazaki et al. 1989). The asymmetric distribution of the Ran regulators produces high concentrations of RanGTP in the nucleus and RanGDP in the cytoplasm. This RanGTP gradient generates the directionality of nucleocytoplasmic transport (Görlich, Pantéet al. 1996; Izaurralde, Kutay et al. 1997). RanGTP binds importins in the nucleus to release import cargos and the RanGTP-importin complexes are recycled back to the cytoplasm(Chook and Blobel 1999; Cingolani, Petosa et al. 1999; Vetter, Arndt et al.

1999; Lee, Cansizoglu et al. 2006). Upon RanGTP hydrolysis with the help of RanGAP1 and RanBP1, RanGDP dissociates from importins and are again available for cargo-loading (Bischoff, Klebe et al. 1994; Becker, Melchior et al. 1995; Bischoff and Ponstingl 1995; Bischoff and Görlich 1997; Floer, Blobel et al. 1997; Gorlich, Dabrowski et al. 1997; Lounsbury and Macara 1997). For nuclear export, RanGTP and export cargos bind exportins cooperatively in the nucleus to form export complexes (Bohnsack, Regener et al. 2002; Dong, Biswas et al. 2009; Dong, Biswas et al. 2009). Upon entering the cytoplasm, conversion of RanGTP to RanGDP disassembles the export complexes and releases export cargos (Bischoff and Görlich 1997; Kutay, Ralf Bischoff et al. 1997). This way, transport receptors can achieve multiple rounds of unidirectional transport. Although nucleocytoplasmic transport is an active process, translocation through the NPC per se does not involve nucleotide hydrolysis (Schwoebel, Talcott et al. 1998; Englmeier, Olivo et al. 1999; Ribbeck, Kutay et al. 1999). Only one GTP is consumed by hydrolysis in the cytoplasm, to regenerate unliganded import-karyopherin for a new round of import and to dissociate export complexes to terminate one round of nuclear export (Gorlich and Kutay 1999; Kehlenbach, Dickmanns et al. 1999).

Ran is a predominantly nuclear (Bischoff and Ponstingl 1991) but the continuous efflux of RanGTP with transport receptors from the nucleus and subsequent release as RanGDP in the cytoplasm depletes nuclear levels of RanGTP. Ran must be reimported into the nucleus rapidly to continue the transport cycle. Nuclear transport factor 2 (NTF2), a 15 KDa homodimeric protein, imports Ran into the nucleus (Grundmann, Nerlich et al. 1988; Moore and Blobel 1994; Paschal and Gerace 1995; Corbett and Silver 1996;

Ribbeck, Lipowsky et al. 1998; Smith, Brownawell et al. 1998). NTF2 binds RanGDP, the predominant form of Ran in cytoplasm, with high affinity (Clarkson, Kent et al. 1996; Nehrbass and Blobel 1996; Paschal, Delphin et al. 1996). The preference for RanGDP is due to a steric clash of NTF2 with the switch regions of Ran in GTP state (Stewart, Kent et al. 1998). NTF2 also binds the FG repeats in FG Nups to mediate the translocation of RanGDP through the NPC (Ribbeck, Lipowsky et al. 1998; Smith, Brownawell et al. 1998). In the nucleus, RanGDP is dissociated from NTF2 for nucleotide exchange simulated by RCC1, which irreversibly terminates Ran import (Renault, Kuhlmann et al. 2001).

Karyopherin Family

Karyopherins are a group of homologous proteins that recognize macromolecular cargos either (or both) the nucleoplasm or the cytoplasm, and aid their transport in or out of the nucleus (Chook and Blobel 2001). The name Karyopherin originates from the Greek "karyon", which means nucleus and "pher(ein)" which means bringing to or carrying from (Radu, Blobel et al. 1995; Wozniak, Rout et al. 1998). The Karyopherinβ (Kapβ) family of transport receptors includes 14 members in yeast and 19 members in human (Chook and Blobel 2001; Fried and Kutay 2003; Mosammaparast and Pemberton 2004). Evolutionary analysis divides them into 15 subfamilies as shown in Table 1-1 (Quan, Ji et al. 2008; Chook and Suel 2010). Kapβs share similar molecular weight (90-150 KDa) and isoelectric points (4.0-5.0), but have low overall sequence similarity (15%-20% identity) (Gorlich, Dabrowski et al. 1997; Chook and Blobel 2001). Kapβs are made of 19-20 multiple tandem helical repeats called HEAT (Huntingtin, elongation factor 3,

PR65/A subunite of protein phosphatase 2A and the TOR lipid kinase) repeats (Hemmings, Adams-Pearson et al. 1990; Madrid and Weis 2006; Suel, Cansizoglu et al. 2006; Cook, Bono et al. 2007). Each HEAT motif consists of a pair of antiparallel α-helices connected with a loop segment and is stacked against each other in a parallel fashion to form superhelical or ring-shaped structures (Madrid and Weis 2006; Suel, Cansizoglu et al. 2006; Cook, Bono et al. 2007). Entire Kapβs form single domains with contiguous hydrophobic cores that can be roughly divided into functional regions such as the Ran binding or cargo binding regions (Cook, Bono et al. 2007). The Ran binding regions at the N-terminal 150 residues are the most conserved regions among Kapβs, indicating that Ran is a general regulator for Kapβ function (Quan, Ji et al. 2008). Depending on the direction of cargo transport, Kapβs can be classified as importins, exportins or bidirectional transporters (Mosammaparast and Pemberton 2004).

Importins bind the cargos in the cytoplasm via their nuclear localization signals or NLSs in the cargos. There are 11 importins in human and 10 in yeast (Mosammaparast and Pemberton 2004). The best-characterized import pathway is the so-called classical import pathway that uses the Impβ/Impα heterodimer (Conti and Izaurralde 2001). Impα functions as an adaptor for Impβ. Impα consists of a flexible N-terminal Importin-b-binding (IBB) domain (Görlich, Henklein et al. 1996),(Weis, Dingwall et al. 1996) and a helical ARM domain with 10 armadillo (ARM) repeats (Herold, Truant et al. 1998). The ARM domain binds classical NLSs (cNLSs), which are short stretches of basic residues. The monopartite cNLS has a single stretch of basic residues (consensus K-K/R-X-K/R, X is any amino acid) (Kalderon, Roberts et al. 1984) and the bipartite cNLS has two

stretches of basic residues connected by a linker (loose consensus (K/R)(K/R)$X_{10–12}$ $(K/R)_{3/5}$, where $(K/R)_{3/5}$ represents three lysine or arginine residues out of five consecutive amino acids) (Robbins, Dilworth et al. 1991). In unliganded Impα, the IBB domain is autoinhibitory as it covers the c-NLS binding site of ARM domain (Cingolani, Petosa et al. 1999; Cingolani, Lashuel et al. 2000). NLS binding displaces the IBB domain to bind Impβ to form a ternary Impβ-Kapα-cNLS import complex (Cingolani, Lashuel et al. 2000) (Cingolani, Petosa et al. 1999).

Kapβ2 or transportin, is a prototypical karyopherin that binds its cargos directly (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007; Imasaki, Shimizu et al. 2007). It is the second characterized import pathway and the mechanism of cargo recognition by Kapβ2 is described in a separate section below. Other importins are currently known to bind the cargos with highly diverse sequences and different conformations (Chook and Suel 2010). It remains extremely difficult to identify the common characteristics of those cargos that can be classified into new NLSs.

Exportins bind their cargos in the nucleus in the presence of RanGTP via nuclear export signals (NESs). There is only one known type of NES so far, which is recognized by the exportin CRM1(Fornerod, Ohno et al. 1997; Fukuda, Asano et al. 1997; Neville, Stutz et al. 1997; Ossareh-Nazari, Bachelerie et al. 1997; Stade, Ford et al. 1997). The so-called leucine-rich NESs are 10-15 residues long and composed of 3-4 regularly spaced hydrophobic residues. The leucine-rich NES can be described by the consensus sequence

of $Ø_1$-X_{2-3}- $Ø_2$ –X_{2-3}- $Ø_3$ -X- $Ø_4$ ($Ø_n$ represents L, V, I, F or M; and X can be any amino acid)(Dong, Biswas et al. 2009; Dong, Biswas et al. 2009).

All Kapβs, including Impβ, can bind their cargos directly and each Kapβ recognizes a subset of cargos to create distinct transport pathways (Chook and Suel 2010). But due to small number of known cargos and the absence of specific inhibitors for individual Kapβs, we know little about other transport pathways. The detailed mechanism of signal recognition by Karyopherins from structural prospect is discussed in chapter 2.

In addition to mediating nuclear transport, Karyopherins have also been found to play important roles in other cellular functions, such as mitosis (Gruss, Carazo-Salas et al. 2001; Nachury, Maresca et al. 2001; Wiese, Wilde et al. 2001), assembly of the nuclear pore complex (Harel, Chan et al. 2003).

Table 1-1 Karyopherinβ Family of Proteins

Subfamily	Human	Yeast
IMB1	Importin-β/Kapβ1	Kap95p
IMB2	Kapβ2/Transportin	Kap104p
IMB3	Importin-5/RanBP5/Kapβ3	Kap121p/Pse1p
IMB4	Importin-4/RanBP4	Kap123p
IPO8	Importin-7/RanBP7	Kap119p/Nmd5p
	Importin-8/RanBP8	Kap108p/Sxm1p
IMB5	Importin-9	Kap114p
KA120	Importin-11	Kap120p
TNPO3	Transportin-SR/SR2/-3/TNPO3	Kap111p/Mtr10p
	Importin-13	Kap122p/Pdr6p
XPO4	Exportin-4	
XPO5	Exportin-5	Kap142p/Msn5p
XPO6	Exportin-6	
XPO7	Exportin-7/RanBP16	
XPOT	Exportin-t/Xpo-t	Los1p
XPO1	CRM1/Exportin1	CRM1p/Xpo1p
XPO2	CAS	Cse1p

Recognition of the PY-NLSs by Kapβ2

Kapβ2 contains 20 HEAT repeats that form a perfect superhelix (Figure 1-3) (Lee, Cansizoglu et al. 2006). More than 20 mRNA binding proteins have been reported as the cargos of Kapβ2, including hnRNPs, A1, D, F, M, HuR, DDX3, YBP1, NXF1) (Pollard, Michael et al. 1996; Bonifaci, Moroianu et al. 1997; Siomi, Eder et al. 1997; Fan and Steitz 1998; Truant, Kang et al. 1999; Kawamura, Tomozoe et al. 2002; Guttinger, Muhlhausser et al. 2004; Rebane, Aab et al. 2004; Suzuki, Iijima et al. 2005; Lee,

Cansizoglu et al. 2006; Chook and Suel 2010). The structure of Kapβ2 in complex with its best-known substrate hnRNP A1-NLS (also called M9NLS) demonstrates that the NLS binds in extended conformation to the concave surface of the C-terminal arch of Kapβ2 (Figure1-3A) (Lee, Cansizoglu et al. 2006). The large flat NLS-binding interface on Kapβ2 is highly acidic and mixed with hydrophobic patches (Figure 1-3B), suggesting the preference of overall positive charged NLSs (Lee, Cansizoglu et al. 2006). Sequence examination of the known NLSs of Kapβ2 identified two conserved regions: 1) the C-terminal PY motif preceded by a basic residue within 2-5 residues; 2) the N-terminal hydrophobic or basic motif (Lee, Cansizoglu et al. 2006). Collectively, these physical characteristics lead to the discovery of a new type of NLSs named PY-NLSs. The PY-NLSs recognized by Kapβ2 are structurally disordered and have overall positive charges. They contain an N-terminal hydrophobic or basic motif followed by a C-terminal R/H/K$X_{(2-5)}$PY consensus motif (Figure 1-3C)(Lee, Cansizoglu et al. 2006). The PY-NLSs can further divided into two subclasses based on the N-terminal motifs: the hydrophobic PY-NLSs (hPY) and basic PY-NLSs (bPY-NLSs) (Figure 1-3C) (Lee, Cansizoglu et al. 2006). However, the NLS-binding site is occupied by the acidic H8 loop of Kapβ2 in the structure of Kapβ2-RanGTP (Chook and Blobel 1999). Thus, the binding of RanGTP in the N-terminal arch of Kapβ2 induces structure changes of Kapβ2 that are incompatible with cargo-binding and causes the dissociation of substrates (Chook and Blobel 1999; Lee, Cansizoglu et al. 2006).

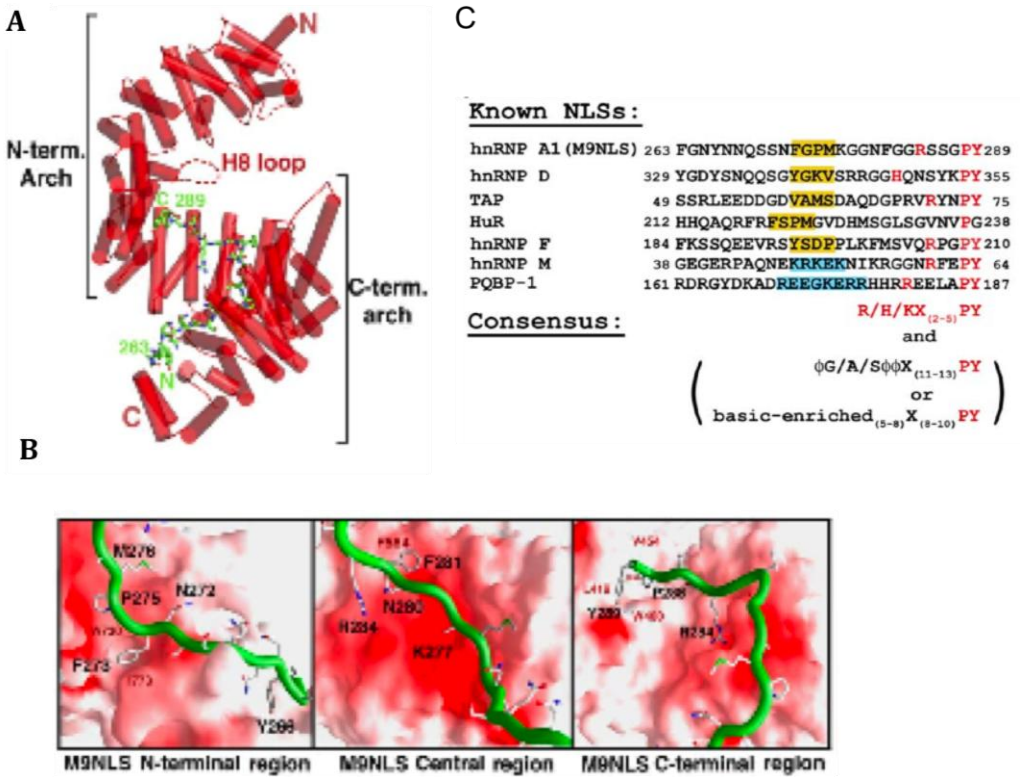

Figure 1-3 The structure of Kapβ2-hnRNP A1-NLS. (A) The ribbon Ribbon diagram of the Kapb2-M9NLS complex with Kapb2 in red (a helices represented as cylinders and structurally disordered loops as red dashes) and M9NLS shown as a stick figure (carbon: green, oxygen: red, nitrogen: blue, and sulfur: orange). (B) The Kapβ2-M9NLS interface. The N-terminal third (left), the central region (middle), and the C-terminal third (right) of M9NLS. Substrate is shown as a green ribbon and the Kapβ2 electrostatic potential is mapped onto its surface, all drawn using GRASP (Nicholls et al., 1991). Red indicates negative electrostatic potential, white neutral, and blue positive. Residues in the hydrophobic patches of Kapβ2 are labeled in red and M9NLS residues labeled in black. (C) The consensus sequence of PY-NLSs. (Lee, Cansizoglu et al. 2006)

mRNA export

Export of messenger RNAs (mRNAs) is more complex than nuclear transport of proteins since the former is linked with the upstream and downstream events (Erkmann and Kutay 2004; Vinciguerra and Stutz 2004; Cole and Scarcelli 2006; Kohler and Hurt 2007; Stewart 2010). Following transcription from DNA templates, nascent pre-mRNAs associate with numerous proteins to form mRNP complexes that then undergo a series of processing such as 5'-capping, splicing, and 3'–polyadenylation (Erkmann and Kutay 2004; Vinciguerra and Stutz 2004; Cole and Scarcelli 2006; Kohler and Hurt 2007; Stewart 2010). Only mature mRNPs that have completed these remodeling processes are ready to form export complexes. The exact mechanism of how the mature mRNPs are recognized is still unclear, but it is obvious that the recruitment of NXF1/NXT1 (or Tap/p15 in human, Mex67/Mtr2 in yeast) heterodimer is critical for mRNA export (Segref, Sharma et al. 1997; Grüter, Tabernero et al. 1998; Herold, Klymenko et al. 2001). NXF1 is the major mRNA export factor and is highly conserved from yeast to human (Herold, Suyama et al. 2000). NXF1 is not related to Karyopherins. Instead, it is a modular protein with four globular domains: the RNA binding (RBD), Leucine-rich (LRR), NTF2-like (NTF2-L) and ubiquitin-associated (UBA) domains (Liker, Fernandez et al. 2000; Fribourg, Braun et al. 2001; Grant, Hurt et al. 2002; Ho, Coburn et al. 2002; Fribourg and Conti 2003; Senay, Ferrari et al. 2003; Stutz and Izaurralde 2003). Even though NXF1 has an RNA binding domain (Braun, Rohrbach et al. 1999), it is recruited to the mRNPs by adaptor proteins such as REF/Aly/Sub2, EJC components or SR proteins (Bachi, Braun et al. 2000; Strasser, Bassler et al. 2000; Stutz, Bachi et al. 2000;

Huang, Gattoni et al. 2003; Aguilera 2005; Reed and Cheng 2005; Hautbergue, Hung et al. 2008) (Figure 1-4). The NXF1/NXT1 heterodimer also binds FG Nups for translocation of the mRNP export complex through the NPC (Santos-Rosa, Moreno et al. 1998; Katahira, Strasser et al. 1999; Fribourg, Braun et al. 2001; Grant, Hurt et al. 2002; Senay, Ferrari et al. 2003) (Figure 1-4). In the cytoplasm, DEAD-box helicase Dbp5, Gle1 and inositol phosphate IP_6 cooperate to remove NXF1 from mRNPs to end the mRNA export process (Tseng, Weaver et al. 1998; York, Odom et al. 1999; Lund and Guthrie 2005; Alcázar-Román, Tran et al. 2006; Weirich, Erzberger et al. 2006) (Figure 1-4). NXF1/NXT1 is then reimported into the nucleus for a new round of mRNA export (Bear, Tan et al. 1999; Braun, Rohrbach et al. 1999; Kang and Cullen 1999; Katahira, Strasser et al. 1999; Truant, Kang et al. 1999; Bachi, Braun et al. 2000). Even though mRNA export is distinct from Karyopherin-mediated transport, both processes four common steps: 1) cargo recognition and transport complex assembly in the initial compartment; 2) translocation through the NPC; 3) disassembly of the transport complex in the target complex followed by removal of the carrier; 4) recycling of the carrier back to the initial compartment for a new round of transport.

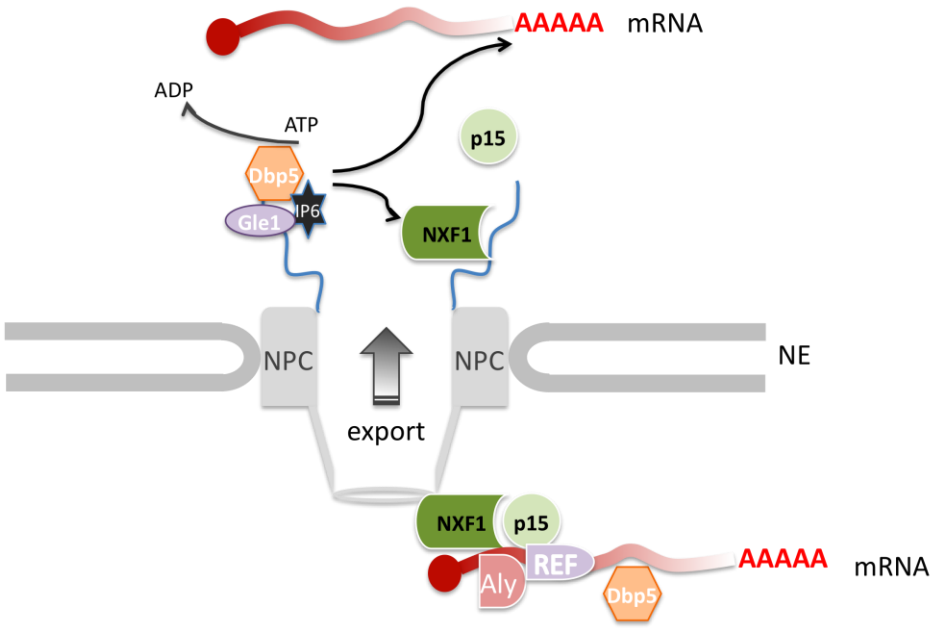

Figure 1-4 mRNA export pathway.

Conclusion

Despite the obvious importance of nucleocytoplasmic transport in cellular function, we still lack understanding of the mechanistic aspects of this fundamental process. Most of the known cargos are for Impβ/a, Kapβ2 and CRM1. Cargos for many Kapβs remain undiscovered. Cargo recognition mechanisms for most Kapbs remain unclear. Even for the better studied Kaps like Impβ/α, Kapβ2 and CRM1, large binding interfaces and the flexible nature of the receptors allow them to accommodate diverse cargos in different ways, suggesting that there must be more than one recognition mechamism for each Kapβ. Some cargos are also transported by more than one Kap and the existence of redundant pathways for individual Kaps adds further complexity to the

nucleocytoplasmic transport process. Increasing cargo repertoires for individual Kaps and developing pathway-specific inhibitors will greatly help elucidate the cargo recognition mechanisms. A combination of bioinformatics, biochemistry, biophysics and cell biological approaches will be required to achieve new goals.

CHAPTER TWO

STRUCTURE BASED DESIGN OF A PATHWAY SELECTIVE NUCLEAR IMPORT INHIBITOR*

Abstract

Kapβ2 recognizes PY nuclear localization signal (NLS), a new class of NLS with a R/H/K$x_{(2–5)}$PY motif. The structural and biochemical studies of Kapβ2 with hnRNP A1 and M NLSs led to the design of the M9M peptide, a Kapβ2-specific inhibitor. In this chapter, I demonstrated that M9M specifically mislocalized the Kapβ2 cargos, hnRNP A1, HuR and hnRNP M into the cytoplasm, but has no effect on HDAC1, a cargo for Impα/β pathway. As the first pathway-specific inhibitor for nuclear import, M9M is a valuable tool to study Kapβ2-mediated nuclear import or other cellular functions.

Introduction and Background

Ten different importins mediate trafficking of human proteins into the cell nucleus through recognition of distinct NLSs. Large panels of import substrates are known only for Impβ and Kapβ2 (Mosammaparast and Pemberton 2004; Lee, Cansizoglu et al. 2006). The substrate repertoire of each Kapβ and the functional consequences of pathway

---

* Originally published in Nat Struct Mol Biol, (2007) 14(5):452-4. Copyright by Nature Publishing Group

specificities are some of the main challenges in understanding intracellular signaling and trafficking. In the case of nuclear export, CRM1 inhibitor leptomycin B has been crucial for identifying many CRM1 substrates. Such specific inhibitors of nuclear import could be invaluable for proteomic analyses to map extensive nuclear traffic, but none has been found.

Two classes of NLS are currently known: short, basic classical NLSs that bind the heterodimer Impα/β (Dingwall and Laskey 1991; Mosammaparast and Pemberton 2004), and newly identified PY-NLSs that bind Kapβ2 (Lee, Cansizoglu et al. 2006). PY-NLSs are 20- to 30-residue signals with intrinsic structural disorder, overall basic character, C-terminal R/K/H$X_{2–5}$PY motifs (where $X_{2–5}$ is any sequence of 2–5 residues) and N-terminal hydrophobic or basic motifs. These weak but orthogonal characteristics have provided substantial limits in sequence space, enabling the identification of over 100 PY-NLS–containing human proteins (Lee, Cansizoglu et al. 2006). Two subclasses, hPY-NLSs and bPY-NLSs, are defined by their N-terminal motifs: hPY-NLSs contain ϕG/A/Sϕϕ motifs (where ϕ is a hydrophobic residue), whereas bPY-NLSs are enriched with basic residues.

Structural comparison of Kapβ2 with hnRNP A1 and M NLSs

The structures of human Kapβ2 bound to the hPY-NLS of heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and the bPY-NLS of human hnRNP M have been solved (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007) to understand how diverse hydrophobic or basic N-terminal motifs are recognized by Kapβ2. The two NLSs

trace different paths while lining a common interface on the structurally invariant Kapβ2 C-terminal arch (Figure 2-1, Kapβ2$_{435–780}$ Cα r.m.s. deviation is 0.9 Å). Upon Kapβ2 superposition, the NLSs converge structurally at three sites: the N-terminal motif and the arginine and proline-tyrosine residues of the R/H/Kx$_{(2–5)}$PY motif (Figure 2-1B). At the N-terminal motifs, hnRNP M residues 51–54 in the basic ^{50}KEKNIKR^{56} motif and hnRNP A1 residues 274–277 in the hydrophobic motif overlap (main chain r.m.s. deviation 1.3 Å). Residues 51–64 of hnRNP M and residues 273–289 of hnRNP A1 contact a common Kapβ2 surface, with the highest overlap at their PY motifs. R.m.s. deviations for all PY atoms and for arginine guanido group atoms in the R/H/Kx$_{(2–5)}$PY motifs are 0.9 Å and 1.2 Å , respectively (Figure 2-1B). In contrast, intervening segments ^{61}FE^{62} in hnRNP M and ^{285}SSG^{287} in hnRNP A1, and those between the N-terminal and R/H/Kx$_{(2–5)}$PY motifs, diverge up to 4.0 Å and 7.2 Å , respectively (Figure 2-1B). Thus, these sites are key binding epitopes, confirming their designation as consensus sequences, and the structurally variable linkers vary in both sequence and length across the PY-NLS family. The multivalent nature of the PY-NLS–Kapβ2 interaction probably allows modulation of binding energy at each site to tune overall affinity to a narrow range suitable for regulation by nuclear RanGTP. NLSs.

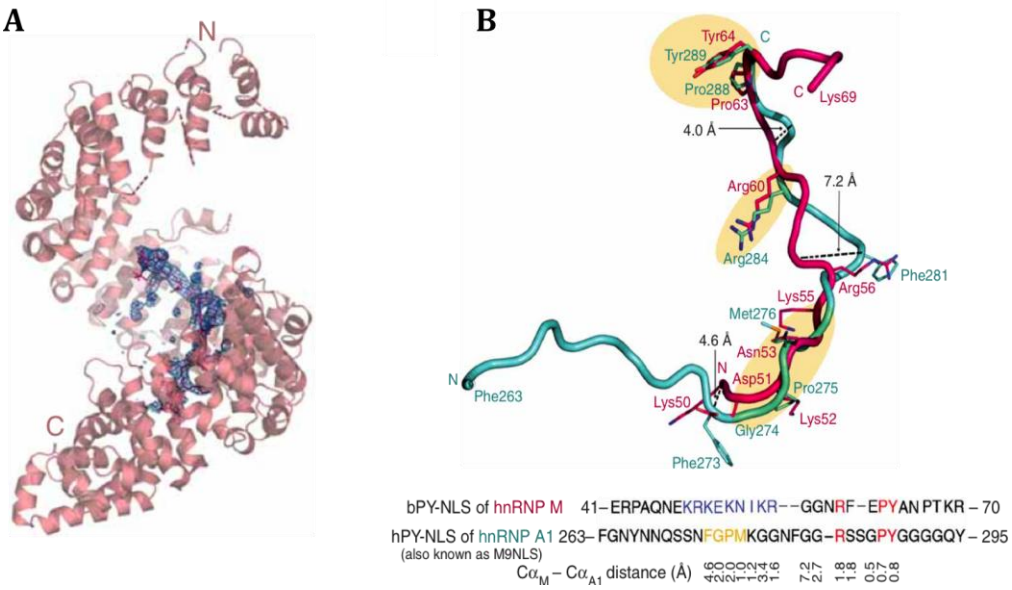

Figure 2-1 Kapβ2 bound to bPY-NLS of hnRNP M. (A) Ribbon model of Kapβ2 (pink), hnRNP M NLS (magenta) and the 2.5 σ Fo – Fc map (blue). (B) NLSs of hnRNP M (magenta) and hnRNP A1 (2H4M; blue) upon superposition of Kapβ2 residues 435–780. Regions of structural similarity are highlighted in yellow. Structurally aligned NLS sequences, Cα–Cα distances and inhibitor M9M sequence are shown. (Cansizoglu, 2007)

Distribution of Binding Energy along PY NLSs

Despite structural conservation of key motifs, the distribution of binding energy along PY-NLSs is very different. In hnRNP A1, Gly274 is the only binding hot spot (Nakielny, Siomi et al. 1996; Fridell, Truant et al. 1997; Bogerd, Benson et al. 1999), and the

energetic contribution from the C-terminal PY is modest (Iijima, Suzuki et al. 2006). In contrast, the only hnRNP M NLS hot spot is at its PY motif (Figure 2-2).

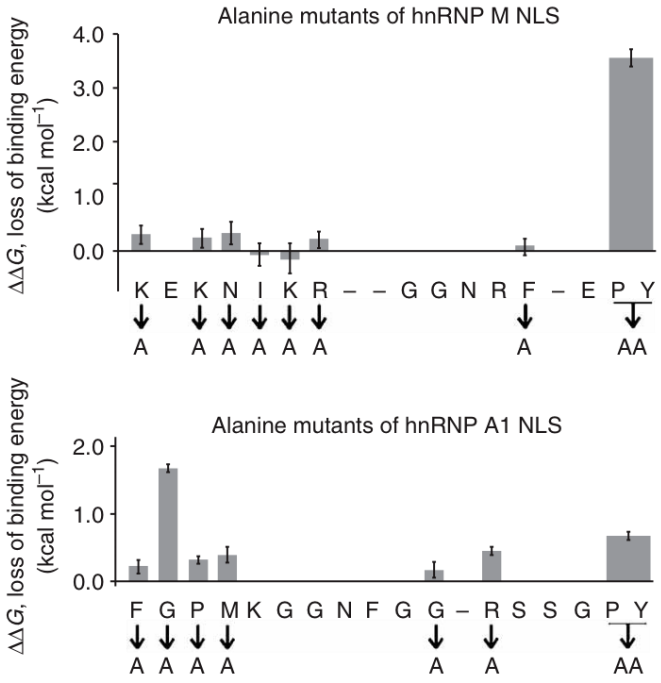

Figure 2-2 Loss of Kapβ2- binding energy in alanine mutants of hnRNP A1 (Lee, Cansizoglu et al. 2006) **and hnRNP M** ($\Delta\Delta G = -RT\ln(K_d(WT)/K_d(mutant))$; Kds determined by ITC).(Cansizoglu, Lee et al. 2007).

Design of pathway selective inhibitor M9M

Asymmetric locations of NLS hot spots in hnRNP A1 and hnRNPM, and the presence of variable linkers between the sites, allowed the design of chimeric peptides with enhanced Kapβ2-binding affinities. We designed a peptide named M9M, which fuses the N-terminal half of the hnRNP A1 NLS to the C-terminal half of the hnRNP M NLS and thus contains both binding hot spots (Figure 2-2 and 2-3). When bound to Kapβ2, M9M

shows decreased dissociation by RanGTP, competes effectively with wild-type NLS and binds specifically to Kapβ2 but not Impβ (Figure 2-4), thus behaving like a Kapβ2-specific inhibitor. The mechanism of inhibition is explained by the 200-fold tighter binding of M9M to the PY-NLS binding site of Kapβ2 (competition ITC shows K_d of 107 pM, compared with 20 nM for hnRNP A1 NLS) (Cansizoglu, Lee et al. 2007).

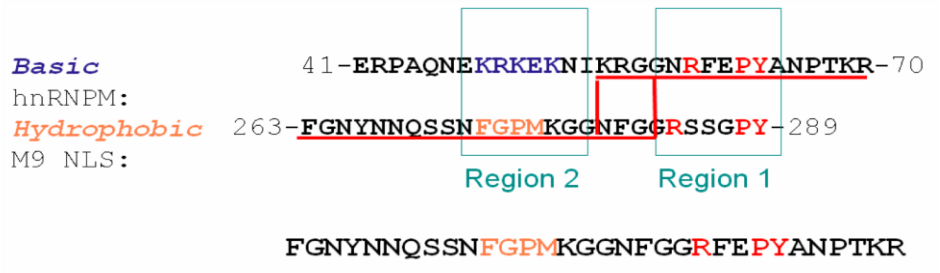

Figure 2-3 A chimeric peptide carrying both of the hotspots from hnRNP A1 and hnRNP M NLS sequences is constructed. Red lines correspond to two different chimeric peptides tested. Bottom panel is the sequence of the successful chimeric peptide

In this chapter, I describe my contribution to the discovery of M9M as a Kap β 2-specific inhibitor. I tested the inhibitory activity of M9M in the cells. The localization of several Kap β 2 cargos and an Imp α / β pathway cargo were examined by immunofluorescence after M9M transfection.

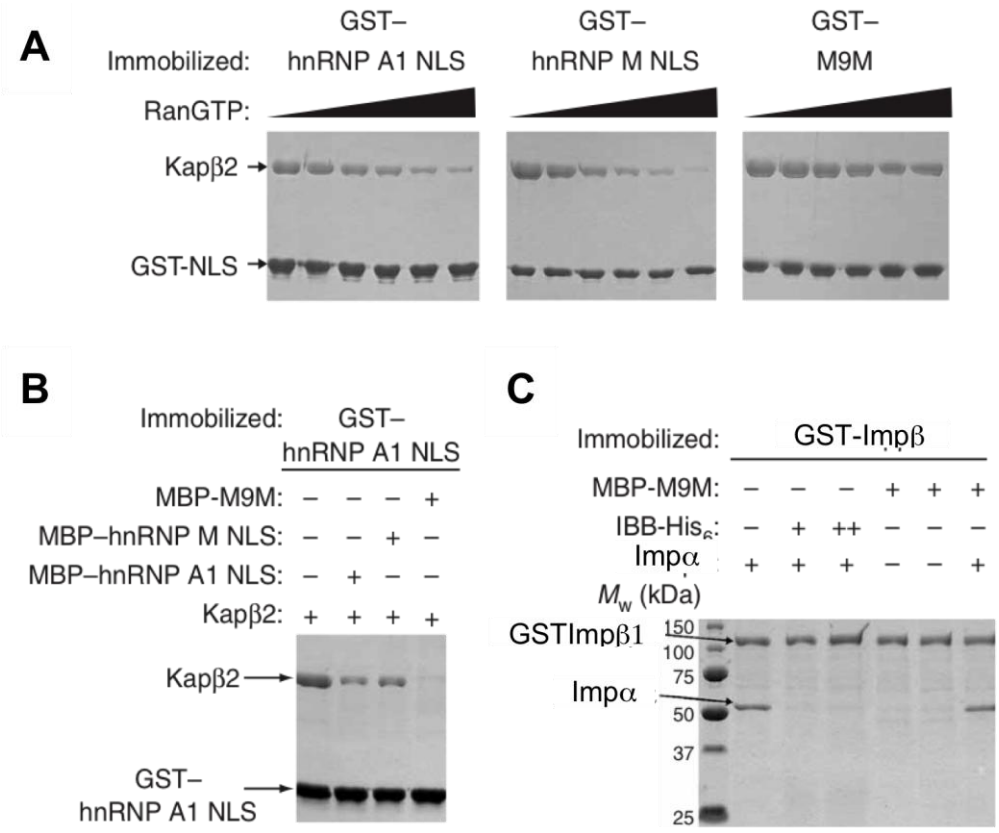

Figure 2-4 Competition Binding Assays for M9M. (A – C) Coomassie-stained gels of (A) immobilized GST fusions of hnRNP A1 NLS, hnRNP M NLS and M9M bound to Kapβ2 and then dissociated by 0.3 – 1.6 mM RanGTP; (B) immobilized GST-hnRNP A1-NLS bound to Kapβ2 and displaced by MBP-hnRNP A1-NLS, MBP-hnRNP M-NLS and MBP-M9M; (C) immobilized GST-Impβ1 bound to Impα and then competed with IBB-His_6 and MBP-M9M.

Materials and Methods

Cloning

The fragments of MBP, MBP-hnRNP A1-NLS and MBP-M9M from previous pMALTEV constructs were amplified and subcloned into the modified pCS2-MT mammalian vector at Sal I and Not I sites. The resulting constructs contain a 6-Myc tag at the N-terminal of the MBP fusion inserts.

Western blotting

For western blot analysis, MBP-hnRNP A1-NLS, MBP-hnRNP M-NLS, MBP-M9M proteins or HeLa lysates were resolved on SDS-PAGE, transferred to PVDF membrane and probed with monoclonal antibody 4C2 (a gift from Dr. Michael Matunis, John Hopkins Univ) diluted at 1:2000 and antibody 2A6 diluted at 1:1000 (a gift from Dr. Maurice Swanson, Univ of Florida.) Secondary horseradish peroxidase-conjugated anti-mouse antibody (diluted 1:10000, Amersham, NJ, USA) and the ECL system (Amersham, NJ, USA) were used to visualize the blots.

Cell transfection and immunofluorescence

HeLa cells were maintained in DMEM (GIBCO BRL, MD, USA) with 10% fetal bovine serum (Gemini Bio-Products, CA, USA). Cells were grown on 12 mm coverslips placed in 24-well cell culture and transfected using Effectene (Qiagen, CA, USA) according to the manufacturer's instructions. After 16 hours, cells were fixed with 4% formaldehyde in PBS for 10 minutes at room temperature, permeabilized with 0.2% Triton X-100 in

PBS for 5 minutes at room temperature, and blocked in 1%BSA/PBS. Cells were incubated with primary antibodies in 1% BSA/PBS for one hour at room temperature followed by secondary antibodies, and stained with 4,6-diamidino-2-phenylindole (DAPI). Goat-anti-myc-FITC polyclonal antibody (Bethyl Laboratories, TX, USA) diluted to 5 ug/ml was used to detect the myc-MBP-peptides.

The monoclonal antibody 4C2 at 1:1000 dilution detected endogenous hnRNP A1 when incubated with goat-anti-mouse-Cy3 (Jackson ImmunoResearch Laboratories, PA, USA) antibody at 1:400 dilution. Monoclonal antibody 2A6 was used at 1:1000 dilution to detect endogenous hnRNP M. Mouse anti-HuR antibody was purchased from Zymed and was used at 1:100 dilution. Mouse anti-HDAC1 monoclonal antibody 2E10 (Upstate Biotechnology, MA, USA; diluted 1:500) was used. Cells were then examined in a Zeiss Axiovert 200M microscope with De-convolution and Apotome systems. Images were acquired with the AxioVision software (Carl Zeiss Image Solutions) and processed with Image J software (National Institutes of Health, Bethesda, MD). HuR and hnRNP M images were acquired using a Leica TCS SP5 confocal microscope and the Leica LAS AF software (Leica Microsystems Inc).

Results and Discussion

Antibodies for hnRNP A1 and M do not recognize the chimeric peptide M9M

The chimeric peptide M9M contains the 21 residues from hnRNP A1-NLS and 11 residues from hnRNP M-NLS and can possibly be recognized by antibodies against either hnRNP A1 or M, which interferes the detection of hnRNP A1 and M in

immunofluorescence. Hence anti-hnRNP A1 and M antibodies were first tested for their reactions to M9M in western blotting. Same amounts of purified recombinant MBP-hnRNP A1-NLS, MBP-hnRNP M-NLS, and MBP-M9M as well as the Hela cell lysate were loaded and probed with either 4C2 or 2A6 antibodies. The monoclonal antibody 4C2 has been previously shown to recognize human hnRNP A1, A2, B1 and B2 (Matunis, Matunis et al. 1992) I show by western blot that 4C2 recognizes both the recombinant MBP-hnRNP A1-NLS and the hnRNP A1 in cell lysate , but not the chimeric inhibitory peptide M9M (Figure 2-5A). The monoclonal antibody 2A6 against hnRNP M (Datar, Dreyfuss et al. 1993) only recognizes the endogenous hnRNP M in HeLa cell lysate, and it reacts with neither recombinant MBP-hnRNP M-NLS nor MBP-M9M (Figure 2-5B). Thus these two antibodies can be used to detect the endogenous hnRNP A1 and M in the presence of M9M.

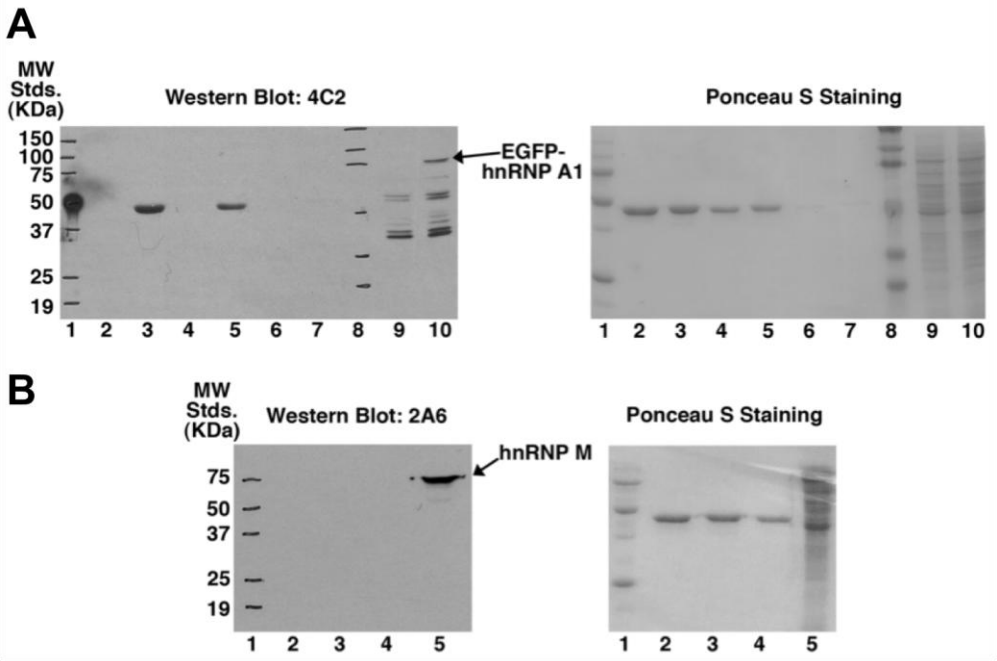

Figure 2-5 Western blots using antibodies against hnRNPs A1 and M. (A) Western Blot with antibody 4C2 (left), which recognizes human hnRNPs A1, A2 and B1, and visualization of proteins by Ponceau staining (right). Lanes 2, 4 and 6 contain 2 ug, 1 ug, and 0.1 ug of MBP-M9M; lanes 3, 5 and 7 contain 2 ug, 1 ug and 0.1 ug of MBP-hnRNP A1- NLS; Lane 9 contains control HeLa cell lysate and lane 10 has lysate from myc-EGFP-A1- transfected HeLa cells. Lanes 1 and 8 are molecular weight standards. (B) Western Blot with antibody 2A6 (left), which recognizes human hnRNP M, and visualization of proteins by Ponceau staining (right). Lane 1 contains molecular weight standards; Lane 2 contains 1 ug of MBP-M9M; Lane 3 contains 1 ug of MBP-hnRNP A1-NLS; Lane 4 contains 1 ug of MBP-hnRNP M-NLS; Lane 5 contains HeLa cell lysate.

M9M mislocalizes Kapβ2 cargos

The M9M peptide with super high affinity to Kapβ2 efficiently competes with nature PY-NLSs and even prevents the dissociation by RanGTP in in vitro binding assays. It may act as a specific inhibitor that blocks Kapβ2-mediated nuclear import and causes the mislocalization of Kapβ2 cargos in the cells. In order to test the effect of M9M in cells, Myc-tagged MBP-M9M was transfected into HeLa cells and the subcellular localization of endogenenous cargos for Kapβ2, hnRNP A1, HuR and hnRNP M were examined by immunofluorescence. As an mRNP binding protein, hnRNP A1 shuttles between the nucleus and cytoplasm and is predominantly nuclear (Michael, Choi et al. 1995; Siomi and Dreyfuss 1995). In the control cells transfected with only MBP, hnRNP A1 accumulates in the nucleus as expected. However, more than 50% of the cells transfected with MBP-M9M showed significant cytoplasmic staining of hnRNPA1 (Figure 2-6 and 2-7). HuR is also a nuclear protein containing a noncanonical hPY-NLS where the conserved PY motif is replaced with PG (Fan and Steitz 1998; Fan and Steitz 1998; Peng, Chen et al. 1998; Lee, Cansizoglu et al. 2006) (Figure 1-3C). Over 70% of the cells with MBP-M9M have altered HuR localization in the cytoplasm (Figure 2-6 and 2-7). Similarly, in about 50% of the cells with MBP-M9M, hnRNP M was mislocalized into the cytoplasm (Figure 2-6 and 2-7). Thus, expressing M9M in the cells resulted in mislocalization of multiple cargos of Kapβ2, which is possibly due to the inhibition of Kapβ2-mediated nuclear import by M9M.

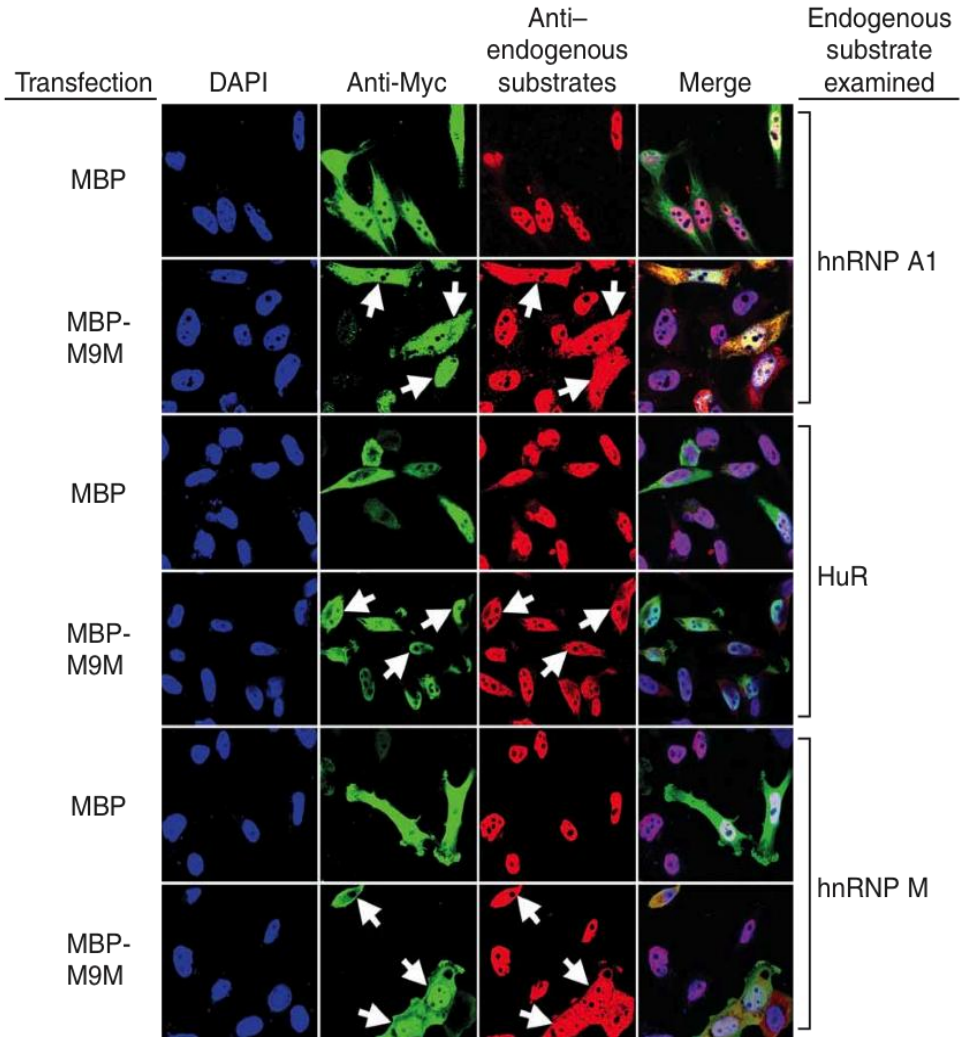

Figure 2-6 M9M mislocalizes endogenous Kapβ2 substrates. Immunofluorescence and deconvolution microscopy of HeLa cells transfected with plasmids encoding Myc-tagged MBP or MBP-M9M, using anti-Myc and antibodies to hnRNP A1, hnRNP M and HuR. The arrows indict the cells transfected with MBP-M9M.

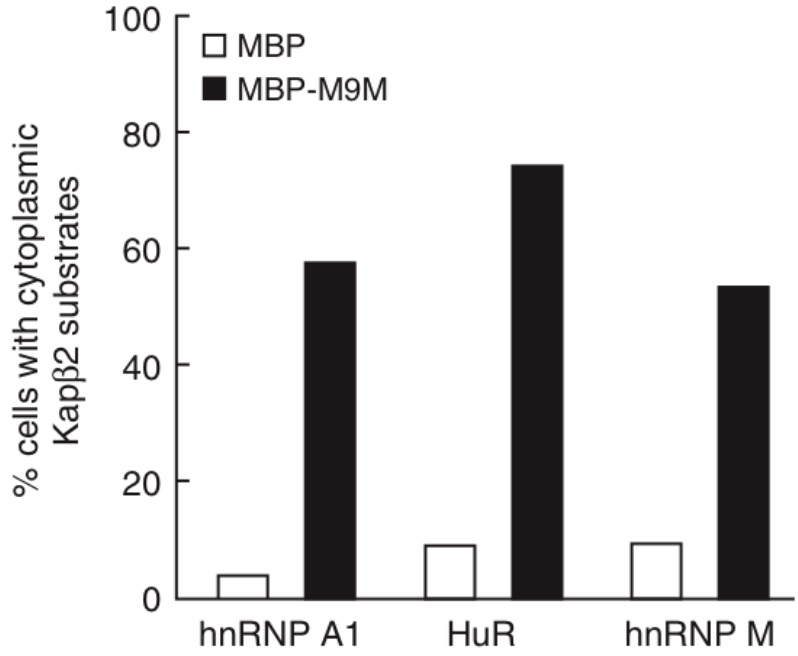

Figure 2-7 Quantification of transfected cells that with cytoplasmic Kapβ2 substrates in Figure 2-6.

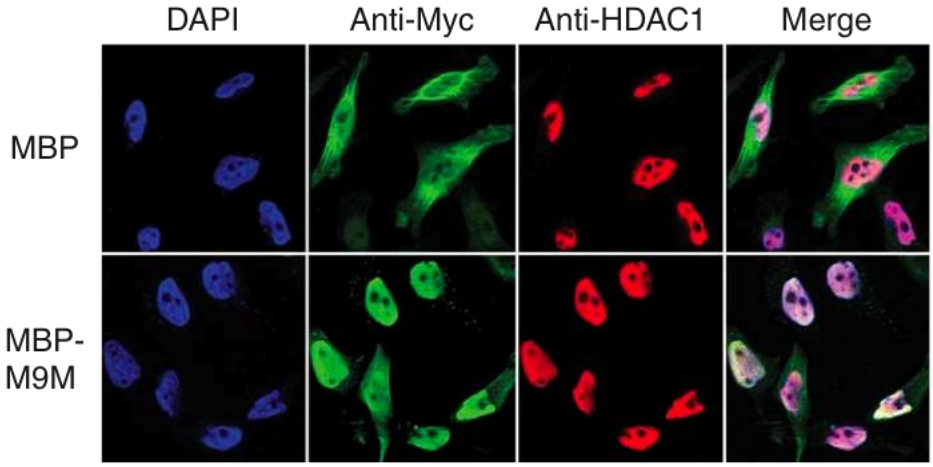

Figure 2-8 M9M does not mislocalize HDAC1. Immunofluorescence and deconvolution microscopy of HeLa cells transfected with plasmids encoding Myc-tagged MBP or MBP-M9M, using anti-Myc and antibodies to. HDAC1 (Impα–Impβ1 substrate).

M9M has no effect on Impα/β cargo HDAC1

In order to demonstrate that M9M is a Kapβ2-specific inhibitor that does not affect other import pathways, we would like to test the effect of M9M on other non-Kapβ2 cargos. HDAC1 was previously reported to be imported into the nucleus by Impα/Impβ (Smillie, Llinas et al. 2004). We have confirmed by in vitro binding assays that recombinant HDAC1 binds Impα but not Kapβ2 (data not shown). The endogenous HDAC1 accumulated in the nucleus no matter whether the cells were transfected with MBP only or MBP-M9M (Figure 2-8). Thus, M9M has no effect on Impα/β mediated nuclear import.

Conclusions

In summary, the interactions between PY-NLSs Kapβ2 are multivalent and structurally conserved in at the arginine and proline-tyrosine residues of their C-terminal $R/K/HX_{2-5}PY$ motifs and at their N-terminal basic or hydrophobic motifs. The discovery of asymmetric NLS binding hot spots in hnRNP M and hnRNP A1 led to the design of the M9M peptide, which binds Kapβ2 200-fold tighter than natural NLSs. This M9M peptide can specifically inhibits the interaction of Kapβ2 with its cargos both in vitro and in the cells, but does no affect other import pathways. It is the first pathway-specific inhibitor for nuclear import and will be a valuable tool used to either identify new cargos for Kapβ2, or study other important cell functions involving Kapβ2.

CHAPTER THREE

EVOLUTIONARY DEVELOPMENT OF REDUNDANT NUCLEAR LOCALIZATION SIGNALS IN THE MRNA EXPORT FACTOR NXF1∗

Abstract

In human cells, the mRNA export factor NXF1 resides in the nucleoplasm and at nuclear pore complexes. Karyopherinβ2 or Transportin is known to recognize a PY-NLS in the N-terminal tail of NXF1 and imports it into the nucleus. Here, biochemical and cellular studies to understand the energetic organization of the NXF1 PY-NLS have revealed unexpected redundancy in the nuclear import pathways used by NXF1. Human NXF1 can be imported into *via* Importinβ, Karyopherinβ2, Importin4, Importin11 and Importinα. Two NLS epitopes within the N-terminal tail, an N-terminal basic segment and a C-terminal R-X_{2-5}-P-Y motif, provide the majority of binding energy for all five Karyopherins. Mutation of both NLS epitopes abolished binding to the Karyopherins, mislocalized NXF1 to the cytoplasm and significantly compromised its mRNA export function. The understanding of how different Karyopherins recognize human NXF1, the examination of NXF1 sequences from divergent eukaryotes and the interactions of NXF1

∗ This work is submitted to Mol Biol Cell and under revision.

homologs with various Karyopherins have revealed the evolutionary development of redundant NLSs in NXF1 of higher eukaryotes. Redundancy of nuclear import pathways for NXF1 increased progressively from fungi to nematodes and insects to chordates, potentially paralleling the increasing complexity in mRNA export regulation and the evolution of new nuclear functions for NXF1.

