

STRUCTURAL AND BIOCHEMICAL STUDIES OF NUCLEAR IMPORT
FACTOR KARYOPHERIN BETA-2

APPROVED BY SUPERVISORY COMMITTEE

Yuh Min Chook, Ph.D.

Elizabeth J. Goldsmith, Ph.D.

David R. Corey, Ph.D.

Margaret A. Phillips, Ph.D.

Mischa Machius, Ph.D.

DEDICATION

I would like to express my appreciation and thanks to my mentor Yuh Min Chook for her continuous support throughout my Ph.D. education. She has been a great exemplary scientist in guiding me in my studies. I also want to thank every member of the Chook Lab past and present for making the lab a very enjoyable environment especially Brittany for her friendship and support during difficult times. Thanks to Katie, Hongmei and ZiChao for being wonderful lab mates. I would like to thank the members of my committee Elizabeth Goldsmith, Margaret Phillips, David Corey, and Mischa Machius for their time and invaluable guidance throughout my education as a scientist. I would especially like to thank Dr. David Corey for being the reason for my coming to UT Southwestern and introducing me to the fantastic science going on. I would also like to thank Dr. Mischa Machius in particular. He has taught me a lot about X-ray Crystallography in theory as well as in practice. I would like to thank my friends without whose support I would not have made it through the graduate school, especially Naci, Abdullah and Hadi. I would like to express my thanks to my mother, my brother Fatih and my sister Emine. Without their love and support I wouldn't become who I am. Finally, I would like to thank my wife Esra, for her continued support and patience.

STRUCTURAL AND BIOCHEMICAL STUDIES OF NUCLEAR IMPORT
FACTOR KARYOPHERIN BETA-2

by

AHMET ERTUGRUL CANSIZOGLU

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

September, 2007

Copyright

by

AHMET ERTUGRUL CANSIZOGLU, 2007

All Rights Reserved

STRUCTURAL AND BIOCHEMICAL STUDIES OF NUCLEAR IMPORT
FACTOR KARYOPHERIN BETA-2

Ahmet Ertugrul Cansizoglu, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2007

Supervising Professor: Yuh Min Chook, Ph.D.

Karyopherin β (Kap β) proteins recognize nuclear localization and export signals (NLS/NES) and mediate the transport of macromolecules across the nuclear envelope, a process regulated by Ran GTPase through its nucleotide cycle. Diversity and few number of available signal sequences recognized by Kap β s have prevented prediction of new Kap β substrates. The structure of Kap β 2 (also known as Transportin) in complex with its most well studied substrate, the NLS of hnRNP A1 (M9NLS), elucidates the mechanism of substrate release by Ran GTPase. Analyses of the structure in conjunction with available NLS sequences

reveal three general rules for NLS recognition by Kap β 2: NLSs are structurally disordered in substrates, they have overall basic character, and they carry an N-terminal hydrophobic or basic sequence motif followed by a C-terminal R/H/KX₍₂₋₅₎PY consensus sequence. These rules successfully identify NLSs in seven previously known Kap β 2 substrates. These studies define and validate a new type of NLS, we term as PY-NLSs that could not be predicted by primary sequence analysis alone. After solving the structure of the basic-PY NLS (M9NLS) complex structure, I solved the structure of a basic-PY NLS substrate (hnRNP-M NLS) in complex with Kap β 2. Kap β 2 complexes carrying hydrophobic (M9NLS) and basic PY-NLSs converge in structure only at previously identified consensus motifs, which explains the ligand diversity. Using complementary biochemical experiments, we designed a chimeric Kap β 2 substrate, M9M, which acts as a pathway specific nuclear import inhibitor. To complete the Kap β 2 import cycle picture, I solved the structure of Kap β 2 in its unliganded form. The structure of the unliganded Kap β 2 together with structures of hydrophobic, basic PY-NLSs and Ran complexes provides understanding of conformational heterogeneity that accompanies ligand binding. The Kap β 2 superhelix is divided into three major segments. Two of them (HEAT repeats 9-13 and 14-18), which constitute the substrate binding site, are rigid elements that rotate relative to each other about a flexible hinge. The third (HEAT repeats 1-8), which constitute the Ran binding site, exhibits conformational changes throughout

its length. An analogous segmental architecture is also observed in Importin β suggesting that it is functionally significant and may be conserved in other import Karyopherins.

TABLE OF CONTENTS

CHAPTER ONE

Introduction.....	1
Nucleocytoplasmic Transport.....	1
Nuclear Pore Complex.....	2
Karyopherins and Nuclear Transport.....	9
Nuclear Localization and Nuclear Export Signals (NLS/NES)	12
Ran Mediated Nuclear Transport.....	14
Karyopherin β 2 and Nuclear Import.....	17
Bidirectional Transport Karyopherins and Msn5.....	18

CHAPTER TWO

Biochemical and structural studies on Karyopherin beta-2dloop.hnRNPA1-NLS (M9NLS) structure.....	22
Abstract.....	22
Introduction.....	23
Materials and Methods.....	25
Protein expression, purification and complex formation.....	25
Crystallization and Data Collection.....	26
Phasing of Kap β 2dloop.M9NLS structure.....	27
NLS-mapping, site directed mutagenesis and Kap β 2 binding assays.....	29

Quantitation of binding affinity with Isothermal Titration Calorimetry (ITC)...	30
Results and Discussions.....	31
Kap β 2-M9NLS Complex Structure Overview.....	31
The Kap β 2-M9NLS binding interface.....	34
Distribution of binding energy along M9NLS.....	38
Rule 1: NLS is structurally disordered in substrate.....	41
Rule 2: Overall positive charge for NLS is preferred.....	42
Rule 3: Consensus sequences for the NLS.....	45
The NLS rules are predictive.....	48
Mechanism of Ran-mediated substrate dissociation from Kap β 2.....	50

CHAPTER THREE

Structural and Biochemical studies on Kap β 2.basic-PY NLS Complex.....	60
Abstract.....	60
Introduction.....	60
Materials and Methods.....	62
Expression and Purification of human Kap β 2.....	62
Expression of human hnRNPM NLS Fragment.....	66
Kap β 2:hnRNP M-NLS Complex Formation.....	69
Crystallization and Data Collection	70
Phasing and Structure Determination	72

Expression and purification of MBP tagged NLS fragments.....	75
Alanine mutants of NLS fragments.....	76
Quantitation of binding affinity with Isothermal Titration Calorimetry.....	76
Competition ITC Experiments.....	77
Qualitative Binding assays.....	78
Subcellular localization of proteins in HeLa cells.....	80
Results and Discussions.....	82
The structure of the bPY-NLS Cargo hnRNP-M and Kap β 2 Complex.....	84
Binding energy determination using ITC experiments.....	91

CHAPTER FOUR

Structural studies on Unliganded Kap β 2.....	104
Abstract.....	104
Introduction.....	105
Materials and Methods.....	107
Expression and purification of Kap β 2.....	107
Crystallization, data collection and structure determination.....	108
Structure refinement.....	113
Analyses of conformational heterogeneity.....	120
Results and discussions.....	122
Structure determination and overall structure of unliganded Kap β 2.....	122

Conformational flexibility of unliganded Kap β 2.....	125
NLS recognition: a hinge in the Kap β 2 C-terminal arch	133
Kap β 2-Ran interaction: the flexible N-terminal arch and H8 loop.....	137
Segmental architecture in Imp β and Kap95p.....	146
Conclusions.....	149
CHAPTER FIVE	
Studies of the Kap β 2 acidic H8 loop.....	151
Introduction.....	151
Materials and Methods.....	156
Alanine Loop Mutants.....	156
Lysine Loop Mutants.....	158
Results and Discussion.....	159
CHAPTER SIX	
Biochemical and structural studies on bi-directional nuclear transporter	
Msn5p.....	162
Introduction.....	162
Materials and Methods.....	164
Expression and Purification of S.Cerevisiae Msn5p Native and Se-met	
Proteins.....	164

Crystallization of Msn5p.....	167
Expression and Purification of the Substrates.....	168
Complex formation with RPA.....	173
Results and Discussions.....	173
BIBLIOGRAPHY.....	176

PRIOR PUBLICATIONS

Cansizoglu AE, Chook YM. *Conformational heterogeneity of Karyopherin β 2 is segmental*. Structure, 2007 (*in press*)

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 Z, Chook YM. *Rules for nuclear localization sequence recognition by karyopherin beta 2*. Cell. 2006 Aug 11;126(3):543-58.

Suel KE, Cansizoglu AE, Chook YM. *Atomic resolution structures in nuclear transport*. Methods. 2006 Aug;39(4):342-55.

Kaihatsu K, Braasch DA, Cansizoglu A, Corey DR. *Enhanced strand invasion by peptide nucleic acid-peptide conjugates*. Biochemistry. 2002 Sep 17;41(37):11118-25.

LIST OF FIGURES

FIGURE 1-1.....	3
FIGURE 1-2.....	7
FIGURE 1-3.....	11
FIGURE 1-4.....	13
FIGURE 2-1.....	35
FIGURE 2-2.....	37
FIGURE 2-3.....	41
FIGURE 2-4.....	44
FIGURE 2-5.....	49
FIGURE 2-6.....	52
FIGURE 2-7.....	54
FIGURE 2-8.....	56
FIGURE 2-9.....	58
FIGURE 3-1.....	64
FIGURE 3-2.....	64
FIGURE 3-3.....	65
FIGURE 3-4.....	66
FIGURE 3-5.....	68
FIGURE 3-6.....	70
FIGURE 3-7.....	71
FIGURE 3-8.....	83
FIGURE 3-9.....	86

FIGURE 3-10	87
FIGURE 3-11	89
FIGURE 3-12	92
FIGURE 3-13	94
FIGURE 3-14	96
FIGURE 3-15	97
FIGURE 3-16	98
FIGURE 3-17	100
FIGURE 3-18	101
FIGURE 3-19	102
FIGURE 4-1	113
FIGURE 4-2	125
FIGURE 4-3	130
FIGURE 4-4	132
FIGURE 4-5	136
FIGURE 4-6	139
FIGURE 4-7	141
FIGURE 4-8	143
FIGURE 4-9	144
FIGURE 4-10	145
FIGURE 4-11	148
FIGURE 5-1	152
FIGURE 5-2	155

FIGURE 5-3.....	160
FIGURE 5-4.....	161
FIGURE 6-1.....	165
FIGURE 6-2.....	166
FIGURE 6-3.....	169
FIGURE 6-4.....	170
FIGURE 6-5.....	171
FIGURE 6-6.....	174

LIST OF TABLES

TABLE 1-1	4
TABLE 1-2	8
TABLE 1-3	20
TABLE 2-1	33
TABLE 2-2	39
TABLE 3-1	72
TABLE 3-2	74
TABLE 3-3	91
TABLE 4-1	111
TABLE 4-2	115
TABLE 4-3	118
TABLE 4-4	128

LIST OF ABBREVIATIONS

APS	Advanced Photon Source
ATP	Adenosine triphosphate
β -ME	β -mercaptoethanol
CCP4	Collaborative Computing Project IV
CNS	Crystallography and NMR System
DTT	Dithiothreitol
E.coli	Escherichia coli
EDTA	Ethylenediamine tetra-acetic acid
GAP	GTPase activating protein
GDP	Guanosine 5'-diphosphate
GEF	Guanine nucleotide exchange factor
GMPPNP	Guanosine 5'-[β,γ -imido]triphosphate
GST	Glutathione S-transferase
GTP	Guanosine 5'-triphosphate
HEAT	Huntingtin, Elongation factor 3, 'A' subunit of protein phosphatase-2A, and TOR1
HEPES	4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid
hnRNP	heterogeneous nuclear ribonucleoprotein
IPTG	Isopropyl β -D-thiogalactoside
ITC	Isothermal Titration Calorimetry
Kap α	Karyopherin Alpha
Kap β 1	Karyopherin Beta-1
Kap β 2	Karyopherin Beta-
K _D	Dissociation constant
kDa	kilo Dalton
LB	Luria Bertoni
MBP	Maltose Binding Protein
MDa	Mega daltons
MES	2-(N-morpholino)ethanesulfonic acid
MR	Molecular replacement
NCS	Non-crystallographic Symmetry
NES	Nuclear Export Signal
NPC	Nuclear Pore Complex
OD ₆₀₀	Optical density at 600nm
PCR	Polymerase chain reaction
PDB	Protein databank
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride

RPA	Replication Protein A
r.m.s.d.	Root Mean Square Deviation
s200	Superdex 200
SAD	Single wavelength anomalous dispersion
SAXS	Small Angle X-ray Scattering
SBC	Structural Biology Center
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel eletrophoresis
TCEP	Tris(2-carboxyethyl)phosphine
TLS	Translation, Libration, Screw
TLSMD	Translation/Libration/Screw Motion Determination

CHAPTER ONE

Introduction

Nucleocytoplasmic Transport

Characteristic to eukaryotic organisms, cytoplasm is separated from the nucleoplasm by a double layer membrane nuclear envelope. The nucleus and the contents are physically separated from the cytoplasm. DNA and RNA synthesis is localized to the nucleus whereas the protein synthesis is localized to the cytoplasm. This provides the cell with an opportunity of extensive control over these critical processes inside the cell. A transcription factor would be inactive if located inside the cytoplasm. However, upon activation, it translocates into the nucleus where it will activate the transcription of designated target genes.

The transport of the RNAs and proteins between these compartments is carried out through the nuclear pore complexes (NPC) present on the nuclear envelope(Akey and Goldfarb 1989). The passage through the nuclear pore can be via either passive or active/facilitated diffusion(Gorlich and Kutay 1999) Small molecules such as ions, cofactors, metabolites and proteins of smaller size (<20-30 kDa) can passively diffuse through the nuclear pore(Feldherr, Cohen et al. 1983; Adam, Marr et al. 1990). However, larger molecules require signal-mediated transport(Adam, Marr et al. 1990).

Macromolecules that carry a nuclear localization signal (NLS) or nuclear export signal (NES) are actively transported through the NPC by soluble transport factors known as karyopherins. This transport mechanism is completely different than protein transport into mitochondria or endoplasmic reticulum (ER), where the proteins are unfolded during this process. During nuclear transport, the proteins maintain their fully folded structure(Gorlich and Kutay 1999).

Nuclear Pore Complex

The NPC is formed by soluble and membrane associated factors that are assembled in a large, multimeric structure. The nuclear pore complex is a rather sizeable assembly, connecting the cytoplasm to the nuclear periphery (~66MDa in yeast and ~125MDa in vertebrates(Reichelt, Holzenburg et al. 1990; Rout and Blobel 1993)). As revealed by cryoelectron microscopy, the NPC follows an approximate 822 symmetry with a ring structure composed of large spoke domains (Figure 1-1)(Akey and Radermacher 1993; Yang, Rout et al. 1998). It consists of an unequal distribution of ~30 different repeating proteins belonging to the Nucleoporin family of proteins (Nups)(Hinshaw, Carragher et al. 1992).

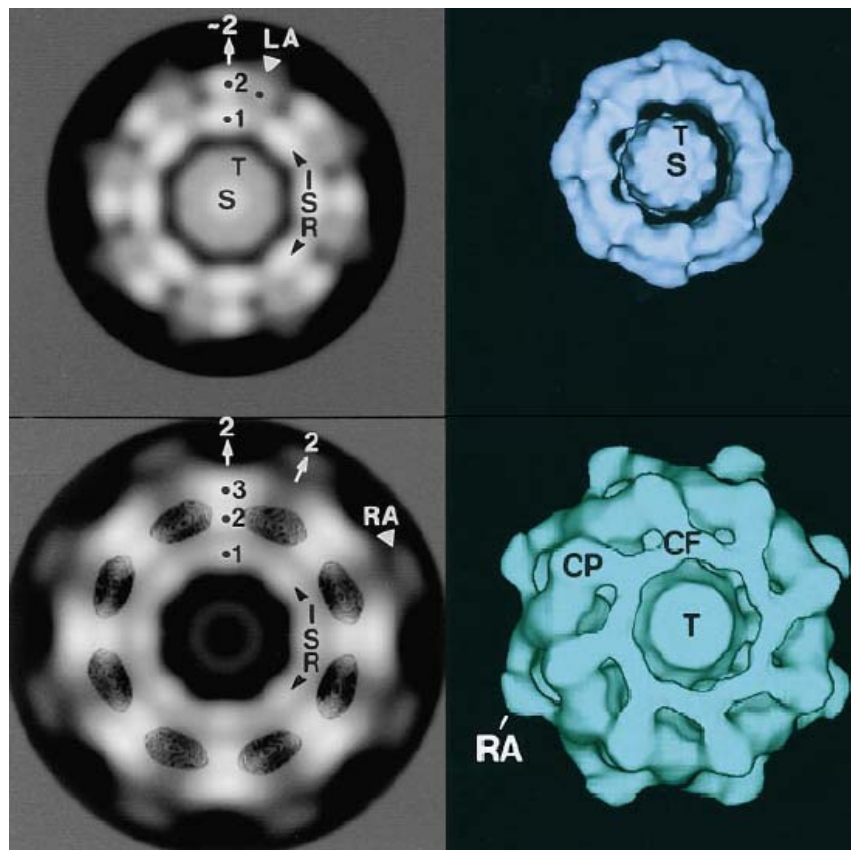


Figure 1-1) Yeast (top panel) and vertebrate (bottom panel) NPCs revealed by cryo-EM studies (on the left) and built 3D models (on the right). ISR: inner spoke ring

Numbered: Spoke domains Arrows: Putative symmetry axes LA: Linear Arm RA: Radial Arm Putative: T: Transporter S: Substrate. Adapted from (Yang, Rout et al. 1998).

Table 1-1) Domain Positions and Dimensions in Yeast and Vertebrate NPCs (adapted from (Yang, Rout et al. 1998)): NPC dimensions are from (Akey and Goldfarb 1989; Akey and Radermacher 1993; Yang, Rout et al. 1998) Di/Dc/Do are inner, central and outer diameters of the NPC.

	Yeast NPC				Vertebrate NPC			
Structural Domains	Di (Å)	Dc (Å)	Do (Å)	Height (Å)	Di (Å)	Dc (Å)	Do (Å)	Height (Å)
Transporter			340-380	300			320-420	625
Inner Spoke Domain	490	580	680	350-380	480	600	720	~340
Second Spoke Domain	730	840	960	160-220	720	860	940	
Membrane Ring		860		50-70				
Nuclear Membrane Pore	780	820	870			840-850		
Entire NPC			960	350-380			1200-1450	800-1000

It is thought that all karyopherins bind transiently to nucleoporins as they translocate through the NPC. There are several types of nucleoporins. These include transmembrane proteins that anchor the NPC in the nuclear membrane,

FG-containing nucleoporins and nucleoporins that contain WD repeats or seven-blade propeller motif. Transmembrane proteins anchor the assembly into the double layer membrane. FG-containing nups are found to be particularly rich in Phe-Gly (FG) repeats. FxFG and GLFG Nups that contain multiple copies of FG repeats are thought to act as binding sites for transport receptors. Although the FG Nups are shown to interact with the nuclear transport receptors, it has been shown that at least half the FG-repeat mass of the NPC in yeast can be deleted with little effect on transport (Strawn, Shen et al. 2004).

The NPC is roughly symmetrical along the double layer nuclear envelope with asymmetrical distribution of proteins only at its periphery. This asymmetry may be critical for the translocation process. However the exact mechanism of translocation still remains unclear to date.

In EM structures of NPCs, it appears as if there are variable passageways along the radial axis of the cylindrical architecture of the complex, possibly allowing several different molecules to pass simultaneously using alternative routes in and out. However, there is no other experimental evidence to support this suggestion. An interesting feature of the NPC is the presence of a visible “plug” of considerable size in the center of the cylindrical empty structure in cryo-EM studies (see Figure 1-1 and Table 1-1). This plug has been speculated to be the soluble protein fraction of the NPC ‘caught in the action’. The constituents of this

plug may contain the karyopherins active in the transport process(Akey and Radermacher 1993; Yang, Rout et al. 1998).

The NPC has filamentous projections on the inner and the outer surface of the nuclear envelope. Eight ~ 50 nm fibrils extend into the cytoplasm (Figure 1-2) on the cytoplasmic side of the complex. On the nucleoplasmic side of the NPC, eight nuclear fragments protrude into the nuclear side forming the nuclear basket. These projections are thought to play a role in interactions with the nuclear transport receptors(Bednenko, Cingolani et al. 2003).

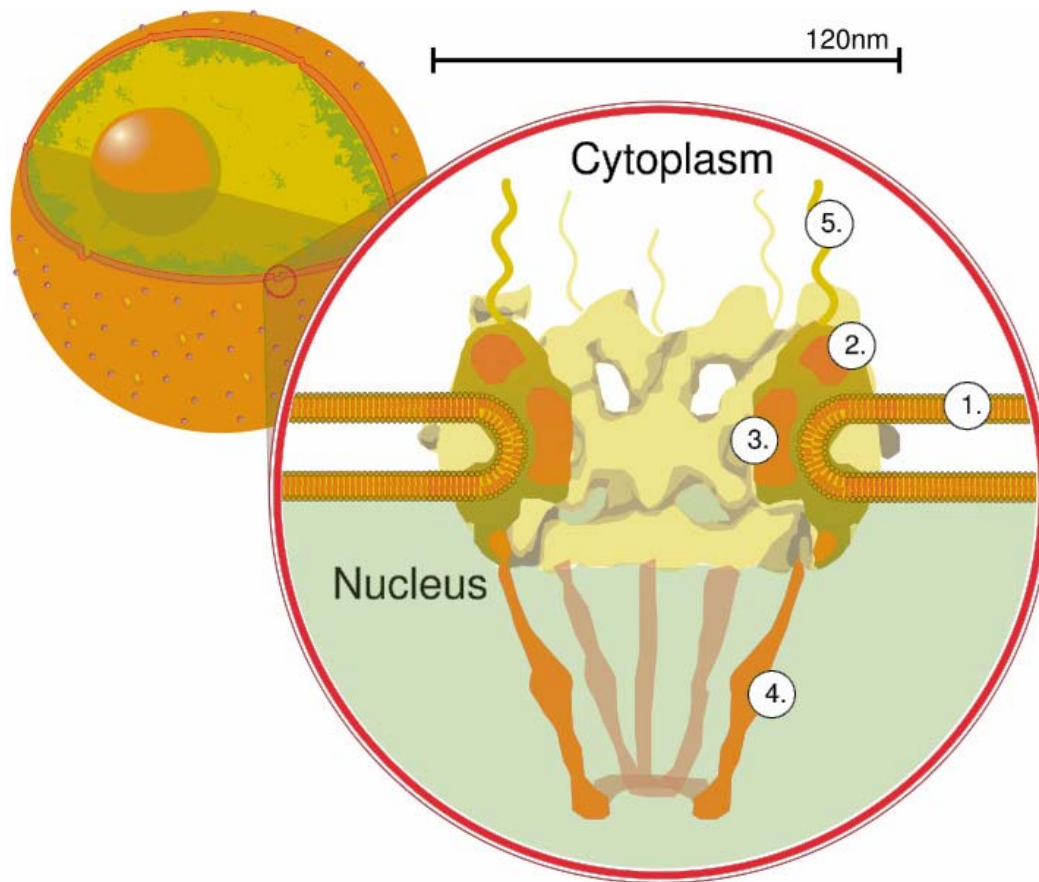


Figure 1-2) Nuclear pore. Side view: 1. Nuclear envelope. 2. Outer ring. 3. Spokes. 4. Nuclear basket. 5. Cytoplasmic filaments. (Drawing is based on electron microscopy images. Figure Taken from Wikipedia(created by Mike Jones at http://en.wikipedia.org/wiki/Image:NuclearPore_crop.svg.png licensed under the Creative Commons Attribution ShareAlike 2.5))

Table 1-2) Karyopherin β Family of Proteins (Modified from (Pemberton and Paschal 2005)).

Human	Cargo	Yeast	Cargo	Essential?
Import				
Karyopherin β 1	Cargoes with basic NLSs via karyopherin α , srebp, UsnRNPs via snurportin	Kap95	Cargoes including those with basic NLS via karyopherin α	Yes
Karyopherin β 2	hnRNPA1,D, F, M, DDX3, histones, TAP, HuR, ribosomal proteins	Kap104	Nab2, Hrp1	ts
Transportin SR1	SR proteins	Mtr10/Kap111	Npl3, Hrb1	ts
Transportin SR2	HuR			
Importin 4	Histones, ribosomal proteins	Kap123	Ribosomal proteins, histones	No
Importin 5	Histones, ribosomal proteins	Kap121	Ribosomal proteins, histones, Pho4, others	Yes
Importin 9	Histones, ribosomal proteins	Kap114	TBP, histones, Nap1p	No
Importin 7	HIV RTC, Glucocorticoid receptor, ribosomal proteins	Nmd5/Kap119	TFIIS, Hog1	No
		Sxml/Kap108	Lhp1, ribosomal proteins	No
Importin 8	SRP19			
Importin 11	UbcM2, rpL12			No
		Kap122	TFIIA	
Export				
Crm1	Leucine rich NES cargoes	Crml	Leucine rich NES cargoes	Yes
Exportin-t	tRNA	Los1	tRNA	No
CAS	Karyopherin α	Cse1	Karyopherin α	Yes
Exportin 4	eIF5 A			
Exportin 5	microRNA precursors			
Exportin 6	Profilin, actin			
Exportin 7	p50RhoGAP, 14338			
Import/Export				
Importin 13	Rbm8, Ubc9, Pax6 (import) eIF1 A (export)			
		Msn5	Pho4, Mig1p, Crz1p, ... (import) Replication protein A (export)	No
Uncharacterized				
RanBP6	undefined			
RanBP17	undefined			
		Kap120	undefined	No

Karyopherins and Nuclear Transport

The major players of transport between cytoplasm and nucleus are members of a family of proteins called Karyopherins (Kaps) or Importins and Exportins (Radu, Blobel et al. 1995; Gorlich, Dabrowski et al. 1997; Mattaj and Englmeier 1998; Pemberton and Paschal 2005). There are at least 14 members of karyopherin family in yeast and 20 known in humans (Table 1-2). The proteins in this family can be classified into three groups depending on their known function: import factors, export factors and bidirectional transporters. However, an import factor may well prove to be an export factor depending on newly identified cargoes.

Karyopherins share an overall low sequence identity (15-20%). They are relatively large proteins with molecular weights of around 95-145KDa and they have a rather acidic pI (4-5) (Chook and Blobel 2001). Another feature of Kaps is the presence of *HEAT* repeats (Huntingtin, Elongation factor 3, 'A' subunit of protein phosphatase 2A, and TOR1) (Hemmings, Adams-Pearson et al. 1990). These HEAT repeats are composed of multiple repeats of two antiparallel α -helices connected with a loop segment and stacked against each other. This helical stack forms a right handed superhelix (Figure 1-3). Therefore they do not have a globular domain structure. So far known Kaps carry 19-20 HEAT repeats. The superhelical structure can be divided into 'regions' which interact with Ran

and/or the cargo substrates. However, there is not a clear distinction of domain separation between these functional units. The proteins more likely resemble a single domain structure with functional interfaces mapping to different HEAT repeat regions.

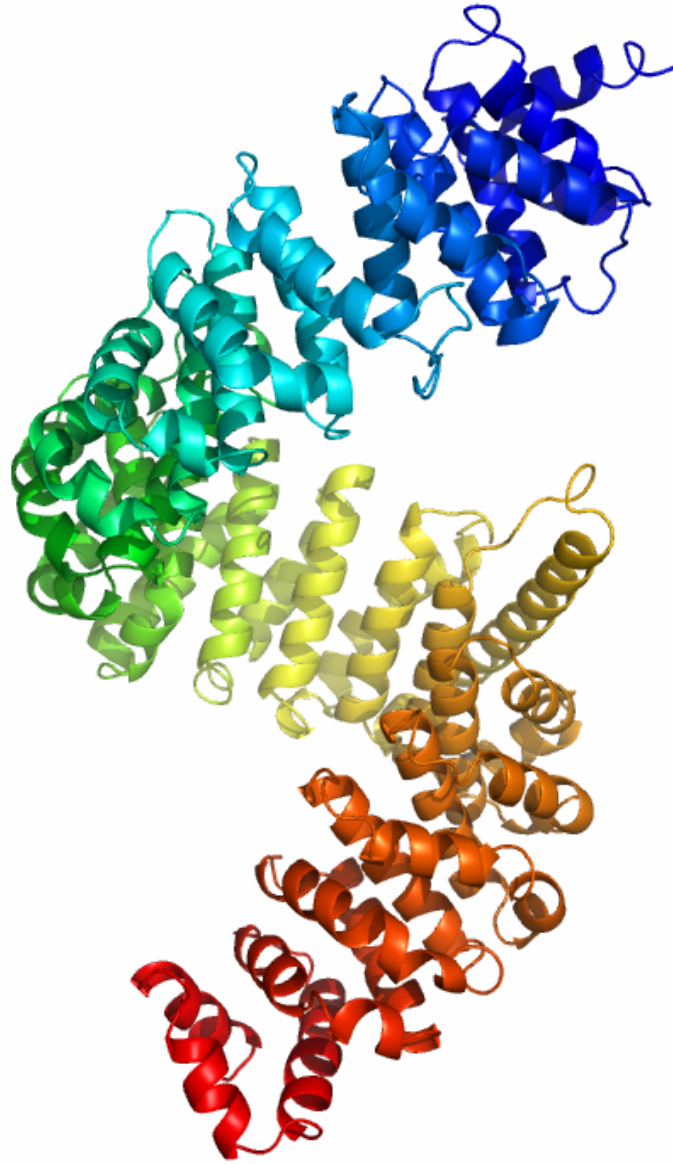


Figure 1-3) Karyopherins have a superhelical structure. Figure was made using PYMOL (DeLano 2002).

Nuclear Localization and Nuclear Export Signals (NLS/NES)

Karyopherins recognize and bind to target proteins targeted for the nucleus directly or indirectly through these nuclear localization signals (NLS) and nuclear export signals (NES) translocate them into or out of the nucleus by interacting with Nups (Chook and Blobel 2001).

The interactions of import Karyopherins with the particular cargo substrates occur via the NLS sequences. The most well studied nuclear import factor so far is Kap β 1 (importin β), which forms a heterodimer with Karyopherin α (Kap α). The NLS sequence recognizable by Kap α is a short stretch of basic residue segment termed as 'classical' NLS. This basic stretch of residues can be monopartite (SV40 large T antigen(Kalderon, Roberts et al. 1984)) or bipartite (Nucleoplasmin NLS(Robbins, Dilworth et al. 1991), Figure 1-4). Kap α is in its autoinhibited state when not bound to Kap β 1. IBB domain of Kap α binds to NLS binding site on Kap α itself. Kap α binds to Kap β 1 through its IBB domain rendering the NLS interaction site available for high affinity interaction with NLSs. NLSs can bind to the heterodimeric Kap β 1/Kap α complex for transport into the nucleus.

Kap α consists of a super helix composed of 10 armadillo (ARM) repeats. Kap α heterodimerizes with Kap β 1 forming the trimeric import complex. In this case, Kap α acts as a mediator for binding to Kap β 1. This trimeric import complex is

then transported into the nucleus. Inside the nucleus, Ran.GTP binds and dissociates the import cargo complex.

SV40 large-T antigen NLS	¹²⁶ <u>PKKKRRV</u> ¹³²
Nucleoplasmin NLS	¹⁵⁰ <u>KRPAATKKAGQAKK</u> ¹⁷⁰

Figure 1-4) Classical NLS sequences from SV40 Large T antigen and Nucleoplasmin NLS(Kalderon, Richardson et al. 1984; Robbins, Dilworth et al. 1991)

No other import karyopherin is known to utilize adaptor proteins like Kap α . All other proteins bind to their substrates directly. Furthermore, it is also known that even Kap β 1 can recognize and bind to structural and sequence motifs directly without Kap α . The relatively large inner surface area of karyopherins could potentially harbor distinct binding regions for different cargo proteins. For example, three crystal structures of Kap β 1-substrate complexes show substrates binding to different sites on Kap β 1 (Cingolani, Petosa et al. 1999; Cingolani, Bednenko et al. 2002; Lee, Sekimoto et al. 2003). Thus the potential substrate spectrum of Kaps with multiple substrate sites should be very large.

On the other hand, the NLS sequence conservation is not well described for the remaining karyopherins except for Kap β 2. This is mainly due to the lack of a sufficient number of substrate cargoes which prevents the identification of the consensus sequence. It is evidently possible that the recognition motif could also be a three dimensional fold rather than a sequence consensus. It is essential that all the proteins functioning inside the nucleus be carried into the nucleus if larger than a particular size defined by the size of the nuclear pore. This will include all proteins inside the nucleus; all the transcription factors and related activators and DNA interacting proteins as well as the RNA processing pathway proteins. This is a large target protein pool. Thus multiple Karyopherins may also interact with a substrate thus providing a further redundancy and complexity to this system.

Ran Mediated Nuclear Transport

For larger macromolecules, nuclear transport is regulated by the Ran GTPase. Ran is a member of Ras super family of small G proteins. According to the current nuclear transport model, all Karyopherins have high affinity for Ran in the GTP bound form only. Ran is asymmetrically distributed between the nucleoplasm and cytoplasm with GTP-bound form in almost 1000-fold in excess

in the nucleus(Gorlich, Seewald et al. 2003). Whereas Ran is present extensively in GDP-bound form on the cytoplasmic side.

The asymmetric distribution of Ran is achieved by the asymmetrical localization of Ran-GEF (Guanidine Nucleotide Exchange Factor- RCC1) into the nucleus and Ran-GAP (Guanidine Nucleotide Activating Protein) into the cytoplasm. Ran-GEF catalyzes the exchange of GDP to GTP bound to Ran(Bischoff and Ponstingl 1991). Cytoplasmic Ran-GAP catalyzes the hydrolysis of Ran.GTP to Ran.GDP resulting in cytoplasmic Ran-GDP(Bischoff, Klebe et al. 1994).

In Ran mediated import of nuclear proteins, import Kap β is able to bind to substrate targeted for the nucleus. In the cytoplasm, import Karyopherins cannot bind Ran because it is in its GDP bound form. The Kap β .substrate complex binds to the NPC and is translocated through the NPC via an undetermined mechanism. In the nucleus, Ran, which is present in its GTP bound form, can readily bind to the Kap β .substrate complex and dissociates the substrate from Kap β . After the release of substrate, Kap β , which is now bound to Ran.GTP, is returned to the cytoplasm again by a mechanism not very well understood. Once in cytoplasm, Ran.GTP is converted into Ran.GDP by Ran-GAP, which catalyzes the hydrolysis of GTP to GDP. This results in loss of affinity of Ran for the Karyopherin and the Karyopherin-Ran complex dissociates inside the cytoplasm allowing another round of import.

In nuclear export, an export substrate is able to bind to an export Kap β cooperatively with Ran.GTP in the nucleus. The trimeric complex is then translocated to the cytoplasm. On the cytoplasmic side, Ran.GTP is again converted into Ran.GDP by Ran-GAP, resulting in the complete dissociation of the export complex.

This mechanism is well supported by the crystal structures of Kap β 1 and Kap β 2(Chook and Blobel 1999; Vetter, Arndt et al. 1999). The structure of Kap β 2 shows a superhelical molecule with two arches, the N-terminal one of which is occupied by RanGppPnP(Chook and Blobel 1999). Biochemical studies suggest that import substrate binds to the C-terminal arch when Ran is absent(Chook, Jung et al. 2002). The three dimensional structure of Kap β 1 also has a superhelical shape with similar binding regions analogous to the Ran and substrate binding pockets in Kap β 2.

Continuous unidirectional transport of Ran from nucleus towards the cytoplasm by the karyopherins is balanced by import of Ran.GDP back into the nucleus by a small import receptor NTF2(Ribbeck, Lipowsky et al. 1998) which binds Ran.GDP and FG nups directly.

Karyopherin $\beta 2$ and Nuclear Import

The consensus sequence motif for the Kap $\beta 1$ /Kap α system has been well studied and characterized. Collectively, NLS sequences other than the “classical” NLS have been termed as “non-classical” NLS sequences. Non-classical NLSs are an extremely diverse group with no apparent consensus sequence pattern.

The second best studied import factor is Karyopherin $\beta 2$ (Kap $\beta 2$). For Kap $\beta 2$, more than 20 proteins have been identified earlier as experimentally known substrates. Some examples are hnRNPs A1, D, F, M, and TAP, Y-box binding protein 1, DDX3, and HuR (Siomi and Dreyfuss 1995; 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, Muhlhauser et al. 2004; Rebane, Aab et al. 2004; Suzuki, Iijima et al. 2005). Nevertheless, earlier attempts for the identification of a sequence consensus motif have been unsuccessful. For instance, hnRNPA1, which is an import substrate for Kap $\beta 2$ was found to have an NLS fragment composed of 39 amino acids, termed m9. the other substrates share no sequence homology to hnRNP-A1(Siomi and Dreyfuss 1995). Moreover, it appeared likely that the M9 sequence does not have regular secondary structure either(Pollard, Michael et al. 1996; Chook, Jung et al. 2002). The few NLS sequences mapped for substrates hnRNP D, HuR and TAP

have sequences that are different(Siomi and Dreyfuss 1995; 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).

Bidirectional Transport Karyopherins and Msn5

One interesting type of Kap β s is the bi-directional transporters which import a set of proteins going into the nucleus and export another set of proteins going into the cytoplasm. So far this group includes Msn5p and importin13 (Mingot, Kostka et al. 2001; Yoshida and Blobel 2001). Msn5p (Multi-copy suppressor of snf1-ts mutations) was initially found to be important in the nuclear export of various proteins(Kaffman, Rank et al. 1998; Blondel, Alepuz et al. 1999; DeVit and Johnston 1999; Boustany and Cyert 2002). However, strong evidence has been discovered supporting the idea that Msn5p is a bidirectional transporter(Yoshida and Blobel 2001). This evidence suggests that Msn5p takes part in the import of

Replication Protein A(RPA) into the nucleus and exports several others out into the cytoplasm (Table 1-3).

In this system, an import substrate can bind Msn5p in the cytoplasm and is dissociated from Msn5p in the nucleus in the presence of Ran in GTP bound form. On the other hand, an export cargo can bind Msn5p cooperatively in the presence of Ran.GTP in the nucleus and is dissociated in the cytoplasm upon GTP hydrolysis. Therefore, Msn5p has the distinct function in that it can bind different set of cargo substrates depending upon the nucleotide state of association with Ran. In this aspect, this protein will bear valuable information for both export and import pathways. An import substrate in complex with Msn5p would give information about the substrate binding characteristics of the import pathway, on the other hand, an export substrate in complex with Ran.GTP and Msn5p would provide valuable information about the export substrate binding nature of the protein.

Table 1-3) Known import and export cargoes of Msn5p(Kaffman, Rank et al. 1998; Blondel, Alepuz et al. 1999; DeVit and Johnston 1999; Boustany and Cyert 2002).

