

**Dissertation**

for the award of the degree

"Doctor rerum naturalium"

Division of Mathematics and Natural Sciences  
of the Georg-August-Universität Göttingen

**Structural and Functional Analysis of  
Exportin-Cargo Recognition**

submitted by

**Thomas Güttler**

from Löbau (Germany)

Göttingen, August 2010

## **Members of the Thesis Committee**

Prof. Dr. Dirk Görlich  
(thesis mentor and referee)      Max-Planck-Institut für biophysikalische Chemie  
Abteilung Zelluläre Logistik  
Göttingen

Prof. Dr. Ralf Ficner  
(co-referee)      Georg-August-Universität Göttingen  
Institut für Mikrobiologie und Genetik  
Abteilung für Molekulare Strukturbioogie

Prof. Dr. Reinhard Lührmann      Max-Planck-Institut für biophysikalische Chemie  
Abteilung Zelluläre Biochemie  
Göttingen

## **Additional Members of the Examination Board**

Prof. Dr. Detlef Doenecke      Georg-August-Universität Göttingen  
Abteilung Biochemie I

PD Dr. Ralph Kehlenbach      Georg-August-Universität Göttingen  
Abteilung Biochemie I

Prof. Dr. Kai Tittmann      Georg-August-Universität Göttingen  
Albrecht-von-Haller-Institut für Pflanzenwissenschaften  
Abteilung Bioanalytik

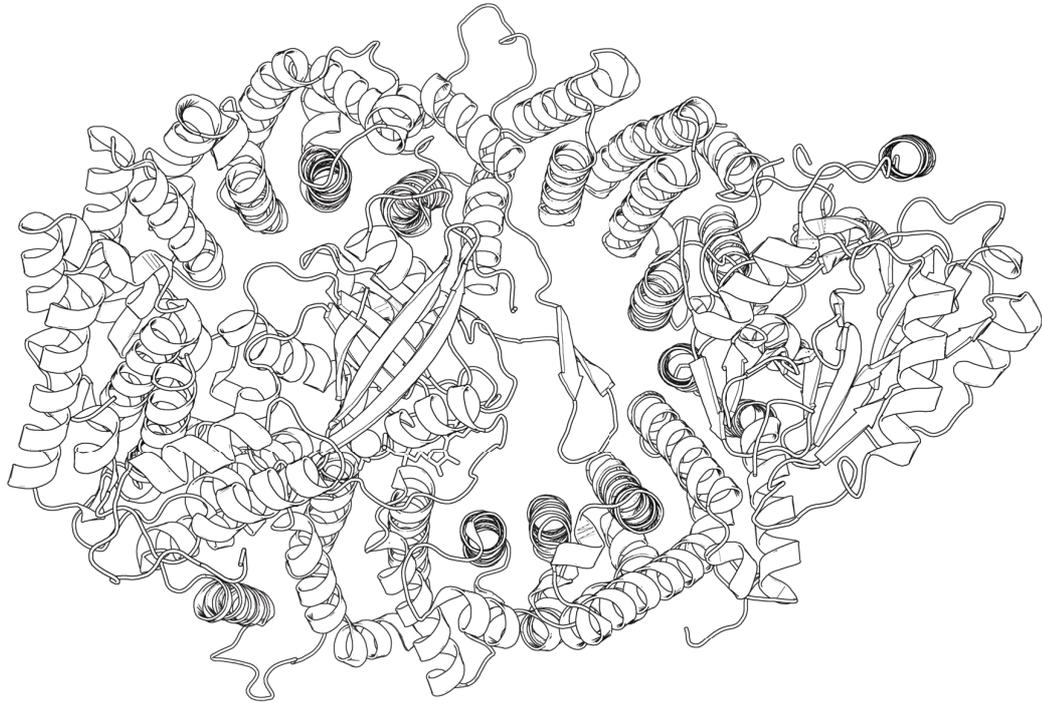
**Date of the disputation:**      September 17, 2010

The research detailed in this thesis has been performed in the laboratory of Prof. Dirk Görlich at the Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH, in the period from January 2004 to April 2007), and at the Max Planck Institute for Biophysical Chemistry (in the time from April 2007 to August 2010). Large parts of the presented work were done in collaboration, as indicated.

I hereby declare that I completed my thesis entitled  
**"Structural and Functional Analysis of Exportin-Cargo Recognition"**  
independently and with no other sources and aids than quoted.

Thomas Güttler  
Göttingen, August 2010

**To my parents and grandparents**



*"Alles, was die reichste Phantasie und höchste Geistesbildung zu ersinnen vermag, erscheint, gegen die Wirklichkeit gehalten, wie eine bunte, schillernde Seifenblase."*

*Justus von Liebig*

## Acknowledgements

First of all, I would like to thank my mentor, Dirk Görlich, for his input to this work, his contagious enthusiasm and for many inspiring discussions. Thank you also for everything you taught me about doing science, and for all your support. The time in your lab has been very enjoyable.