Introduction

The transport of mRNA from the site of transcription in the nucleus to the site of translation in the cytoplasm is an essential process in eukaryotic gene expression. In human cells, the mRNA export factor NXF1 (also known as TAP) escorts mRNA transcripts out of the nucleus by simultaneously binding mRNA, mRNA adaptor proteins and phenyalanine-glycine (FG) repeats of the nuclear pore complex (NPC) (Stutz and Izaurralde 2003; Erkmann, Sanchez et al. 2005; Reed and Cheng 2005; Kohler and Hurt 2007; Hautbergue, Hung et al. 2008; Carmody and Wente 2009; Kelly and Corbett 2009). NXF1 is unique among nuclear transport factors as it is a multi-domain protein that bears no structural or mechanistic resemblance to the Karyopherin proteins that transport protein cargos, tRNAs and micro-RNAs through the NPC. mRNA export by NXF1 is a process that occurs independently of the GTPase Ran (Grüter, Tabernero et al. 1998).

Human NXF1 (*hs*NXF1) contains a 110-residue N-terminal tail that precedes four well-characterized globular domains (Figure 3-1A) (Liker, Fernandez et al. 2000; Fribourg, Braun et al. 2001; Grant, Hurt et al. 2002; Ho, Coburn et al. 2002; Fribourg and Conti 2003; Senay, Ferrari et al. 2003; Stutz and Izaurralde 2003). The RNA-binding (RBD)

and leucine-rich repeat (LRR) domains bind constitutive transport element (CTE) containing viral RNAs (Braun, Rohrbach et al. 1999). The two domains are also involved in binding cellular mRNAs, likely with the help of adaptor proteins (Bachi, Braun et al. 2000; Strasser, Bassler et al. 2000; Stutz, Bachi et al. 2000; Huang, Gattoni et al. 2003; Hautbergue, Hung et al. 2008). Beyond the two domains that bind RNA are the NTF2-like and UBA domains. The heterodimer of NTF2-like domain with NXT1 and the UBA domain bind FG repeats of nucleoporins to target NXF1 to the NPC for translocation (Santos-Rosa, Moreno et al. 1998; Katahira, Strasser et al. 1999; Fribourg, Braun et al. 2001; Grant, Hurt et al. 2002; Senay, Ferrari et al. 2003). The N-terminal tail is the least well-characterized region of *hs*NXF1. The tail is predicted to be structurally disordered and contains a 10-residue segment that is critical for targeting *hs*NXF1 to the nucleus (Bear, Tan et al. 1999; Braun, Rohrbach et al. 1999; Kang and Cullen 1999; Katahira, Strasser et al. 1999; Truant, Kang et al. 1999; Bachi, Braun et al. 2000). This segment was later identified as part of a proline-tyrosine nuclear localization signal (PY-NLS) that binds the Importin Karyopherin β2 (Kapβ2 or Transportin) (Lee, Cansizoglu et al. 2006; Imasaki, Shimizu et al. 2007). The N-terminal tail also contributes to interactions with adaptor proteins E1B-AP5, ALY/REF, SR proteins and the NS1, the influenza virus protein, which inhibits mRNA export (Bachi, Braun et al. 2000; Stutz, Bachi et al. 2000; Huang, Gattoni et al. 2003; Satterly, Yarbrough et al. 2011).

PY-NLSs are generally 15-30 amino acids long, are basic in character, found in structurally disordered regions of proteins and usually contains an N-terminal basic or hydrophobic motif and a C-terminal R-X_{2-5}-P-Y motif (Lee, Cansizoglu et al. 2006;

Cansizoglu, Lee et al. 2007; Suel, Gu et al. 2008; Suel and Chook 2009). Kapβ2 binds PY-NLSs with high affinity (K_Ds ~10-50 nM) to target import cargos for translocation through the NPC. RanGTP releases PY-NLSs from Kapβ2 in the nucleus. The crystal structure of Kapβ2 bound to a 30-residue fragment of the *hs*NXF1 PY-NLS showed interactions with only 10 residues immediately surrounding the C-terminal R-x_{2-5}-P-Y motif but not with an N-terminal basic/hydrophobic motif (Imasaki, Shimizu et al. 2007).

Here, we report that biochemical and cellular studies to understand the energetic organization of the *hs*NXF1 PY-NLS have unexpectedly revealed that the mRNA export factor is imported into the nucleus *via* five different Karyopherin pathways. *hs*NXF1 can be imported into the nucleus through the interactions of its N-terminal tail with Impβ, Kapβ2, Imp4, Imp11 and Impα. Within the N-terminal tail of *hs*NXF1, an N-terminal basic NLS epitope spanning residues 21-30 is important for binding Impα and for direct interactions with Impβ, Imp4 and Imp11, whereas the R-X_{2-5}-P-Y motif at residues 71-75 is important for Kapβ2 binding. Mutation of both NLS epitopes abolished binding to all five Karyopherins, mislocalized *hs*NXF1 to the cytoplasm and significantly compromised its functions in gene expression. The understanding of how different Karyopherins recognize *hs*NXF1, how different Karyopherins bind NXF1 proteins from various organisms and the examination of diverse NXF1 sequences have revealed the evolutionary development of redundant NLSs in the mRNA export factors. The redundancy of nuclear import pathways for NXF1 increases with the complexity of the eukaryote, suggesting parallel evolution of new nuclear functions for NXF1.

Materials and Methods

Plasmids

GST fusion constructs were generated by inserting PCR fragments corresponding to the regions of the genes of interest into pGEXTEV vectors (modified pGEX4T3 (GE Healthcare, UK) with TEV site) (Chook and Blobel 1999). The GST fusion constructs include full length human Impβ, Kapβ2, Imp4, Imp5, Imp9, Imp11, Trn-SR, Imp13, RanBP1; mouse Impα2-ARM (residues 75-496); full length human NXF1 or *hs*NXF1; *hs*NXF1 fragments *hs*NXF1-N (residues 1-109), RBD (residues 115-200), LRR (residues 201-365), NTF2-like (residues 368-554), UBA (residues 563-619) and *hs*NXF1(1-40), *hs*NXF1(40-80), *hs*NXF1(30-80), *hs*NXF1(1-80), *hs*NXF1(70-109), *hs*NXF1(80-109); N-terminal tails of *X. tropicalis* NXF1 (residues 1-115), *D. rerio* NXF1 (residues 1-136), *D. melanogaster* NXF1 (residues 1-109), C. elegans NXF1 (residues 1-87) and *S. pombe* Mex67p (residues 1-31). Synthetic oligonucleotides corresponding to residues 1-87 from C.elegans and residues 1-31 from S.pombe were annealed and inserted into the pGEXTEV vector. MBP fusion constructs of *hs*NXF1 (full length and fragments) were subcloned from pGEXTEV-*hs*NXF1 constructs into pMALTEV (modified pMAL (New England BioLabs) with TEV site (Chook, Jung et al. 2002) vectors. Mouse Impα2 without the IBB domain (Impα2-ΔIBB, residues 75-529) was cloned into pET21a vector (EMD Biosciences). p10, Ran (Chook, Jung et al. 2002). Mammalian expressing vectors pEGFP-c1-NXF1 and pCMV-Luc were kindly provided by E. Izaurralde (Max Planck Institute, Tubingen, Germany and D. Levy (New York University, USA), respectively. The Kapβ2 pathway inhibitor vector pCS2-MT-MBP-M9M and the control vector pCS2-

MT-MBP were described in (Cansizoglu, Lee et al. 2007). NXF1 Mutations were made by site-directed mutagenesis using Quikchange site-directed mutagenesis kit (Stratagene, La Jolla, CA) and all constructs were sequenced before use.

Recombinant Protein Preparation

All recombinant proteins were expressed in BL21 (DE3) *E. coli* cells by induction with 0.5 mM IPTG overnight at 25 ℃. For pull-down binding assays, bacteria expressing GST fusion proteins were lyzed by sonication and centrifuged. The supernatants were incubated with glutathione (GSH) sepharose (GE Healthcare, NJ, USA) followed by extensive washes using transfer buffer TB (20 mM HEPES pH7.3, 110 mM KOAc, 2 mM DTT, 2 mM MgOAc, 1 mM EGTA) with 20% glycerol. Immobilized GST fusion proteins were stored in TB buffer with 40% glycerol at -20 ℃ before use. Bacteria expressiong GST fusions of Impβ, Kapβ2, Imp4, Imp5, Imp9, Imp11, Trn-SR and Imp13 were lyzed using cell homogenizer EmulsiFlex-C5 (Avestin Inc, Ontario, Canada) and after centrifugation, cell lysates were purified by GSH affinity chromatgraphy. GST-Imp4 and GST-Imp11 were used for nuclear import assays. For all other experiments, the GST-Kapβs cleaved with TEV protease and further purified by ion-exchange (HiTrap Q; GE Healthcare, NJ, USA) and gel filtration (Superdex 200; GE Healthcare, NJ, USA) chromatography. Mouse Impα2-ARM and RanBP1 were purified similar ways(Chook, Jung et al. 2002).

To purify MBP fusion proteins, bacterial lysates were incubated with amylose beads (New England Biolabs, MA, USA) and the fusion proteins eluted with 20mM Hepes

pH7.5, 50mM NaCl, 2mM EDTA, 2mM DTT, 10%glycerol and 10mM Maltose. For the binding assays with *hs*NXF1-N alanine scanning mutants, MBP-*hs*NXF1-N proteins were concentrated and dialyzed against TB buffer with 20% glycerol before use. For all other experiments, MBP fusion proteins were further purified by ion-exchange chromatography.

Human Ran and mouse Impα2-ΔIBB were expressed as His-tagged proteins and purified by affinity and ion-exchange chromatography (Chook, Jung et al. 2002; Dong, Biswas et al. 2009; Dong, Biswas et al. 2009). For RanGTP-mediated dissociation assay, recombinant Ran was loaded with GTP analog GMPPNP before use, as previously described (Suel, Gu et al. 2008; Suel and Chook 2009) and the His_6-NTF2 used in this assay was purified by affinity chromatography using Talon beads followed by gel filtration (Chook, Jung et al. 2002).

In vitro pull-down binding assays

In vitro pull down binding assays were performed by incubating immobilized GST-fusion proteins with potential binding partners in TB buffer with 20% glycerol at 4℃ for 30min, followed by extensive washing with the same buffer. Bound proteins were visualized using SDS-PAGE and Coomassie Blue staining. ~ 5 μg of immobilized GST-NXF1 proteins or fragments were incubated with ~ 20 μg of purified Karyopherins. About half of the bound proteins were loaded for gel analysis. ~10-20 μg of immobilized GST-Karyopherins were incubated with ~20 μg of MBP-NXF1 fragments and ~ 25% of bound proteins were loaded for gel analysis.

RanGTP-mediated dissociation assay

~5 μg of immobilized GST-NXF1 were first incubated with ~20 μg of Kapβs for 30 min at 4 ℃ followed by extensive washing. A second incubation was done with either 112 μg of RanGMPPNP or buffer. After extensive washing, half of the bound proteins were separated by SDS-PAGE and visualized with Coomassie blue staining.

Cell culture, transfection and fluorescence microscopy

HeLa Tet-on cells and 293T cells were maintained in Dulbecco's modified Eagle's medium (Invitrogen, CA, USA) supplemented with 10% fetal bovine serum (Gemini Bio-Products, CA, USA) at 37 ℃ in a humidified atmosphere containing 5% CO_2 in air. Transfections were performed with Lipofectamine 2000 (Invitrogen, CA, USA) according to the manufacturer's instructions. After 16 hours of transfection, HeLa Tet-on cells were subjected to standard immunostaining procedures as described in (Cansizoglu *et al.*, 2007) with goat-anti-myc-FITC (Bethyl Laboratories, TX, USA), mouse monoclonal antibody 4C2 (a gift from Dr. M. Matunis, Johns Hopkins University), goat-anti-mouse-Cy3 (Jackson ImmunoResearch Laboratories, PA, USA), mouse anti-NXF1 monoclonal antibody 53H8 (Sigma-Aldrich, MO, USA). Cells were stained with 4,6-diamidino-2-phenylindole (DAPI) and then mounted onto slides for imaging. Cells transfected with EGFP fusion proteins were directly stained with DAPI for imaging after fixation and permeabilization. Cells were examined in an Applied Precision Deltavision RT Deconvolution microscope using 60X oil objective lens. Images were acquired by

SoftWoRx software (Applied Prevision Inc, WA, USA) and processed with Image J software (National Institutes of Health, MD, USA).

Nuclear import assays

HeLa Tet-on cells were grown to 50% confluency on coverslips, washed in cold TB buffer, and permeabilized with 35 μg/mL digitonin on ice for 5 min. Permeabilized cells were incubated with import reaction mixture (5 μM of MBP-*hs*NXF1, Ran mix [3 μM Ran, 0.3 μM RanBP1, 0.3 μM p10, 1 mM GTP, 8 mM magnesium acetate, with or without 5 μM of the individual recombinant Karyopherins) for 30 min at room temperature followed by washing and fixing. The MBP proteins were detected by immunofluorescence using mouse anti-MBP monoclonal antibody.

Isothermal Titration Calorimetry (ITC)

Binding affinities of MBP-*hs*NXF1-N proteins to Impβ and Kapβ2 were quantitated using ITC as described in (Cansizoglu, Lee et al. 2007; Suel, Gu et al. 2008). ITC experiments were performed with a MicroCal Omega VP-ITC calorimeter (MicroCal Inc., MA, USA). Proteins were dialyzed against buffer containing 20 mM Tris pH 7.5, 100 mM NaCl, and 2 mM β-mercaptoethanol. 100–300 μM MBP-*hs*NXF1-N proteins were titrated into a sample cell containing 10–20 μM recombinant Impβ or Kapβ2. Most ITC experiments were performed at 20 ℃ with 35 rounds of 8 μl injections. Data were plotted and analyzed using MicroCal Origin software (version 7.0).

Western blotting

293T cells were transfected with either wild type or mutant pEGFP-C1-*hs*NXF1 using Lipofectamine™ 2000 (Invitrogen, CA, USA) according to the manufacturer's instructions. After 12 hours, cells were lyszed with CelLytic™ M (Sigma-Aldrich, MO, USA). The protein concentration were measured by Bradford methods and ~50 μg of proteins were loaded for each lane on SDS-PAGE gel. The proteins were transferred onto PVDF membrane and probed with mouse monoclonal anti-NXF1 antibody (Sigma-Aldrich, MO, USA) at 1:2000 dilution. Signals were detected using ECL detection reagent (GE Healthcare, NJ, USA) after incubation with HRP-labeled anti-mouse antibody (GE Healthcare, NJ, USA) at 1:5000 dilution.

Luciferase reporter gene assay

The experiments were performed according to (Chakraborty, Satterly et al. 2006). Briefly, 293T cells grown on 30-mm six-well plates were co-transfected with pCMV-Luc (2 μg) and either wild type or mutant pEGFP-C1-*hs*NXF1 (2 μg) using Lipofectamine™ 2000 (Invitrogen, CA, USA) according to the manufacturer's instructions. After 12 hours of transfection, cells were lysed and luciferase activities of each sample were measured using luciferase assay reagent (Promega, WI, USA) in triplicate. Cell-titer Glo assays were performed similarly with Cell-titer Glo reagent (Promega, WI, USA) according to manufacturer's instructions. Averages of the luciferase signals (S_{Luc}) were divided by the average of Cell-titer Glo signals (S_{Cell}) to diminish the difference of cell numbers between

samples. And the ratios (S_{Luc}/ S_{Cell}) were normalized to that of EGFP control (100%) and represented as percentages in the bar graph.

Sequence alignment

Multiple sequence alignments were performed using ClustalW (Chenna, Sugawara et al. 2003) with manual adjustment. Uniprot accession numbers for the NXF1 or Mex67p sequences are Q9Y8G3 (S. pombe), B6JXN8 (S. japonicas), Q9XVS7 (C. elegans), A8WY32 (C. briggsae), Q9U1H9 (D. melanogaster), B4JKG4 (D. grimshawi), Q7QK79 (A. gambiae), Q17MK6 (A. aegypti), Q9UBU9 (H. sapiens), Q28C94 (X. tropicalis), Q5CZT0 (D. rerio). Genbank accession numbers: XP_002589241 (B. floridae) and XP_002129680 (C. intestinalis).

Results

Multiple Karyopherins mediate nuclear import of human NXF1.

In human cells, *hs*NXF1 is localized mostly to nucleoplasm and the NPC (Bear, Tan et al. 1999; Katahira, Strasser et al. 1999; Bachi, Braun et al. 2000). Despite the ability of *hs*NXF1 to interact with the NPC through its C-terminal NTF2-like and UBA domains (Santos-Rosa, Moreno et al. 1998; Katahira, Strasser et al. 1999; Fribourg, Braun et al. 2001; Grant, Hurt et al. 2002; Senay, Ferrari et al. 2003), a minimal non-classical NLS spanning residues 61-102 in the N-terminal tail was found to be critical for its nuclear localization through nuclear import by Kapβ2 (Bear, Tan et al. 1999; Kang and Cullen 1999; Katahira, Strasser et al. 1999; Truant and Cullen 1999; Bachi, Braun et al. 2000).

Consistent with these previous findings, we showed that full-length *hs*NXF1 was localized in the nucleus but a mutant lacking the N-terminal tail was cytoplasmic (Flag-*hs*NXF1(115-619); Figure 3-1). Since *hs*NXF1 is a well-established Kapβ2 cargo (Truant *et al.*, 1999; Bachi *et al.*, 2000; Lee *et al.*, 2006; Imasaki *et al.*, 2007), we expressed the Kapβ2-specific peptide inhibitor M9M (Cansizoglu, Lee et al. 2007)in HeLa cells to determine if Kapβ2 is the main nuclear import factor for *hs*NXF1. Surprisingly, myc-MBP-M9M failed to mislocalize *hs*NXF1 to the cytoplasm even though the inhibitor mislocalized other Kapβ2 cargos such as hnRNP A1 (Figure 3-2), hnRNP M, HIV-1 Rev and FUS to the cytoplasm (Cansizoglu, Lee et al. 2007; Hutten, Walde et al. 2009; Dormann, Rodde et al. 2010). These results suggested that Kapβ2 is not the sole nuclear importer of *hs*NXF1.

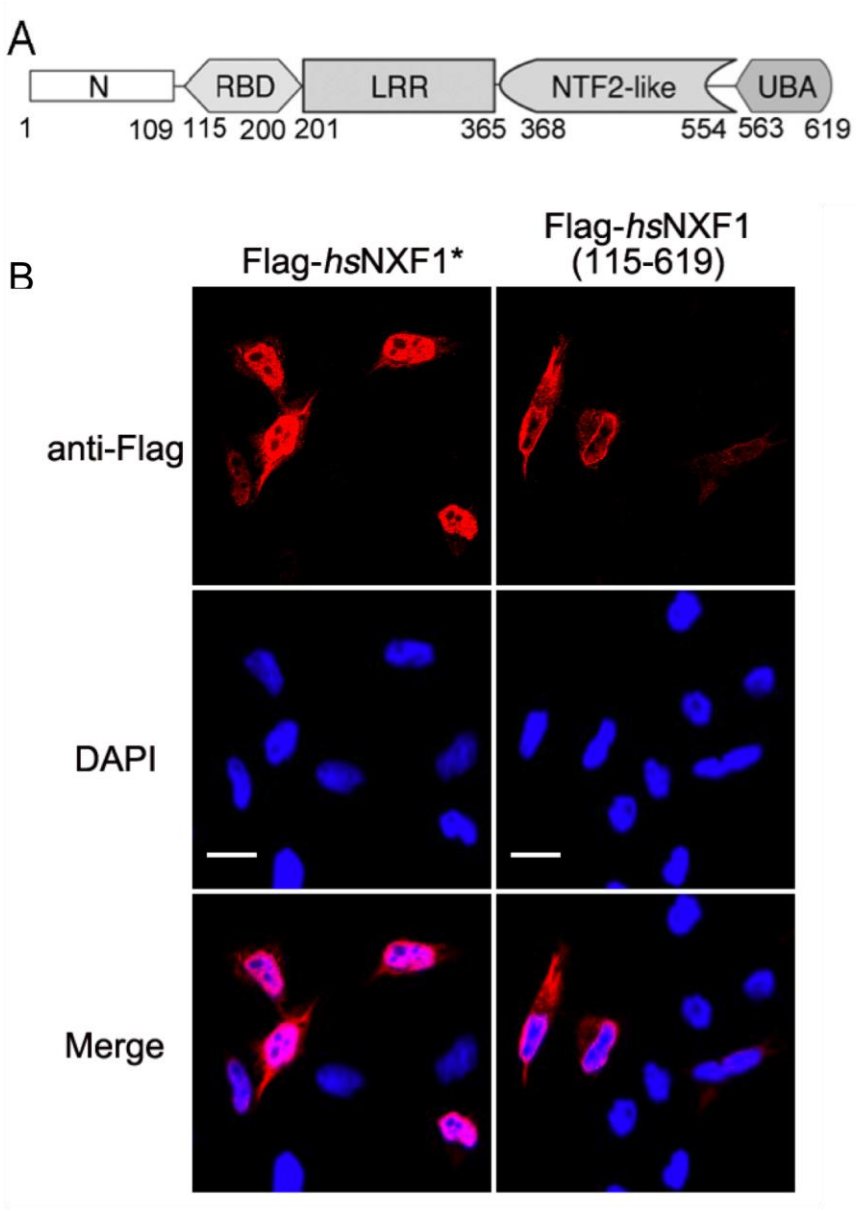

Figure 3-1 The N-terminal tail of hsNXF1 is necessary for its nuclear localization. (A) The domain organization of hsNXF1. (B) *hs*NXF1 and deletion mutant *hs*NXF1(115-619) were cloned into pFLAG-CMV2 vectors and transfected into HeLa cells. The overexpressed proteins were detected by immunofluorescence using anti-Flag antibodies. Scale bar, 10 μm

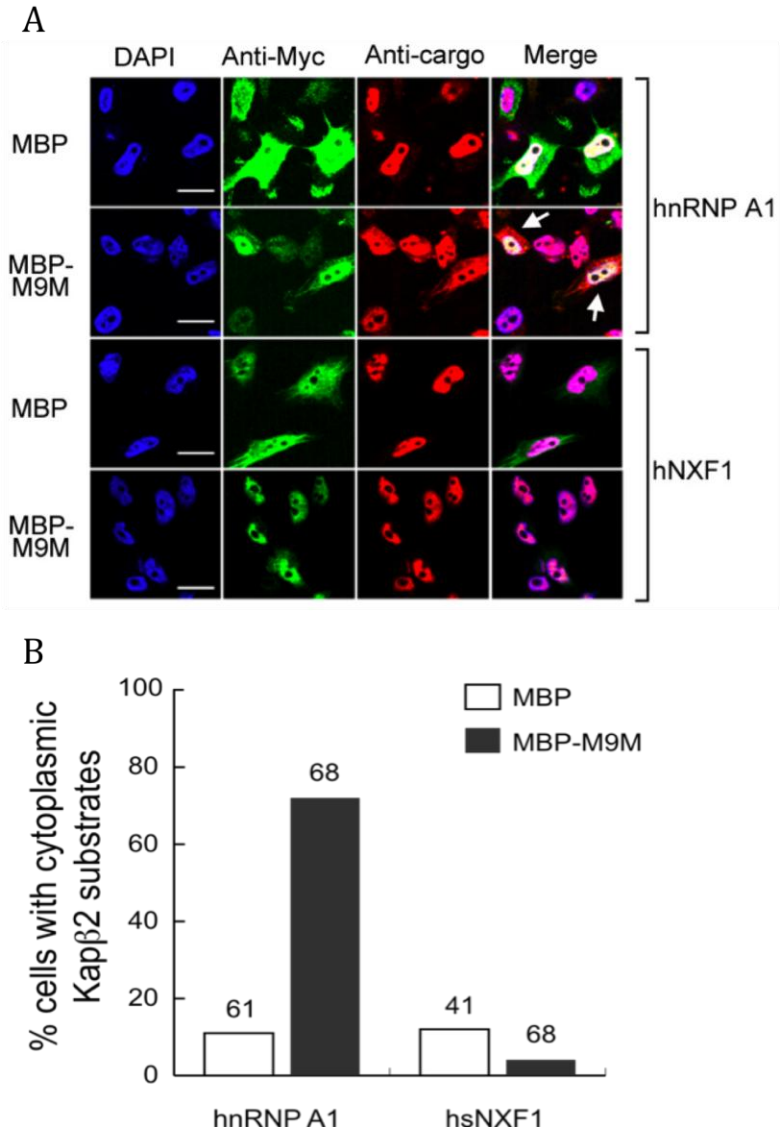

Figure 3-2 Endogenous hsNXF1 is not mislocalized by Kapβ2-specific inhibitor M9M. (A) Kapβ2 inhibitor M9M did not alter the subcellular localization of hsNXF1. HeLa cells were transfected with myc-tagged MBP or MBP-M9M and endogeneous Kapβ2 cargos hnRNP A1 and hsNXF1 were detected by immunofluorescence. Scale bars, 10 μm. (B) Histogram of shows percentages of transfected cells with cytoplasmci Kapβ2 substrates. The numbers of the cells counted are on top of each bar.

To identify additional nuclear import factors for *hs*NXF1, we tested its binding to most of the known human import-Karyopherins. Immobilized GST-*hs*NXF1 bound recombinant Impβ, Kapβ2, Imp4, Imp11 and Impα with significant affinity as shown by strong Coomassie-stained bands of the five Karyopherins (Figure 3-3A). *hs*NXF1 did not bind recombinant Imp5, Imp9, Trn-SR or Imp13 (Figure 3-3A). Interactions with Impβ, Kapβ2, Imp4, Imp11 were Ran-sensitive as subsequent incubations with RanGTP released *hs*NXF1 from the Karyopherins (Figure 3-3B). Impβ, Kapβ2, Imp4, Imp11 also mediated nuclear import of MBP-*hs*NXF1 in digitonin-permeabilized HeLa cells (Figure 3-4). Impα was not tested in the nuclear import assays since its effect cannot be distinguished from that of direct *hs*NXF1-Impβ interactions. Results of the Karyopherin-binding and nuclear import assays suggested that in addition to the well-established Kapβ2 pathway, *hs*NXF1 can be imported into the nucleus through direct interactions with Impβ, Imp4, Imp11 and *via* the classical Impα/β pathway.

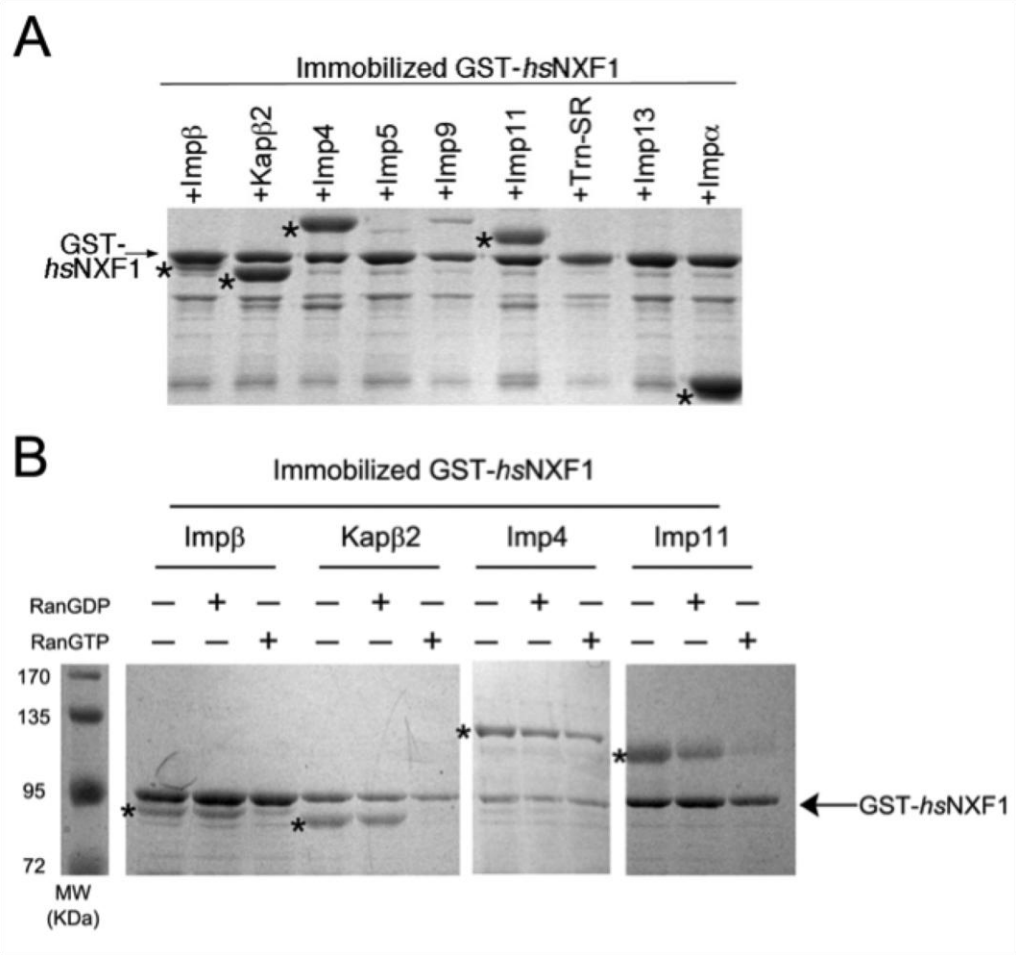

Figure 3-3 hsNXF1 interacts with multiple Kapβs. (A) hsNXF1 interacts with Karyopherins Impβ, Kapβ2, Imp4, Imp11 and impα in pull-down binding assays. Immobilized GST-hsNXF1 was incubated with purified recombinant Karyopherins. Bound proteins were visualized by SDS-PAGE and Coomassie staining. (B) Kapβ-hsNXF1 interactions are RanGTP sensitive. Immobilized GST-hsNXF1 were first incubated with Karyopherins, washed extensively and then incubated with buffer, RanGDP or RanGTP. Bound proteins in (A) and (B) were visualized using Coomassie staining.

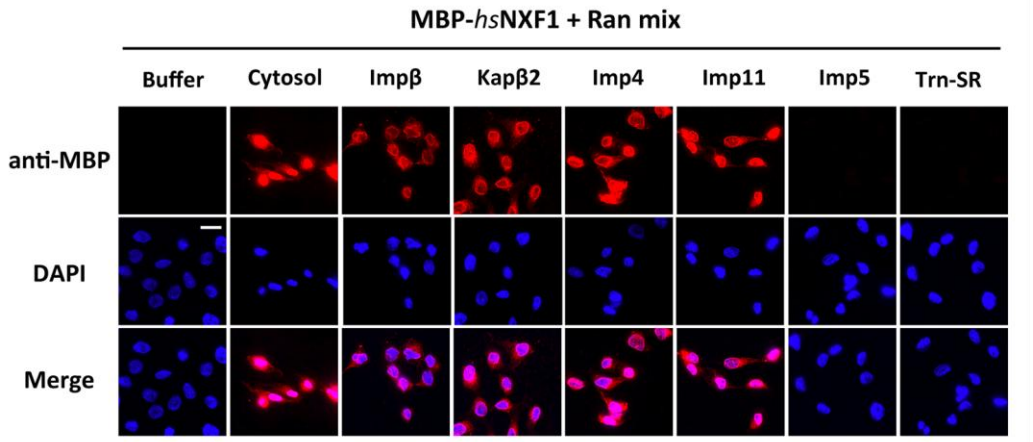

Figure 3-4 Impβ, Kapβ2, Imp4, Imp11 are able to import hsNXF1 into HeLa cell nucleus. Nuclear import assays were performed in digitonin-permeabilized HeLa cells with MBP-hsNXF1 in the presence of purified Kapβs or buffer. Samples were fixed and stained with anti-MBP antibody and Alexa546-anti-mouse secondary antibody, then subjected to immunofluorescence analysis. Scale bar, 10 μm.

NLSs for Impβ, Kapβ2, Imp4, Imp11 and Impα reside within the hs*NXF1 N-terminal tail.*

We divided the multi-domain *hs*NXF1 into its modular domains based on available structural information (Liker, Fernandez et al. 2000; Fribourg, Braun et al. 2001; Grant, Hurt et al. 2002; Ho, Coburn et al. 2002; Fribourg and Conti 2003; Senay, Ferrari et al. 2003). Immobilized GST fusions of the N-terminal tail (*hs*NXF1-N; residues 1-109), the RBD (residues 115-200), LRR (residues 201-365), NTF2-like (residues 368-554) and UBA (residues 563-619) domains were tested for binding to Impβ, Kapβ2, Imp4, Imp11 and Impα (Figure 3-5 and 3-6). All five karyopherins bound strongly to *hs*NXF1-N but not to the other domains. Impβ, Kapβ2, Imp4, Imp11 mediated nuclear import of *hs*NXF1-N into the nucleus of digitonin permeabilized HeLa cells (Figure 3-7A).

*hs*NXF1-N also targeted pyruvate kinase to the HeLa cell nuclei (Figure 3-7B) whereas *hs*NXF1 lacking its N-terminal tail was cytoplasmic (Figure 3-1). These results suggested that all the NLSs in *hs*NXF1 that are recognized by Impβ, Kapβ2, Imp4, Imp11 and Impα are located within its N-terminal tail.

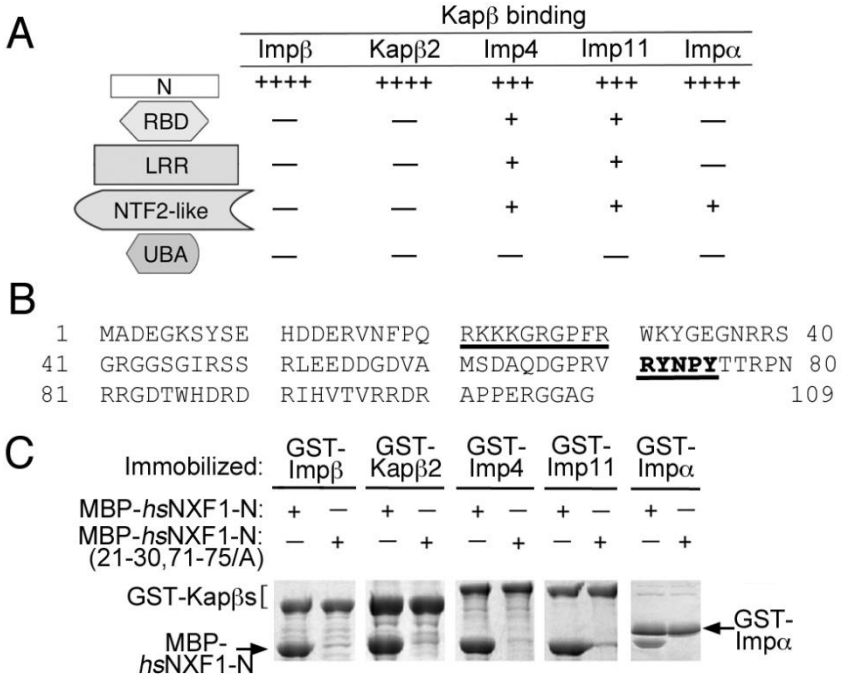

Figure 3-5 The NLSs of hsNXF1 for Impβ, Kapβ2, Imp4, Imp11 and impα are all located in the N-terminal tail (hsNXF1-N). (A) Summary of the pull-down binding assays of hsNXF1 domains with Impβ, Kapβ2, Imp4, Imp11 and impα (data shown in Figure S2). The number of "+" indicates the relative binding strength, and "-" indicates no significant binding. (B) The sequence of hsNXF1-N. The two NLS epitopes identified by alanine scanning mutagenesis and ITC (Figure 3-8 and Table 3-1) are underlined. (C) Alanine mutations at both NLS epitopes of hsNXF1 eliminated binding to Impβ, Kapβ2, Imp4, Imp11 and Impα. Immobilized GST-Karyopherins were incubated with MBP-hsNXF1-N or mutant MBP-hsNXF1-N(21-30, 71-75/A). Bound proteins were visualized by SDS-PAGE and Coomassie staining. GST-Impα* refers to Impα without its N-terminal IBB domain.

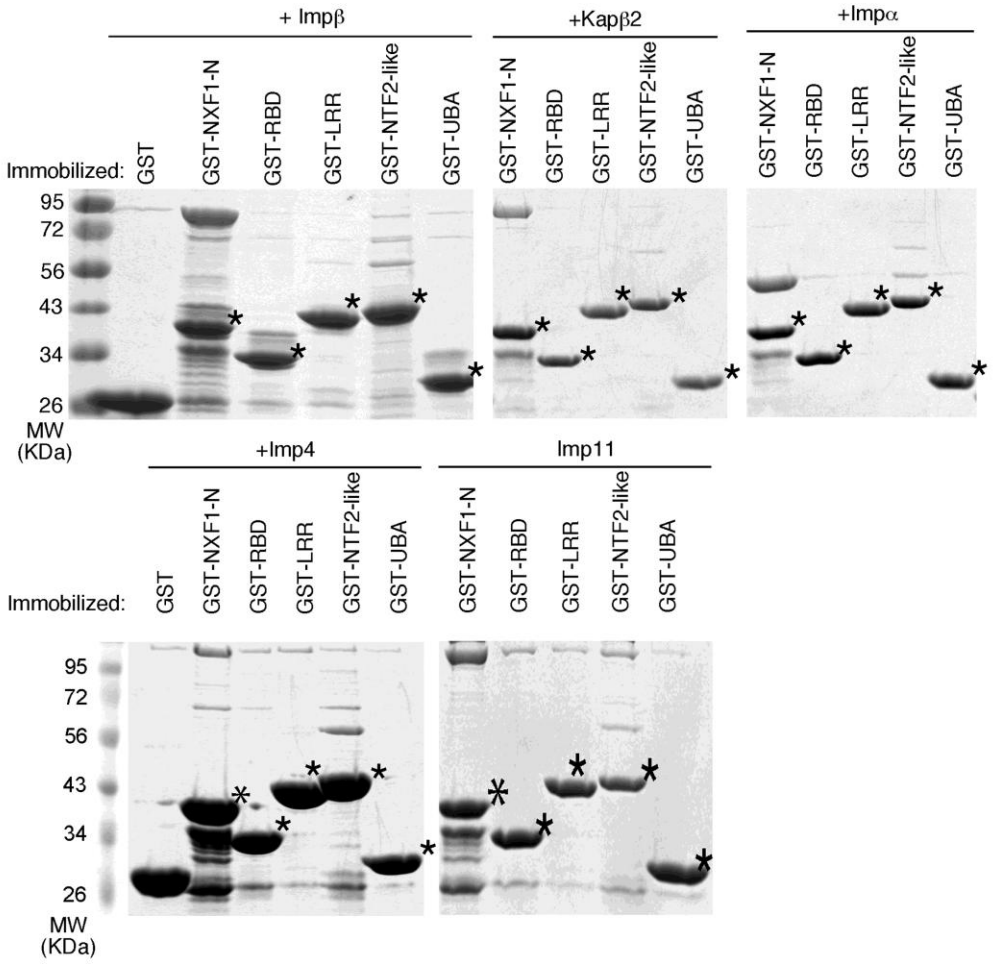

Figure 3-6 *hs*NXF1 binds different Karyopherins through its N-terminal tail or *hs*NXF1-N. Individual domains of *hs*NXF1 were expressed as GST fusion proteins, immobilized onto GSH sepharose and then incubated with purified recombinant Karyopherins. Bound proteins were visualized using Coomassie staining.

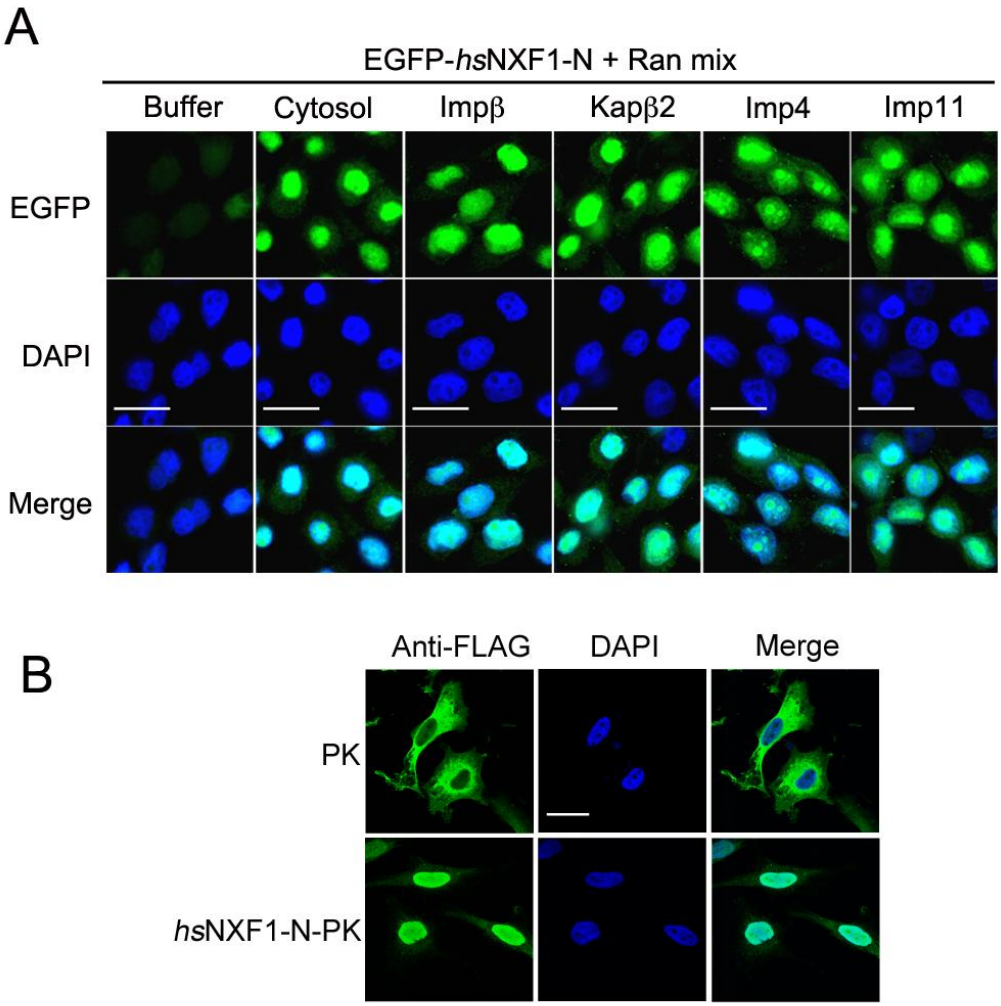

Figure 3-7 *hs*NXF1-N is sufficient for nuclear import. *(A)* Nuclear import assays were performed in digitonin-permeabilized HeLa cells with MBP-EGFP-*hs*NXF1-N in the presence of purified Kapβs or buffer. Samples were fixed and stained with DAPI, then subjected to immunofluorescence analysis *(B) hs*NXF1-N was fused to the N-terminus of pyruvate kinase (PK) gene and cloned into pFLAG-CMV vector and transfected into HeLa cells. *hs*NXF1-N-PK was detected by immunofluorescence using anti-Flag antibodies. Scale bar, 10 μm.

Two NLS epitopes contribute differently to interactions with Impβ, Kapβ2 and Impα.

*hs*NXF1 contains a PY-NLS that interacts with Kapβ2 (Lee, Cansizoglu et al. 2006; Imasaki, Shimizu et al. 2007). Interactions between *hs*NXF1 residues 68-79, which contains a R-X_{2-5}-P-Y motif, was observed in the crystal structure of Kapβ2 bound to residues 53-82 of *hs*NXF1 (Imasaki, Shimizu et al. 2007). The absence of electron density for residues 53-67 of *hs*NXF1 in their structure suggested that a previously predicted hydrophobic motif at ^{59}VAMS^{62} contributed little to *hs*NXF1-Kapβ2 interactions and may not be the N-terminal hydrophobic motif of the PY-NLS. We used *in vitro* pull-down binding assays, isothermal calorimetry (ITC), deletion and scanning alanine mutagenesis to study the energetic organization of the *hs*NXF1 PY-NLS. *hs*NXF1-N bound Kapβ2 with a K_D of 40.5 nM (Table 3-1and Figure 3-8; the *hs*NXF1-N sequence is shown in Figure 3-5B). N- and C-terminal truncations mapped residues 1-92 as the smallest *hs*NXF1 fragment that maintains the high affinity Kapβ2-binding of *hs*NXF1-N (K_D of 54 nM; Table 3-2). We then used scanning alanine mutagenesis and qualitative pull-down binding assays to identify binding determinants or NLS epitopes in the *hs*NXF1 PY-NLS (Figure 3-9). The results suggested binding hotspots at residues 71-75 and at residues 21-30 (Figure 3-9). We then used ITC to measure the energetic contributions of these potential NLS epitopes. Mutation of residues 71-75 to alanines reduced *hs*NXF1-N-Kapβ2 affinity by ~ 5-fold while mutation of the basic patch spanning *hs*NXF1 residues 21-30 resulted in an ~ 3-fold affinity reduction (Table 3-1and Figure 3-8). These results confirmed that the C-terminal R-X_{2-5}-P-Y motif at ^{71}RYNPY^{75}

as a hotspot for binding Kapβ2 and that hsNXF1 contains a PY-NLS of the basic subclass with its N-terminal basic motif at residues 21-30.

Table 3-1 Binding affinity of hsNXF1-N proteins for Kapβ2 and Impβ

Kapβ	*hs*NXF1-N	K_D[a] (nM)	ΔH (kcal/mol)	TΔS[b] (kcal/mol/K)	$K_{Dmutant}$/ $K_{Dwild\ type}$
Kapβ2	Wild type	40.5±12.6	-17.21±1.25	-7.31±1.42	-
	21-30/A	127.9±19.5	-13.93±2.65	-4.68±2.56	3.2
	71-75/A	215.3.8±61.4	-12.91±1.12	-3.98±1.27	5.4
	21-30, 71-75/A	n.d.	n.d.	n.d.	>>6[c]
Impβ	Wild type	K_{D1} 6±5 K_{D2} 1519±225	-7.51±0.07 -4.60±0.55	3.40±0.58 3.20±0.49	-
	21-30/A	2234±52	-2.48±0.52	5.09±0.53	372[d]
	71-75/A	K_{D1} 6±2 K_{D2} 2461±723	-7.73±0.28 -3.28±0.58	3.26±0.48 4.22±0.60	1 (K_{D1}); 2 (K_{D2})
	21-30, 71-75/A	n.d.	n.d.	n.d.	>400[d,e]

[a] Stoichiometry = 0.9-1.1.
[b] TΔS= ΔH – ΔG.
[c] The lowest measurable K_D of 215 nM in the Kapβ2-*hs*NXF1-N series was used to estimate $K_{Dmutant}$/ $K_{Dwild\ type}$.
[d] ratio taken using $K_{D1,wildtype}$
[e] The lowest measurable K_D of 2.23 μM in the Impβ-*hs*NXF1-N series was used to estimate $K_{Dmutant}$/ $K_{Dwild\ type}$.
n.d., not detectable; All experiments were performed 3-5 times (± standard deviation)

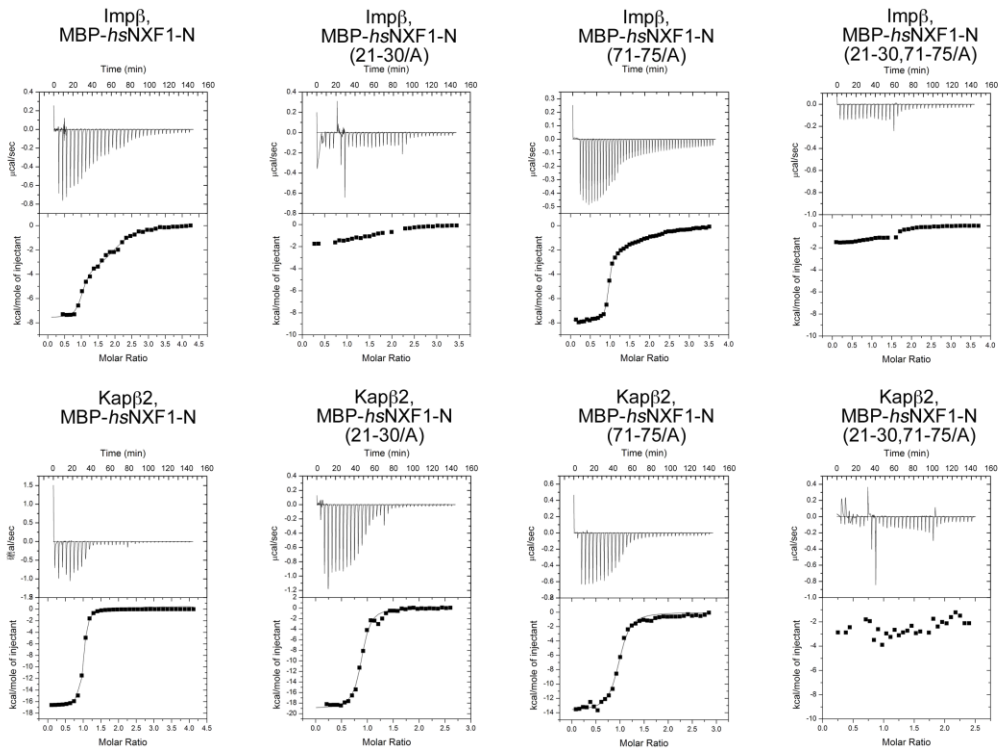

Figure 3-8 Selected ITC measurements of MBP-hsNXF1-N proteins biniding to Impβ and Kapβ2. After dialyzed against the same buffer, about 100–300 μM MBP-hsNXF1-N proteins were titrated into a sample cell containing 10–20 μM recombinant Impβ or Kapβ2 . The experiments were performed at 20°C with either 35 rounds of 8 μl injections or 56 rounds of 6 μl injections. Data were plotted and analyzed using MicroCal Origin software (version 7.0).

Table 3-2 Binding affinity of hsNXF1-N fragments for Kapβ2

Karyopherin	*hs*NXF1-N fragments	K_D[a] (nM)	ΔH (kcal/mol)	TΔS[b] (kcal/mol/K)
Kapβ2	1-109	40±13	-17.21±1.25	-7.31±1.42
	30-109	91±13	-18.51±1.13	-9.16±1.12
	1-92	54±4	-19.57±0.05	-9.82±0.03
	1-80	109±33	-17.74±0.89	-8.40±0.88
	30-80	204±29	-19.16±1.31	-10.17±1.40

[a] Stoichiometry = 0.9-1.1.
[b] $T\Delta S = \Delta H - \Delta G$.
All experiments were performed 3-5 times (± standard deviation)

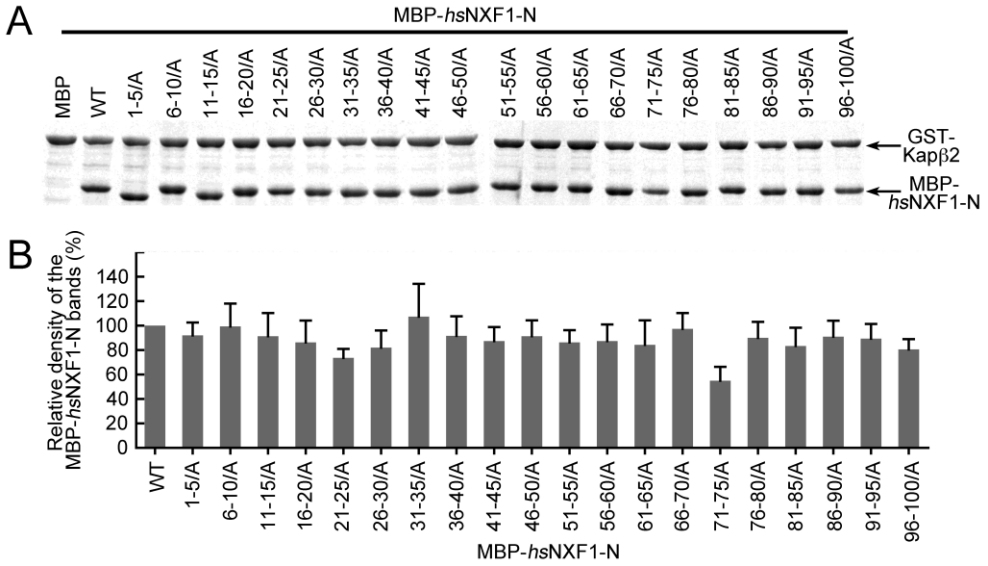

Figure 3-9 Mapping hsNXF1-N for Kapβ2 binding determinants. (A) Every five residues of hsNXF1-N (MBP fusion protein) were mutated into alanines and incubated with immobilized GST-Kapβ2. Bound proteins were visualized using Coomassie staining. (B) Gels in (A) were subjected to densitometry analysis. The density of the MBP-hsNXF1-N band in each lane was divided by the density of GST-Kapβ2 in the same lane (DMBP-hsNXF1-N/DGST-Kapβ2). The ratios were then normalized to the ratios of MBP-hsNXF1-N(WT) vs. GST-Kapβ2. Averages of 3 densitometry scans of the gels in (A) are shown in the histogram.