Known Msn5p substrates:		
Pho4	Transcription factor. Phosphorylated Pho4 is exported to cytoplasm by Msn5p	Export
Mig1p	Transcription factor. Phosphorylated Mig1p is exported to cytoplasm by Msn5p	Export
Crz1p	Transcription factor. Phosphorylated Crz1p is exported to cytoplasm by Msn5p	Export
Ste5p	Scaffold protein of MAPK cascade	Export
Far1p	Far1p is required in the nucleus to arrest the cell cycle, Inhibitor of G1 cyclin/CDK	Export
Cdh1p	Functions as an E3 ubiquitin ligase to degrade the mitotic cyclin Clb2p and other substrates during the G1 phase of the cell cycle	Export
Haa1	Transcriptional activator	Export
RPA	Replication protein A Takes part in replication, mismatch	Import

The transport in and out of the nuclear envelope is a good candidate to become a critical step for regulation of gene expression. Understanding the molecular

interactions taking place at this step will help to devise new techniques for regulating access to nucleus which is a critical step for gene regulation. It is important to investigate the interactions of the Karyopherins with the cargo substrates at the molecular level since they display rather specific cargo recognition. Even for the import pathway, the substrate interactions occurring at the molecular level are not well understood. Therefore, understanding the interactions taking part at the molecular level will give valuable information on the regulation of the proteins at this “gateway” step.

One can in principle, easily block the expression of a particular gene. This can be achieved by sequestering the corresponding transcription factor inside the cytoplasm, thereby controlling its access to nucleus by devising inhibitory molecules that will interfere with their binding to their carrier molecules. These target proteins very well include the viral proteins which need to access to the nucleus to carry out certain functions. The presence of certain level of distinction of cargo substrate recognition among the substrates could be a powerful tool to manipulate this to advantage by close examination of the structures in the future.

Likewise, the control of the export of proteins is a potential drug target as well. Enhancing or inhibiting the transport of proteins could be a next generation drug target for gene expression and gene regulation.

CHAPTER TWO

Biochemical and structural studies on Karyopherin beta-2dloop.hnRNPA1-NLS (M9NLS) structure

RULES FOR NUCLEAR LOCALIZATION SEQUENCE RECOGNITION BY KARYOPHERIN BETA 2

Abstract

Karyopherin β (Kap β) proteins bind nuclear localization and export signals (NLSs and NESs) to mediate nucleocytoplasmic trafficking, a process regulated by Ran GTPase through its nucleotide cycle. Diversity and complexity of signals recognized by Kap β s have prevented prediction of new Kap β substrates. The structure of Kap β 2 (also known as Transportin) bound to one of its substrates, the NLS of hnRNP A1, that we report here explains the mechanism of substrate displacement by Ran GTPase. Further analyses reveal three rules for NLS recognition by Kap β 2: NLSs are structurally disordered in free substrates, have overall basic character, and possess a central hydrophobic or basic motif followed by a C-terminal R/H/KX₍₂₋₅₎PY consensus sequence. We demonstrate the predictive nature of these rules by identifying NLSs in seven previously known Kap β 2 substrates and uncovering 81 new candidate substrates, confirming five

experimentally. These studies define and validate a new NLS that could not be predicted by primary sequence analysis alone.

Introduction

Our knowledge of substrate/import factor repertoire is very limited. In humans, ten import Kap β s have been shown to carry a diverse set of macromolecular substrates into the nucleus (Mosammaparast and Pemberton 2004). The diverse set of NLS sequences and the corresponding nuclear import or export cargo proteins seem to be far more complicated than classical NLS sequences.

In principle, the proteins functioning inside the nucleus need to be carried into the nucleus. This requires a much more sophisticated balance and distribution of cargo proteins being carried into the nucleus. The whole picture is far from complete.

The identification of very few substrates for each karyopherin prevents identification of consensus recognition motifs for these proteins. Only a few substrates have been determined for most import karyopherins except for Kap β 1 and Kap β 2. The lack of large numbers of substrates and the lack of sequence homology amongst the sequences prevent the identification of a consensus NLS recognition motif for most karyopherins.

The NLS for Kap β 1 has been well studied. On the other hand, for Kap β 2, so far, more than 20 proteins have been identified as known substrates. These proteins include hnRNPs A1, D, F, M, and TAP, Y-box binding protein 1, DDX3, and HuR (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, Muhlhauser et al. 2004; Rebane, Aab et al. 2004; Suzuki, Iijima et al. 2005). Nevertheless, identification of NLS consensus motif have been unsuccessful for Kap β 2 so far.

The best characterized NLS of a Kap β 2 substrate is that of RNA splicing factor hnRNP-A1. The NLS sequence for hnRNP A1 is mapped to residues 268-305, a 38 residue sequence termed M9 sequence (Pollard, Michael et al. 1996; Bonifaci, Moroianu et al. 1997). NLSs in HuR(Fan and Steitz 1998), TAP (Truant, Kang et al. 1999), and hnRNP D and JKTPB (Kawamura, Tomozoe et al. 2002; Suzuki, Iijima et al. 2005), have also been mapped. The NLSs of these proteins show either marginal or very poor sequence similarity to M9. The poor sequence similarity prevented identification of a consensus NLS recognition motif for Kap β 2 even though several substrates are known.

In this chapter, we determined the crystal structure of Kap β 2 bound to the hnRNP A1 NLS to gain insight into substrate recognition by Kap β 2. The structure in concord with extensive complementary biochemical studies revealed a set of rules for NLS recognition by Kap β 2.

Rule 1: Kap β 2 recognizes a stretch of more than 30 residues with intrinsic structural disorder.

Rule 2: The NLS has overall basic character.

Rule 3: There's a consensus motif which consists of a loose N terminal either basic or hydrophobic stretch of residues followed by a more conserved R/K/H-X₂₋₅-PY motif at the C terminus. These rules were shown to be predictive and successfully predicted NLSs in 7 known Kap β 2 cargo substrates.

Materials and Methods

Protein expression, purification and complex formation

Protein expression and purification for the Kap β 2.hnRNP A1 Fragment (M9NLS) cargo complex has been carried out by Brittany Lee in Chook Lab. In the crystallographic studies Kap β 2 residues 337-367 (A long loop in HEAT8) were replaced with a GGSGGSG linker sequence. This protein was expressed in *E. coli* BL21 (DE3) as a GST-fusion from pGEX-Tev vector and purified as described in chapter 3. M9NLS fragment was expressed in *E. coli* as an N terminal GST-fusion of hnRNP A1 residues 257-305, and purified as previously described (Chook, Jung et al. 2002). Two-fold molar excess of GST-M9NLS was

mixed with purified Kap β 2 and cleaved with Tev protease for 4 hours at room temperature. The complex was purified by size exclusion chromatography using superdex200 gel filtration column. Selenomethionine-Kap β 2 and selenomethionine-M9NLS derivatives were purified and assembled as for the native proteins. All complexes were concentrated to 25 mg/ml for crystallization.

Crystallization and Data Collection

Crystallization of the complex has been done by Brittany Lee in Chook Lab with help from Katie Suel. Native Kap β 2-M9NLS complex was crystallized by vapor diffusion in condition: 40 mM MES pH 6.5, 3M potassium formate and 10% glycerol. The crystals were cryoprotected in crystallization condition and flash frozen in liquid propane. These crystals initially diffracted to 3.5 Å resolution.

It is proposed that karyopherins bind to small hydrophobic peptide sequences found extensively on nucleoporins (Proteins of the nuclear pore complex) on the nuclear pore complex (see chapter 1). These repeat sequences are selectively present as hydrophobic FXFG sequence repeats. To enhance the stability of the crystals, the crystals were soaked into reservoir solution which contained 0.7 mM of a 12-residue FXFG-peptide (sequence: TGGFTFGTAKTA). These soaks gave improved diffraction to 3.05 Å. Brittany Lee, Katie Suel and I participated in data collection of these crystals. Data from an FXFG-soaked crystal was collected on

the SBC-CAT beamline 19-ID at the Advanced Photon Source at the Argonne National Laboratory. Data was processed using HKL2000 (Otwinowski and Minor 1997).

Phasing of Kap β 2dloop.M9NLS structure

Selenomethionine derivative crystals were also obtained by vapor diffusion in a similar condition. (Reservoir solution: 0.1M Tris 8.0, 3M potassium formate and 15% glycerol), soaked in FXFG-peptide and diffracted to 3.3 Å. Single-wavelength anomalous dispersion (SAD) data was collected on SBC-19-ID and processed with HKL2000 (Otwinowski and Minor 1997). Initial phasing of the data was done with SAD data from se-met derivative crystals. Model building was done extensively by Dr. Yuh Min Chook. I performed molecular replacement with program Phaser using the partially built se-met model from the unliganded Kap β 2. The solution produced decent looking map. The rest of the model was completed with good refinement statistics.

Native Kap β 2-M9NLS crystals (space group C2, unit cell parameters a=152.0 Å, b=154.1 Å, c=141.7 Å and β =91.7°) contain two complexes in the asymmetric unit. Selenomethionine Kap β 2-M9NLS also crystallized space group C2, but has a significantly different unit cell length in its a axis (unit cell parameters: a=155.6

Å, $b=154.6$ Å, $c=141.6$ Å and $\beta=91.6^\circ$; Table 2-1). Native Patterson maps indicate that the two complexes in the asymmetric unit are related by pseudo-translation along the crystallographic c axis. Molecular replacement trials using the Kap β 2-Ran structure were unsuccessful but SAD phasing followed by solvent flipping, both using the program CNS produced interpretable electron density maps (Brunger, D. et al. 1998). A model comprising 90% of Kap β 2 was built using O (Jones, Cowan et al. 1991) but electron density for the substrate remained uninterpretable even though M9NLS residue M276 could be clearly placed using a selenium site. The partial SAD-phased models of substrate complex and unliganded Kap β 2 were used as a search model for molecular replacement using the program Phaser with the higher resolution native dataset (McCoy, Grosse-Kunstleve et al. 2005). Positional refinement using REFMAC5 (CCP4 1994) followed by solvent flipping using CNS (Brunger, D. et al. 1998) yielded electron density maps that allowed 97% of Kap β 2 to be built. The density was further improved by rigid body, positional and simulated annealing refinement of Kap β 2 alone, using the programs CNS (Brunger, D. et al. 1998). The F_o-F_c map plotted at 2.5 sigma clearly showed strong density for M9NLS residues 267-289 in the complex I, and residues 263-289 in complex II (Figure 2-1C). Even though soaking the crystals in FXFG peptide improved diffraction, no density was observed for the FXFG peptide. The final refined model shows good stereochemistry with R_{factor} of 24.2% and R_{free} of 27.2%.

NLS-mapping, site directed mutagenesis and Kap β 2 binding assays

Biding experiments were performed by Brittany Lee and Yuh Min Chook. cDNA for hnRNPs F, M, PQBP-1, EWS, SAM68, HMBA-inducible protein, YBP1, FUS, DDX3, Clk3, Sox14 and WBS16 were obtained from Open Biosystems. cDNA for HCC1 and RB15B were obtain by PCR from a human fetal thymus cDNA library (Clontech). The full-length proteins as well as fragments of the proteins were sub-cloned using PCR into pGEX-Tev vector. Expression constructs for NLSs of cyclin T1 and CPSF6 were generated using synthetic complementary oligonucleotides coding for the 28-mer peptides. Single, double and triple mutations to alanine residues were performed using the Quickchange method (Stratagene), and all constructs were confirmed by nucleotide sequencing. Substrate proteins were expressed in *E. coli* BL21 (DE3) cells. GST-M9NLS was expressed at 37°C, GST- Kap β 2 was expressed at 30°C and the other substrates were expressed at 25°C, and all were purified using glutathione sepharose (GE Healthcare).

In each binding reaction involving new NLSs, mutant NLSs and new Kap β 2 substrates, approximately 18 μ g of Kap β 2 was added to 5-10 μ g of GST-substrate immobilized on glutathione sepharose followed by extensive washing of the beads with buffer containing 20 mM Hepes pH 7.3, 110 mM potassium acetate, 2 mM

DTT, 1 mM EGTA, 2 mM Magnesium acetate and 20% glycerol. Immobilized proteins were visualized using SDS-PAGE and Coomassie Blue staining. 3-5 fold molar excess of RanGTP (compared to Kap β 2) is also used in some binding assays. Binding assays involving mutants of Kap β 2 were performed similarly, with each reaction using approximately 10 μ g of MBP-M9NLS added to 5-10 μ g of GST-Kap β 2.

Quantitation of binding affinity with Isothermal Titration Calorimetry (ITC)

Binding affinities of wild type and mutant MBP-M9NLS to Kap β 2 were measured using standard ITC experiments. I performed all ITC experiments with help from Katie Suel. The ITC experiments were done using a MicroCal Omega VP-ITC calorimeter (MicroCal Inc., Northampton, MA). Proteins were dialyzed overnight against buffer containing 20 mM Tris pH 7.5, 100 mM NaCl and 2 mM β -mercaptoethanol. 100-500 μ M Wild type and mutant MBP-M9NLS proteins were titrated into a sample cell containing 10-100 μ M full length Kap β 2. Most ITC experiments were done at 20°C with 35 rounds of 8 μ l injections. ITC experiments involving wild type M9NLS were similar, but with 56 rounds of 5 μ l injections to provide better resolution of the curve. Data was plotted and analyzed using MicroCal Origin software version 7.0, with a single binding site model.

Results and Discussions

Kap β 2-M9NLS Complex Structure Overview

Kap β 2 is a superhelical protein with 20 HEAT repeats. It is almost exclusively α -helical except for a 62-residue loop in repeat 8 (H8 loop; Figure 2-1A). Each repeat consists of two antiparallel helices A and B, each lining the convex and concave side of the superhelix respectively (Chook and Blobel 1999; Chook, Jung et al. 2002). The Kap β 2-M9NLS crystals contain a Kap β 2 mutant with a truncated H8 loop bound to residues 257 to 305 of hnRNP A1 (Figure 2-1B). Biochemical studies showed that the loop neither hinders nor is necessary for substrate binding. However, it is sensitive to proteolytic degradation in substrate-bound Kap β 2, suggesting structurally flexibility (Chook, Jung et al. 2002). In the final Kap β 2 construct, the H8 loop was truncated (a GGSGGSG linker replaces residues 337-367) to minimize disorder in the crystal. The Kap β 2-M9NLS crystal structure was solved to 3.05 Å resolution (Table 2-1, PDB ID code 2H4M).

The asymmetric unit of the crystal contains two Kap β 2-M9NLS complexes (I and II). All residues in both Kap β 2s are modeled except for three short loops at the N-termini, H8 loop residues 320-337 and the engineered GGSGGSG H8 loop

linker (disordered regions are indicated by dashes in Figures 2-1A, 2-5A and 2-5B). Substrate residues 267-289 are observed in complex I, while additional substrate residues 263-266 are modeled in complex II (Figure 2-1C). Thus, the latter is used in structural analysis and discussion below. HEAT repeats 5-20 share similar conformations in both complexes (rmsd 1.7 Å). In contrast, HEAT repeats 1-4 diverge to a distance of 7 Å at their N-termini with high average B-factors (99 Å^2 for complex I and 135 Å^2 for complex II), suggesting inherent conformational flexibility in this region of Kap β 2.

The 20 HEAT repeats of the Kap β 2-M9NLS complex form an almost perfect superhelix (pitch $\sim 72 \text{ Å}$, diameter $\sim 60 \text{ Å}$ and length $\sim 111 \text{ Å}$; Figure 2-1A). The superhelix can also be described as two overlapping arches, with the N-terminal arch spanning HEAT repeats 1-13 and the C-terminal arch spanning repeats 8-20. In the Kap β 2-Ran complex, RanGTP binds in the N-terminal arch (Chook and Blobel 1999). Here, we observe that M9NLS binds in the C-terminal arch (Figures 2-1A and 2-1C).

Table 2-1) Data collection and refinement statistics for crystals of full length native and selenomethionine derivatives of Kap β 2.M9NLS complex

Space group:	C2	C2
Unit Cell Parameters:		
a, b, c (Å)	152.0,154.1,141.7	155.7, 154.6, 141.6
α , β , γ (°)	90.0, 91.75, 90.0	90.00, 91.56, 90.00
Resolution (Å)	100-3.05	100-3.3
Rsym	0.055 (0.429)*	0.103 (0.50)*
I / σ _I	24.7 (2.0)*	21.5 (2.1)*
Completeness (%)	99% (92.8%)*	98.5% (91.5%)*
Redundancy	4.6 (4.1)	4.9 (4.7)*
Refinement:		
Rfactor† / Rfree†	0.242/0.272	
rmsd from ideal bond lengths (Å)	1.136	
rmsd from ideal bond angles (°)	0.0074	
Ramachandran Plot:		
Most favoured regions (%)	90. 4	
Allowed regions (%)	9.6	
Model:	Average B Factor (Å ²):	
Kapβ2 Chain A (Residues: 6-36, 44-77, 80-319, 368-890)	70.0	
M9NLS Chain C (263-289)	73.7	
Kapβ2 Chain B (6-36, 44-55, 59-75, 80-319, 369-890)	81.5	
M9NLS Chain D (266-289)	67.4	

$R_{sym} = \sum_h \sum_i | (I_i(h) - \langle I(h) \rangle) | / \sum_h \sum_i I_i(h)$; $I_i(h)$ is the i^{th} measurement of reflection h and $\langle I(h) \rangle$ is the weighted mean of all measurements of h .

* values in parentheses are calculated for data in the highest resolution shell.

[†] R-factor = $\sum_h | |F_{obs}(h)| - |F_{calc}(h)| | / \sum_h |F_{obs}(h)|$. R_{free} is calculated with 5% of the data.

The Kap β 2-M9NLS binding interface

M9NLS binds in extended conformation to line the concave surface of C-terminal arch of Kap β 2 (Figure 2-1A). Its peptide direction is antiparallel to that of the karyopherin superhelix, and substrate buries 3432 Å² of surface area in both binding partners. Tracing M9NLS from N- to C-terminus, residues 263-266 interact with helices H18A, H19A and H20B of Kap β 2 while residues 267-269 drape over the intra-HEAT 18 loop into the C-terminal arch of the karyopherin. The rest of M9NLS follows the curvature of the C-terminal arch to contact B helices of repeats 8-17 (Figures 2-1A and 2-2A). The substrate interface on Kap β 2 comprises ~30% of the concave surface of the C-terminal arch, which is relatively flat and devoid of deep pockets or grooves. Most of this surface, which includes the M9NLS interface, is also highly acidic (Figure 2-2B).

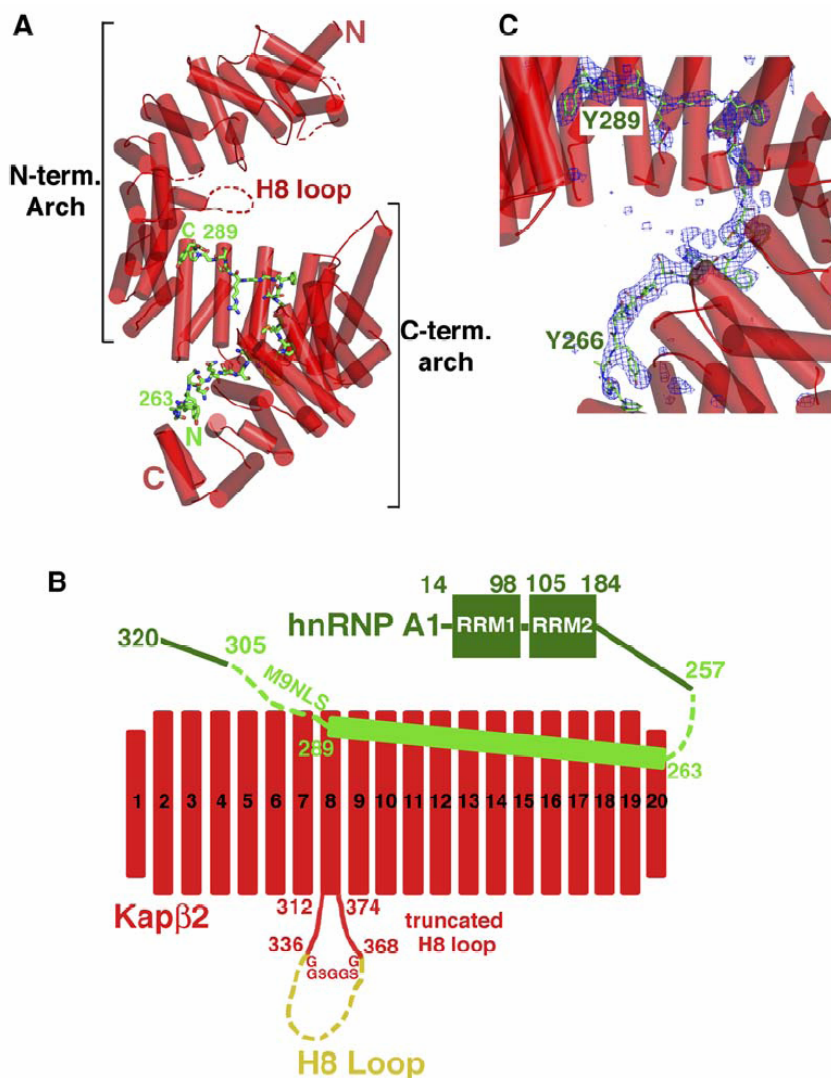


Figure 2-1) Crystal Structure of the Kap β 2-M9NLS Complex (Adapted from (Lee, Cansizoglu et al. 2006)) **A)** Ribbon diagram of the Kap β 2-M9NLS complex with Kap β 2 in red (α helices represented as cylinders and structurally disordered loops as red dashes) and M9NLS shown as a stick figure (carbon: green, oxygen: red, nitrogen: blue, and sulfur: orange). **B)** The 20 HEAT repeats and H8 loop of Kap β 2 used in structural analyses (red) and M9NLS (light green) within hnRNP A1 (green). The deleted portion of the H8 loop is in yellow. **C)** The M9NLS binding site with Fo-Fc map (2.5 σ) calculated using Kap β 2 alone (blue mesh), drawn with PYMOL(DeLano 2002).

M9NLS forms an extensive network of polar and hydrophobic interactions with Kap β 2, involving both the main chain and sidechains of the substrate (Figure 2-2A). Most of the substrate interface on Kap β 2 is acidic with the exception of several scattered hydrophobic patches. At the N-terminus of M9NLS, residues 263-266 contact a hydrophobic patch on Kap β 2 helices H19A and H20B (Figure 2-2B, left). In the central region, a hydrophobic stretch ²⁷³FGPM²⁷⁶ contacts hydrophobic Kap β 2 residues I773 and W730 (Figures 2-2B and 2-2C). Farther C-terminus, F281 binds near a hydrophobic patch formed by Kap β 2 residues F584 and V643 (Figure 2-2B, center) and finally, the C-terminal ²⁸⁸PY²⁸⁹ residues bind a large hydrophobic swath that includes Kap β 2 residues A380, A381, L419, I457 and W460 (Figures 2-2B, right and 2-2D). Despite the extensive acidic interface on Kap β 2, there are only two basic residues in M9NLS. R284 forms salt links with Kap β 2 residues E509 and D543, and the sidechain of K277 is not observed.

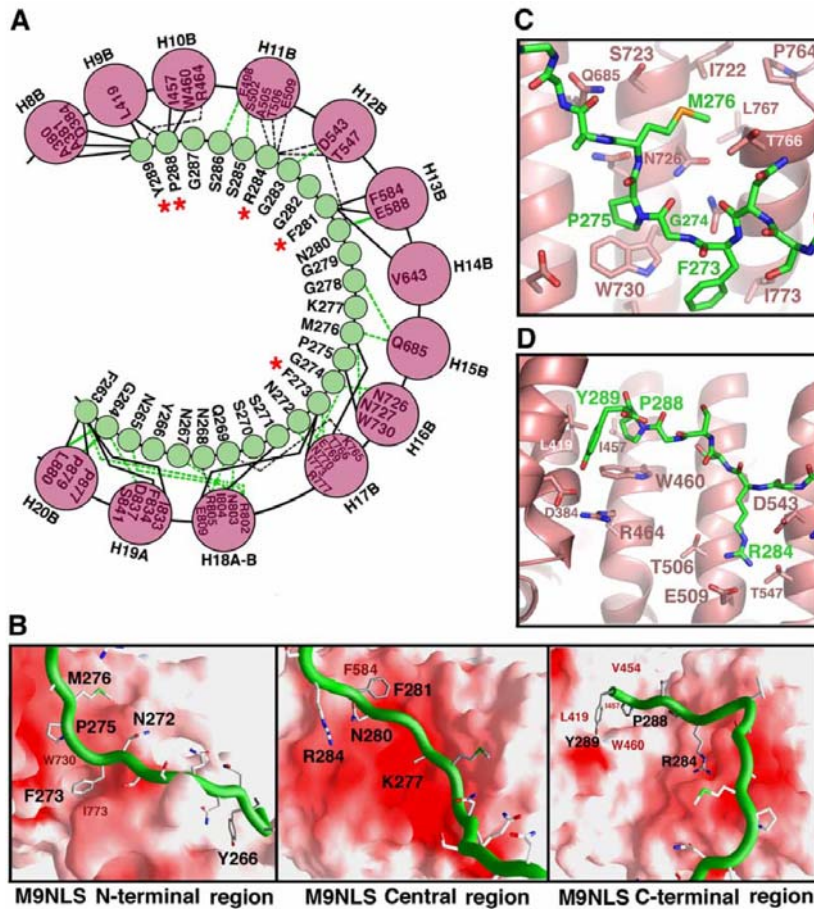


Figure 2-2) Kap β 2-M9NLS Interactions **A)** Kap β 2-M9NLS contacts (<4.0 Å) with M9NLS residues in green circles and Kap β 2 helices as pink circles. Contacts involving main chain atoms of M9NLS are shown with green lines. Contacts involving M9NLS side chains are shown with black lines. Solid lines are hydrophobic contacts and dashed lines are polar contacts. Red asterisks label M9NLS residues that make two or more side-chain contacts in both complexes in the asymmetric unit. **B)** The Kap β 2-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 Kap β 2 electrostatic potential is mapped onto its surface, all drawn using GRASP(Nicholls, Sharp et al. 1991). Red indicates negative electrostatic potential, white neutral, and blue positive. Residues in the hydrophobic patches of Kap β 2 are labeled in red and M9NLS residues labeled in black. **C)** Interactions between Kap β 2 (red) and substrate at M9NLS (green) residues 273FGPM276, drawn using PYMOL(DeLano 2002). **D)** Interactions between Kap β 2 (red) and M9NLS (green) at the C terminus of the substrate, drawn using PYMOL(DeLano 2002).

Distribution of binding energy along M9NLS

In order to understand the distribution of binding energy along M9NLS, we measured dissociation constants (K_D s) of a series of M9NLS mutants binding to Kap β 2 using isothermal titration calorimetry (ITC). The results of the binding studies using MBP-fusion proteins of M9NLS residues 257-305 and wild type Kap β 2 are summarized in Table 2-2 and Figure 2-8. Wildtype M9NLS binds Kap β 2 with a K_D of 42 nM. This ITC-measured affinity is somewhat lower than the previous K_D of 2 nM measured by fluorescence titration, but may be explained by the presence of both a covalently attached aromatic fluorophore and a significantly longer M9NLS spanning residues 238-320 in the earlier studies (Chook, Jung et al. 2002). Substrate residues that make two or more sidechain contacts with Kap β 2 (F273, F281, R284, P288 and Y289) were systematically mutated to alanines. Additional residues G274, P275 and M276 were also mutated given their implied importance in yeast-two-hybrid studies (Bogerd, Benson et al. 1999).

Table 2-2) Kap β 2 Binding to M9NLS Mutants: Dissociation Constants by Isothermal Calorimetry

MBP-M9NLS(257-305) proteins	KD
Wild type	42 \pm 2 nM
F273A	61 \pm 10 nM
G274A	746 \pm 63 nM
P275A	74 \pm 5 nM
M276A	83 \pm 17 nM
F281A	56 \pm 11 nM
R284A	92 \pm 9 nM
P288A	158 \pm 20 nM
Y289A	133 \pm 21 nM
P288A/Y289A	136 \pm 8 nM
R284A/P288A/Y289A	461 \pm 27 nM
G274A/P288A/Y289A	5.9 \pm 0.7 μ M

G274A is the only single mutant that shows significant (18-fold) decrease in Kap β 2 binding (Table 2-2). Single mutants of C-terminal residues P288 and Y289 follow with modest decreases of 3-4 fold. Thus, it appears that M9NLS binds Kap β 2 in a mostly distributive fashion, with a strict requirement for glycine at position 274 and modest though possibly important energetic contributions from C-terminal residues P288 and Y289. The importance of the PY motif is

suggested in the R284/P288/Y289 and G274/P288/Y289 triple mutants where 10-fold and 140-fold decreases were observed, respectively. Both triple mutants show non-additivity in their binding energies when compared with single G274A, R284A and the double PY mutants, suggesting cooperativity between the C-terminal PY motif and both upstream binding sites at R284 and G274. The significance the G274A mutation had previously been reported in both Kap β 2-binding and nuclear import assays (Nakielny, Siomi et al. 1996; Fridell, Truant et al. 1997). The alpha carbon of G274 is in close proximity to neighboring substrate sidechains F273 and P275 as well as Kap β 2 residue W730, such that a sidechain in position 274 may result in a steric clash (Figure 2-2C).

The important energetic contributions of the substrate's C-terminal PY motif and its central G274 residue are also supported by mutations of interacting residues in Kap β 2. Double and triple Kap β 2 mutants, W460A/W730A and I457A/W460A/W730A, both show significant decreases in Kap β 2 binding (Figure 2-9). I457 and W460 interact with the substrate PY motif while W730 makes a hydrophobic contact with substrate P275 and is also close to G274 (Figures 2-2C and 2-2D).

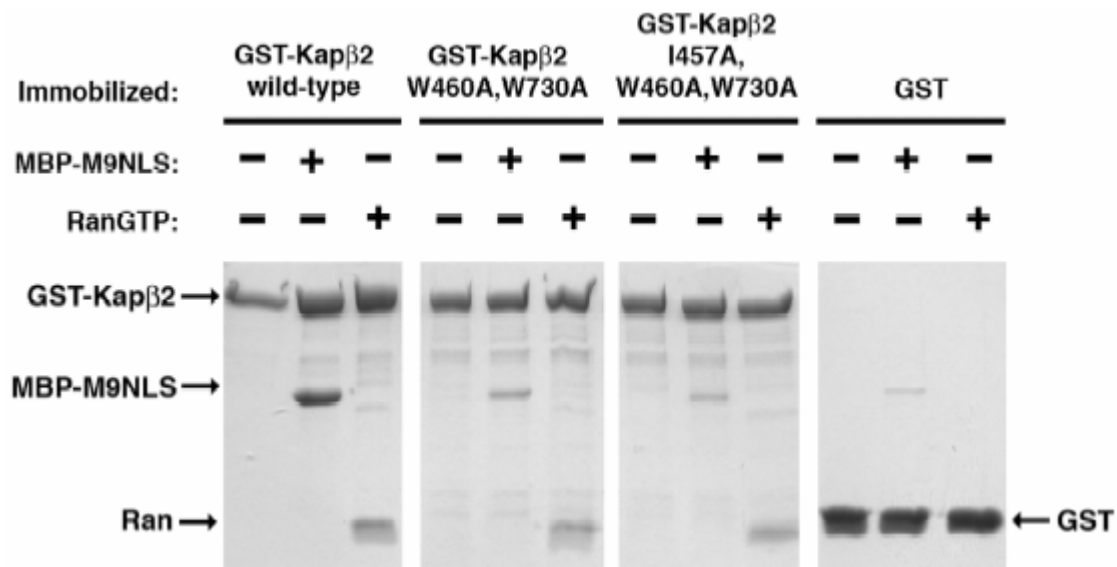


Figure 2-3) Binding studies of MBP-M9NLS and immobilized Kap β 2 mutants. Control experiments were also performed using immobilized Kap β 2 proteins and RanGTP.

Rule 1: NLS is structurally disordered in substrate

The extended conformation of the 26-residue M9NLS results in a linear epitope that traces a path of ~ 110 Å. The structure of the bound substrate suggests that an NLS recognized by Kap β 2 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 substrate. The prediction of this NLS requirement is further supported by the fact that all seven known NLSs in Kap β 2

substrates occur within sequences with high probability of structural disorder (> 0.7) calculated by the program DISEMBL(Linding, Jensen et al. 2003). All seven NLSs are found either in loop regions between the RNA binding or other folded domains or at the termini of the substrates.

Rule 2: Overall positive charge for NLS is preferred

A second requirement for an NLS recognized by Kap β 2 emerges from the observation that Kap β 2's substrate interface is highly negatively charged. An acidic peptide would likely not bind due to electrostatic repulsion, while an NLS with overall positive charge would most likely be favored. Examination of all known Kap β 2 NLSs indicates overall basic character spanning at least 30 residues in six of seven cases (Figure 2-4A). In addition, regions that flank the NLSs most likely also contribute favorably to electrostatics. For example, although the TAP-NLS sequence delineated in Figure 2-4A has slightly more acidic than basic residues, flanking regions are highly basic and may ultimately contribute to overall basic character to promote Kap β 2 binding. The importance of basic flanking regions is also observed in hnRNP A1. Here, the entire 135-residue C-terminal tail of the substrate has overall positive charge. A recent study showed that following osmotic shock stress in cells, four serine residues C-terminally adjacent to the M9NLS are phosphorylated, resulting in decreased binding to

Kap β 2 and accumulation of hnRNP A1 in the cytoplasm (Allemand, Guil et al. 2005). Phosphorylation of the M9NLS-flanking serines may decrease the basic character of M9NLS and thus modulate interactions with Kap β 2.

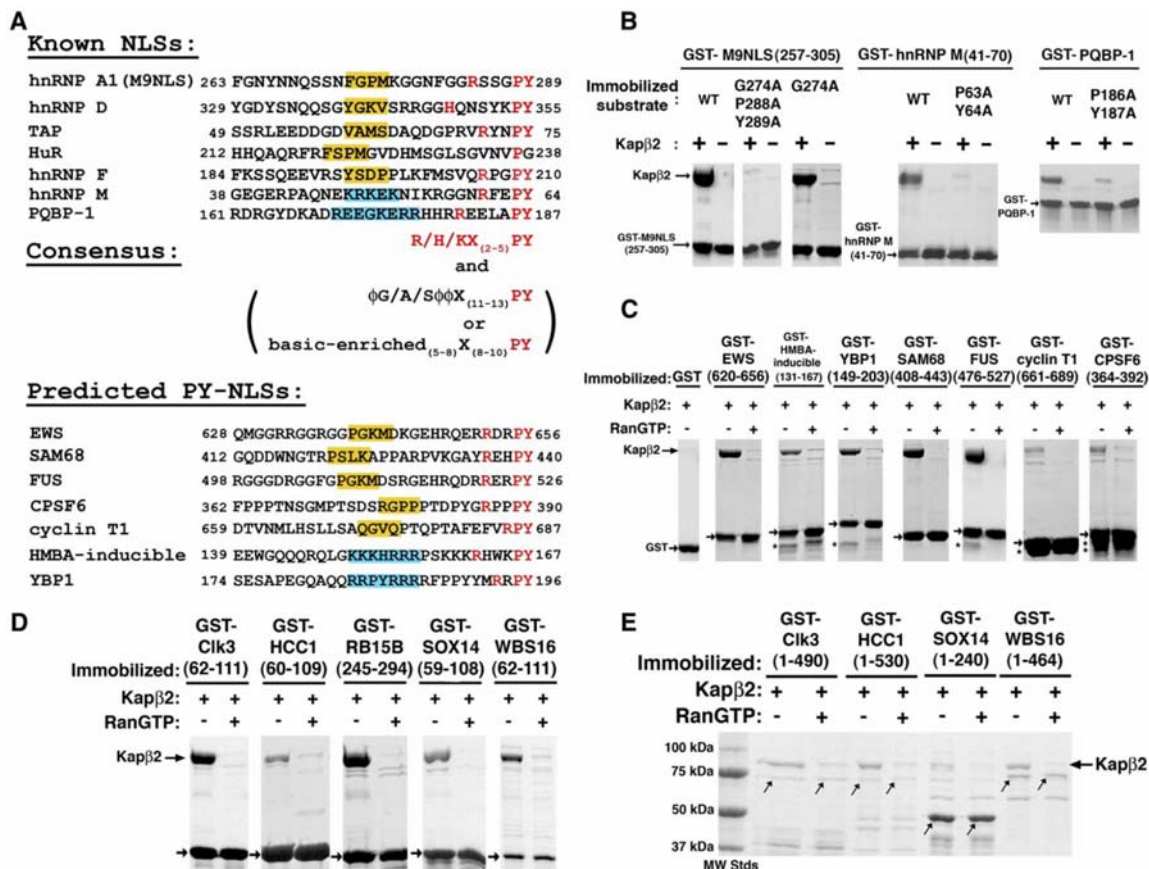


Figure 2-4) Consensus Sequences of NLSs Recognized by Kap β 2: **A)** Alignment of all known (top) and predicted (bottom) NLSs recognized by Kap β 2 at conserved PY residues. NLSs in known Kap β 2 substrates are predicted by the presence of the R/K/H-X₍₂₋₅₎-P-Y C-terminal motifs (red) within structurally disordered and positively charged regions of 30 amino acids. Central hydrophobic motifs G/A/S (is a hydrophobic side chain) are shaded yellow. Central basic motifs are shaded blue. **B)** Binding assays of Kap β 2 and immobilized alanine mutants of M9NLS, PQBP-1, and NLS-containing fragments of hnRNP M. Bound proteins are visualized with Coomassie blue. **C)** Binding assays of predicted NLSs from known Kap β 2 substrates EWS, HMBA-inducible protein, YBP1, SAM68, FUS, Cyclin T1 and CPSF6. Kap β 2 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. **D)** Five predicted Kap β 2 substrates (Clk3, HCC1, RB15B, Sox14, and WBS16) are validated experimentally. GST NLSs (arrows) are immobilized on glutathione sepharose. **E)** Binding assays of full-length substrates Clk3, HCC1, Sox14, and WBS16 to Kap β 2. Expression of recombinant full-length RB15B was not successful. Coomassie-stained bands at the size of the GST substrates are labeled with arrows. Lower-molecular-weight proteins are likely degraded substrates.

Rule 3: Consensus sequences for the NLS

All seven characterized NLSs recognized by Kap β 2 exist in structurally disordered regions suggesting that this class of NLS is represented by linear epitopes and not folded domains. However, apparent sequence diversity among previously characterized NLSs from hnRNP A1, HuR, TAP and JKTBP homologs had prevented delineation of a consensus sequence that could be used to identify new NLSs or substrates. However, despite apparent NLS diversity, mutagenesis, structural and sequence analysis have resulted in identification of two regions of conservation within the sequences.

The first region of conservation is found at the C-terminus of the NLSs. Mutagenesis of M9NLS suggested the importance of its C-terminal PY motif (Table 2-2). Sequence examination of previously characterized NLSs from hnRNP D, HuR and TAP as well as the newly characterized NLSs of hnRNP F, M and PQBP-1, identified consecutive PY residues in six of the seven sequences (Figure 2-4A). Mutations of the PY residues in PQBP-1 and hnRNP M also decreased Kap β 2 binding suggesting that they make energetically important contacts (Figure 2-4B). Mutations of the PY motif in JKTBP proteins and M9NLS were also previously shown to inhibit nuclear import (Suzuki, Iijima et al. 2005; Iijima, Suzuki et al. 2006). In addition, we observe that a basic residue

is always found several residues N-terminal of the PY sequence, consistent with an adjacent acidic surface on Kap β 2 (Figures 2-2B, 2-2D and 2-3A). Based on these observations, we propose a C-terminal consensus sequence R/K/H-X₍₂₋₅₎-P-Y (where X is any residue) for NLSs recognized by Kap β 2. We refer to this class of NLSs as PY-NLSs.

A second region of conservation within the PY-NLSs is found in the central region of the peptides. Examination of the central region divides the seven PY-NLSs into two sub-classes. The first sub-class includes M9NLS and NLSs of hnRNP D, F, TAP and HuR, where four consecutive predominantly hydrophobic residues are located 11-13 residues N-terminal to the PY residues (Figure 2-4A). We refer to this sub-class of sequences as hydrophobic PY-NLSs or hPY-NLSs. In contrast, the central regions of NLSs from hnRNP M and PQBP-1 are virtually devoid of hydrophobic residues but are instead enriched in basic residues. They appear to represent a distinct sub-class of PY-NLSs that we call the basic PY-NLSs or bPY-NLSs.

The central hydrophobic motif in M9NLS spans residues ²⁷³FGPM²⁷⁶ previously found in yeast two-hybrid and mutagenesis analysis to be important for import by Kap β 2, and a consensus sequence of Z-G-P/K-M/L/V-K/R (where Z is a hydrophobic residue) was previously suggested (Bogerd, Benson et al. 1999). The mutagenesis-derived consensus holds in the context of the M9NLS sequence, but does not describe NLSs in other Kap β 2 substrates. A loose consensus of ϕ -

G/A/S- ϕ - ϕ (where ϕ is a hydrophobic sidechain) seems more appropriate upon comparison of the five central hydrophobic motifs in hnRNPs A1, D, F, TAP and HuR (Figure 2-4A). The Kap β 2-M9NLS structure explains preferences for hydrophobic sidechains in positions 1, 3 and 4 as well as small or no sidechain in position 2. Position 1 in M9NLS is F273, which occupies a hydrophobic pocket formed by Kap β 2 residues W730 and I773 (Figure 2-2C). Position 3 is occupied by P275, which stacks on top of the indole ring of Kap β 2 W730, and M276 in position 4 binds a small hydrophobic patch on Kap β 2 formed by I722, P764, L766 and the C β of S767. Thus, hydrophobic or long aliphatic sidechains at positions 1, 3 and 4 in other hydrophobic hPY-NLSs would provide energetically favorable hydrophobic contacts with Kap β 2. Mutagenesis of M9NLS suggests a strict requirement for glycine at position 2 (residue G274 in M9NLS) of the central hydrophobic motif. G274 is surrounded by adjacent substrate residues F273, P275 and Kap β 2 residue W730, suggesting that the strict requirement for glycine is likely heavily dependent on the identity of adjacent substrate residues. Nevertheless, hydrophobic neighbors, even those not as bulky as F273 and P275 in M9NLS, will likely still not accommodate large sidechains in position 2.

The Kap β 2-M9NLS structure provides some suggestion for the how the central basic motif in the bPY-NLSs could be accommodated. In the structure, the M9NLS hydrophobic motif interacts with Kap β 2 hydrophobic residues that are surrounded by numerous acidic residues (Figures 2-2B and 2-2C). Thus, the

highly acidic substrate interface on Kap β 2 that contacts the central region of an NLS should also be able to interact favorably with numerous basic sidechains. It is possible that the central basic and hydrophobic motifs in the two sub-classes of PY-NLSs may take slightly different paths on Kap β 2. Structures of Kap β 2 bound to bPY-NLSs will be necessary to understand the difference between the two subclasses of PY-NLSs.

The NLS rules are predictive

We have examined the sequences of eight recently identified Kap β 2 substrates: Ewing Sarcoma protein (EWS), HMBA-inducible protein, Y-box binding protein 1 (YBP1), SAM68, FUS, DDX3, CPSF6 and cyclin T1 (Guttinger, Muhlhauser et al. 2004), and found the C-terminal R/K/H-X₍₂₋₅₎-P-Y consensus within structurally disordered and positively charged regions of seven of them. The predicted NLSs for EWS, HMBA-inducible protein, YBP1, SAM68, FUS, CPSF6 and cyclin T1 are listed in the bottom half of Figure 2-4A. The predicted signals in EWS, SAM68, FUS, CPSF6 and Cyclin T1 are hPY-NLSs and those from HMBA-inducible protein and YBP1 are bPY-NLSs (Figure 2-5). The easily-detected PY motif is absent from DDX3, and we have not been able to show direct binding of DDX3 to Kap β 2 (data not shown). Thus, DDX3 may not be a substrate of Kap β 2, but may enter the nucleus by binding to a bona-fide Kap β 2

substrate. All seven predicted NLSs bind Kap β 2 and are dissociated from the karyopherin by RanGTP, consistent with NLSs imported by Kap β 2 (Figure 2-4C). The NLSs of cyclin T1 and CPSF6 bind Kap β 2, but more weakly than other substrates. It is not clear if this is due to proteolytic degradation of the substrates or to poor central hydrophobic motifs (Figures 2-4A, 2-4C and 2-5). Confirmation of these seven NLSs indicates that the three rules for NLS recognition by Kap β 2 described above are predictive.

hPY-NLSs:

	central hydrophobic motif	C-term. PY motif
hnRNP A1 (M9NLS)	FGPM-X ₁₁ -PY	
hnRNP D	YGKV-X ₁₁ -PY	
TAP	VAMS-X ₁₁ -PY	
HuR	FSPM-X ₁₃ -PG	
hnRNP F	YSDP-X ₁₁ -PY	
EWS	PGKM-X ₁₂ -PY	
SAM68	PSLK-X ₁₄ -PY	
FUS	PGKM-X ₁₃ -PY	
CPSF6	RGPP-X ₉ -PY	
cyclin T1	QGVQ-X ₁₁ -PY	
consensus sequence	ϕ G/A/S ϕ ϕ X ₍₉₋₁₄₎ PY	

bPY-NLSs:

	central basic motif	C-term. PY motif
hnRNP M	KRKEK-X ₁₀ -PY	
PQBP-1	REEGKERR-X ₈ -PY	
HMBA-inducible	KKKHRRR-X ₉ -PY	
YBP1	RRPYRRRR-X ₈ -PY	
	basic-enriched ₍₅₋₈₎ X ₍₈₋₁₀₎ PY	

Figure 2-5) Summary of determined hydrophobic and basic PY-NLS motifs found on different proteins

Mechanism of Ran-mediated substrate dissociation from Kap β 2

The interaction of RanGTP with Kap β 2 to dissociate substrates in the nucleus is a crucial step in nuclear import. Structural comparison of Kap β 2s in the M9NLS and RanGTP complexes (Chook and Blobel 1999) show large differences in their H8 loops (Figure 2-6A), and finally reveal the mechanism of Ran-mediated substrate dissociation. In the Kap β 2-Ran structure, the H8 loop makes extensive contacts with both Ran and the Kap β 2 C-terminal arch (Figures 2-6A and 2-6B) (Chook and Blobel 1999). In fact, much of the H8 loop is sequestered in the C-terminal arch such that loop residues 338-350 occupy the same binding site as M9NLS residues 268-281. In contrast, proteolysis studies have suggested that the loop is exposed when Ran is absent (Chook, Jung et al. 2002) and this is confirmed by the Kap β 2-M9NLS structure. Even though the H8 loop in the M9NLS complex is truncated, only 14 of its 32 residues are observed, indicating disorder in much of the loop. Ordered loop residues include 312-319 that emerge from helix H8A and residues 369-374 that precede helix H8B (Figures 2-7A and 2-7B). The former are in similar positions in both complexes, but the latter has shifted to direct the loop away from the arch in the substrate complex (Figures 2-6A and 2-6B). In summary, the concave surface of the C-terminal arch is free to bind substrate when Ran is absent, but the H8 loop occupies the substrate binding

site when Ran is present. Interestingly, most of the substrate binding site remains unchanged in both ligand-bound states with repeats 9-17 superimposing well at rmsd of 1.2 Å (Figure 2-6A). The mechanism of Ran-mediated substrate dissociation described here is a thermodynamic one. Ran may increase the dissociation rate of substrate, thus accelerating its release from Kap β 2. Alternatively, the system is limited by the intrinsic dissociation rate of the substrate, and Ran-induced changes in the loop prevent substrate rebinding once dissociation has occurred.

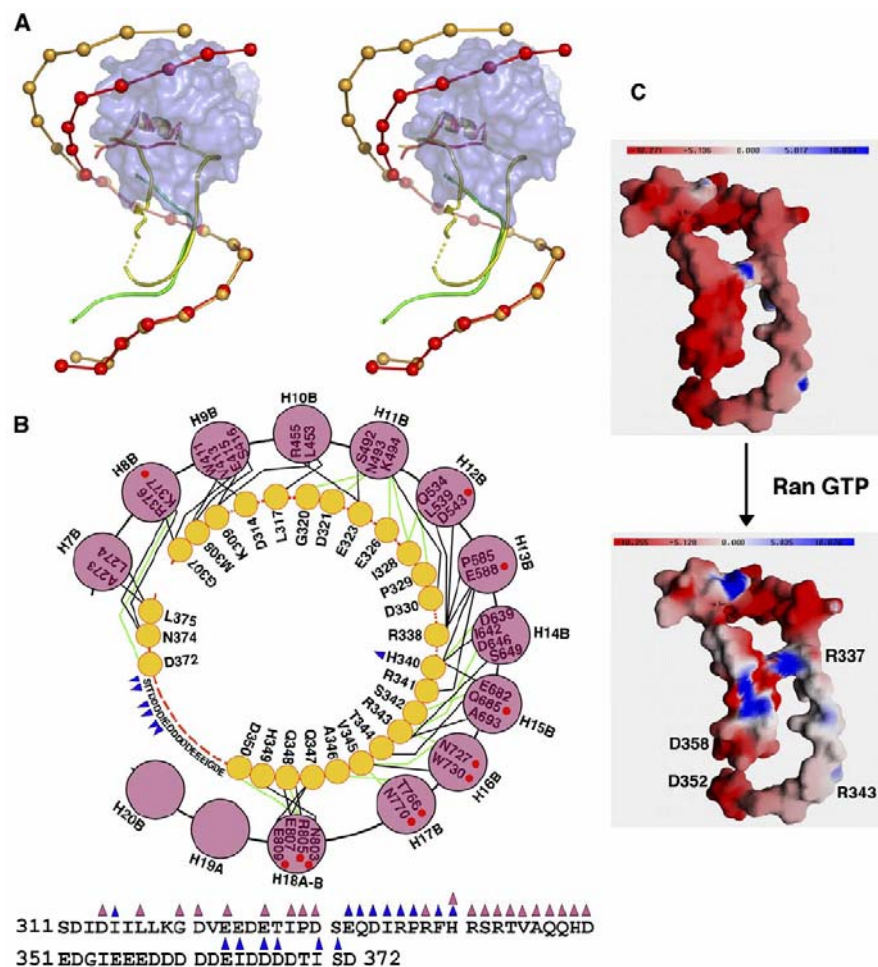


Figure 2-6) Comparison of Kapβ2-M9NLS and Kapβ2-Ran Complexes: **A)** Stereo diagram of Kapβ2-M9NLS complex superimposed on the Ran complex. Kapβ2s are drawn as spheres at the geometric center of each HEAT repeat, and the H8 loops are drawn as ribbons. In the M9NLS complex, Kapβ2 is red and substrate green. In the Ran complex, Kapβ2 is light brown and its H8 loop is yellow. The molecular surface of RanGTP is shown in blue. **B)** Contacts (<4.0 Å) between the H8 loop and the C-terminal arch of Kapβ2 in the Ran state with the sequence of the H8 loop shown at bottom. Yellow circles are loop residues that contact the Kapβ2 arch and pink circles are Kapβ2 helices. Red dashed lines indicate intervening loop residues that do not contact the Kapβ2 arch. Blue triangles label residues that contact Ran, pink triangles label residues that contact the Kapβ2 arch, and red circles label Kapβ2 residues that also contact M9NLS. Polar contacts are shown with green lines and hydrophobic contacts with black lines. **C)** Electrostatic surface potential of the H8 loop in the presence and absence of RanGTP, drawn with GRASP(Nicholls, Sharp et al. 1991). Top shows molecular surface of the isolated H8 loop (Kapβ2 residues 310–372, Kapβ2-Ran complex 1QBK). RanGTP and the Kapβ2 superhelix are omitted from the electrostatic calculation to approximate charges of the loop in the absence of Ran. Bottom shows molecular surface of the H8 loop with electrostatic surface potential calculated using both RanGTP and the H8 loop to represent the Ran bound state.

Despite extensive spatial overlap between the Ran-bound H8 loop and M9NLS, they share no obvious sequence similarity. This is not surprising since they bind in antiparallel direction to each other and their backbones deviate in path even where spatial overlap is greatest (loop residues 338-350 and M9NLS residues 268-281; Figure 2-9). However, the H8 loop obviously contains a linear epitope that binds Kap β 2 and raises the possible existence of a different class of NLSs.

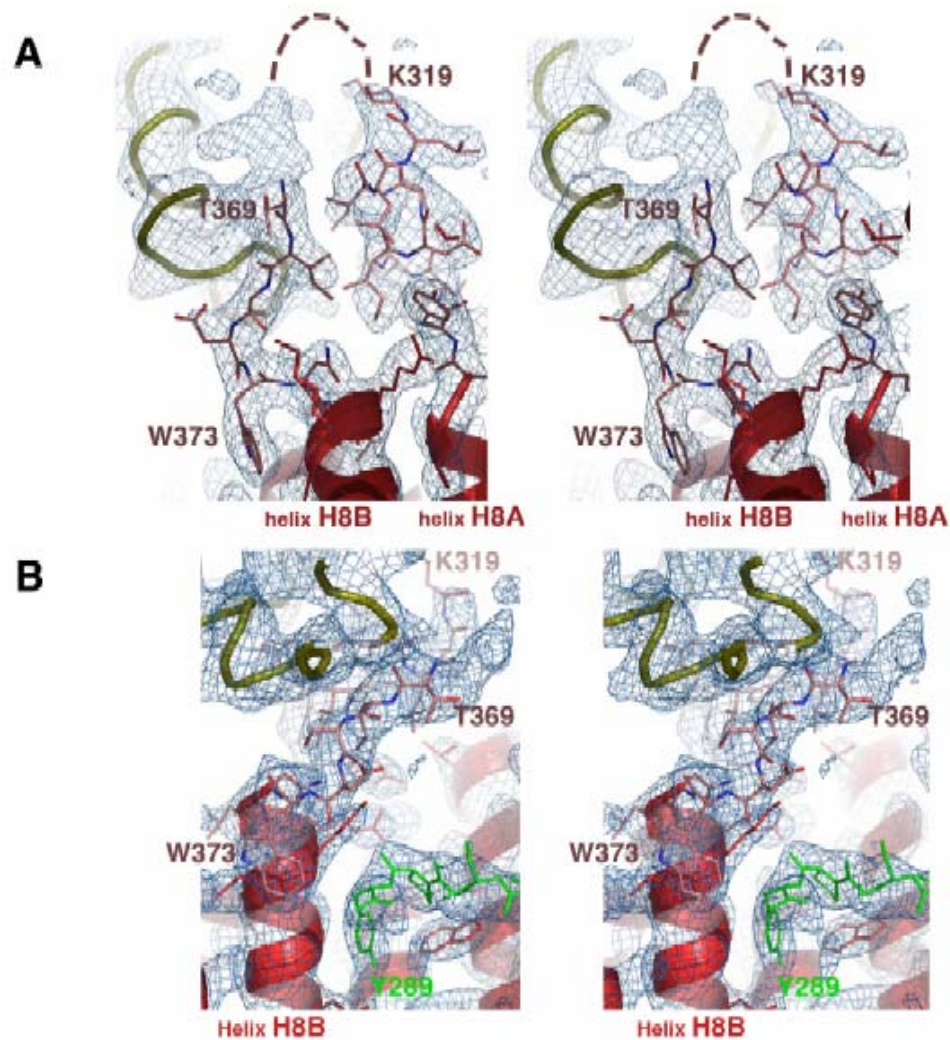


Figure 2-7) Loop density is discontinuous in Kap β 2.M9NLS complex structure. **A)** Stereo diagram of the 2Fo-Fc map (1.0σ , blue mesh) drawn at Kap β 2 (red) HEAT repeat 8, showing H8 loop residues 312-319 connecting to H8A and residues 369–374 to H8B. A neighboring Kap β 2 in the crystal is shown as a yellow ribbon. Red dashes represent the disordered connection between loop residues 319 and 369. **B)** Same as in (A), rotated $\sim 90^\circ$ about the vertical axis (M9NLS is depicted in green).

Why does the H8 loop only bind the C-terminal arch in the presence of Ran? The calculated electrostatic surface potential of the H8 loop in the presence and absence of RanGTP is distinct (Figure 2-6C). The H8 loop contains many acidic residues, particularly through ³⁵¹EDGIEEEDDDDEIDDD³⁶⁸ directly C-terminal to residues 338-350 which overlap with M9NLS. Negative charges here may prevent binding of the loop to the acidic C-terminal arch (Figure 2-6C, top). When Ran binds Kap β 2, its basic patch (K127, R129, K132, K134, R140, K141 and K159) interacts with H8 loop residues 332-340 and 363-371. Again, long-range electrostatic effects of the basic interface of Ran may substantially decrease the negative charge of the loop, converting residues 338-350 into a more suitable ligand for the Kap β 2 substrate binding site (Figure 2-6C, bottom). Ran probably also imparts conformational constraints to orient the H8 loop in the substrate site. The relative importance of electrostatic versus conformational effects of Ran binding is not known. Biophysical studies of H8 loop mutants with varying charge and H8 loop peptides *in trans* will be crucial to parse the different effects of Ran on the loop.

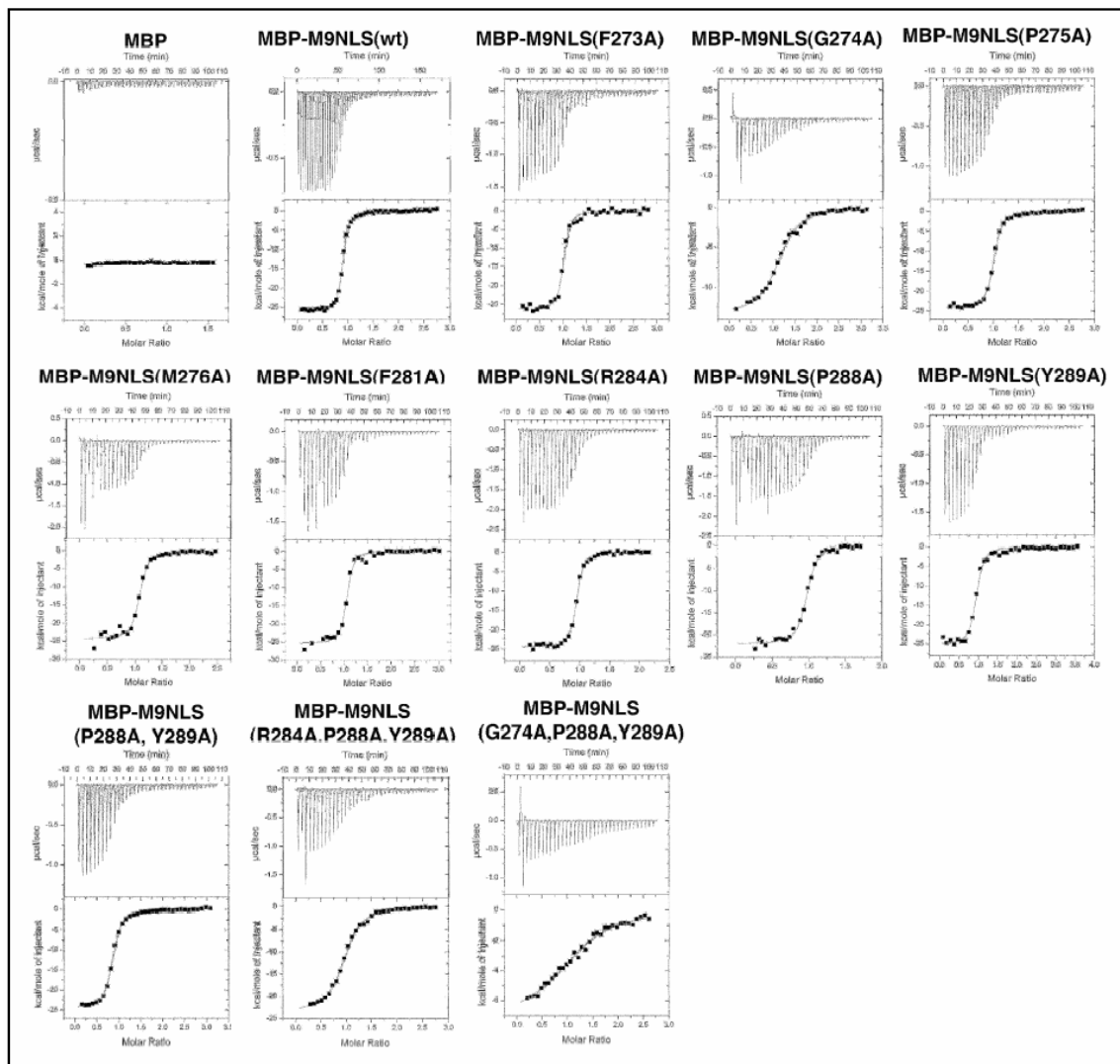


Figure 2-8) ITC profiles of MBP, MBP fusions of wild type M9NLS and various alanine mutants interacting with full length Kap β 2. Nonlinear least squares fit with the single binding site model was used to fit the ITC profiles (closed squares).

Another structural difference between the Kap β 2-M9NLS and Kap β 2-Ran complexes is found at the N-terminal arches (Figure 2-6A). Small changes in the orientation of α -helices within and between HEAT repeats 1-10 result in a maximum displacement of over 23 Å at the N-terminus. The M9NLS complex in the crystal cannot accommodate RanGTP but biochemical studies had shown that Kap β 2 can adopt a Ran-competent conformation when bound to substrate in solution (Chook, Jung et al. 2002). The two Kap β 2-M9NLS complexes in the asymmetric unit also diverge structurally with high B-factors at the N-terminal four repeats, suggesting inherently flexibility in that region. Many Kap β s have been shown to exhibit structural plasticity and adopt multiple conformations (Fukuhara, Fernandez et al. 2004). The Kap β 2-M9NLS crystals have trapped a conformation of the N-terminal arch that is incompetent for Ran-binding.

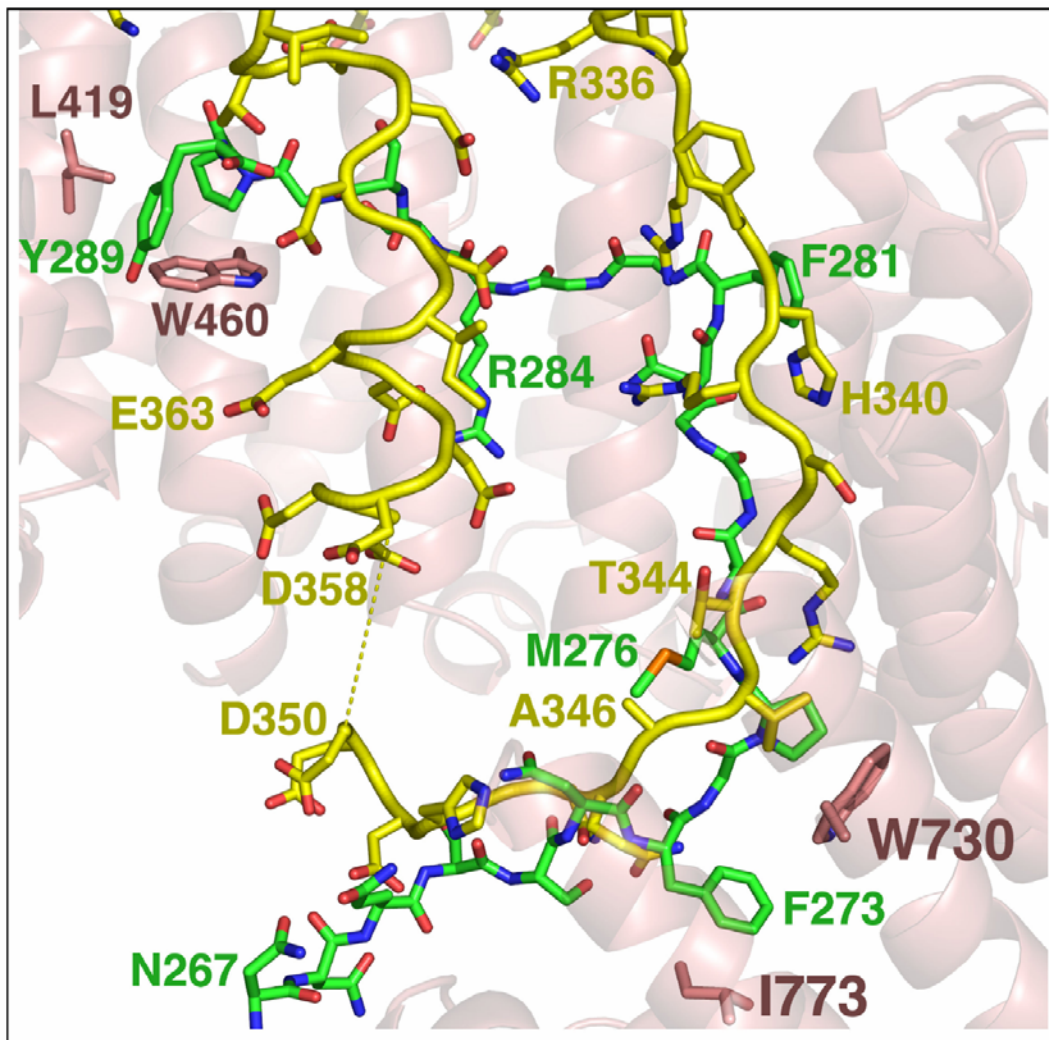


Figure 2-9) Superposition of the Kap β 2-M9NLS and Kap β 2-Ran complexes, showing the spatial overlap between the Kap β 2 H8 loop in the Ran state (yellow) and M9NLS (green)

Many other Kap β s contain large insertions like the Kap β 2 H8 loop. Kap β 1 has a short 15-residue acidic loop in repeat 8 (Cingolani, Petosa et al. 1999; Lee,

Matsuura et al. 2005), Cse1 has a 2-helix insertion in repeat 8 (Matsuura and Stewart 2004; Cook, Fernandez et al. 2005) and Crm1, Kap β 3, Imp4, Imp7, Imp8, Imp9 and Imp11 are all predicted to have large insertions in their central repeats. Mutational studies of the predicted Crm1 insertion suggest that it also directly couples Ran and substrate binding (Petosa, Schoehn et al. 2004). However, in Kap β 1 and Cse1, the mechanisms of substrate dissociation appear distinct from those in Kap β 2 and Crm1. Kap β 1 binds three different substrates in three different binding sites, and RanGTP causes a drastic change in superhelical shape that distorts binding sites of substrates Kap α and SREBP-2 while directly displacing substrate PTHrP from the N-terminal arch (Cingolani, Petosa et al. 1999; Cingolani, Bednenko et al. 2002; Lee, Sekimoto et al. 2003; Lee, Matsuura et al. 2005). Similarly, the Cse1 insertion is a pivot point for global conformational change like that in Kap β 1 (Cook, Fernandez et al. 2005). Trends for coupling Ran and substrate binding in the Kap β family are emerging. Kap β 2 and probably Crm1 employ a large insertion to directly couple the two ligands with little conformational change in the substrate binding site. In contrast, Kap β 1 and Cse1 use large-scale conformational changes to transition from closed substrate-free to open substrate-bound conformations.

CHAPTER THREE
Structural and Biochemical studies on Kap β 2.basic-PY NLS Complex
STRUCTURE BASED DESIGN OF A PATHWAY SELECTIVE NUCLEAR
IMPORT INHIBITOR

Abstract

Kapbeta2 (also called transportin) recognizes PY nuclear localization signal (NLS), a new class of NLS with a R/H/KX₂₋₅-PY motif. Here we show that Kapbeta2 complexes containing hydrophobic and basic PY-NLSs, as classified by the composition of an additional N-terminal motif, converge in structure only at consensus motifs, which explains ligand diversity. On the basis of these data and complementary biochemical analyses, we designed a Kapbeta2-specific nuclear import inhibitor, M9M.

Introduction

Transport of proteins into the human cell nucleus is mediated by 10 different import factors, all members of the Karyopherin β family of proteins (Mosammaparast and Pemberton 2004). These proteins are believed to recognize distinct nuclear localization signals (NLSs). Numerous substrates are known for Kap β 1/Importin β and Kap β 2/Transportin (Mosammaparast and Pemberton 2004; Lee, Cansizoglu et al. 2006), but few are known for other

import-Kap β s. It is also known that there is some redundancy in nuclear import of certain proteins.

In the case of nuclear export, the availability of the Crm1 inhibitor Leptomycin B has been critical for identifying many Crm1 substrates(Hamamoto, Gunji et al. 1983; Yashiroda and Yoshida 2003). Such inhibitors of specific import pathways do not exist, but could be invaluable for proteomic analyses to reveal the extensive traffic map into the nucleus.

Only two classes of NLSs are currently known: The short basic classical-NLSs that bind the Kap α /Kap β 1 heterodimer(Dingwall and Laskey 1991; Mosammaparast and Pemberton 2004) , and the recently identified PY-NLSs that bind Kap β 2(Lee, Cansizoglu et al. 2006). PY-NLSs are defined by 20-30 residue sequences that are structurally disordered, positively charged overall and contain C-terminal R/K/H-PY and N-terminal either hydrophobic or basic motifs. These orthogonal rules have provided substantial limits in sequence space to identify over 100 PY-NLS-containing human proteins(Lee, Cansizoglu et al. 2006). Two subclasses of PY-NLSs, hPY and bPY, are defined by N-terminal motifs: hPY-NLSs contain ϕ G/A/S $\phi\phi$ motifs (ϕ , hydrophobic residue) whereas the equivalent region in bPY-NLSs is enriched with basic residues. From our previous structure of human Kap β 2 bound to the hPY-NLS of heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1)(Lee, Cansizoglu et al. 2006), it was unclear if the loose consensus sequences are structurally conserved between the NLS

subfamilies or how Kap β 2 recognizes basic versus hydrophobic N-terminal motifs.

In this chapter, the structure of Kap β 2 in complex with the basic-PY NLS cargo (human hnRNPM NLS (Datar, Dreyfuss et al. 1993; Gattoni, Mahe et al. 1996; Hase, Yalamanchili et al. 2006; Lee, Cansizoglu et al. 2006)) is presented. Biochemical studies are performed to map the energetics of the interaction. Comparative studies in correlation with the previously solved hydrophobic-PY NLS complex structure, more insight was put into the recognition mechanism of the PY NLSs by Kap β 2. The biochemical studies enabled the design of a pathway specific inhibitor of Kap β 2. The inhibitory peptide binds to Kap β 2 with a substantially high affinity which cannot be released by natural mechanism which involves dissociation by Ran.GTP. The inhibitory peptide works both *in vitro* and *in vivo*.

Materials and Methods

Expression and Purification of human Kap β 2

Human Kap β 2 was expressed in pGEX-Tev vector (pGEX-4T3 with a Tev cleavage site) as an N terminal GST fusion protein. Kap β 2dloop that as the

particular truncation construct used in these experiments is a specific loop deletion mutant of Kap β 2. There's an unusually long sequence insertion between the HEAT repeats 8A and 8B (A being on the inner surface helix of Kap β 2 – See figure 3-1)(Chook and Blobel 1999). It is composed of ~52 residues (318-370) which are significantly acidic in composition (318-LKGDVEEDETIPDSEQDIRPRFHRSTVAQQHDEDGIEEEDDDDDDEIDDDDTI-370). It is thought that this highly acidic loop takes part in dissociation of the substrate in presence of Ran.GTP. For crystallization purposes, this loop sequence spanning residues 324-366 is replaced by a shorter linker fragment (GGSGG). It is thought that the presence of the loop sequence causes instability of the crystals producing floppy projections thus reducing the quality of the crystals. This truncation does not affect import cargo binding to Kap β 2. Kap β 2 was purified following the same protocol from chapter 2 (See figure 3-2, 3-3 and 3-4).

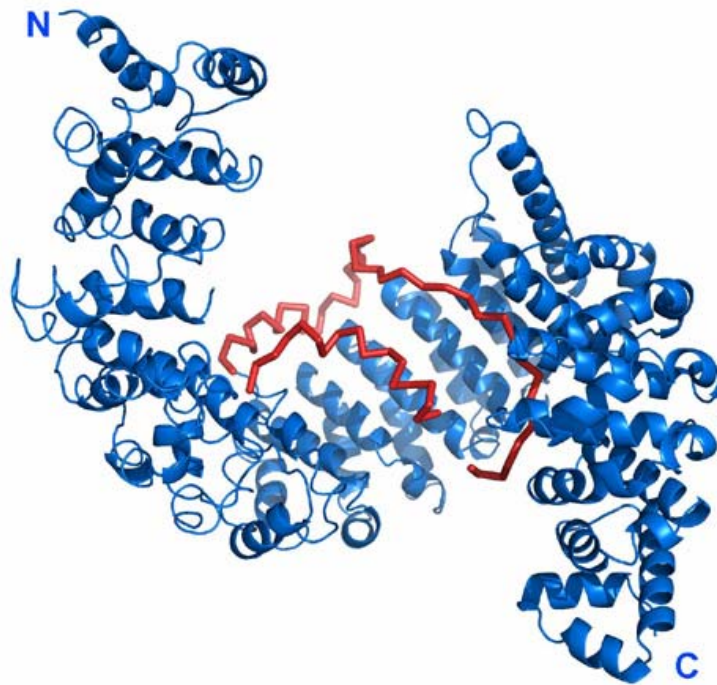


Figure 3-1) Kap β 2 loop insertion (shown in red; residues 310-375) between HEAT repeats 8A and 8B (PDB ID 1QBK).

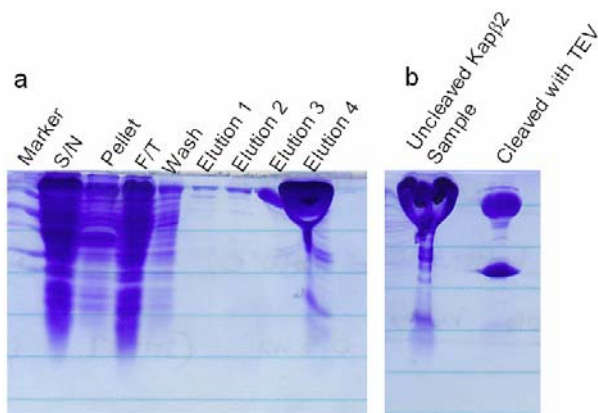


Figure 3-2) Kap β 2dloop purification 1: **a)** The protein lysate is put over beads and washed extensively using lysis buffer and ATP buffer. The protein is eluted from the beads using 20mM glutathione in lysis buffer. **b)** Kap β 2dloop purification 2. The elutions are pooled together and cleaved off GST tag.

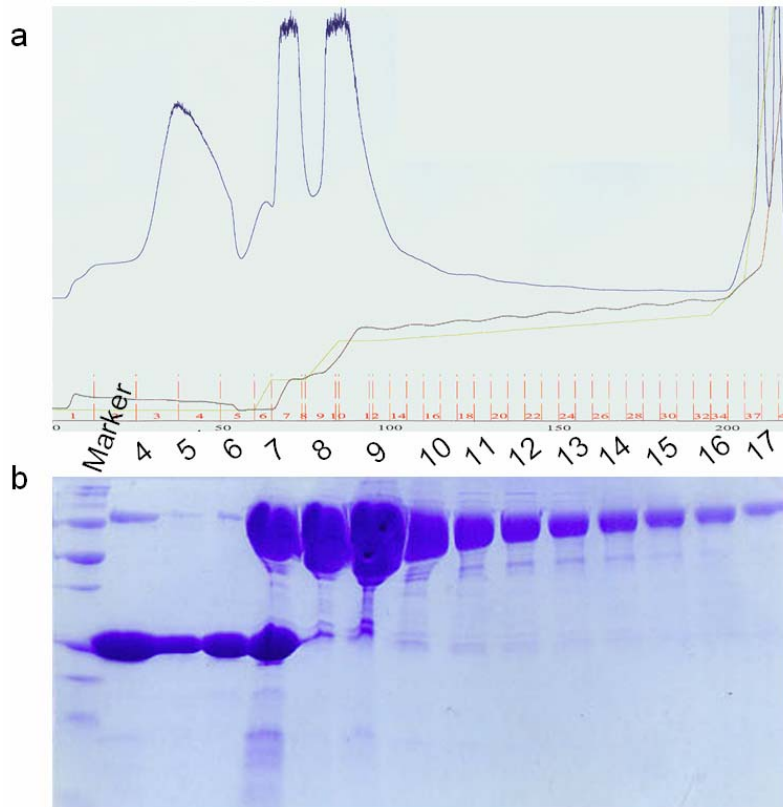


Figure 3-3) Kap β 2dloop purification 2: **a)** The sample is put through anion exchange chromatography (HiTrapQ column). **b)** The fractions (4-17) are run on 10% SDS PAGE.

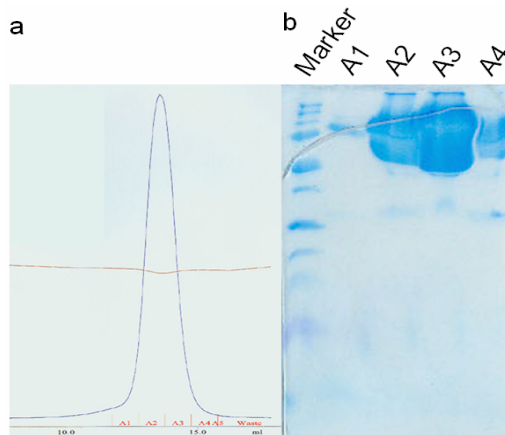


Figure 3-4) Kap β 2dloop purification 3: **a)** The fractions from Anion exchange chromatography step are pooled together and put through gel filtration chromatography (Superdex s200). **b)** Superdex s200 fractions (A1-4) are run on 10% SDS PAGE.

Expression of human hnRNPM NLS Fragment

The NLS fragment for human hnRNPM corresponding to residues 41-70 was cloned into pGEX-Tev vector (pGEX 4T-3 modified with a Tev site N terminal to cloning site) by Brittany Lee in Chook Lab. The protein was over expressed in BL21 (DE3) *E. coli* cells as an N terminal GST-fusion protein spanning residues 41-70.

The overnight grown starter culture was diluted 10ml to 1 liter LB media supplemented with 10mg/ml of ampicillin. The cells were grown at 37°C until OD_{600nm}~0.8. The cells were induced with 0.5mM final concentration of IPTG.