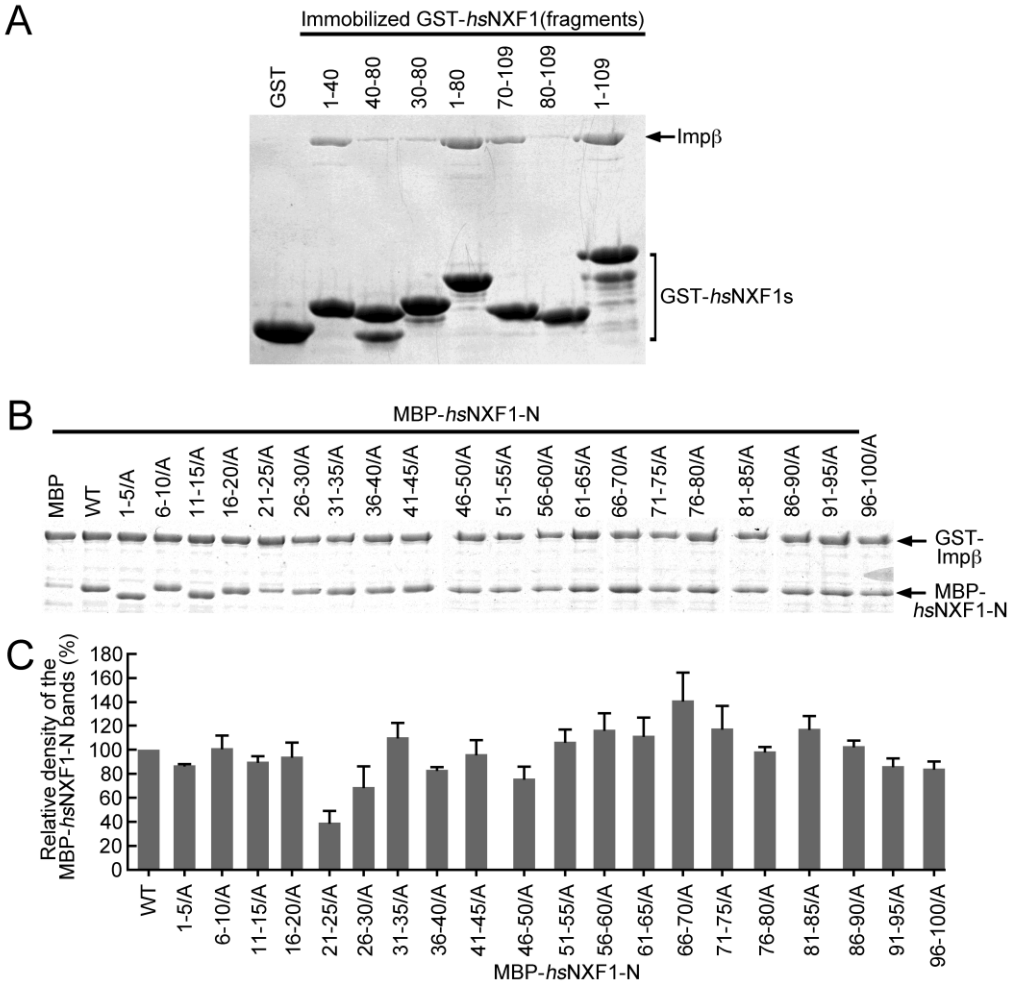

Figure 3-10 Mapping hsNXF1-N for Impβ binding determinants. (A) Immobilized GST fusion proteins of hsNXF1-N fragments were incubated with purified recombinant Impβ. (B) Every five residues of hsNXF1-N (MBP fusion protein) were mutated into alanines and incubated with immobilized GST-Impβ. Bound proteins in (A) and (B) were visualized using Coomassie staining. (C) Gels in (B) were subjected to densitometry analysis. The density of the MBP-hsNXF1-N band in each lane was divided by the density of GST-Impβ in the same lane (DMBP-hsNXF1-N/DGST-Impβ). The ratios were then normalized to the ratios of MBP-hsNXF1-N(WT) vs. GST-Impβ. Averages of 3 densitometry scans of the gels in (B) are shown in the histogram.

*hs*NXF1-N binds Impβ with high affinity. The Impβ binding isotherm of MBP-*hs*NXF1-N fitted a two-site binding model ($Chi^2 \sim 1.78X10^4$), with K_Ds of 6 nM and 1.5 μM, respectively (Table 3-1 and Figure 3-8. N- and C-terminal deletion mutants and alanine scanning mutagenesis of *hs*NXF1-N suggested that the binding energy for Impβ was likely concentrated in the first 40 residues of *hs*NXF1 with small contributions from residues 70-109 (Figure 3-11). MBP-*hs*NXF1-N(21-30/A) showed no detectable binding by ITC, suggesting that the ^{21}RKKKGRGPFR^{30} basic patch was indeed essential for interactions with Impβ (Table 3-1 and Figure 3-8). Mutations of ^{71}RYNPY^{75} to alanines had no effect on Impβ-binding (Table 3-1 and Figure 3-8).

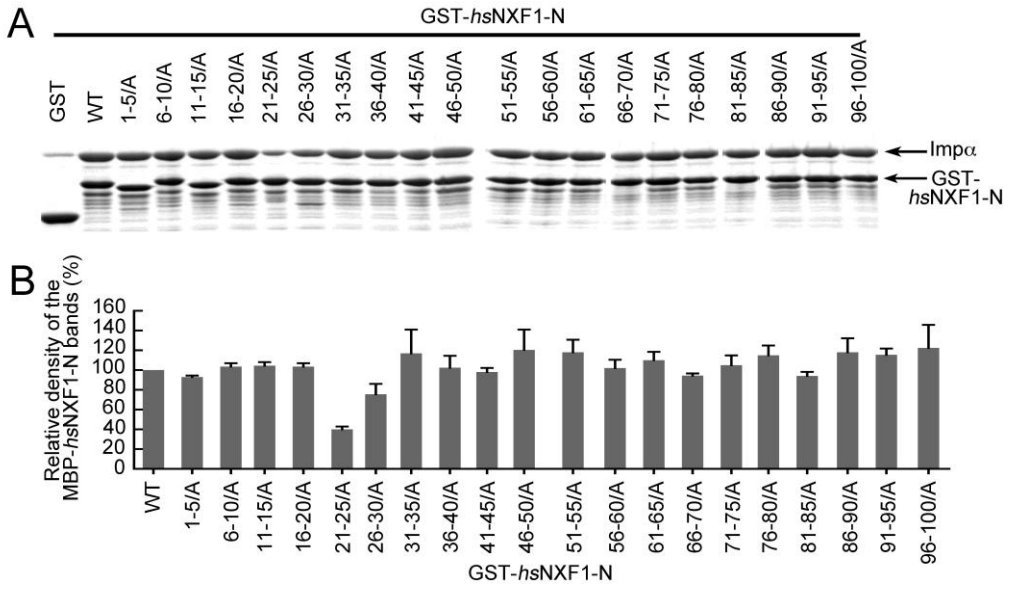

Figure 3-11 Mapping hsNXF1-N for Impα binding determinants. (A) Every five residues of hsNXF1-N (MBP fusion protein) were mutated into alanines and incubated with immobilized GST-Impα (the Impα construct used is missing its N-terminal IBB domain). Bound proteins were visualized using Coomassie staining. (B) Gels in (A) were subjected to densitometry analysis. The density of the MBP-hsNXF1-N band in each lane was divided by the density of GST-Impα in the same lane (DMBP-hsNXF1-N/DGST-Impα). The ratios were then normalized to the ratios of MBP-hsNXF1-N(WT) vs. GST-Impα. Averages of 3 densitometry scans of the gels in (A) are shown in the histogram.

The ^{21}RKKKGR^{26} segment of *hs*NXF1 matches the K-K/R-X-K/R consensus sequence for the monopartite classical-NLS (Chelsky *et al.*, 1989; Hodel *et al.*, 2001; Lange *et al.*, 2007; Yang *et al.*, 2010). Furthermore, scanning alanine mutagenesis revealed an Impα binding hotspot at residues 21-30 (Figure 3-11), suggesting that ^{21}RKKKGR^{26} might indeed be a bona fide monopartite classical-NLS. Scanning alanine mutagenesis of MBP-*hs*NXF1-N also suggested that the basic patch at residues 21-30 might contribute

significantly to interactions with Imp11 (Figure 3-12). Similar experiments were unsuccessful with Imp4 due to instability of the immobilized GST-Imp4.

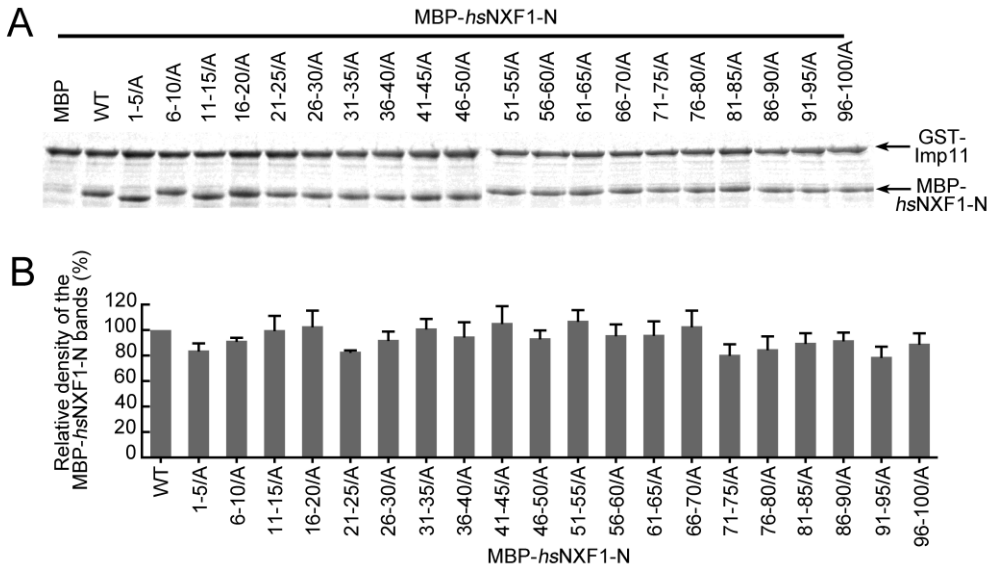

Figure 3-12 Mapping hsNXF1-N for Imp11 binding determinants. (A) Every five residues of hsNXF1-N (MBP fusion protein) were mutated into alanines and incubated with immobilized GST-Imp11. Bound proteins were visualized using Coomassie staining. (B) Gels in (A) were subjected to densitometry analysis. The density of the MBP-hsNXF1-N band in each lane was divided by the density of GST-Imp11 in the same lane (DMBP-hsNXF1-N/DGST-Imp11). The ratios were then normalized to the ratios of MBP-hsNXF1-N(WT) vs. GST-Imp11. Averages of 3 densitometry scans of the gels in (A) are shown in the histogram.

Collectively, the above results showed that interactions of *hs*NXF1 with Kapβ2, Impβ, Impα and Imp11 were differentially mediated by two distinct NLS epitopes. The R-X_{2-5}-P-Y motif at residues 71-75 of *hs*NXF1 is important for Kapβ2 binding whereas the ^{21}RKKKGRGPFR^{30} basic patch contributes significantly to interactions with Impβ, Impα and Imp11.

Mutation of the two NLS epitopes abolished Karyopherin-binding, mislocalized hs*NXF1 in cells and compromised gene expression.*

Mutations of the two NLS epitopes, ^{21}RKKKGRGPFR^{30} and ^{71}RYNPY^{75}, to alanines in MBP-*hs*NXF1-N(21-30,71-75/A) abolished binding to all five Karyopherins (Figure 3-5C). In order to determine the importance of the NLS epitopes for nuclear import, we transfected pyruvate kinase and EGFP fusions of full-length *hs*NXF1 proteins into HeLa cells (Figure 3-14A and 3-15). Pyruvate kinase (PK) alone localized to the cytoplasm whereas PK-*hs*NXF1 appeared exclusively in the nucleus. Mutations of the individual NLS epitopes in PK-*hs*NXF1(21-30/A) showed some cytoplasmic NXF1, PK-*hs*NXF1(71-75/A) showed more cytoplasmic mislocalization and mutation of both epitopes in PK-*hs*NXF1(21-30,71-75/A) showed extensive cytoplasmic mislocalization. NLS epitope mutants of *hs*NXF1-N-PK and EGFP-*hs*NXF1 showed similar mislocalization patterns as the PK-*hs*NXF1 mutants in HeLa cells (Figure 3-14A and 3-15).

To determine if nuclear import of *hs*NXF1 is important for mRNA export or NXF1-mediated gene expression, we examined how overexpressed *hs*NXF1 and its import mutants affected stimulation of Luciferase reporter gene expression. The expression levels of transfected EGFP-hsNXF1 proteins were similar (Figure 3-14C).. As expected, without significant overexpression of EGFP-*hs*NXF1 stimulated gene expression (Figure 3-14B) (Gruter *et al.*, 1998; Satterly *et al.*, 2007). *hs*NXF1-mediated stimulation of gene expression was decreased when either the *hs*NXF1 basic patch (^{21}RKKKGRGPFR^{30}) or its R-X_{2-5}-P-Y motif at residues 71-75 were mutated. The latter mutation had a larger

effect than the former, suggesting that of the five Karyopherins that import *hs*NXF1, Kapβ2 likely played the most important role. Simultaneous mutation of both ^{21}RKKKGRGPFR^{30} and ^{71}RYNPY^{75} epitopes lowered gene expression to the level of the EGFP control. These results showed that nuclear import of *hs*NXF1 is critical for its activity in mediating gene expression.

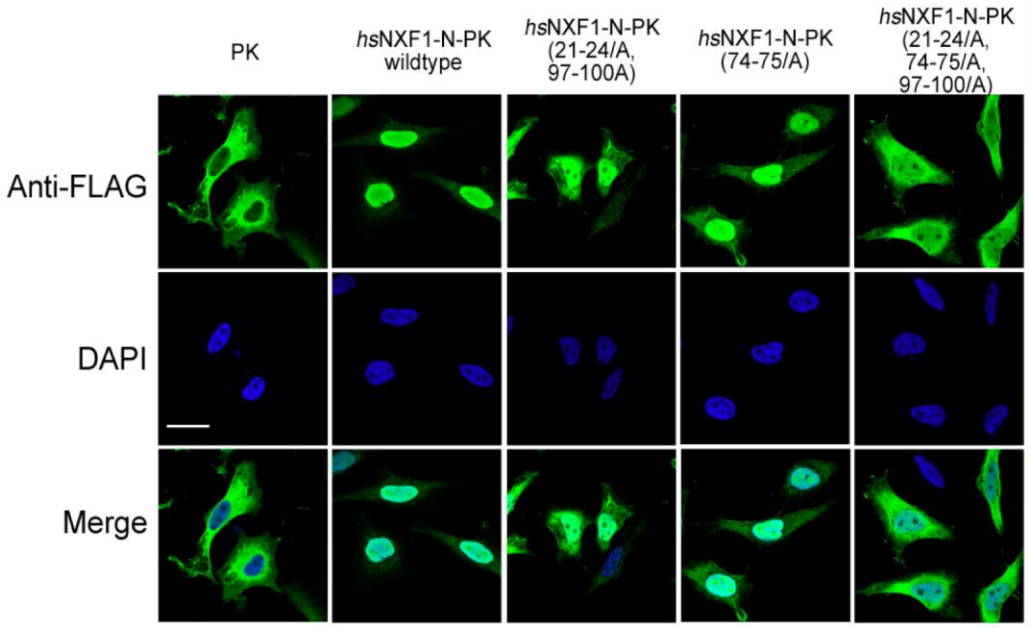

Figure 3-13 Mutations in the N-terminal basic NLS epitope and the C-terminal R-x2-5-P-Y NLS epitope diminish nuclear localization of hsNXF1-N. Pyruvate kinase (PK) fused to the C-terminus of hsNXF1-N proteins were cloned into the pFLAG-CMV2 vector and transfected into HeLa cells. hsNXF1-N-PK proteins were detected by immunofluorescence using anti-Flag antibodies. Scale bar, 10 μm**.**

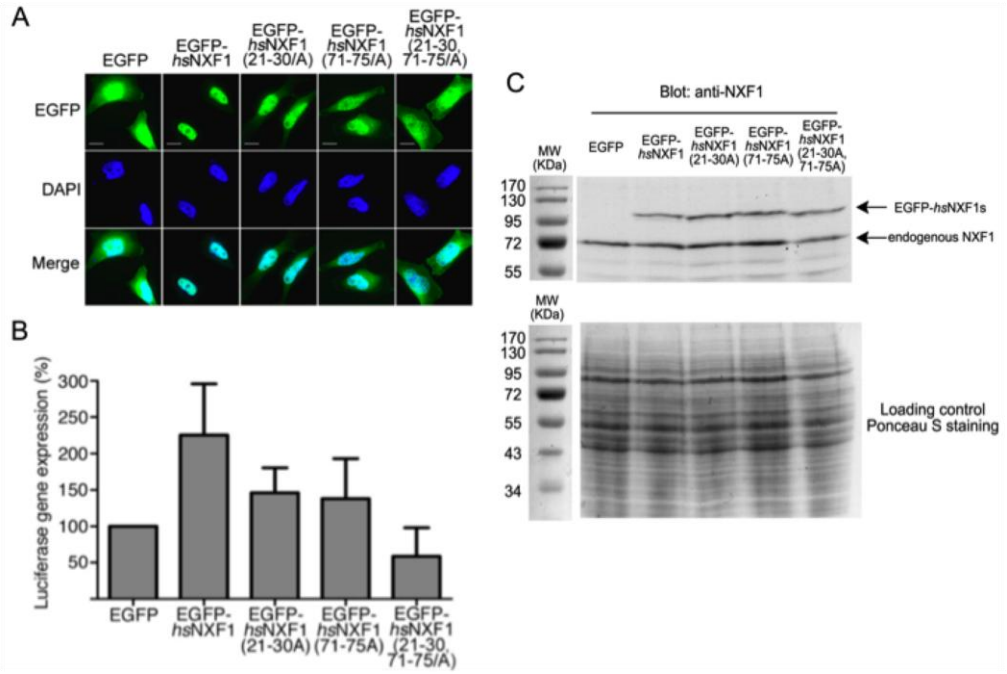

Figure 3-14 NLS mutations impair nuclear localization of hsNXF1 and its ability to activate Luciferase gene expression. (A) EGFP-hsNXF1 and its NLS mutants were transfected into HeLa cells. Localization of EGFP fusion proteins were detected by deconvolution microscope. Scale bars, 10 μm. (B) Luciferase reporter gene expression assays of hsNXF1 and its NLS mutants. EGFP-hsNXF1 proteins were cotransfected with pCMV-Luc vector and the expression levels of the Luciferase gene were calculated by normalizing the Luciferase signals that were detected by Luciferase Assay System to Celltiter-Glo signals. The results are averages of six independent experiments ± standard deviation. (C) The expression levels of transfected EGFP-hsNXF1 proteins and endogenous NXF1 were examined by western blotting.

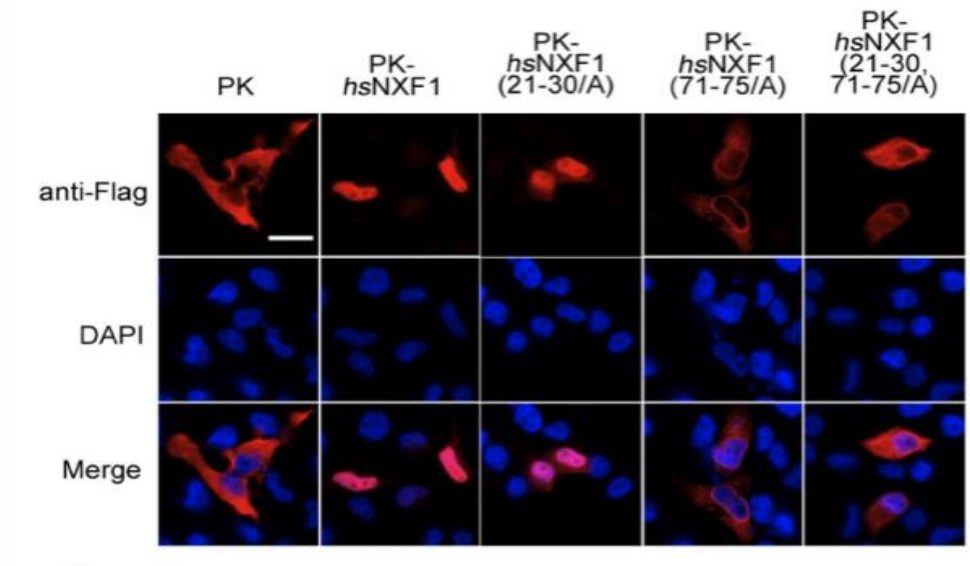

Figure 3-15 NLS mutations impair nuclear localization of EGFP-hsNXF1. Flag-tagged Pyruvate kinase (PK)-hsNXF1 and its NLS mutants were transfected into HeLa cells. Localization of PK fusion proteins was detected by deconvolution microscope. Scale bars, 20μm.

Potential NLS epitopes of NXF1 proteins in diverse eukaryotes.

Understanding how different Karyopherins recognize *hs*NXF1 was a necessary prerequisite to the identification of potential NLS epitopes in the N-terminal tails of different eukaryotic NXF1s. Residues 1-200 of *hs*NXF1 were used to identify NXF1 homologs by BLAST (Altschul, Gish et al. 1990). Sequences were available for NXF1 homologs from vertebrate, lancelets, tunicates, echinoderms, nematodes, insects and fungi. We examined the sequences of NXF1s from fungi (budding yeast *S. cerevisiae*; fission yeast *S. pombe* and *S. japonicus*) and animals (nematodes *C. elegans* and *S. briggsae*; insects *D. melanogaster*, *D. grimshawi*, *A. gambiae* and *A. aegypti*; chordates *H. sapiens, X. tropicalis, D. rerio, B. floridae* and *C. intestinalis*). Although these NXF1 homologs share ~30% sequence identities and have the same domain organization, their N-terminal tails shared no significant sequence homology (Altschul, Gish et al. 1990). In fact, their NXF1-Ns vary considerably in lengths. For example, the NXF1 homolog in *S. cerevisiae* Mex67p has a short 20-residue N-terminal tail whereas NXF1s of fission yeast *S. pombe* (*sp*Mex67p) and *S. japonicus* contain N-terminal tails that are 40-50 residues long. N-terminal tails of animal NXF1s are generally longer than 100 residues (Figure 3-16).

Instead of generating an alignment of all the very diverse NXF1-Ns, we aligned groups of closely related NXF1s from budding and fission yeasts, nematodes, insects and chordates (Figure 3-16) (Chenna, Sugawara et al. 2003; Dunn, Hejnol et al. 2008). We examined the NXF1-N groups for sequence/motif trends that are similar to the *hs*NXF1 NLS epitopes that we have characterized. In particular, we looked for basic patches that

resembled the basic NLS epitope of *hs*NXF1-N, segments that resembled the R-X_{2-5}-P-Y epitope of *hs*NXF1-N and the 7-residue acidic region that resides between the two NLS epitopes.

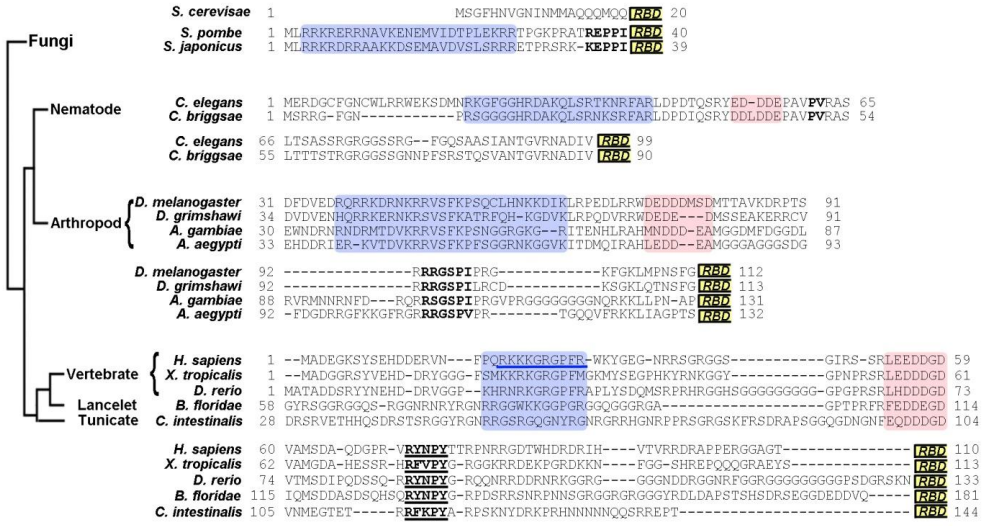

Figure 3-16 Potential NLS epitopes of NXF1 proteins in diverse eukaryotes. Residues 1-200 of hsNXF1 were used to identify NXF1 homologs by BLAST. Sequences were available for NXF1 homologs from vertebrate, lancelets, tunicates, echinoderms, nematodes, insects and fungi. Since the NXF1-Ns of divergent organisms shared no significant sequence homology and vary considerably in lengths, closely related NXF1s from within the groups fission yeast, nematodes, insects and chordates were aligned by ClustalW. The NXF1-N groups were examined for sequence trends similar to the NLS epitopes in the hsNXF1-N. The four divergent groups show similar organizations of motifs. N-terminal basic patches are shaded blue shades, with the N-terminal basic NLS epitope in hsNXF1 underlined in blue; central acidic patches are shaded pink; the C-terminal R/K/P-x2-5-PΦ motifs (Φ is a hydrophobic amino acid) are in bold, with the R-x2-5-P-Y motifs of chordate NXF1s underlined. RNP boxes indicate the beginning of RNP domains.

Interactions between Karyopherins and the N-terminal tails of chordate NXF1s.

NXF1s of the five chordates that we examined (*H. sapiens*, *X. tropicalis*, *D. rerio*, *B. floridae* and *C. intestinalis*) shared basic patches homologous to the ^{21}RKKKGRGPFR^{30} basic patch of *hs*NXF1, acidic regions that aligned with ^{53}LEEDDGD^{59} of *hs*NXF1 and the R-X_{2-5}-P-Y motifs (Figure 3-16). In fact, the R-X_{2-5}-P-Y motifs of all five chordate NXF1s matched the R-Y/F-X-P-Y consensus that is characteristic of energetically strong R-X_{2-5}-P-Y motifs (Suel, Gu et al. 2008; Suel and Chook 2009). Pull-down binding assays with recombinant NXF1-Ns showed that Kapβ2 bound human, *X. tropicalis* and *D. rerio* NXF1-Ns (*hs*NXF1-N, *xt*NXF1-N and *dr*NXF1-N, respectively) and their R-X_{2-5}-P-Y motifs were critical for the interactions (Figure 3-17A-C and Table 3-1). Impβ bound strongly to *hs*NXF1-N but weaker to *X. tropicalis* and *D. rerio* NXF1-Ns (Figure 3-17A, B and Table 3-1). Imp4 bound *hs*NXF1-N but not *X. tropicalis* and *D. rerio* NXF1-Ns, and all three vertebrate NXF1-Ns bound Imp11 (Figure 3-17A).

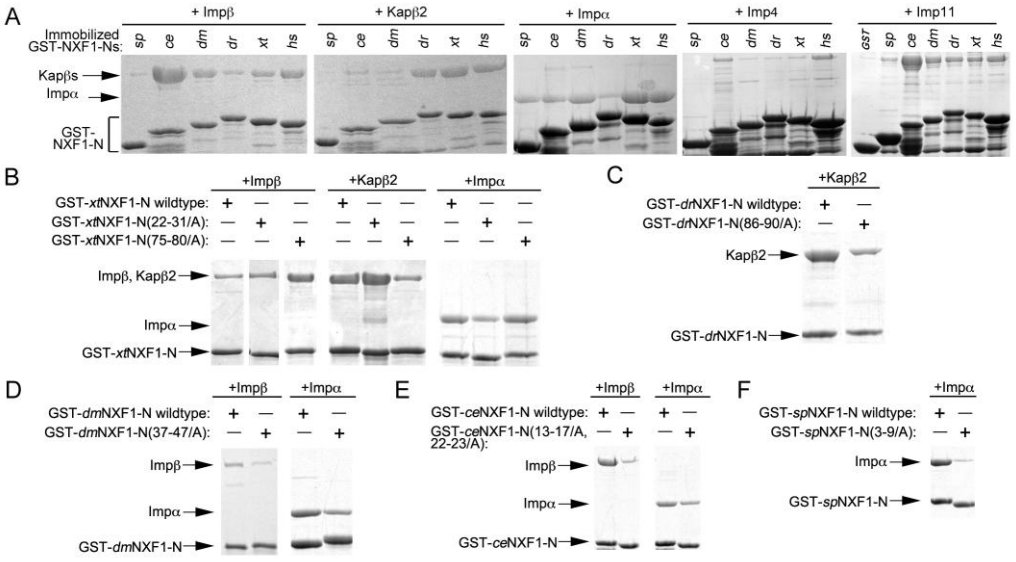

Figure 3-17 Interactions of NXF1-Ns from different organisms with Karyopherins. (A) Immobilized GST-NXF1-Ns of S. pombe, C. elegans, D. melanogaster, D. rerio, X. tropicalis and H. sapiens were incubated with purified recombinant Impβ, Kapβ2, Impα, Imp4 or Imp11. Bound proteins were visualized by SDS-PAGE and Coomassie staining. Mutations within the N-terminal basic patches and C-terminal R-x2-5-P-Y motifs of NXF1-Ns from X. tropicalis (B), D. rerio (C), D. melanogaster (D), C. elegans (E) and S. pombe (F) were tested for interactions with Impβ, Kapβ2 or Impα. Bound proteins in (A)-(F) were visualized using Coomassie staining.

The monopartite classical-NLS in ^{21}RKKKGR^{26} of *hs*NXF1 is preserved in ^{22}KKRKGR^{27} of *xt*NXF1 but no monopartite or bipartite classical NLSs are evident in the N-terminal tails of *D. rerio*, *B. floridae* or *C. intestinalis* NXF1s (Figure 3-16). These observations were supported by pull-down binding assays that showed binding of Impα to the basic patch of the human and *X. tropicalis* NXF1-Ns but not to that of the *D. rerio* NXF1-N (Figure 3-17A and B). These results suggested that many chordate NXF1s are likely Kapβ2 and Imp11 cargos and some are also imported into the nucleus through Imp4 and direct Impβ-binding and/or by the classical Impα/β pathway.

Interactions between Karyopherins and insect and nematode NXF1s.

Nematode (*C. elegans* and *S. briggsae*) and insect (*D. melanogaster*, *D. grimshawi*, *A. gambiae* and *A. aegypti*) NXF1s appear to all have N-terminal basic patches followed by small acidic regions but not R-X_{2-5}-P-Y motifs (Figure 3-16). Instead, insect NXF1s have PY-like R-X_3-P-I/V motifs C-terminal of their acidic regions that could potentially bind Kapβ2. The two nematodes have PAVPV segments that showed poor resemblance to the R-X_{2-5}-P-Y motif (Suel, Gu et al. 2008). Pull-down binding assays showed that neither *C. elegans* nor *Drosophila* NXF1-Ns (*ce*NXF1-N and *dm*NXF1-N, respectively) bound Kapβ2 (Figure 3-17A), suggesting that the PY-like R-X_3-P-I/V motifs in insect NXF1s are poor substitutes for the R-X_{2-5}-P-Y motif of PY-NLSs. In contrast, the basic patches in NXF1-Ns from both *C. elegans* and *Drosophila* contribute to direct interactions with Impβ (Figure 3-17A, D and E). Imp4 bound *ce*NXF1-N but not *dm*NXF1-N, and Imp11 bound to both *ce*NXF1-N and *dm*NXF1-N (Figure 3-17A).

Although there were no obvious monopartite classical NLSs in the *C. elegans* and *Drosophila* NXF1-Ns, the sequences of their basic patches matched the bipartite classical-NLS consensus sequence of K/R-K/R-$X_{10–12}$-K/$R_{3/5}$, where K/$R_{3/5}$ represents three lysine or arginine residues out of five consecutive amino acids (Figure 3-16) (Dingwall and Laskey 1991). These observations were supported by pull-down binding assays that showed binding of Impα to *C. elegans* and *Drosophila* NXF1-Ns (Figure 3-17A). Mutations of residues in the basic patches of both NXF1s decreased Impα-binding (Figure 3-17D and E). Collectively, these results showed that the classical Impα/β and direct Impβ pathways rather than the Kapβ2 pathway likely mediate nuclear import of NXF1 in nematodes and insects.

Interactions between Karyopherins and the N-terminal tails of S. pombe *NXF1. T*

The shorter N-terminal tails of fission yeast (*S. pombe* and *S. japonicus*) Mex67p contained N-terminal basic patches but no acidic regions or R-X_{2-5}-P-Y motifs (Figure 3-16). R/K-X_2-P-I segments at the C-terminal end of the tail most closely resembled the R-X_{2-5}-P-Y motif of a PY-NLS. The basic patches appeared to contain bipartite classical-NLSs (Figure 3-16). Pull-down binding assays showed binding of the *S. pombe* Mex67p N-terminal tail or *sp*NXF1-N to Impα, very weakly to Imp11 and not to Impβ, Kapβ2 or Imp4 (Figure 3-17A and F). These results suggested that nuclear import of NXF1 in *S. pombe* is most likely to be mediated by the classical Impα/β pathway.

Summary of Karyopherin-binding to NXF1-Ns from diverse organisms.

The trend of NXF1-Ns binding to human Karyopherins is conserved in binding assays using *S. cerevisiae* Kap95p, Kap60p and Kap104p (Figure 3-18). Binding analysis of diverse NXF1-Ns showed that the numbers of redundant NLSs in NXF1s and the Karyopherins that mediate their nuclear localization increase progressively from fungi to nematodes and insects to chordates (Table 3-3). The *S. cerevisiae* NXF1 contained neither N-terminal tail nor NLS. The *S. pombe* NXF1 appeared to use the classical Impα/β pathway. Nematodes and insects employed the classical Impα/β, direct Impβ and Imp11 pathways, and chordates employed 3-5 different nuclear import pathways to target their NXF1s to the nucleus.

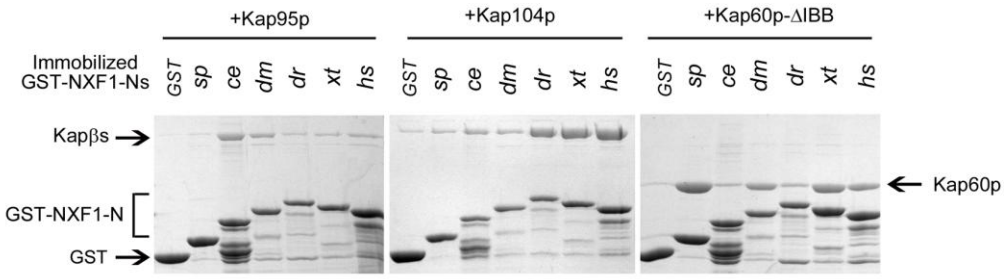

Figure 3-18 Interactions of NXF1-Ns from different organisms with S. cerevisiae Kap95p, Kap104p and Kap60p. Immobilized GST-NXF1-Ns of S. pombe, C. elegans, D. melanogaster, D. rerio, X. tropicalis and H. sapiens were incubated with purified recombinant Kap95p, Kap104p and Kap60p-ΔIBB. Bound proteins were visualized using Coomassie staining.

Table 3-3 Summary of interactions between Karyopherins and the N-terminal tails of NXF1s from diverse eukaryotes.

NXF1-N[a]	Kapβ2	Impβ	Imp4	Imp11	Impα
human	++++	++++	+++	+++	++++
X. tropicalis	++++	+++	-	+++	++++
D. rerio	++++	+	-	+++	+
D. melanogaster	-	+++	-	+++	+++
C. elegans	-	+++++	+	+++++	+++
S. pombe	-	-	-	+	+++

[a] binding data shown in Figure 3-17.

Discussion

*hs*NXF1 is a well-established nuclear import cargo of Kapβ2 (Truant, Kang et al. 1999; Bachi, Braun et al. 2000; Lee, Cansizoglu et al. 2006; Imasaki, Shimizu et al. 2007). The Karyopherin binds a PY-NLS in the N-terminal tail of *hs*NXF1 (Lee, Cansizoglu et al. 2006). Through extensive mutagenesis, qualitative and quantitative binding assays, we have shown that the PY-NLS of *hs*NXF1 spans residues 1-92, binds Kapβ2 with a K_D of 40 nM, and is a member of the basic and not the previously predicted hydrophobic subclass of PY-NLSs. We have identified binding determinants or NLS epitopes in two distinct segments of *hs*NXF1 that correspond to an N-terminal basic epitope at residues 21-30 and the R-X_{2-5}-P-Y motif at residues 71-75. The latter is a marginal hotspot whereby mutation of the entire 5-residue motif decreased Kapβ2 binding by 5-fold while mutation of the former decreased affinity by 3-fold. The basic/hydrophobic and R-X_{2-5}-P-Y epitopes of previously identified PY-NLSs are connected by 3-11 residues long linkers (Lee, Cansizoglu et al. 2006). The unusually long 40-residue PY-NLS linker in *hs*NXF1 significantly extends previous limits for linker lengths without compromising high affinity interactions with Kapβ2.

Surprisingly, inhibition of Kapβ2 by the M9M peptide inhibitor did not mislocalize endogenous *hs*NXF1 in HeLa cells, suggesting that Kapβ2 is not its sole nuclear import factor. We have shown that the N-terminal tail of *hs*NXF1 contains multiple redundant and overlapping NLSs that are recognized by Kapβ2, Impβ, Imp4, Imp11 and Impα. The five Karyopherins differentially bind the same two NLS epitopes that are recognized by Kapβ2. The basic patch at residues 21-30 is used in interactions with all five Karyopherins whereas the R-X_{2-5}-P-Y motif at residues 71-75 is used only for binding Kapβ2. The overlapping nature of the NLSs suggests that a single molecule of *hs*NXF1 likely binds only one Karyopherin molecule at a time. Mutations of both NLS epitopes greatly diminished nuclear localization of *hs*NXF1 and perturbed NXF1-mediated gene expression as observed by the significant decrease in reporter gene expression.

Our biochemical and biophysical characterization of the *hs*NXF1 NLS epitopes that bind Kapβ2, Impβ and Impα allowed extension of these studies to other eukaryotes. The N-terminal tails of NXF1s from fission yeasts, nematodes, insects and chordates share similar sequence/motif organizations even though they are very diverse in sequence and length. The N-terminal tails of nematode, insect and chordate NXF1s contain N-terminal basic patches of 10-30 residues, followed by acidic patches of about 6-8 residues and C-terminal R/K/P-X_{2-5}-P-Φ motifs. N-terminal tails of two fission yeast NXF1s show similar trends but lack the central acidic patches. No basic, acidic patches or R/K/P-X_{2-5}-P-Φ motifs are present in the N-terminal tail of *S. cerevisiae*. The N-terminal basic patches of the NXF1s are reminiscent of the N-terminal basic NLS epitope of *hs*NXF1

while their C-terminal R/K/P-X_{2-5}-P-Φ motifs resemble the R-X_{2-5}-P-Y motif of the *hs*NXF1 PY-NLS. Functions of the central acidic patches are currently not known.

Individual Karyopherins are highly conserved in eukaryotes, both in their sequences and cargo recognition (Enenkel, Blobel et al. 1995; Lange, Mills et al. 2008; Suel, Gu et al. 2008; Marfori, Mynott et al. 2010). The diverse NXF1s N-terminal tails bound similarly to human and *S. cerevisiae* Karyopherins, suggesting that Karyopherin specificities for their NLSs are conserved from human to yeast. We found that the number of Karyopherins that can mediate nuclear import of NXF1s increased steadily from fungi to nematodes and insects to chordates. Mex67p of *S. cerevisiae* has NLS and is known to be localized not to the nucleoplasm but to NPCs (Segref, Sharma et al. 1997; Katahira, Strasser et al. 1999). NXF1s from *S. pombe*, *C. elegans*, *drosopila* and human are known to be nuclear (Bear, Tan et al. 1999; Katahira, Strasser et al. 1999; Bachi, Braun et al. 2000; Tan, Zolotukhin et al. 2000; Yoon, Love et al. 2000; Herold, Klymenko et al. 2001; Wilkie, Zimyanin et al. 2001). Mex67p of *S. pombe* bound mostly Impα while the Karyopherin repertoires for *C. elegans* and *D. melanogaster* NXF1s were expanded to include Impα, Imp11 and direct interactions with Impβ. The complexity of nuclear import is further increased in chordates with the use of at least four Karyopherins: Impβ, Kapβ2, Imp11 and Impα.

The NLS epitopes recognized by Impβ and Impα are all located within the N-terminal basic patches of the NXF1 proteins while Kapβ2 recognized the R-X_{2-5}-P-Y motifs in chordate (*H. sapiens*, *X. tropicalis* and *D. rerio*) NXF1s. Interestingly, the slightly divergent R/K-X_2-P-I, P-X_2-P-V and R-X_{2-3}-P-I/V motifs in *S. pombe*, *C. elegans* and *D.*

melanogaster, respectively, were unsuitable for Kapβ2 binding. Therefore, it appears that strong R-X_{2-5}-P-Y motifs evolved only in chordates to expand nuclear import to Kapβ2. The motif, in combination with the more primitive basic patch, produced functional basic PY-NLSs in the NXF1s of these higher eukaryotes, resulting in a total of 3-5 different nuclear import pathways that target NXF1s to the nuclei of human cells.

It is puzzling that the means of transporting NXF1 into the nucleus are different from *S. cerevisiae* to humans even though its mRNA export function is conserved. What are the advantages of increased complexity in NXF1 nuclear import or increased redundancy of NXF1 nuclear import pathways in higher eukaryotes? The simplistic suggestion that redundant nuclear import pathways are necessary to ensure correct localization of NXF1 to the nucleus for the crucial process of mRNA export is rather unsatisfactory given that *S. cerevisiae* Mex67p has no NLSs and does not need to be localized to the cell nucleus at all. It is more likely that redundant NLSs in NXF1s are important to regulate mRNA export and its coupling to the upstream and downstream gene expression processes of transcription, splicing and/or translation.

NXF1 binds mRNAs weakly, but the interaction is significantly enhanced by adaptor proteins REF and SR proteins (Hautbergue, Hung et al. 2008). In higher eukaryotes, adaptor proteins couple mRNA export to upstream processes of capping and splicing (Izaurralde and Mattaj 1995; Zhou, Luo et al. 2000; Masuda, Das et al. 2005; Cheng, Dufu et al. 2006). Interactions with mRNA and adaptor proteins were mapped to *hs*NXF1 residues 61-118 and 1-362, respectively (Bachi, Braun et al. 2000; Stutz, Bachi et al. 2000; Huang, Gattoni et al. 2003), thus overlapping significantly with Karyopherin

binding. In the nucleus, the termination of NXF1 import is likely coupled to its interactions with mRNA, adaptor proteins and to upstream processes of capping and splicing. In the cytoplasm, the Karyopherins that import NXF1 may contribute to its release from adaptor proteins and mRNA prior to translation. Furthermore, differential binding of Kapβ2, Impβ, Imp4, Imp11 and Impα to the N- and C-terminal NLS epitopes of *hs*NXF1 may affect its interactions with various subsets of adaptor proteins, thus providing a means of regulating assembly and disassembly of diverse populations of mRNA export complexes.

Finally, the striking difference in nuclear localization of NXF1 in higher eukaryotes but not in *S. cerevisiae* may reflect new and still undetermined functions of NXF1 in the nucleus of higher eukaryotes. The discovery of the mRNA poly(A) processing factor CPSF30 as a direct binding partner of *hs*NXF1 and a mediator of a crosstalk between the NXF1- and CRM1-mediated mRNA export pathways may represent an intranuclear regulatory and compensatory step acquired by higher eukaryotes (Satterly, Yarbrough et al. 2011). Knockdown of CPSF30 by siRNA rescued the inhibition of mRNA export induced by NXF1 siRNAs and the observed mRNA export release occurred via CRM1. This connection between poly(A) processing and mRNA export is possibly a checkpoint in which CPSF30 would be released from NXF1 only upon proper polyadenylation, which would then allow NXF1 to promote mRNA export. The increasing complexity of NXF1 nuclear import in higher eukaryotes may be correlated with similar complexity in nuclear functions of NXF1 The architecture of modular NLS epitopes within the flexible and structurally disordered N-terminal tail of NXF1 may have allowed significant

evolvability to form multiple NLSs (Suel, Gu et al. 2008). This in turn could have provided a path for NXF1 to switch from using one Karyopherin to another and ultimately from one cellular process to another.

CHAPTER FOUR

CRYSTALLIZATION OF KAPβ2-NXF1-NLS COMPLEX

Abstract

Karyopherinβ2 (Kapβ2) imports numerous mRNA binding proteins and it recognizes a class of nuclear localization signals (NLSs) called PY-NLSs. Human NXF1, which is the major mRNA export factor, was reported as a cargo of Kapβ2. Our biochemical analysis revealed that NXF1 has a more complex PY-NLS, which is longer than other known PY-NLSs and some binding epitopes were missing in previous mapped NXF1 NLSs. In order to understand how Kapβ2 accommodates this longer PY-NLS and visualize the interactions between NXF1 and Kapβ2, we crystallized the complete PY-NLS of NXF1 (residues 1-92) in complex with Kapβ2 to solve the structure of the complex.

Introduction

Kapβ2 imports numerous mRNA binding proteins into the nucleus. Crystal structures of unliganded Kapβ2, Kapβ2 complexes with NLSs of substrates hnRNPs A1, M and D, JKTBP, NXF1 as well as a Kapβ2 complex with RanGTP have been solved (Chook and Blobel 1999; Lee, Cansizoglu et al. 2006; Cansizoglu and Chook 2007; Cansizoglu, Lee et al. 2007; Imasaki, Shimizu et al. 2007). The structures of Kapβ2 bound to its cargos show that these PY-NLSs are structurally conserved only at the consensus motifs and the

linkers that connect these motifs are structurally variable (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007). The linkers also vary in both sequence and length across the PY-NLS family (Lee, Cansizoglu et al. 2006). The distribution of binding energies among these sites is also quite variable in different PY-NLSs (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007; Imasaki, Shimizu et al. 2007; Suel, Gu et al. 2008). The large surface of Kapβ2 and its inherent flexibility raise the possibility that Kapβ2 may recognize other types of NLSs. It will be interesting to examine the structures of Kapβ2 in complex with other substrates, which may lead the discovery of new classes of NLSs.

Previous studies have shown that human NXF1 (hsNXF1) is a cargo of Kapβ2 and its NLS, which shows no homology to the NLSs of hnRNPs A1 and M, is located within the N-terminal 120 residues (Bear, Tan et al. 1999; Kang and Cullen 1999; Truant, Kang et al. 1999; Lee, Cansizoglu et al. 2006). This region possesses the common features of PY-NLSs recognized by Kapβ2, such as structural disorder, overall positive charge and the $RX_{2\text{-}5}PY$ motif (Lee, Cansizoglu et al. 2006). However, crystal structure of Kapβ2 bound to residues 53-82 of hsNXF1 shows electron density only for hsNXF1 residues 68-79, a short fragment at the PY motif (Imasaki, Shimizu et al. 2007). This finding is consistent with our previous unpublished structures of Kapβ2 with hsNXF1(40-80) and hsNXF1(67-102), where strong 2Fo-Fc and Fo-Fc electron densities are observed only around the PY motif (Cansizoglu. A.E. & Chook, Y.M., unpublished data). Weak electron density N-terminal to this region, extending towards the region analogous to the N-terminal motif in hnRNP A1 and M-NLSs can be observed but not modeled (Dr. Yuh

Min Chook's observation; data not shown). Furthermore, quantitative binding data suggests that hsNXF1-NLS fragments used in previous structural studies were likely missing energetically significant binding determinants/epitopes for Kapβ2 (Table 3-2 and 4-1). The predicted N-terminal hydrophobic epitope ^{59}VAMS^{62} does not have significant contribution to binding energy (Figure 3-7). Collectively, these results suggest that the PY-NLS of hsNXF1 may be more complex with a longer linker and still undefined N-terminal binding epitopes. This chapter describes the effort to crystallize the complex of Kapβ2 with the complete PY-NLS of hsNXF1 and solve the structure so that we can compare the interactions of hsNXF1-PY-NLS with other classic PY-NLSs.

Materials and Methods

Protein purification and complex formation

In this crystallographic study, residues 337-367 of human Kapβ2 (accession number AAB58254) were replaced with a GGSGGSG linker because the acidic loop region causes instability of the crystals. The resulting deletion mutant Kapβ2Δloop3 was expressed in pGexTev vector as a N-terminal GST fusion protein at 25℃ for 16 hours. The cells were resuspended in Tris buffer (50 mM Tris pH 7.5, 100 mM NaCl, 2 mM EDTA, 2 mM DTT, 20% glycerol) with protease inhibitors (100 μg/ml of pefabloc, 157 μg/ml of benzamidine, 10 μg/ml of leupeptin) and lysed using cell disruptor EmulsiFlex-C5 (Avestin, Inc, Ontario, Canada). The clarified supernatants were loaded onto 10 ml GSH sepharose beads (GE Healthcare, NJ, USA) and passed through twice, then the beads were washed with 50 ml of Tris buffer 5 times, 20 ml of ATP buffer (50mM Tris

pH7.5, 2mM DTT, 5mM ATP, 10mM MgAC, 2mM EGTA, 20% glycerol) 5 times at room temperature, and 50 ml of Tris buffer twice. The bound proteins were eluted with 15 ml of Tris buffer containing 20 mM glutathione (pH8.0) 5 times. The eluates were concentrated to 10 ml and added 1 ml of TEV protease to cleave the GST tag off at room temperature overnight before they were loaded onto 5 ml Hitrap Q column (GE Healthcare, NJ, USA) (Figure 4-1) Fractions containing Kapβ2Δloop3 were collected and injected onto Superdex S200 column (GE Healthcare, NJ, USA) after concentration (Figure 4-2). The purified Kapβ2Δloop3 were flash frozen and stored at -80 ℃ for future use.

Human NXF1 (Uniprot: Q9UBU9) residues 1-92 were cloned into pGexTev vector and expressed at 25 ℃ for 16 hours. The fusion proteins were purified by GSH affinity chromatography and the eluates were loaded onto 5 ml Hitrap SP column (GE Healthcare) (Figure 4-3). The fractions containing GST-hsNXF1-(1-92) were collected, flash frozen and stored at -80 ℃ for future use.

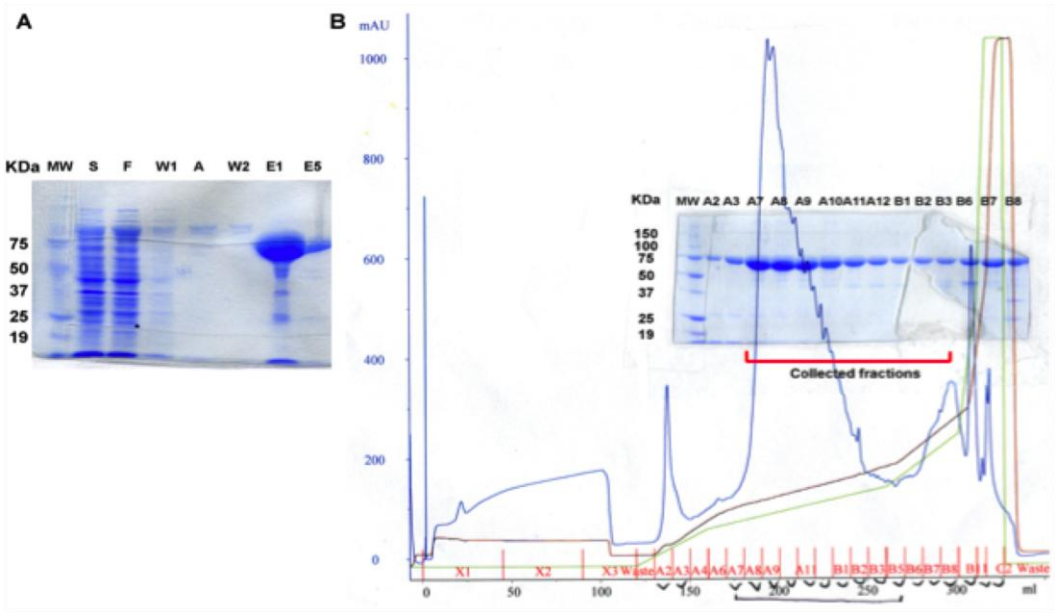

Figure 4-1 Purification of GST-Kap®2 Δ loop3. (A) The gel samples of GSH affinity purification of GST-Kap®2 Δ loop3., S (supernatant), F (flowthough), W1 (first wash with Tris buffer), A (fifth wash with ATP buffer), W2 (second wash with Tris buffer), E1 (first fraction of elution), E5 (fifth faction of elution); (B) The chromatograph of Kap β 2 Δ loop3 on 5 ml Hitrap Q column. The gels samples from the indicated fractions were run on 12% SDS-PAGE gel.

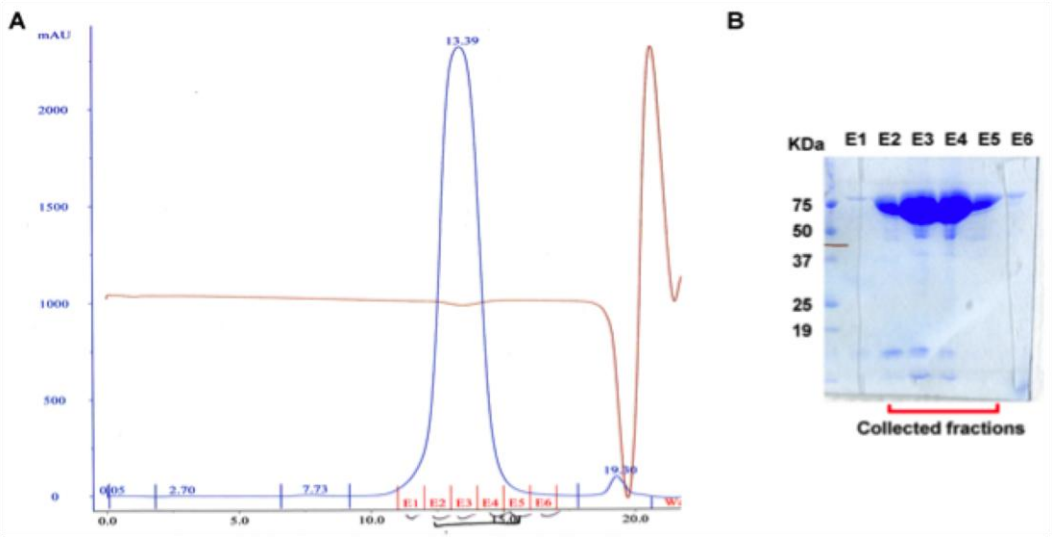

Figure 4-2 Purification of Kapβ2 Δ loop3 by gel filtration. (A)The chromatogram of Kapβ2 Δ loop3 on Superdex S200 column. (B)The gel samples from each fraction were run on 12% SDS-PAGE gel. The indicated fractions were collected, concentrated and stored at -80˚C for further experiments.