The induction is continued at 37°C for 3 hours. The cells were harvested by spinning down at 4000rpm for 10 minutes. The cells are resuspended in lysis buffer (10mls for each liter of cells grown) containing 50mM HEPES pH7.3, 100mM NaCl, 2mM DTT 1mM EDTA and 20% Glycerol supplemented with protease inhibitors (Pefabloc, Benzamidine, Leupeptine and Aprotinin). The cells were passed through the cell disrupter twice. The lysate was put over Glutathione Sepharose (GS) beads pre-equilibrated with lysis buffer. The GS beads were extensively washed with lysis buffer and eluted with elution buffer (Lysis buffer supplemented with 10mM glutathione and pH adjusted to 8.1).

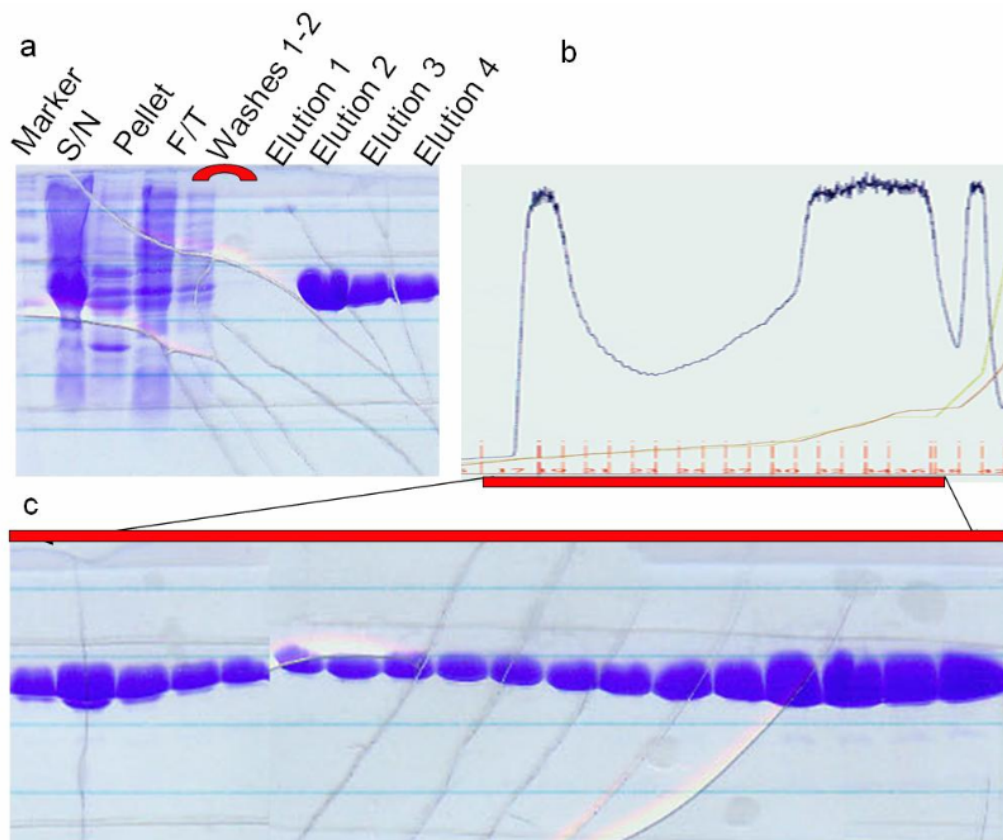


Figure 3-5) Purification of GST-tagged hnRNP-M protein: **a)** GST tagged hnRNPM protein lysate is put over GS beads and washed extensively using lysis buffer. The protein is eluted using 20mM Glutathione in lysis buffer. **b)** The elutions are pooled together and put over anion exchange (HiTrap Q) column. **c)** The fractions containing the protein (17-38) were run on a 12% SDS-PAGE gel and appropriate fractions containing the protein were pooled.

For anion exchange chromatography run, the elutions are pooled together and concentrated using Millipore spin filters. The sample is diluted with low salt buffer containing 20mM Tris pH 7.0 20mM NaCl, 2mM DTT, 1mM EDTA and

20% Glycerol. The sample is put over 5ml HiTrap-Q column (Anion exchange chromatography column). High salt buffer containing 1M NaCl, 2mM DTT, 1mM EDTA and 20% Glycerol was used as a gradient against low salt buffer for the ion exchange chromatography run.

The fractions were run on a 12% SDS-PAGE gel and appropriate fractions containing the protein were pooled. Final sample was concentrated to beyond 40mg/ml using Millipore spin filters.

Kap β 2:hnRNP M-NLS Complex Formation

GST-hnRNPM-NLS was mixed with Kap β 2 in a 3:1 mass ratio (corresponds to >6 fold molar ratio). The complex was mixed with Tev protease (70 μ l added for 1 ml of mixture) to cleave the substrate from the GST. The reaction was allowed to continue at room temperature for 4 hours. The cleavage was checked on SDS-PAGE gel. The resulting complex is put through gel filtration chromatography (Superdex s200). The peaks corresponding to the complex were pooled. To avoid GST contamination, the final sample was put over 2 consecutive 1ml GS beads equilibrated with TB buffer. The complex was concentrated to 20 mg/ml for crystallization.

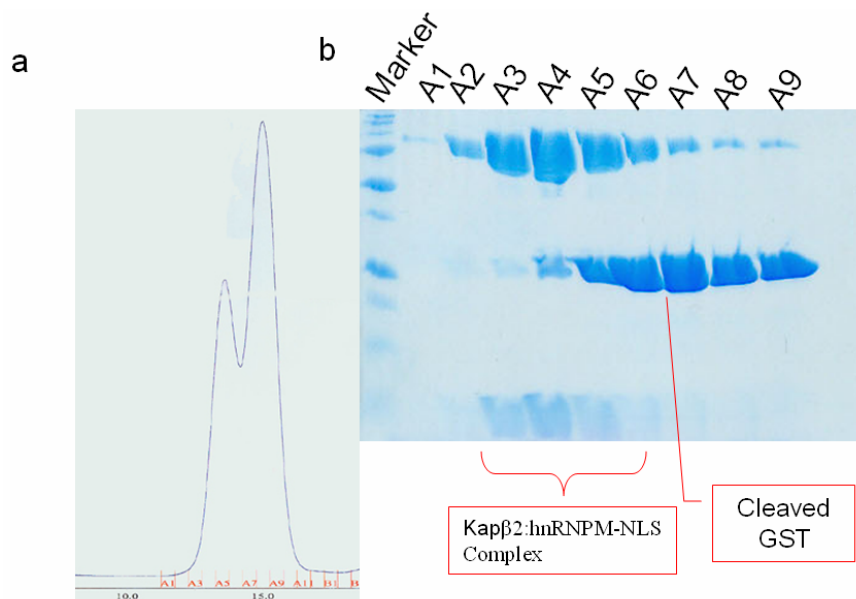


Figure 3-6) Kapβ2: hnRNPM-NLS Complex purification: **a)** The complex is cleaved off the GST tag of the NLS cargo fragment using TEV protease. Cleaved complex is put through gel filtration chromatography (Superdex s200). **b)** The fractions are run on 15%SDS PAGE.

Crystallization and Data Collection

Kapβ2-hnRNP M-NLS complex was crystallized in the condition containing 2.7 M potassium formate 100 mM HEPES pH 7.0, and 10% glycerol in the reservoir solution. The crystals were obtained using vapor diffusion. Crystals were grown

to beyond 100X100 μ m in size to avoid rapid crystal decay. The crystals were already cryo-protected in mother liquor, thus were flash frozen in liquid propane.

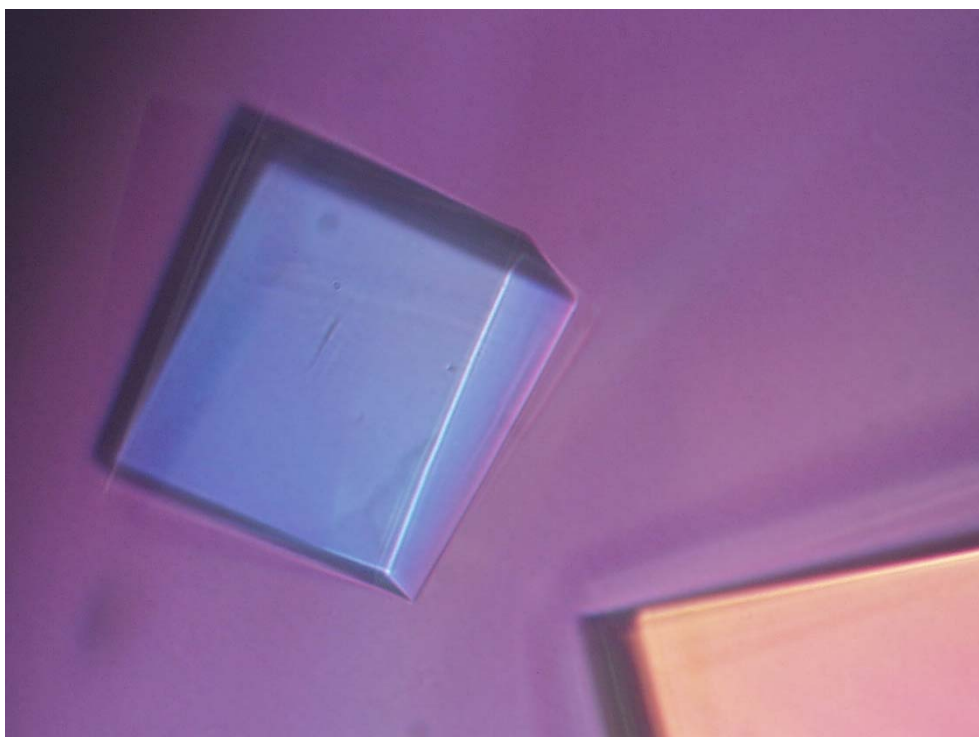


Figure 3-7) Kap β 2-hnRNP M-NLS complex crystals: Kap β 2-hnRNP M-NLS complex was crystallized in the condition containing 2.7 M potassium formate 100 mM HEPES pH 7.0, and 10% glycerol.

Crystals were initially screened at home source. Decent diffracting crystals were taken to the synchrotron for data collection. Data from these crystals were collected at the Advanced Photon Source at beamline 19-ID, Argonne National

Laboratory at X-ray wavelength 12.66 keV and temperature 100 K. Severe crystal decay was a serious concern for these crystals characteristic to Kap β 2 crystals.

Phasing and Structure Determination

Data was processed using HKL2000. Kap β 2-hnRNP M-NLS crystals were in a very similar space group as the Kap β 2-hnRNP A1-NLS crystals (PDB ID: 2h4m), with space group C2, unit cell parameters $a=152.0$ Å, $b=154.1$ Å, $c=141.7$ Å and $\beta=91.7^\circ$ (Table 3-1). The asymmetric unit contains two complexes cargo complexes.

Table 3-1) Data collection statistics for Kap β 2-hnRNP M-NLS complex

Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	153.2, 155.0, 141.5
α , β , γ ($^\circ$)	90.0, 92.6, 90.0
Resolution (Å)	50-3.0 (3.1-3.0)*
R_{sym} or R_{merge}	0.068 (0.65)
$I / \sigma I$	20 (1.5)
Completeness (%)	98.8 (92.3)
Redundancy	3.6 (3.1)

The Kap β 2-hnRNP A1-NLS model was used as a search model for molecular replacement using the program Phaser. Initial positional refinement was done using Refmac5. Density modification by solvent flipping using CNS produced electron density map that allowed most of the model to be built. Model building was done using Coot. The map was further improved using rigid body, positional and simulated annealing refinement of Kap β 2 alone, using CNS. The same test data set (5%) was used throughout the entire refinement process. Although the relative B factors are high, the Fo-Fc map plotted at 2.5σ shows interpretable density for hnRNP M-NLS residues 51-53 and 55-68 in chain C, and residues 50-69 in chain D. The final refined model shows good stereochemistry with R_{factor} of 25.4% and R_{free} of 29.1%. Ramachandran plot for final model: 90.7% in most favored and 9.3% in allowed regions (Table 3-2).

Table 3-2) Refinement statistics for Kap β 2-hnRNP M-NLS complex

Refinement	
Resolution (Å)	50-3.1
No. reflections	56,210
$R_{\text{work}} / R_{\text{free}}$	0.263/0.294
No. atoms	
Protein	12,799
Water	none
<i>B</i> -factors	
Protein	Kap β 2 <i>B</i> -factors Chain A: 88.8 Å ² Chain B: 98.5 Å ²
	hnRNP M NLS Chain C: 98.5 Å ² (51-58:121.5 Å ² , 59-64:79.0 Å ² , 65-68:97.6 Å ²)
	hnRNP M NLS Chain D: 112.1 Å ² (49-58:142.6 Å ² , 59-64:74.4 Å ² , 65-69:112.6 Å ²)
Water	none
R.m.s deviations	
Bond lengths (Å)	1.179
Bond angles (°)	0.008

Expression and purification of MBP tagged NLS fragments

Various NLS fragments were cloned into pMAL-Tev vector (modified from pMAL vector by insertion of a Tev site following the MBP and preceding the cloning site). The fragments were expressed as N terminal MBP fusion proteins in BL21 (DE3) cells. The overnight starter culture was diluted 30ml into 3L and grown until $OD_{600nm} \sim 0.8$ the cells were harvested by spinning down at 4000rpm for 10minutes. The cells were resuspended in lysis buffer containing 50mM HEPES pH7.3, 100mM NaCl, 2mM DTT 1mM EDTA and 20% Glycerol with protease inhibitors. The cells were lysed by passing twice through the cell disrupter. The lysate was spun down at 15000rpm for 50minutes. The supernatant was loaded onto 15ml amylose beads (New England Biolabs) pre equilibrated with lysis buffer.

The beads were washed extensively (8X15mls) with the lysis buffer. The protein was eluted with lysis buffer with 10mM Maltose. Last 3 washes and three elutions were pooled together and concentrated using Millipore spin filter. The sample was diluted 2.5 times to reduce the salt concentration and loaded onto Anion exchange (Q) column.

The fractions are run on 12%SDS PAGE and the fractions containing the MBP protein peak were pooled and concentrated using the spin filter beyond $\sim 30mg/ml$ and frozen at $-80^{\circ}C$.

Alanine mutants of NLS fragments

Alanine mutants of the NLS fragments were produced following site directed mutagenesis protocol (Stratagene). Complementary primers containing appropriate mutations were designed for PCR amplification of the template plasmid DNA. The mutant proteins were expressed and purified following the same purification protocol for the wild type MBP tagged protein.

Quantitation of binding affinity with Isothermal Titration Calorimetry

Binding affinities for wild type and mutant MBP-hnRNP M-NLS were determined using Isothermal titration calorimetry (ITC). The experiments were performed using a MicroCal Omega VP-ITC calorimeter (MicroCal Inc., Northampton, MA). MBP-NLS proteins were dialyzed against 2X2L buffer containing 20 mM Tris pH 7.5, 100 mM NaCl and 2 mM β -mercaptoethanol overnight. 100-300 μ M Wild type and mutant MBP-hnRNP A1-NLS proteins were titrated into the sample cell containing 10-100 μ M full-length Kap β 2. All ITC experiments were done at 20°C with 35 rounds of 8 μ l injections. Data were

plotted and analyzed using the single binding site model of MicroCal Origin software version 7.0. The ITC runs are shown in figure 3-13.

Competition ITC Experiments

The detection of K_D values of a binding experiment for the standard ITC experiments, provided that there's sufficient amount of heat release that can be detected upon binding is in low nanomolar to millimolar range. Good substrates for Kap β 2 bind at the high affinity edge of the binding detection limit (low nanomolar, K_D of 42 nM and 10 nM respectively, by standard ITC). The inhibitory M9M peptide appears to bind Kap β 2 with higher affinity than the natural NLSs, because the measurements by standard ITC yielded inconsistent results at the low nanomolar range of the detection limit (~ 10 nM). To be able to detect possible high K_D value of the inhibitor peptide, we performed competition ITC to extend the range of sub-nanomolar range ($K_D < 10^{-9}$ M) affinities. hnRNP A1-NLS R284A/P288A/Y289A mutant (K_D of 461 nM, measured by standard ITC) was used as the competition displacement ligand. This mutant was used because the displacement of a tighter complex would require a lot higher protein concentration and at this tight binding range, the displacement may not work. The calorimetry cell containing 12 μ M Kap β 2 and 18 μ M R284A/P288A/Y289A

mutant of MBP-hnRNP A1-NLS was titrated with syringe protein sample solution of 108 μM MBP-M9M inhibitor (or 154 μM wildtype hnRNP A1-NLS as control). The experiment was repeated using 20 μM of the competition displacement ligand. Data were analyzed with the competition model in MicroCal Origin software version 7.0. Competition ITC experiments yielded K_D values of 107 pM and 111 pM for inhibitor M9M peptide and K_D of 20 nM for wildtype hnRNP A1-NLS.

Qualitative Binding assays

RanGTP-mediated dissociation experiments: approximately 20-40 μg of GST-hnRNP A1-NLS, GST-hnRNP M-NLS and GST-M9M were immobilized on glutathione sepharose (Amersham, NJ, USA). 20 μg of Kap β 2 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 increasing concentrations of RanGTP (0.32 μM , 0.64 μM , 0.96 μM , 1.28 μM , 1.6 μM), each in 100 μL solution. After extensive washing, a quarter of the bound proteins were analyzed by SDS-PAGE and visualized with Coomassie staining.

Competition NLS dissociation experiments: approximately 20-40 μg of GST-hnRNP A1-NLS was immobilized on glutathione sepharose (Amersham, NJ, USA) and incubated with 10 μg of Kap β 2 and 7 μg of either MBP-hnRNP M-NLS, MBP-hnRNP A1NLS or MBP-M9M. Samples were washed extensively and a quarter of each reaction was subjected to SDS-PAGE and Coomassie staining.

Kap β 1 binding experiments: approximately 1 μg GST-Kap β 1 immobilized on glutathione sepharose (Amersham, NJ, USA) and incubated with Kap α (5 μg), Kap α (5 μg) and IBB-His₆ (~50 μg), Kap α (5 μg) and MBP-M9M (7 μg) or MBP-M9M (7 μg). Samples were washed extensively and a quarter of each reaction was subjected to SDS-PAGE and Coomassie staining.

Kap β 2 mutants binding experiments: approximately 30 μg of GST-hnRNP A1-NLS, GST-hnRNP M-NLS and GST-M9M were immobilized on glutathione sepharose (Amersham, NJ, USA), followed by addition of 20 μg of Kap β 2 or Kap β 2 W460A/W730A mutant(Lee, Cansizoglu et al. 2006). Samples were washed extensively and a quarter of each reaction was subjected to SDS-PAGE and Coomassie staining.

Subcellular localization of proteins in HeLa cells

In vivo and western blot analyses were performed by Zi Chao Zhang in Chook Lab guided by Dr. Beatriz Fontoura in Cell Biology Department. MBP, MBP-hnRNP A1-NLS and MBP-M9M were subcloned into the modified pCS2-MT mammalian vector at Sal I and Not I sites. HeLa cells were maintained in DMEM (GIBCO BRL, Gaithersburg, MD) with 10% fetal bovine serum (Gemini Bio-Products, West Sacramento, CA). Cells were grown on 12 mm coverslips placed in 24-well cell culture and transfected using Effectene (Qiagen, Valencia, CA) 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, Montgomery, TX) diluted to 5 ug/ml was used to detect the myc-MBP-peptides.

The monoclonal antibody 4C2 (a gift from Dr. M. Matunis) at 1:1000 dilution detected endogenous hnRNP A1 when incubated with goat-anti-mouse-Cy3 (Jackson ImmunoResearch Laboratories, West Grove, PA) antibody at 1:400 dilution. 4C2 has been previously shown to recognize human hnRNP A1, A2, B1

and B2(Matunis, Matunis et al. 1992). We show by western blot (Figure 3-19) that 4C2 recognizes the hnRNP A1 fragment 257-305 but not the chimeric inhibitory peptide M9M (Figure 3-12). Monoclonal antibody 2A6 (a gift from Dr. M. Swanson) 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. HDAC1 has previously been reported to be imported into the nucleus by Kap α /Kap β 1(Smillie, Llinas et al. 2004). We have confirmed by *in vitro* binding assays that recombinant HDAC1 binds Kap α but not Kap β 2 (data not shown). To detect endogenous HDAC1, mouse anti-HDAC1 monoclonal antibody 2E10 (Upstate Biotechnology; 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). 52-157 transfected cells were analyzed for each of the experiments, and percentages with cytoplasmic substrates are shown in a histogram (Figure 3-18).

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 diluted at 1:2000 and antibody 2A6 diluted at 1:1000. Secondary horseradish peroxidase-

conjugated anti-mouse antibody (diluted 1:10000, Amersham) and the ECL system (Amersham) was used to visualize the blots.

Results and Discussions

Structure determination of the hPY-NLS in complex with Kap β 2 enabled the discovery of certain rules generalized for the classification of Kap β 2 cargo substrates. A detailed bioinformatics analysis in conjunction with biochemical experiments, revealed the presence of the second type of PY-NLS fragments carrying a basic N-terminal motif rather than the hydrophobic motif present in hnRNP-A1-NLS. The most significant difference of these two subtypes was the composition the N-terminal motif of the NLS peptide recognition consensus sequence. The N-terminal motif in the second class of NLSs is composed of highly basic residues (Figure 3-9).

hnRNP A1 (M9NLS) 263 FGNYNNQSSNFGPMKGGNFGGRSSGPY289
 hnRNP D 329 YGDYSNQQSGYGKVSRRGGHQNSYKPY355
 TAP 49 SSRLEEDDGDVAMSDAQDGPRVRYNPY 75
 HuR 212 HHQAQRFRFSPMGVDHMSGLSGVNVPG238
 hnRNP F 184 FKSSQEEVRSYSDPPLKFMSVQRPGPY210
 hnRNP M 38 GEGERPAQNEKRKEKNIKRGGNRFEPY 64
 PQBP-1 161 RDRGYDKADREEGKERRHHRREELAPY187

Consensus :

$$\begin{array}{c}
 \text{R/H/KX}_{(2-5)} \text{PY} \\
 \text{and} \\
 \left(\begin{array}{c}
 \phi\text{G/A/S}\phi\phi\text{X}_{(11-13)} \text{PY} \\
 \text{or} \\
 \text{basic-enriched}_{(5-8)} \text{X}_{(8-10)} \text{PY}
 \end{array} \right)
 \end{array}$$

Figure 3-8) Previous biochemical and experimental studies revealed the consensus sequence to consist of 2 major recognition motifs(Lee, Cansizoglu et al. 2006). These are N-terminal basic or hydrophobic stretch followed by an R-PY motif.

It is important to elucidate the binding interactions with Kap β 2 since although the PY motif is conserved, it is still possible that binding may occur through a different interface. Kap β 2 has a relatively flat surface on the concave side. hnRNP-A1-NLS binds to Kap β 2 in a completely extended conformation spanning about 100Å along the inner surface of the C-terminal arch. Kap β 2 has a wide surface area giving it the ability to accommodate large possible alternative paths that the diverse cargo substrates may bind to. This could be a means to enable it to

bind to a wide variety of cargo substrates. However, it was found that the hPY and bPY NLSs follow the same path along the inner surface of the protein.

Nevertheless, this does not specify that Kap β 2 recognizes all of its cargo substrates only through this mechanism. It is known that, the most common and well defined classical NLS pathway (Kap β 1) protein can bind to completely different cargo proteins in a completely different way without Kap α in a completely independent mode. In a particular case, the NLS recognition motif appears to be a three dimensional fold rather than a simple primary consensus sequence.

The structure of the representative bPY-NLS (hnRNP-M-NLS) bound to Kap β 2 was solved to understand how the diverse hydrophobic and basic N-terminal motifs are recognized by Kap β 2 surface.

The structure of the bPY-NLS Cargo hnRNP-M and Kap β 2 Complex

Here we report the 3.1 Å crystal structure of human Kap β 2 bound to the bPY-NLS of human splicing factor heterogeneous nuclear ribonucleoprotein M (hnRNP M). The complex was crystallized in approximately the same space group with similar unit cell parameters compared to hnRNP-A1-NLS Kap β 2 cargo complex structure.

The asymmetric unit consists of two complexes of Kap β 2.NLS with minor differences. Most of the cargo NLS peptide was successfully traced (residues 50-69 in chain D and residues 51-53 and 55-68 in chain C).

The NLSs of hnRNPs M and A1 follow different paths while lining a common interface on the structurally invariant Kap β 2 C-terminal arch. The C α root mean square deviation of Kap β 2 (residues 435 through 780) is 0.9Å. A detailed examination of the structure revealed that the binding site of the NLS peptide was mostly conserved on the inner surface of Kap β 2. The previously determined sequence consensus was well conserved in terms of binding interface. The most conserved and important PY residues determined from the binding experiments were located at exactly the same pocket, Pro63 making strong hydrophobic interactions with Trp460, Leu419 of Kap β 2 and Tyr64 making decent contacts with Ala423, Ala380 and Lys377 of Kap β 2. Residues 51-64 of hnRNP M and residues 273-289 of hnRNP A1 contact a common Kap β 2 surface with highest overlap at their Pro-Tyr motifs. Upon superposition of Kap β 2, r.m.s.d. for all atoms in the NLS Pro-Tyr and for guanido-group atoms of the Arg residues in the R/H/K-PY motifs are 0.9 Å and 1.2 Å, respectively. The C α -C α distances are significantly low in the conserved sequence regions structurally as well (Figure 3-9). The second conserved residue, Arg60, in the common consensus was also well observed in bPY structure making the same contacts with Asp543, Glu509 and Ser502 on Kap β 2.

When the consensus motif at the N-terminal region of the NLSs is examined, it can be observed that the N-terminal consensus motif is structurally conserved on the Kap β 2 binding interface as well. Although the interactions involve basic residues rather than hydrophobic residues in comparison to the hPY N-Terminus, the interactions of the basic residues taking part in the interactions at the same interface also contribute significant hydrophobic contacts. Therefore, although the residues present at the interaction surface are highly basic in composition, the interactions involving the NLS on Kap β 2 consist of number of hydrophobic contacts. This is achieved via the elongated aliphatic side chains of the basic residues. For example the aliphatic chain of Lys52 makes hydrophobic contacts with the hydrophobic side chain of Trp730 of Kap β 2 (Figure 3-10).

It should be noted however that there are significant number of polar interactions with the highly acidic Kap β 2 interface. Lys52 of the NLS interacts with Asp693 and Glu653. Glu51 of the NLS interacts with Asn770 at the very N-terminus of the NLS peptide.

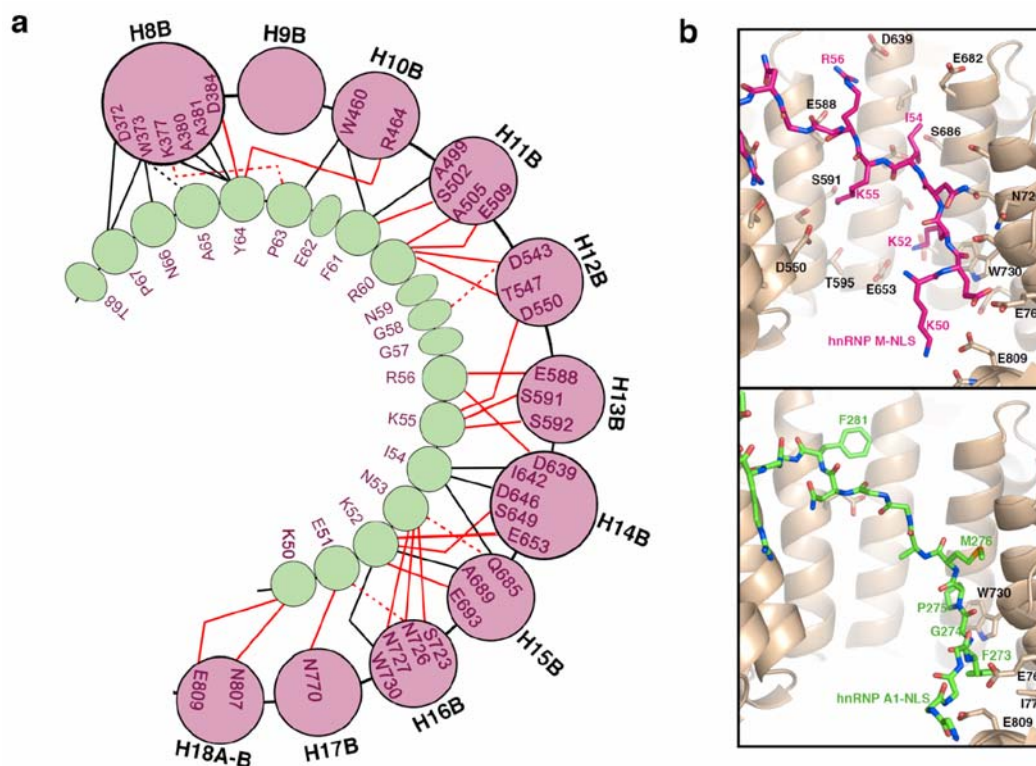


Figure 3-10) a) Kap β 2-hnRNP M NLS contacts ($< 4.0 \text{ \AA}$) hnRNP M NLS residues are shown as green circles and Kap β 2 helices as pink circles. Contacts involving the main chain and sidechains of hnRNP M NLS are shown with dashed and solid lines, respectively. Hydrophobic contacts are in black and polar contacts in red. **b)** Interactions between Kap β 2 (light brown) and the central NLS motifs of hnRNP M (magenta) and A1 (green)(Cansizoglu, Lee et al. 2007)

The N-terminus of hnRNP A1-NLS exits the Kap β 2 arch with residues 263-266 binding the convex side of Kap β 2, whereas the N-terminus of hnRNP M proceeds towards the opening of the Kap β 2 arch. The branching of the very N-terminal basic region takes an alternative path compared to the hydrophobic motif. This is approximately 90° to the hydrophobic path taken by the hnRNP-A1-NLS (Figure 3-11).

hnRNP A1 is disordered beyond its C-terminal Pro288-Tyr289 while hnRNP M extends five residues beyond its Pro-Tyr motif C-terminal to the NLS. This could be an indication of significance of the stability of the C-terminal PY motif of hnRNP-M-NLS. From ITC measurements, it is indicative that PY sequence for hnRNP-M-NLS is contributing energetically much more significant than that of the hnRNP-A1-NLS.

In addition, residues 51-54 of the basic ⁵⁰KEKNIKR⁵⁶ motif in hnRNP M and the hydrophobic motif in hnRNP A1 (residues 274-277) also overlap with main chain r.m.s. deviation of 1.3 Å. In contrast, backbone atoms of intervening ⁶¹FE⁶² in hnRNP M and ²⁸⁵SSG²⁸⁷ in hnRNP A1, and those between the N-terminal and R/H/K-PY motifs, diverge up to 4.0 Å and 7.2 Å, respectively. The structurally variable linkers show significant divergence in amino acid composition and length across the PY-NLS family. This apparent variance makes the consensus determination particularly non-trivial.

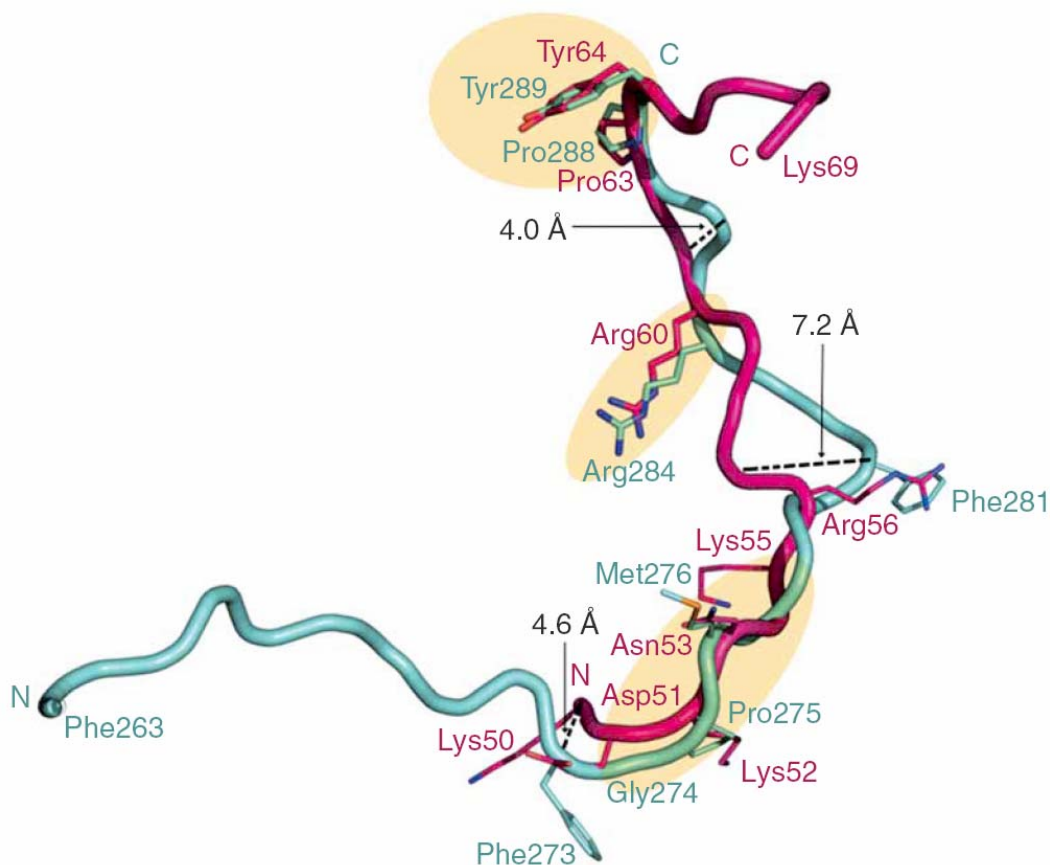


Figure 3-11) NLS fragments of hnRNPs M (2OT8; magenta) and A1 (2H4M; blue) upon superposition of Kap β 2 residues 435-780: Regions of structural similarity are highlighted in orange(Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007).

From the two structures examined together, the NLS structures show three-dimensional convergence at three sites: the N-terminal motif, the arginine and proline-tyrosine residues of the R/H/K-PY motif. These overlapping residues are the key binding epitopes of the NLSs, further structurally supporting their designation as consensus sequences (Figure 3-11).

The multivalent nature of the NLS-karyopherin interaction likely allows modulation of binding energy at each of the three sites to tune the overall affinity to a narrow range suitable for regulation by nuclear RanGTP.

Functional groups in the hnRNP M basic motif ⁵⁰KEKNIKR⁵⁶ are very different from the hnRNP A1 hydrophobic motif ²⁷³FGPM²⁷⁶, despite a common Kap β 2 interface. Most side chain interactions in the former are polar, while those in the latter are entirely hydrophobic. The corresponding Kap β 2 interface is highly acidic with scattered hydrophobic patches. In the hnRNP A1 complex, Phe273 and Pro275 in the hydrophobic motif make hydrophobic contacts with Kap β 2 Ile773 and Trp730, respectively. In the hnRNP M complex, similar hydrophobic contacts occur between the aliphatic portion of the NLS Lys52 sidechain and Kap β 2 Trp730, and between NLS Ile54, Kap β 2 Ile642 and aliphatic portions of Kap β 2 Asp646 and Gln685. Other side chains within hnRNP M ⁵⁰KEKNIKR⁵⁶ make myriad polar and charged interactions with the acidic surface of Kap β 2. Thus, the relatively flat and open NLS binding site on Kap β 2 coupled with its mixed acidic/hydrophobic surface can accommodate diverse sequences, ranging from the hydrophobic segment in hPY-NLSs to basic groups in bPY-NLSs that contribute electrostatic and hydrophobic interactions to Kap β 2.

Binding energy determination using ITC experiments

The binding energy contributions of the surface residues were determined using ITC experiments. The heat release upon titration of MBP-NLS fragment into cell containing Kap β 2 was used to determine the K_D of the interaction. Alanine mutants of the NLS fragment were used to map the binding energy distribution of the NLS binding surface interaction.

Table 3-3) Kap β 2 binding to hnRNP M NLS and mutants: Dissociation constants are measured by Isothermal titration Calorimetry (ITC)(Cansizoglu, Lee et al. 2007) (Figure 3-12)

MBP-hnRNP M(41-70)	K_D
Wild type	10 ± 1.7 nM
K50A	16.4 ± 0.4 nM
K52A	14.6 ± 0.3 nM
N53A	17.1 ± 0.5 nM
I54A	8.8 ± 1.8 nM
K55A	7.6 ± 2.3 nM
R56A	13.9 ± 2 nM
K50A/E51A/K52A/N53A	22.3 ± 4.1 nM
K50A/E51A/K52A/N53A/I54A/K55A/R56A	1.2 ± 0.2 μ M
F61A	11.2 ± 1.6 nM
P63A/Y64A	4.5 ± 0.7 μ M
F61A/ P63A/Y64A	8.6 ± 1.4 μ M
R59A/ P63A/Y64A	ND
P67A	8.7 ± 1.5 nM

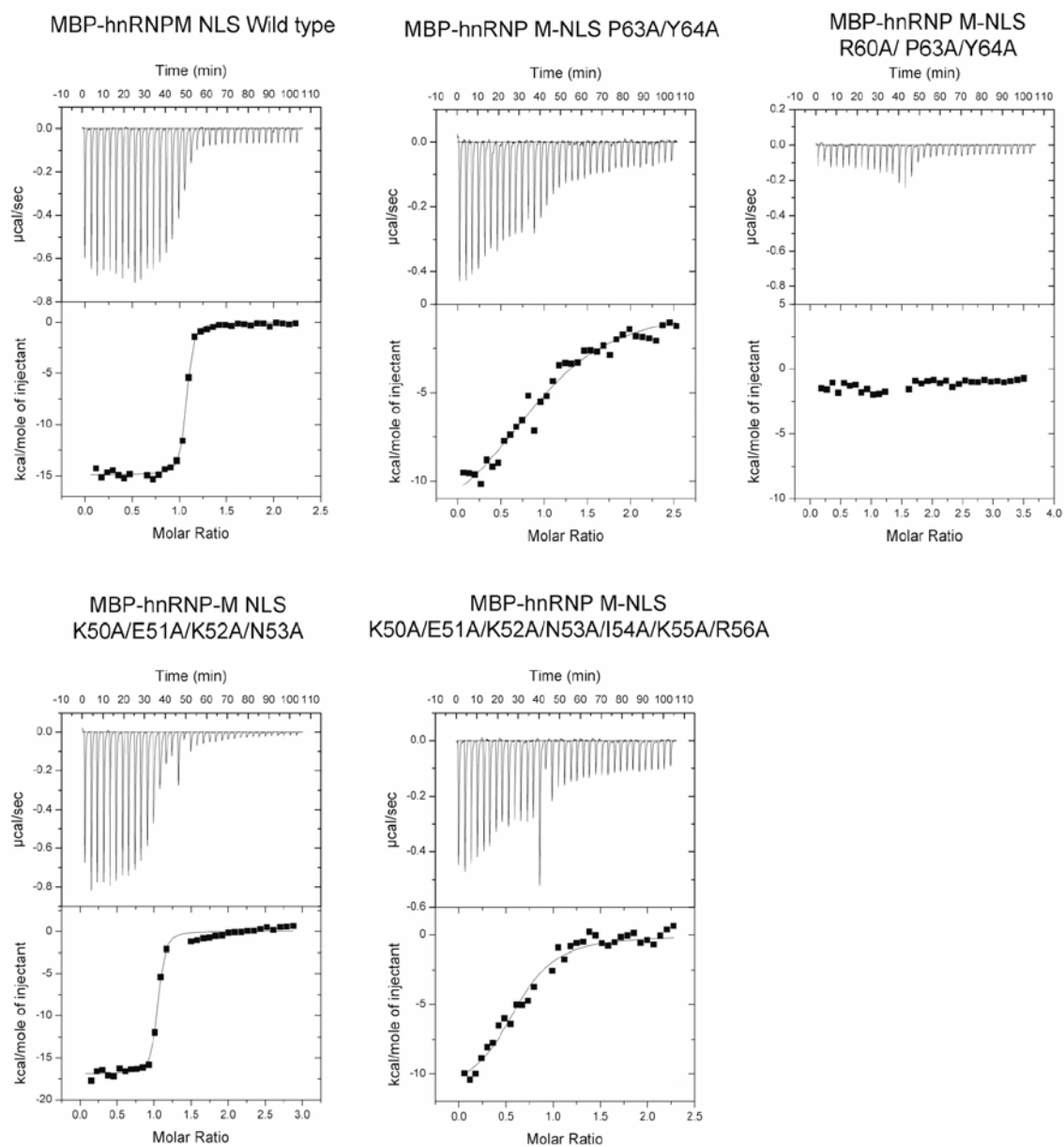


Figure 3-12) Isothermal Titration Calorimetry (ITC) measurements of select hnRNP M-NLSs binding to Kap β 2(Cansizoglu, Lee et al. 2007).

ITC experiments using Alanine mutants revealed the presence of the binding hotspots of the NLS interaction. It is interesting to note that the binding energy distribution of the two NLSs is completely different. Overall, the binding energy appears to be distributive amongst the three binding epitopes of the interaction especially for the hnRNP-A1, however, there are also binding hotspots which contribute significant energy to the interaction.

Despite structural conservation of key motifs, the distribution of binding energy along PY-NLSs is very different. It was observed previously that the PY to alanine mutants of hnRNP-A1 can still bind to Kap β 2. In hnRNP A1, Gly274 is the only binding hotspot (Nakielny, Siomi et al. 1996; Fridell, Truant et al. 1997; Bogerd, Benson et al. 1999; Lee, Cansizoglu et al. 2006) and energetic contribution from its C-terminal Pro-Tyr is modest despite strong sequence conservation (Iijima, Suzuki et al. 2006; Lee, Cansizoglu et al. 2006). Therefore, the binding energy is relatively more distributive between the N-terminal and the C-terminal binding epitopes of the NLS (Figure 3-13).

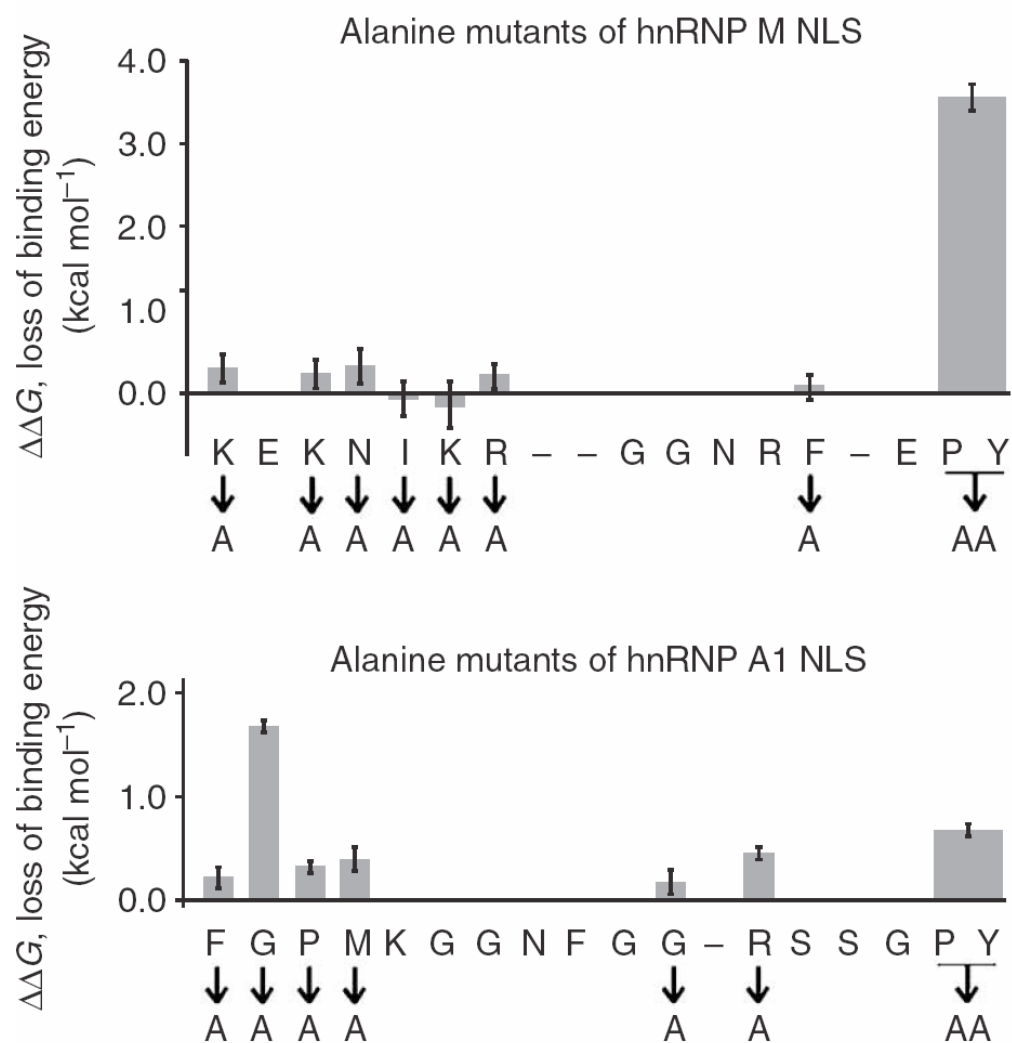


Figure 3-13) Loss of Kap β 2 binding energy in alanine mutants of hnRNPs A1(Lee, Cansizoglu et al. 2006) and M ($\Delta\Delta G = -RT\ln(K_{D,wt}/K_{D,mutant})$; K_{DS} by ITC)(Cansizoglu, Lee et al. 2007).

On the other hand, the only hnRNP M-NLS hotspot is at its Pro-Tyr motif. Neither single alanine mutants within $^{50}\text{KEKNIKR}^{56}$ nor a quadruple $^{50}\text{KEKN}^{53}/\text{AAAA}$ hnRNP M mutant had decreased affinity for Kap β 2 measured by ITC. Affinity decreased substantially only when all seven residues were mutated to alanines (1.2 μM versus 10 nM for wild type-NLS). Therefore, the significance of the highly basic patch flanking the N-terminus of the NLS could be important in general electrostatics to initially position the NLS in the interface. It is also indicative of interactions involving multi-step cooperative binding since the mutants do not seem to show an additive effect in this region.

Asymmetric locations of NLS binding hotspots in hnRNPs A1 and M, and the fact that these sites are joined by variable linker-like intervening sequences, provide opportunity to design of chimeric peptides with possible enhanced binding affinities for Kap β 2. A peptide carrying both of the hotspots from hnRNP-A1 and M could in principle potentially bind to Kap β 2 with sufficiently high affinity to compete with natural substrates and even be resistant to Ran-mediated complex dissociation inside the nucleus (Chook, Jung et al. 2002). This peptide would function as a nuclear import inhibitor. A chimeric peptide named M9M was constructed consisting of the N-terminal half of hnRNP A1-NLS (residues 257-281) fused to the C-terminal half of hnRNP M-NLS (residues 59-69) and thus carrying the binding hotspots for the both signals (Figure 3-14).

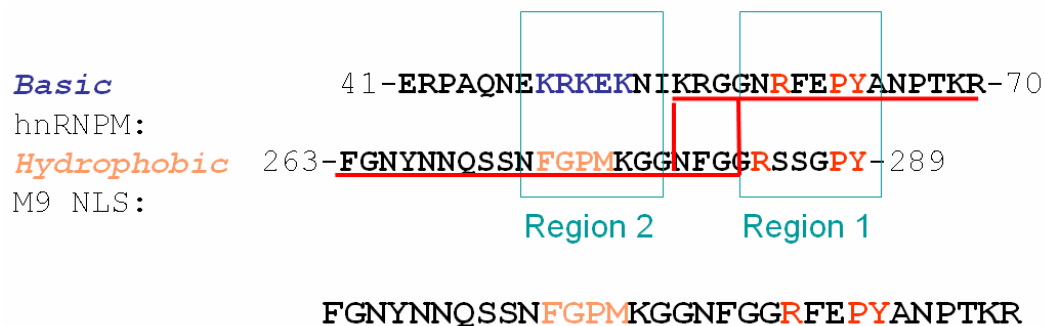


Figure 3-14) a chimeric peptide carrying both of the hotspots from hnRNP A1 and hnRNP M NLS sequences is constructed. Red lines correspond to two different chimeric peptides tested. Bottom panel is the sequence of the successful chimeric peptide.

Determined by on bead binding experiments, M9M cannot be successfully dissociated by RanGTP (Figure 3-16). It successfully competes out the wild type NLS fragments. Furthermore, this peptide binds specifically only to Kap β 2 but not Kap β 1, making it a Kap β 2 pathway specific inhibitor. The affinity of M9M is approximately 200 fold higher than the natural NLS peptides. (K_D of 107 pM for M9M versus 20 nM for hnRNP A1-NLS determined by competition ITC). This avidity effectively explains the inhibition mechanism by M9M peptide (Figure 3-15).

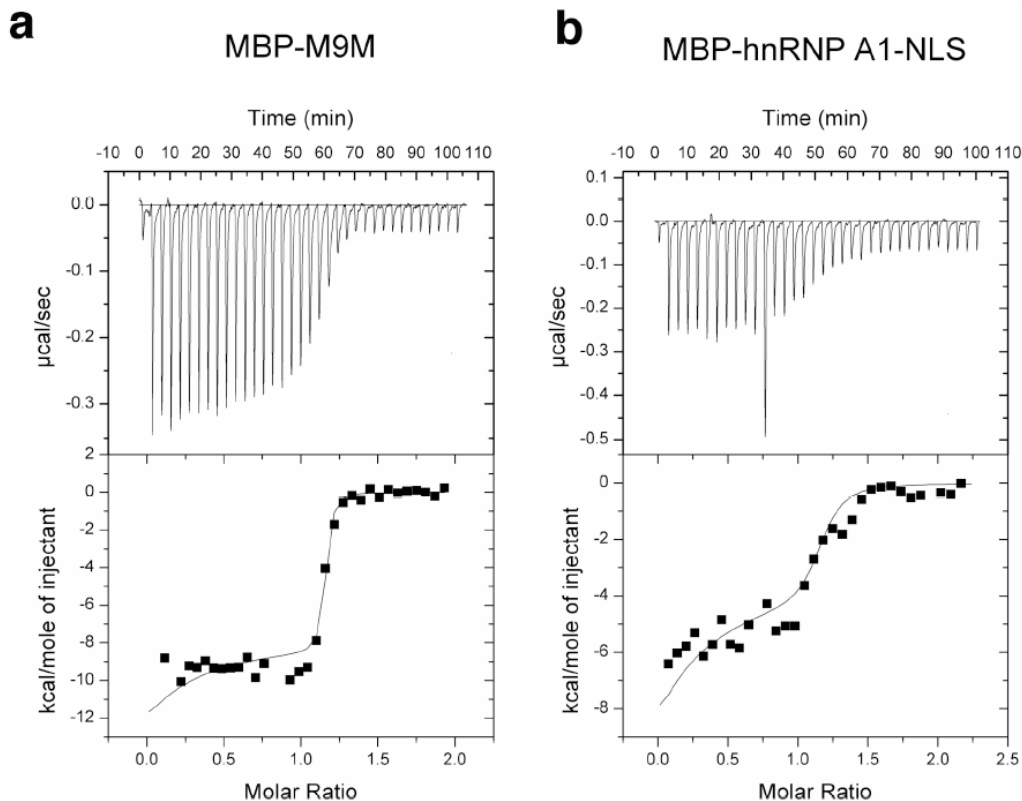


Figure 3-15) Competition ITC experiment data for inhibitor MBP-M9M binding to Kap β 2(Cansizoglu, Lee et al. 2007). **a)** The titration cell containing 12 μ M Kap β 2 and 18 μ M R284A/P288A/Y289A mutant of MBP-hnRNP A1-NLS was titrated with syringe solution containing 108 μ M MBP-M9M inhibitor. The KD obtained for Kap β 2-M9M interaction is 107 pM. **b)** Control ITC experiment was performed with 12 μ M Kap β 2 and 20 μ M R284A/P288A/Y289A mutant of MBP-hnRNP A1-NLS in the calorimetry cell, and titration with syringe solution of 154 μ M of MBP-hnRNP A1-NLS. The KD obtained for Kap β 2-hnRNP A1-NLS interaction by ITC competition is 20 nM, comparable to KD of 42 nM by direct/standard ITC.

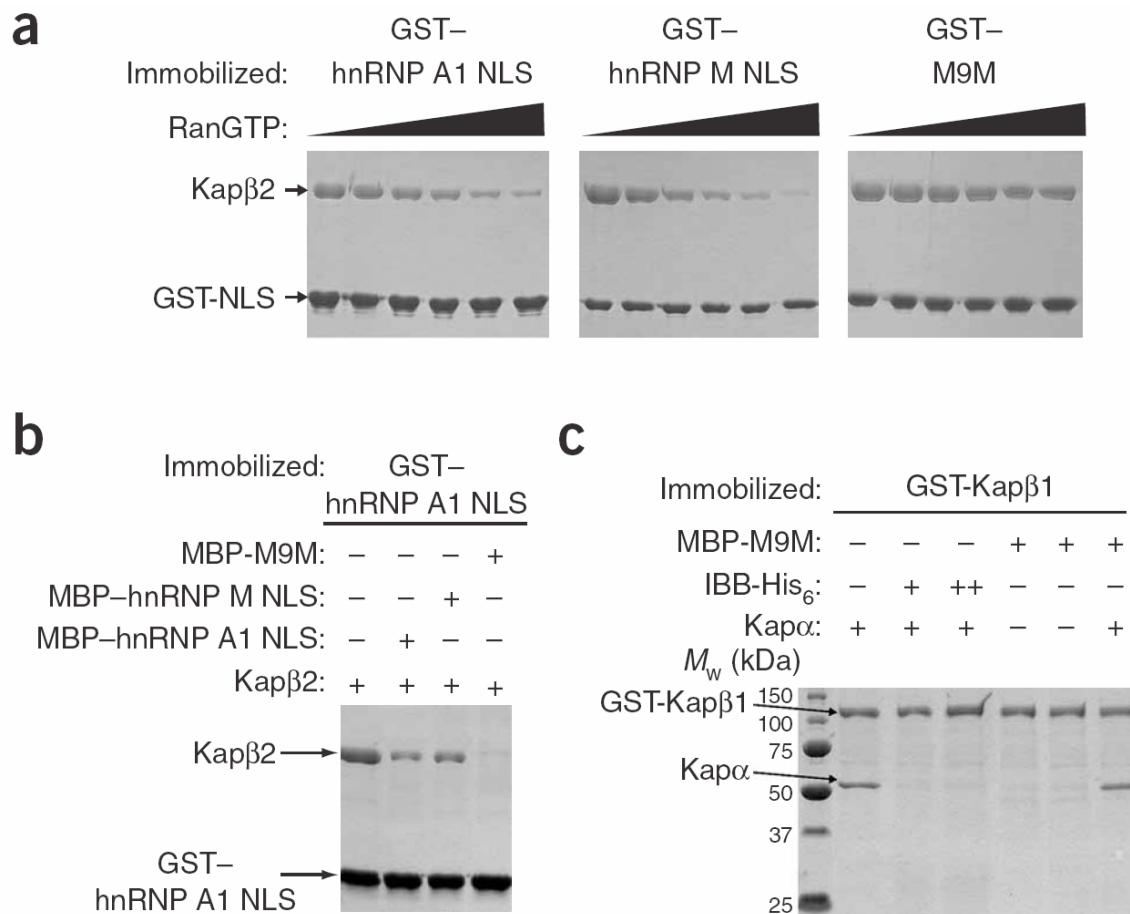


Figure 3-16 Coomassie-stained gels of on beads binding experiments (Cansizoglu, Lee et al. 2007): **a**) GST-fusions of hnRNP A1-NLS, hnRNP M-NLS and M9M bound to Kap β 2 then dissociated by 0.3-1.6 μ M RanGTP **b**) GST-hnRNP A1-NLS bound to Kap β 2 in the presence of buffer, MBP-hnRNP A1-NLS, MBP-hnRNP M-NLS or MBP-M9M **c**) Interactions of GST-Kap β 1 with Kap α , Kap α in the presence of IBB domain, M9M, or Kap α in the presence of M9M.

Transfection of Myc-tagged MBP-M9M in HeLa cells mislocalizes endogenous Kap β 2 substrates hnRNP A1, hnRNP M and HuR from the nucleus to the cytoplasm but not endogenous HDAC1 (Smillie, Llinas et al. 2004), which binds Kap α /Kap β 1. Therefore, M9M can inhibit Kap β 2-mediated nuclear transport in cells in a pathway specific manner (Figure 3-17).

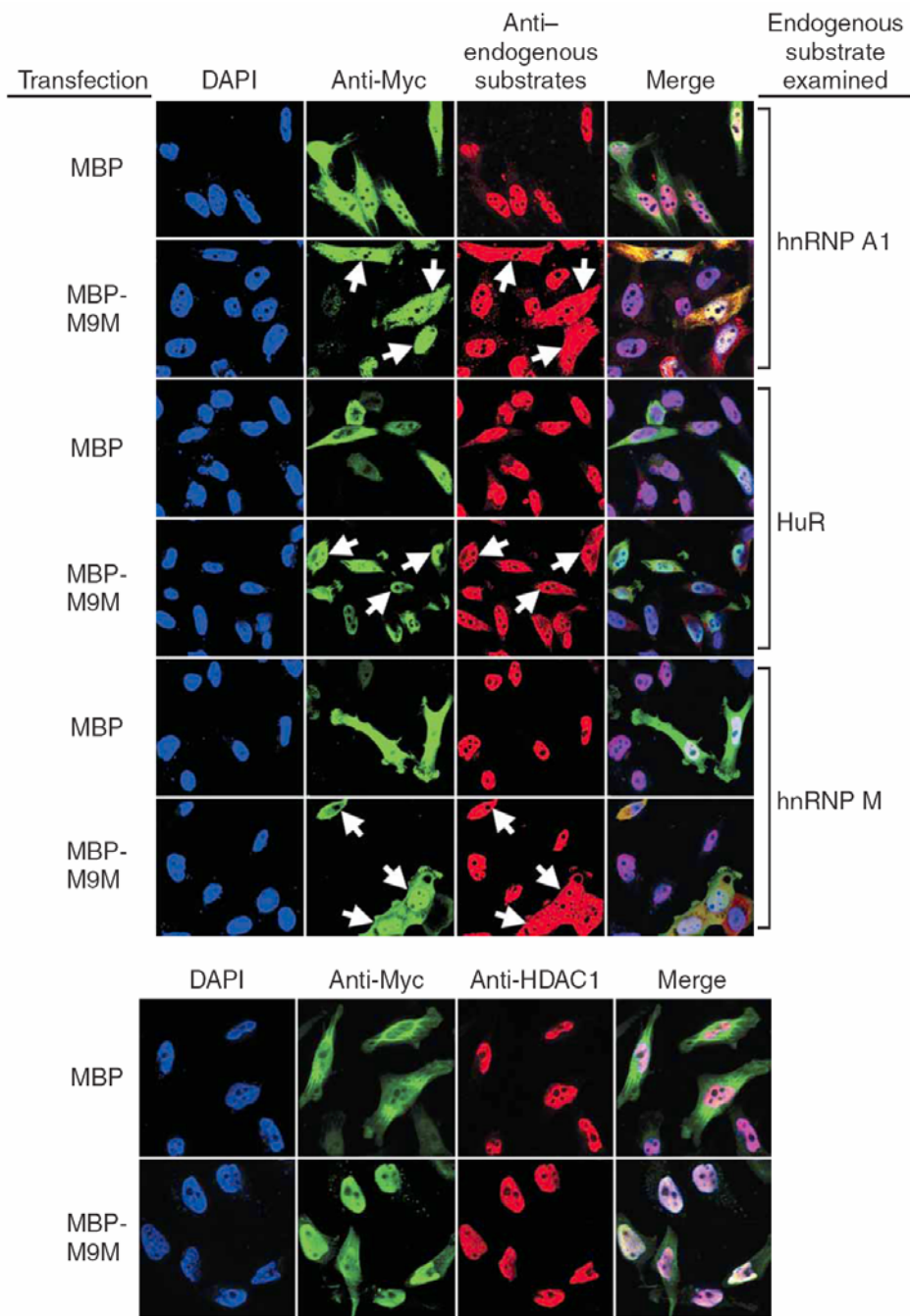


Figure 3-17) Top panel; Immunofluorescence and De-Convolution microscopy of Hela cells transfected with plasmids encoding myc-tagged MBP or MBP-M9M, using anti-myc and hnRNP A1, hnRNP M and HuR antibodies. Bottom panel: localization of endogenous HDAC1 (Kpα/Kpβ1 substrate) is determined as control(Cansizoglu, Lee et al. 2007).

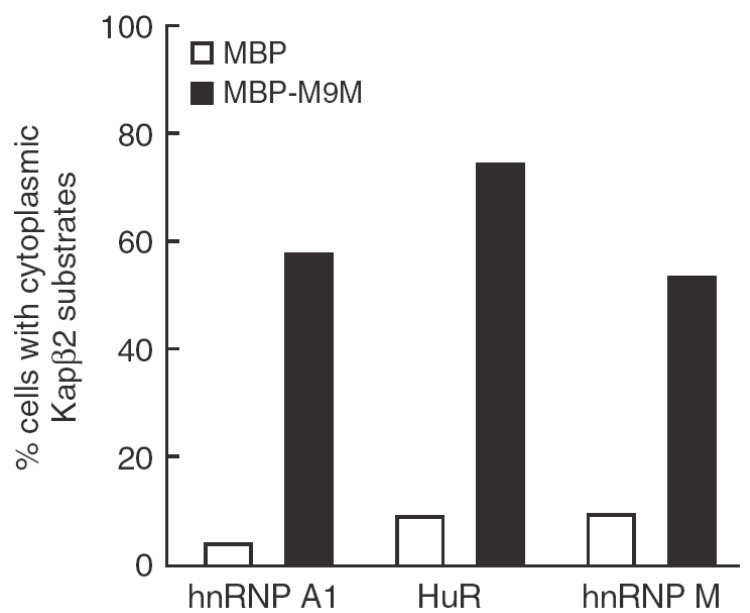


Figure 3-18) Histogram shows percentages of transfected cells with cytoplasmic Kap2 substrates

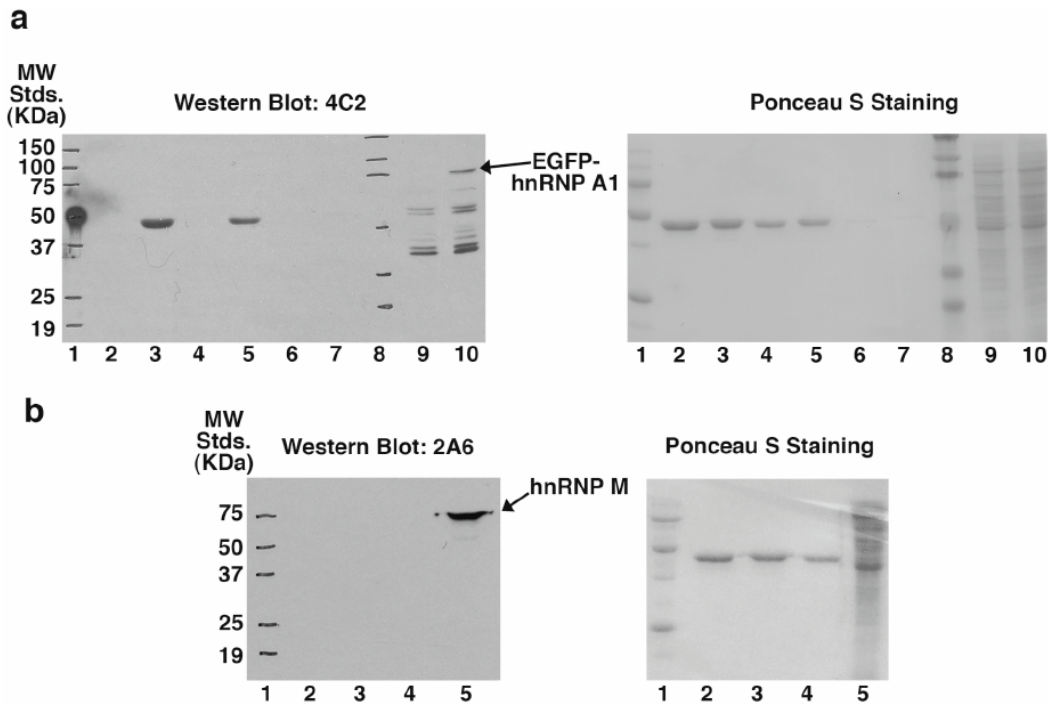


Figure 3-19) Western blots using antibodies against hnRNP A1 and M(Cansizoglu, Lee et al. 2007). **a)** Western Blot with antibody 4C2 (left), which recognizes human hnRNPs A1, A2 and B1, and visualization of proteins by Ponceau staining (right). Lanes 2, 4 and 6 contain 2 ug, 1 ug, and 0.1 ug of MBP-M9M; 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.

In summary, both bPY- and hPY-NLSs bind to Kap β 2 in extended conformation with structural conservation at their Arg and Pro-Tyr residues of the C-terminal R/K/H-PY motifs and at their N-terminal basic/hydrophobic motifs. Segments, flanking and between the consensus motifs, show large structural differences as well as sequence variability. The structure confirms both the requirement for intrinsic structural disorder in PY-NLSs and the identification of N-terminal hydrophobic/basic and C-terminal R/K/H-PY consensus motifs. Finally, our discovery of asymmetric NLS binding energy hotspots in hnRNPs M and A1 allowed us to design the first Kap β 2 specific import inhibitor, M9M peptide, which binds Kap β 2 200-fold tighter than the natural NLSs peptides and thus specifically inhibits Kap β 2-mediated nuclear import in cells.

CHAPTER FOUR
Structural studies on Unliganded Kap β 2

CONFORMATIONAL HETEROGENEITY OF KAPB2 IS SEGMENTAL

Abstract

Karyopherin β 2 (Kap β 2) or Transportin imports numerous RNA binding proteins into the nucleus. Kap β 2 binds substrates in the cytoplasm and targets them through the nuclear pore complex where RanGTP dissociates them in the nucleus. Here we report the 3.0 Å crystal structure of unliganded Kap β 2, which consists of a superhelix of 20 HEAT repeats. Together with previously reported structures of NLS and Ran complexes, this structure provides understanding of conformational heterogeneity that accompanies ligand binding. The Kap β 2 superhelix is divided into three major segments. Two of them (HEAT repeats 9-13 and 14-18), which constitute the substrate binding site, are rigid elements that rotate relative to each other about a flexible hinge. The third (HEAT repeats 1-8), which constitute the Ran binding site, exhibits conformational changes throughout its length. An analogous segmental architecture is also observed in Importin β suggesting that it is functionally significant and may be conserved in other import Karyopherins.

Introduction

Kap β s have high B-factors in crystals indicating that they possibly have flexible structures. The overall architecture consisting of tandem HEAT repeats is thought to enhance its flexibility in adopting different conformations. Multiple crystal structures of four different full length Kap β s and small-angle X-ray scattering (SAXS) models of six different Kap β s have been reported (Fukuhara, Fernandez et al. 2004). Previously solved crystal structures are: 1) human Imp β complexed with substrates Kap α IBB domain and SREBP-2, respectively (Cingolani, Petosa et al. 1999; Lee, Sekimoto et al. 2003), 2) yeast homolog Kap95p complexed with RanGTP and a Nup1p fragment, respectively (Lee, Matsuura et al. 2005; Liu and Stewart 2005), 3) human Kap β 2 or Transportin in complex with RanGppNHp, substrates hnRNP A1-NLS and hnRNP M-NLS, respectively (Chook and Blobel 1999; Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007) and 4) unliganded and substrate-bound Cse1p (Matsuura and Stewart 2004; Cook, Fernandez et al. 2005).

SAXS models are available for multiple states of Imp β , Kap β 2, Crm1, Cse1p, Xpo-t and its yeast homolog Los1p, and comparisons within each Karyopherin series show substantial protein flexibility (Fukuhara, Fernandez et al. 2004). However, low resolution SAXS models limit characterization of flexibility to large scale changes in radius of gyration and description of extended versus

compact conformations(Fukuhara, Fernandez et al. 2004). Crystal structures of Cse1p(Matsuura and Stewart 2004; Cook, Fernandez et al. 2005) and of full-length Imp β complexes(Cingolani, Petosa et al. 1999; Lee, Sekimoto et al. 2003; Lee, Matsuura et al. 2005; Liu and Stewart 2005) show that both Karyopherins undergo large conformational changes upon ligand binding, again showing that Kap β s are indeed quite flexible. However, despite the large number of crystal structures, there are relatively few studies of Kap β conformational flexibility(Cingolani, Petosa et al. 1999; Lee, Sekimoto et al. 2003; Cook, Fernandez et al. 2005; Lee, Matsuura et al. 2005; Conti, Muller et al. 2006). With the exception of a recent review on Kap β flexibility(Conti, Muller et al. 2006), most studies involve qualitative descriptions of differences in superhelical pitches and regions of structural changes. Analyses may be complicated by the unusual non-globular and non-modular architecture of these proteins and the limited number of structures within each Kap β series (Imp β , Kap95p, Kap β 2 and Cse1p). We have solved the 3.0 Å structure of unliganded Kap β 2 to extend the structural map of this import pathway that transports numerous RNA binding proteins into the nucleus. A suite of Kap β 2 structures (unliganded, hnRNP A1-NLS, hnRNP M-NLS and Ran complexes) is now available to explain conformational changes that accompany substrate recognition, Ran binding and substrate release(Chook and Blobel 1999; Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007). Most importantly, through three independent methods of domain motion analyses by

rotation vector clustering(Hayward and Berendsen 1998), B-factor grouping(Painter and Merritt 2006) and structural superpositions, we have identified three major segments of the Kap β 2 superhelix that show intrinsic flexibility and also respond to binding of different ligands. We also observe segmental architecture in Imp β , suggesting generality across the nuclear import factors.

Materials and Methods

Expression and purification of Kap β 2

Full length human Kap β 2 (residues 1-890; accession AAB58254) is in pGEX-Tev vector (Modified from pGEX-4T-3) as a GST fusion protein and purified as previously described(Lee, Cansizoglu et al. 2006) selenomethionine Kap β 2 was expressed BL21 (DE3) cells using M9 minimal media as previously described(Lee, Cansizoglu et al. 2006) and its purification is similar to that for native Kap β 2. The purified native and selenomethionine proteins were concentrated to 20 mg/ml and 35 mg/ml, respectively, for crystallization.

Crystallization, data collection and structure determination

Native Kap β 2 was crystallized by vapor diffusion in hanging drops using 3.2M potassium formate, 0.1 M HEPES pH 7.5 and 10% glycerol in the reservoir solution. The crystals were flash frozen in liquid propane. 3.0 Å data from these crystals (spacegroup P2₁, a = 129.9, b = 169.3, c = 141.1 and β = 93.1° with four molecules in the asymmetric unit) were collected at X-ray wavelength 0.97933 Å and temperature 100 K at beamline 19ID of the Advanced Photon Source, Argonne National Laboratory. Data was processed using HKL2000 (Otwinowski and Minor 1997). MR using Kap β 2 fragments from the Kap β 2-Ran structure (1QBK) as search models failed, indicating significant conformational differences in both N- and C-terminal arches between the two states. The Kap β 2-Ran structure was the only the Kap β 2 structure available at the time of structure determination.

Selenomethionine Kap β 2 was crystallized by vapor diffusion in hanging drops using 2.7 M potassium formate and 20% glycerol in the reservoir solution. Unfortunately, analysis these crystals (spacegroup P2₁, a = c = 108.5, b = 294.0 and β = 92° with four molecules in the asymmetric unit) were pseudomerohedrally twinned. In an effort to obtain single selenomethionine Kap β 2 crystals, mixtures of selenomethionine and native proteins were crystallized. A

1:1 molar mixture of both proteins gave single crystals (spacegroup $P2_1$, $a=108.5$, $b=294.0$, $c=108.3$ and $\beta=92.1^\circ$), and a 3.6 Å SAD dataset ($\lambda_{\text{peak}}=0.97903$ Å) was collected at beamline 19ID, Argonne National Laboratory.

We performed MR using the program Phaser(McCoy, Grosse-Kunstleve et al. 2005) and multiple search models, each comprising many different Kap β 2 segments from the Kap β 2-Ran structure. Only a search model of Kap β 2 residues 396-864 (H9-H19 or C-terminal arch) resulted in a successful molecular replacement solution. This result suggests that the C-terminal arch of selenomethionine unliganded Kap β 2 is more similar to that of the Ran complex than the native unliganded Kap β 2, but the N-terminal arches of both unliganded Kap β 2s are significantly different from the Ran state. However, to avoid model, 72 selenium sites (of the total 104) were extracted from the partial MR model and iterative refinement of these heavy atom sites allowed the heavy atom model to be extended to 100 selenium sites. The last four selenium belonging to N-terminal methionines were not located. Phase refinement followed by density modification with solvent flipping, both using the program CNS(Brunger, Adams et al. 1998), resulted in an electron density map where three of the four Kap β 2s in the asymmetric unit could be modeled using O(Jones, Cowan et al. 1991) and COOT(Emsley and Cowtan 2004). This incomplete and low resolution model was not refined.

Instead, the unrefined selenomethionine Kap β 2 model was used as a MR search model 3.0 Å resolution native Kap β 2 data. Using the program Phaser(McCoy, Grosse-Kunstleve et al. 2005), the Kap β 2 segment corresponding to residues 88-656 (H3-H14 or N-terminal arch) gave a successful molecular replacement solution for all four molecules in the asymmetric unit, suggesting similarities in the N- but not the C-terminal arches between the selenomethionine and native Kap β 2s. Overall superhelical structures of the native unliganded Kap β 2 (refined), the selenomethionine unliganded Kap β 2 (unrefined) and the Kap β 2-Ran complex are compared in Figure 4-1. The resulting electron density map for the native Kap β 2 crystal was interpretable for all four molecules in the asymmetric unit. Refinement of the completed model (residues 6-890 for Kap β 2 chains A, C, D and residues 31-890 for chain B) was done with 50-3.0 Å data using the program Refmac5(Murshudov, Vagin et al. 1997) from the CCP4 package(CCP4 1994). The final model refined to an R_{factor} of 24.6% and R_{free} of 28.4%.

Table 4-1) Data collection and refinement statistics for crystals of full length native and selenomethionine derivatives of unliganded Kap β 2

Data Collection:		Native:	Selenomethionine:
Space group		P2 ₁	P2 ₁
Unit Cell Parameters:			
a, b, c (Å)		129.9, 169.3, 141.1	108.5, 294.0, 108.3
α, β, γ (°)		90.0, 93.1, 90.0	90.00, 92.1, 90.00
Resolution (Å)		50-3.0 (3.1-3.0)*	50-3.6(3.4-3.6)*
Rsym		0.099 (0.45)*	0.126 (0.88)*
I / σ _I		24 (4.8)*	22 (2.7)*
Completeness (%)		100% (99.8%)*	99.9% (100%)*
Redundancy		5.0 (4.9)	4.2 (4.1)*
Refinement:			
No. of unique reflections:		114642	
Rfactor [†] / Rfree [†]		0.261/0.286	
rmsd from ideal bond lengths (Å)		1.144	
rmsd from ideal bond angles (°)		0.008	
Ramachandran Plot:			
Most favoured regions (%)	90.6	Generously allowed regions (%)	0.8
Additional allowed regions (%)	8.6	Disallowed regions (%)	0.0
Model:		Average B Factor (Å²):	
Chain A (Residues: 6-319, 370-890)		70.0	
Chain B (30-158, 167-319, 372-890)		73.7	
Chain C (5-21, 30-158, 164-319, 369-890)		81.5	
Chain D (7-319, 369-890)		67.4	

[#] $R_{\text{sym}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$; $I_i(h)$ is the i^{th} measurement of reflection h and $\langle I(h) \rangle$ is the weighted mean of all measurements of h .

* values in parentheses are calculated for data in the highest resolution shell.

[†] $R\text{-factor} = \sum_h |F_{\text{obs}}(h) - F_{\text{calc}}(h)| / \sum_h F_{\text{obs}}(h)$. R_{free} is calculated with 5% of the data.

MR was not successful with native unliganded Kap β 2 using Kap β 2-Ran as search model because the two states show large conformational differences in both the N- and C-terminal arches (Figures 4-1 and 4-5c). However, we found a MR solution for selenomethionine Kap β 2 that corresponds to the C-terminal arch of Kap β 2 in the Ran state. Figure 4-1 shows conformational differences between selenomethionine, native unliganded and Ran-bound Kap β 2s. The C-terminal arch of selenomethionine Kap β 2 is somewhat intermediate in conformation between native unliganded and Ran-bound Kap β 2s, but is similar enough to the latter for MR. This partial MR model allowed us to locate most of the 104 selenium sites without determining them *de novo* using SHELX(Sheldrick, Dauter et al. 1993). Unfortunately, structure determination of the selenomethionine protein was hampered by low resolution, pseudo-translation and weak SAD phases (Phasing: R_{cullis} 0.76, phasing power 1.06, Se site B-factors 41-85 Å² when occupancies set to 1). Fortunately, the N-terminal arch of selenomethionine Kap β 2 was similar to the native protein (Figure 4-1), and we could use this segment of the unrefined selenomethionine model to bootstrap our way to a MR solution for the higher resolution native dataset. Interestingly, when comparing the selenomethionine and native Kap β 2s, a hinge was also located at H13-H14

that rotates rigid segments H9-H13 and H14-H18 relative to each other.