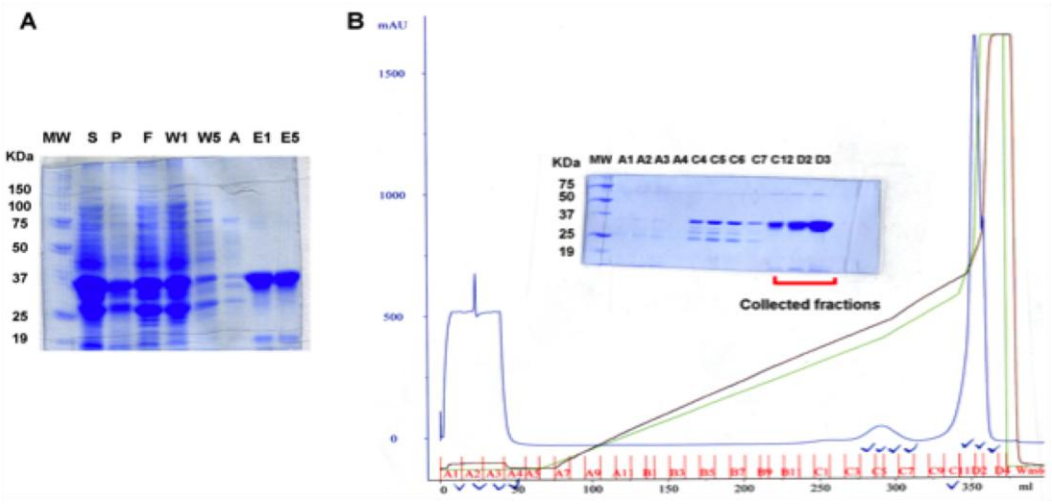

Figure 4-3 Purification of GST-hsNXF1-(1-92). (A) The gel samples of GSH affinity purification of GST-hsNXF1-(1-92). S (supernatant), P (pellet), F (flowthough), W1&W5 (first and fifth wash with Tris buffer), A (first wash with ATP buffer), E1 (first fraction of elution), E5 (fifth faction of elution); (B) The chromatograph of GST-hsNXF1-1-92 on 5 ml Hitrap SP column. The gels samples from the indicated fractions were run on 12% SDS-PAGE gel

To form the complex, Kapβ2Δloop3 and GST-hsNXF1-(1-92) were mixed at 4℃ in a molar ratio of 1: 5 and cleaved with TEV protease overnight, followed by tandem purification with 5 ml Hitrap SP column, Superdex S200 and 2 ml GSH column (Figure 4-4 and 4-5). The purified complexes were concentrated to about 30 mg/ml for crystallization.

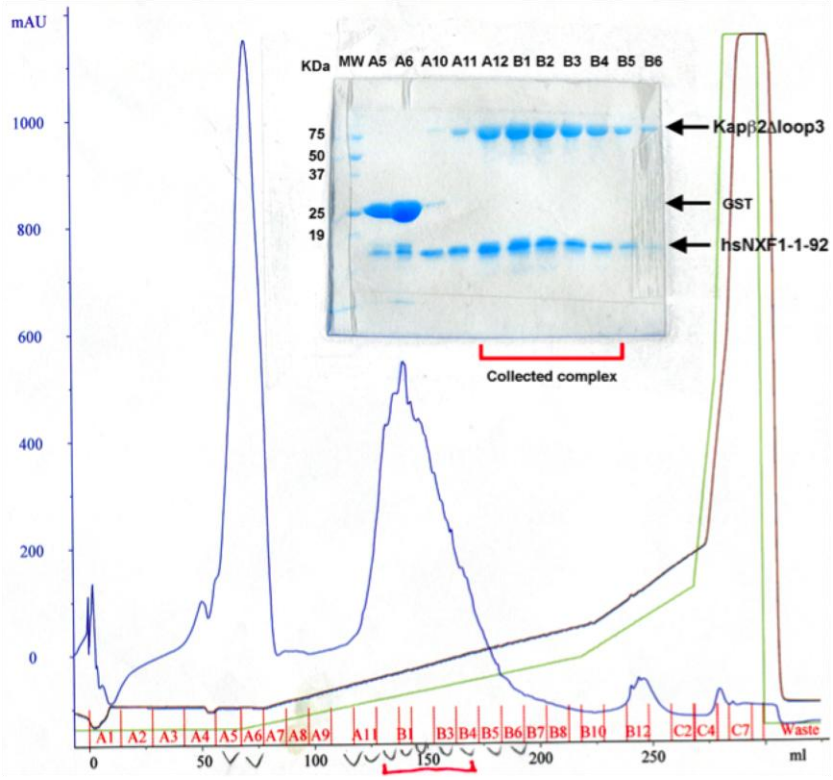

Figure 4-4 Purification of Kap®2 Δ loop3-hsNXF1-(1-92) complex by ion exchange chromatography. The chromatograph and gel samples of Kap®2 Δ loop3-hsNXF1-(1-92) complex on 5 ml HiTrap Q column**.**

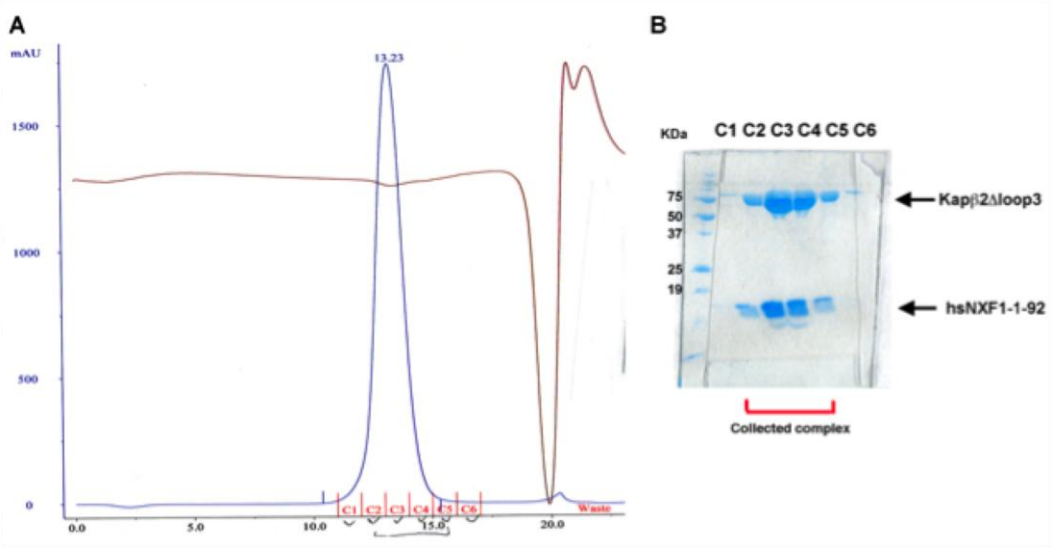

Figure 4-5 Purification of Kapβ2 Δ loop3-hsNXF1-1-92 complex by gel filtration. (A) The chromatograph of Kap®2⊗loop3-hNXF1-1-92 complex on Superdex S200 column. (A) The gel samples from each fraction were run on 15% SDS-PAGE gel and stained with Coomassie Blue R-250. The indicated factions were pooled together, concentrated and stored at -80˚C for further experiments

Crystallization and crystal screen

The Kapβ2Δloop3-hsNXF1-(1-92) complex was crystallized by vapor diffusion in hanging and sitting drops. Based on previous crystallization studies of Kapβ2 and its cargos, potassium formate (KF) was used as precipitant in the presence of glycerol. Detailed optimization were carried out with various concentrations of KF (1.0-3.2 M), glycerol (0-20%) and complex (5-30 mg/ml) in 0.1 M Hepes buffer (pH7.0-7.4) or 0.1 M MES buffer (pH6.2-6.6) at 4ºC, 16ºC, 20ºC, 25ºC. Additive screen HT™ (Hampton Research, CA, USA) and different approaches including seeding, dilution, microbatch, dehydration and annealing were tried in an attempt to improve the quality of the crystals.

Crystals with nice shape and size bigger than 100 μm were flash frozen in liquid nitrogen for screening.

Crystals were first screened at the home source (Rigaku Americas, TX, USA) with the exposure time of 5 min, image width of 1 º, at detector length of 200 mm. Single crystals that diffracted beyond 3.5 Å were saved, four of which were sent to APS for data collection.

Data collection and processing

Data was collected at the Advanced Photon Source at beamline 19-ID, Argonne National Laboratory at X-ray wavelength 12.66 keV and temperature 100 K, and processed with HKL2000 (Otwinowski and Minor 1997) (Table4-2). Kapβ2 from Kapβ2-hnRNP A1-NLS structure (PDB ID code 2H4M, (Lee, Cansizoglu et al. 2006) was used as a search model to solve the structure by molecular replacement using AutoMR in Phenix (Adams, Afonine et al. 2010). The structure was refined against the native dataset by iterative manual model building in Coot(Emsley, Lohkamp et al. 2010) and refinement using the Phenix refinement module (Adams, Afonine et al. 2010)to reduce model bias.

Results and Discussion

Formation of Kapβ2-hsNXF1-NLS complex

The PY-NLS of hsNXF1 resides in its N-terminal disordered tail, which spans residues 1-120. In order to map the smallest fragment that contains the complete NLS recognized by Kapβ2, a series of truncated NXF1s were generated and expressed as MBP fusion

proteins (Figure 4-6). The binding affinities of these fragments were measured by ITC and summarized in Table 4-1. The hsNXF-(1-109) fragment has low-nanomolar affinity (K_d=40nM) that is similar to those of other Kapβ2 cargos hnRNP A1 (K_d=42nM) and hnRNP M (K_d= 10mM) (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007), and most likely contains the complete PY-NLS. Truncation of N-terminal residues 1-30 or C-terminal residues 80-109 reduced binding affinities by 2-3 fold compared to hsNXF1-(1-109), suggesting that these regions contain binding epitopes for Kapβ2. hsNXF1-(1-92) is the shortest fragment that still binds Kapβ2 with high affinity similar to hsNXF1-(1-109) and was chosen as the minimal complete PY-NLS for crystallization.

Table 4-1 Kapβ2 binding to hsNXF1 Fragments
Dissociation constant by isothermal calorimetry

	hsNXF1 N-terminal Truncations	Kd (nM)	ΔH (kcal/mol)	TΔS (kcal/mol/K)	Kd,truncation/ Kd,NXF1-N
Kapβ2	1-109	40±13	-17.21±1.25	-7.31±1.42	
	10-109	134±2	-16.33±0.23	-7.09±0.24	2.4
	20-109	117±8	-16.01±0.23	-6.70±0.26	2.1
	30-109	91±13	-18.51±1.13	-9.16±1.12	1.7
	40-109	176±19	-13.48±0.26	-4.43±0.24	3.2
	70-109	203±14	-9.51±0.18	-0.54±0.23	3.7
	1-92	54±4	-19.57±0.05	-9.82±0.03	1.0
	1-80	109±33	-17.74±0.89	-8.40±0.88	2.0
	30-80	204±29	-19.16±1.31	-10.17±1.40	3.7
	40-80	3112±528	-12.51±0.95	-5.12±0.93	56.6
	53-82	1416±22	-15.27±0.05	-7.42±0.06	25.8

GST-hsNXF1-(1-92) was expressed in *E. coli*. After affinity and ion exchange purification, it was mixed with purified Kapβ2Δloop3 to form the complex. The GST tag was removed by TEV protease after complex formation. The complex was further purified by ion exchange column, gel filtration, and finally put through GSH beads to

remove residual GST proteins. The cleavage product hsNXF1-(1-92) (apparent MW of 10 KDa on SDS-PAGE gel) co-eluted with Kapβ2Δloop3 on during chromatography (Figure 4-4 and 4-5), indicating successful complex formation. The purity of the final complex is more than 95% based on Coomassie blue staining (Figure 4-5).

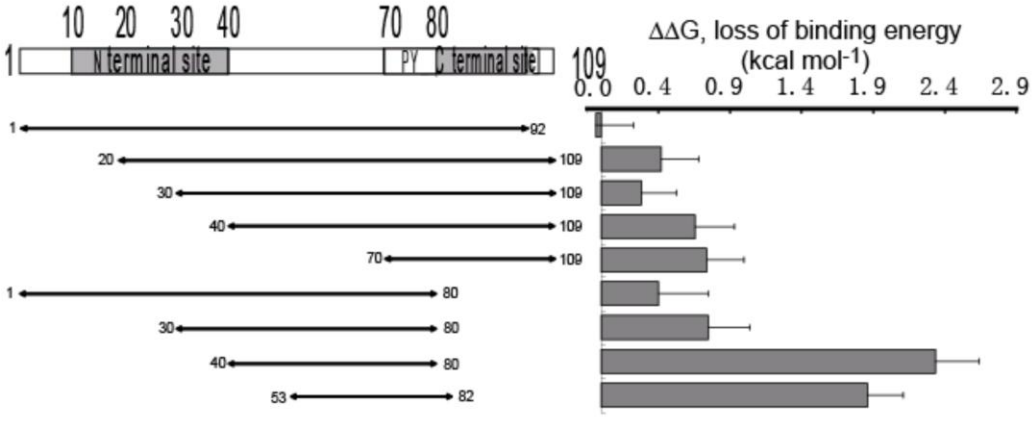

Figure 4-6 Schematic representation of hsNXF1 truncations. The bar graph shows the loss of binding energy of each trunction mutant compared to hsNXF1-(1-109). $\Delta\Delta G = -RT\ln[K_{d(truncation)}-K_{d(1-109)}]$.

Crystallization and optimization

In previous studies, crystals of Kapβ2 bound to the hnRNP A1-NLS and the hnRNP M-NLS were obtained in crystallization conditions containing potassium formate and glycerol (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007). In addition, former graduate student A. Ertugrul Cansizoglu had crystallized complexes of Kapβ2 bound to hsNXF1-(67-102), hsNXF1-(40-80) and hsNXF1-(20-120) in similar conditions. Crystals of Kapβ2-hsNXF1-(1-92) complex were also easily obtained in these conditions. They were shaped like cuboids or plates (Figure 4-7B). The crystals were harvested, washed

extensively, and dissolved in buffer for analysis by SDS-PAGE. They contained both Kapβ2 and hsNXF1-(1-92) (Figure 4-7A). Glycerol helps nucleation, but concentrations higher than 12.5% glycerol produced too many small crystals. The optimal condition for spontaneous crystal growth is 100mM of MES pH6.6, 2.4-2.8 M of KF and 10-12.5% of glycerol. Crystals can spontaneously grow to dimensions of 300 μm × 80 μm. These crystals were very prone to radiation damage and did not diffract beyond 3.5 Å at the home source. Seeding and dilution methods were not successful in producing thicker crystals. Dehydration did not improve the resolution. Annealing method that thaws and refreezes the crystals destroyed the crystals. The crystals growing in microbatch did not show better quality than the normally grown ones. Additive screen were performed and seven reagents were found to help the 3D single crystals grow to bigger than 200μm.: yttrium chloride hexahydrate, potassium sodium tartrate tetradydrate, phenol, galactose, NDSB-195, NDSB-201 and γ-butyrolactone. Among these additives, γ-butyrolactone was chosen for further optimization because it reduced the nucleation and allows the single crystals to grow bigger. But the crystals grown in reservoir buffers containing γ-butyrolactone only have marginal improvement on resolution. The best four crystals were sent to APS for data collection.

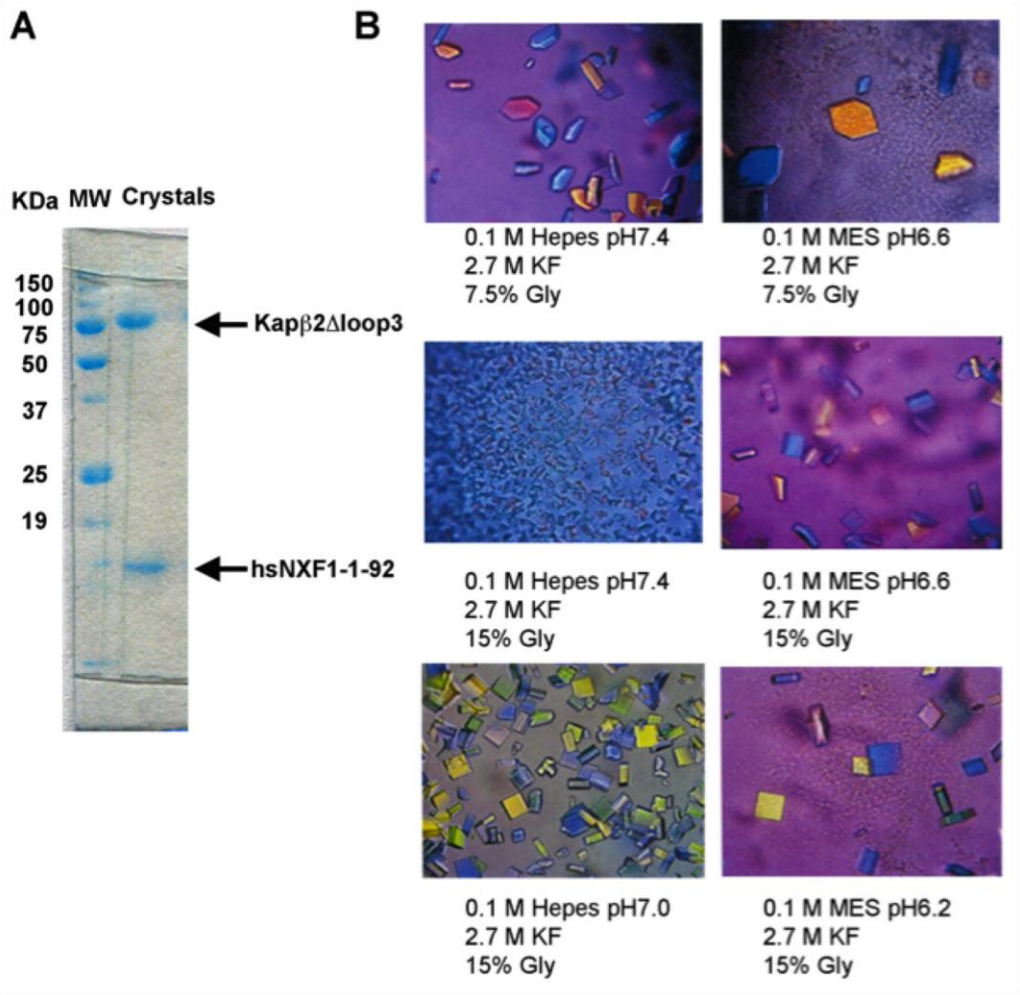

Figure 4-7 Crystallization of Kapβ2 Δ loop3-hsNXF1-1-92 complex. (A) Crystals were harvested, washed and dissolved in SDS sample buffer, then run on 15% SDS-PAGE gel. (B) Images of crystals under microscope (10X objective lens). The reservoir conditions were labeled

Data collection, structure determination and model building

The Kapβ2-hsNXF1-(1-92) crystals have the same C2 space group as previous Kapβ2-hsNXF1-(40-80) and Kapβ2-hsNXF1-(67-102) crystals (Table 4-2). The dimensions of

the unit cell (a=152.959, b=153.771, c=141.802, β=92.670) are also very close to the crystals of Kapβ2-hsNXF1-(40-80) complex. Kapβ2 chains from the structure of Kapβ2-hnRNP A1 complex was used as a search model for molecular replacement using the AutoMR module in the program Phenix (Adams, Afonine et al. 2010). Solutions for rotation and translation functions were found and model building is ongoing with the program Coot (Emsley, Lohkamp et al. 2010). The asymmetric unit contains two Kapβ2-NLS complexes. The relatively low quality of the data set makes it very difficult to trace the NLS peptide in the electron density map. Better crystals with higher resolution will be needed to complete structure determination. In the future, different constructs of hsNXF1-NLS may be tried along with other crystallization conditions.

Table 4-2 Crystallographic Data and Refinement Statistics

	hsNXF1 67-102 Complex	hsNXF1 40-80 Complex	hsNXF1 1-92 Complex
Space group	C2	C2	C2
Cell dimensions	a=154.5Å, b=155.0Å, c=70.9Å, β=92.285°	a=153.4Å, b=154.1Å, c=141.6Å, β=92.3°	a=152.9Å, b=153.8Å, c=141.8Å, β=92.7°
Resolution	50-3.2Å	50-3.1Å	50-3.2Å
Redundancy	3.7 (3.0)	6.2(5.2)	3.3(3.1)
R_{merge}	0.057(0.446)	0.074 (0.64)	0.111(0.99)
Completenes	98.1% (93.0%)	90.6% (98.1%)	97.2% (98.1%)
I/σ	26.8 (1.9)	25.7 (2.1)	19.4 (1.2)
R_{work}/R_{free}	33.4/41.2	27.6/32.5	

Conclusions

Biochemical analysis has identified that residues 1-92 of hsNXF1 is the minimal fragment of the complete PY-NLS of hsNXF1. The complex of Kapβ2 bound to hsNXF1-(1-92) was crystallized in the similar conditions as other Kapβ2-PY-NLS complexes. The resulting crystal also has the C2 symmetry and its unit cell dimensions are very close to the crystal of Kapβ2-hsNXF1-(40-80), suggesting that the longer PY-NLS did not affect the crystal packing and the molecules are in the same orientation as previous crystals. It is very promising to simply solve the structure by molecular replacement. However, the quality of the Kapβ2-hsNXF1-(1-92) crystal needs to be improved for further model building.

CHAPTER FIVE

VALIDATION OF PREDICTED PY-NUCLEAR LOCALIZATION SIGNALS *

Abstract

Bioinformatic search using a set of physical predictive rules from previous studies led to the prediction of 81 candidate PY-NLSs. In this chapter, I describe biochemical studies to validate these putative PY-NLSs. 72 out of 81 predicted PY-NLSs on the list were cloned and tested using for Kapβ2 binding and Ran dissociation. Of the 77 tested PY-NLSs, 13 showed strong binding to Kapβ2, 8 showed moderate binding and 56 have very weak or no binding. Alanine mutagenesis of 7 PY-NLSs revealed that their conserved PY motifs are critical for Kapβ2 binding. The information gathered from this in vitro validation study will be valuable to modify and improve cargo recognition rules for Kapβ2.

---

* Part of this chapter was originally published in Cell 126(3): 543-58.

Introduction and Background

Signal-directed nuclear transport of proteins is the critical regulatory step for gene expression but large sequence diversity among various cargos has prevented identification of NLSs for most Kapβs. It remains extremely difficult to predict NLSs in candidate import cargos. Previous structural and biochemical studies on Kapβ2-hnRNP A1-NLS complex have revealed a set of physical predictive rules for substrate recognition by Kapβ2 (Lee, Cansizoglu et al. 2006), which make it possible for the first time to predict the NLSs for Kapβ2.

Rules for substrate recognition by Kapβ2

Rule 1: NLS Is Structurally Disordered in Substrate. In the structure of Kapβ2-hnRNP A1-NLS, the 26-residue NLS adopts an extended conformation, suggesting that an NLS recognized by Kapβ2 should exist within a stretch of at least 30 residues that lacks secondary structure in its native, unbound state. Thus, the NLS is most likely structurally disordered in the free cargo.

Rule 2: Overall Positive Charge for NLS Is Preferred. The cargo binding site of Kapβ2 is highly acidic and thus favors an NLS with overall positive charges.

Rule 3: Consensus Sequences for the NLS. PY-NLSs share a C-terminal consensus motif R/K/H-$X_{(2-5)}$-P-Y, where X is any residue. PY-NLSs are divided into two subclasses based on their N-terminal conserved regions: 1) The hydrophobic PY-NLS or hPY-NLS with a loose consensus of ϕ-G/A/S-ϕ-ϕ (where ϕ is a hydrophobic side chain) 11-13 residues N-

terminus of the PY motif, and 2) The basic PY-NLS or bPY-NLS which has a basic-enriched region at the N-terminus.

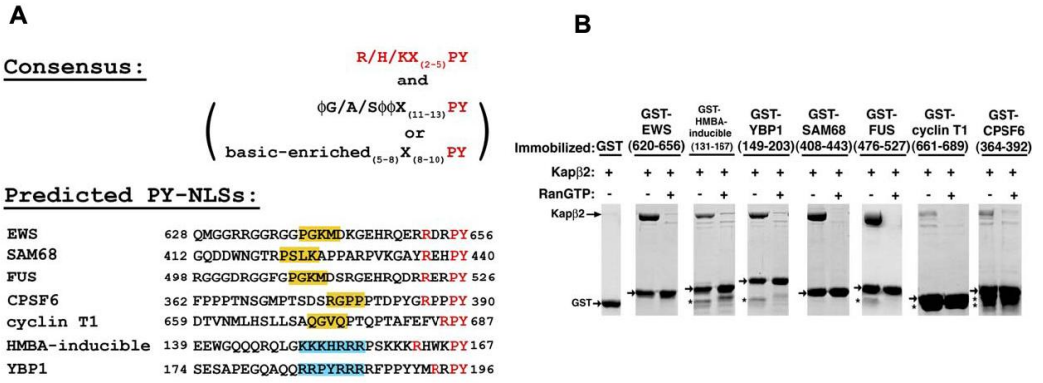

Figure 5-1 Predicted PY-NLSs recognized by Kapβ2. (A) Alignment of predicted NLSs recognized by Kapβ2 at conserved PY residues. NLSs in known Kapβ2 substrates are predicted by the presence of the R/K/H-$X_{(2-5)}$-P-Y C-terminal motifs (red) within structurally disordered and positively charged regions of 30 amino acids. Central hydrophobic motifs ϕG/A/Sϕϕ (ϕ is a hydrophobic side chain) are shaded yellow. Central basic motifs are shaded blue. (B) Binding assays of predicted NLSs from known Kapβ2 substrates EWS, HMBA-inducible protein, YBP1, SAM68, FUS, Cyclin T1 and CPSF6. Kapβ2 is added to immobilized GST-NLSs (arrows) in the presence and absence of excess RanGTP, and bound proteins visualized with Coomassie blue. Asterisks label degraded fragments of substrates. (Lee, Cansizoglu et al. 2006)

The NLS rules are predictive

The C-terminal R/K/H-$X_{(2-5)}$-P-Y consensus within structurally disordered and positively charged regions were found in seven recently identified Kapβ2 cargos: Ewing Sarcoma protein (EWS), hexamethylene bis acetamide (HMBA)-inducible protein, Y-box binding

protein 1 (YBP1), SAM68, FUS, CPSF6, and Cyclin T1 (Guttinger, Muhlhausser et al. 2004)(Figure 5-1A). All seven predicted NLSs bind Kapβ2 and are dissociated from the karyopherin by RanGTP, consistent with NLSs imported by Kapβ2 (Figure 5-1B). Confirmation of these seven NLSs indicates that the three rules for NLS recognition by Kapβ2 described above are predictive.

In the attempt to identify human candidate cargos for Kapβ2, bioinformatic searches were performed by the program ScanProsite (Gattiker, Gasteiger et al. 2002) using motifs ϕ_1-G/A/S-ϕ_3-ϕ_4-X_{7-12}-R/K/H-X_{2-5}-P-Y (where ϕ_1 is strictly hydrophobic, ϕ_3 and ϕ_4 are hydrophobic and also include long aliphatic side chains R and K) and K/R-X_{0-2}-K/R-K/R-X_{3-10}-R/K/H-X_{1-5}-P-Y in the UniProtKB/Swiss-Prot protein database (Bairoch, Boeckmann et al. 2004). All resulting entries were filtered for structural disorder using the program DisEMBL(Linding, Jensen et al. 2003) and for overall positive charges. Eighty-one new candidate cargos were predicted (Tables 5-1 and 5-2). Of the 81 candidate Kapβ2 substrates, 48 contain hPYNLSs (Table 5-1), 28 contain bPY-NLSs (Table 5-2), and 5 contain PY-NLSs with both basic and hydrophobic central motifs. Forty-nine of the new substrates (60%) are involved in transcription or RNA processing, 18 have unknown cellular activity, and the rest are involved in signal transduction (8), cell-cycle regulation (3), and the cytoskeleton (3). Interestingly, information on subcellular localization is available for 62 of the predicted substrates, of which 57 (92%) are annotated to have nuclear localization. Five out of 81 substrates from the lists—protein kinase CLK3 (P49761), transcription factor HCC1 (Q14498), mRNA processing proteins RB15B (Q8NDT2) and SOX14 (O95416), and the Williams-Beuren syndrome

chromosome region 16 protein/WBS16 (Q96I51)—have been tested and bind Kapβ2 in a Ran-dependent manner (Figure 5-2). This chapter describes the validation of the rest predicted cargos for Kapβ2 on the list.

Table 5-1 Predicted Kapβ2 substrates with hPY-NLSs (Lee, Cansizoglu et al. 2006)

Accession number	Name	Localization[a]	N-Term. Residue	Sequences for Candidate Hydrophobic PY-NLS[b]	C-Term. Residue
Q8IZP0	Abl interactor 1	C, N	158	KHGNNQPARTGTLSRTNPPTQKPPSPP MSGRGTLG**R**NT**PY**KTLEPVKPPT	207
Q9UKA4	A-kinase anchor protein 11/AKAP 220	C, Centrosome	385	QRKGHKHGKSCMNPQKFKFDRPALPA NVRKPTP**R**KPES**PY**GNLCDAPDSP	434
P50995	Annexin A11 (Annexin XI) (Calcyclin-associated annexin 50)	C, N	84	PVPPGGFGQPPSAQQPVPPYGMYPP PGGNPPS**R**MPSYP**PY**PGAPVPGQPM	133
Q13625	Apoptosis-stimulating of p53 protein 2	C, N	474	TLRKNQSSEDILRDAQVANKNVAKVP PPVPTKP**KQ**INL**PY**FGQTNQPPSD	523
Q9BXP5	Arsenite-resistance protein 2[c]	not known	53	GEYRDYDRNRRERFSPPRHELSPP QKRMRRDWDE**H**SSD**PY**HSGYEMPYAG	102
Q92560	Ubiquitin carboxyl-terminal hydrolase BAP1(BRCA1-associated protein 1)[c]	N	685	EGMLANLVEQNISVRRRQGVSIGRL HKQRKPDRRK**R**SR**PY**KAKRQ	729
P48634	Large proline-rich protein BAT2 (HLA-B-associated transcript 2)	C, N	690	VPAPQAPPPPPKALYPGALGRPPPM PPMNFDP**R**WMMIP**PY**VDPRLLQGRP	739
O15178	Brachyury protein	N	251	TSTLCPPANPHPQFGGALSLPSTHS CDRYPTLRSH**R**SS**PY**PSPYAHRNNS	300
O60885	Bromodomain-containing protein 4 (HUNK1 protein)	N	1015	QGQQPPHPPPGQQPPPPQPAKPQQV IQHHHSP**R**HHKSD**PY**STGHLREAPSP	1064
Q14004	Cell division cycle 2-like protein kinase 5	not known	376	YERGGDVSPSPYSSSSWRRSRSPYSPVLRRSGKS**R**SRS**PY**SSRHSRSRSR	425
Q9NYV4	Cell division cycle 2-related protein kinase 7	N	256	SSNYDSYKKSPGSTSRRQSVSPPYK EPSAYQSST**R**SPS**PY**SRRQRSVSPY	305
Q5TGI0	Protein C6orf168	not known	94	IDSKDAIILHQFARPNNGVPSLSPF CLKMETYL**R**MADL**PY**QNYFGGKLSA	143
P49761	Dual specificity protein kinase CLK3 (CDC-like kinase 3/Clk3)[c]	N	18	YRWKRRRSYSREHEGRLRYPSRREP PPRRSRS**R**SHDRL**PY**QRRYRERRDS	67
P05997	Collagen alpha-2(V) chain precursor	not known	611	MGLPGPKGSNGDPGKPGEAGNPGVP GQRGAPGKDG**K**VG**PY**GPPGPPGLRG	660
Q03692	Collagen alpha-1(X) chain precursor	not known	84	GYGSPGLQGEPGLPGPPGPSAVGKP GVPGLPG**K**PGERG**PY**GPKGDVGPAG	133
Q8TBR5	Protein C19orf23[c]	not known	70	TWQTRNHTRTGHAYPRFTRPSFPSC NRNGKRRKL**R**LGL**PY**	119
Q96RT6	Protein cTAGE-2	not known	692	PPGTVFGASPDYFSPRDVPGPPRAP FAMRNVYLP**R**GFL**PY**RPPRPAFFPQ	741
Q9NSV4	Protein diaphanous homolog 3 (Diaphanous-related formin-3)	not known	1070	GAAFRDRRKRTPMPKDVRQSLSPM SQRPVLKVCN**H**GNK**PY**L	1110
P56177	Homeobox protein DLX-1	N	44	CLHSAGHSQPDGAYSSASSFSRPLG YPYVNSVSS**H**ASS**PY**ISSVQSYPGS	93
O95147	Dual specificity protein phosphatase 14/MAP kinase phosphatase 6	not known	156	RQLIDYERQLFGKSTVKMVQTPYGIV PDVYEKES**R**HLM**PY**WGI	200
Q9BUP0	EF-hand domain-containing protein 1 (Swiprosin-2)	not known	42	PPARAPTASADAELSAQLSRRLDINE GAARPRRC**R**VFN**PY**TEFPEFSRRL	91
Q6ZV73	FYVE, RhoGEF and PH domain-containing protein 6 (Zinc finger FYVE domain-containing protein 24)	C	269	SSELEALENGKRSTLISSDGVSKKSE VKDLGPLEI**H**LV**PY**TPKFPTPKPR	318
Q92837	Proto-oncogene FRAT1	N	89	PAVPLLLPPALAETVGPAPPGVLRCA LGDRGRVRG**R**AA**PY**CVAELATGPS	138
Q96AE4	FUSE-binding protein 1/DNA helicase V	N	465	PGPHGPPGPPGPGTPMGPYNPAPYNP GPPGPAP**H**GPPA**PY**APQGWGNAYP	514
Q8NEA6	Zinc finger protein GLIS3	N	601	LTAVDAGAERFAPSAPSPHHISPRRV PAPSSILQ**R**TQP**PY**TQQPSGSHLK	650
Q8TEK3	Histone H3-K79 methyltransferase	N	775	SPAKIVLRRHLSQDHTVPGRPAASEL HSRAEHT**K**ENGL**PY**QSPSVPGSMK	824
P35452	Homeobox protein Hox-D12 (Hox-4H)	N	175	AGVASCLRPSLPDGKRCPCSPGRPAVG GGPGEA**R**KKRK**PY**TKQQIAELEN	224
Q13422	DNA-binding protein Ikaros (Lymphoid transcription factor LyF-1)	N	254	CKIGSERSLVLDRLASNVAKRKSSMPQ KFLGD**K**GLSDT**PY**DSSASYEKEN	303
O43474	Kruppel-like factor 4 (Epithelial zinc-finger protein EZF) (Gut-enriched Krueppel-like factor)	N	218	GKFVLKASLSAPGSEYGSPSVISVSKGS PDGS**H**PVVVA**PY**NGGPPRTCPK	267
Q8NEZ4	Histone-lysine N-methyltransferase, H3 lysine-4 specific MLL3	N	2427	NVNQAFTRPPPPYPGNIRSPVAPPLGPR YAVFP**K**DQRG**PY**PPDVASMGMR	2476
Q96G25	Mediator of RNA polymerase II transcription subunit 8 homolog (ARC32)	N	227	GAPSQQQPMLSGVQMAQAGQPGKMPSG IKTNI**K**SASMH**PY**QR	268
Q93074	Mediator of RNA polymerase II transcription subunit 12	N	1854	DLLHHPNPGSITHLNYRQGSIGLYTQN QPLPAGGP**R**VD**PY**RPVRLPMQKL	1903
O43312	Metastasis suppressor protein 1 (Metastasis suppressor YGL-1)	not known	379	LPRVTSVHLPDYAHYYTIGPGMFPSSQ IPSWKDWA**K**PG**PY**DQPLVNTLQR	428
Q13310	Polyadenylate-binding protein 4	C	484	GAAQQGLTDSCQSGGVPTAVQNLAPRA AVAAAAP**R**AVA**PY**KYASSVRSPH	533
Q9Y6V0	Piccolo protein (Aczonin)	C	2874	VVYKLPFGRSCTAQQPATTLPEDRFGYR DDHYQYD**R**SG**PY**GYRGIGGMKP	2923
Q8NFH8	RalBP1-associated Eps domain-containing protein 2 (RalBP1-interacting protein 2)	C	188	PTMSPLASPPSSPPHYQRVPLSHGYSKL RSSAEQM**H**PA**PY**EARQPLVQPE	237
O75177	SS18-like protein 1 (SYT homolog 1)	not known	196	SHYSSAQGGSQHYQGQSSIAMMGQGSQGSSMMGQ**R**PMA**PY**RPSQQGSSQQ	245
Q92922	SWI/SNF complex 155 kDa subunit (BRG1-associated factor 155)	C, N	960	QQQHGQNPQQAHQHSGGPGLAPLGAAGHPGMMP**H**QQPP**PY**PLMHHQMPPP	1009
P09012	U1 small nuclear ribonucleoprotein A (U1 snRNP protein A)	N	123	AVQGGGATPVVGAVQGPVPGMPPMTQAPRIMH**H**MPGQP**PY**MPPPGMIPPP	172
P18583	SON3/Negative regulatory element-binding protein/DBP-5	N	945	GQDPYRLGHDPYRLTPDPYRMSPRPYRI APRSY**R**IAPR**PY**RLAPRPLMLA	994

Table 5-1 Predicted Kapβ2 substrates with hPY-NLSs (continued)(Lee, Cansizoglu et al. 2006)

Accession number	Name	Localization[a]	N-Term. Residue	Sequences for Candidate Hydrophobic PY-NLS[b]	C-Term. Residue
Q8IXZ3	Transcription factor Sp8 (Specificity protein 8)	N	164	GGSSAHSQDGSHQPVFISKVHTSVDGL QGIYP**R**VGMAH**PY**ESWFKPSHPG	213
Q15532	SSXT protein (SYT protein)	not known	214	QYNMPQGGGQHYQGQQPPMGMMGQVNQGNHMMGQ**R**QIP**PY**RPPQQGPPQQ	263
Q9UMS6	Synaptopodin-2 (Myopodin) (Genethonin 2)[c]	C, N	931	PSYPLAALKSQPSAAQPSKMGKKKGKK PLNALDVM**K**HQ**PY**QLNASLFTFQ	980
Q9Y5Q8	General transcription factor 3C polypeptide 5	N	31	GVVRDVAKMLPTLGGEEGVSRIYADPT KRLELYF**R**PKD**PY**CHPVCANRFS	80
Q04206	Transcription factor p65 (Nuclear factor NF-kappa-B p65 subunit)	C, N	310	KSIMKKSPFSGPTDPRPPPRRIAVPSR SSASVP**K**PAPQ**PY**PFTSSLSTIN	359
Q9NRE2	Teashirt homolog 2 (Zinc finger protein 218) (Ovarian cancer-related protein 10-2)	N	558	LPMGSRVLQIRPNLTNKLRPIAPKWKV MPLVSMPT**H**LA**PY**TQVKKESEDK	607
Q9UJT2	Testis-specific serine kinase substrate	not known	275	PAATSQGCPGPPGSPDKPSRPHGLVPA GWGMGP**R**AGEG**PY**VSEQELQKLF	324
Q8TAP9	TTD nonphotosensitive 1 protein	N	15	GPGGGGWGSGSSFRGTPGGGGPRPPSPRDGYGSPH**H**TP**PY**GPRSRPYGSS	64
Q96I51	Williams-Beuren syndrome chromosome region 16 protein (WBS16)	N	62	FVWGFSFSGALGVPSFVVPSSGPGPRAG ARPRR**R**IQPV**PY**RLELDQKISS	111
P19544	Wilms' tumor protein (WT33)	N	94	VHFSGQFTGTAGACRYGPFGPPPPSQAS SGQA**R**MFPNA**PY**LPSCLESQPA	143
P17861	X box-binding protein 1 (XBP-1) (Tax-responsive element-binding protein 5)	N	202	ISCWAFWTTWTQSCSSNALPQSLPAWRS SQRSTQ**K**DPV**PY**QPPFLCQWGR	251
Q8NAP3	Zinc finger and BTB domain-containing protein 38	N	539	HAIDHRLSISKKTANGGLKPSVYPYKLY RLLPMKC**K**RA**PY**KSYRNSSYEN	588
Q9C0A1	Zinc finger homeobox protein 2	N	784	VKPPATATPASLPKFNLLLGKVDDGTGR EAPK**R**EAPAF**PY**PTATLASGPQ	833

[a] As annotated in the UniProtKB/Swiss-Prot entries. C represents cytoplasm and N represents nucleus.
[b] Central hydrophobic motifs are underlined and the R/K/H-PY motifs are in bold.
[c] Substrates also identified using bPY-NLS motif.

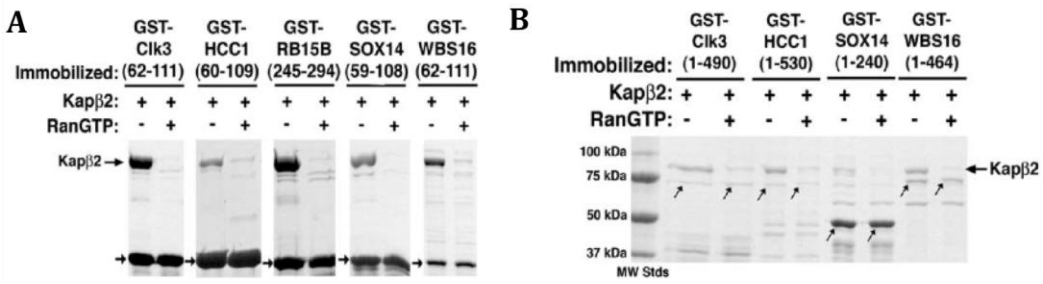

Figure 5-2 Five predicted Kapβ2 substrates (Clk3, HCC1, RB15B, Sox14, and WBS16) are validated experimentally. (A) Binding assays of GST-NLSs (arrows); (B) Binding assays of full-length substrates Clk3,HCC1, Sox14, andWBS16 to Kapβ2. Expression of recombinant full-length RB15B was not successful. Coomasie-stained

bands at the size of the GST substrates are labeled with arrows.Lower-molecular-weight proteins are likely degraded substrates. (Lee, Cansizoglu et al. 2006)

Table 5-2 Predicted Kapβ2 substrates with bPY-NLSs (Lee, Cansizoglu et al. 2006)

Accession number	Name	Localization[a]	N-Term. Residue	Sequences for Candidate Hydrophobic PY-NLS[b]	C-Term. Residue
Q13023	A-kinase anchor protein 6 (AKAP 100)	not available	1851	GSVKRVSENNGNGKNSSHTHELGT KRENKKTIFKVN**K**D**PY**VADMENGNIE	1900
Q9BXP5	Arsenite-resistance protein 2[c]	not available	61	NRRERFSPPRHELSPPQKRMRRDWD EHSSDPY**H**SGYEM**PY**AGGGGGPTYG	110
Q92560	BRCA1-associated protein 1[c]	N	685	EGMLANLVEQNISVRRRQGVSIG RLHKQRKPDRRK**R**SR**PY**KAKRQ	729
Q9NYF8	Bcl-2-associated transcription factor 1	C, N	32	KRYSSRSRSRTYSRSRSRDRMYSRD YRRDYRNNRGM**R**R**PY**GYRGRGRGYY	81
Q9ULD4	Bromodomain and PHD finger-containing protein 3	not available	1	MRKPRRKSRQNAEGR**R**SPS**PY**SLKC SPTRET	31
Q9UK58	Cyclin-L1	N	337	ASKPSSPREVKAEEKSPISINVKTVKK EPEDRQQAS**K**S**PY**NGVRKDSKRS	386
Q9NYF5	Protein C5orf5 (GAP-like protein N61)	not available	531	QRFLHDPEKLDSSSKALSFT RIRRSSFSSKDEKRED**R**T**PY**QLVKKLQKKI	580
P49761	CDC-like kinase 3[c]	N	62	RERRDSDTYRCEERSPSFGEDYYGPS RSRHRRRSRE**R**G**PY**RTRKHAHHCH	111
Q8TBR5	Protein C19orf23[c]	not available	70	TWQTRNHTRTGHAYPRFTRPSFPS CNRNGKRRKL**R**LGL**PY**	109
Q92782	Zinc-finger protein neuro-d4	C, N	156	EDLEDDIPRRKNRAKGKAYGIGGLRKR QDTASLED**R**DK**PY**VCDKFYKELA	205
O00358	Forkhead box protein E1/Thyroid transcription factor 2	N	17	TVKEERGETAAGAGVPGEATGRGAGG RRRKRPLQ**R**GKP**PY**SYIALIAMAI	66
Q13461	Forkhead box protein E3 (FKHL12) (Forkhead-related transcription factor 8)	N	35	AEPGREPEEAAAGRGEAAPTPAPGPG RRRRRPLQ**R**GKP**PY**SYIALIAMAL	84
O75593	Forkhead box protein F1	N	1	MDPASSGPSKAKKTNAGIR**R**PEKP**PY**SYIALIVMAI	36
O75593	Forkhead box protein H1/Forkhead activin signal transducer 1	N	1	MGPCSGSRLGPPEAESPSQPPKRRKKR YL**R**HDKP**PY**TYLAMIALVI	46
Q9UPW0	Forkhead box protein J3	N	142	SKDDPGKGSYWAIDTNPKEDALPT RPKKRARSVE**R**AST**PY**SIDSDSLGME	191
P55317	Hepatocyte nuclear factor 3-alpha (Forkhead box protein A1).	N	135	MNPCMSPMAYAPSNLGRSRAGGGGDAKTFKRSYPHA**K**P**PY**SYISLITMAI	184
P55318	Hepatocyte nuclear factor 3-gamma (Forkhead box protein A3)	N	81	LGVSGGSSSSGYGAPGPGLVHGKEMP KGYRRPLAHA**K**P**PY**SYISLITMAI	130
Q9Y483	Metal-response element-binding transcription factor 2	N	370	HEFKIKGRKASKPISDSREVSNGIE KKGKKKSVG**R**PPG**PY**TRKMIQKTAE	419
O95644	NFAT transcription complex cytosolic component	C, N	238	PSTSPRASVTEESWLGARSSRPASP CNKRKYSLNG**R**QP**PY**SPHHSPTPSP	287
Q9ULL1	Pleckstrin homology domain-containing family G member 1	not available	1304	SKFVDADFSDNVCSGNTLHSLNSP RTPKKPVNS**K**LGLS**PY**LTPYNDSDKL	1353
Q99575	Ribonucleases P/MRP protein subunit POP1	N	372	QTELPDEKIGKKRKRKDDGENAKPIKK IIGDGT**R**DPCL**PY**SWISPTTGII	421
Q8NEY8	Periphilin 1/Gastric cancer antigen Ga50	C, N	84	YRWTRDDHSASRQPEYRDMRDGFRRKS FYSSHYARE**R**S**PY**KRDNTFFRES	133
Q8NDT2	RNA-binding protein 15B	N	245	SRSGERWGADGDRGLPKPWEERRKRR SLSSDRG**R**TTHS**PY**EERSRTKGSG	294
Q14498	Splicing factor HCC1	N	60	DRERKKSKSRERKRSRSKERRRSRSRSRDRRFRG**R**YRS**PY**SGPKFNSAIR	109
P62241	40S ribosomal protein S8	N	1	GISRDNWHKRRKTGGK**R**K**PY**HKKRKYELGR	30
O95416	Transcription factor SOX-14	N	59	DEAKRLRAQHMKEHPDYKYRPRRKPKNLLKKD**R**YVFPL**PY**LGDTDPLKAA	108
Q9Y651	Transcription factor SOX-21 (SOX-A)	N	59	DEAKRLRAMHMKEHPDYKYRPRRKPKTLLKKD**K**FAFPV**PY**GLGGVADAEH	108
O00267	Transcription elongation factor SPT5	N	678	GGQRGGFGSPGGGSGGMSRGRGRRDNELIGQTV**R**ISQG**PY**KGYIGVVKDA	727
Q9UMS6	Synaptopodin-2 (Myopodin) (Genethonin 2)[c]	C, N	931	PSYPLAALKSQPSAAQPSKMGKKKGKKPLNALDVM**K**HQ**PY**QLNASLFTFQ	980
Q8IWR0	Zinc finger CCCH-type domain-containing protein 7A	N	464	ANIDHKCKKDILIGRIKNVEDKSWKKIRPRPT**K**TNYEG**PY**YICKDVAAEE	513
Q9H091	Zinc finger MYND domain-containing protein 15	not available	522	RDSLEVSVRPGSGISARPSSGTKEKGGRRDLQI**K**VSAR**PY**HLFQGPKPDL	571
Q9H116	Zinc finger protein 336	N	177	LTDSLDYPGERASNGMSSDLPP KKSKDKLDKKKEVV**K**P**PY**PKIRRASGRL	226
Q8N895	Zinc finger protein 366	N	49	RGPFSQFRYEPPPGDLDGFPGVFEGAGS RKRKSMPT**K**M**PY**NHPAEEVTLA	98

[a] As annotated in the UniProtKB/Swiss-Prot entries. C represents cytoplasm and N represents nucleus.
[b] Central basic-enriched regions are underlined and the R/K/H-PY motifs are in bold.
[c] Substrates also identified using hPY-NLS motif.

Materials and Methods

Cloning and protein purification

The PY-NLS fragments were amplified by PCR from human brain cDNA library (BD Biosciences, MD, USA) or annealed as synthetic oligos, cloned into pGEXTEV vector, and expressed as GST fusion proteins. Mutations were generated using Quikchange site-directed mutagenesis kit. For immunofluerescence study, the PY-NLSs were subcloned into pFLAG-CMV2 vector with a human pyruvate kinase gene at their C-terminus. The correct inserts were confirmed by DNA sequencing.

In vitro binding assays

The in vitro binding assays and Ran dissociation assays were done similarly as described in Chapter 3. Approximately 20-40 μg of GST-PY-NLSs were immobilized on glutathione sepharose (GE Healthcare, NJ, USA). 20 μg of Kapβ2 was added to the peptide bound sepharose for 10 minutes followed by extensive washing (TB Buffer: 20 mM HEPES pH7.3, 110 mM KAc, 2 mM DTT, 2 mM MgAc, 1 mM EGTA and 20% Glycerol). A second incubation was done with 40 μl of RanGTP (2.8 mg/ml) or 40 μl of MBP-M9M (3 mg/ml). After extensive washing, a quarter of the bound proteins were analyzed by SDS-PAGE and visualized with Coomassie staining.

ITC

ITC experiments were done similarly as decribed in Chapter 3. Binding affinities of wild type MBP-Sam68-NLS or MBP-HuR-NLS to Kapβ2 were quantitated using ITC. The ITC experiments were done using a MicroCal Omega VP-ITC calorimeter (MicroCal Inc., Northampton, MA). Proteins were dialyzed against buffer containing 20 mM Tris pH 7.5, 100 mM NaCl and 2 mM ®-mercaptoethanol. 100-500 ⌈M MBP-NLS proteins were titrated into a sample cell containing 10-100 ⌈M full-length Kap®2. Most ITC experiments were done at 20° C with 35 rounds of 8 ⌈l injections. Data was plotted and analyzed using MicroCal Origin software version 7.0, with a single binding site model.

Results and Discussion

Validation of predicted PY-NLSs

In previous studies, five predicted PY-NLSs, CLK3, HCC1, RB15B, SOX14 and WBS16, were shown to bind Kapβ2 in a Ran dependent manner. (Figures 5-2D and E). In order to test the remaining 76 putative PY-NLSs in Tables 5-1 and 5-2, I cloned the fragments either by PCR or using synthetic oligonucleotides into the pGEXTEV vector. I successfully got 72 new constructs. Cloning of the FGD6, MTF2, PPHLN, and SOX21 fragments were unsuccessful. The binding assays that I performed using the immobilized GST-PY-NLSs showed that six hPY-NLSs (BRAC, CDK12, CDK13, MED8, SON, ZBT38) and seven bPY-NLSs (BRPF3, CLK3, FA13B, KHDR3, NFAC1, PABP2, RB15B) bound strongly to Kapβ2 (Figure 5-3 and 4). Three hPY-NLSs (EFHD1,

TTDN1, WBS16) and five bPY-NLSs (BAP1, GZF1, HCC1, SOX14, RS8) showed moderate binding (Figure 5-5) including the five previous tested PY-NLSs (CLK3, HCC1, RB15B, SOX14, WBS16, Lee, 2006). All of the bound PY-NLSs are dissociated from Kapβ2 by RanGTP or by the Kapβ2 inhibitor MBP-M9M (Figure 5-3, 4, 5 and 6; Table 5-4). Twelve of the 77 putative PY-NLSs bound Kapβ2 weakly and 44 showed no binding (Figure 5-7 and 8). So far, I have tested all 81 predicted PY-NLSs in Tables 6-1 and 6-2, and found that 21 of them bind Kapβ2 and are dissociated by RanGTP, thus behaving like Kapβ2 cargos (Table 5-3).