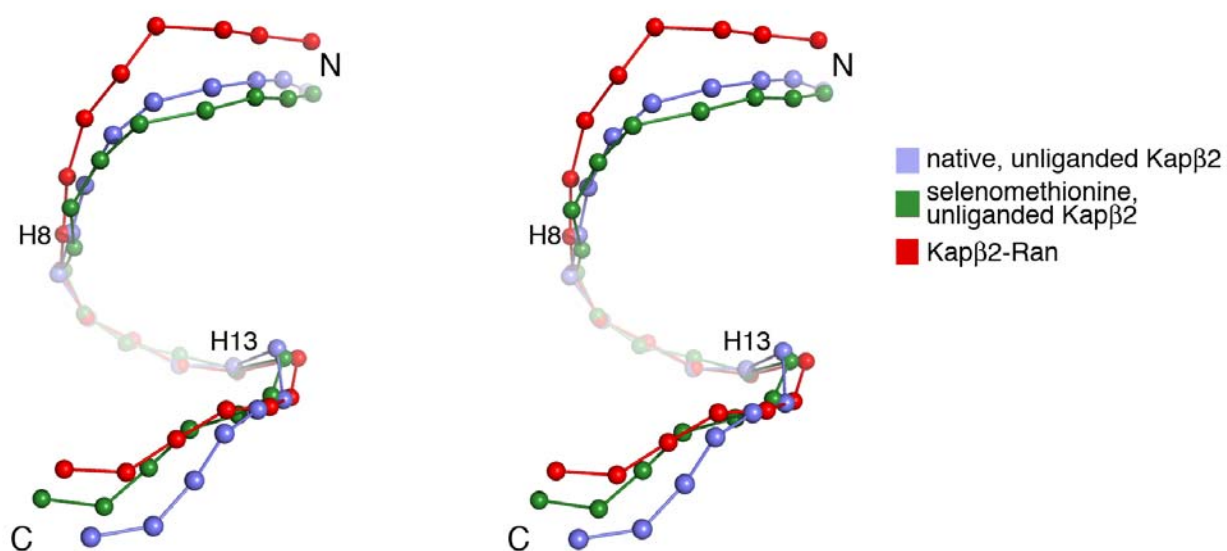


Figure 4-1) Comparison of native and selenomethionine unliganded and Ran-bound Kap β 2s: All three Kap β 2s are superimposed at H8-H13. The chains are drawn as spheres at the geometric center of each HEAT repeat. The selenomethionine model is not refined and thus its coordinates are not deposited in the PDB.

Structure refinement

There are four unliganded Kap β 2 chains in the asymmetric unit and the protein

has an unusual non-globular helical repeat architecture. Therefore, NCS restraints may be used in many different ways when refining the structure. Refinement of the model (residues 6-890 for Kap β 2 chains A, C, D and residues 31-890 for chain B) was done with 50-3.0 Å data, using the program Refmac5(Murshudov, Vagin et al. 1997) from the CCP4 package(1994). Results of 16 different refinement protocols are listed in Tables 4-2 and 4-3. Without NCS restraints (refinement #1, Table 4-2), Rfree is 28.4% and Rfactor is 24.6%. However, without the use of NCS, refinement is not efficient as ratio of observables to unknowns is low at 3.0 Å resolution and a large asymmetric unit. Tight NCS restraints of all four chains resulted in very high R factors (Rfree = 42.6% and Rfactor = 40.5%; Table 4-3 refinement #7), suggesting significant conformational heterogeneity amongst the four chains and NCS restraints should be relaxed.

Table 4-2) Summary of refinement using NCS to constrain Kap β 2 chains A to D and chains B to C

<i>Refinement #</i>	<i>Number of NCS Groups</i>	<i>NCS Groups</i>	<i>NCS Restraints</i>	<i>R_{free} (%)</i>	<i>R_{factor} (%)</i>
1	0	None	None	28.4	24.6
2	2	A,D*(H1-20) [§] B,C*(H1-20)	Tight Tight	31.4	29.1
3	8	A,D(H1-4,H5-12,H14-17, H19-20) B,C(H1-4,H5-12,H14-17, H19-20)	Tight, Tight, Tight, Tight Tight, Tight, Tight, Tight	28.4	25.9
4	6	A,D(H1-12,H14-17, H19-20) B,C(H1-12,H14-17, H19-20)	Tight, Tight, Tight Tight, Tight, Tight	29.2	26.5
5	6	A,D(H1-4,H5-17, H19-20) B,C(H1-4,H5-17, H19-20)	Tight, Tight, Tight Tight, Tight, Tight	29.1	26.8
6	6	A,D(H1-4,H5-12,H14-20) B,C(H1-4,H5-12,H14-20)	Tight, Tight, Tight Tight, Tight, Tight	28.6 [†]	26.1 [†]

* A, D indicates that Kap β 2 chains A and D are constrained by NCS; B, C denotes Kap β 2 chains B and C are constrained by NCS.

[§] (H1-20) denotes Kap β 2 HEAT repeats 1 through 20.

[†] Reported in Table 4-1.

Analysis of the sans-NCS restraints structure (refinement #1, Table 4-2) suggested that chain A is very similar to D, and chain B is very similar to C. When entire chains AD and BC are refined with tight NCS restraints, R_{free} dropped by ~11% (Table 4-2). However, R factors are still significantly higher than without NCS restraints suggesting that the chains may need to be divided into multiple NCS groups. Structural comparison of chains A and C from refinement #1 suggested that Kap β 2 could be divided according to conformational heterogeneity into four segments H1-H4, H5-H13, H14-H18 and H19-H20. Segment junctions H13 and H18 show clear conformational differences. This information guided our division of the molecule into NCS groups. We experimented with dividing of each chain into four NCS groups (H1-H4, H5-H12, H14-H17 and H19-H20) and using tight NCS restraints in Refmac5 (Murshudov, Vagin et al. 1997) to constrain respective groups within the AD and BC pairs (refinement #3). The 3% drop in R_{free} indicated that relaxation of NCS constraints is warranted and that there is conformational heterogeneity of segmental nature even within the two similar pairs of Kap β 2. When contiguous segments are combined to generate tri-segmented Kap β 2 (refinements #4-6), R factors increased except when the last two segments are combined. We chose that structure from refinement #6 (R_{free} = 28.6% and R_{factor} = 26.1%, Table 4-1) for analysis. Comparison of this structure with refinement #1 shows that they are

virtually identical, with C_α r.m.s.d. ~ 0.3 Å.

We also experimented with NCS restraints on all four chains (Table 4-3). Dividing Kap β 2 into four segments (H1-4, H5-12, H14-17 and H19-20) for refinement of all groups with tight NCS restraints decreased R_{free} by 10%, again consistent with segmental conformational heterogeneity within the four chains. Relaxing NCS restraints of the H5-H12 and H19-H20 groups decreased R_{free} further by about 4% suggesting that conformational differences within these amongst the four unliganded Kap β 2 chains. Interestingly, superposition of H5-H13 of chains A and C gives C_α r.m.s.d. 1.3 Å suggesting that differences may occur in sidechains or the H8 loop (always omitted when superimposing structures) rather than in helical arrangements.

Table 4-3) Summary of refinement using NCS to constrain Kap β 2 chains A, B, C and D

<i>Refinement #</i>	<i>Number of NCS Groups</i>	<i>NCS Groups</i>	<i>NCS Restraints</i>	<i>R_{free} (%)</i>	<i>R_{factor} (%)</i>
7	1	A-D*(H1-20) [§]	Tight	42.6	40.5
8	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Tight, Tight, Tight	32.3	30.4
9	4	A-D(H1-4,H5-12,H14-17, H19-20)	Medium, Tight, Tight, Tight	32.4	30.0
10	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Medium, Tight, Tight	30.6	27.8
11	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Tight, Medium , Tight	32.8	30.0
12	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Tight, Tight, Medium	30.7	28.8
13	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Medium, Tight, Medium	28.7	26.0
14	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Loose, Tight, Medium	28.5	25.7
15	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Medium, Tight, Loose	28.5	25.9
16	4	A-D(H1-4,H5-12,H14-17, H19-20)	Tight, Loose, Tight, Loose	28.3	25.6

* A-D indicates that Kap β 2 chains A, B, C and D are all subjected to NCS restraints.

§ (H1-20) represents HEAT repeats 1 through 20.

Superposition of H19-H20 of chains A and C (C_{α} r.m.s.d. 2.2 Å) suggests that this C-terminal segment is indeed conformationally heterogeneous in all four chains. Comparison of structures from refinements #1 and #16 shows that they are virtually identical, with C_{α} r.m.s.d. ~ 0.2 Å. We also experimented with dividing Kap β 2 into NCS groups other than H1-4, H5-12, H14-17, and H19-20. These include four 5-HEAT repeat groups and combining two of our contiguous segments at a time to make three groups. In all cases, R_{free} is around $\sim 36\%$.

TLS refinement(Painter and Merritt 2006) with Refmac5(Murshudov, Vagin et al. 1997) using 12 TLS groups (three groups for each molecule) resulted in $R_{\text{free}} = 27.0\%$ and $R_{\text{factor}} = 23.3\%$. However, since there was no improvement in the electron density map and the structure is similar to that in refinement #6, we did not deposit this TLS-refined structure in the PDB. Finally, the weakest density occurs at the N-terminal region of Kap β 2 molecules, which is also part of the MR search model. To reduce model bias, we calculated simulated annealing omit (SA-omit) maps using CNS(Brunger, Adams et al. 1998) to omit each of the first four HEAT repeats of Kap β 2.

Analyses of conformational heterogeneity

We used three independent/orthogonal methods to analyze flexibility. The first method involved pair wise superpositions of HEAT repeats. We performed comprehensive superpositions: 1) scanned the length of Kap β 2 chains for groups of 1-19 HEAT repeats, 2) superimposed A-helices of individual repeats and examined the B helices for reorientations within each HEAT repeat, and 3) superimposed B helices and examined A-helices for changes between HEAT repeats. The Superpose program from CCP4 package(CCP4 1994) was used to determine C $_{\alpha}$ - C $_{\alpha}$ distance r.m.s.deviation. Helix Packing Pair program(Dalton, Michalopoulos et al. 2003) was used to determine angles between helices.

The second method to analyze conformational flexibility involved domain motion analyses by clustering rotation vectors (that relates two protein chains) using the program DynDom(Hayward and Berendsen 1998). Domain motion analysis to locate hinged motions was performed using coordinates for unliganded Kap β 2 (chains A-D) and Kap β 2s in the Ran, hnRNP A1-NLS (chains A and B; 2H4M) and hnRNP M-NLS (chains A and B; 2OT8). The nine Kap β 2 molecules extracted from four coordinate files were subject to pairwise analyses for the presence of hinged conformational differences using the program

DynDom(Hayward and Berendsen 1998). The window length for these analyses was set at 5 residues and the minimum ratio for inter-domain to intra-domain displacement was set to 1.0.

The final method involved analyses of the spatial distribution of individual atomic B-factors in a single protein chain using the program in TLSMD (TLS Motion Determination) for TLS group analysis(Painter and Merritt 2006). Individual structures with refined B-factors were used. This procedure identifies the portions of the protein with similar atomic displacement/thermal factor characteristics and determines the optimal TLS segments that behave like a pseudo-rigid body. For each protein chain, this analysis partitions the molecule into TLS groups and further analyzes the pseudo-rigid body translational and rotational motion of each group and how well these parameters fit the refined atomic thermal parameters in the crystal structure.

Methods 1 and 2 locate conformational differences by comparing pairs of structures, whereas method 3 locates pseudo-rigid segments using a single protein chain. Finally, all figures were prepared using the program PYMOL(DeLano 2002).

Results and discussions

Structure determination and overall structure of unliganded Kap β 2

Kap β 2 has 20 HEAT repeats, each consisting of two antiparallel helices A and B. The A helix of HEAT repeat 1 (H1) is abbreviated to H1A. All helices are connected by short loops or small helices except for H8A and H8B, which are connected by a 62-residue acidic loop named the H8 loop (Lee, Cansizoglu et al. 2006). The Kap β 2 superhelix can be divided into two arches. The N-terminal arch is composed of H1-H13, and the C-terminal is composed of H9-H20. Ran binds in the N-terminal arch and substrates have been observed so far to bind in the C-terminal arch.

We have solved the 3.0 Å crystal structure of unliganded Kap β 2 using a combination of single-wavelength anomalous dispersion (SAD) and molecular replacement (MR). Full-length native and selenomethionine Kap β 2s (residues 1-890) were crystallized. The native crystals (space group P2₁, a = 129.9 Å, b = 169.3 Å, c = 141.1 Å and β = 93.1°, four Kap β 2 molecules in the asymmetric unit; Table 4-1) diffracted to 3.0 Å, and MR using the Kap β 2-Ran structure (1QBK; the only Kap β 2 structure available at the time of structure determination)

was unsuccessful, suggesting significant conformational differences between the unliganded and Ran states.

Selenomethionine Kap β 2 crystals (space group P2₁, a = 108.5 Å, b = 294.0 Å, c = 108.5 Å and β = 92.0°, four molecules in the asymmetric unit) were pseudomerohedrally twinned, but a 1:1 molar mixture of native and selenomethionine proteins resulted in a single crystal (space group P2₁, a = 108.5 Å, b = 294.0 Å, c = 108.3 Å and β = 92.0°) from which 3.6 Å SAD data was obtained. Even though MR using Kap β 2 residues 396-864 from the Kap β 2-Ran structure was successful, we calculated experimental SAD phases to avoid model bias. To facilitate determination of the 104 selenium sites in the asymmetric unit (26 methionines per Kap β 2), 72 selenium sites were extracted from the partial MR model followed by determination of the next 28 by iterative phase refinement and difference Fourier methods. The last four selenium sites that correspond to N-terminal methionines were not located. Phase refinement with 100 selenium sites followed by density modification with solvent flipping resulted in an electron density map where three of the four Kap β 2s in the asymmetric unit could be traced. This incomplete and low resolution SAD model was not refined, but instead used for MR with the 3.0 Å native data. Residues 88-656 gave a MR solution, resulting in an interpretable electron density map. Differences between the SAD and native models are discussed in Experimental Procedures. Given the four molecules in the asymmetric unit and the unusual helical repeat structure of Kap β 2, the final

3.0 Å model of residues 6-890 for chains A, C, D and residues 31-890 for chain B was refined using many combinations of non-crystallographic symmetry (NCS) constraints (described in methods section, Tables 4-2 and 4-3). Statistics of a refinement using NCS to constrain chains A to D and chains B to C, both in three segments, are reported in Table 4-1 ($R_{\text{free}} = 28.6\%$ and $R_{\text{factor}} = 26.1\%$).

All four copies of Kap β 2s are arranged with their superhelical axes parallel to the crystallographic y-axis, and two pairs (Chains A-B and C-D) are related to each other by a pseudo-translation of approximately half unit cell along the crystallographic z-axis. Approximately 4700 Å² or 12 % of each molecule's total surface area is buried in crystal contacts. The low percentage of surface area involved crystal contact compared to other protein crystals is consistent with the low resolution and high solvent content (> 70 %) of the Kap β 2 crystal (Carugo and Argos 1997). All NCS related Kap β 2s have similar crystal contacts and similar overall structure.

The unliganded Kap β 2 molecule is a rather symmetrical superhelix with a 77 Å pitch, a 115 Å length and a 65 Å diameter (Figure 4-2). Its overall structure is similar to those of Kap β 2-substrate complexes. Both unliganded and substrate-bound Karyopherins are more elongated along their superhelical axes compared to the Ran-bound molecule, consistent with previous SAXS studies of Kap β 2 (Fukuhara, Fernandez et al. 2004).

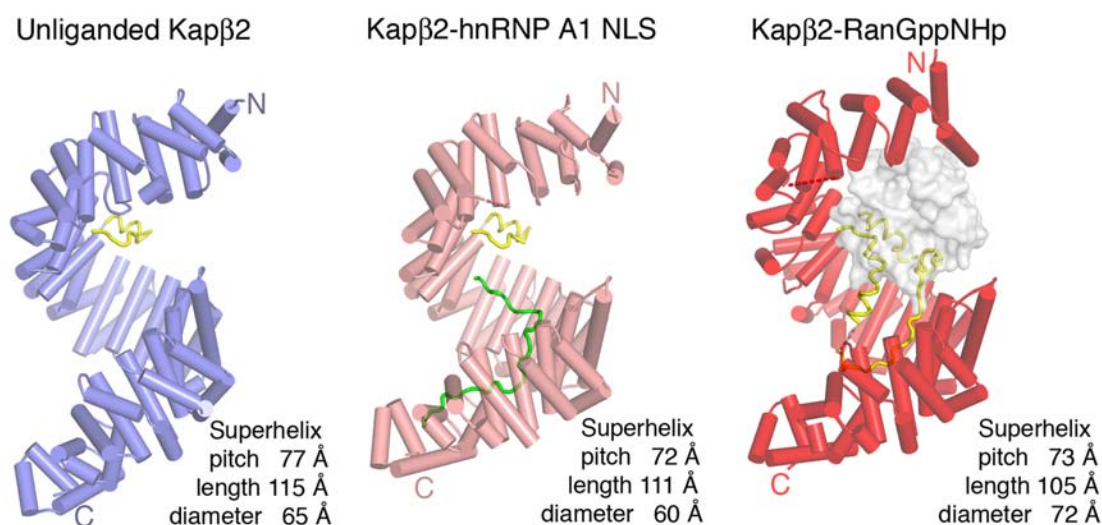


Figure 4-2) Cartoon diagrams of unliganded, substrate- and Ran-bound Kap β 2s. α -helices are represented as cylinders and structurally disordered loops as dashed lines. Unliganded Kap β 2 is blue, Kap β 2 bound to substrate is pink, Kap β 2 bound to Ran is red and H8 loops in all three structures are yellow. Substrate hnRNP A1-NLS is in green and Ran is drawn as surface representation in grey.

Conformational flexibility of unliganded Kap β 2

The four unliganded Kap β 2s in the asymmetric unit show obvious conformational heterogeneity. We used three independent methods to analyze conformational

heterogeneity. The first involves several different pair wise superpositions: 1) we scanned the length of Kap β 2 chains superimposing groups of 1-20 HEAT repeats, 2) we superimposed A helices of each repeat and examined B helices for changes within HEAT repeats, and 3) we superimposed B helices and examined A helices of the next repeat for changes between HEAT repeats. The second method is domain motion analyses by clustering rotation vectors using the program DynDom (Hayward and Berendsen 1998). The third method involves analyses of the spatial distribution of individual atomic B-factors in a single protein chain using the program TLSMD (Translation/Libration/Screw Motion Determination; (Painter and Merritt 2006)). The first two methods compare pairs of structures, whereas the third analyzes single protein chains.

When the structure was refined without using any NCS constraints (Table 4-2, refinement #1), unliganded Kap β 2 chains A and D are similar (C_{α} r.m.s.d. 1.2 Å) and chains B and C are also similar (C_{α} r.m.s.d. 1.0 Å). In contrast, other pairwise comparisons give C_{α} r. m. s. d. of 2.6 - 3.2 Å. Given similarity of the AD and BC pairs, the two pairs were constrained by NCS in the final structure reported in Table 4-1. Structural comparisons below are therefore limited to chains A and C.

Unliganded Kap β 2 chains A and D are similar (C_{α} r.m.s.d. 1.2 Å). Chains B and C are also similar (C_{α} r.m.s.d. 1.0 Å). In contrast, other pairwise comparisons give C_{α} r. m. s. d. of 2.6 - 3.2 Å. Given similarity of the AD and BC pairs,

structural comparisons below for unliganded Kap β 2s are limited to chains A and C.

Unliganded Kap β 2s are similar in their central regions. Exclusion of terminal repeats H1-H4, and H19-H20 decreased C α r. m. s. d. from 2.9 Å for all residues to 1.8 Å (Table 4-4). Moreover, when superimposed at H8-H13 (C α r. m. s. d. < 1.0 Å), their termini deviate up to 11 Å (Figure 4-3a). HEAT repeats at the N- and C-termini also have higher B-factors despite ~2200 Å² surface area buried by crystal contacts in these regions. For example, average B-factors for H1-H4, H5-H18 and H19-H20 in chain B are 112 Å², 65 Å² and 84 Å², respectively (electron density maps for the N- and C-termini are shown in Figures 4-4a and 4-4b). Thus, it appears that unliganded Kap β 2, like Ran- and substrate-bound Kap β 2, other Kap β s and Kap α , is flexible at its termini (Conti, Uy et al. 1998; Chook and Blobel 1999; Cingolani, Petosa et al. 1999; Kobe 1999; Cook, Fernandez et al. 2005; Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007).

Table 4-4) Summary of C_{α} r.m.s.d. for Kap β 2 superpositions

<i>HEAT Repeats</i>	<i>Unliganded Kapβ2 Chains A,C</i>	<i>Unliganded Kapβ2 Kapβ2- Substrate</i>	<i>Unliganded Kapβ Kapβ2- RanGppNHp</i>	<i>Kapβ2- Substrate Kapβ2- RanGppNHp</i>
1-20	2.9	4.2	7.7	6.1
1-13	2.3	2.2	6.8	6.0
1-7	2.4	2.5	6.6	5.5
1-4	1.1	3.0	6.3	5.4
5-20	2.6	4.1	5.8	4.1
5-18	1.8	3.3	5.1	4.1
5-13	1.3	1.3	4.1	4.0
8-20	2.3	3.6	4.5	2.4
8-13	0.8	1.2	2.0	1.7
14-20	2.5	0.9	1.7	2.2
14-18	0.5	0.6	1.0	1.3
19-20	2.2	0.5	1.7	1.9

Repeats H5-H13 of all four unliganded Kap β 2s are virtually identical, but their superhelical paths deviate at H14 (Figure 4-3a). To characterize conformational flexibility the C-terminal arch, we analyzed each unliganded chain for pseudo-

rigid segments using the program TLSMD (Translation/Libration/Screw Motion Determination; (Painter and Merritt 2006)). The molecule is partitioned through analysis of the spatial distribution of individual atomic B-factors into 1-20 contiguous pseudo-rigid segments termed TLS (Translation/Libration/Screw) groups. We used these multi-group TLS models as a starting point for analysis of large-scale motion in Kap β 2. The junction for the 2-group TLS model of unliganded Kap β 2 mapped to residues 560-610 (H13-H14; Figure 4-3b). As partitioning progressed, the H13-H14 junction persisted and additional TLS groups include the two flexible N- and C-termini segments (H1-H4 and H18-H20) discussed above.

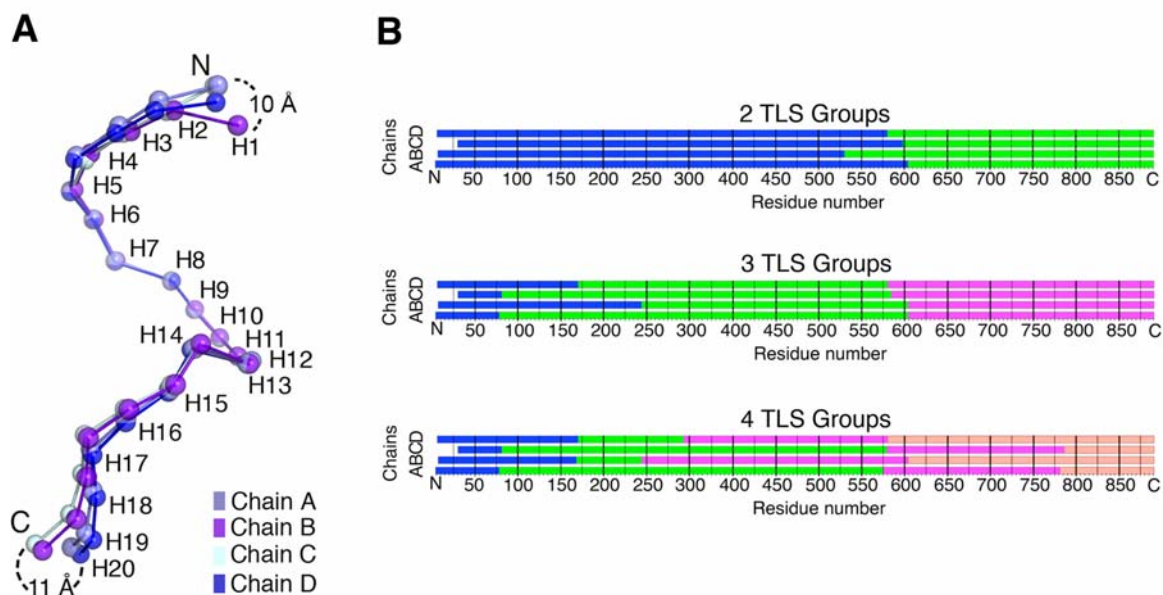


Figure 4-3) The four molecules of unliganded Kapβ2 in the asymmetric unit show conformational heterogeneity. **A)** Four non-crystallographically related Kapβ2s (chains A-D) are superimposed at H8-H13. The chains (each in a different shade of blue) are drawn as spheres at the geometric center of each HEAT repeat. Maximum deviations between geometric centers at the termini are shown. **B)** Results of TLSMD analysis for partitioning of the four Kapβ2 chains into 2-4 TLS groups (each in a different color).

Kapβ2 flexibility was also characterized by another independent method of clustering rotation vectors between two different conformations using the program DynDom (Hayward and Berendsen 1998). Analyses of these domains in

full-length unliganded Kap β 2 molecules identified a hinge in H18, consistent with the flexible C-terminal segment (H19-H20). Removal of terminal flexible segments (H1-H4 and H19-H20) revealed a second flexible hinge spanning H13-H14. The location of this hinge corresponds to the prominent TLS junction above. A 9° rotation of two rigid segments, H5-H13 and H14-H18, relative to each other about this hinge axis explains the small conformational difference in C-terminal arches of chains A and C. This hinge-like motion of two rigid segments is consistent with a decrease in C_{α} r. m. s. d. for individual versus combined segments (Table 4-4; C_{α} r. m. s. d.: H5-H13 is 1.3 Å, H14-18 is 0.5 Å and H5-H18 is 1.8 Å).

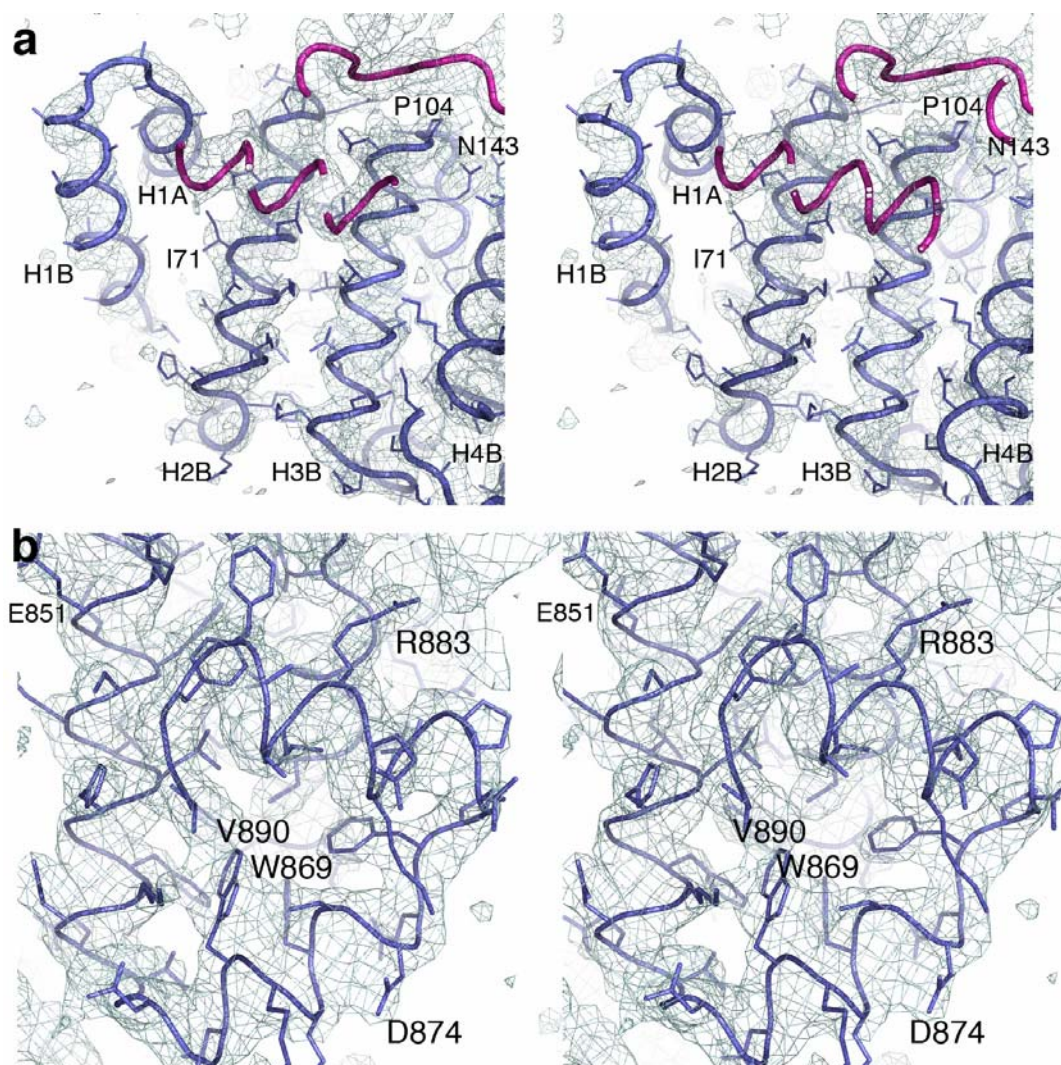


Figure 4-4) Electron density maps of the N- and C-termini of unliganded Kap β 2: **a)** 2Fo-Fc map (stereo diagram, 1.0 σ , blue mesh) of unliganded Kap β 2 is shown at its N-terminal H1-4 segment (blue). H1 is modeled as polyalanines and a neighboring molecule is in magenta. **b)** 2Fo-Fc map (stereo diagram, 1.0 σ , blue mesh) of unliganded Kap β 2 is shown at its C-terminal H19-20 segment (blue).

Based on structural superpositions and two independent methods of identifying pseudo-rigid bodies in a molecule, the intrinsically flexible unliganded Kap β 2 can be divided into four moving segments. Two small segments, H1-H4 and H19-H20, are found at the termini. Two larger central segments, H5-H13 and H14-H18, are somewhat rigid bodies that rotate relative to each other about a flexible hinge. Interestingly, results of structure refinements using NCS restraints are also consistent with the division of unliganded Kap β 2 into the four particular segments (Tables 4-2 and 4-3).

NLS recognition: a hinge in the Kap β 2 C-terminal arch

Kap β 2 binds RNA binding proteins through recognition of a signal in the substrates known as PY-NLS (Lee, Cansizoglu et al. 2006). The 20-30 residues long PY-NLSs are sequentially diverse, but share three common characteristics: structural disorder, overall basic character, and a set of weakly conserved hydrophobic or basic N-terminal motif and a C-terminal R/K/Hx₂₋₅PY motif. Crystal structures of Kap β 2 bound to two diverse PY-NLSs, from splicing factors hnRNPs A1 and M (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007), show that both NLSs bind a common interface (B helices of H8-H18) on the structurally invariant Kap β 2 C-terminal arch (H2-20 C α r.m.s.d. is 0.8 Å). Since

the Karyopherins in both substrate complexes are virtually identical, we will use the hnRNP A1-NLS complex to represent the substrate-bound state. Chain A in the unliganded Kap β 2 crystal represents the unliganded state.

Although the overall structures of unliganded and substrate bound Kap β 2s appear similar, a 4.2Å C α r.m.s.d. suggests substantial conformational difference (Table 4-4). With the exception of the first four HEAT repeats, most of their N-terminal arches are quite similar (H5-H13 C α r.m.s.d. is 1.3 Å). Thus, conformational differences mostly map to the C-terminal arch. Like unliganded Kap β 2, domain motion analysis using DynDom (Hayward and Berendsen 1998) to compare unliganded and substrate states also identified a flexible hinge at H13-H14. Subdomains on both sides of the hinge, H5-H13 and H14-H20, behave like rigid bodies (C α r.m.s.d. for H5-H13, H14-H20 and H5-H20 are 1.3 Å, 0.9 Å and 4.1 Å, respectively; Table 4-4). NLS binding rotates one rigid segment relative to the other 16° about the hinge axis (Figure 4-5). This movement originates from the cumulative effect of small rotations between helices H12B, H13A, H13B, H14A and H14B.

Interestingly, even though RanGTP binding results in large conformational changes in the Kap β 2 N-terminal arch, the C-terminal arch in both Ran- and substrate-bound states are quite similar (H9-H18 C α r.m.s.d. is 1.4 Å; Table 4-4; (Lee, Cansizoglu et al. 2006)). Furthermore, comparison of H8-H20 (N-terminal arches are removed due to large conformational changes) in the unliganded and

Ran states again identified a flexible hinge in H13-H14 that rotates rigid segments H9-H13 and H14-H18 $\sim 18^\circ$ relative to each other (C_α r.m.s.d. for H9-H13, H14-H18 and H9-H18 are 1.5 Å, 1.0 Å and 3.2 Å, respectively; Table 4-4). The hinge axes that relate unliganded Kap β 2 to substrate- and to Ran-bound conformations are also somewhat similar (Figure 4-5). Similar C-terminal arches in Ran and substrate states is due to occupation of Kap β 2's long H8 acidic loop in that arch, much of which overlaps spatially with bound NLSs (Chook and Blobel 1999; Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007). Thus, conformational similarity in the C-terminal arch in both substrate and Ran states supports the idea that the H8 loop is a pseudo-NLS (Lee, Cansizoglu et al. 2006).

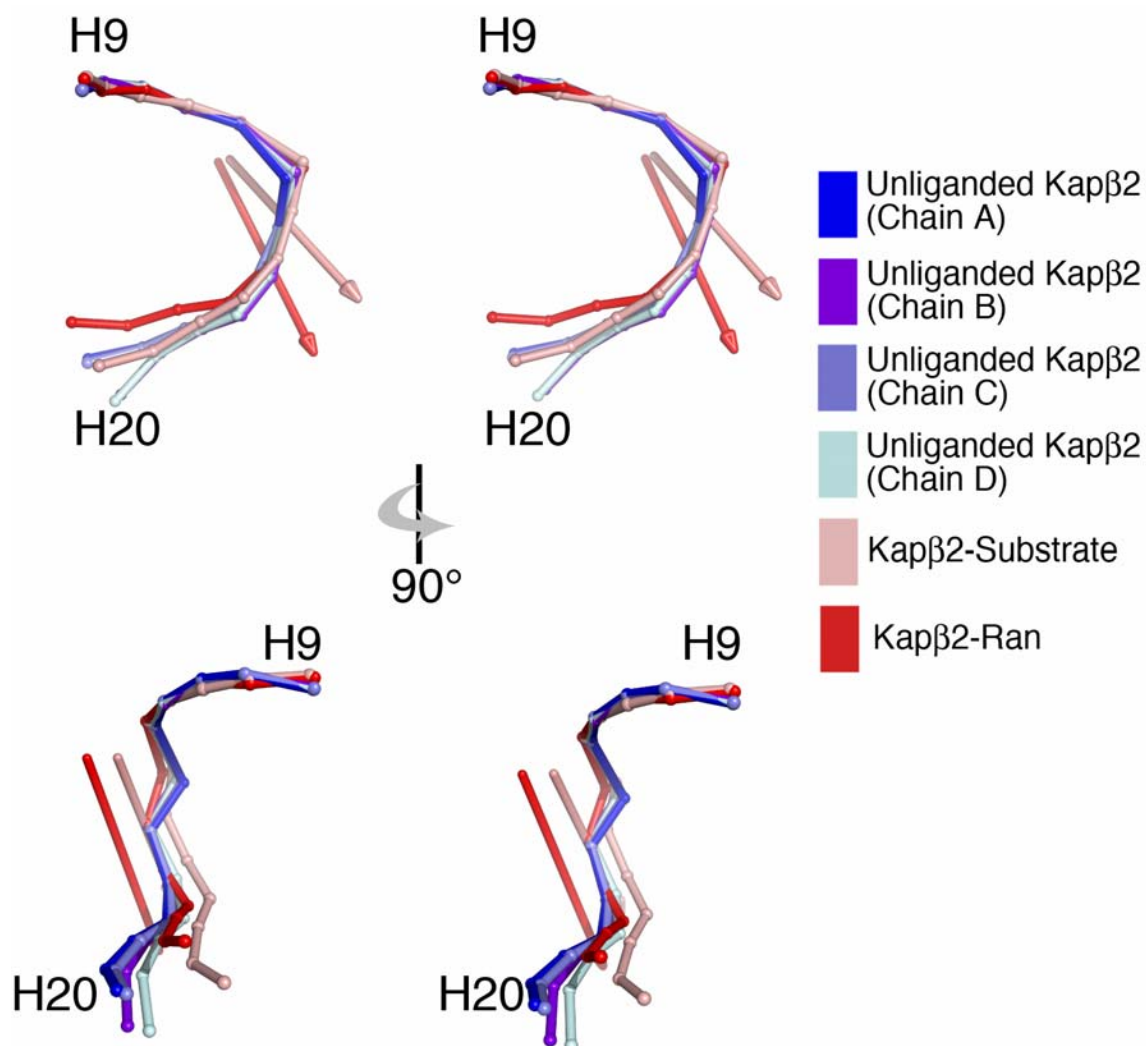


Figure 4-5) Hinge motion in the C-terminal arch of Kap β 2: Stereo diagrams of the C-terminal arch (H9-H20) of chains A-D of unliganded Kap β 2, substrate- and Ran-bound Kap β 2s superimposed at H9-13 and drawn as spheres at the geometric center of each HEAT repeat. The hinge axes that rotate unliganded Kap β 2 with respect to substrate- and Ran-bound Kap β 2s are in pink and red, respectively.

The same rigid segments, H9-H13 and H14-H18, rotate relative to each other in the different unliganded Kap β 2 chains and between unliganded, substrate- and Ran-bound Kap β 2s. Intrinsic segmental flexibility observed in the unliganded chains suggests varying degrees of rotation about the hinge. Paradoxically, similarity of the C-terminal arches in the NLS- and pseudo-NLS/Ran-bound Kap β 2s suggests a discrete energetically favorable arch conformation that binds ligands. What is the range of motion about this hinge? Does a range of rotation about the hinge allow fine-tuning of the substrate binding site to accommodate diverse PY-NLSs that vary significantly in length and composition? Or accommodate entirely new classes of undiscovered NLSs? Or accommodate interactions with nucleoporins? Crystal structures of Kap β 2 in complex with longer and more diverse PY-NLSs such as those in HuR and TAP (Fan and Steitz 1998; Truant, Kang et al. 1999; Lee, Cansizoglu et al. 2006), and structures of Kap β 2-nucleoporin complexes will shed light on these questions.

Kap β 2-Ran interaction: the flexible N-terminal arch and H8 loop

Ran binds in the N-terminal arch of Kap β 2, contacting B helices of H1-H4, H7-H8 and the H8 acidic loop (Chook and Blobel 1999). Dimensions of the Kap β 2 N-terminal arch change upon Ran binding (Figure 4-6c). The arch opens as the

width of its base increases by 13 Å compared to unliganded and substrate bound structures, enabling Ran to fit between the interfaces with H1-H4 and the H8 loop (Figures 4-6a and 4-6b). The C-terminal third of the N-terminal arch (H9-H13) does not contact Ran, and is similar between unliganded and Ran states (C_{α} r.m.s.d. 1.5 Å; Figure 4-6c). However, large changes in helical orientations occur within and between each HEAT repeat from H1-H8. These large helical reorientations occur as helical content of the A, B and connector helices changes, particular at helix termini, and connector loops also change in conformation. HEAT repeats move both laterally towards the dome of the arch and along the superhelix towards the N-terminus to form a larger N-terminal arch.