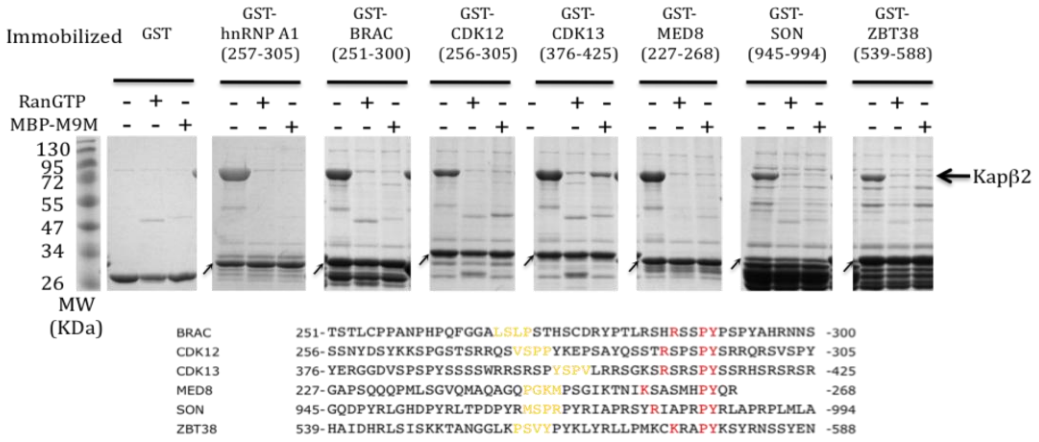

Figure 5-3 Predicted hPY-NLSs show strong binding to Kapβ2. Immobilized GST-NLSs were first incubated with Kapb2, then with RanGTP or MBP-M9M. The bound proteins were resolved on SDS-PAGE gels and visualized by Coomassie blue staining. Arrows, GST-NLSs .The NLS sequences are shown under the gels.Yellow, predicted hydrophobic motif; red, predicted $RX_{(2-5)}PY$ motif**.**

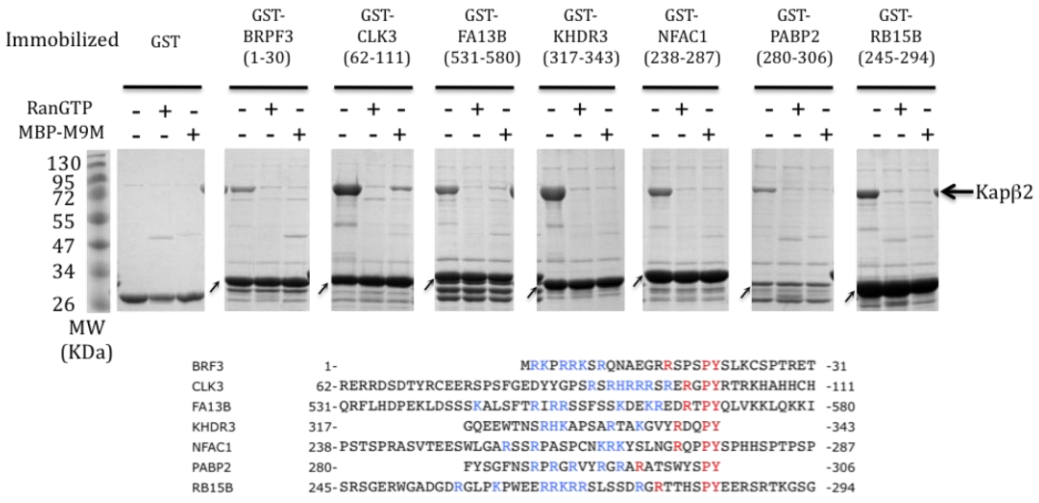

Figure 5-4 Predicted bPY-NLSs show strong binding to Kapβ2. Immobilized GST-NLSs were first incubated with Kapβ2, then with RanGTP or MBP-M9M. The bound proteins were resolved on SDS-PAGE gels and visualized by Coomassie blue staining. Arrows, GST-NLSs .The NLS sequences are shown under the gels. Blue, predicted basic-enriched motif; red, predicted $RX_{(2-5)}PY$ motif**.**

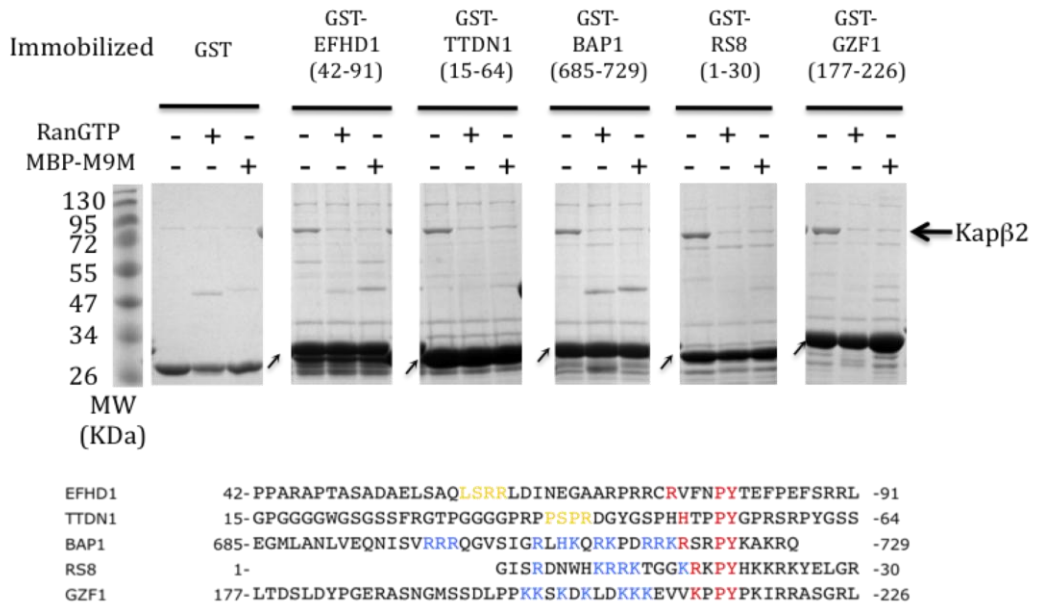

Figure 5-5 Predicted hPY-NLSs and bPY-NLSs show moderate binding to Kapβ2. Immobilized GST-NLSs were first incubated with Kapβ2, then with RanGTP or MBP-M9M. The bound proteins were resolved on SDS-PAGE gels and visualized by Coomassie blue staining. Arrows, GST-NLSs. The NLS sequences are shown under the gels. Yellow, predicted hydrophobic motif; blue, predicted basic-enriched motif; red, predicted $RX_{(2-5)}PY$ motif**.**

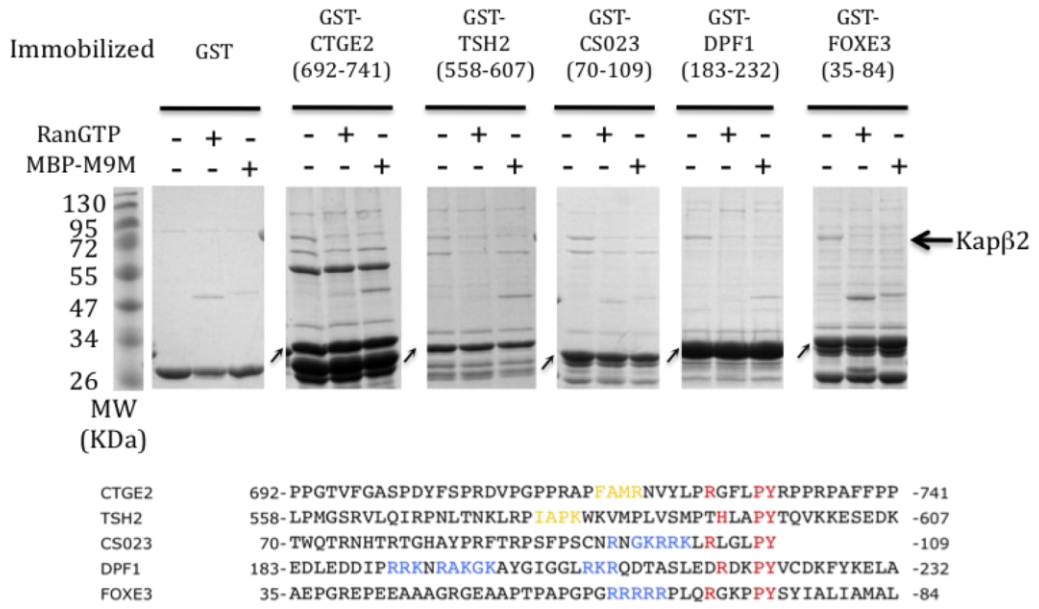

Figure 5-6 Examples of predicted hPY-NLSs and bPY-NLSs show weak binding to Kapβ2. Arrows, GST-NLSs. The NLS sequences are shown under the gels.

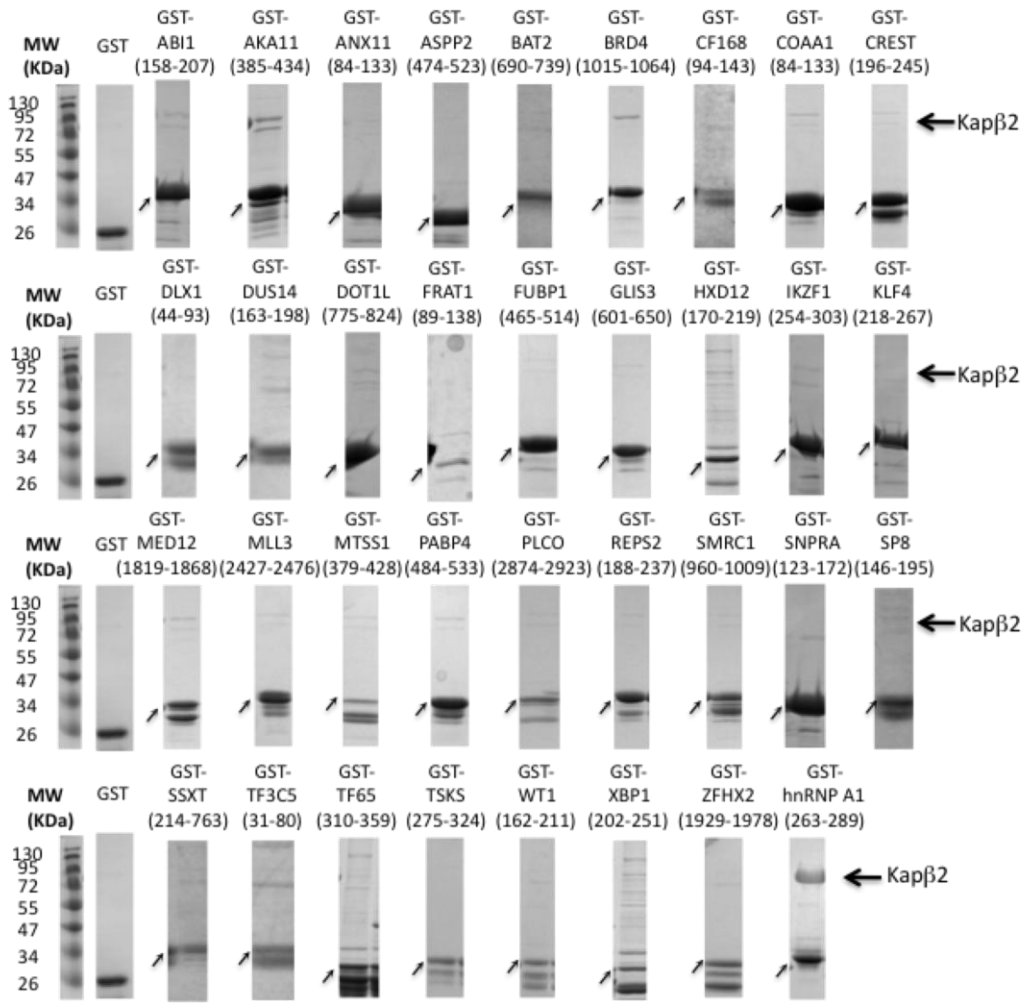

Figure 5-7 Predicted hPY-NLSs show weak or no binding to Kapβ2. Immobilized GST-NLSs were first incubated with Kapβ2, then with RanGTP or MBP-M9M. The bound proteins were resolved on SDS-PAGE gels and visualized by Coomassie blue staining. Arrows, GST-NLSs .The NLS sequences are shown on next page.

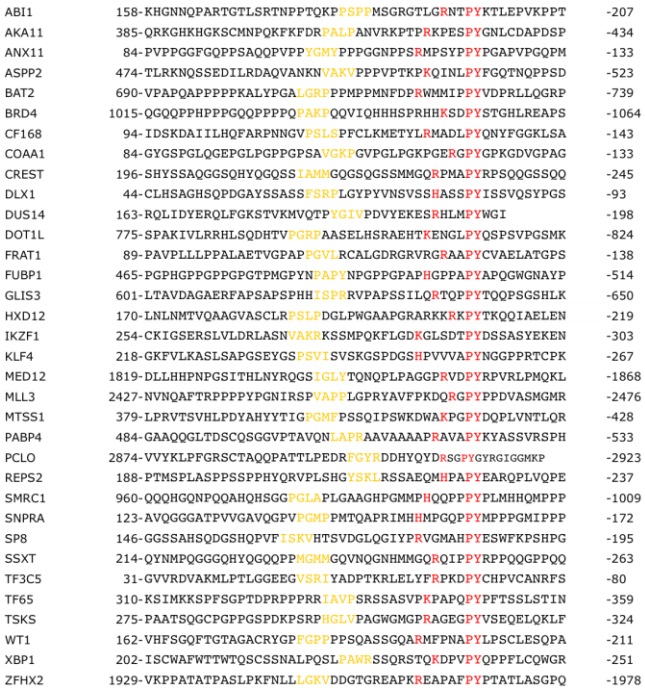

Figure 5-7 (Continued). The hPY-NLS sequences have weak or no binding to Kapβ2. Yellow, predicted hydrophobic motif; red, predicted $RX_{(2-5)}PY$ motif**.**

Table 5-3 Summary of binding assays of predicted PY-NLSs

	Strong	Moderate	Weak or no binding		
hPY-NLSs	BRAC	EFHD1	ABI1	DOT1L	TSH2
	CDK12	TTDN1	AKA11	FRAT	REPS2
	CDK13	WBS16	ANX11	FUBP1	SMRC1
	MED8		ASPP2	GLIS3	SNPRA
	SON		BAT2	HXD12	SP8
	ZBT38		BRD4	IKZF1	SSXT
			CF168	KLF4	TF3C5
			COAA1	MED12	TF65
			CTGE2	MLL3	TSKS
			CREST	MTSS1	WT1
			DLX1	PABP4	XBP1
			DUS14	PCLO	ZFHX2
bPY-NLSs	BRPF3	BAP1	AKAP6	FOXE3	
	CLK3	GZF1	ARS	FOXH1	
	FA13B	HCC1	BCLF1	FOXJ3	
	KHDR3	SOX14	CCLN1	PKAG1	
	NFAC1	RS8	CS023	POP1	
	PABP2		DPF1	SPT5H	
	RB15B		FOXA1	SYNP2	
			FOXA3	Z3H7A	
			FOXE1	ZN366	
				ZN655	

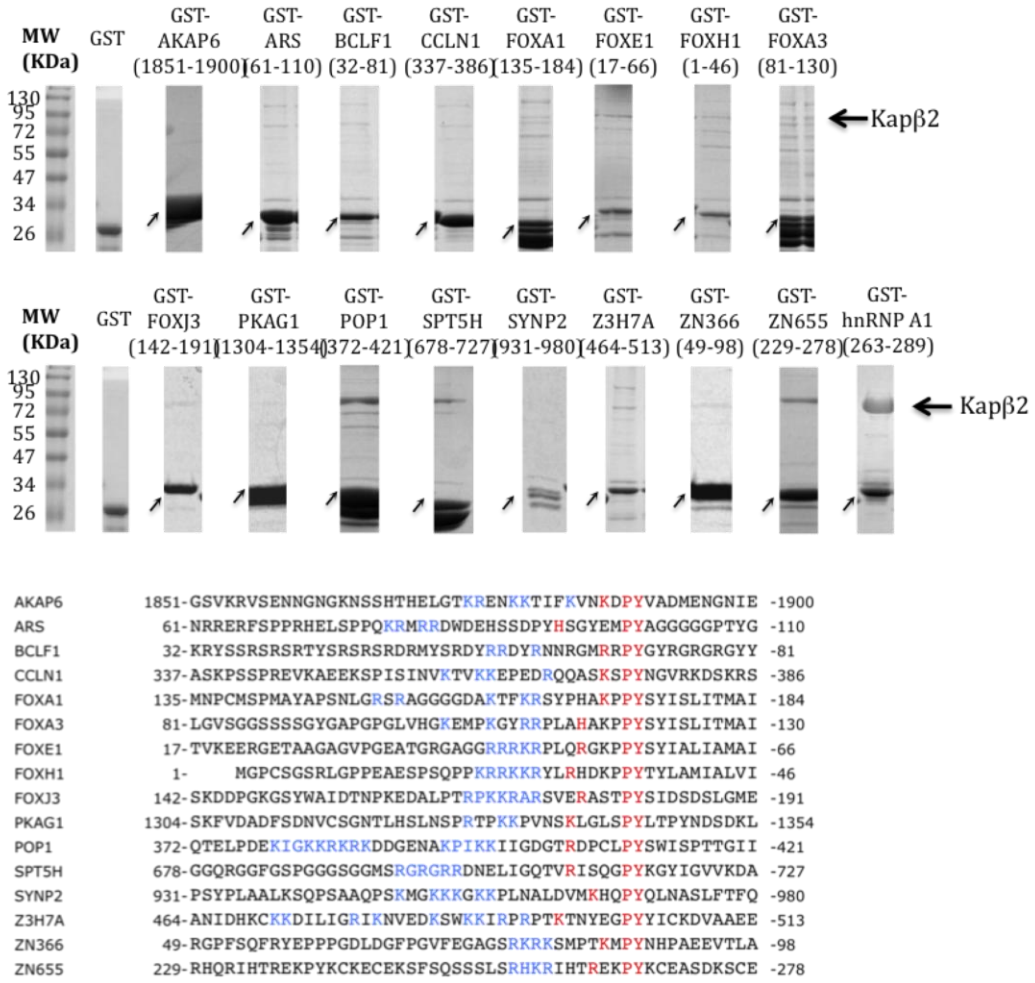

Figure 5-8 Predicted bPY-NLSs show weak or no binding to Kapβ2. Immobilized GST-NLSs were first incubated with Kapβ2, then with RanGTP or MBP-M9M. The bound proteins were resolved on SDS-PAGE gels and visualized by Coomassie blue staining. Arrows, GST-NLSs .The NLS sequences are shown under the gels. Blue, predicted basic-enriched motif; red, predicted $RX_{(2-5)}PY$ motif.

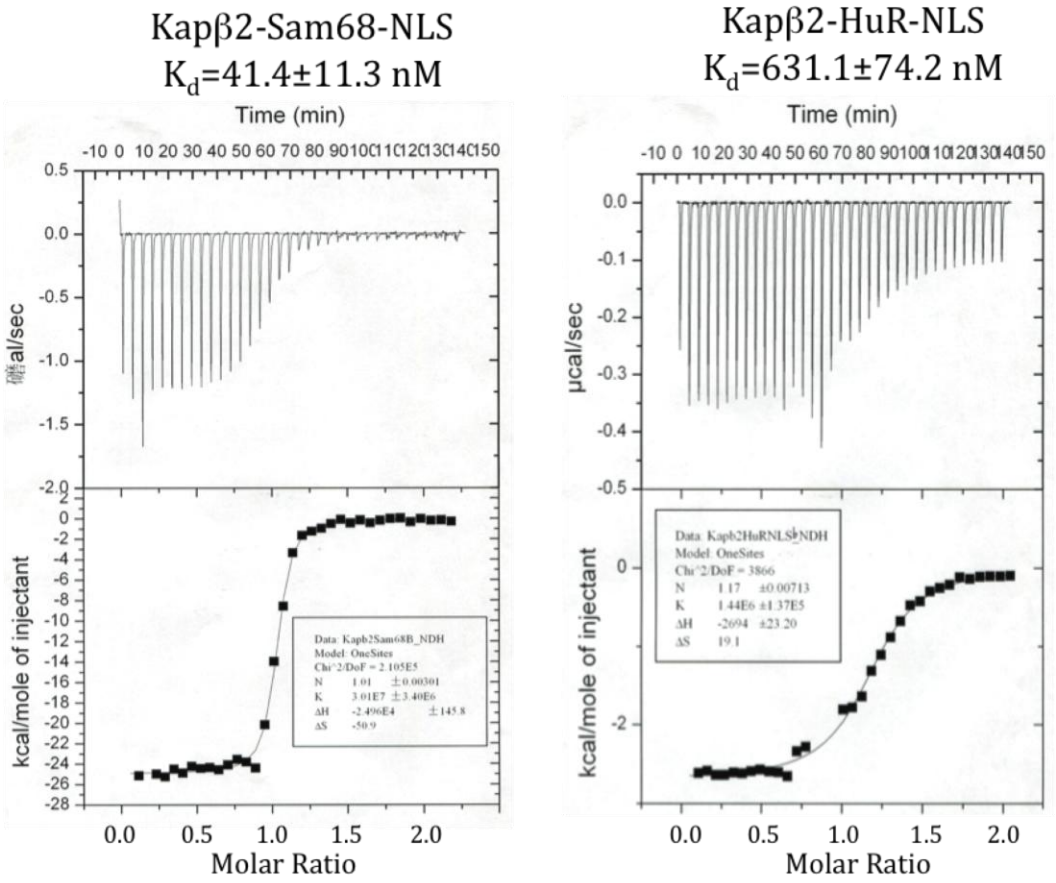

Figure 5-9 ITC profiles of MBP-Sam68-NLS and MBP-HuR-NLS with full-length Kapβ2. Nonlinear least squares fits to the single binding site model were used to fit the ITC profiles (closed squares).

PY motifs of 7 PY-NLSs are critical for Kapβ2 binding

The R-X_{2-5}-P-Y consensus motif appears to be conserved among the 15 Kapβ2 cargos that were experimentally identified (Lee, Cansizoglu et al. 2006). Structural analysis of Kapβ2-PY-NLS complexes also explained the importance of this motif for interactions

with Kapβ2 (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007; Imasaki, Shimizu et al. 2007). One exception is the cargo HuR, which has a "PG" motif instead of "PY" at its C-terminus. I have analyzed the interactions of the HuR PY-NLS with Kapβ2 by ITC and determined the K_D to be 631 nM. I have also measured the affinity of Kapβ2 binding to the Sam68 PY-NLS and the K_D is 41 nM (Figure 5-9). Other known PY-NLSs also have low-nanomolar affinity (hnRNP A1, Kd~40nM, hnRNP M, Kd~ 10 nM, NXF1, 46 nM) (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007, Chapter 3). The lower affinity of HuR may suggest that "PY" is more favored at this position. Alternatively, like NXF1 (Chapter 4), the PY-NLS in HuR may be significantly longer and we may not have located all its binding determinants for Kapβ2. Future studies including replacement of the "PG" in the HuR PY-NLS with a more typical "PY" motif. More thorough mapping of the HuR NLS will also be necessary to resolve these mechanistic questions.

To test the energetic contribution of the PY motif to Kapβ2 binding, I mutated the PY motifs in the PY-NLSs of CLK3, RB15B, SOX14, RS8, GZF1, EFHD1 and TTDN1 to alanines. The PY to AA mutants showed significantly decreased binding to Kapβ2, suggesting that the PY motifs in these PY-NLSs are key Kapβ2-binding epitopes (Figure 5-10).

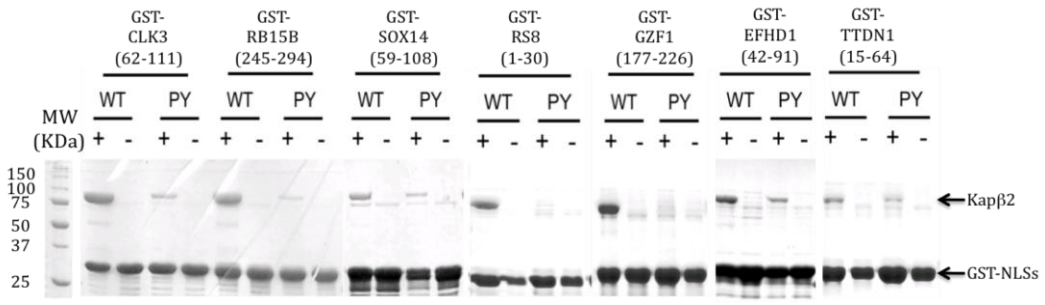

Figure 5-10 PY mutants show disrupted binding to Kapβ2. The PY motifs in the NLSs were mutated into alanines. Immobilized wild type (WT) and mutants (PY) were incubated with purified recombinant Kapβ2. The bound proteins were resolved on SDS-PAGE gels and visualized by Coomassie blue staining.

Conclusions

An initial bioinformatics application of the PY-NLS recognition rules to the Swissprot database led to the prediction of 81 new PY-NLSs (Lee, Cansizoglu et al. 2006). I have tested a 77 of the 81 putative PY-NLSs for Kapβ2-binding. 13 showed strong Ran-sensitive binding to Kapβ2, 8 showed moderate binding to Kapβ2 and 56 showed very weak or no Kapβ2-binding. Although the positive rate is disappointing, this approach of in vitro validation present limitations, which needs to be further addressed before conclusions can be drawn. PY-NLSs are structurally disordered peptides that are prone to proteolysis. Severe degradation was observed in numerous GST-PY-NLSs and proteolysis in others may not be detectable. Degradation may hinder the binding of Kapβ2. PY-NLSs that showed no Kapβ2-bindin will have to be tested by mass spectrometry to assess for the extent of degradation. Lengths of the NLSs tested are also

of concern. All the PY-NLSs tested are about 30-50 residues long. However, we have recently learned that some PY-NLSs may be significantly longer, such as the 109-residues PY-NLS of NXF1 that I discovered. Increasing the length of the PY-NLS peptides may improve Kapβ2-binding. The PY-NLSs that I have validated here for Kapβ2-binding will also need to be tested in cells for their ability to target a heterologous protein to the nucleus. The PY-NLS-containing cargos will also need to be tested in cells using Kapβ2-specific inhibitor M9M. The information gathered from this in vitro validation study will be very useful to modify and improve cargo recognition rules for Kapβ2.

CHAPTER SIX

NUCLEAR IMPORT MEDIATED BY TRANSPORTIN-SR AND IMPORTIN-5

Abstract

Numerous cargos have been identified for Transportin-SR (Trn-SR) and Importin-5 (Imp5), making them ideal to study the Kap-cargo recognition process in order to discover new classes of NLSs. In this chapter, preliminary protein purification and cargo cloning for the crystallographic studies of Trn-SR and Imp5 were performed. Both Trn-SR and Imp5 behaved well as recombinant proteins during purification and the interaction between Imp5 and its cargo p35 was also verified. These preliminary data suggest that Trn-SR and Imp5 are good candidates for structural studies to elucidate the mechanisms of Kap-cargo recognition.

Introduction

Transportin-SR

Transportin-SR (Trn-SR, also known as Transportin-3 or Trn-3 or TNPO3) is the homolog of Kap111p (also known as Mtr10p) in human. There are two splicing variants, Trn-SR and Trn-SR2 (Figure 6-1A). Like Kapβ2, it imports many RNA binding proteins, especially splicing factors. Cargos of Trn-SR proteins include human SR proteins ASF/SF2, SC35, TRA2-alpha, TRA2-beta and drosophila splicing factors 9G8, Rbp1 and

RSF1(Kataoka, Bachorik et al. 1999; Lai, Lin et al. 2000; Lai, Lin et al. 2001; Allemand, Dokudovskaya et al. 2002).These SR proteins usually contain one or two RNA binding domains (RBDs) and a C-terminal RS domain (Zahler, Lane et al. 1992). The latter domain is at least 50-residue long and composed of many arginine-serine dipeptide repeats. They often contain multiple phosphorylation sites and are actively regulated during RNA biogenesis (Figure 6-1B). Trn-SRs bind SR proteins through their RS domains (Kataoka, Bachorik et al. 1999; Lai, Lin et al. 2000; Lai, Lin et al. 2001; Allemand, Dokudovskaya et al. 2002; Yun, Velazquez-Dones et al. 2003). In certain cases, Trn-SRs only recognize phosphorylated cargos. For instance, both Trn-SRs bind only phosphorylated ASF/SF2 and Trn-SR2 also imports TRA2-beta in a phosphorylation-dependent manner (Yun, Velazquez-Dones et al. 2003). However, phosphorylation is not required for the recognition of TRA2-alpha by both Trn-SRs and TRA2-beta by Trn-SR. RS domains have low complexity sequences, and are likely to be structurally disordered (Haynes and Iakoucheva 2006). Thus they may present the third class of linear NLS. In fact, the first four RS dipeptides of ASF/SF2 bind the kinase SRPK1 in extended conformation (Ngo, Giang et al. 2008). However, molecular dynamics simulations predicted unphosphorylated RS repeats are likely helical whereas the phosphorylated repeats may change to extended or to helical-strand conformations (Hamelberg, Shen et al. 2007). Structures of phosphorylated and unphosphorylated RS repeats bound to Trn-SR and Trn-SR2 will be important to understand how Trn-SR recognizes different cargos.

Besides SR proteins, Trn-SR2 also imports non-RS containing splicing regulator RBM4. It interacts with Trn-SR2 via its C-terminal alanine-rich (CAD) domain (Lai, Kuo et al. 2003). Trn-SR2 may also mediate the nuclear import of the HIV-1 preintegration complex (PIC) (Brass, Dykxhoorn et al. 2008).

Importin-5

Importin-5 (Imp5, also known as Kapβ3 or RanBP5) imports ribosomal proteins, core histones and numerous proteins of other functions into the nucleus (Yaseen and Blobel 1997; Jäkel and Görlich 1998; Baake, Bauerle et al. 2001; Mühlhäusser, Müller et al. 2001). Twenty different cargos have been identified for Importin-5. Many of these cargos are also imported by other Kapβs (Chook and Suel 2010). One of Imp5's cargos is the CDK5 activator p35, which is mostly found in neurons and muscle cells. The Cdk5-p35 complex functions in cytoskeletal dynamics, cell adhesion, axonal guidance, cell signaling and synaptic plasticity (Dhavan and Tsai 2001; Lim, Qu et al. 2003). A fraction of CDK5-p35 complex was found in the nucleus (Ino and Chiba 1996; Nikolic, Dudek et al. 1996; Qu, Li et al. 2002; Gong, Tang et al. 2003), which is important for its function. Imp5, Impβ and Imp7 have been identified as the import receptors for p35 and they interact with the residues 31-98 of p35 (Fu, Choi et al. 2006).

In this chapter, I describe preliminary protein purifications for Transportin-SR and Importin-5. Several cargos were also cloned to test interactions with the Kaps and map the NLSs. These studies are preparation for the long-term crystallographic studies of these two import pathways

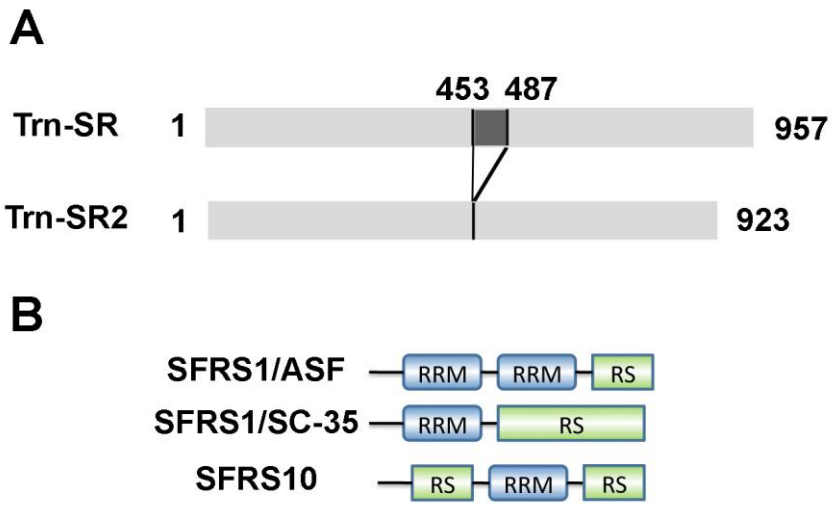

Figure 6-1 Schematic representation of Trn-SRs and their cargos. (A) Human Trn-SR and Trn-SR2. Dark grey block represents the insert in Trn-SR due to alternative splicing. (B) Schematic domain organizations of human SFRS1, SFRS1 and SFRS10.

Materials and Methods

Constructs and protein expression

Human Trn-SR (Access number: NM_012470.3) and Imp5 (Access number: NM_002271) were cloned into pGexTEV vector by former technician Alex D'Brot. Mouse p35 cDNA was a gift from Dr. James Bibbs (UT Southwestern). The full-length and putative NLS (residue 31-98) of p35 were subcloned into pGexTEV. Human full-

length Ran in the pET15b vectort was constructed by former technician Tom Louis. Proteins were expressed at 25°C for 16 hours, otherwise indicated in the figures.

Protein purification

For human Trn-SR, the cells were lysed in Tris300 buffer (50 mM Tris pH7.5, 2 mM EDTA, 2 mM DTT, 20%glycerol, 300mM NaCl) with P.I. (100 μg/ml of Pefabloc, 157 μg/ml of benzaimidine, 50 μg/ml of leupeptin) by cell disruptor and the clarified supernatants were loaded onto 20 ml GSH beads (GE Healthcare, NJ, USA). The unbound proteins were washed out by 75 ml of Tris300 buffer for 3 times, 20 ml of ATP buffer (50 mM Tris pH7.5, 100 mM NaCl, 10 mM MgAc, 20% glycerol, 2 mM EGTA, 2 mM DTT, 5 mM ATP) for 5 times at RT and then 15 ml of Tris20 buffer (same as Tris 300 except NaCl 20 mM). Bound GST-Trn-SR was cleaved with TEV in 15 ml of Tris20 buffer at 4°C for overnight and then eluted with 15 ml of Imidazole buffer (20 mM Imidazole pH6.5, 2mM EDTA, 2 mM DTT, 20%glycerol, 20 mM NaCl) for 3 times. The eluates were concentrated and subjected to 1ml Hitap Q column (GE Healthcare, NJ, USA)) (Figure 6-2) followed by Superdex 200 (GE Healthcare, NJ, USA). Proteins were concentrated, flash frozen and stored at -80°C.

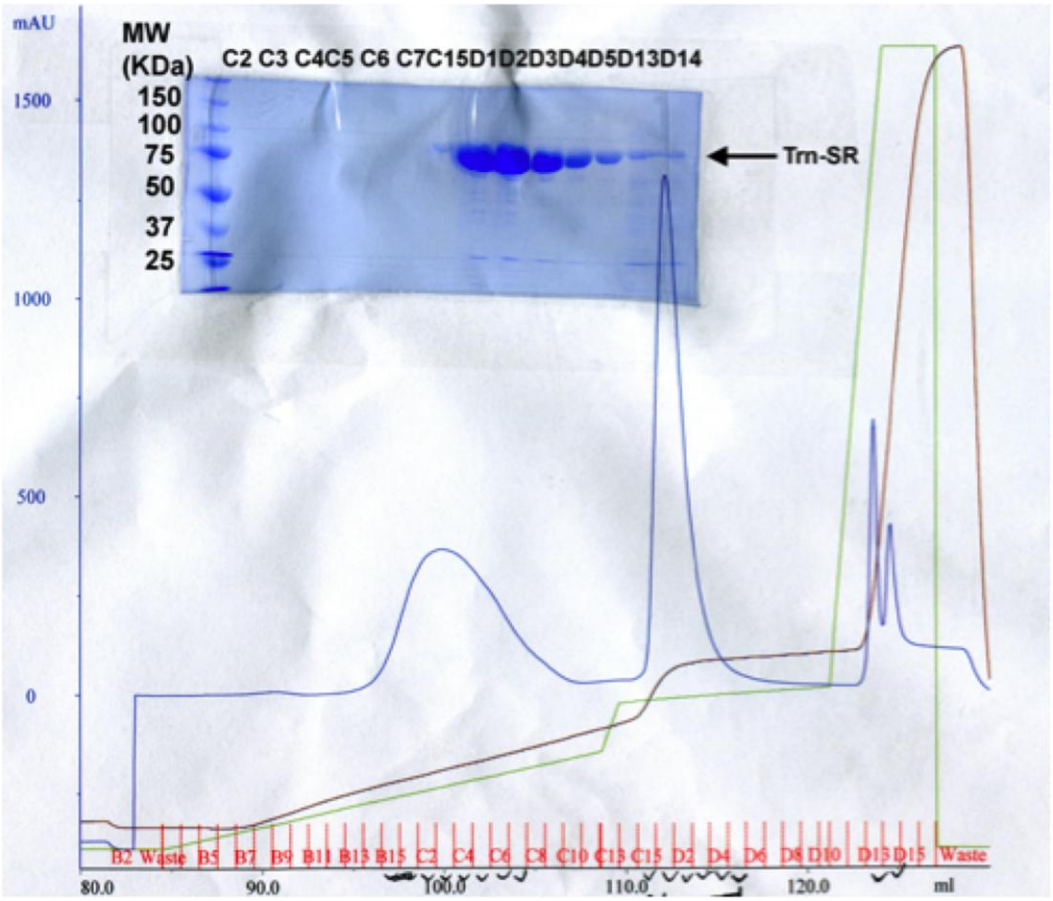

Figure 6-2 Purification of Transportin-SR. The chromatograph of Transportin-SR on 1 ml Hitrap Q column after purified by glutathione affinity column and cleaved by TEV overnight. Gel samples from the indicated fractions were run on 12% SDS-PAGE gel**.**

Human Ran was expressed as a His-tag fusion protein. Cells were lysed in Tris buffer (50 mM Tris pH7.5, 2mM MgAC, 2 mM β-mercaptoethanol, 2.5%glycerol, 500mM NaCl) with 20 mM Imidazole and the supernatants were loaded onto 1 ml Histrap column (GE Healthcare, NJ, USA) and eluted with imidazole gradiant (20 mM- 250 mM) (Figure 6-3). The collected fractions were desalted and subjected to 5 ml Hitrap SP column (Figuer 7-4). Proteins were concentrated, flash frozen and stored at -80°C.

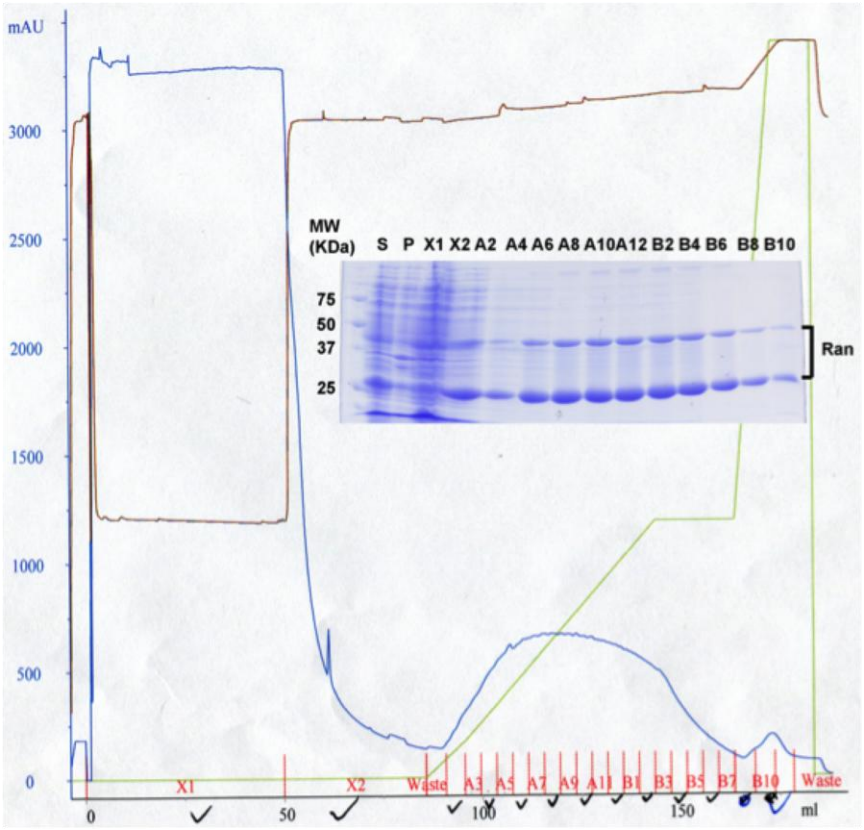

Figure 6-3 Affinity purification of human full-length Ran. The chromatograph of Ran on 5ml HisTrap column. Gel samples were run on 12% SDS-PAGE. S, supernatant; P, pellet.

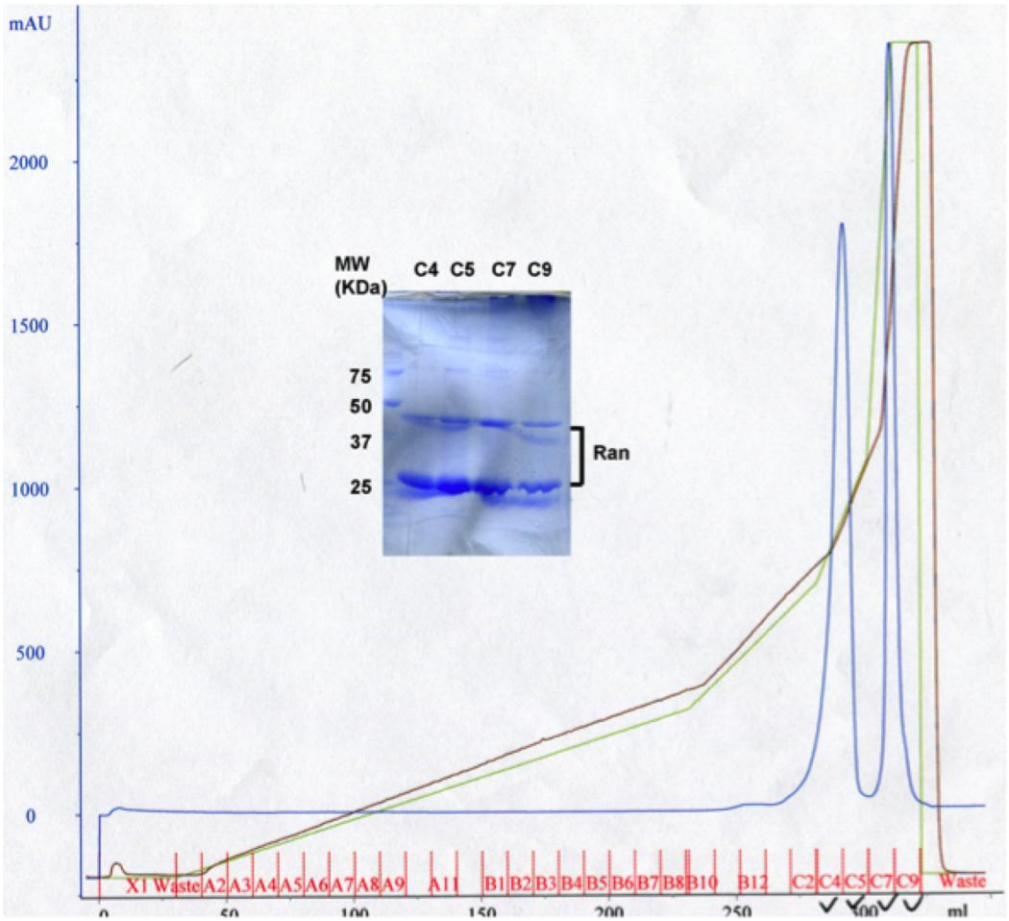

Figure 6-4 Ion exchange purification of human full-length Ran. The chromatograph of Ran on 5 ml HiTrap SP column. Gel samples were run on 12% SDS-PAGE gel.

Human Imp5 was expressed in E. coli as a GST fusion protein. In order to optimize the purification conditions of Imp5, three different induction temperatures (16°C, 25°C and 30°C) were first tested. Then the binding buffer for affinity chromatography was optimized using phosphate buffers at pH 6.0-8.0 with various concentrations of glycerol (10% or 20%) and NaCl (150-300 mM) (Figure 6-8). Imp5 was tandem purified using

GSH sepharose, Hitrap Phenyl, Hitrap Q and Superdex S200 columns. Proteins were concentrated, flash frozen and stored at -80°C.

Crystallization

Purified Trn-SR were used to set up four crystal screens: PACT (Qiagen Inc., CA, USA), JCSG+ (Molecular Dimension, FL, USA), Index (Hampton Research, CA, USA), Winzard I&II (Emerald Biosystems, WA, UAS). The protein (7 mg/ml) and reservoir buffers were mixed up in drops of 0.3 µl : 0.3 µl in 3-well INTELLI-PLATE™ 96 (Art Robbins Instruments, CA, USA) using Phoenix Liquid Handling System (Art Robbins Instruments, CA, USA). The plates were checked for crystals after 24-hour incubation at 20°C.

RanGMPPnP loading

About 500 µl of purified human Ran (~16 mg/ml) were mixed with 30 µl of GMPPnP (50 mg/ml, Sigma) and 4 µl of EDTA (0.5 M) on ice for 1 hour, then added 4 µl of MgAc (2M) and 1 µl of DTT (1M) for another 30 min on ice. The reaction mixture was loaded onto Superdex S75 column (GE Healthcare, NJ, USA) to separate the proteins and excess GMPPnP. The loaded RanGMPPnP was used to make complex with Trn-SR in a molar ratio of 5:1.

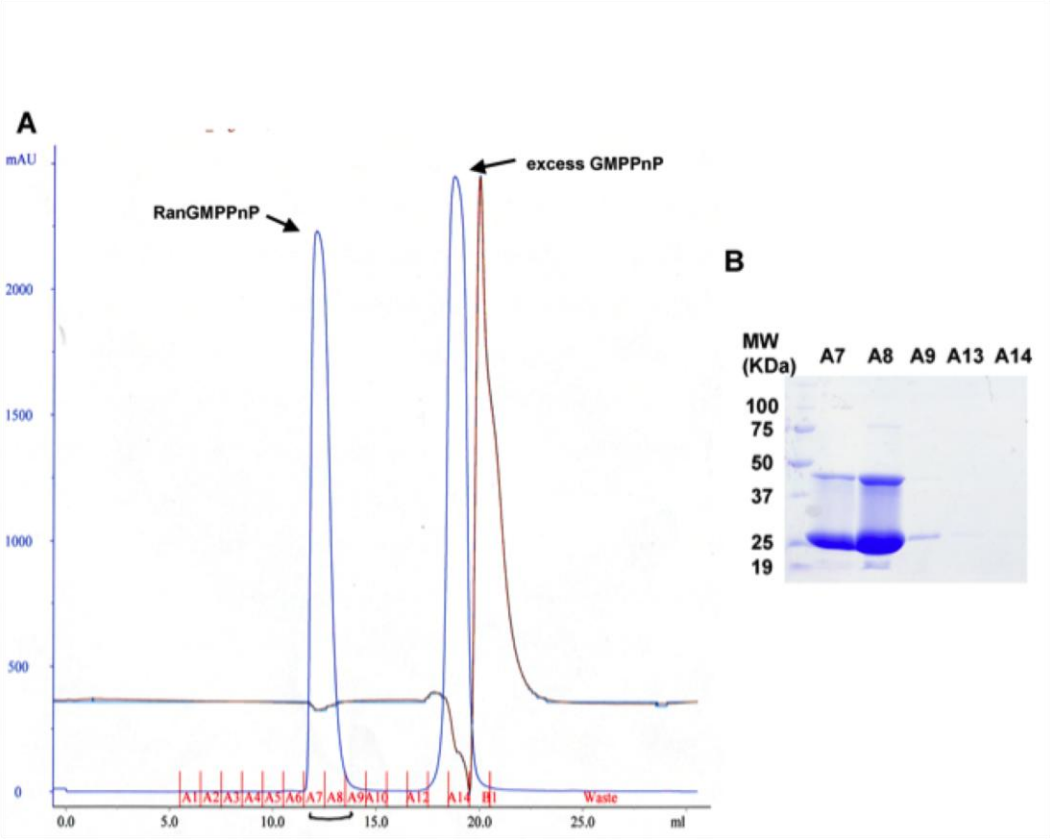

Figure 6-5 Separation of RanGMPPnP from excess GMPPnP after nucleotide exchange reaction. (A) The chromatograph of RanGMPPnP on Superdex S75 column. (B) Gel samples from the indicated fractions were run on 12% SDS-PAGE gel.

In vitro binding assay and ran-dissociation assay

Approximately 20 µg of GST-p35-NLS were immobilized on glutathione sepharose (Amersham, NJ, USA). 20 µg of Impα, Impβ, Kapβ2 and Imp5 were added to the peptide bound sepharose for 10 minutes followed by extensive washing TB Buffer (20 mM HEPES pH7.3, 110 mM KAc, 2 mM DTT, 2 mM MgAc, 1 mM EGTA and 20%

Glycerol). For Ran-dissociation assay, a second incubation was done with or without 40 µl of RanGTP (2.8 mg/ml). After extensive washing, a fifth of the bound proteins were analyzed by SDS-PAGE and visualized with Coomassie staining.

Results and Discussions

Transportin-SR expression and crystallization

Trn-SR expressed well in *E. coli* and its purification resulted in good yield and purity for crystallization. Four popular crystal screen kits were used to crystallize the cargo-free full-length Trn-SR. About 20% of the conditions in Index HT™, 40% in Wizard I&II, 40% in JCSG+, and 12% in PACT show precipitation after 1-3 days of incubation at 20 °C. No crystals resulted from any of these screens. The low percentage of drops with precipitates suggests that the concentration of Trn-SR used in these screens (7 mg/ml) was too low. Increasing protein concentration and crystallizing different constructs of Trn-SR should be tried in future experiments. Limited protease treatment may be also performed to detect regions prone to degradation, which affects the crystallization. Due to the intrinsic flexibility of karyopherins, which are composed of multiple HEAT repeats, it may be difficult to crystallize cargo-free Trn-SR. For example, several research groups have avidly tried for decades but still could not crystallize cargo-free exportin CRM1.

Trn-SR bound to cargos or other binding partners may crystallize more readily. cDNAs of three Trn-SR cargos (ASF/ and SC-35, TRA2-alpha) were obtained from Dr. Kristen Lynch (Univ Penn). Trn-SR transports its cargo also in a Ran-dependent manner (Kataoka, Bachorik et al. 1999). Full-length human Ran was purified and loaded with

RanGMPPnP to form a complex with Trn-SR. But the results showed that Trn-SR and Ran came out as two separate peaks on gel filtration and no complex were detected (Figure 6-6). Ran is a small GTPase and easily inactivated during purification. The activity of RanGMPPnP needs to be further confirmed.

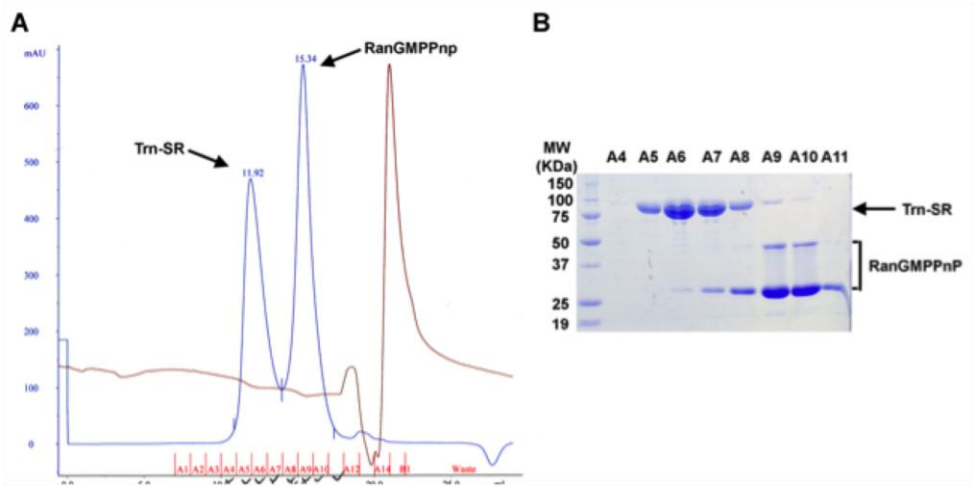

Figure 6-6 Gel filtration after mixing Trn-SR and RanGMPPnP together. (A) The chromatograph of Trn-SR with RanGMPPmP on Superdex S200 column. (B) Gel samples from the indicated fractions were run on 12% SDS-PAGE gel.

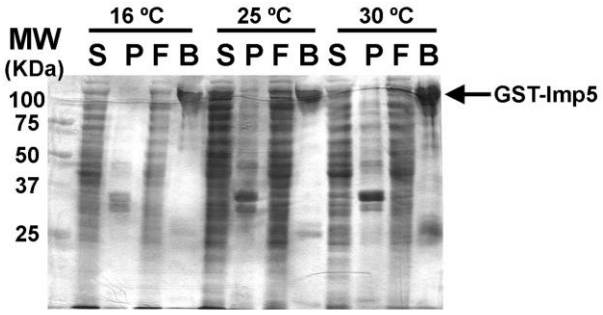

Figure 6-7 Expression optimization of Imp5. Imp5 was expressed at indicated temperatur, lysed and put through glutathione beads. The supernatant (S), pellet (P), flowthrough (F) and beads (B) from each sample were run on 12% SDS-PAGE gel.

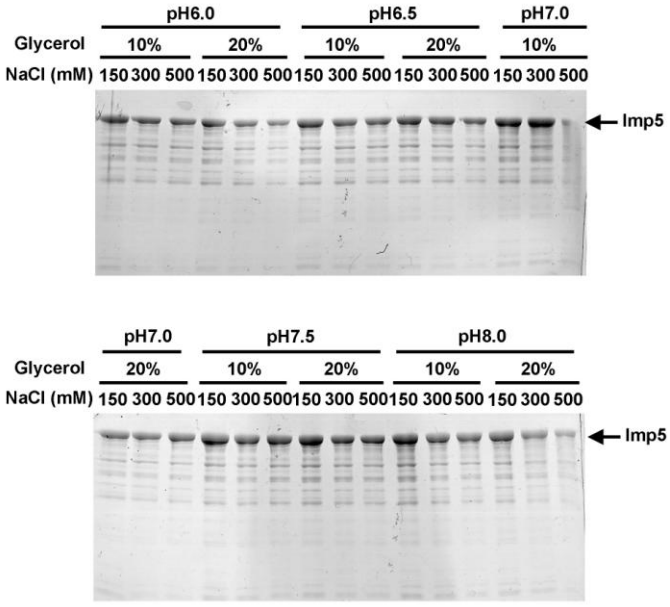

Figure 6-8 Buffer optimization for Imp5. Phosphate buffers containing various concentrations of glycerol and NaCl at different pHs were used as binding buffer for affinity purification of Imp5. After extensive washing, the samples of eluates were run on 12% SDS-PAGE gel.