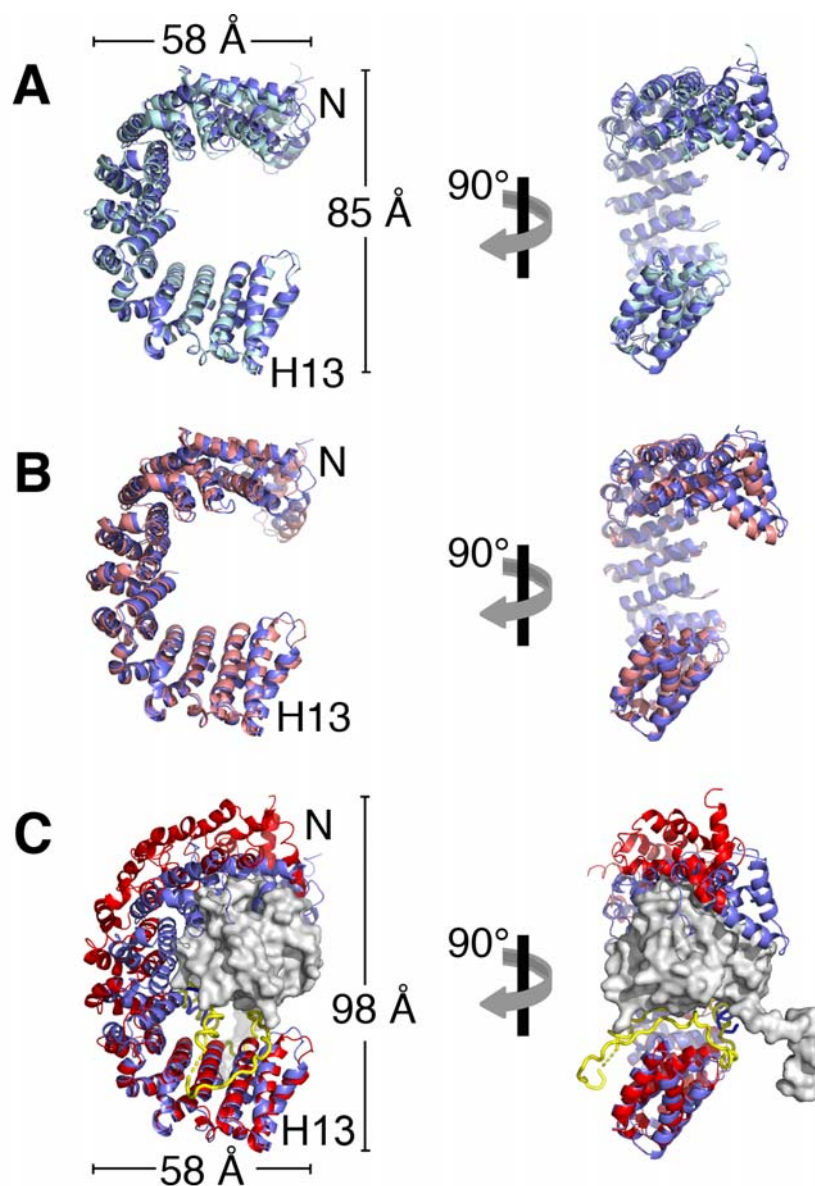


Figure 4-6) Conformational change in the Kap β 2 N-terminal arch: **A)** Ribbon diagram of the N-terminal arches of chains A (blue) and C (light blue) of unliganded Kap β 2, superimposed at H9-H13. **B)** Same as **A)**, except chain A of unliganded Kap β 2 is superimposed on substrate-bound Kap β 2 (pink). **C)** Same as **A)**, except chain A of Kap β 2 is superimposed on Ran-bound Kap β 2 (red), Ran is shown as a surface representation in grey and the H8 loop of the Ran complex is in yellow.

A second obvious conformational difference between unliganded and Ran Kap β 2s is found at the H8 loop. Proteolysis studies have suggested that the loop is exposed in unliganded and substrate bound Kap β 2s (Chook, Jung et al. 2002). This was confirmed by both the Kap β 2-substrate structures (Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007) and is now also confirmed in the unliganded Kap β 2 structure. The H8 loop in the unliganded Kap β 2 crystals is biochemically intact, but only 13 of its 62 residues are observed, indicating that most of the loop is indeed mobile and disordered. Electron density is present only for residues 312-319 and 370-375 (Figure 4-7). Ordered loop residues 312-319 emerge from helix H8A and residues 370-374 precede helix H8B (Figure 4-8), and these have similar structures in unliganded and substrate bound Kap β 2s. Residues 312-319 are in similar positions in all Kap β 2 structures, but residues 370-374 have shifted to direct the loop away from the arch in the unliganded and substrate structures (Figure 4-2) (Lee, Cansizoglu et al. 2006). In contrast, in the Ran complex, H8 loop residues 332-340 and 363-373 form a platform that interacts with the basic patch of Ran, while the rest of the loop resides in the C-terminal arch (Chook and Blobel 1999). In summary, the concave surface of the unliganded C-terminal arch is free to bind substrate. The structure of unliganded Kap β 2 provides additional evidence that the H8 loop is flexible and does not reside in the C-terminal arch

until Ran is present. Interactions with Ran change the conformation of the loop, converting it into a pseudo-NLS to displace substrate from the C-terminal arch.

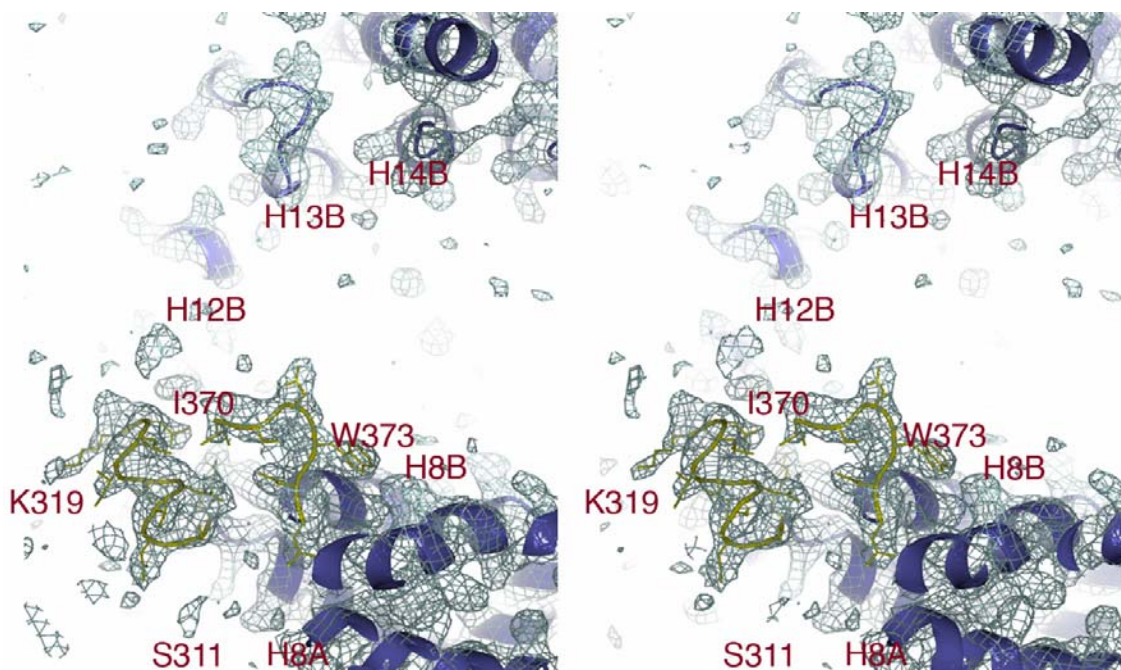


Figure 4-7) The 2Fo-Fc map (stereo diagram, 1.0 σ , blue mesh) of unliganded Kap β 2 is shown at its H8 loop and C-terminal arch (blue). H8 loop (yellow) residues 311-319 are shown connecting to H8A and residues 370-375 to H8B. No continuous density is observed beyond residue K319 until residue I370, suggesting that the rest of the H8 loop is structurally disordered. There is also no additional density bound to the C-terminal arch concave surface.

When Ran is absent, the N-terminal arches have almost identical dimensions (height 85 Å and base width 58 Å) that are too small to accommodate Ran. Segments H5-13 have similar conformations (C_{α} r.m.s.d. ~ 1.1 Å) but segments H1-4 show conformational heterogeneity (Table 4-4 and Figures 4-6a and 4-6b). H1-4 is intrinsically flexible, with high B-factors (>100 Å²), weak electron density and a different conformation in every structure (Figure 4-9; (Chook and Blobel 1999; Lee, Cansizoglu et al. 2006; Cansizoglu, Lee et al. 2007)). Structural here are of a continuous nature, with changes in helical orientations within and between HEAT repeats. Since the disordered H8 loop and the H1-H4 segment appear to be most flexible in the N-terminal arch, we speculate that either or both regions may serve as capture sites for initial interactions to “reel” in Ran. This is then followed by helical rearrangements in H5-H8 to optimally position the GTPase in the N-terminal arch and the rest of the H8 loop in the C-terminal arch. Kinetic studies of Ran binding and substrate dissociation will be necessary to investigate this model.

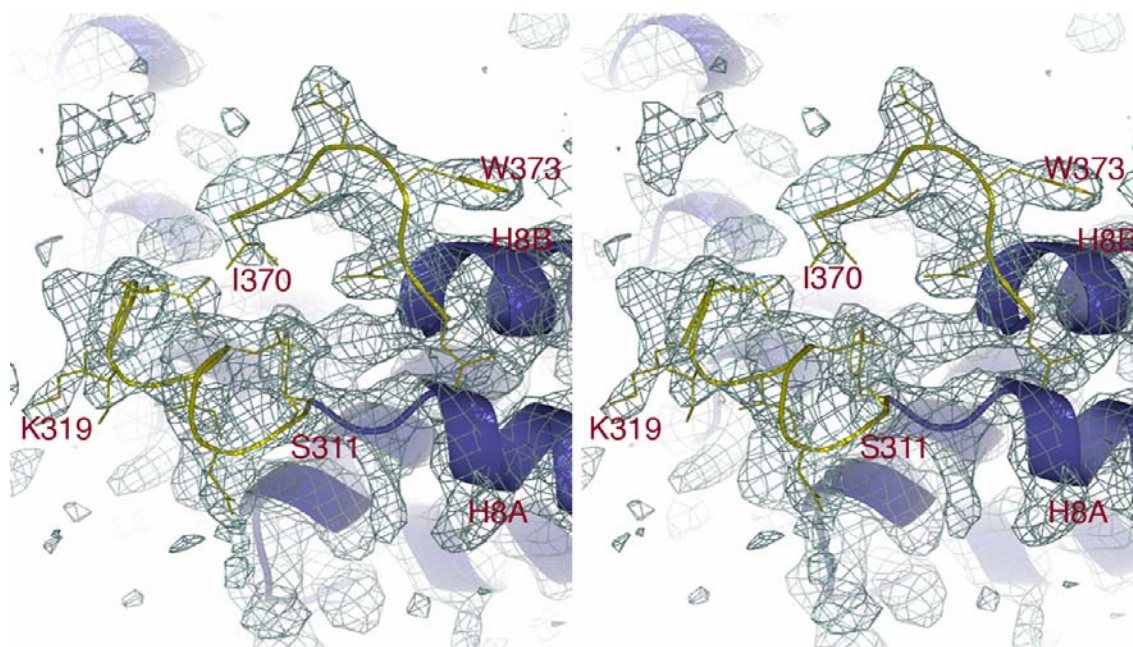


Figure 4-8) The 2Fo-Fc map (stereo diagram, 1.0σ, blue mesh) of unliganded Kapβ2 is shown at its H8 loop (yellow), with residues 311-319 shown connecting to H8A and residues 370-375 to H8B.

The ability of the N-terminal arch to change conformation along most of its length may also enable binding of additional ligands. Impβ binds substrate PTHrP in its N-terminal arch(Cingolani, Bednenko et al. 2002). Conformational flexibility of Kapβ2's N-terminal arch may also enable binding undiscovered substrates in that region. Finally, N-terminal arch flexibility is also likely important for interactions with nucleoporins.

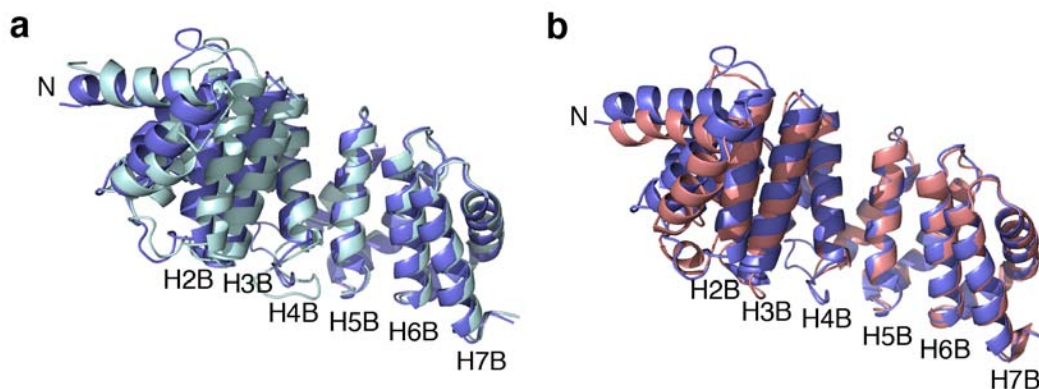


Figure 4-9) Conformational flexibility at the N-terminal H1-H4 segment of Kap β 2: **a)** Ribbon diagram showing H1-H7 of chains A (blue) and C (light blue) of unliganded Kap β 2, superimposed at H5-H13. **b)** Same as **a)**, except chain A of unliganded Kap β 2 is superimposed with substrate-bound Kap β 2 (pink).

In summary, conformational heterogeneity in Kap β 2 can be organized into three major segments. The N-terminal H1-H8 segment shows large changes along its entire length upon Ran binding. Segments H9-H13 and H14-H18 in the C-terminal arch are rigid bodies that rotate about a flexible hinge to bind NLSs and the H8 loop. Three additional small flexible segments are also observed. Flexible C-terminal H19-H20 segment is detected in most structural comparisons (Figure 4-10). The N-terminal H1-H8 segment can be further divided: intrinsically

flexible H1-H4 is different in every structure whereas H5-H8 changes conformation only when bound to Ran.

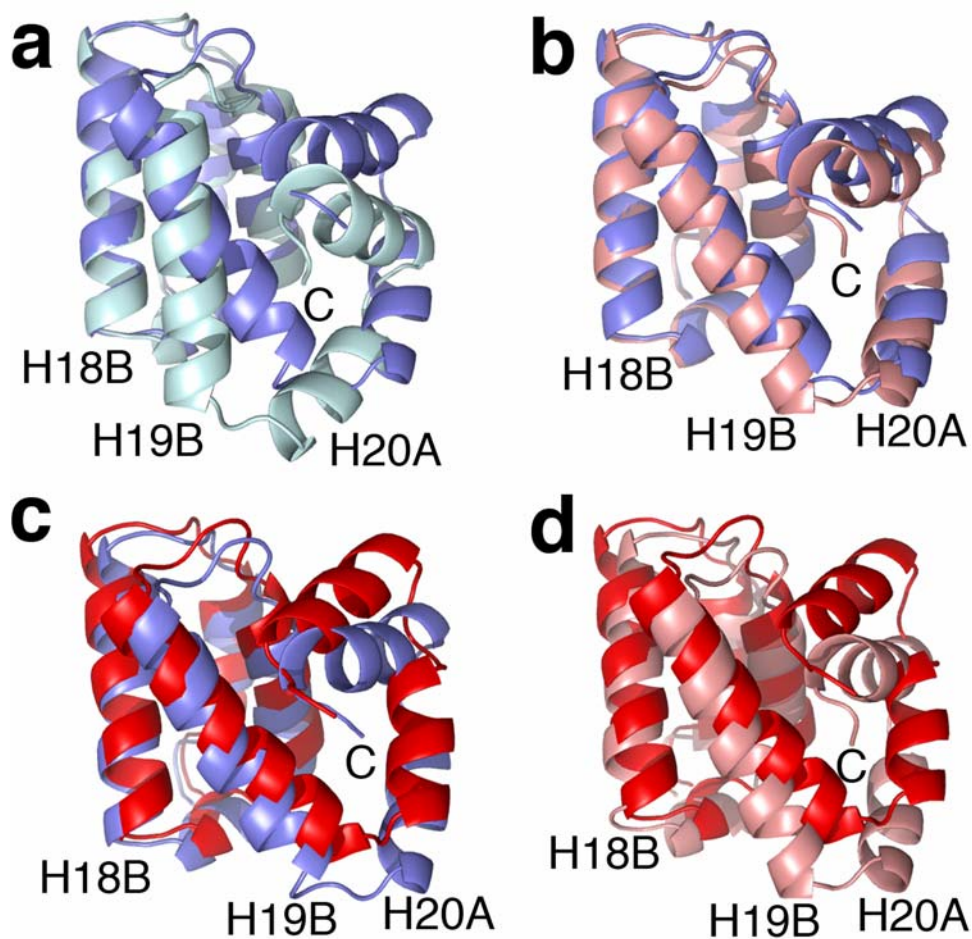


Figure 4-10) Conformational flexibility at the C-terminal H19-H20 segment of Kap β 2: **a)** Ribbon diagram showing H18-H20 of chains A (blue) and C (light blue) of unliganded Kap β 2, superimposed at H14-H18. **b)** Same as **a)**, except chain A of unliganded Kap β 2 is superimposed with substrate-bound Kap β 2 (pink). **c)** Same as **a)**, except chain A of unliganded Kap β 2 is superimposed with Ran-bound Kap β 2 (red). **d)** Same as **a)**, except substrate-bound Kap β 2 (pink) is superimposed with Ran-bound Kap β 2 (red).

Segmental architecture in Imp β and Kap95p

Imp β and its *S. cerevisiae* homolog Kap95p are the best studied Kap β s. Many structures of Imp β have been determined, but most are of the N-terminal half of the molecule (Vetter, Arndt et al. 1999; Bayliss, Littlewood et al. 2000; Lee, Imamoto et al. 2000; Bayliss, Littlewood et al. 2002; Cingolani, Bednenko et al. 2002), thus unsuitable for analysis of conformational flexibility along the superhelix. Three crystal structures of full length Imp β are available: two crystal forms of Imp β bound to the IBB domain of Kap α and a SREBP2 complex (Cingolani, Petosa et al. 1999; Lee, Sekimoto et al. 2003). Two crystal structures are also available for full length Kap95p, which is 33% identical to Imp β . These are Kap95p bound to RanGTP and Kap95p bound to a Nup1p fragment (Lee, Matsuura et al. 2005; Liu and Stewart 2005). We performed domain motion and TLS group analyses on both pairs of structures to examine potential segmental architecture analogous to those in Kap β 2.

Hinges in Imp β were previously reported to rotate H1-H11 by 10° with respect to H12-H17, and to rotate the latter by 10° with respect to H18-H18 in different Imp β -IBB crystal forms (Cingolani, Petosa et al. 1999). Hinge analysis of the IBB and SREBP complexes located a flexible hinge in H13 (Figure 4-11a). The N-terminal H1-H13 segment (C α r.m.s.d. is 1.9 Å) swings 22° about the hinge

axis with respect to the C-terminal H14-H19 segment (C_{α} r.m.s.d. is 1.1 Å), changing the superhelical pitch by 18Å to bind conformationally diverse substrates. The N-terminal segment appears not to be rigid, with small changes in helical orientations along its length.

Similar analyses of the Kap95p structures located a flexible hinge at H14-H15. In this case, subdomains on either sides of the hinge are H6-H14 (C_{α} r.m.s.d. is 1.5 Å) and H15-H19 (C_{α} r.m.s.d. is 1.2 Å). These segments rotate 38° relative to each other, resulting in a superhelical pitch difference of 10Å (Figure 4-11b). TLS group analysis maps a junction between TLS groups to H14 (residues 598-600) consistent with the position of the flexible hinge. Finally, examination of the Kap95p N-terminal arch alone (H1-H14) identified another hinge at H5, where segment H1-H4 (C_{α} r.m.s.d. is 0.7 Å) rotates 11° relative to segment H6-H14 (C_{α} r.m.s.d. is 1.5 Å).

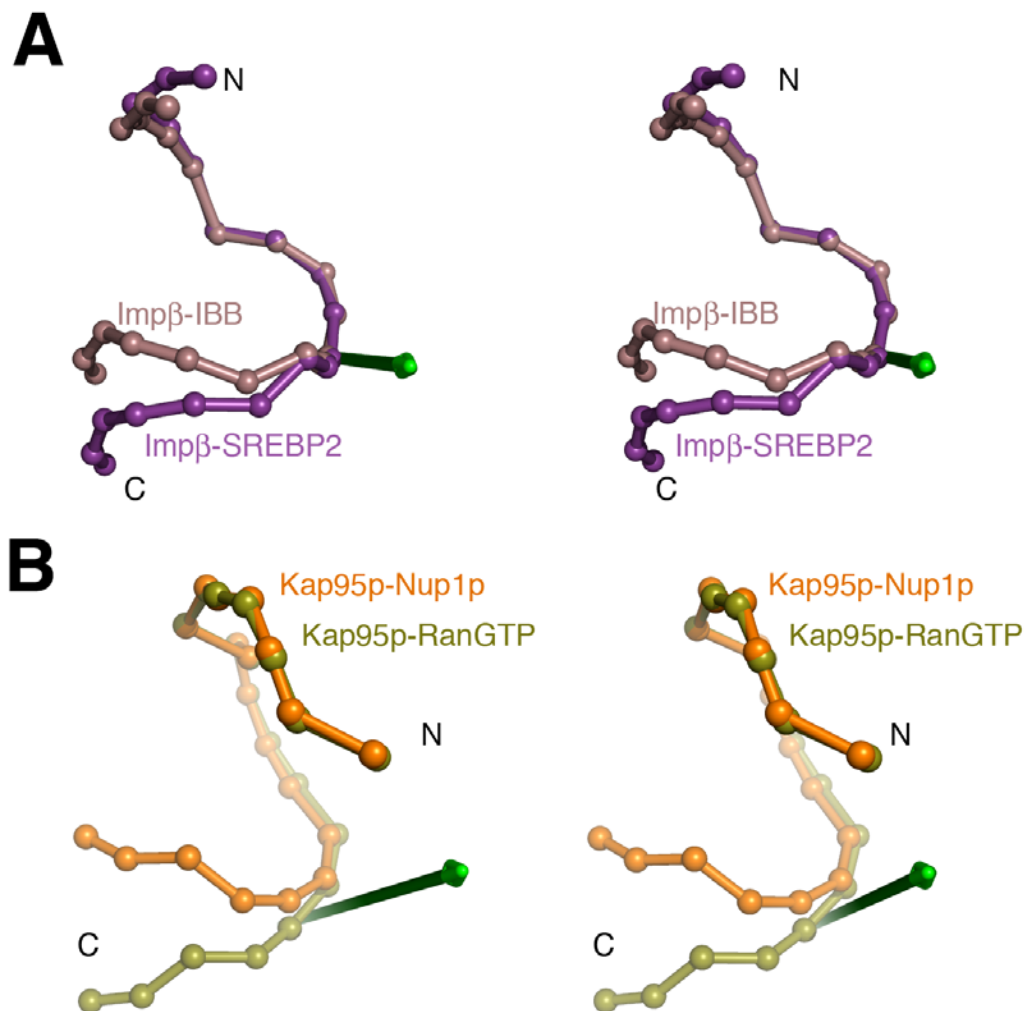


Figure 4-11) Hinge motion in Imp β and Kap95p structures: **A)** IBB-bound Imp β (brown) and SREBP2-bound Imp β (purple) are superimposed at HEAT repeats 5-10, and shown as spheres at the geometric center of each HEAT repeat. The hinge axis that rotates H1-H12 with respect to H14-H19 is shown in green. **B)** RanGTP-bound Kap95p (light green) and Nup1p-bound Kap95p (orange) are superimposed at HEAT repeats 6-13, and shown as spheres at the geometric centers of their HEAT repeats. The hinge axis that rotates H1-H14 with respect to H15-H19 is shown in green.

In summary, like Kap β 2, conformational heterogeneity in Imp β and Kap95p is also segmental. Imp β can be divided into three segments that rotate about two flexible hinges in H12-H13 and H18 (Cingolani, Petosa et al. 1999) and Kap95p can also be divided into three segments, with flexible hinges at H5 and H14-H15. Structures of full length unliganded Imp β and Kap95p are not available, but a SAXS model of unliganded Imp β show this state to be significantly more extended than the Ran or substrate states (Fukuhara, Fernandez et al. 2004). Crystal structures of full length unliganded Imp β /Kap95 and additional structures of the full length ligand-bound proteins will be necessary for more comprehensive analyses conformational heterogeneity in this nuclear import pathway.

Conclusions

The crystal structure of unliganded Kap β 2 has an overall superhelical structure similar to those of the Ran, hnRNP A1-NLS and hnRNP M-NLS complexes, confirming previous SAXS studies that the different Kap β 2 states adopt extended S-shaped structures. More importantly, the four Kap β 2 chains in the asymmetric unit show conformational heterogeneity, allowing characterization of intrinsic flexibility. Comparison with substrate- and Ran-bound Kap β 2s also showed

significant conformational differences. Conformational flexibility analyses using three independent methods of structural superpositions, clustering of rotation vectors and B-factor analysis revealed the moving parts. Kap β 2 can be divided into three major segments. Rigid segments H9-H13 and H14-H18 in the C-terminal substrate-binding arch rotate relative to each other about a flexible hinge in H13-H14. H1-H8 in the N-terminal arch show continuous changes along its length upon Ran binding. Using the same approaches, Imp β and its yeast homolog Kap95p can also be divided into three major segments that rotate about flexible hinges, suggesting that conformational heterogeneity in import-Kap β s may be generally segmental.

CHAPTER FIVE

Studies of the Kap β 2 acidic H8 loop

Introduction

The H9 loop of Kap β 2 is thought to play a critical role in binding and release of substrates. This long loop is inserted between Helix A and B of HEAT 8 (H8A and H8B – see Figure 3-1). It is composed of approximately 62 residues with significant acidic character (318-

LKGDVEEDETIPDSEQDIRPRFHRSRTVAQQHDEDGIEEEDDDDDDEIDDDD

TI -370). It has been shown that conformational changes in this loop leads to the release of substrate inside the nucleus[1]. In Ran bound state, the loop occupies the substrate binding site of Kap β 2 (Figure 5-1a).

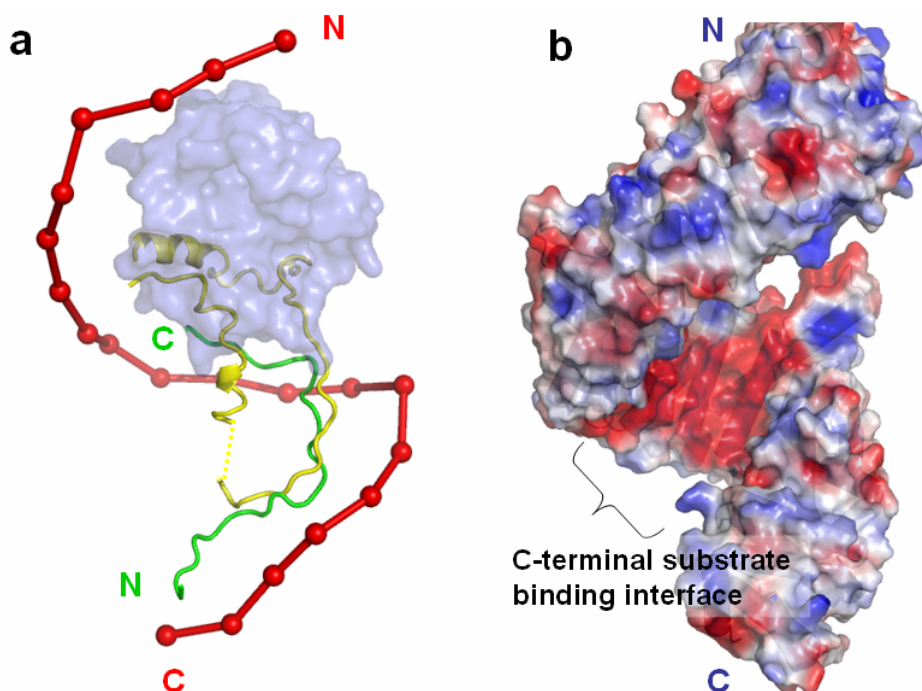


Figure 5-1) Loop segment from Kap β 2 sits directly on the substrate binding interface in the Ran bound state of Kap β 2. **a)** Kap β 2 is shown in red with each red sphere representing the geometric center of consecutive HEAT repeats. Kap β 2 loop region is shown in yellow and Ran is shown in blue transparent surface representation. M9NLS is depicted in green. Figure is produced by superimposing Kap β 2 main chains for Kap β 2.RanGPPMP (PDB ID 1QBK) and Kap β 2.M9NLS (PDB ID 2H4M) complexes. **b)** Unliganded Kap β 2 is shown in electrostatic surface potential representation in similar orientation. Inner substrate binding interface is highly acidic.

Kap β 2 can be divided into two known functional regions, the N terminal arch is H1-H13 (residues 1-600) and C terminal arch is H9-H20 (residues 390-890). The protein can be cleaved into two structurally intact fragments as thrombin cleaves at the H8 loop[2]. Proteolytic cleavage of the H8 loop does not affect binary interactions with substrates or Ran, but, dissociation of the import substrate upon binding to Ran is completely abolished[2]. Therefore, the loop participates in Ran-mediated substrate dissociation by displacing substrate in the C-terminal arch of Kap β 2. Replacing this loop with a short linker, thus substantially truncating the loop, also does not affect substrate binding but Ran.GTP mediated dissociation of the substrate is lost.

The detailed biophysical mechanism of substrate dissociation from Kap β 2 C-terminal arch still needs investigation. It is however interesting to note that the NLS fragments often have a basic character while the loop itself is highly acidic in composition. The inner concave surface of the substrate binding arch at the C-terminus of Kap β 2 is also very acidic (Figure 5-1b). When the substrate is not bound to Kap β 2, the loop segment, which is also highly acidic in composition, remains unstructured (See Chapter 4). We propose that this is possibly due to electrostatic repulsion between the loop and the inner Kap β 2 surface of the C-terminal arch. Therefore, we also suggested that the basic character of Ran residues interacting with Kap β 2 act to neutralize the acidic character of this loop

to allow the loop to interact with the substrate binding region to displace the cargo from Kap β 2[1] (Figure 5-2).

There are several basic Ran residues contacting Kap β 2 loop residues in Ran bound state. These are R129, K132, K134 and K159. To interfere with Ran binding acidic loop residues that interact with basic Ran residues were mutated to alanines.

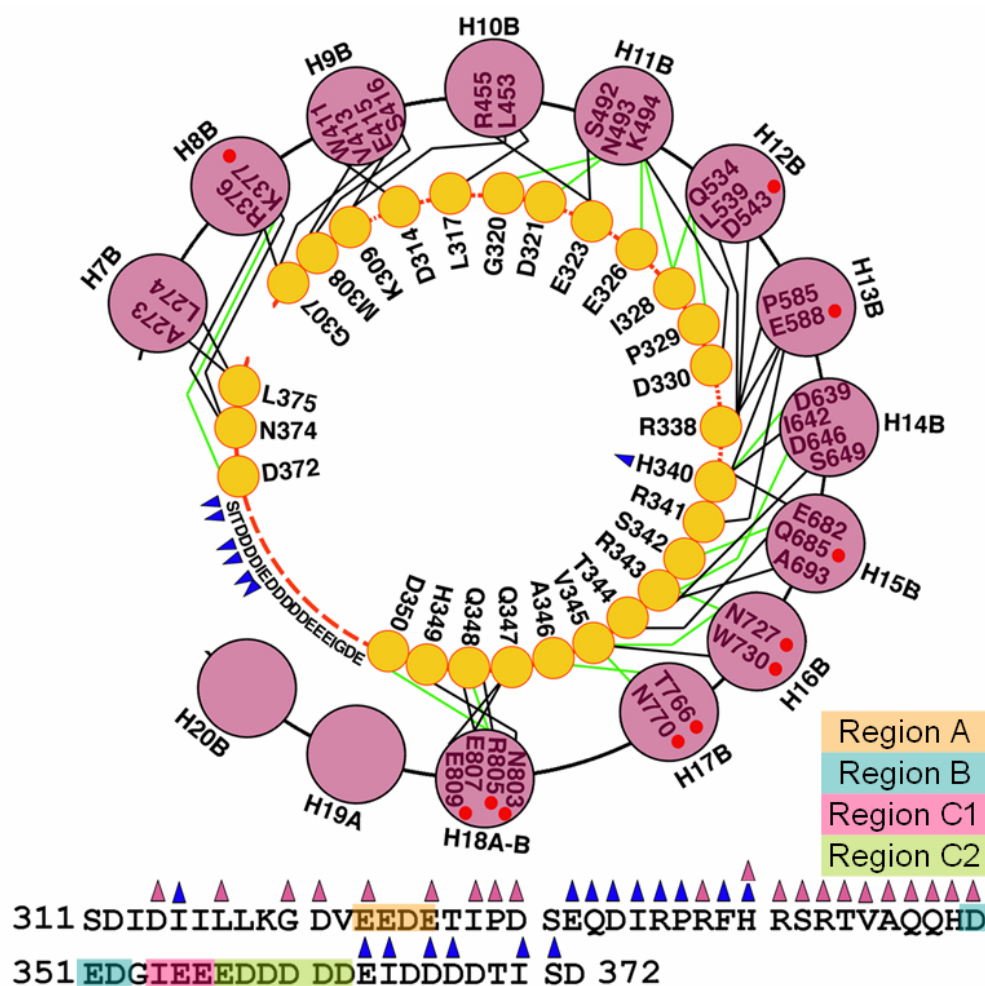


Figure 5-2) Contacts (<4.0 Å) between the H8 loop and the C-terminal arch of Kap β 2 in the Ran state with the sequence of the H8 loop are shown at bottom. Yellow circles are loop residues that contact the Kap β 2 arch and pink circles are Kap β 2 helices. Red dashed lines indicate intervening loop residues that do not contact the Kap β 2 arch. Blue triangles label residues that contact Ran, pink triangles label residues that contact the Kap β 2 arch, and red circles label Kap β 2 residues that also contact M9NLS. Polar contacts are shown with green lines and hydrophobic contacts with black lines (adapted from [1]). Regions selected for lysine mutations are shown in boxed color (Region A-C) on loop sequence.

Furthermore, to test the electrostatic switch model, it is essential to neutralize the acidic character of the loop. According to this model, the loop should be repelled from the Kap β 2 arch since both the loop and the Kap β 2 surface are highly negatively charged. However if the loop is not acidic, it may occupy the negatively charged substrate binding arch and auto-inhibit Kap β 2 and thus decrease affinity for substrate. Therefore, acidic loop residues were systematically mutated into positively charged lysines.

Materials and Methods

Alanine Loop Mutants

Prevalently acidic H8 loop residues that interact with basic Ran residues were selected for site directed mutagenesis (Figure 5-2). Initially four mutants were made (E332+D334, E363, I364, D366+D367). Residues were mutated to Alanines by site directed mutagenesis. Multiple (E332+D334+E363, E332+D334+I364, E363+I364, I364+D366+D367) site mutants to alanine

residues were performed using the Quickchange method (Stratagene) and all constructs were confirmed by nucleotide sequencing.

PCR mixture used for site directed mutagenesis:

1 μ l Template DNA

1 μ l Primer mixture (Forward and reverse mutation primer to 200ng each)

1 μ l 10mM dNTP mixture

5 μ l 10X Pfu reaction Buffer

1 μ l Pfu Turbo

41 μ l ddH₂O

The PCR cycle used for template amplification:

	95°	30 seconds
18 cycles	{ 95	30 seconds
	{ 55	1 minute
	{ 68	8 minutes
	68	15 minutes

After PCR reaction completion, 1 μ l DpnI enzyme was added to PCR tube directly and the sample was incubated at 37°C for 1 hour to cleave the template DNA. 5 μ l

sample was used to transform into DH10- α cells. Resulting colonies were picked and plasmid amplified for confirming by sequencing.

Alanine mutants of Kap β 2 were purified in small preps. 100ml cultures were grown at 37°C until OD_{600nm}=0.6 and induced at 30°C for 5 hours with 0.5mM IPTG. The cells were harvested and resuspended in lysis buffer containing 50mM Tris pH=7.5, 100mM NaCl, 2mM DTT, 1mM EDTA and 20% Glycerol. The cells were lysed using pulse sonicator. The lysate was incubated with 500 μ l GS beads pre-equilibrated with lysis buffer on the rotator for ~2 hours. The beads bound to Kap β 2 mutants were washed extensively with lysis buffer and stored at 4°C for binding experiments with NLS substrates.

50 μ l bead slurry was incubated with 2 μ l of MBP-M9NLS (4mg/ml) with or without Ran (10 μ l in excess amount) for 10 minutes on ice. The sample was washed 3 times with 500 μ l lysis buffer. The beads and the flow through (F/T) were run on 15% SDS PAGE gel (Figure 5-3 and 5-4).

Lysine Loop Mutants

Regions of acidic patches (Regions A-C in figure 5-2) were selected for site directed mutagenesis into positively charged lysine residues. Kap β 2 Lysine loop mutants spanning regions (See figure 5-2) A, B, C1, C2, D, A+B, A+C1, A+C2, A+C, A+B+C and A+B+C+D were made for further testing for substrate binding

and Ran mediated substrate dissociation. Mutants were made using the same site directed mutagenesis protocol utilized above using the same conditions.

Results and Discussion

The purified Kap β 2 alanine mutants (E332+D334, E363, I364, D366+D367) were tested for qualitative binding to the well-known NLS substrate M9NLS. Binding was not affected by the alanine loop mutations (Figure 5-3). These few mutations did not sufficiently decrease the acidic character of the loop to cause autoinhibition.