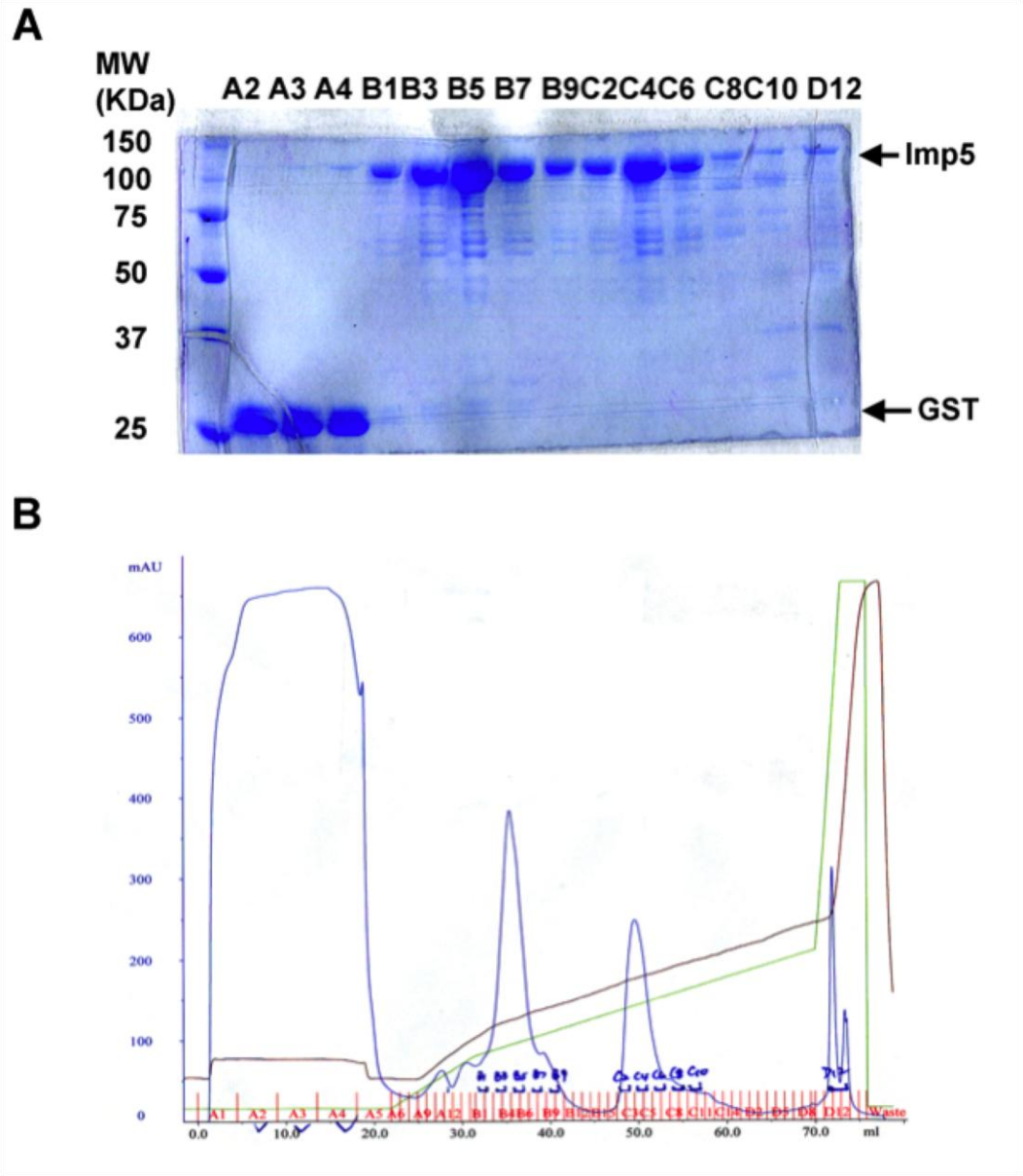

Figure 6-9 Ion exchange purification of Imp5. (A) Gel samples of the indicated fractions from ion exchange purification of were run on 12% SDS-PAGE gel. (B) The chromatograph of Imp5 on 5 ml HiTrap Q column.

Optimization of Imp5 Purification

Imp5 was also expressed in *E. coli.* In order to optimize the expression condition, three different temperatures were tested: 16°C for 16 hours, 25°C for 16 hours and 30°C for 5 hours. The expression level at 30°C was higher than 16°C and 25°C but also had more non-specific bands at low-molecular weigh (Figure 6-7). Thus I chose 25°C for 16 hours to express Imp5. To increase the binding of GST-Imp5 to glutathione beads in affinity chromatography, systematic tests of binding buffer solutions were performed. Since the useful pH range of Tris buffer is limited at 7.0-9.0, I chose phosphate buffers to test the pH effect on binding at the range of 6.0-8.0. The binding buffers also contained 10% or 20% of glycerol and 150-300 mM of NaCl. After extensive washing, the proteins were eluted with Tris buffer (50 mM Tris pH7.5, 2 mM EDTA, 2 mM DTT, 20%glycerol, 300mM NaCl) containing 20 mM of glutathione and samples were run on 12% SDS-PAGE gels. The purity of each eluate is similar, however the yield decreases with the increase of salt (150 mM> 300 mM> 500mM) and glycerol (10%>15%) concentration (Figure 6-8). Thus the optimal pH is pH7.5 (Figure 6-8). In the following purification, Tris buffer (50 mM Tris pH7.5, 2 mM EDTA, 2 mM DTT, 20%glycerol, 150mM NaCl) was used as binding buffer. When loaded onto the Q column, Imp5 came out at two peakes with different impurities. In order to improve the purification, a Phenyl column (GE Healthcare, NJ, USA) was tried before loading Imp5 onto the Q column. Imp5 bound strongly on the Phenyl column and was eluted only at the end of NaCl gradient (no salt buffer). This procedure did remove some impurities at ~25 KDa but the majority of impurities were persistent (Figure 6-11). When the eluates were run on the Q column

again, the first peak containing more impurities at ~10 KDa was much lower compared to previous purification (Figure 6-11). Both peak fractions ran at the same position on gel filtration (Figure 6-12).

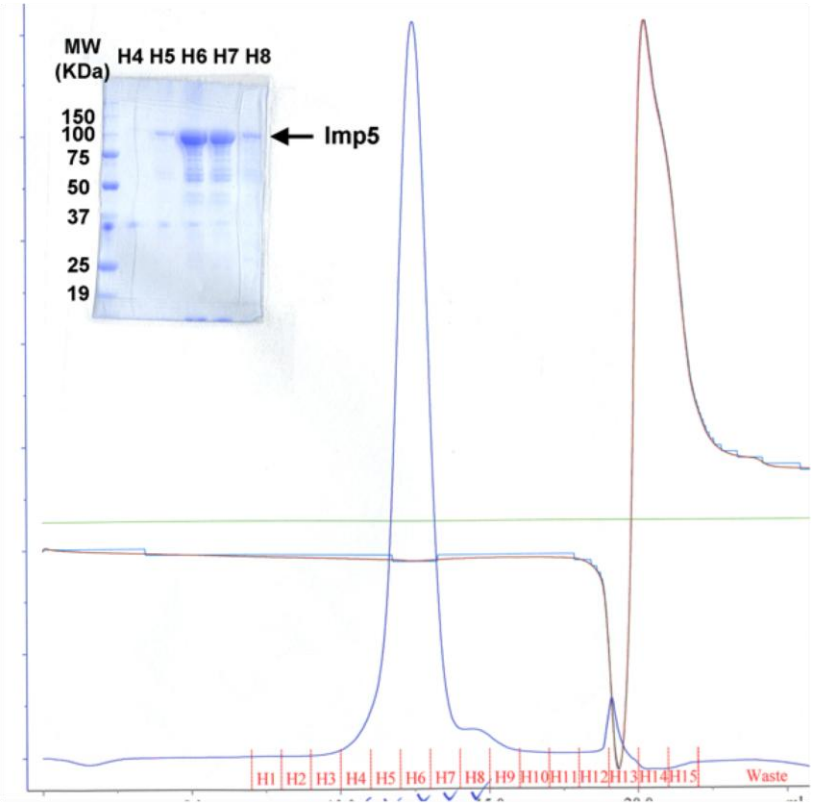

Figure 6-10 Chromatograph of Imp5 on Superdex S200 column. Gel samples were run on 12% SDS-PAGE gel**.**

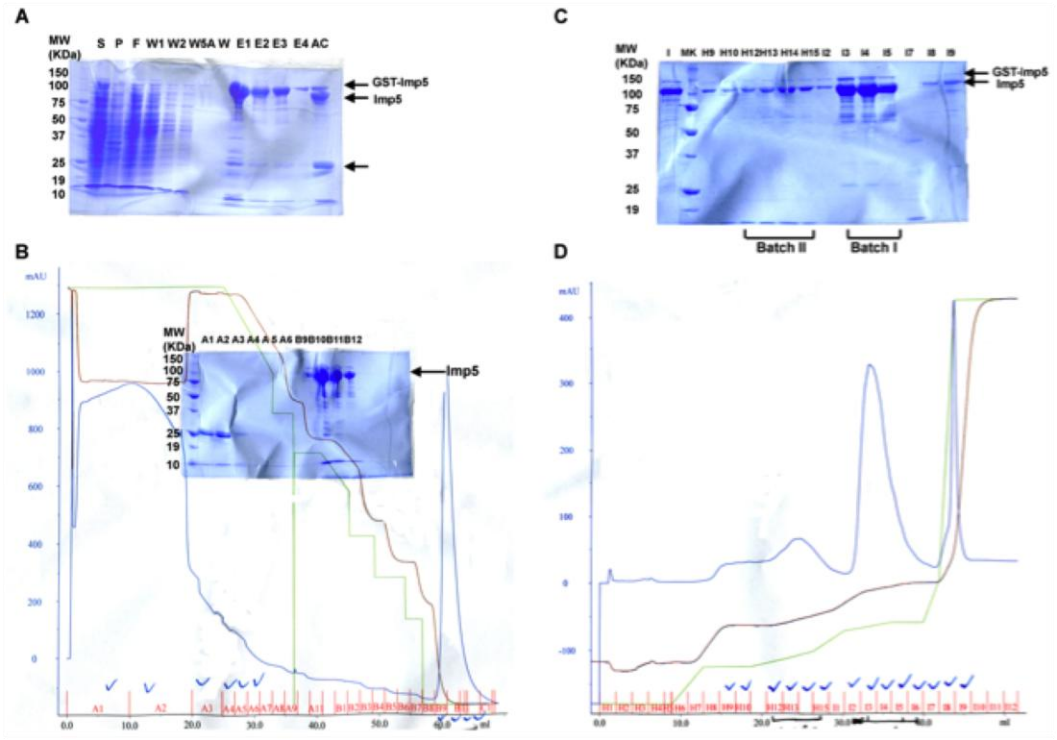

Figure 6-11 Purification of Imp5. (A) Gel samples from affinity purification of Imp5. S, supernatant; P, pellet; F, flowthrough; W1, W2, W5 and W, Tris buffer washes; A, ATP buffer wash; E1-E4, eluates; AC, after TEV cleavage. (B) The chromatograph and gels samples of Imp5 on Phenyl column. (C) and (D) Gel samples and the chromatograph of Imp5 on 5 ml HiTrap Q column after Phenyl column.

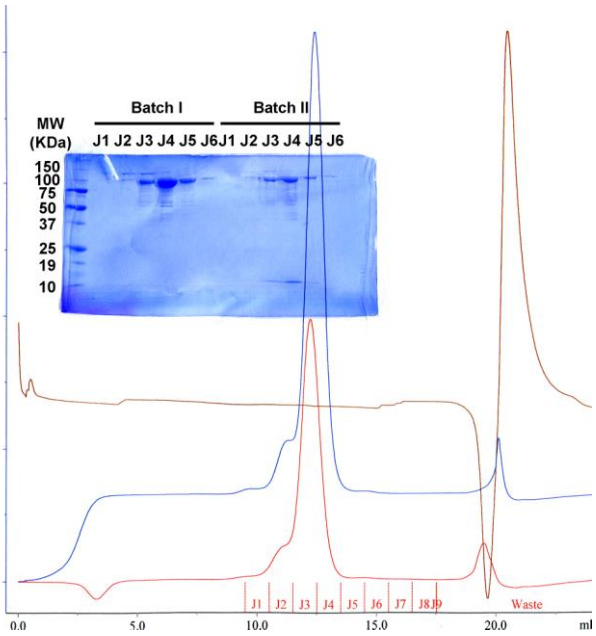

Figure 6-12 The chromatograph and gel samples of Imp5 on Superdex S200 after Phenyl and Q columns. Batch I and Batch II are two peak fractions from the Q column.

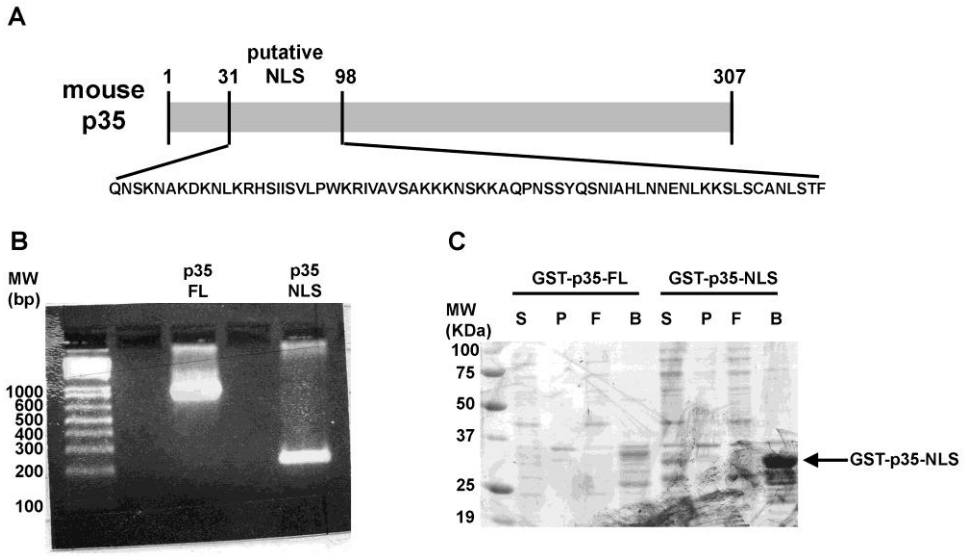

Figure 6-13 Cloning and expression of mouse p35. (A) Schematic diagram of full-length mouse p35 and the putative NLS sequence. (B) PCR products of full-length and NLS of p35. (C) Expression of GST-p35-FL and GST-p35-NLS. S, supernatant; P, pellet; F, flowthrough; B, beads.

Interaction of Imp5 with p35

The CDK5 activator p35 are was reported as a cargo of Imp5 and its NLS was mapped to within residues 31-98 (Fu, Choi et al. 2006). Mouse p35 (full-length) and a fragment containing residues 31-98 were cloned into pGexTev and expressed as N-terminal GST fusion proteins (Figure 6-13). Full-length p35 did not express at 25 °C but the GST-p35 NLS expressed well under the same condition. The GSTp35-NLS was tested for binding with Impα, Impβ, Kapβ2 and Imp5. GSTp35-NLS bound only Imp5 and is dissociated by RanGTP. I have verified that p35 possesses a NLS for Imp5 between residues 31-98. Further biochemical and structural analyses are needed to study the interaction between Imp5 and p35.

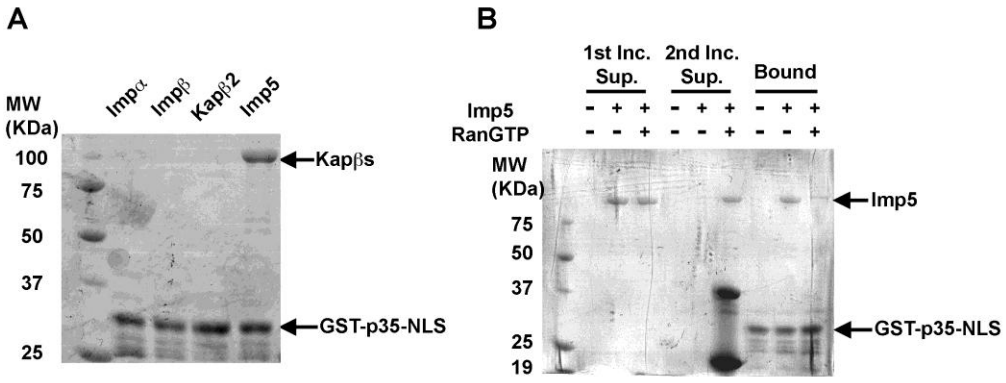

Figure 6-14 Interaction of p35 with Imp5. (A) Binding assays of p35-NLS with Impα, Impβ, Kapβ2 and Imp5. Immoblized GST-p35-NLS was incubated with purified recombinant Impα, Impβ, Kapβ2 and Imp5. (B) Ran dissociation assay of p35 and Imp5. Immoblized GST-p35-NLS was first incubated with Imp5 and then incubated with buffter or RanGTP. 1st Inc. Sup, supernatants from the first incubation; 2nd Inc. Sup, supernatants from the second incubation; Bound, proteins bound on glutathione beads.

Conclusions

In this chapter, Trn-SR and Imp5 were expressed as recombinant proteins for structural studies. Both proteins behaved well during purification and resulted in good purity and decent yield. Although the first attempt to crystallize cargo-free Trn-SR was not successful, several cargos of Trn-SR have been cloned to form Trn-SR-cargo complexes, which may be readily for crystallization. I have verified direct interaction between Imp5 and its cargo p35. These preliminary data suggest that Trn-SR and Imp5 are good candidates for structural studies of new karyopherin import pathways.

CHAPTER SEVEN

NUCLEAR IMPORT OF CRTC PROTEINS

Abstract

Hyperglycaemia is a long-known risk factor of cardiovascular diseases and tight control of glucose metabolism lowers the incidence of the diseases. CREB-regulated transcription coactivators or CRTCs, also known as Transducers of regulated CREB activity or TORCs are key regulators of fasting glucose metabolism. Stimuli-induced nuclear transport is an essential and conserved step of the CRTC signaling pathway, but little is known as for. This chapter describes biochemical and cellular studies pertaining to nuclear-cytoplasmic localization of CRTC1. We identified a conserved putative PY-NLS within residues 19-163 of CRTC1 and demonstrated its direct interaction with Kapβ2, suggesting that Kapβ2 may mediate nuclear import of CRTC1. Surprisingly, wild type CRTC1 is excluded from the nucleus of HeLa cells in the absence of stimuli, but mutant S151A tends to accumulate in the nucleus. Nuclear import of CRTC1(S151A) was not affected by Kapβ2-specific inhibitor M9M. Combined with the evidence that other Kapβs can also bind CRTC1 via the same region, we propose that nuclear import of CRTC1 involves multiple Kapβ pathways. The detailed mechanism of CRTC1 localization needs further investigation.

Introduction and Background

Hyperglycaemia is a long-known risk factor of cardiovascular diseases (CVDs) and leads to a series of maladaptive stimuli that result in myocardial fibrosis and collagen deposition (Epstein 1967; Aneja, Tang et al. 2008; Dokken 2008). Tight control of glucose metabolism lowers the outcome of CVDs. The CREB regulated transcription coactivators (CRTCs or Transducers of regulated CREB activity/TORCs) are key regulators of fasting glucose metabolism (Liu, Dentin et al. 2008). In response to fasting stimuli, cytoplasmic CRTCs are dephosphorylated and transported into the nucleus where they activate CREB-dependent transcription and enhance gluconeogenic program (Screaton, Conkright et al. 2004; Katoh, Takemori et al. 2006; Takemori, Kajimura et al. 2007; Jansson, Ng et al. 2008). Even though stimuli-induced nuclear transport is an essential and conserved step of the CRTC signaling pathway, little is known about the mechanism of this critical step. The goal of the project described in this chapter was to identify Kaps that import CRTCs, characterize their NLSs and study the mechanism of Kap-CRTC recognition through structural analysis. These studies will provide the first molecular details for nuclear import of key regulators in glucose metabolism and help develop drugs that modulate glucogenesis and protect cardiovascular systems of patients. Studies to understand the mechanisms of glucose metabolism are important because glycemic control is critical for therapeutic approaches to reduce the incidence of devastating complications (Shaw, Cardenas et al. 2005; Anselmino, Mellbin et al. 2008).

The CRTC signaling pathway

CRTCs were first identified in mucoepidermoid carcinoma (Tonon, Modi et al. 2003)and later were found to interact with and activate CREB (Conkright, Canettieri et al. 2003; Iourgenko, Zhang et al. 2003). CRTC1, CRTC2 and CRTC3 were identified in human cells (Conkright, Canettieri et al. 2003) and shown to play important roles in glucose metabolism and energy balance (Liu, Dentin et al. 2008). In the resting state, CRTCs are phosphorylated at critical serine sites (S151 in CRTC1 and S171 in CRTC2) by salt-inducible kinase (SIK) (Katoh, Takemori et al. 2006; Takemori, Kajimura et al. 2007). Phosphorylated CRTCs bind 14-3-3 proteins and are sequestered in cytoplasm (Screaton, Conkright et al. 2004; Jansson, Ng et al. 2008). In response to extracellular stimuli such as hormones and glycogens, SIK is inhibited by elevated cellular cyclic AMP and the phosphatase Calcineurin is activated by elevated calcium, leading to dephosphorylation of CRTCs (Screaton, Conkright et al. 2004). Dephosphorylated CRTCs are imported into the nucleus and activate CREB-dependent transcription for glucogenic genes such as PGC1α (Wu, Huang et al. 2006)(Figure 7-1). Drosophila CRTC also translocates in response to Calcineurin activation (Bittinger, McWhinnie et al. 2004). Thus, nuclear import of CRTCs is an essential and conserved step in CRTC-regulated signaling. However, the detailed mechanisms of this crucial step are still poorly understood.

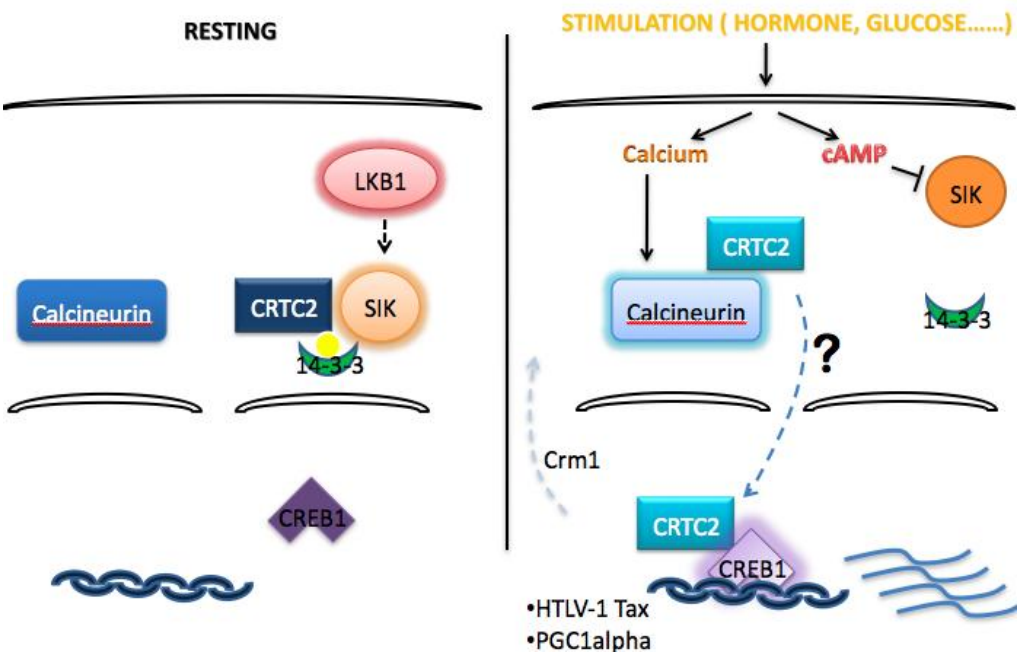

Figure 7-1 CRTC signal cascade.

Identification of Kapβ2 as a binding partner of human CRTC1

Our collaborator, Dr. Frederic Kaye performed a yeast two-hybrid screen to identify candidate CRTC1 binding partners, The Kaye lab used a series of overlapping CRTC1 baits corresponding to N-terminal, middle, and C-terminal fragments (Figure 7-2). No true positive prey clones was identify with baits mapping to the center of the CRTC1 ORF. The C-terminal baits of CRTC1 were self-activating in yeast binding assay. This region is similar to the transcription activation domain of CRTC2 (Conkright, Canettieri et al. 2003) and the activity was expected. However, a bait of CRTC1 residues 1-180 resulted in a positive clone of a Kapβ2 fragment (residues 314-891), which includes its

PY-NLS binding site but missing its inhibitory RanGTP binding site (Figure 7-2)(Kaye). The isolated Kapβ2-AD prey clone was purified and co-transformed into fresh yeast, and it confirmed the binding to the CRTC1-BD clone but not to a series of unrelated baits (data not shown). Since the activity of CRTC1 is regulated by its accessibility to the nucleus, the identification of Kapβ2 suggests that CRTC1 may contain the NLS for Kapβ2 at its N-terminal region and is transported into the nucleus by Kapβ2 pathway.

The goal of this project is to understand the molecular mechanisms for nuclear import of CRTCs using biochemical, structural and cell biological approaches. Knowledge of this important regulatory step of glucose metabolism will lead to further insight of glycemic control and provide a foundation to facilitate future therapeutic efforts for cardiovascular diseases.

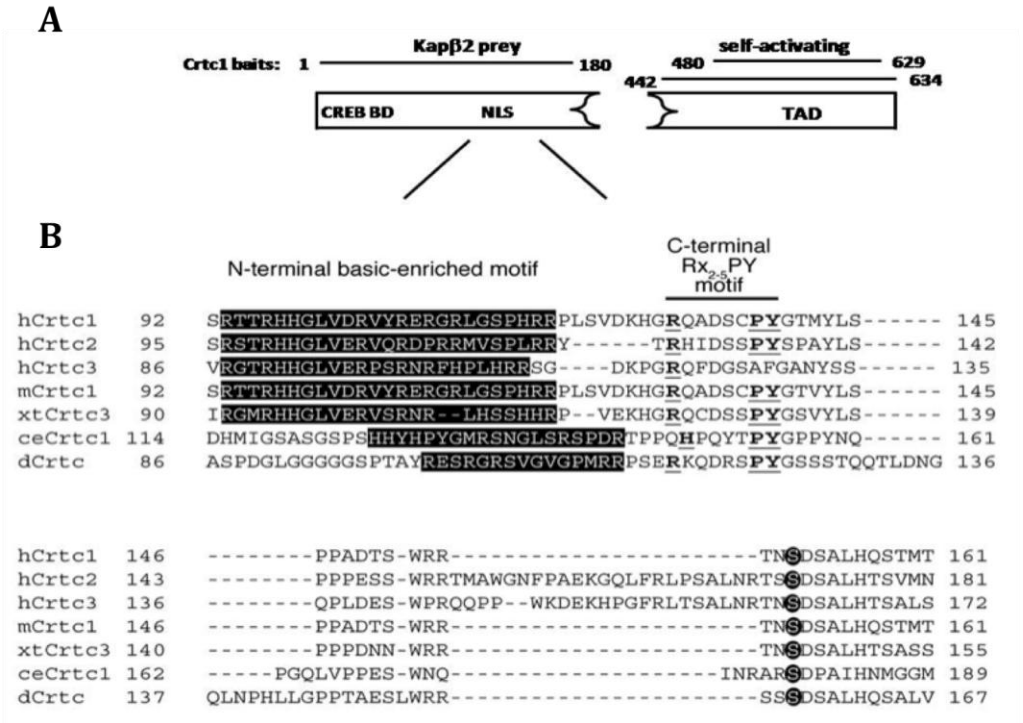

Figure 7-2 Yeast two hybridization and sequence alignment of human CRTRs. (A) The schematic diagram of yeast hybridization assay using human CRTC1 fragments as baits. The N-terminal fragment of CRTC1 got a hit of Kapb2, while the C-terminal fragment is self-activating. (B) Sequence alignment of CRTCs across species. h, Homo sapiens; m, Mus musculus; xt, Xenopus tropicalis; ce, Caenorhabditis elegans; d, Drosophila melanogaster. Black blocks, the basic-enriched regions; black cycles, the critical phosphorylation site.

Materials and Methods

Constructs

Human full-length wild type CRTC1, CRTC2, CRTC3 and mutants CRTC1-S11A, CRTC1-S11D, CRTC1-S151A, CRTC1-S151D, CRTC1-Δ246-249 were constructed into pFLAG-CMV2 vector by Dr. Frederic Kaye's lab (Univ. Florida). The bacteria expressing construct pGEX2TK-hCRTC1-(19-163) was also obtained from Dr. Kaye's lab. The constructs pGEX2TK-hCRTC19-163—PY(P132AY133A) and pFLAG-CMV2-hCRTC1-PY(P132AY133A) were generated using Quikchange ® multiple site-directed mutagenesis kit (Stratagene). A series of truncation mutants of human CRTC1 were cloned into pGEXTEV vector to map the binding regions for Kapβs (Figure 7-3).

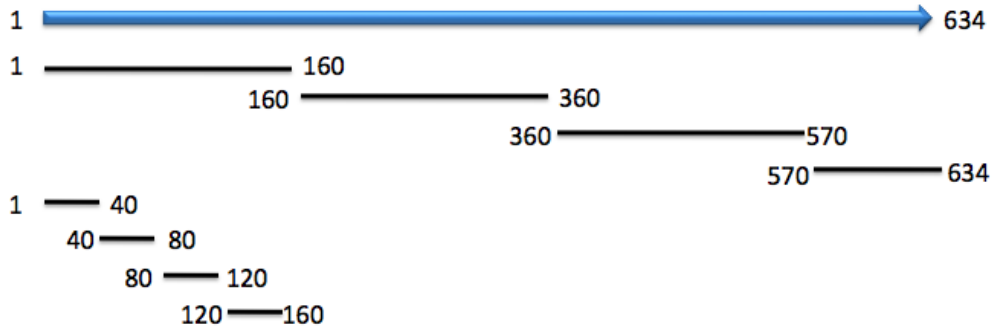

Figure 7-3 Schematic representations of human CRTC1 and its truncation mutants. The numbers indicate the starting and ending residues of each fragment.

In vitro binding, Ran-dissociation assay and competition binding assays

Approximately 20-40 μg of GST-hCRTC1-(19-163) or GST-hCRTC1-(19-163)-PY mutant, were immobilized on glutathione sepharose (Amersham, NJ, USA), and 20 μg of Impβ, Kapβ2, Imp5, Ipo9, Ipo13, Trn-SR, Msn5p, Impα-ARM was added to the peptide bound sepharose for 30 minutes followed by extensive washing (TB Buffer: 20 mM HEPES pH7.3, 110 mM KAc, 2 mM DTT, 2 mM MgAc, 1 mM EGTA and 20% Glycerol). For Ran dissociation aassay, a second incubation was done with 8 mg/ml of RanGTP in 50 μl solution. For competition assay, a second incubation was done with 2 mg/ml of MBP-M9M in 50 μl solution After extensive washing, a quarter of the bound proteins were analyzed by SDS-PAGE and visualized with Coomassie staining.

Subcellular localization.

About 0.2 μg of pFLAG-CMV2 plasmids with inserts of human CRTCs or mutants were transfected into HeLa cells in 24-well plate. After 12-16 hours, cells were fixed by 4% formaldehyde, and permeabilized with 0.2% Triton-X in PBS. The cells were then incubated with primary antibody anti-FLAG (1:400) for 1hour at RT, Alexa546-labeled-goat-anti-rabbit secondary antibody (1:500) for 30 min at RT, followed by DAPI staining. The cells were examined in a Leica TCS SP5 confocal microscope and images were processed with Image J (NIH, MD, USA).

Results and Discussion

Putative PY-NLS of human CRTC1 for Kapβ2

Interestingly, sequence analysis of CRTC1 shows that the sequence flanking the regulatory serine151 has a candidate PY-NLS between amino acids 19 and 163 (Figure 7-2B). This putative signal exists within a larger structurally disordered region spanning amino acids 50-170. The signal also contains signature N-terminal basic-enriched and C-terminal $Rx_{2-5}PY$ motifs typical of PY-NLSs (Lee, Cansizoglu et al. 2006). The putative PY-NLS of CRTC1 overlaps with the previously mapped nuclear localization signal at amino acids 56-144 of CRTC2/Torc2 (Screaton, Conkright et al. 2004) and shows conservation across species (Figure 7-3A). To validate the putative PY-NLS of CRTC1, we purified and tested GST-CRTC1-(19-163) for interactions with recombinant Kap2 in the absence or presence of RanGTP (Figure 7-4). We detected strong binding to Kap2 that was blocked by the addition of RanGTP. This indicates that the NLS of CRTC1, like all other PY-NLS, binds Kap®2 in a Ran-sensitive manner. In a competition binding assay, the binding of CRTC1-(19-163) to Kap®2 was competed by the inhibitor M9M (Figure 7-5), suggesting that CRTC1-(19-163) likely occupies the same binding site as other PY-NLSs. These results suggest that Kap2 is most likely a nuclear import factor that transports CRTC1 into the nucleus.

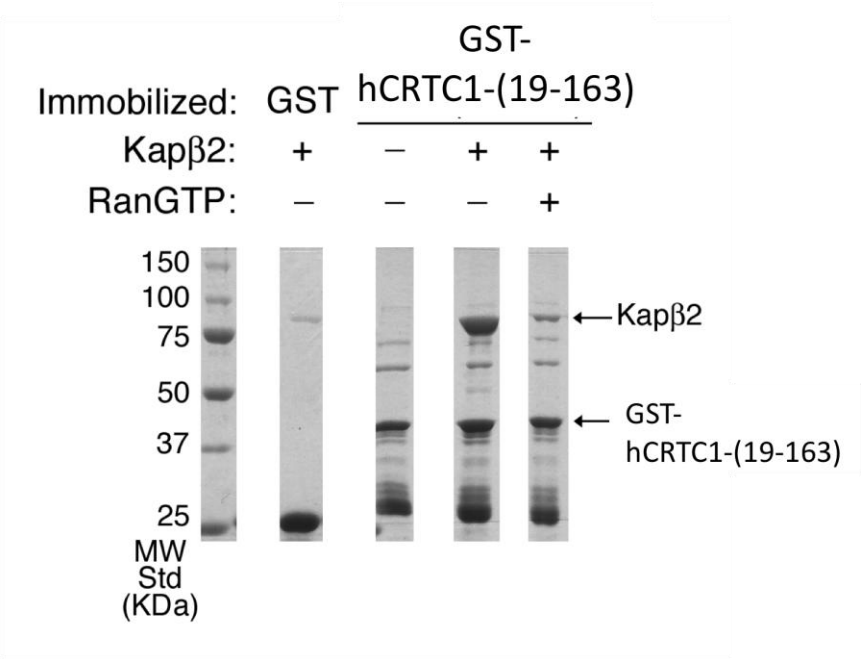

Figure 7-4 Human CRTC1 interacts with Kapb2 in Ran-sensitive manner. Immobilized GST-hCRTC1-(19-163) fragment was first incubated with purified recombinant Kapβ2, and then incubated with buffer or RanGTP. The proteins bound to GSH beads were resolved on SDS-PAGE and visualized by Coomassie blue staining.

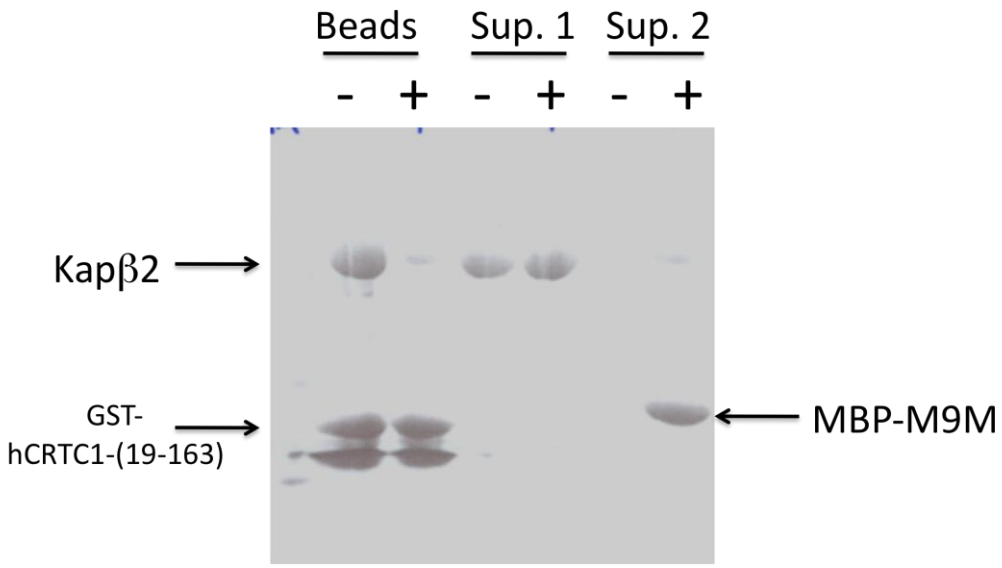

Figure 7-5 Competition binding assays of hCRTC1 and Kapb2 with MBP-M9M. Immobilized GST-hCRTC1-(19-163) fragment was first incubated with purified recombinant Kpab2, and then incubated with buffer or Kapβ2 inhibitor MBP-M9M. The proteins bound to GSH beads were resolved on SDS-PAGE and visualized by Coomassie blue staining.

Previous studies have shown that the PY-NLSs consist of three energetically independent binding epitopes: 1) the N-terminal hydrophobic/basic motif, 2) the arginine residue of the C-terminal $RX_{2-5}PY$ sequence motif, and 3) the PY of the C-terminal $RX_{2-5}PY$ motif (Lee, Cansizoglu et al. 2006). The binding energy contribution of each epitope can be different in PY-NLSs (Cansizoglu, Lee et al. 2007; Suel, Gu et al. 2008). To investigate the energy contribution of the conserved PY motif in CRTC1, I mutated the $^{132}PY^{133}$ into alanines and tested the binding of the mutant to Kap®2. The result shows that PY mutant binds Kap®2 equally well as the wild type (Figure 7-6), suggesting most of the binding energy comes from other parts of the NLS or evenly distributed among the epitopes. Mutations at other epitopes of CRTC1-NLS are needed for further investigation.

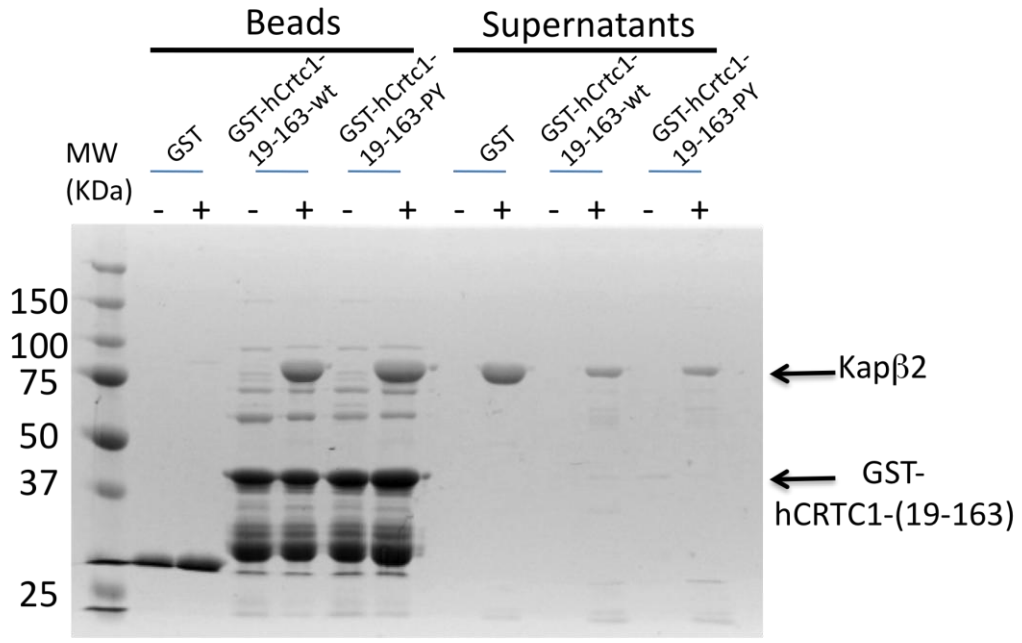

Figure 7-6 CRTC1 PY mutant has similar Kapβ2 binding affinity as the wild type. GST-hCRTC1-(19-163) wt and PY mutant were immobilized on glutathione beads and incubated with Kapb2. The proteins bound to the beads and in the supernatants were resolved on SDS-PAGE gel.

Subcellular localization of human CRTCs

Next, FLAG tagged full-length CRTCs were transfected into HeLa cells and their subcellular localizations were detected by immunofluorescence. In the absence of stimulation, wild type CRTC1 is excluded from the nucleus, and CRTC2 is primarily in the cytoplasm (Figure 7-7 and 7-10). CRTC3, which does not have the conserved PY motif (Figure 7-2), accumulates in the nucleus (Figure 7-9), indicating the nuclear import

of CRTC3 may be different from its homologs CRTC1 and CRTC2. Mutations of CRTC1 (S11A, S11D, S151D, Δ246-249, P132AY133A) and Kapβ2 inhibitor M9M did not affect the localization of CRTC1 (Figure 7-7).

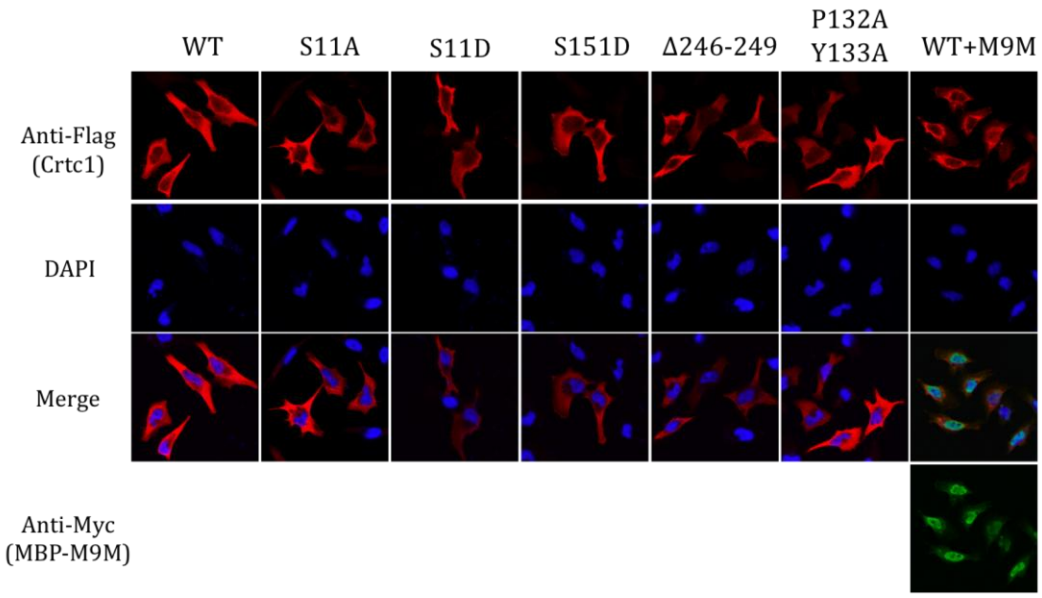

Figure 7-7 Subcellular localization of hCRTC1 wild type and mutants.

Several groups have reported that phosphorylation at a critical serine site (S151 of CRTC1, S171 of CRTC2, respectively) regulates nuclear import of CRTCs (Screaton, Conkright et al. 2004; Katoh, Takemori et al. 2006; Takemori, Kajimura et al. 2007; Jansson, Ng et al. 2008). Dephosphorylation of CRTC1 at S151 was shown to promote nuclear import of CRTC1. In our experiment, the S151A mutant of CRTC1, which cannot be phosphorylated at residue 151, accumulates in the nucleus even without any stimulation (Figure 7-8). These results suggested that phosphorylation at S151 might

mask the NLSs and dephosphorylation makes the NLS accessible to transport receptors. However, the co-transfection of Kapβ2 inhibitor M9M did not mislocalize the S151A mutant into the cytoplasm (Figure 7-8). It is possible that other import pathways are involved in the n uclear import of CRTC1.

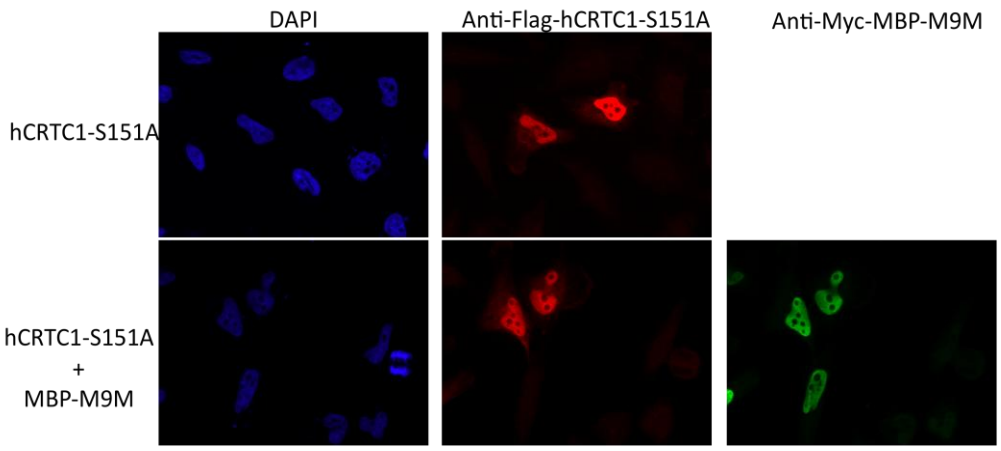

Figure 7-8 Subcelluar localization of hCRTC1-S151A mutant with or without Kapb2 inhibitor MBP-M9M.

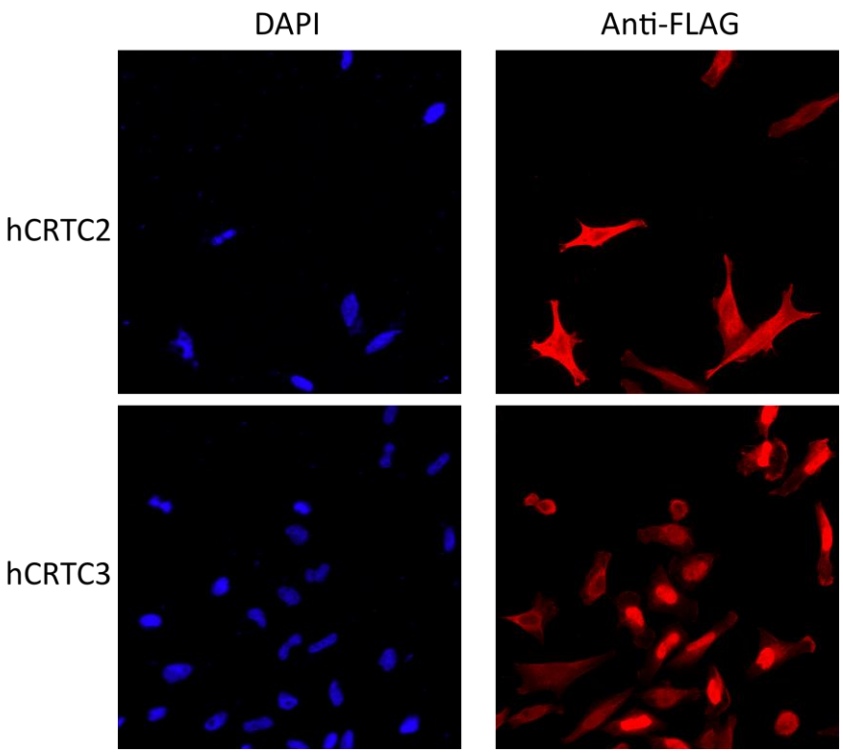

Figure 7-9 Subcellular localization of hCRTC2 and hCRTC3.

Other Kapβs involved in CRTC1 nuclear import

In order to test if other Kapβs can also import CRTC1, immobilized GST-CRTC1-(9-163) was incubated with recombinant Kapβs. Besides Kapβ2, Trn-SR shows stoichiometric binding to CRTC1 (Figure 7-10), suggesting that Trn-SR may also import CRTC1. Impβ, Imp5, Ipo9 and Ipo13 also bound to the CRTC1 fragment, but much more weakly compared to Kapβ2 and Trn-SR (Figure 7-10). They may not be the major import factors for CRTC1. Msn5p and Impa did not bind to CRTC1. Although these in vitro

binding assays using recombinant proteins showed direct interactions between CRTC1 and multiple Kapβs, the ability of these various Kapβs in the nuclear import of CRTC1 has to be tested in future cell-based assays.

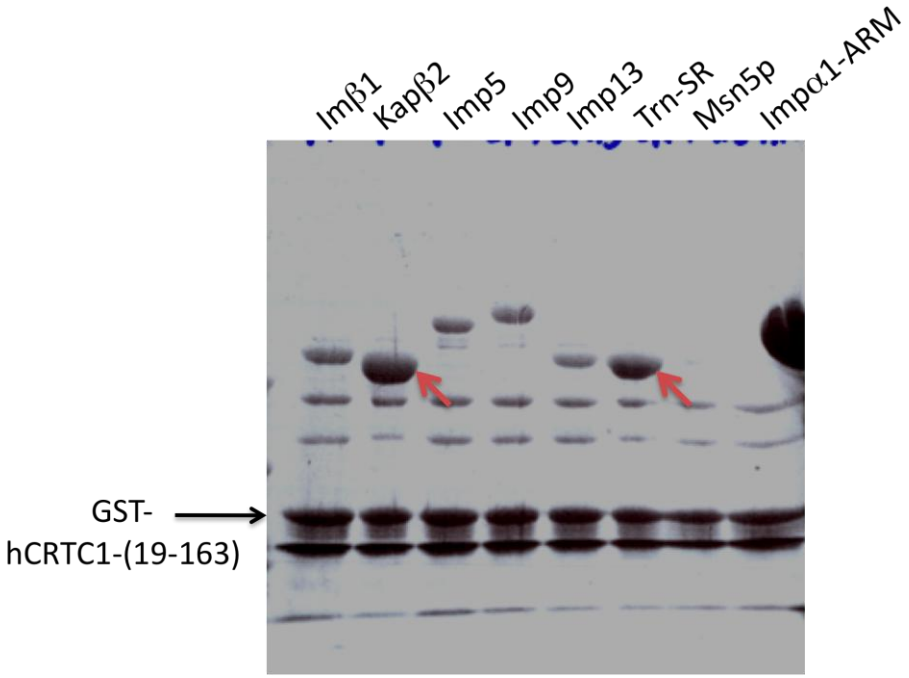

Figure 7-10 Binding assays of hCRTRC1-(19-163) with Kapβs. The red arrows mark the Kapβ2 and Trn-SR that have strong binding to hCRTC1-(19-163).

Conclusions

In summary, we found that the N-terminus of CRTC1 contains a putative PY-NLS that can interact directly with Kapβ2. This region is conserved among CRTC1 homologs. Nuclear accumulation of CRTC1 requires dephosphorylation of residue S151 and

blocking the Kapb2 pathway using inhibitor M9M does not alter the nuclear localization of the CRTC1(S151A) mutant. Trn-SR also binds CRTC1 via the same region. The detailed mechanism of CRTC1 import needs to be further investigated in the future.

CHAPTER EIGHT

STRUCTURAL ANALYSIS OF KARYOPHERIN-MEDIATED NUCLEOCYTOPLASMIC TRANSPORT (LITERATURE REVIEW)

Abstract

In human cells, the majority of nucleocytoplasmic transport is mediated by 19 members of the Karyopherinβ (Kapβs/Importins/Exportins) protein family. Thus, Kapβs are critically involved in cellular processes such as gene expression, signal transduction, immune response, oncogenesis and viral propagation, all of which require proper nucleocytoplasmic targeting. Despite the importance of nucleocytoplasmic transport, the mechanisms of transport particularly of nuclear export and the distinctions in targeting signals recognized by the different Kapβ pathways remain poorly understood. Many crystal structures of two different import pathways involving Impβ and Kapβ2 are available and they provide structural explanations for the different steps of nuclear import such substrate recognition, nucleoporin binding and Ran-mediated substrate dissociation. In contrast, the only available export-Kapβ structures are of Cse1p and of a fragment of Crm1. In this chapter we will review structures of karyopherins complexed with transport substrates, nucleoporins and the Ran GTPase in both import and export systems, and the resulting mechanistic insights from comparative analysis of the current collection of atomic resolution nuclear transport structures.

Introduction

Proteins in the Karyopherinβ (Kapβ/Importin/Exportin) family mediate the majority of macromolecular nucleocytoplasmic transport in eukaryotic cells. Nucleocytoplasmic transport is signal-mediated: nuclear localization signals (NLSs) and nuclear export signals (NESs) in macromolecules direct them in and out of the nucleus, respectively. Kapβ proteins recognize these signals and target transport substrates to the nuclear pore complex (NPC) for translocation(Gorlich and Kutay 1999; Chook and Blobel 2001; Conti and Izaurralde 2001; Macara 2001; Stewart, Baker et al. 2001; Fahrenkrog and Aebi 2003; Weis 2003; Mosammaparast and Pemberton 2004; Pemberton and Paschal 2005; Conti, Muller et al. 2006; Madrid and Weis 2006; Cook, Bono et al. 2007)

There are 19 known human Kapβs and 14 known *S. cerevisiae* Kapβs.(Fried and Kutay 2003; Mosammaparast and Pemberton 2004) Each Kapβ functions in distinct nuclear import, export or bi-directional transport. The proteins share similar molecular weights (90-150 kDa) and isoelectric points (pI = 4.0-5.0), low sequence identity (8-15%) and all made up of almost entirely of helical repeats. Kapβs recognize multiple classes of ligands. Each member of the family binds unique sets of proteins or RNA. In addition, Kapβs also bind phenylalanine-glycine (FG) repeats in nucleoporins to target Kapβ-substrate complexes to the NPC.(Gorlich and Kutay 1999; Chook and Blobel 2001; Stewart, Baker et al. 2001; Fahrenkrog and Aebi 2003; Pemberton and Paschal 2005; Cook, Bono et al. 2007; Stewart 2007). Finally, all Kapβs bind the Ran GTPase, which regulates Kapβ-substrate interactions and transport directionality through its nucleotide

cycle (Gorlich and Kutay 1999; Chook and Blobel 2001; Conti and Izaurralde 2001; Macara 2001; Stewart, Baker et al. 2001; Fahrenkrog and Aebi 2003; Weis 2003; Mosammaparast and Pemberton 2004; Pemberton and Paschal 2005; Conti, Muller et al. 2006; Madrid and Weis 2006; Cook, Bono et al. 2007) RanGTP is concentrated in the nucleus, while RanGDP is concentrated in the cytoplasm. In nuclear import pathways, RanGTP and substrates bind Kapβs competitively, allowing substrate binding in the cytoplasm and RanGTP-mediated release in the nucleus. In contrast, export-Kapβs bind RanGTP and substrates cooperatively, resulting in substrate binding in the nucleus and release in the cytoplasm as Ran-bound GTP is hydrolyzed (Gorlich and Kutay 1999; Chook and Blobel 2001; Conti and Izaurralde 2001; Macara 2001; Stewart, Baker et al. 2001; Fahrenkrog and Aebi 2003; Weis 2003; Mosammaparast and Pemberton 2004; Pemberton and Paschal 2005; Conti, Muller et al. 2006; Madrid and Weis 2006; Cook, Bono et al. 2007).