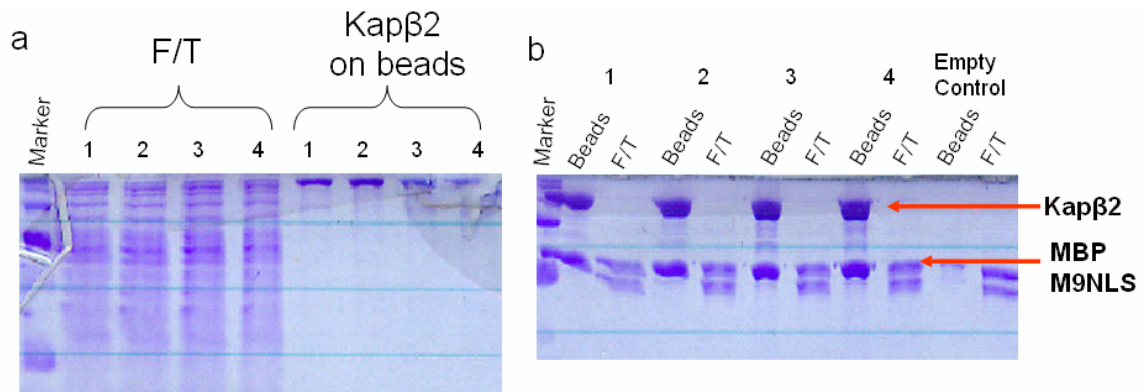


Figure 5-3) Kap β 2 binding experiments: **a)** Kap β 2 was purified using 500 μ l GS beads with extensive washing. Lanes 1-4 denote alanine mutants of Kap β 2 (E332+D334, E363, I364, D366+D367). **b)** Binding experiments performed using Kap β 2 alanine loop mutants: 50 μ l GS beads containing bound Kap β 2 mutants (Empty control contains clean GS beads.) from purification steps were incubated with 2 μ l of MBP-M9NLS (4mg/ml) for 10 minutes on ice, washed 3 times with 500 μ l lysis buffer and the beads and the flow through (F/T) were run on 15% SDS PAGE gel.

Dissociation of the import complex by Ran.GTP was also tested for all of the Kap β 2 mutants. Import complexes were formed on beads using All 4 Kap β 2 mutants (E332+D334, E363, I364, D366+D367) using MBP-M9NLS. The complexes were all successfully dissociated by the addition of Ran.GTP to the beads with bound Kap β 2 and MBP-M9NLS (Figure 5-4).

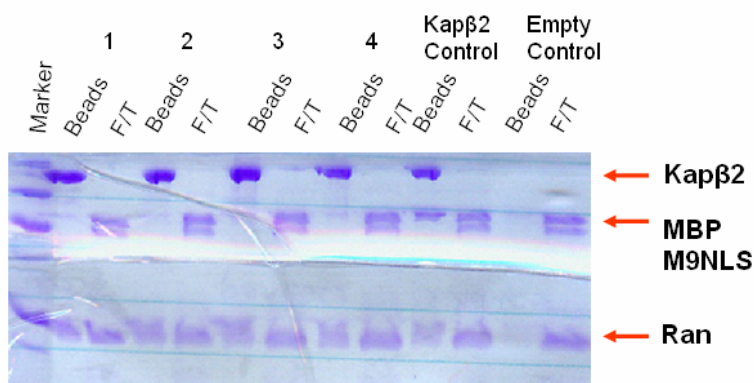


Figure 5-4) Kap β 2.MBP-M9NLS import complexes were formed on beads using All 4 mutants (E332+D334, E363, I364, D366+D367) using MBP-M9NLS. The complexes were all successfully dissociated by the addition of Ran.GTP.

There are 23 acidic residues in the H8 loop. Mutating two or three at a time to alanine will not change its electrostatic character. Therefore, more extensive mutations, especially to lysine residues will be necessary to change loop electrostatics to give further insight into the biophysical mechanism of substrate dissociation.

The H8 loop in Kap β 2 structure is evidently taking part in substrate dissociation with the help of Ran.GTP. However, the negatively charged loop in principle should not allow it to adopt pseudo-substrate conformation that is adopted in

Ran.GTP bound state. Electrostatic properties of NLSs are important for substrate-Kap β 2 association. This situation surely raises the questions about the electrostatic properties of the loop. Is the loop still going to be repelled from the substrate binding region when its acidic residues are mutated into neutral or basic amino acids? Furthermore, in the Ran state, the loop spans the same surface that the NLS binds in an anti-parallel fashion. Therefore, is it possible to have an entirely different class of NLSs binding to the same region of the loop in an opposite orientation? A fragment of the loop devoid of acidic residues or even made basic in overall composition may also be used *in Trans* as a starting point to design a new class of NLS. If such peptide could be found further support will be given to the possibility that the NLS binding interface on Kap β 2 is also able to accommodate sequences that are completely different from the PY-NLSs.

CHAPTER SIX

Biochemical and structural studies on bi-directional nuclear transporter Msn5p

Introduction

The members of the third group of the Karyopherin-family of nuclear transport factors are bidirectional transporters. So far, this group includes Msn5p and importin13 (Mingot, Kostka et al. 2001; Yoshida and Blobel 2001). Msn5p (*multi-copy suppressor of snf1-ts mutation 5*) was initially found to be important in the nuclear export of various proteins(Kaffman, Rank et al. 1998; Blondel, Alepuz et al. 1999; DeVit and Johnston 1999; Boustany and Cyert 2002). However, strong evidence has been discovered supporting the idea that Msn5p is a bidirectional transporter(Yoshida and Blobel 2001). The results strongly suggest that Msn5p takes part in the trafficking of a protein in and different set of proteins out of the nucleus via a non-classical NLS pathway.

In bi-directional transport system, the import cargo protein can only bind to Msn5p in the presence of Ran.GDP in the cytoplasm and is dissociated from Msn5p in the nucleus in the presence of Ran.GTP. On the other hand, the nuclear export cargo can bind to Msn5p in the presence of Ran.GTP bound and is dissociated in the cytosol upon Ran.GTP hydrolysis. Therefore, Msn5p has the distinct function in that it can bind import or export substrates depending upon the nucleotide state of Ran. This pathway would involve both import and export

mechanisms with a double functional player. In this aspect, this protein will bear valuable information for both export and import pathways assuming that a bidirectional transporter uses the same mechanism as the import and/or export pathway.

Nuclear transport of proteins has important implications in various regulated systems. The structure of Msn5p will not only help to understand the critical molecular interactions to support the current model for Ran mediated nuclear transport, but also give insight into the import pathway previously examined through the structures of Karyopherins $\beta 1$ and $\beta 2$. Msn5p would provide invaluable structural information to investigate both export and import pathways. The structure could help to examine the similarities or differences between the previously determined import factors and in comparison, the structural characteristics that are required to perform an export function as well as an import function.

The recognition of various cargo proteins, rather specifically by a single protein, is another aspect of the study. Characterization of the interactions of Msn5p with various import and export cargo proteins could give information about the degeneracy in binding to the protein itself. The recognition specificity is also important since different Karyopherins recognize different proteins.

Materials and Methods

Expression and Purification of S.Cerevisiae Msn5p Native and Se-met Proteins

Full length Msn5p cloned into pQE60 vector was obtained from the O'Shea Lab. Msn5p was expressed as a C-terminal his-tagged protein in BL21 (Codonplus) cells. The induction was done at 20°C for 20hrs using 0.5mM IPTG. The cells were lysed in buffer containing: 50mM Tris pH=7.5, 100mM NaCl, 2mM β -ME and 10% Glycerol using two passes from the cell disrupter. The sample was centrifuged at 12K rpm for 50 minutes. The supernatant was loaded onto His-Trap column using peristaltic pump.

The protein was eluted with elution buffer with increasing concentrations (2mM-500mM) of Imidazole pH=7.5 with 100mM NaCl 10% Glycerol and 5mM DTT (Figure 6-1). The NaCl in the final elution sample was increased to >350mM and loaded onto Phenyl-HP HIC column (Buffered with Tris, pH=7.5). Decreasing NaCl concentration was used to elute the protein from the Phenyl-HP column.

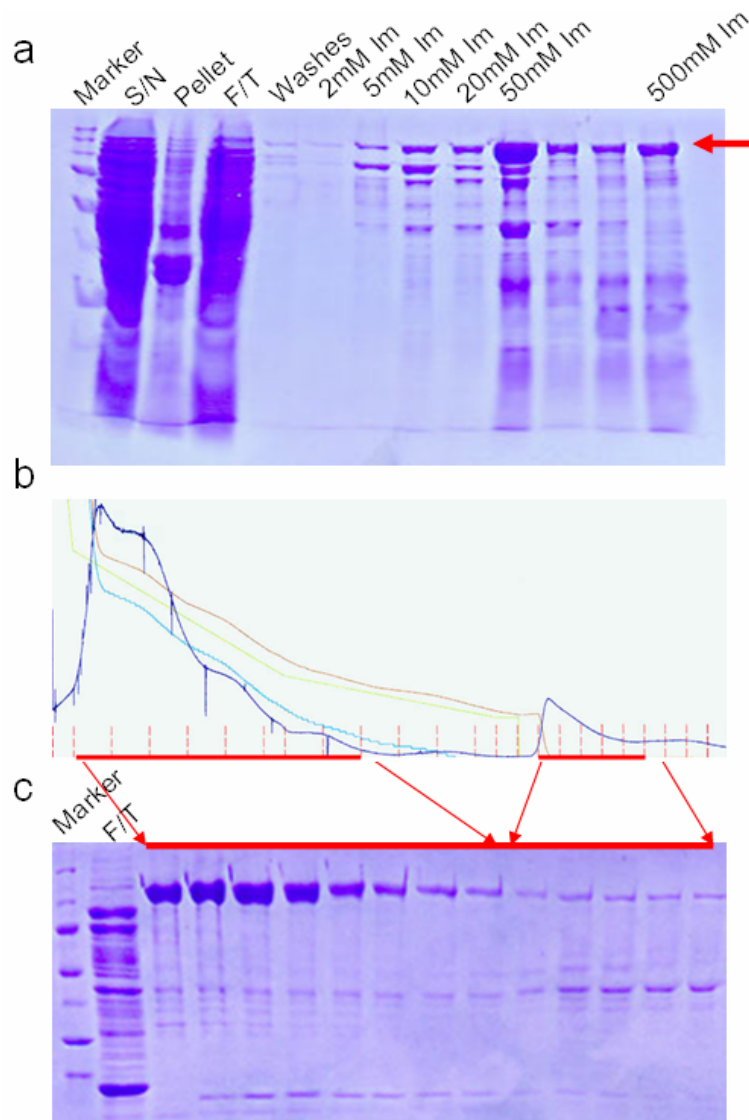


Figure 6-1) Purification of Msn5p: **a)** Lysed sample is put through His-Trap column and after extensive washing, eluted with Imidazole gradient (2mM-500mM). **b)** Elutions containing the protein are pooled and put through Phenyl-HP column. **c)** Fractions containing the protein are run on 10% SDS PAGE gel.

The appropriate fractions containing the protein were pooled together and put through a final gel filtration column using Superdex s200 column (Figure 6-2). The final sample was concentrated beyond 25mg/ml for crystallization. The final sample was subjected to N-terminal sequencing analysis for confirmation of the correct sample. GST-tagged Msn5p was also cloned into pGEX-TEV vector and purified using the same protocol used for purification of Kap β 2 from Chapter 3 using 10% glycerol instead of 20% glycerol in the buffers. However the expression and the purity of the his-tagged protein was significantly better than the GST-tagged protein.

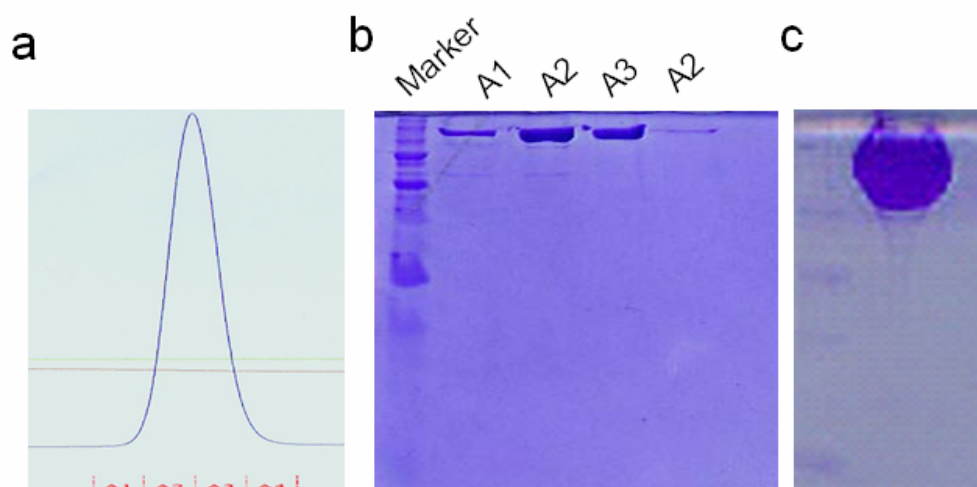


Figure 6-2) Purification of Msn5p: **a)** The appropriate fractions containing the protein were pooled together and put through a final gel filtration column using Superdex s200 column **b)** Fractions containing the protein are run on a 10% SDS PAGE gel. **c)** Final Msn5p sample was concentrated beyond 25mg/ml for crystallization (1 μ l loaded on gel)

Se-met Msn5p protein was expressed in M9 minimal media as described in (Lee, Cansizoglu et al. 2006) and purified following the same protocol used for the native protein with the addition of 10mM DTT to all purification buffers. The yield of the Se-met protein was significantly lower than the native protein.

Crystallization of Msn5p

Available crystal screens were tested on the protein in hanging and sitting drop techniques with varying drop: reservoir volume ratios. The protein was crystallized in the condition containing 2.0M Ammonium Sulfate 2% PEG400 and 0.1M HEPES pH=7.5. In further trials, Ammonium Sulfate was substituted into Sodium Malonate. Alternative crystallization condition was: 1.8M NaMalonate 3-4% PEG400 0.1M MES pH=6.0 (Figure 6-3).

Further candidate crystallization conditions were identified from Wizard-Emerald screen:

#11: 0.1M Imidazole pH=8.0 0.2M CaAc 20%PEG1000

#13: 0.1M NaCacodylate pH=6.5 1.0M NaCitrate *

#28: 0.1M CAPS pH=10.5 0.2M Li2SO4 2.0MAS

#31: 0.1M Imidazole pH=8.0 1.0 M NaCitrate

These conditions however did not improve the size and shape of the crystals. The crystals produced in these conditions were much smaller (~ 10 microns).

Expression and Purification of the Substrates

There are several candidate cargo proteins which can be used for crystallization purposes for Msn5p (see table 4-3). We have selected the most well studied export cargo Pho4 and the only import substrate, Replication Protein A (RPA) for crystallization purposes. RPA is a single strand DNA binding protein composed of three subunits of 70, 32 and 14KDa. RPA trimer can be expressed as a polycistronic vector expressing all three subunits at the same time with stoichiometric amounts(Henricksen, Umbricht et al. 1994), as well as in separate vectors(Bochkareva, Korolev et al. 2002).

RPA was expressed and purified from polycistronic p11d-tsc-RPA construct (Obtained from Wold Lab at Univ. of Iowa) expressing all three subunits. (RPA was found to be toxic to the bacteria.) The sample was grown overnight without shaking in 100ml LB at 37°C and diluted to 1L the next day. The cells were grown at 37°C until $OD_{600nm} = 0.8$ and induced with 0.3mM IPTG for 2 hours. The cells were harvested and washed with 150mM ice cold NaCl and resuspended

in lysis buffer containing HI buffer (30mM HEPES pH=7.8, 0.25mM EDTA, 0.5% (w/v) Inositol, 0.01% T-20) and lysed by 3 passes from the cell disrupter.

The lysate was put through BIO-RAD blue column (5ml) and washed using sequential washes of 3 column volumes (15mls) of HI +80mM KCl, HI +800mM KCl and 0.5M NaSCN. The protein was eluted using 1.5M NaSCN (Figure 6-3).

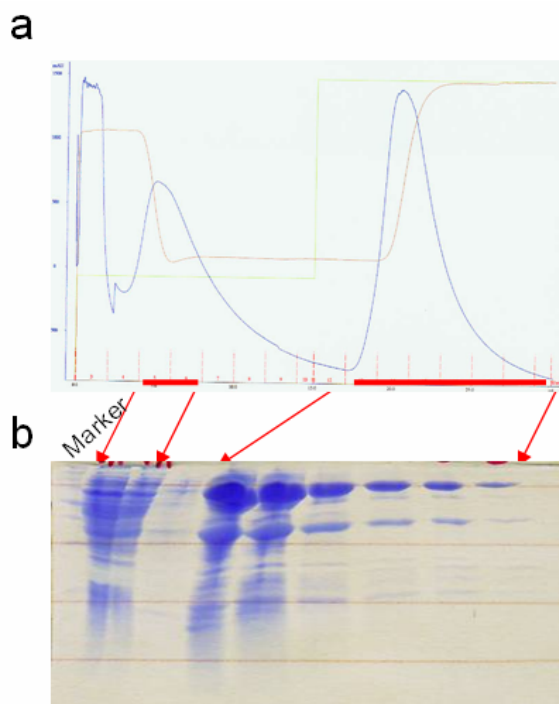


Figure 6-3) RPA purification 1: a) The protein is loaded onto BIO-RAD Blue column and washed extensively with HI buffer +KCl. The 3 peaks on the UV trace correspond to HI buffer +80mM KCl, HI buffer +800mM KCl (First 2 lanes on the gel) and the elution step with 1.5M NaSCN at the last step (last 8 fractions). **b)** Peak fractions are run on 10% SDS PAGE gel. The last peak corresponds to the trimeric RPA complex.

Fractions containing the RPA complex are then pooled and put through a final superdex s200 gel filtration column (Figure 6-4)

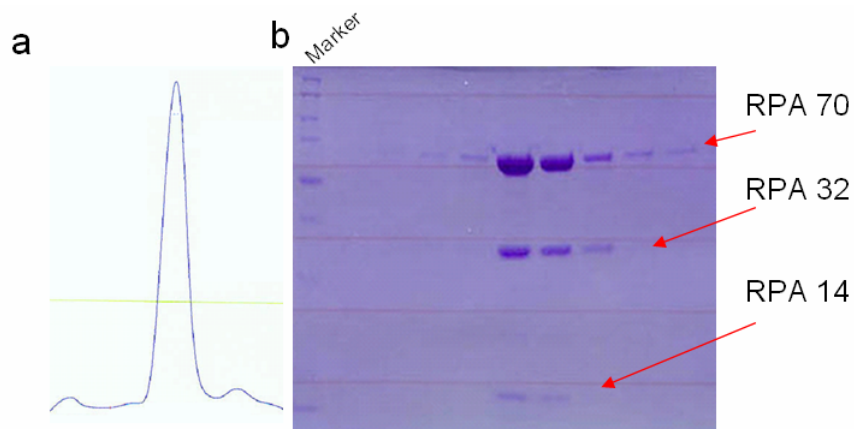


Figure 6-4) RPA purification 2: **a)** The RPA complex is put through a final superdex s200 gel filtration column run. **b)** The fractions containing the final trimeric complex are run on 10% SDS PAGE gel.

There are several possible candidates for the export complex, including transcription factors Pho4, Crz1, and Mig1. These proteins are all transcription factors in yeast. We selected Pho4 for crystallization purposes because it is one of the most well studied transcription factors. Pho4 is a member of phosphate

response pathway in yeast. It is regulated by phosphorylation by the cyclin/CDK complex Pho80/Pho85(O'Neill, Kaffman et al. 1996). Various functional domains have been mapped in the overall structure of the protein(O'Neill, Kaffman et al. 1996; Komeili and O'Shea 1999) (Figure 6-5). For phosphorylation of Pho4 and fragments, Pho80 and Pho85 were PCR amplified and cloned from the yeast genome using standard cloning techniques into pGEX-TEV vector for expression.

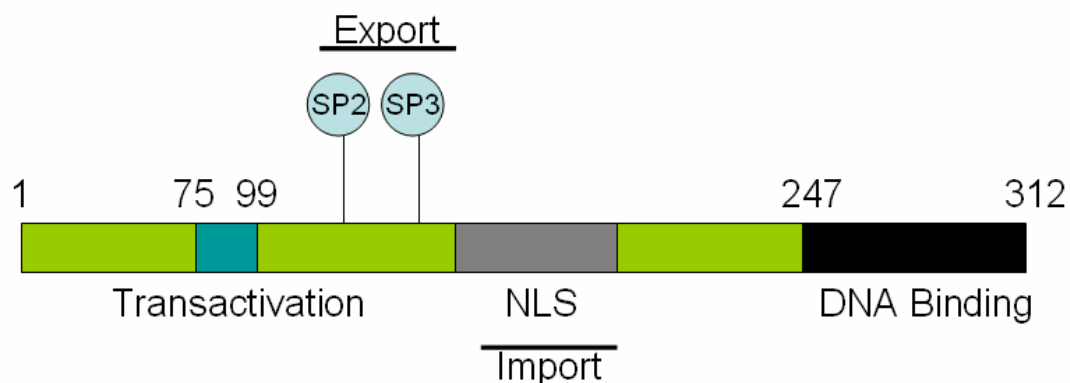


Figure 6-5) Overall functional domain structure of Pho4 (Adapted from(Komeili and O'Shea 1999)) SP2 and SP3 denotes phosphorylation sites for recognition by Msn5p for export purposes.

Full length, 1-200, 1-170 and 1-140 fragments of Pho4 were cloned into pGEX-Tev vector in E.coli using classical PCR and cloning techniques. The protein was expressed in BL21 (DE3). The cells were grown at 37°C until $OD_{600nm} = 0.6$ induced with 0.5mM IPTG, continued for 3 hours and harvested. Cells were lysed in lysis buffer containing 50mM Tris pH=7.5, 100mM NaCl, 1mM EDTA and 2mM DTT. Cell lysate was put through GST beads after lysing the cells using 2 passes from the cell disrupter. The protein was be eluted from the beads using lysis buffer containing 20mM glutathione (pH=8.1). The eluted fractions were concentrated and cleaved from the GST tag using TEV protease. TEV protease was added to the concentrated elution sample and the protein was incubated at room temperature overnight. The protein was put through HiTrap Q column at the final step. Purification of Pho4 and fragments failed due to severe proteolytic degradation problems of full length and the fragments of the protein. The degradation was probably occurring at the overnight incubation step with TEV protease.

Complex formation with RPA

Msn5p and RPA final purified samples were mixed together at varying concentrations however the hetero-tetrameric Msn5p.RPA complex was not successfully purified from the single protein peaks. The individual protein peaks were not distinguishable at the gel filtration (superdex s200) runs. The resolution limit of the s200 was not sufficient to separate a complex (adding up to 250KDa) from the non-complex samples. However, it is also possible that the complex did not form since most of the peaks overlap.

Results and Discussions

Msn5p protein was crystallized in its unliganded form. The crystals diffracted to 6.5 - 7Å at home source and to 4.0Å at the synchrotron. Limited diffraction data was collected for indexing purposes however the crystals decayed rapidly. The crystals were P6 with unit cell parameters: $a=212.4\text{\AA}$ $b=212.4\text{\AA}$ $c=303.5\text{\AA}$ (Figure 6-6). Additive screens yielded ZnCl_2 as a possible condition. Although the appearance of the crystals improved slightly, they did not improve in terms of resolution.

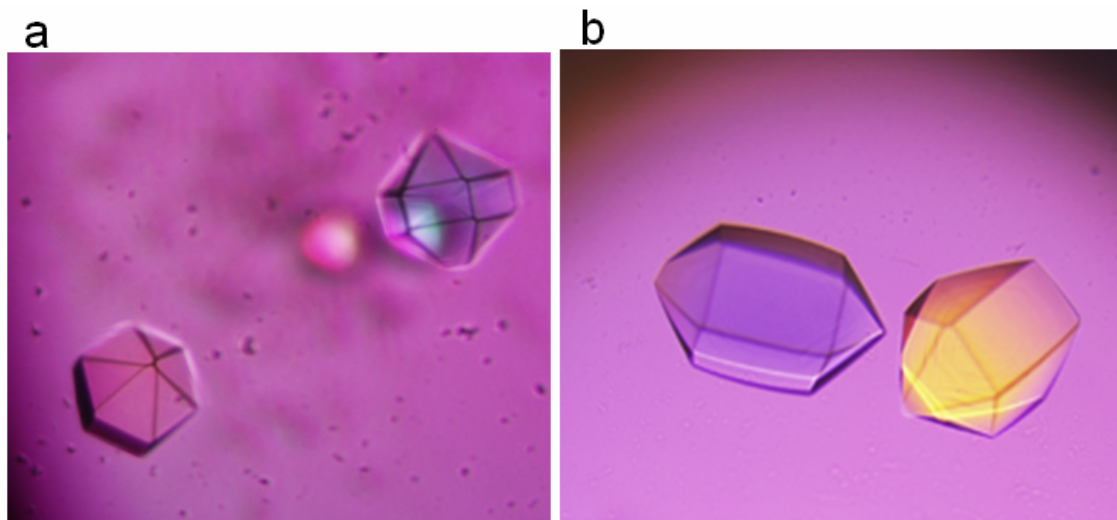


Figure 6-6) Msn5p crystals in: **a)** Ammonium Sulfate 2% PEG400 and 0.1M HEPES pH=7.5 **b)** 1.8M NaMalonate 3-4% PEG400 0.1M MES pH=6.0

The Msn5p.RPA complex trials did not give convincing results. It is possible that the complex was not purified consistently due to resolution limits of the gel filtration columns used. Another possibility is the condition of the RPA trimeric complex. The purification protocol for RPA contains harsh conditions to elute the protein from the column (see materials and methods section). It is possible that the high concentration of NaSCN utilized during the purification steps results in unfolding of the RPA complex, thus preventing to form appropriate complex and

crystals due to failure of re-folding of the RPA complex properly. It is however interesting to note that the final RPA sample runs at the appropriate size range for a ~120KDa protein complex (70+32+14KDa) on the gel filtration column (Figure 6-4). It is possible that not the entire protein complex is folded properly forming a non-homogeneous final sample in terms of tertiary structure.

Another unique feature of Msn5p is that all of the export substrates are recognized in phosphorylated forms. Thus, one challenge in terms of forming and crystallizing the export complex will be the phosphorylation of the export substrate. It is not clear if the phosphorylation exposes a buried NES sequence or the phosphorylated fragment itself is the actual NES. Therefore, the specificity determinants for the binding may or may not be located around the phosphorylated residues. Each cargo substrate should have a minimal binding sequence to Msn5p. However, export substrates for non-classical NLS pathways are very diverse. So far, no recognizable consensus NLS signal has been successfully determined in protein alignments.

BIBLIOGRAPHY

- (1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallogr D Biol Crystallogr **50**(Pt 5): 760-3.
- Adam, S. A., R. S. Marr, et al. (1990). "Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors." J Cell Biol **111**(3): 807-16.
- Akey, C. W. and D. S. Goldfarb (1989). "Protein import through the nuclear pore complex is a multistep process." J Cell Biol **109**(3): 971-82.
- Akey, C. W. and M. Radermacher (1993). "Architecture of the *Xenopus* nuclear pore complex revealed by three-dimensional cryo-electron microscopy." J Cell Biol **122**(1): 1-19.
- Allemand, E., S. Guil, et al. (2005). "Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock." Proc Natl Acad Sci U S A **102**(10): 3605-10.
- 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): 50597-606.
- Bednenko, J., G. Cingolani, et al. (2003). "Nucleocytoplasmic transport: navigating the channel." Traffic **4**(3): 127-35.
- 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): 2587-91.
- Bischoff, F. R. and H. Ponstingl (1991). "Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1." Nature **354**(6348): 80-2.
- Blondel, M., P. M. Alepuz, et al. (1999). "Nuclear export of Far1p in response to pheromones requires the export receptor Msn5p/Ste21p." Genes Dev **13**(17): 2284-300.
- Bochkareva, E., S. Korolev, et al. (2002). "Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA." Embo J **21**(7): 1855-63.
- Bogerd, H. P., R. E. Benson, et al. (1999). "Definition of a consensus transportin-specific nucleocytoplasmic transport signal." J Biol Chem **274**(14): 9771-7.
- 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): 5055-60.

- Boustany, L. M. and M. S. Cyert (2002). "Calcineurin-dependent regulation of Crz1p nuclear export requires Msn5p and a conserved calcineurin docking site." Genes Dev **16**(5): 608-19.
- Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54**(Pt 5): 905-21.
- Brunger, A. T., A. P. D., et al. (1998). "Crystallography & NMR System: A new software suite for macromolecular structure determination." Acta Cryst. A **D54**: 905-921.
- 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): 452-4.
- Carugo, O. and P. Argos (1997). "Protein-protein crystal-packing contacts." Protein Sci **6**(10): 2261-3.
- CCP4 (1994). "The CCP4 suite: programs for X-ray crystallography." Acta Crystallogr. D **50**(760-763).
- Chook, Y. and G. Blobel (1999). "Structure of the nuclear transport complex karyopherin-beta2-Ran.GppNHp." Nature **399**: 230-237.
- Chook, Y. M. and G. Blobel (1999). "Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp." Nature **399**(6733): 230-7.
- Chook, Y. M. and G. Blobel (2001). "Karyopherins and nuclear import." Curr Opin Struct Biol **11**(6): 703-15.
- Chook, Y. M., A. Jung, et al. (2002). "Uncoupling Kap β 2 substrate dissociation and Ran binding." Biochemistry **41**: 6955-6966.
- Chook, Y. M., A. Jung, et al. (2002). "Uncoupling Kap β 2 substrate dissociation and ran binding." Biochemistry **41**(22): 6955-66.
- 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): 1345-53.
- Cingolani, G., C. Petosa, et al. (1999). "Structure of importin-beta bound to the IBB domain of importin-alpha." Nature **399**(6733): 221-9.
- Conti, E., C. W. Muller, et al. (2006). "Karyopherin flexibility in nucleocytoplasmic transport." Curr Opin Struct Biol **16**(2): 237-44.
- 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., 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): 355-67.
- Dalton, J. A., I. Michalopoulos, et al. (2003). "Calculation of helix packing angles in protein structures." Bioinformatics **19**(10): 1298-9.

- 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): 439-46.
- DeLano, W. L. (2002). Pymol. (DeLano Scientific, San Carlos, CA).
- DeVit, M. J. and M. Johnston (1999). "The nuclear exportin Msn5 is required for nuclear export of the Mig1 glucose repressor of *Saccharomyces cerevisiae*." *Curr Biol* **9**(21): 1231-41.
- Dingwall, C. and R. A. Laskey (1991). "Nuclear targeting sequences: a consensus?" *Trends Biol. Sci.* **16**: 178-181.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2126-32.
- Enarson, P., J. B. Rattner, et al. (2004). "Autoantigens of the nuclear pore complex." *J Mol Med* **82**(7): 423-33.
- 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): 15293-8.
- Feldherr, C. M., R. J. Cohen, et al. (1983). "Evidence for mediated protein uptake by amphibian oocyte nuclei." *J Cell Biol* **96**(5): 1486-90.
- 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): 1325-31.
- Fukuhara, N., E. Fernandez, et al. (2004). "Conformational variability of nucleocytoplasmic transport factors." *J Biol Chem* **279**(3): 2176-81.
- Fukuhara, N., E. Fernandez, et al. (2004). "Conformational variability of nucleocytoplasmic transport factors." *J Biol Chem* **279**(32): 176-81.
- Gattoni, R., D. Mahe, et al. (1996). "The human hnRNP-M proteins: structure and relation with early heat shock-induced splicing arrest and chromosome mapping." *Nucleic Acids Res* **24**(13): 2535-42.
- Gorlich, D., M. Dabrowski, et al. (1997). "A novel class of RanGTP binding proteins." *J Cell Biol* **138**(1): 65-80.
- Gorlich, D. and U. Kutay (1999). "Transport between the cell nucleus and the cytoplasm." *Annu Rev Cell Dev Biol* **15**: 607-60.
- Gorlich, D., M. J. Seewald, et al. (2003). "Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation." *Embo J* **22**(5): 1088-100.
- Guttinger, S., P. Muhlhauser, et al. (2004). "Transportin2 functions as importin and mediates nuclear import of HuR." *Proc Natl Acad Sci U S A* **101**(9): 2918-23.
- Hamamoto, T., S. Gunji, et al. (1983). "Leptomycins A and B, new antifungal antibiotics. I. Taxonomy of the producing strain and their fermentation, purification and characterization." *J Antibiot (Tokyo)* **36**(6): 639-45.

- Hase, M. E., P. Yalamanchili, et al. (2006). "The drosophila hnRNP M protein, HRP59, regulates alternative splicing and controls the production of its own mRNA." J Biol Chem.
- Hayward, S. and H. J. Berendsen (1998). "Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme." Proteins **30**(2): 144-54.
- 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): 3166-73.
- Henricksen, L. A., C. B. Umbricht, et al. (1994). "Recombinant replication protein A: expression, complex formation, and functional characterization." J Biol Chem **269**(15): 11121-32.
- Hinshaw, J. E., B. O. Carragher, et al. (1992). "Architecture and design of the nuclear pore complex." Cell **69**(7): 1133-41.
- 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): 1365-70.
- Jones, T. A., S. W. Cowan, et al. (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta Crystallogr. A **47**: 110-119.
- Kaffman, A., N. M. Rank, et al. (1998). "The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus." Nature **396**(6710): 482-6.
- 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.
- 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): 2732-9.
- Kobe, B. (1999). "Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha." Nat Struct Biol **6**(4): 388-97.
- Komeili, A. and E. K. O'Shea (1999). "Roles of phosphorylation sites in regulating activity of the transcription factor Pho4." Science **284**(5416): 977-80.
- Lee, B. J., A. E. Cansizoglu, et al. (2006). "Rules for nuclear localization sequence recognition by karyopherin beta 2." Cell **126**(3): 543-58.

- Lee, S. J., N. Imamoto, et al. (2000). "The adoption of a twisted structure of importin-beta is essential for the protein-protein interaction required for nuclear transport." *J Mol Biol* **302**(1): 251-64.
- Lee, S. J., Y. Matsuura, et al. (2005). "Structural basis for nuclear import complex dissociation by RanGTP." *Nature* **435**(7042): 693-6.
- Lee, S. J., Y. Matsuura, et al. (2005). "Structural basis for nuclear import complex dissociation by RanGTP." *Nature* **435**(7042): 693-6.
- 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): 1571-5.
- Linding, R., L. J. Jensen, et al. (2003). "Protein disorder prediction: implications for structural proteomics." *Structure (Camb)* **11**(11): 1453-9.
- 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): 515-25.
- Matsuura, Y. and M. Stewart (2004). "Structural basis for the assembly of a nuclear export complex." *Nature* **432**(7019): 872-7.
- Mattaj, I. W. and L. Englmeier (1998). "Nucleocytoplasmic transport: the soluble phase." *Annu Rev Biochem* **67**: 265-306.
- Matunis, M. J., E. L. Matunis, et al. (1992). "Isolation of hnRNP complexes from *Drosophila melanogaster*." *J Cell Biol* **116**(2): 245-55.
- McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2005). "Likelihood-enhanced fast translation functions." *Acta Cryst. D* **61**: 458-464.
- McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2005). "Likelihood-enhanced fast translation functions." *Acta Crystallogr D Biol Crystallogr* **61**(Pt 4): 458-64.
- Mingot, J. M., S. Kostka, et al. (2001). "Importin 13: a novel mediator of nuclear import and export." *Embo J* **20**(14): 3685-94.
- Mosammaparast, N. and L. F. Pemberton (2004). "Karyopherins: from nuclear-transport mediators to nuclear-function regulators." *Trends Cell Biol.* **14**(10): 547-56.
- Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." *Acta Crystallogr D Biol Crystallogr* **53**(Pt 3): 240-55.
- Nakielnny, S., M. C. Siomi, et al. (1996). "Transportin: nuclear transport receptor of a novel nuclear protein import pathway." *Exp Cell Res* **229**(2): 261-6.
- Nicholls, A., K. A. Sharp, et al. (1991). "Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons." *Proteins: Struct. Funct. Genet.* **11**: 281-296.

- O'Neill, E. M., A. Kaffman, et al. (1996). "Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex." Science **271**(5246): 209-12.
- Otwinowski, Z. and W. Minor (1997). "Processing of X-ray Diffraction Data Collected in Oscillation Mode." Methods in Enzymology **276**: 307-326.
- Painter, J. and E. A. Merritt (2006). "Optimal description of a protein structure in terms of multiple groups undergoing TLS motion." Acta Crystallogr D Biol Crystallogr **62**(Pt 4): 439-50.
- Pemberton, L. F. and B. M. Paschal (2005). "Mechanisms of receptor-mediated nuclear import and nuclear export." Traffic **6**(3): 187-98.
- 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): 761-75.
- Pollard, V. W., W. M. Michael, et al. (1996). "A novel receptor-mediated nuclear protein import pathway." Cell **86**(6): 985-94.
- 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): 1769-73.
- Rebane, A., A. Aab, et al. (2004). "Transportins 1 and 2 are redundant nuclear import factors for hnRNP A1 and HuR." Rna **10**(4): 590-9.
- 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): 883-94.
- Ribbeck, K., G. Lipowsky, et al. (1998). "NTF2 mediates nuclear import of Ran." Embo J **17**(22): 6587-98.
- 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): 615-23.
- Rout, M. P. and G. Blobel (1993). "Isolation of the yeast nuclear pore complex." J Cell Biol **123**(4): 771-83.
- Sheldrick, G. M., Z. Dauter, et al. (1993). "The application of direct methods and Patterson interpretation to high-resolution native protein data." Acta Crystallogr. D **49**: 18 - 23.
- Siomi, H. and G. Dreyfuss (1995). "A nuclear localization domain in the hnRNP A1 protein." J Cell Biol **129**(3): 551-60.
- Siomi, M. C., P. S. Eder, et al. (1997). "Transportin-mediated nuclear import of heterogeneous nuclear RNP proteins." J Cell Biol **138**(6): 1181-92.
- 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): 1857-66.

- 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.
- 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): 3975-87.
- 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): 32167-71.
- 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): 32167-71.
- Vetter, I. R., A. Arndt, et al. (1999). "Structural view of the Ran-Importin beta interaction at 2.3 Å resolution." Cell **97**(5): 635-46.
- Vetter, I. R., A. Arndt, et al. (1999). "Structural view of the Ran-Importin beta interaction at 2.3 Å resolution." Cell **97**(5): 635-46.
- 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): 223-34.
- Yashiroda, Y. and M. Yoshida (2003). "Nucleo-cytoplasmic transport of proteins as a target for therapeutic drugs." Curr Med Chem **10**(9): 741-8.
- Yoshida, K. and G. Blobel (2001). "The karyopherin Kap142p/Msn5p mediates nuclear import and nuclear export of different cargo proteins." J Cell Biol **152**(4): 729-40.

VITAE

Ahmet Ertugrul Cansizoglu was born in Eskisehir, Turkey, on February 22nd, 1979, the son of Fatma Cansizoglu and Veli Cansizoglu. He received the degree of Bachelor of Science with a major in molecular biology and genetics from Bilkent University in Ankara, Turkey, in June, 2001. In June, 2001 he entered the Graduate School of Division of Cell and Molecular Biology at the University of Texas Southwestern Medical Center at Dallas. He continued his research at the Department of Pharmacology in the laboratory of Dr. Yuh Min Chook. In July, 2007 he married Esra Cansizoglu. He was awarded the degree of Doctor of Philosophy in Molecular Biophysics in September, 2007.

Permanent Address: Gundogdu Mh. Filizer Sk. Fatih Sitesi No: 6/7
26100, Eskisehir, TURKEY