Ten Kapβs have been shown to function in nuclear import (Mosammaparast and Pemberton 2004). The best-characterized nuclear import pathway is known as the classical Impα/Impβ (also known as Kapα/Kapβ1) pathway (Conti and Izaurralde 2001). Impβ binds its adaptor protein Impα (also known as Kapα), which in turn recognizes the classical short basic NLS (Conti and Izaurralde 2001). Impβ also binds directly to a distinct set of import substrates, without using Impα or another adaptor protein. In fact, none of the other nine import-Kapβs uses adaptor proteins for substrate binding. Surprisingly, most import pathways have not been well-charcaterized as only a few substrates have been identified for most of the import-Kapβs, and large panels of

substrates are currently known for only Impβ and Kapβ2 (also known as Transportin)(Chook and Blobel 2001; Lee, Cansizoglu et al. 2006). Accordingly, high-resolution structures are currently available only for these two import pathways. Numerous crystal structures of different ligand-bound states of Impα, Impβ and Kapβ2 now provide atomic level explanations for the different steps of nuclear import such substrate recognition, nucleoporin binding and Ran-mediated substrate dissociation.(Conti, Uy et al. 1998; Chook and Blobel 1999; Cingolani, Petosa et al. 1999; Kobe 1999; Vetter, Arndt et al. 1999; Bayliss, Littlewood et al. 2000; Conti and Kuriyan 2000; Fontes, Teh et al. 2000; Lee, Imamoto et al. 2000; Bayliss, Littlewood et al. 2002; Cingolani, Bednenko et al. 2002; Fontes, Teh et al. 2003; Fontes, Teh et al. 2003; Lee, Sekimoto et al. 2003; Matsuura, Lange et al. 2003; Matsuura and Stewart 2004; Petosa, Schoehn et al. 2004; Chen, Ben-Efraim et al. 2005; Lee, Matsuura et al. 2005; Liu and Stewart 2005; Matsuura and Stewart 2005; Lee, Cansizoglu et al. 2006; Cansizoglu and Chook 2007; Cansizoglu, Lee et al. 2007; Imasaki, Shimizu et al. 2007; Mitrousis, Olia et al. 2008).

Even less structural information is available for nuclear export. Although there are seven known export pathways, structures of full-length export-Kapβ are available only for Cse1p, the yeast homolog of human export-Kapβ CAS (Matsuura and Stewart 2004; Cook, Fernandez et al. 2005). Cse1p is a specialized exporter with a single known substrate, Kap60p, which is the yeast homolog of Impα(Kutay, Izaurralde et al. 1997; Solsbacher, Maurer et al. 1998). Crystal structures of unliganded Cse1p and of the Cse1p-Kap60p-RanGTP export substrate complex provide structural explanations for substrate

recognition, positive cooperativity between Kap60p and RanGTP in export complex assembly and substrate dissociation in the cytoplasm (Nishinaka, Masutani et al. 2004; Cook, Fernandez et al. 2005).

CRM1 is the most general and versatile export-Kapβ, with >200 export substrates identified to date. Most of these export substrates contain short leucine-rich (LR) NESs, which conform loosely to the L-X_{2-3}-[LIVFM]-X_{2-3}-L-X-[LI] consensus (Fischer, Huber et al. 1995; Wen, Meinkoth et al. 1995; Fornerod, Ohno et al. 1997), but CRM1 also binds substrates without recognizable LR-NES such as Snurportin1 (SPN1)(Paraskeva, Izaurralde et al. 1999). Despite the importance and prevalence of Crm1-substrate recognition in cells, the only crystal structure available for this system is of the C-terminal third of Crm1 (Petosa, Schoehn et al. 2004). Nevertheless, high resolution structure of this fragment when combined with a low resolution electron microscopy (EM) reconstructed image of unliganded full length CRM1 and mutagenesis analysis, has led to a model of how export substrate binding may be regulated by Ran through a large internal Crm1 loop (Petosa, Schoehn et al. 2004).

Structural organization of the Karyopherins

Impα

Impα contains a positively charged N-terminal domain known as the Impβ binding (IBB) domain (Görlich, Henklein et al. 1996),(Weis, Dingwall et al. 1996), a central armadillo (ARM) domain and a small hydrophilic C-terminal tail(Herold, Truant et al. 1998). The central ARM domain has 10 α-helical repeats known as ARM repeats, which were first

identified in the gene product of the Drosophila gene *armadillo* and its human ortholog β-catenin. (Wieschaus and Riggleman 1987; Riggleman, Wieschaus et al. 1989);(Peifer, McCrea et al. 1992; Peifer, Berg et al. 1994; Huber, Nelson et al. 1997). An ARM repeat has approximately 40 amino acids that form three α-helices H1, H2 and H3 (Weis, Dingwall et al. 1996; Conti, Uy et al. 1998; Conti and Kuriyan 2000). Consecutive ARM repeats form a cylindrical superhelical structure with a shallow groove along the superhelical axis that is lined by the H3 helices (Fig. 1a) (Conti, Uy et al. 1998).

Kapβs

Kapβs are generally made up of almost entirely α-helical HEAT repeats. All known Kapβs are predicted to have 19-20 HEAT repeats, which were first identified in the proteins Huntingtin, elongation factor 3, the Protein phosphatase 2A PR65/A subunit and TOR1 Kinase, hence the term "HEAT"(Andrade and Bork 1995; Groves, Hanlon et al. 1999; Chook and Blobel 2001). A HEAT repeat consists of two antiparallel helices A and B. Individual helices in Kapβs are named according to their position in the HEAT repeat such that the A helix of HEAT repeat 1 is abbreviated to H1A. The helices are connected by either loops or short helices such that each Kapβ HEAT repeat contain either two or three helices. The Kapβ HEAT repeats stack in a parallel manner to produce the single contiguous hydrophobic core of a superhelical structure with A and B helices lining the concave and convex surfaces, respectively (Figs. 2-5)(Chook and Blobel 2001; Cook, Bono et al. 2007).

ARMs versus HEATs

Superposition of Impα ARM repeats with Impβ HEAT repeats show significant structural similarities (rmsd <1 Å), supporting suggestions that both types of α-helical repeat motifs are highly related (Malik, Eickbush et al. 1997; Cingolani, Petosa et al. 1999; Vetter, Arndt et al. 1999; Chook and Blobel 2001). Furthermore, both types of repeats have similar overall protein architecture: the repeats stack to form two ribbons of parallel α-helices with B or H3 helices lining the concave side of the proteins and the A or H1-H2 ribbon lining their convex sides. However, ARM repeats contain conserved consensus residues whereas Kapβ HEAT repeats show almost no sequence conservation (Huber, Nelson et al. 1997; Malik, Eickbush et al. 1997; Conti, Uy et al. 1998). Curvature of the two types of proteins are also different: regular 30 ° rotations of ARM repeats generate a smooth and elongated Impα whereas variable 10 °-60 ° HEAT repeat rotations bend Kapβ superhelices to generate coils and arches(Chook and Blobel 2001).

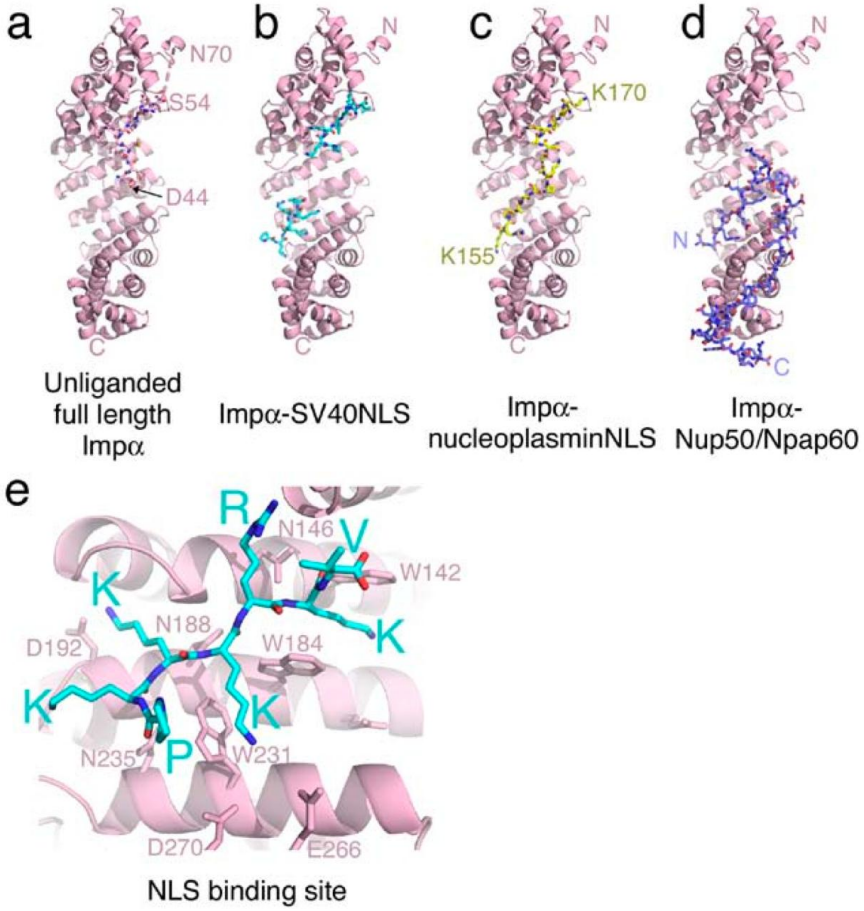

Figure 8-1 Crystal structures of Impα. a) Unliganded full- length Impα (1IAL). The ARM domain of Impα (residues N70- F496) is represented by a ribbon drawing with its N- terminal IBB domain (only residues D44- S54 modeled) shown as a stick figure. b- c) Impα in complex with classical NLSs: monopartite SV40- NLS (b, 1EJL) and bipartite nucleoplasmin- NLS (c, 1EJY). Impα ARM domains are shown as in a and the NLSs shown as stick figures. First (K155) and last (K170) residues of the nucleoplasmin- NLS are labeled. d) Impα in complex with residues 1- 50 of Nup50 (2C1M). The Impα ARM domain is shown as in a- c and Nup1 is shown as stick figure. e. Interactions between Impα and SV40- NLS at major NLS binding site (ARM2- 4). Both proteins are represented as in b) NLS residues are labeled in large font and selected Impα residues in this binding site are labeled in a smaller font

Structural analysis of nuclear import

Substrate recognition in the classical Impα/Impβ pathway

In the classical Impα/Impβ pathway, Impα is the adaptor protein that binds both the classical-NLS and Impβ (Gorlich and Kutay 1999; Chook and Blobel 2001; Pemberton and Paschal 2005; Cook, Bono et al. 2007). Two classes of classical-NLSs have been characterized: monopartite and bipartite NLSs. Monopartite NLSs contain a single cluster of basic residues whereas bipartite sequences contain two clusters of basic residues separated by a 10-12 amino acid linker (Dingwall, Sharnick et al. 1982; Kalderon, Richardson et al. 1984; Dingwall and Laskey 1991; Pemberton, Blobel et al. 1998).

The Impα NLS binding site was first observed in the crystal structure of N-terminally truncated yeast Impα (yImpα) bound to the SV40 T antigen NLS peptide, which is a typical monopartite classical-NLS (Conti, Uy et al. 1998). Since then, structures of Impα have been solved with monopartite NLSs from proteins c-myc (Conti and Kuriyan 2000) and phopholipid scramblase 1 (PLSCR1) (Chen, Ben-Efraim et al. 2005). Due to autoinhibition of Impα by its N-terminal IBB domain (Kobe 1999) (Fig. 1a), all crystal structures of Impα-NLS complexes were obtained using N-terminally truncated karyopherin. (Conti, Uy et al. 1998; Conti and Kuriyan 2000; Fontes, Teh et al. 2003; Matsuura, Lange et al. 2003; Chen, Ben-Efraim et al. 2005; Matsuura and Stewart 2005). Impα is conformationally invariant in these structures, suggesting that its ARM domain is relatively rigid and little structural change accompanies NLS binding.

General features of NLS binding are conserved in all Impα-NLS structures (Conti, Uy et al. 1998; Conti and Kuriyan 2000; Fontes, Teh et al. 2000; Fontes, Teh et al. 2003; Matsuura, Lange et al. 2003; Chen, Ben-Efraim et al. 2005; Matsuura and Stewart 2005). The monopartite NLS cores bind at two sites, at the concave faces of ARM2-4 (major site) and ARM6-8 (minor site), respectively (Conti, Uy et al. 1998; Conti and Kuriyan 2000; Fontes, Teh et al. 2000). Acidic residues at the periphery of these binding grooves form electrostatic interactions with the basic NLS side chains, conserved tryptophan residues on H3 helices form hydrophobic interactions with long aliphatic portions of the basic NLS side chains and asparagine residues on H3 helices form hydrogen bonds with the NLS mainchains. Structural placement of the conserved tryptophans and acidic residues that interact with the NLS sidechains has provided rationale for mutagenesis, thermodynamic studies and for the previously defined consensus K-K/R-X-K/R that defines the core of monopartite NLSs (Fig. 1b, e) (Kalderon, Richardson et al. 1984; Dingwall and Laskey 1991; Conti, Uy et al. 1998).

Crystal structures of Impα bound to bipartite NLSs from proteins nucleoplasmin (Conti and Kuriyan 2000), retinoblastoma and N1N2 (Fontes, Teh et al. 2003) proteins have been solved. In these structures, a single bipartite NLS spans both monopartite NLS major and minor binding sites, with a connecting linker between them (Fig. 1c). Similar structural determinants in the monopartite NLS complexes also apply to bipartite NLS binding. The bipartite NLS linker interacts with Impα only through mainchain contacts. One example of a consensus for the bipartite NLS is $KRX_{10\text{-}12}KRRK$ (Conti and Kuriyan 2000). In general, the consensus for this larger NLS is less well defined than for the small

monopartite NLS. This may be due to the lack of mutagenic and thermodynamic analyses coupled with multivalency of a larger signal, which can accommodate larger sequence diversity (Suel, Gu et al. 2008).

The Impα/Impβ heterodimer has significantly higher affinity for NLS than Impα alone (Moroianu, Blobel et al. 1996; Fanara, Hodel et al. 2000). Structure of the full-length mouse Impα shows that residues 44-54(DEQMLKRRNVS) in the IBB domain bind the N-terminal major NLS binding site (ARM2-4) in the same manner as an exogenous NLS peptide (Fig. 1a) (Kobe 1999). This structural information confirms previous biochemical findings that the IBB domain autoinhibits Impα through an internal pseudo-NLS sequence (Moroianu, Blobel et al. 1996; Kobe 1999). Autoinhibition is relieved by removal of the IBB domain such as by N-terminal truncation of Impα (Conti, Uy et al. 1998; Conti and Kuriyan 2000; Fontes, Teh et al. 2000; Fontes, Teh et al. 2003; Matsuura, Lange et al. 2003; Chen, Ben-Efraim et al. 2005; Matsuura and Stewart 2005) or by Impβ binding the IBB domain for nuclear import (Görlich, Henklein et al. 1996; Weis, Ryder et al. 1996; Kobe 1999). Autoinhibition is restored in the nucleus as RanGTP dissociates the Impα/Impβ heterodimer (Vetter, Nowak et al. 1999; Lee, Matsuura et al. 2005). The IBB domain is then freed for intramolecular competition at the NLS binding site, resulting in release of the exogenous NLS.

A final class of Impα ligands consists of metazoan nucleoporin Nup50/Npap60[22] and the functionally analogous yeast nucleoporin Nup2p (Matsuura, Lange et al. 2003). The N-termini of both nucleoporins contain 50 residues that bind Impα with higher affinity than

NLS peptides and thus accelerate NLS dissociation from Impα (Matsuura, Lange et al. 2003; Matsuura and Stewart 2005). Even though these nucleoporin fragments contain stretches of basic residues reminiscent of a bipartite NLS, they bind Impα in entirely distinct manners, hence their classification as non-NLS ligands. Structures of Impα bound to Nup50 and Nup2p are similar (Matsuura, Lange et al. 2003; Matsuura and Stewart 2005). The nucleoporin fragments show bipartite interactions with Impα: their N-terminal basic cluster bind the minor NLS site in a manner similar to NLS peptides, but the rest of the nucleoporin peptides extend towards the C-terminus of Impα with critical interactions of another basic cluster contacting the outer surface of ARM9-10 (Fig. 1d) (Matsuura, Lange et al. 2003; Matsuura and Stewart 2005). Bipartite Impα-Nup interaction appears to be critical for the ability of the nucleoporins to actively displace NLS from Impα in the nucleus.

Direct substrate recognition by Impβ

Impβ is the most widely studied member of the Kapβ family. Although it was first identified in the classical nuclear import pathway using Impα as an adaptor to recognize the classical-NLSs (Adam and Adam 1994; Chi, Adam et al. 1995; Görlich, Vogel et al. 1995; Imamoto, Tachibana et al. 1995; Radu, Blobel et al. 1995), Impβ also binds a different set of substrates directly. These Impβ substrates include retroviral proteins Rev and Tat (Truant and Cullen 1999), ribosomal proteins L23a and L5 (Jäkel and Görlich 1998), transcription factors SREBP-2 (Nagoshi, Imamoto et al. 1999), GAL4 (Chan, Hübner et al. 1998), CREB, Jun, fos (Forwood, Lam et al. 2001), and Smad3(Xiao, Liu et

al. 2000) and other proteins such as parathyroid hormone-related protein (PTHrP) (Lam, Briggs et al. 1999) and cyclin B (Moore, Yang et al. 1999). Furthermore, since Impβ binds and transports Impα into the nucleus, the latter is considered a direct Impβ substrate (Cingolani, Petosa et al. 1999). No consensus recognition sequence has been defined this collection of substrates. Structural studies of Impβ bound to the IBB domains of Impα(Cingolani, Petosa et al. 1999) and Snurportin 1 (SPN1) (Mitrousis, Olia et al. 2008), SREBP-2 (Lee, Sekimoto et al. 2003) and PTHrP (Cingolani, Bednenko et al. 2002) explain substrate recognition mechanisms of Impβ and how the karyopherin can recognize diverse substrates.

Impβ binds the N-terminal IBB domain of Impα (αIBB) (Chi, Adam et al. 1997; Chi and Adam 1997; Kutay, Izaurralde et al. 1997). In the free protein, αIBB is in extended conformation and binds the major NLS binding site on the ARM domain (Fig. 1a)(Kobe 1999). However, when bound to Impβ, the structure of the αIBB is entirely different. Impβ-bound αIBB is an L-shaped molecule with N-terminal residues 11-23 in extended conformation followed by a perpendicular C-terminal helix (Cingolani, Petosa et al. 1999) (Fig. 2a). When bound to the αIBB, full-length Impβ adopts a compact snail-like helicoidal shape of ~50 Å in diameter. The αIBB is wrapped at the center of the superhelix, binding Impβ's inner concave surface. Its extended N-terminal portion interacts with H7-11 and the H8 acidic loop while its C-terminal helix contact H12-19. The complex is primarily stabilized by electrostatic interactions, with basic residues in the αIBB interacting with acidic residues of Impβ.

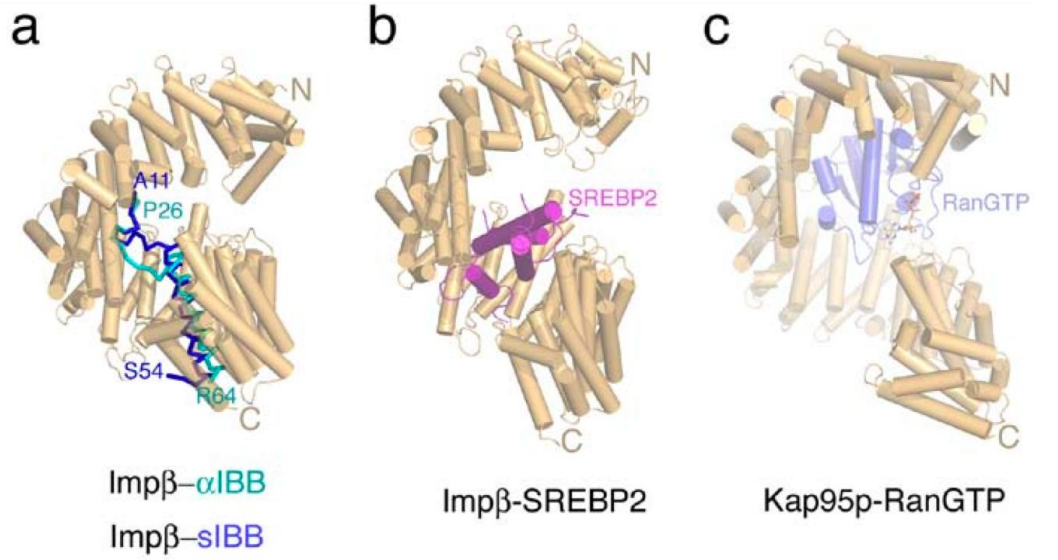

Figure 8-2 Structures of Impβ complexes. a) Impβs bound to the IBB domains of Imp and SPN1 (αIBB and sIBB; 1QGK and 2P8Q), respectively. Impβ is shown in a ribbon representation with α—helices drawn as gold cylinders. Impα is shown as a green ribbon and sIBB as a blue ribbons. b) The Impβ- SREBP2 complex (1UKL). Impβ is drawn as in a and the ribbon diagram of the SREBP2 dimer in magenta. c) The Kap95p- RanGTP complex (2BKU). Kap95p is drawn similar to Impβ in a and b and the ribbon drawing of RanGTP is in blue.

SPN1 recognizes and imports snRNPs into the nucleus (Palacios, Hetzer et al. 1997; Huber, Cronshagen et al. 1998). SPN1 also has an IBB domain (denoted sIBB) similar to that in Impα (Strasser, Dickmanns et al. 2005). Structure of the Impβ-sIBB complex shows that Impβ is virtually identical compared to that in complex with αIBB (Cingolani, Petosa et al. 1999; Mitrousis, Olia et al. 2008). sIBB residues 25-65 are homologous to the αIBB and adopt a similar helical conformation when bound to Impβ. Within this epitope, a short 3_{10} helix (residue 27-30) is connected perpendicularly to a long α-helix (residues 41-65) by a 7-residue linker (Fig. 2a). As with the αIBB, all basic residues in

this helical region of sIBB are critical for binding Impβ. Thermodynamic studies suggest that the sIBB-Impβ interaction is bipartite in nature (Mitrousis, Olia et al. 2008). Residues 1-24 in sIBB shares sequence similarity with residues 1011-1035 of Nup153 (Mitrousis, Olia et al. 2008). This region of the sIBB is predicted to adopt an extended conformation to bind H1-10 of Impβ. The bipartite nature of sIBB-Impβ binding is probably critical to confer high affinity interaction as the sIBB 1-65 fragment binds Impβ 7-fold tighter than the 25-65 fragment. Finally, similarity between the N-terminal epitope of sIBB to Nup153 suggests a possible mechanism for SPN1 release from Impβ in the nucleus (Mitrousis, Olia et al. 2008). Nup153 displaces sIBB residues 1-24 from Impβ, destabilizing the Impβ-substrate complex to release SPN1 into the nucleus.

PTHrP is a secretory hormone, which regulates cell apoptosis and proliferation (Clemens, Cormier et al. 2001). It contains an NLS at residues 66-94, which binds Impβ with dissociation constant (K_D) of 2 nM (Lam, Briggs et al. 1999). The crystal structure of a truncated Impβ (H1-11) bound to PTHrP-NLS shows the peptide binding to the concave surface of the Impβ N-terminal arch in an extended conformation with its mainchain running parallel to the superhelical axis of Impβ (Cingolani, Petosa et al. 1999). The PTHrP-NLS binding site is entirely distinct from that for αIBB and sIBB, which is mostly in the C-terminal arch of Impβ (Cingolani, Petosa et al. 1999; Mitrousis, Olia et al. 2008). Instead, PTHrP-NLS occupies the Ran binding site in the N-terminal arch of Impβ, suggesting that direct competition with Ran is sufficient to release it in the nucleus.

SREBP-2 is a transcription factor of genes that control cholesterol metabolism (Brown and Goldstein 1997). It binds Impβ directly through its basic helix-loop-helix leucine zipper (bHLHZ) domain (Nagoshi, Imamoto et al. 1999; Nagoshi and Yoneda 2001; Lee, Sekimoto et al. 2003). The structure of full length Impβ bound to the SREBP-2 bHLHZ domain shows that SREBP-2 binds Impβ as a dimer(Lee, Sekimoto et al. 2003). The HLHZ dimer inserts into the central portion of Impβ between H7 and H17, in a direction perpendicular to the central axis of the Impβ superhelix (Fig. 2b). Compared to the αIBB-bound Impβ (Cingolani, Petosa et al. 1999), the SREBP-2-bound superhelix adopts a more twisted open conformation to accommodate the HLHZ dimer. Two long Impβ helices H7B and H17B are the major binding sites for the HLHZ dimer.

Impβ substrates are structurally very diverse (Cingolani, Petosa et al. 1999; Cingolani, Bednenko et al. 2002; Lee, Sekimoto et al. 2003; Mitrousis, Olia et al. 2008). Furthermore, other than PTHrP-NLS,(Cingolani, Bednenko et al. 2002) substrates recognized by Impβ are three dimensional epitopes,(Cingolani, Petosa et al. 1999; Lee, Sekimoto et al. 2003; Mitrousis, Olia et al. 2008) which cannot be defined by consensus sequences. These substrates also bind to different sites on Impβ. PTHrP-NLS shares the N-terminal arch of Impβ with Ran while SREBP-2 binds in the central region between the N- and C-terminal arches and IBB domains bind mainly to the C-terminal arch of Impβ. The flexible nature of Impβ molecule allows it to adopt different conformations to accommodate these structurally diverse substrates (Conti, Muller et al. 2006; Cansizoglu and Chook 2007).

Substrate recognition by Kapβ2

Kapβ2 is a prototypical import-Kapβ. It binds import substrates and nucleoporins simultaneously to target substrates to the NPC. It also binds RanGTP with high affinity to release substrates in the nucleus (Chook and Blobel 1999; Chook, Jung et al. 2002). Numerous structures of Kapβ2 have been determined, including the unliganded karyopherin (Cansizoglu, Lee et al. 2007), Kapβ2 bound to RanGTP (Chook and Blobel 1999) and five different Kapβ2-substrate complexes (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007; Imasaki, Shimizu et al. 2007). The latter include NLSs from mRNA binding proteins hnRNPs A1(Lee, Cansizoglu et al. 2006), D(Imasaki, Shimizu et al. 2007) and M,(Cansizoglu, Lee et al. 2007) JKTBP-1(Imasaki, Shimizu et al. 2007) and mRNA export factor TAP/NXF1(Imasaki, Shimizu et al. 2007).

Prior to availability of Kapβ2-substrate structures, ~20 mRNA processing proteins (such as hnRNPs A1, D, F and M, HuR, DDX3, Y-box binding protein 1 and TAP) had been identified as Kapβ2 import substrates (Pollard, Michael et al. 1996; Bonifaci, Moroianu et al. 1997; Siomi, Eder et al. 1997; Fan and Steitz 1998; Truant, Kang et al. 1999; Kawamura, Tomozoe et al. 2002; Guttinger, Muhlhausser et al. 2004; Rebane, Aab et al. 2004; Suzuki, Iijima et al. 2005). Kapβ2 binds its best-characterized substrate, splicing factor hnRNP A1, through the 38-residue M9 sequence (Pollard, Michael et al. 1996; Bonifaci, Moroianu et al. 1997). We now refer to the M9 sequence more generally as the hnRNP A1-NLS (Cansizoglu, Lee et al. 2007). NLSs in HuR, TAP, hnRNP D and its homologs, the JKTPB proteins had previously been mapped but showed marginal or no

sequence homology to the hnRNP A1-NLS. NLSs recognized by Kapβ2 appeared very diverse.(Fan and Steitz 1998; Truant, Kang et al. 1999; Kawamura, Tomozoe et al. 2002; Suzuki, Iijima et al. 2005).

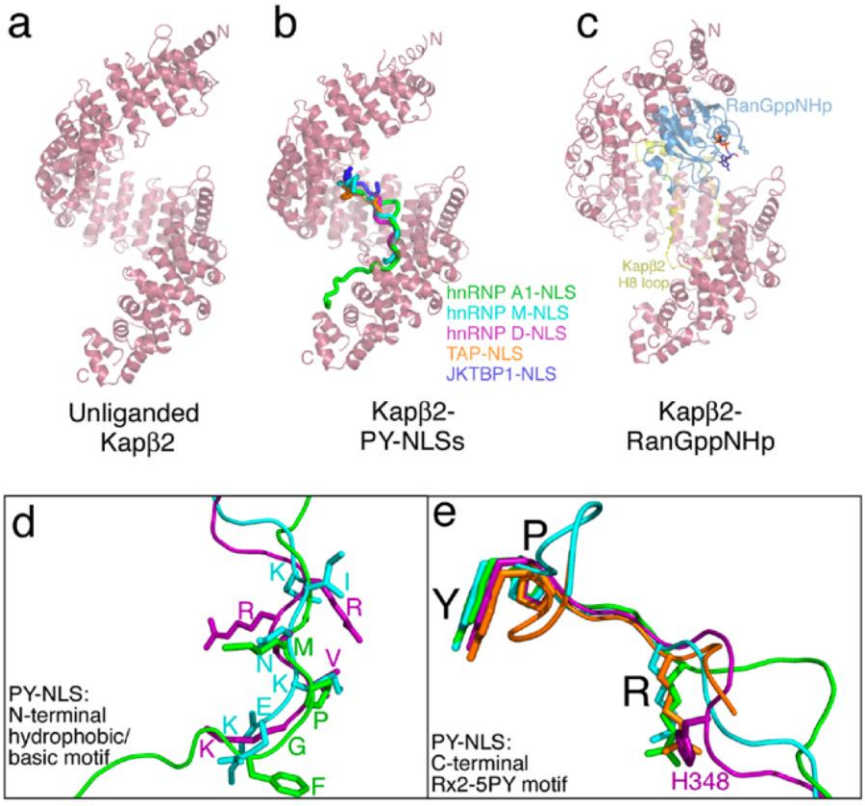

Figure 8-3 Structures of Kapβ2 complexes. a) Unliganded Kapβ2 (2QMR). All Kapβ2 molecules (a- d) are represented as pink ribbons. b) PY- NLSs bind in extended conformation to the C- terminal arch of Kapβ2. Mainchains of hnRNP A1- NLS (green; 2H4M), hnRNP M- NLS (cyan; 2OT8), hnRNP D- NLS (magenta; 2Z5N), TAP- NLS (orange; 2Z5K) and JKTBP1- NLS (dark blue; 2Z5O) as H10- H17 of the Kapβ2s are superimposed. c) The Kapβ2- Ran complex (1QBK). The ribbon diagram of RanGppNHp is in light blue ribbons and the Kapβ2 H8 loop in yellow. d- e) Details of the PY- NLSs in b, focusing on the N- terminal hydrophobic/basic motif (d) and the C- terminal Rx2- 5PY motif (e) of the PY- NLSs

In the Kapβ2-hnRNP A1-NLS complex (Lee, Cansizoglu et al. 2006), 20 Kapβ2 HEAT repeats form an almost perfect superhelix, which can also be described as two overlapping arches. The N-terminal arch spans H1-13 and the C-terminal arch spans H8-20. 26 residues of substrate hnRNP A1-NLS bind in extended conformation to the concave surface of Kapβ2 C-terminal arch with the NLS running antiparallel to the karyopherin superhelix (Fig. 3b). The substrate interface on Kapβ2 is relatively flat without deep pockets or grooves(Lee, Cansizoglu et al. 2006). Most of this Kapβ2 interface is acidic except for a few several small hydrophobic patches that contact the N-terminal FGPM and the C-terminal PY motifs of hnRNP A1-NLS. Despite the highly acidic Kapβ2 interface, the NLS contains only two basic residues: Arg284 forms salt bridges with Kapβ2 and the K277 sidechain is not disordered (Lee, Cansizoglu et al. 2006).

Structure of the Kapβ2-hnRNP A1-NLS complex in combination with biochemical analyses revealed physical rules that describe Kapβ2's recognition of a diverse set of NLSs (Lee, Cansizoglu et al. 2006). These rules or requirements are 1) structural disorder of a 30-residue or larger peptide segment, 2) overall basic character, and 3) weakly conserved sequence motifs composed of a loose N-terminal hydrophobic or basic motif and a C-terminal $RX_{2-5}PY$ motif (Lee, Cansizoglu et al. 2006). This last rule led to the term "PY-NLS" to describe diverse sequences that are recognized by Kapβ2. The composition of N-terminal motifs divides PY-NLSs into hydrophobic and basic subclasses (hPY- and bPY-NLSs). hPY-NLSs contain four consecutive predominantly hydrophobic residues while the equivalent region in bPY-NLSs is enriched in basic

residues (Lee, Cansizoglu et al. 2006). The physical rules that describe PY-NLS recognition are predictive and uncovered 81 new candidate substrates(Lee, Cansizoglu et al. 2006). These new putative NLSs are complex signals, discovered using a collection of individually weak rules rather than just a strongly restrictive sequence motif. Many uncharacterized NLSs/NESs are poorly defined in sequence, and many still unidentified signals across the Kapβ family will likely be similarly ill-defined in sequence. More generally, the concept of signals as a collection of physical rules rather than specific sequence motifs alone may be applicable across organelle systems for the numerous obscure targeting signals in eukaryotic cells.

Structural comparison of Kapβ2-hnRNP A1-NLS (hPY-NLS) and Kapβ2-hnRNP M-NLS (bPY-NLS) complexes explained recognition of the two types of chemically diverse motifs (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007). The Kapβ2 molecules in these structures are conformationally invariant while the PY-NLSs trace different paths. The hydrophobic and basic PY-NLSs converged structurally only at consensus sequence motifs, confirming the consensus designations and suggesting multipartite interaction (Fig. 3d, e) (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007). General features of PY-NLS binding are also observed in structures of Kapβ2 bound to hnRNP D, its homolog JKTBP-1 and mRNA exporter TAP/NXF1 (Fig. 3b) (Imasaki, Shimizu et al. 2007). Electron density is observed only at the C-terminal PY motifs but not at the N-terminal motifs of the latter two structures (Fig. 3d, e) (Imasaki, Shimizu et al. 2007).

PY-NLSs are sequentially and structurally diverse. Studies using yeast Kap104p (*S. cerevisiae* Kapβ2 homolog) suggest how Kapβ2 recognizes such diverse sequences (Suel, Gu et al. 2008). Kap104p binds specifically only to bPY- but not hPY-NLSs and Kap104p-NLS thermodynamics studies confirm the three energetically significant linear PY-NLS epitopes (N-terminal basic motif, the arginine and lastly the proline-tyrosine of the C-terminal $Rx_{2-5}PY$ motif)(Suel, Gu et al. 2008). Each of these epitopes accommodates substantial sequence diversity and interestingly, the epitopes are energetically quasi-independent and a given epitope can contribute differently to total binding energy in different PY-NLSs (Suel, Gu et al. 2008). This last property likely amplifies signal diversity through combinatorial mixing of energetically weak and strong motifs{Suel, 2008 #4122

The ability to recognize diverse substrates may be shared by most Kapβs. Like Impβ, Kapβ2 appears to bind more than one class of substrates, each at different binding sites on the karyopherin. Kehlenbach and colleagues have found Kapβ2 imports HIV1-REV and c-Fos into the nucleus {Arnold, 2006 #573}(Arnold, Nath et al. 2006). Their binding sites on Kapβ2 are different from that for PY-NLS such as the hnRNP A1-NLS thus leading to the suggestion that Kapβ2 uses different binding sites to recognize multiple classes of substrates.

Substrate dissociation by RanGTP

The interaction of RanGTP with import-Kapβs to dissociate substrates in the nucleus is a crucial step in nuclear import.(Gorlich and Kutay 1999; Chook and Blobel 2001; Weis

2002). Unique substrate repertoires for Kapβs suggest significant differences in their mechanisms of substrate recognition and therefore also differences in their regulation by Ran. The latter is seen in two different mechanisms of Ran-mediated substrate dissociation in the Impβ and Kapβ2 pathways.

Impβ binds RanGTP in its N-terminal arch (Fig. 2c). The switch 1 region of Ran contacts H12-15 of Impβ, the switch 2 region contacts H1-4 and the basic patch of the GTPase contacts the 15-residue acidic loop that connects H8 helices (H8 loop) (Vetter, Arndt et al. 1999; Lee, Matsuura et al. 2005). Impβ binds structurally diverse substrates at different sites (Cingolani, Petosa et al. 1999; Cingolani, Bednenko et al. 2002; Lee, Sekimoto et al. 2003; Lee, Sekimoto et al. 2003; Mitrousis, Olia et al. 2008). The shapes and pitches of the Impβ superhelices are also different in these substrate complexes. In order to effectively dissociate structurally diverse proteins, Ran binding globally changes the superhelical structure of Impβ, distorting the different substrate binding sites to release them (Cingolani, Petosa et al. 1999; Cingolani, Bednenko et al. 2002; Lee, Sekimoto et al. 2003; Lee, Sekimoto et al. 2003; Mitrousis, Olia et al. 2008). In addition, Ran and the αIBB helix also contact overlapping sites on the Impβ H8 loop such that occupation of one ligand on the loop is incompatible with the other (Cingolani, Petosa et al. 1999; Vetter, Arndt et al. 1999). Finally, since substrate PTHrP binds in the Impβ N-terminal arch (Cingolani, Petosa et al. 1999), Ran can simply displace it directly. Therefore, RanGTP causes substrate release from Impβ via a combination of global Impβ conformational change and direct displacement.

RanGTP displaces substrates from Kapβ2 through a different mechanism. In the RanGTP state, the two arches of the Kapβ2 superhelix are orthogonal (Fig. 3c) (Chook and Blobel 1999). Ran binds in the N-terminal arch, with its switch regions contacting Kapβ2 H1-3 and its basic patch contacting the 62-residue internal acidic loop (H8 loop), H7-8 and H14-15 of Kapβ2. The H8 loop is sequestered in the Kapβ2 C-terminal arch with a significant portion occupying the PY-NLS binding site (Chook and Blobel 1999; Lee, Cansizoglu et al. 2006; Cansizoglu and Chook 2007; Cansizoglu, Lee et al. 2007). Thus, the substrate binding site is no longer accessible when RanGTP is bound. In contrast, when Ran is absent, biochemical and structural studies show that the H8 loop is exposed and disordered, and the C-terminal arch is empty and free to bind substrate (Fig. 3a) (Lee, Cansizoglu et al. 2006; Cansizoglu and Chook 2007; Cansizoglu, Lee et al. 2007). The Ran and substrate binding sites of Kapβ2 do not overlap. Furthermore, the substrate binding site remains relatively unchanged as Ran and NLS are exchanged. Thus, substrate dissociation from Kapβ2 in the presence of RanGTP is executed by the long internal acidic H8 loop of Kapβ2. When Ran is bound, the Kapβ2 H8 loop occupies the NLS site to displace substrate. Therefore, it appears that the two best-known nuclear import pathways utilize RanGTP to dissociate substrates in different manners.

Many other Kapβs have large insertions in their HEAT repeats like the Kapβ2 H8 loop. Impβ has a 15-residue acidic H8 loop, Cse1 has a 2-helix insertion in H8 and Crm1, Kapβ3, Imp4, Imp7, Imp8, Imp9 and Imp11 all appear to have large insertions in their central repeats (Cingolani, Petosa et al. 1999; Vetter, Arndt et al. 1999; Petosa, Schoehn et al. 2004; Lee, Matsuura et al. 2005). Mutational studies of Crm1 suggest that, like

Kapβ2, a large internal loop may couple Ran and substrate binding directly (see below) (Petosa, Schoehn et al. 2004). Trends for coupling Ran and substrate binding in the Kapβ family are emerging (Lee, Cansizoglu et al. 2006). Kapβ2 and probably Crm1 use a large insertion to directly couple the two ligands with little conformational change in the substrate binding site. In contrast, Kapβ1 and Cse1 (see below) use large-scale conformational changes to transition from closed substrate-free to open substrate-bound conformations.

Interactions with nucleoporins

Kapβs interact with nucleoporins and carry their substrates through the NPC (Gorlich and Kutay 1999; Chook and Blobel 2001; Stewart, Baker et al. 2001; Fahrenkrog and Aebi 2003; Pemberton and Paschal 2005; Cook, Bono et al. 2007; Stewart 2007). About one third of the nucleoporins contain domains with repeating FG motifs interspersed with flexible linkers. There are three classes of FG repeats: FG, GLFG or FxFG, based on their hydrophobic cores (Rout and Wente 1994). FG repeats are natively unstructured and predicted to line the inner surface of the NPC channel (Denning, Patel et al. 2003; Tran and Wente 2006). Several models have been proposed to explain the mechanisms of translocation through the NPC. First, the 'virtual gate' model proposes that the highly disordered and mobile FG repeats form an entropic barrier for large molecules (Rout, Aitchison et al. 2003). Kapβ-nucleoporin interactions lower energy barrier to favor passage of Kapβ-substrate complexes. Next, the 'selective phase' model suggests that hydrophobic interactions of FG motifs result in a three-dimensional gel-like meshwork

and Kapβ binding dissolves this structure to allow the translocation (Macara 2001). The third model, which is based on atom force microscopy studies, suggests that FG regions adopt brush-like structures that simultaneously function as both an entropic barrier and a medium with reversible collapse capability that selectively traps for large molecules (Lim, Huang et al. 2006). Despite their differences, all three models are based on Kapβ-nucleoporin FG repeat interactions. Three crystal structures of Impβ-nucleoporin studies have been solved. These structures of Impβ bound to the Nsp1p FxFG repeats (Bayliss, Littlewood et al. 2000), Impβ bound to a synthetic GLFG peptide (DSGGLFGSK; sequence similar to GLFG motifs in GLFG Nups such as Nup116, Nup49, Nup54) (Bayliss, Littlewood et al. 2000), and Kap95p (yeast homolog of Impβ) bound to the Nup1p FG domain (Liu and Stewart 2005), provide molecular details of intermolecular interactions that are critical for translocation through the NPC.

The structure of Impβ residues 1-442 (H1-10) with a fragment of yeast nucleoporin Nsp1p containing five FxFG repeats has been solved (Bayliss, Littlewood et al. 2000). Only two FxFG stretches were seen in the structure, both binding to the convex face of the Impβ N-terminal arch. The primary binding site is between helices H5A and H6A and the secondary one between helices H6A and H7A. Impβ binds Nsp1p mostly through hydrophobic interactions of the two phenylalanines of the FxFG cores (Fig. 4b) (Bayliss, Littlewood et al. 2000).

The crystal structure of an Impβ-GLFG peptide complex shows that a GLFG motif binds Impβ in a similar manner as the FxFG motif (Bayliss, Littlewood et al. 2000). The GLFG

core is buried in a hydrophobic pocket in a similar conformation as the FxFG cores, between the Impβ helices H5A and H6A.

Nup1p binds Kap95p much tighter than other FG-containing nucleoporins (K_D of 7.9 nM for Nup1p-Kap95p vs 1.5 μM for Kap95p-Nup42p) (Pyhtila and Rexach 2003). The central part of Nup1p (residues 333-962) contains 24 FxFG repeats and its C-terminal region (residues 963-1076) has three FG repeats(Pyhtila and Rexach 2003). The C-terminal FG domain of Nup1p, which binds with higher affinity to Kap95p than its central FxFG domain, was used in the Kap95p-Nup1p crystal structure (Liu and Stewart 2005). There are three Nup1p FG binding sites on the convex face of Kap95p between H5, 6, 7 and 8 (Fig. 4a). In all three sites, phenylalanine residues of Nup1p are buried in hydrophobic pockets between the adjacent HEAT repeats. Again, the dominant interactions here are hydrophobic contacts. Nup1p adopts extended conformations two of the three sites and forms a small helix at the third. Interestingly, the linker between two FG motifs also contacts Kap95p. Therefore, linker composition and length may also be critical for interactions with Kapβs. The more extensive FG repeat contact (three versus two) and substantial linker contribution may explain the higher affinity of Kap95p for Nup1p compared to Impβ for FxFG or GLFG (Liu and Stewart 2005).

Although all structures of Impβ bound to nucleoporin FG motifs show binding to the Impβ N-terminal arch, Bednenko and colleagues reported a second nucleoporin binding region at the C-terminus of Impβ (Bednenko, Cingolani et al. 2003). H8-19 of Impβ binds the FG region of Nup153. They presented a model in which the N- and C-terminal

arches of Impβ bind FG domains of nucleooporin in succession to promote movement of transport complexes through the NPC. Thus, although the binding affinity of FG repeats to the C-terminal arch of Impβ is lower than to the N-terminal arch, the C-terminal interactions may be equally important in the translocation process (Bednenko, Cingolani et al. 2003).

Structural analysis of nuclear export

Cse1p: Unliganded versus substrate-bound states

Cse1p contains 20 HEAT repeats. In unliganded Cse1p, H1-20 are organized into a compact ring structure, with H1-3 (in the N-terminal arch) contacting H14-16 (in the C-terminal arch) and H17-20 protruding perpendicular to the plane of the ring (Fig. 5a) (Cook, Fernandez et al. 2005). Short loops connect all helices except in H8 and H19. Helices H8A and H8B are connected by a 29-residue insert of two α-helices whereas H19A and H19B are connected by a long 48-residue loop.

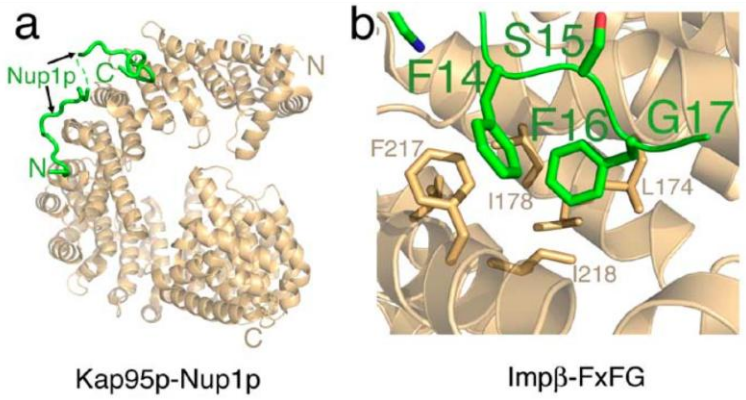

Figure 8-4 Structures of Impβ/Kap95p in complex with nucleoporin fragments. Imp β /Kap95p is represented by ribbon drawings. a) Both proteins in the Kap95- Nup1p

complex (2BPT) are drawn as ribbon representations. Structurally disordered residues in the C- terminal FG domain of Nup1p are represented by dashes. b) The N- terminal fragment of Imp β bound to FxFG repeats of Nsp1p (1F59). Details of a FXFG binding site is shown. Nsp1p residues are labeled with large font and hydrophobic Impβ residues in the binding site are labeled with smaller font.

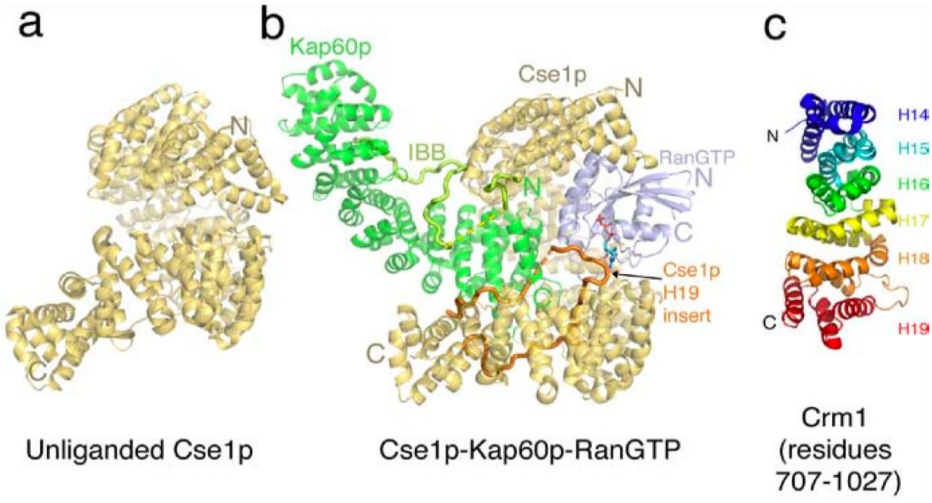

Figure 8-5 Structures of export- Kapβs. a) Unliganded Cse1p (1Z3H). Cse1p is represented by a ribbon drawing. b) The Cse1- Kap60p- RanGTP complex (1WAS). All three proteins are represented in a ribbon diagram. The IBB domain of Kap60p and the H19 insert of Cse1p are shown as thick ribbons and structurally disordered regions are represented with dashes. c) The C- terminal region (H14- H19) of Crm1 (1W9C).

Comparison of unliganded (Cook, Fernandez et al. 2005) and substrate (Kap60p or Impα)-bound Cse1p (Matsuura and Stewart 2004) structures shows dramatic conformational differences. Upon binding Kap60p and RanGTP, the N- and C-terminal arches twist apart and Cse1p opens up into the helicoidal shape that is also adopted by Impβ and Kapβ2 (Fig. 5b) (Matsuura and Stewart 2004). All three proteins in the ternary export complex make extensive interactions with each other, providing rationale for

positive cooperativity between substrate and RanGTP in binding an export-Kapβ (Matsuura and Stewart 2004).

Cse1p-bound Kap60p adopts the autoinhibitory conformation (Kobe 1999; Matsuura and Stewart 2004). Its ARM domain remains structurally invariant while basic stretches in its IBB domain become extended and bind the major and minor NLS sites on the ARM domain. The elongated shape of Kap60p is extended in the Cse1p complex as RanGTP packs against its last ARM repeat (ARM10) (Matsuura and Stewart 2004). Ran and ARM8-10 of Kap60p are sandwiched between the N-and C-terminal arches of Cse1p while ARM1-7 extend away from the Karyopherin (Fig. 5b).

Kap60p contacts Cse1p at several sites including inter-HEAT loops of H2-7 and inter-HEAT loops of H9-12 in the N- and C-terminal arches, respectively (Matsuura and Stewart 2004). The IBB domain also makes extensive contacts with the long internal H19 loop of Cse1p (Fig. 5b). Cse1p-IBB interaction locks Kap60p in its autoinhibited conformation, preventing exogenous NLS from binding, thus ensuring NLS release in the nucleus and export only of empty Kap60p.

Finally, RanGTP contacts Cse1p at two sites (Matsuura and Stewart 2004). The Ran switch 2 region binds H1-3 of Cse1p while its switch 1 region interacts with H13-14 and the long H19 loop of Cse1p. Binding of Cse1p to the switch regions of Ran explains specificity of the export-Kapβ for its GTP state.

In the absence of substrate, RanGTP binds Cse1p with very low affinity (Kutay, Ralf Bischoff et al. 1997). Similarly, Kap60p binds Cse1p with very low affinity when Ran is

absent. Comparison of the two Cse1p structures and mutagenesis show that most of the Ran binding sites are inaccessible when Kap60p is absent (Cook, Fernandez et al. 2005). Similarly, unliganded Cse1p is also incompatible with Kap60p binding as multiple Kap60p binding sites are oriented differently from those in the Cse1-Kap60p-Ran complex. What then is the kinetic mechanism of Kap60p and Ran binding to Cse1p and the accompanying drastic conformational switch? Cook et al suggest that breathing motions may loosen the Cse1p ring to allow transient binding of RanGTP and/or Kap60p (Cook, Fernandez et al. 2005). Interactions of either ligand alone with a subset of their total interaction sites would be insufficient to switch Cse1p from its closed ring conformation to its open helicoidal conformation. They propose that the rare simultaneous occupancy of partially interacting Ran and Kap60p on a Cse1p will be needed to destabilize the Cse1p ring, thus favoring the open conformation.

CRM1: A model for regulation of substrate binding

Even though Crm1 is the most general and versatile export-Kapβ, structural knowledge of this system is limited to a low resolution EM reconstruction of full length unliganded Crm1 and a crystal structure of the C-terminal third of the protein (Petosa, Schoehn et al. 2004). EM single particle analysis resulted in a Crm1 reconstruction of ring-like tubular density, much like the ring-like crystal structure of unliganded Cse1p (Fig. 5) (Petosa, Schoehn et al. 2004; Cook, Fernandez et al. 2005). These finding are consistent with small-angle X-ray scattering (SAXS) data of the unliganded states of Crm1 and Cse1p (Fukuhara, Fernandez et al. 2004).

The crystal structure of full length Crm1 is not yet available, but the crystal structure of a proteolytically stable fragment spanning residues 707-1034 of the 1071-residue Crm1 has been solved (Petosa, Schoehn et al. 2004). This Crm1 fragment forms six HEAT repeats (Fig. 5c). The first five repeats are typical pairs of antiparallel HEAT repeats that wind in a right-handed manner, but the last repeat consists of three similarly sized α-helices that are arranged with a left-handed twist. These six HEAT repeats were denoted H14-19 (Crm1 is predicted to have 19 HEAT repeats) based on homology to a small region of Kapβ2.

Petosa *et al.* also generated homology models of the N-terminal and central regions of Crm1 based on pair-wise sequence alignment with Impβ and Kapβ2(Petosa, Schoehn et al. 2004). They then docked the crystal structure and homology models into the EM map to produce a pseudoatomic model of full length Crm1, which showed that H1 contacts H17-19 to close Crm1 into a ring-like structure. Like Cse1p, the closure of unliganded Crm1 may result in occlusion of Ran binding sites, thus explaining the low intrinsic affinity of Crm1 for RanGTP.

Homology modeling of Crm1 also suggested the presence of a large 65-residue insert or loop in H8 (residues 385-450) (Petosa, Schoehn et al. 2004), much like the acidic H8 loops of Impβ and Kapβ2 or the helical H8 insert of Cse1p (Vetter, Arndt et al. 1999) (Chook and Blobel 1999; Cingolani, Petosa et al. 1999; Lee, Imamoto et al. 2000; Chook, Jung et al. 2002; Cook, Fernandez et al. 2005). Proteolysis and mutagenesis studies are consistent with the prediction of a large H8 loop (Petosa, Schoehn et al. 2004).

Furthermore, disruption of this predicted loop abrogated Ran binding in the presence of substrate. This finding led to the suggestion that the loop may be critical in mediating positive cooperativity between Ran and export substrate, in a manner analogous to the Kapβ2 H8 loop mediating negative cooperativity between Ran and import substrate.

Conclusion

Structures of Impα and Kapβ2 bound to their respective import substrates have confirmed or revealed sequence and structural requirements for recognition of two NLS classes, the classical-NLS and the PY-NLS. In contrast, structures of Impβ-substrate complexes have shown little structural homology in the direct substrates that also bind different Impβ sites. Thus, general features among substrates that bind Impβ directly cannot be inferred at this time. Additional structures of import-Kapβs bound to different classes of transport substrates will inform on the extent of versatile recognition and more importantly will reveal requirements for NLS recognition by the other eight currently unexplored import-Kapβs. Similarly, structures of Ran bound to these other import-Kapβs will be interesting since we already observe that differences in substrate recognition mechanisms is accompanied by differences in their regulation by Ran. Finally, only structures of a very specialized single-substrate export-Kapβ, Cse1p, are available. Therefore, molecular recognition of different NESs, and mechanisms of their assembly and disassembly remain unclear. Structures of other export pathways will reveal the mechanisms of NES recognition, positive cooperativity between NES and

RanGTP for export complex assembly in the nucleus and export substrate dissociation in the cytoplasm.

BIBLIOGRAPHY

Adam, E. J. H. and S. A. Adam (1994). "Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope." Journal of Cell Biology **125**(3): 547-555.

Adams, P. D., P. V. Afonine, et al. (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 213-221.

Aebi, M., M. W. Clark, et al. (1990). "A yeast mutant, PRP20, altered in mRNA metabolism and maintenance of the nuclear structure, is defective in a gene homologous to the human gene RCC1 which is involved in the control of chromosome condensation." Molecular and General Genetics **224**(1): 72-80.

Aguilera, A. (2005). "Cotranscriptional mRNP assembly: From the DNA to the nuclear pore." Current Opinion in Cell Biology **17**(3): 242-250.

Alber, F., S. Dokudovskaya, et al. (2007). "The molecular architecture of the nuclear pore complex." Nature **450**(7170): 695-701.

Alcázar-Román, A. R., E. J. Tran, et al. (2006). "Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export." Nature Cell Biology **8**(7): 711-716.

Allemand, E., S. Dokudovskaya, et al. (2002). "A conserved Drosophila transportin-serine/arginine-rich (SR) protein permits nuclear import of Drosophila SR protein splicing factors and their antagonist repressor splicing factor 1." Mol Biol Cell **13**(7): 2436-2447.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Andrade, M. A. and P. Bork (1995). "HEAT repeats in the Huntington's disease protein." Nature Genetics **11**(2): 115-116.

Aneja, A., W. H. Tang, et al. (2008). "Diabetic cardiomyopathy: insights into pathogenesis, diagnostic challenges, and therapeutic options." Am J Med **121**(9): 748-757.

Anselmino, M., L. Mellbin, et al. (2008). "Early detection and integrated management of dysglycemia in cardiovascular disease: a key factor for decreasing the likelihood of future events." Rev Cardiovasc Med **9**(1): 29-38.

Arnold, M., A. Nath, et al. (2006). "Transportin is a major nuclear import receptor for c-Fos: a novel mode of cargo interaction." J Biol Chem **281**(9): 5492-5499.

Arts, G. J., M. Fornerod, et al. (1998). "Identification of a nuclear export receptor for tRNA." Current Biology **8**(6): 305-314.

Atkinson, N. S., R. W. Dunst, et al. (1985). "Characterization of an essential Saccharomyces cerevisiae gene related to RNA processing: Cloning of RNA1 and generation of a new allele with a novel phenotype." Molecular and Cellular Biology **5**(5): 907-915.

Baake, M., M. Bauerle, et al. (2001). "Core histones and linker histones are imported into the nucleus by different pathways." Eur J Cell Biol **80**(11): 669-677.

Bachi, A., I. C. Braun, et al. (2000). "The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates." RNA **6**(1): 136-158.

Bairoch, A., B. Boeckmann, et al. (2004). "Swiss-Prot: juggling between evolution and stability." Brief Bioinform **5**(1): 39-55.

Bayliss, R., S. W. Leung, et al. (2002). "Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats." Embo Journal **21**(12): 2843-2853.

Bayliss, R., T. Littlewood, et al. (2000). "Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking." Cell **102**(1): 99-108.

Bayliss, R., T. Littlewood, et al. (2002). "GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta." J Biol Chem **277**(52): 50597-50606.

Bear, J., W. Tan, et al. (1999). "Identification of novel import and export signals of human TAP, the protein that binds to the constitutive transport element of the type D retrovirus mRNAs." Mol Cell Biol **19**(9): 6306-6317.

Becker, J., F. Melchior, et al. (1995). "RNA1 encodes a GTPase-activating protein specific for Gsp1p, the Ran/TC4 homologue of Saccharomyces cerevisiae." Journal of Biological Chemistry **270**(20): 11860-11865.

Beddow, A. L., S. A. Richards, et al. (1995). "The Ran/TC4 GTPase-binding domain: Identification by expression cloning and characterization of a conserved sequence motif." Proceedings of the National Academy of Sciences of the United States of America **92**(8): 3328-3332.

Bednenko, J., G. Cingolani, et al. (2003). "Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport." J Cell Biol **162**(3): 391-401.

Belhumeur, P., A. Lee, et al. (1993). "GSP1 and GSP2, genetic suppressors of the prp20-1 mutant in Saccharomyces cerevisiae: GTP-binding proteins involved in the maintenance of nuclear organization." Molecular and Cellular Biology **13**(4): 2152-2161.

Bischoff, F. R. and D. Görlich (1997). "RanBP1 is crucial for the release of RanGTP from importin β-related nuclear transport factors." FEBS Letters **419**(2-3): 249-254.

Bischoff, F. R., C. Klebe, et al. (1994). "RanGAP1 induces GTPase activity of nuclear Ras-related Ran." Proc Natl Acad Sci U S A **91**(7): 2587-2591.

Bischoff, F. R., H. Krebber, et al. (1995). "Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport." Proceedings of the National Academy of Sciences of the United States of America **92**(5): 1749-1753.

Bischoff, F. R. and H. Ponstingl (1991). "Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1." Nature **354**(6348): 80-82.

Bischoff, F. R. and H. Ponstingl (1991). "Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide." Proc Natl Acad Sci U S A **88**(23): 10830-10834.

Bischoff, F. R. and H. Ponstingl (1995). "Catalysis of guanine nucleotide exchange of Ran by RCC1 and stimulation of hydrolysis of Ran-bound GTP by Ran-GAP1." Methods in Enzymology **257**: 135-144.

Bischoff, R., H. Krebber, et al. (1995). "Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1." Embo Journal **14**(4): 705-715.

Bittinger, M. A., E. McWhinnie, et al. (2004). "Activation of cAMP response element-mediated gene expression by regulated nuclear transport of TORC proteins." Curr Biol **14**(23): 2156-2161.

Bogerd, H. P., R. E. Benson, et al. (1999). "Definition of a consensus transportin-specific nucleocytoplasmic transport signal." J Biol Chem **274**(14): 9771-9777.

Bohnsack, M. T., K. Regener, et al. (2002). "Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm." Embo Journal **21**(22): 6205-6215.

Bonifaci, N., J. Moroianu, et al. (1997). "Karyopherin beta2 mediates nuclear import of a mRNA binding protein." Proc Natl Acad Sci U S A **94**(10): 5055-5060.

Bourne, H. R., D. A. Sanders, et al. (1990). "The GTPase superfamily: a conserved switch for diverse cell functions." Nature **348**(6297): 125-132.

Bourne, H. R., D. A. Sanders, et al. (1991). "The GTPase superfamily: conserved structure and molecular mechanism." Nature **349**(6305): 117-127.

Brass, A. L., D. M. Dykxhoorn, et al. (2008). "Identification of host proteins required for HIV infection through a functional genomic screen." Science **319**(5865): 921-926.

Braun, I. C., E. Rohrbach, et al. (1999). "TAP binds to the constitutive transport element (CTE) through a novel RNA-binding motif that is sufficient to promote CTE-dependent RNA export from the nucleus." EMBO J **18**(7): 1953-1965.

Breeuwer, M. and D. S. Goldfarb (1990). "Facilitated nuclear transport of histone H1 and other small nucleophilic proteins." Cell **60**(6): 999-1008.

Brohawn, S. G., J. R. Partridge, et al. (2009). "The nuclear pore complex has entered the atomic age." Structure **17**(9): 1156-1168.

Brown, M. S. and J. L. Goldstein (1997). "The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor." Cell **89**(3): 331-340.

Butler, G. and K. H. Wolfe (1994). "Yeast homologue of mammalian Ran binding protein 1." Biochimica et Biophysica Acta - Gene Structure and Expression **1219**(3): 711-712.

Cai, Y., Y. Gao, et al. (2002). "Characterization and potential function of a novel testis-specific nucleoporin BS-63." Mol Reprod Dev **61**(1): 126-134.

Cansizoglu, A. E. and Y. M. Chook (2007). "Conformational Heterogeneity of Karyopherinβ2 Is Segmental." Structure **15**(11): 1431-1441.

Cansizoglu, A. E., B. J. Lee, et al. (2007). "Structure-based design of a pathway-specific nuclear import inhibitor." Nat Struct Mol Biol **14**(5): 452-454.

Carmody, S. R. and S. R. Wente (2009). "mRNA nuclear export at a glance." J Cell Sci **122**(Pt 12): 1933-1937.

Chakraborty, P., N. Satterly, et al. (2006). "Nuclear export assays for poly(A) RNAs." Methods **39**(4): 363-369.

Chan, C. K., S. Hübner, et al. (1998). "Mutual exclusivity of DNA binding and nuclear localization signal recognition by the yeast transcription factor GAL4: Implications for nonviral DNA delivery." Gene Therapy **5**(9): 1204-1212.

Chen, M. H., I. Ben-Efraim, et al. (2005). "Phospholipid scramblase 1 contains a nonclassical nuclear localization signal with unique binding site in importin alpha." J Biol Chem **280**(11): 10599-10606.

Cheng, H., K. Dufu, et al. (2006). "Human mRNA Export Machinery Recruited to the 5′ End of mRNA." Cell **127**(7): 1389-1400.

Chenna, R., H. Sugawara, et al. (2003). "Multiple sequence alignment with the Clustal series of programs." Nucleic Acids Res **31**(13): 3497-3500.

Chi, N. C., E. J. H. Adam, et al. (1995). "Sequence and characterization of cytoplasmic nuclear protein import factor p97." Journal of Cell Biology **130**(2): 265-274.

Chi, N. C., E. J. H. Adam, et al. (1997). "Different binding domains for Ran-GTP and Ran-GDP/RanBP1 on nuclear import factor p97." Journal of Biological Chemistry **272**(10): 6818-6822.

Chi, N. C. and S. A. Adam (1997). "Functional domains in nuclear import factor p97 for binding the nuclear localization sequence receptor and the nuclear pore." Molecular Biology of the Cell **8**(6): 945-956.

Chook, Y. M. and G. Blobel (1999). "Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp." Nature **399**(6733): 230-237.

Chook, Y. M. and G. Blobel (2001). "Karyopherins and nuclear import." Curr Opin Struct Biol **11**(6): 703-715.

Chook, Y. M., G. Cingolani, et al. (1999). "Pictures in cell biology structures of nuclear-transport components." Trends in Cell Biology **9**(8): 310-311.

Chook, Y. M., A. Jung, et al. (2002). "Uncoupling Kapbeta2 substrate dissociation and ran binding." Biochemistry **41**(22): 6955-6966.

Chook, Y. M. and K. E. Suel (2010). "Nuclear import by karyopherin-betas: Recognition and inhibition." Biochim Biophys Acta.

Cingolani, G., J. Bednenko, et al. (2002). "Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta." Mol Cell **10**(6): 1345-1353.

Cingolani, G., H. A. Lashuel, et al. (2000). "Nuclear import factors importin α and importin β undergo mutually induced conformational changes upon association." FEBS Letters **484**(3): 291-298.

Cingolani, G., C. Petosa, et al. (1999). "Structure of importin-beta bound to the IBB domain of importin-alpha." Nature **399**(6733): 221-229.

Clarkson, W. D., H. M. Kent, et al. (1996). "Separate binding sites on nuclear transport factor 2 (NTF2) for GDP-Ran and the phenylalanine-rich repeat regions of nucleoporins p62 and Nsp1p." Journal of Molecular Biology **263**(4): 517-524.

Clemens, T. L., S. Cormier, et al. (2001). "Parathyroid hormone-related protein and its receptors: nuclear functions and roles in the renal and cardiovascular systems, the placental trophoblasts and the pancreatic islets." Br J Pharmacol **134**(6): 1113-1136.

Cole, C. N. and J. J. Scarcelli (2006). "Transport of messenger RNA from the nucleus to the cytoplasm." Curr Opin Cell Biol **18**(3): 299-306.

Conkright, M. D., G. Canettieri, et al. (2003). "TORCs: transducers of regulated CREB activity." Mol Cell **12**(2): 413-423.

Conti, E. and E. Izaurralde (2001). "Nucleocytoplasmic transport enters the atomic age." Curr Opin Cell Biol **13**(3): 310-319.

Conti, E. and J. Kuriyan (2000). "Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha." Structure **8**(3): 329-338.

Conti, E., C. W. Muller, et al. (2006). "Karyopherin flexibility in nucleocytoplasmic transport." Curr Opin Struct Biol **16**(2): 237-244.

Conti, E., M. Uy, et al. (1998). "Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha." Cell **94**(2): 193-204.

Cook, A., F. Bono, et al. (2007). "Structural biology of nucleocytoplasmic transport." Annu Rev Biochem **76**: 647-671.

Cook, A., E. Fernandez, et al. (2005). "The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding." Mol Cell **18**(3): 355-367.

Cook, A. G. and E. Conti (2010). "Nuclear export complexes in the frame." Curr Opin Struct Biol **20**(2): 247-252.

Corbett, A. H., D. M. Koepp, et al. (1995). "Rna1p, a Ran/TC4 GTPase activating protein, is required for nuclear import." Journal of Cell Biology **130**(5): 1017-1026.

Corbett, A. H. and P. A. Silver (1996). "The NTF2 gene encodes an essential, highly conserved protein that functions in nuclear transport in vivo." Journal of Biological Chemistry **271**(31): 18477-18484.

Coutavas, E., M. Ren, et al. (1993). "Characterization of proteins that interact with the cell-cycle regulatory protein Ran/TC4." Nature **366**(6455): 585-587.

D'Angelo, M. A. and M. W. Hetzer (2006). "The role of the nuclear envelope in cellular organization." Cell Mol Life Sci **63**(3): 316-332.

D'Angelo, M. A. and M. W. Hetzer (2008). "Structure, dynamics and function of nuclear pore complexes." Trends Cell Biol **18**(10): 456-466.

Damelin, M., P. A. Silver, et al. (2002). "Nuclear protein transport." Methods Enzymol **351**: 587-607.

Datar, K. V., G. Dreyfuss, et al. (1993). "The human hnRNP M proteins: identification of a methionine/arginine-rich repeat motif in ribonucleoproteins." Nucleic Acids Res **21**(3): 439-446.

DeGrasse, J. A., K. N. DuBois, et al. (2009). "Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor." Mol Cell Proteomics **8**(9): 2119-2130.

Denning, D. P., S. S. Patel, et al. (2003). "Disorder in the nuclear pore complex: The FG repeat regions of nucleoporins are natively unfolded." Proceedings of the National Academy of Sciences of the United States of America **100**(5): 2450-2455.

Devos, D., S. Dokudovskaya, et al. (2004). "Components of coated vesicles and nuclear pore complexes share a common molecular architecture." PLoS Biology **2**(12).

Devos, D., S. Dokudovskaya, et al. (2006). "Simple fold composition and modular architecture of the nuclear pore complex." Proceedings of the National Academy of Sciences of the United States of America **103**(7): 2172-2177.

Dhavan, R. and L. H. Tsai (2001). "A decade of CDK5." Nat Rev Mol Cell Biol **2**(10): 749-759.

Dingwall, C. and R. A. Laskey (1991). "Nuclear targeting sequences--a consensus?" Trends Biochem Sci **16**(12): 478-481.

Dingwall, C., S. V. Sharnick, et al. (1982). "A polypeptide domain that specifies migration of nucleoplasmin into the nucleus." Cell **30**(2): 449-458.

Dokken, B. B. (2008). "The pathophysiology of cardiovascular disease and diabetes: Beyond blood pressure and lipids." Diabetes Spetrum **21**: 160-165.

Dong, X., A. Biswas, et al. (2009). "Structural basis for assembly and disassembly of the CRM1 nuclear export complex." Nat Struct Mol Biol **16**(5): 558-560.

Dong, X., A. Biswas, et al. (2009). "Structural basis for leucine-rich nuclear export signal recognition by CRM1." Nature **458**(7242): 1136-1141.

Dormann, D., R. Rodde, et al. (2010). "ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import." EMBO J **29**(16): 2841-2857.

Drivas, G. T., A. Shih, et al. (1990). "Characterization of four novel ras-like genes expressed in a human teratocarcinoma cell line." Molecular and Cellular Biology **10**(4): 1793-1798.

Dunn, C. W., A. Hejnol, et al. (2008). "Broad phylogenomic sampling improves resolution of the animal tree of life." Nature **452**(7188): 745-749.

Emsley, P., B. Lohkamp, et al. (2010). "Features and development of Coot." Acta Crystallogr D Biol Crystallogr **66**(Pt 4): 486-501.

Enenkel, C., G. Blobel, et al. (1995). "Identification of a yeast karyopherin heterodimer that targets import substrate to mammalian nuclear pore complexes." J Biol Chem **270**(28): 16499-16502.

Englmeier, L., J. C. Olivo, et al. (1999). "Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis." Curr Biol **9**(1): 30-41.

Epstein, F. H. (1967). "Hyperglycemia. A risk factor in coronary heart disease." Circulation **36**(4): 609-619.

Erkmann, J. A. and U. Kutay (2004). "Nuclear export of mRNA: from the site of transcription to the cytoplasm." Exp Cell Res **296**(1): 12-20.

Erkmann, J. A., R. Sanchez, et al. (2005). "Nuclear export of metazoan replication-dependent histone mRNAs is dependent on RNA length and is mediated by TAP." RNA **11**(1): 45-58.

Fagotto, F., U. Gluck, et al. (1998). "Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin." Curr Biol **8**(4): 181-190.

Fahrenkrog, B. and U. Aebi (2003). "The nuclear pore complex: nucleocytoplasmic transport and beyond." Nat Rev Mol Cell Biol **4**(10): 757-766.

Fan, F., C. P. Liu, et al. (1997). "cDNA cloning and characterization of Npap60: a novel rat nuclear pore-associated protein with an unusual subcellular localization during male germ cell differentiation." Genomics **40**(3): 444-453.

Fan, X. C. and J. A. Steitz (1998). "HNS, a nuclear-cytoplasmic shuttling sequence in HuR." Proc Natl Acad Sci U S A **95**(26): 15293-15298.

Fan, X. C. and J. A. Steitz (1998). "Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs." EMBO J **17**(12): 3448-3460.

Fanara, P., M. R. Hodel, et al. (2000). "Quantitative analysis of nuclear localization signal (NLS)-importin α interaction through fluorescence depolarization: Evidence for auto-inhibitory regulation of NLS binding." Journal of Biological Chemistry **275**(28): 21218-21223.

Fischer, U., J. Huber, et al. (1995). "The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs." Cell **82**(3): 475-483.

Floer, M., G. Blobel, et al. (1997). "Disassembly of RanGTP-karyopherin beta complex, an intermediate in nuclear protein import." J Biol Chem **272**(31): 19538-19546.

Fontes, M. R., T. Teh, et al. (2003). "Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha." J Biol Chem **278**(30): 27981-27987.

Fontes, M. R., T. Teh, et al. (2000). "Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha." J Mol Biol **297**(5): 1183-1194.

Fontes, M. R., T. Teh, et al. (2003). "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 **375**(Pt 2): 339-349.

Fornerod, M., M. Ohno, et al. (1997). "CRM1 is an export receptor for leucine-rich nuclear export signals." Cell **90**(6): 1051-1060.

Forwood, J. K., M. H. Lam, et al. (2001). "Nuclear import of Creb and AP-1 transcription factors requires importin-beta 1 and Ran but is independent of importin-alpha." Biochemistry **40**(17): 5208-5217.

Fribourg, S., I. C. Braun, et al. (2001). "Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor." Mol Cell **8**(3): 645-656.

Fribourg, S. and E. Conti (2003). "Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15." EMBO Rep **4**(7): 699-703.

Fridell, R. A., R. Truant, et al. (1997). "Nuclear import of hnRNP A1 is mediated by a novel cellular cofactor related to karyopherin-beta." J Cell Sci **110 (Pt 11)**: 1325-1331.

Fried, H. and U. Kutay (2003). "Nucleocytoplasmic transport: Taking an inventory." Cellular and Molecular Life Sciences **60**(8): 1659-1688.

Fu, X., Y. K. Choi, et al. (2006). "Identification of nuclear import mechanisms for the neuronal Cdk5 activator." J Biol Chem **281**(51): 39014-39021.

Fukuda, M., S. Asano, et al. (1997). "CRM1 is responsible for intracellular transport mediated by the nuclear export signal." Nature **390**(6657): 308-311.

Fukuhara, N., E. Fernandez, et al. (2004). "Conformational variability of nucleo-cytoplasmic transport factors." J Biol Chem **279**(3): 2176-2181.

Gattiker, A., E. Gasteiger, et al. (2002). "ScanProsite: a reference implementation of a PROSITE scanning tool." Appl Bioinformatics **1**(2): 107-108.

Gong, X., X. Tang, et al. (2003). "Cdk5-mediated inhibition of the protective effects of transcription factor MEF2 in neurotoxicity-induced apoptosis." Neuron **38**(1): 33-46.

Gorlich, D., M. Dabrowski, et al. (1997). "A novel class of RanGTP binding proteins." J Cell Biol **138**(1): 65-80.

Görlich, D., P. Henklein, et al. (1996). "A 41 amino acid motif in importin-α confers binding to importin-β and hence transit into the nucleus." Embo Journal **15**(8): 1810-1817.

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

Görlich, D., N. Pantéé, et al. (1996). "Identification of different roles for RanGDP and RanGTP in nuclear protein import." Embo Journal **15**(20): 5584-5594.

Görlich, D., F. Vogel, et al. (1995). "Distinct functions for the two importin subunits in nuclear protein import." Nature **377**(6546): 246-248.

Grant, R. P., E. Hurt, et al. (2002). "Structure of the C-terminal FG-nucleoporin binding domain of Tap/NXF1." Nat Struct Biol **9**(4): 247-251.

Grant, R. P., D. Neuhaus, et al. (2003). "Structural basis for the interaction between the Tap/NXF1 UBA domain and FG nucleoporins at 1A resolution." J Mol Biol **326**(3): 849-858.

Griffis, E. R., B. Craige, et al. (2004). "Distinct functional domains within nucleoporins Nup153 and Nup98 mediate transcription-dependent mobility." Mol Biol Cell **15**(4): 1991-2002.

Groves, M. R., N. Hanlon, et al. (1999). "The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs." Cell **96**(1): 99-110.

Grundmann, U., C. Nerlich, et al. (1988). "Isolation of cDNA coding for the placental protein 15 (PP15)." Nucleic Acids Res **16**(10): 4721.

Gruss, O. J., R. E. Carazo-Salas, et al. (2001). "Ran induces spindle assembly by reversing the inhibitory effect of importin α on TPX2 activity." Cell **104**(1): 83-93.

Grüter, P., C. Tabernero, et al. (1998). "TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus." Molecular Cell **1**(5): 649-659.

Guttinger, S., P. Muhlhausser, et al. (2004). "Transportin2 functions as importin and mediates nuclear import of HuR." Proc Natl Acad Sci U S A **101**(9): 2918-2923.

Hamelberg, D., T. Shen, et al. (2007). "A proposed signaling motif for nuclear import in mRNA processing via the formation of arginine claw." Proc Natl Acad Sci U S A **104**(38): 14947-14951.

Harel, A., R. C. Chan, et al. (2003). "Importin β Negatively Regulates Nuclear Membrane Fusion and Nuclear Pore Complex Assembly." Molecular Biology of the Cell **14**(11): 4387-4396.

Hautbergue, G. M., M. L. Hung, et al. (2008). "Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP." Proc Natl Acad Sci U S A **105**(13): 5154-5159.

Haynes, C. and L. M. Iakoucheva (2006). "Serine/arginine-rich splicing factors belong to a class of intrinsically disordered proteins." Nucleic Acids Res **34**(1): 305-312.

Hemmings, B. A., C. Adams-Pearson, et al. (1990). "alpha- and beta-forms of the 65-kDa subunit of protein phosphatase 2A have a similar 39 amino acid repeating structure." Biochemistry **29**(13): 3166-3173.

Herold, A., T. Klymenko, et al. (2001). "NXF1/p15 heterodimers are essential for mRNA nuclear export in Drosophila." RNA **7**(12): 1768-1780.

Herold, A., M. Suyama, et al. (2000). "TAP (NXF1) belongs to a multigene family of putative RNA export factors with a conserved modular architecture." Mol Cell Biol **20**(23): 8996-9008.

Herold, A., R. Truant, et al. (1998). "Determination of the functional domain organization of the importin α nuclear import factor." Journal of Cell Biology **143**(2): 309-318.

Ho, D. N., G. A. Coburn, et al. (2002). "The crystal structure and mutational analysis of a novel RNA-binding domain found in the human Tap nuclear mRNA export factor." Proc Natl Acad Sci U S A **99**(4): 1888-1893.

Huang, Y., R. Gattoni, et al. (2003). "SR splicing factors serve as adapter proteins for TAP-dependent mRNA export." Mol Cell **11**(3): 837-843.

Huber, A. H., W. J. Nelson, et al. (1997). "Three-dimensional structure of the armadillo repeat region of β-catenin." Cell **90**(5): 871-882.

Huber, J., U. Cronshagen, et al. (1998). "Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure." Embo Journal **17**(14): 4114-4126.

Hutten, S., S. Walde, et al. (2009). "The nuclear pore component Nup358 promotes transportin-dependent nuclear import." J Cell Sci **122**(Pt 8): 1100-1110.

Iijima, M., M. Suzuki, et al. (2006). "Two motifs essential for nuclear import of the hnRNP A1 nucleocytoplasmic shuttling sequence M9 core." FEBS Lett **580**(5): 1365-1370.

Imamoto, N., T. Tachibana, et al. (1995). "A karyophilic protein forms a stable complex with cytoplasmic components prior to nuclear pore binding." Journal of Biological Chemistry **270**(15): 8559-8565.

Imasaki, T., T. Shimizu, et al. (2007). "Structural basis for substrate recognition and dissociation by human transportin 1." Mol Cell **28**(1): 57-67.

Ino, H. and T. Chiba (1996). "Intracellular localization of cyclin-dependent kinase 5 (CDK5) in mouse neuron: CDK5 is located in both nucleus and cytoplasm." Brain Res **732**(1-2): 179-185.

Iourgenko, V., W. Zhang, et al. (2003). "Identification of a family of cAMP response element-binding protein coactivators by genome-scale functional analysis in mammalian cells." Proc Natl Acad Sci U S A **100**(21): 12147-12152.

Isgro, T. A. and K. Schulten (2005). "Binding dynamics of isolated nucleoporin repeat regions to importin-beta." Structure **13**(12): 1869-1879.

Izaurralde, E., U. Kutay, et al. (1997). "The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus." Embo Journal **16**(21): 6535-6547.

Izaurralde, E. and I. W. Mattaj (1995). "RNA export." Cell **81**(2): 153-159.

Jäkel, S., W. Albig, et al. (1999). "The importin β/importin 7 heterodimer is a functional nuclear import receptor for histone H1." Embo Journal **18**(9): 2411-2423.

Jäkel, S. and D. Görlich (1998). "Importin β, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells." Embo Journal **17**(15): 4491-4502.

Jansson, D., A. C. Ng, et al. (2008). "Glucose controls CREB activity in islet cells via regulated phosphorylation of TORC2." Proc Natl Acad Sci U S A **105**(29): 10161-10166.

Kadowaki, T., D. Goldfarb, et al. (1993). "Regulation of RNA processing and transport by a nuclear guanine nucleotide release protein and members of the Ras superfamily." Embo Journal **12**(7): 2929-2937.

Kalderon, D., W. D. Richardson, et al. (1984). "Sequence requirements for nuclear location of simian virus 40 large-T antigen." Nature **311**(5981): 33-38.

Kalderon, D., B. L. Roberts, et al. (1984). "A short amino acid sequence able to specify nuclear location." Cell **39**(3 Pt 2): 499-509.

Kang, Y. and B. R. Cullen (1999). "The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences." Genes Dev **13**(9): 1126-1139.

Katahira, J., K. Strasser, et al. (1999). "The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human." EMBO J **18**(9): 2593-2609.

Kataoka, N., J. L. Bachorik, et al. (1999). "Transportin-SR, a nuclear import receptor for SR proteins." J Cell Biol **145**(6): 1145-1152.

Katoh, Y., H. Takemori, et al. (2006). "Silencing the constitutive active transcription factor CREB by the LKB1-SIK signaling cascade." FEBS J **273**(12): 2730-2748.

Kawamura, H., Y. Tomozoe, et al. (2002). "Identification of the nucleocytoplasmic shuttling sequence of heterogeneous nuclear ribonucleoprotein D-like protein JKTBP and its interaction with mRNA." J Biol Chem **277**(4): 2732-2739.

Kaye, F. J. unpublished data.

Kehlenbach, R. H., A. Dickmanns, et al. (1999). "A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export." Journal of Cell Biology **145**(4): 645-657.

Kelly, S. M. and A. H. Corbett (2009). "Messenger RNA export from the nucleus: a series of molecular wardrobe changes." Traffic **10**(9): 1199-1208.

Klebe, C., H. Prinz, et al. (1995). "The kinetic mechanism of Ran-nucleotide exchange catalyzed by RCC1." Biochemistry **34**(39): 12543-12552.

Klebe, C., F. Ralf Bischoff, et al. (1995). "Interaction of the nuclear GTP-binding protein ran with its regulatory proteins RCC1 and ranGAP." Biochemistry **34**(2): 639-647.

Kobe, B. (1999). "Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha." Nat Struct Biol **6**(4): 388-397.

Kohler, A. and E. Hurt (2007). "Exporting RNA from the nucleus to the cytoplasm." Nat Rev Mol Cell Biol **8**(10): 761-773.

Kuersten, S., M. Ohno, et al. (2001). "Nucleocytoplasmic transport: Ran, beta and beyond." Trends Cell Biol **11**(12): 497-503.

Kutay, U., E. Izaurralde, et al. (1997). "Dominant-negative mutants of importin-β block multiple pathways of import and export through the nuclear pore complex." Embo Journal **16**(6): 1153-1163.

Kutay, U., G. Lipowsky, et al. (1998). "Identification of a tRNA-specific nuclear export receptor." Molecular Cell **1**(3): 359-369.

Kutay, U., F. Ralf Bischoff, et al. (1997). "Export of importin α from the nucleus is mediated by a specific nuclear transport factor." Cell **90**(6): 1061-1071.

Lai, M. C., H. W. Kuo, et al. (2003). "A novel splicing regulator shares a nuclear import pathway with SR proteins." EMBO J **22**(6): 1359-1369.

Lai, M. C., R. I. Lin, et al. (2000). "A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins." J Biol Chem **275**(11): 7950-7957.

Lai, M. C., R. I. Lin, et al. (2001). "Transportin-SR2 mediates nuclear import of phosphorylated SR proteins." Proc Natl Acad Sci U S A **98**(18): 10154-10159.

Lam, M. H. C., L. J. Briggs, et al. (1999). "Importin β recognizes parathyroid hormone-related protein with high affinity and mediates its nuclear import in the absence of importin α." Journal of Biological Chemistry **274**(11): 7391-7398.

Lange, A., R. E. Mills, et al. (2008). "A PY-NLS nuclear targeting signal is required for nuclear localization and function of the Saccharomyces cerevisiae mRNA-binding protein Hrp1." J Biol Chem **283**(19): 12926-12934.

Lee, B. J., A. E. Cansizoglu, et al. (2006). "Rules for nuclear localization sequence recognition by karyopherin beta 2." Cell **126**(3): 543-558.

Lee, S. J., N. Imamoto, et al. (2000). "The adoption of a twisted structure of importin-β is essential for the protein-protein interaction required for nuclear transport." Journal of Molecular Biology **302**(1): 251-264.

Lee, S. J., Y. Matsuura, et al. (2005). "Structural basis for nuclear import complex dissociation by RanGTP." Nature **435**(7042): 693-696.

Lee, S. J., T. Sekimoto, et al. (2003). "The structure of importin-beta bound to SREBP-2: nuclear import of a transcription factor." Science **302**(5650): 1571-1575.

Lee, S. J., T. Sekimoto, et al. (2003). "Crystallization and preliminary crystallographic analysis of the importin-beta-SREBP-2 complex." Acta Crystallogr D Biol Crystallogr **59**(Pt 10): 1866-1868.

Liker, E., E. Fernandez, et al. (2000). "The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain." EMBO J **19**(21): 5587-5598.

Lim, A. C., D. Qu, et al. (2003). "Protein-protein interactions in Cdk5 regulation and function." Neurosignals **12**(4-5): 230-238.

Lim, R. Y. and B. Fahrenkrog (2006). "The nuclear pore complex up close." Curr Opin Cell Biol **18**(3): 342-347.

Lim, R. Y., K. S. Ullman, et al. (2008). "Biology and biophysics of the nuclear pore complex and its components." Int Rev Cell Mol Biol **267**: 299-342.

Lim, R. Y. H., N. P. Huang, et al. (2006). "Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport." Proceedings of the National Academy of Sciences of the United States of America **103**(25): 9512-9517.

Linding, R., L. J. Jensen, et al. (2003). "Protein disorder prediction: implications for structural proteomics." Structure **11**(11): 1453-1459.

Liu, S. M. and M. Stewart (2005). "Structural basis for the high-affinity binding of nucleoporin Nup1p to the Saccharomyces cerevisiae importin-beta homologue, Kap95p." J Mol Biol **349**(3): 515-525.

Liu, Y., R. Dentin, et al. (2008). "A fasting inducible switch modulates gluconeogenesis via activator/coactivator exchange." Nature **456**(7219): 269-273.

Lounsbury, K. M. and L. G. Macara (1997). "Ran-binding protein 1 (RanBP1) forms a ternary complex with ran and karyopherin β and reduces ran GTPase-activating protein (RanGAP) inhibition by karyopherin β." Journal of Biological Chemistry **272**(1): 551-555.

Lund, M. K. and C. Guthrie (2005). "The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim." Mol Cell **20**(4): 645-651.

Macara, I. G. (2001). "Transport into and out of the nucleus." Microbiology and Molecular Biology Reviews **65**(4): 570-594.

Madrid, A. S. and K. Weis (2006). "Nuclear transport is becoming crystal clear." Chromosoma **115**(2): 98-109.

Mahajan, R., C. Delphin, et al. (1997). "A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2." Cell **88**(1): 97-107.

Malik, H. S., T. H. Eickbush, et al. (1997). "Evolutionary specialization of the nuclear targeting apparatus." Proceedings of the National Academy of Sciences of the United States of America **94**(25): 13738-13742.

Marfori, M., A. Mynott, et al. (2010). "Molecular basis for specificity of nuclear import and prediction of nuclear localization." Biochim Biophys Acta.

Masuda, S., R. Das, et al. (2005). "Recruitment of the human TREX complex to mRNA during splicing." Genes and Development **19**(13): 1512-1517.

Matsuura, Y., A. Lange, et al. (2003). "Structural basis for Nup2p function in cargo release and karyopherin recycling in nuclear import." Embo Journal **22**(20): 5358-5369.

Matsuura, Y. and M. Stewart (2004). "Structural basis for the assembly of a nuclear export complex." Nature **432**(7019): 872-877.

Matsuura, Y. and M. Stewart (2005). "Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling." Embo Journal **24**(21): 3681-3689.

Matunis, M. J., E. Coutavas, et al. (1996). "A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex." Journal of Cell Biology **135**(6): 1457-1470.

Matunis, M. J., E. L. Matunis, et al. (1992). "Isolation of hnRNP complexes from Drosophila melanogaster." J Cell Biol **116**(2): 245-255.

Michael, W. M., M. Choi, et al. (1995). "A nuclear export signal in hnRNP A1: A signal-mediated, temperature- dependent nuclear protein export pathway." Cell **83**(3): 415-422.

Mitrousis, G., A. S. Olia, et al. (2008). "Molecular basis for the recognition of snurportin 1 by importin beta." J Biol Chem **283**(12): 7877-7884.

Moore, J. D., J. Yang, et al. (1999). "Nuclear import of Cdk/cyclin complexes: Identification of distinct mechanisms for import of Cdk2/cyclin E and Cdc2/cyclin B1." Journal of Cell Biology **144**(2): 213-224.

Moore, M. S. and G. Blobel (1994). "Purification of a ran-interacting protein that is required for protein import into the nucleus." Proceedings of the National Academy of Sciences of the United States of America **91**(21): 10212-10216.

Moroianu, J., G. Blobel, et al. (1996). "The binding site of karyopherin alpha for karyopherin beta overlaps with a nuclear localization sequence." Proc Natl Acad Sci U S A **93**(13): 6572-6576.

Mosammaparast, N. and L. F. Pemberton (2004). "Karyopherins: from nuclear-transport mediators to nuclear-function regulators." Trends Cell Biol **14**(10): 547-556.

Mühlhäusser, P., E. C. Müller, et al. (2001). "Multiple pathways contribute to nuclear import of core histones." EMBO Reports **2**(8): 690-696.

Nachury, M. V., T. J. Maresca, et al. (2001). "Importin β is a mitotic target of the small GTPase ran in spindle assembly." Cell **104**(1): 95-106.

Nagoshi, E., N. Imamoto, et al. (1999). "Nuclear import of sterol regulatory element-binding protein-2, a basic helix-loop-helix-leucine zipper (bHLH-Zip)-containing transcription factor, occurs through the direct interaction of importin beta with HLH-Zip." Mol Biol Cell **10**(7): 2221-2233.

Nagoshi, E. and Y. Yoneda (2001). "Dimerization of sterol regulatory element-binding protein 2 via the helix-loop-helix-leucine zipper domain is a prerequisite for its nuclear localization mediated by importin beta." Mol Cell Biol **21**(8): 2779-2789.

Nakielny, S., M. C. Siomi, et al. (1996). "Transportin: nuclear transport receptor of a novel nuclear protein import pathway." Exp Cell Res **229**(2): 261-266.

Nehrbass, U. and G. Blobel (1996). "Role of the nuclear transport factor p10 in nuclear import." Science **272**(5258): 120-122.

Neville, M., F. Stutz, et al. (1997). "The importin-beta family member Crm1 p bridges the interaction between Rev and the nuclear pore complex during nuclear export." Current Biology **7**(10): 767-775.

Ngo, J. C., K. Giang, et al. (2008). "A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1." Mol Cell **29**(5): 563-576.

Nikolic, M., H. Dudek, et al. (1996). "The cdk5/p35 kinase is essential for neurite outgrowth during neuronal differentiation." Genes Dev **10**(7): 816-825.

Nilsson, J., K. Weis, et al. (2002). "The C-terminal extension of the small GTPase Ran is essential for defining the GDP-bound form." J Mol Biol **318**(2): 583-593.

Nishinaka, Y., H. Masutani, et al. (2004). "Importin alpha1 (Rch1) mediates nuclear translocation of thioredoxin-binding protein-2/vitamin D(3)-up-regulated protein 1." J Biol Chem **279**(36): 37559-37565.

Ohtsubo, M., R. Kai, et al. (1987). "Isolation and characterization of the active cDNA of the human cell cycle gene (RCC1) involved in the regulation of onset of chromosome condensation." Genes & development **1**(6): 585-593.

Ohtsubo, M., H. Okazaki, et al. (1989). "The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA." Journal of Cell Biology **109**(4 I): 1389-1397.

Olsson, M., S. Scheele, et al. (2004). "Limited expression of nuclear pore membrane glycoprotein 210 in cell lines and tissues suggests cell-type specific nuclear pores in metazoans." Exp Cell Res **292**(2): 359-370.

Ossareh-Nazari, B., F. Bachelerie, et al. (1997). "Evidence for a role of CRM1 in signal-mediated nuclear protein export." Science **278**(5335): 141-144.

Paine, P. L., L. C. Moore, et al. (1975). "Nuclear envelope permeability." Nature **254**(5496): 109-114.

Palacios, I., M. Hetzer, et al. (1997). "Nuclear import of U snRNPs requires importin β." Embo Journal **16**(22): 6783-6792.

Paraskeva, E., E. Izaurralde, et al. (1999). "CRM1-mediated recycling of snurportin 1 to the cytoplasm." Journal of Cell Biology **145**(2): 255-264.

Paschal, B. M., C. Delphin, et al. (1996). "Nucleotide-specific interaction of Ran/TC4 with nuclear transport factors NTF2 and p97." Proceedings of the National Academy of Sciences of the United States of America **93**(15): 7679-7683.

Paschal, B. M. and L. Gerace (1995). "Identification of NTF2, a cytosolic factor for nuclear import that interacts with nuclear pore complex protein p62." Journal of Cell Biology **129**(4): 925-937.

Peifer, M., S. Berg, et al. (1994). "A repeating amino acid motif shared by proteins with diverse cellular roles." Cell **76**(5): 789-791.

Peifer, M., P. D. McCrea, et al. (1992). "The vertebrate adhesive junction proteins beta-catenin and plakoglobin and the Drosophila segment polarity gene armadillo form a multigene family with similar properties." J Cell Biol **118**(3): 681-691.

Pemberton, L. F., G. Blobel, et al. (1998). "Transport routes through the nuclear pore complex." Curr Opin Cell Biol **10**(3): 392-399.

Pemberton, L. F. and B. M. Paschal (2005). "Mechanisms of receptor-mediated nuclear import and nuclear export." Traffic **6**(3): 187-198.

Peng, S. S., C. Y. Chen, et al. (1998). "RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein." EMBO J **17**(12): 3461-3470.

Peters, R. (2005). "Translocation through the nuclear pore complex: Selectivity and speed by reduction-of-dimensionality." Traffic **6**(5): 421-427.

Petosa, C., G. Schoehn, et al. (2004). "Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex." Mol Cell **16**(5): 761-775.

Pollard, V. W., W. M. Michael, et al. (1996). "A novel receptor-mediated nuclear protein import pathway." Cell **86**(6): 985-994.

Pyhtila, B. and M. Rexach (2003). "A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex." J Biol Chem **278**(43): 42699-42709.

Qu, D., Q. Li, et al. (2002). "The protein SET binds the neuronal Cdk5 activator p35nck5a and modulates Cdk5/p35nck5a activity." J Biol Chem **277**(9): 7324-7332.

Quan, Y., Z. L. Ji, et al. (2008). "Evolutionary and transcriptional analysis of karyopherin beta superfamily proteins." Mol Cell Proteomics **7**(7): 1254-1269.

Rabut, G., V. Doye, et al. (2004). "Mapping the dynamic organization of the nuclear pore complex inside single living cells." Nature Cell Biology **6**(11): 1114-1121.

Radu, A., G. Blobel, et al. (1995). "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 **92**(5): 1769-1773.

Rebane, A., A. Aab, et al. (2004). "Transportins 1 and 2 are redundant nuclear import factors for hnRNP A1 and HuR." RNA **10**(4): 590-599.

Reed, R. and H. Cheng (2005). "TREX, SR proteins and export of mRNA." Curr Opin Cell Biol **17**(3): 269-273.

Reichelt, R., A. Holzenburg, et al. (1990). "Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components." J Cell Biol **110**(4): 883-894.

Renault, L., J. Kuhlmann, et al. (2001). "Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1)." Cell **105**(2): 245-255.

Ribbeck, K. and D. Görlich (2001). "Kinetic analysis of translocation through nuclear pore complexes." Embo Journal **20**(6): 1320-1330.

Ribbeck, K., U. Kutay, et al. (1999). "The translocation of transportin-cargo complexes through nuclear pores is independent of both Ran and energy." Current Biology **9**(1): 47-50.

Ribbeck, K., G. Lipowsky, et al. (1998). "NTF2 mediates nuclear import of Ran." EMBO J **17**(22): 6587-6598.

Richards, S. A., K. M. Lounsbury, et al. (1996). "A nuclear export signal is essential for the cytosolic localization of the Ran binding protein, RanBP1." J Cell Biol **134**(5): 1157-1168.

Richards, S. A., K. M. Lounsbury, et al. (1995). "The C terminus of the nuclear Ran/TC4 GTPase stabilizes the GDP-bound state and mediates interactions with RCC1, Ran-GAP, and HTF9A/RanBP1." Journal of Biological Chemistry **270**(24): 14405-14411.

Riggleman, B., E. Wieschaus, et al. (1989). "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 **3**(1): 96-113.

Robbins, J., S. M. Dilworth, et al. (1991). "Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence." Cell **64**(3): 615-623.

Rout, M. P., J. D. Aitchison, et al. (2003). "Virtual gating and nuclear transport: The hole picture." Trends in Cell Biology **13**(12): 622-628.

Rout, M. P. and G. Blobel (1993). "Isolation of the yeast nuclear pore complex." J Cell Biol **123**(4): 771-783.

Rout, M. P. and S. R. Wente (1994). "Pores for thought: Nuclear pore complex proteins." Trends in Cell Biology **4**(10): 357-365.

Ryan, K. J. and S. R. Wente (2000). "The nuclear pore complex: A protein machine bridging the nucleus and cytoplasm." Current Opinion in Cell Biology **12**(3): 361-371.

Santos-Rosa, H., H. Moreno, et al. (1998). "Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores." Mol Cell Biol **18**(11): 6826-6838.

Satterly, N., M. Yarbrough, et al. (2011). "The Cleavage and Polyadenylation factor CPSF30 Inhibits NXF1 Mediated mRNA Export." Mol Biol Cell **Under revision**.

Scheffzek, K., C. Klebe, et al. (1995). "Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form." Nature **374**(6520): 378-381.

Schlenstedt, G., D. H. Wong, et al. (1995). "Mutants in a yeast Ran binding protein are defective in nuclear transport." Embo Journal **14**(21): 5367-5378.

Schwoebel, E. D., B. Talcott, et al. (1998). "Ran-dependent signal-mediated nuclear import does not require GTP hydrolysis by Ran." J Biol Chem **273**(52): 35170-35175.

Screaton, R. A., M. D. Conkright, et al. (2004). "The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector." Cell **119**(1): 61-74.

Segref, A., K. Sharma, et al. (1997). "Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores." EMBO J **16**(11): 3256-3271.

Senay, C., P. Ferrari, et al. (2003). "The Mtr2-Mex67 NTF2-like domain complex. Structural insights into a dual role of Mtr2 for yeast nuclear export." J Biol Chem **278**(48): 48395-48403.

Shaw, M. L., W. B. Cardenas, et al. (2005). "Nuclear localization of the Nipah virus W protein allows for inhibition of both virus- and toll-like receptor 3-triggered signaling pathways." J Virol **79**(10): 6078-6088.

Siomi, H. and G. Dreyfuss (1995). "A nuclear localization domain in the hnRNP A1 protein." J Cell Biol **129**(3): 551-560.

Siomi, M. C., P. S. Eder, et al. (1997). "Transportin-mediated nuclear import of heterogeneous nuclear RNP proteins." J Cell Biol **138**(6): 1181-1192.

Smillie, D. A., A. J. Llinas, et al. (2004). "Nuclear import and activity of histone deacetylase in Xenopus oocytes is regulated by phosphorylation." J Cell Sci **117**(Pt 9): 1857-1866.

Smith, A., A. Brownawell, et al. (1998). "Nuclear import of Ran is mediated by the transport factor NTF2." Curr Biol **8**(25): 1403-1406.

Solsbacher, J., P. Maurer, et al. (1998). "Cse1p is involved in export of yeast importin α from the nucleus." Molecular and Cellular Biology **18**(11): 6805-6815.

Stade, K., C. S. Ford, et al. (1997). "Exportin 1 (Crm1p) is an essential nuclear export factor." Cell **90**(6): 1041-1050.

Stewart, M. (2007). "Molecular mechanism of the nuclear protein import cycle." Nat Rev Mol Cell Biol **8**(3): 195-208.

Stewart, M. (2010). "Nuclear export of mRNA." Trends Biochem Sci **35**(11): 609-617.

Stewart, M., R. P. Baker, et al. (2001). "Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import." FEBS Lett **498**(2-3): 145-149.

Stewart, M., H. M. Kent, et al. (1998). "Structural basis for molecular recognition between nuclear transport factor 2 (NTF2) and the GDP-bound form of the Ras-family GTPase Ran." Journal of Molecular Biology **277**(3): 635-646.

Strambio-De-Castillia, C., M. Niepel, et al. (2010). "The nuclear pore complex: bridging nuclear transport and gene regulation." Nat Rev Mol Cell Biol **11**(7): 490-501.

Strasser, A., A. Dickmanns, et al. (2005). "Structural basis for m3G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1." Embo Journal **24**(13): 2235-2243.

Strasser, K., J. Bassler, et al. (2000). "Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export." J Cell Biol **150**(4): 695-706.

Strawn, L. A., T. Shen, et al. (2004). "Minimal nuclear pore complexes define FG repeat domains essential for transport." Nat Cell Biol **6**(3): 197-206.

Stutz, F., A. Bachi, et al. (2000). "REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export." RNA **6**(4): 638-650.

Stutz, F. and E. Izaurralde (2003). "The interplay of nuclear mRNP assembly, mRNA surveillance and export." Trends Cell Biol **13**(6): 319-327.

Suel, K. E., A. E. Cansizoglu, et al. (2006). "Atomic resolution structures in nuclear transport." Methods **39**(4): 342-355.

Suel, K. E. and Y. M. Chook (2009). "Kap104p imports the PY-NLS-containing transcription factor Tfg2p into the nucleus." J Biol Chem **284**(23): 15416-15424.

Suel, K. E., H. Gu, et al. (2008). "Modular organization and combinatorial energetics of proline-tyrosine nuclear localization signals." PLoS Biol **6**(6): e137.

Suzuki, M., M. Iijima, et al. (2005). "Two separate regions essential for nuclear import of the hnRNP D nucleocytoplasmic shuttling sequence." FEBS J **272**(15): 3975-3987.

Takemori, H., J. Kajimura, et al. (2007). "TORC-SIK cascade regulates CREB activity through the basic leucine zipper domain." FEBS J **274**(13): 3202-3209.

Tan, W., A. S. Zolotukhin, et al. (2000). "The mRNA export in Caenorhabditis elegans is mediated by Ce-NXF-1, an ortholog of human TAP/NXF and Saccharomyces cerevisiae Mex67p." RNA **6**(12): 1762-1772.

Terry, L. J., E. B. Shows, et al. (2007). "Crossing the nuclear envelope: hierarchical regulation of nucleocytoplasmic transport." Science **318**(5855): 1412-1416.

Tonon, G., S. Modi, et al. (2003). "t(11;19)(q21;p13) translocation in mucoepidermoid carcinoma creates a novel fusion product that disrupts a Notch signaling pathway." Nat Genet **33**(2): 208-213.

Tran, E. J. and S. R. Wente (2006). "Dynamic nuclear pore complexes: life on the edge." Cell **125**(6): 1041-1053.

Truant, R. and B. R. Cullen (1999). "The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals." Mol Cell Biol **19**(2): 1210-1217.

Truant, R., Y. Kang, et al. (1999). "The human tap nuclear RNA export factor contains a novel transportin-dependent nuclear localization signal that lacks nuclear export signal function." J Biol Chem **274**(45): 32167-32171.

Tseng, S. S. I., P. L. Weaver, et al. (1998). "Dbp5p, a cytosolic RNA helicase, is required for poly(A)+ RNA export." Embo Journal **17**(9): 2651-2662.

Vetter, I. R., A. Arndt, et al. (1999). "Structural view of the Ran-Importin beta interaction at 2.3 A resolution." Cell **97**(5): 635-646.

Vetter, I. R., C. Nowak, et al. (1999). "Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: Implications for nuclear transport." Nature **398**(6722): 39-46.

Vinciguerra, P. and F. Stutz (2004). "mRNA export: an assembly line from genes to nuclear pores." Curr Opin Cell Biol **16**(3): 285-292.

Weirich, C. S., J. P. Erzberger, et al. (2006). "Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export." Nature Cell Biology **8**(7): 668-676.

Weis, K. (2002). "Nucleocytoplasmic transport: cargo trafficking across the border." Curr Opin Cell Biol **14**(3): 328-335.

Weis, K. (2003). "Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle." Cell **112**(4): 441-451.

Weis, K., C. Dingwall, et al. (1996). "Characterization of the nuclear protein import mechanism using Ran mutants with altered nucleotide binding specificities." Embo Journal **15**(24): 7120-7128.

Weis, K., U. Ryder, et al. (1996). "The conserved amino-terminal domain of hSRP1α is essential for nuclear protein import." Embo Journal **15**(8): 1818-1825.

Wen, W., J. L. Meinkoth, et al. (1995). "Identification of a signal for rapid export of proteins from the nucleus." Cell **82**(3): 463-473.

Wente, S. R. and M. P. Rout (2010). "The nuclear pore complex and nuclear transport." Cold Spring Harb Perspect Biol **2**(10): a000562.

Wieschaus, E. and R. Riggleman (1987). "Autonomous requirements for the segment polarity gene armadillo during Drosophila embryogenesis." Cell **49**(2): 177-184.

Wiese, C., A. Wilde, et al. (2001). "Role of importin-β in coupling ran to downstream targets in microtubule assembly." Science **291**(5504): 653-656.

Wilkie, G. S., V. Zimyanin, et al. (2001). "Small bristles, the Drosophila ortholog of NXF-1, is essential for mRNA export throughout development." RNA **7**(12): 1781-1792.

Wozniak, R. W., M. P. Rout, et al. (1998). "Karyopherins and kissing cousins." Trends in Cell Biology **8**(5): 184-188.

Wu, J., M. J. Matunis, et al. (1995). "Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region." J Biol Chem **270**(23): 14209-14213.

Wu, Z., X. Huang, et al. (2006). "Transducer of regulated CREB-binding proteins (TORCs) induce PGC-1alpha transcription and mitochondrial biogenesis in muscle cells." Proc Natl Acad Sci U S A **103**(39): 14379-14384.

Xiao, Z., X. Liu, et al. (2000). "Acclerated Publication: Importin β mediates nuclear translocation of Smad 3." Journal of Biological Chemistry **275**(31): 23425-23428.

Yang, Q., M. P. Rout, et al. (1998). "Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications." Mol Cell **1**(2): 223-234.

Yaseen, N. R. and G. Blobel (1997). "Cloning and characterization of human karyopherin beta3." Proc Natl Acad Sci U S A **94**(9): 4451-4456.

Yokoyama, N., H. Hayashi, et al. (1995). "A giant nucleopore protein that binds Ran/TC4." Nature **376**(6536): 184-188.

Yoon, J. H., D. C. Love, et al. (2000). "Mex67p of Schizosaccharomyces pombe interacts with Rae1p in mediating mRNA export." Mol Cell Biol **20**(23): 8767-8782.

York, J. D., A. R. Odom, et al. (1999). "A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export." Science **285**(5424): 96-100.

Yun, C. Y., A. L. Velazquez-Dones, et al. (2003). "Phosphorylation-dependent and -independent nuclear import of RS domain-containing splicing factors and regulators." J Biol Chem **278**(20): 18050-18055.

Zahler, A. M., W. S. Lane, et al. (1992). "SR proteins: a conserved family of pre-mRNA splicing factors." Genes Dev **6**(5): 837-847.

Zasloff, M. (1983). "tRNA transport from the nucleus in a eukaryotic cell: Carrier-mediated translocation process." Proceedings of the National Academy of Sciences of the United States of America **80**(21 I): 6436-6440.

Zhou, Z., M. J. Luo, et al. (2000). "The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans." Nature **407**(6802): 401-405.