Furthermore, I would like to thank Prof. Dirk Görlich, Prof. Ralf Ficner and Prof. Reinhard Lührmann for serving on my thesis committee. I am grateful to my committee members and to Prof. Detlef Doenecke, Dr. Ralph Kehlenbach and Prof. Kai Tittmann for reviewing my thesis. I would also like to thank Dr. Ralf Tolle for coordinating the ZMBH doctoral program. Much appreciated organizational support for finishing the work summarized in this thesis came from the graduate school in Göttingen (GGNB), especially from Dr. Steffen Burkhardt, Kirsten Pöhlker and Elisa Reckmann-Heinrich. Thank you!

Many collaborators contributed tremendously to the projects presented on the following pages. I thank Ralf Ficner, Thomas Monecke, Piotr Neumann and Achim Dickmanns for our very productive collaboration on CRM1. Special thanks go to Piotr, whose contribution was essential for the success of this work. I also thank Thomas and Piotr for a great time in the lab, at synchrotrons and in front of computer screens. Thanks to all of you for the many things I learned about crystallography. I thank Michael Sattler, Tobias Madl and Lorenzo Corsini for our terrific collaboration on the CRM1 NMR project. Many thanks go to Cristina Gontan, Raymond Poot and Robbert Rottier for finding that Exportin 4 is strange and for starting a very exciting joint venture. ("Finally, the cows have come home.") I thank Oliver Hantschel, Michael Sattler and Giulio Superti-Furga for our collaboration to resolve the mysteries of the Abl C-terminal domain, which inspired me to focus on the cell's most exciting export receptor. I also wish to thank Thomas Ruppert, Armin Bosserhoff, Henning Urlaub, Uwe Pleßmann and Monika Raabe for mass spectrometry.

Sincere thanks are given to everyone in the department for making the time of my doctoral studies happy, instructive and productive. Thank you for the many fruitful seminars. Special thanks go to Steffen, Matthias, Bastian, Fabrizia and Volker for numerous inspiring conversations "in passing" and to Cathrin for her help. I also thank Markus, José and Theis for the "good old days". I am most grateful to Danilo for his invaluable assistance during all of our projects, including the famous elevator concerts. I wish to thank Petra, Heike, Uschi, Heinz-Jürgen, Jürgen and Jens for juggling plasmids, proteins, cells and big machines. Many

thanks go to Gaby and Jürgen for all the gels, and to Elisabeth, Gaby and Uwe for driving the cycle of lab glassware. I also wish to thank Connie for all her help.

I thank Markus Wahl, Vladimir Pena and Jürgen Wawrzinek for setting up the crystallization facility at the MPI and the members of the Lührmann and Wahl labs (especially Homa Ghalei, Alexey Rozov and Christian Stegmann) for help and advice. I am grateful to the instructors of the 2008 Cold Spring Harbor workshop on "X-ray Methods in Structural Biology" for facilitating this work. Raphael, Winfried, Holger and Matthias should be mentioned for their great IT support. Your help is much appreciated! Many thanks also go to the very supportive MPI and ZMBH institute services.

Last but not least I thank Dirk, Cathrin and Lena for helpful comments on this thesis.

A big "thank you" to many friends who made the time of my doctoral studies special. Particular thanks go to Cathrin for her support during the various "countdowns" of this work (including innumerable incidents of "Schadfraß"). I thank Markus for all our ornithological excursions (including those to local restaurants), Danilo & Indro for our running events, Oliver for the hiking trips, and Anja & Ania for the music. Many thanks also go to Sandra, Cathrin, Aksana, Sonja, Jörg, Lena, Ufuk, Indro and Michael for many nice evenings outside the lab.

Most importantly, I am grateful to my family, especially my parents and grandparents, for all their curiosity and support. I also thank Sebastian for always being inspiring, critical and helpful.

This work has been generously supported by the Max-Planck-Gesellschaft, the Boehringer Ingelheim Fonds and the Alfried Krupp Foundation. Special thanks go to the B.I.F. for all the interesting seminars and great meetings.

# Table of Contents

<b>List of abbreviations used in this work .....</b>	<b>1</b>
<b>Summary .....</b>	<b>3</b>
<b>CHAPTER 1.....</b>	<b>5</b>
<b>Introduction to nucleocytoplasmic transport .....</b>	<b>5</b>
1.1 Compartmentation of eukaryotic cells.....	5
1.2 Overview of nucleocytoplasmic exchange .....	6
1.3 The structure of Importin $\beta$ -like nuclear transport receptors.....	15
1.4 About this work .....	23
<b>CHAPTER 2.....</b>	<b>26</b>
<b>Exportin 4 mediates a novel nuclear import pathway .....</b>	<b>26</b>
<b>CHAPTER 3.....</b>	<b>38</b>
<b>Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with     Snurportin1 and RanGTP .....</b>	<b>38</b>
3.1 Supplementary Figures and Tables.....	44
3.2 Materials and Methods .....	53
<b>CHAPTER 4.....</b>	<b>56</b>
<b>The NES consensus redefined by structures of PKI-type and Rev-type nuclear export     signals bound to CRM1 .....</b>	<b>56</b>
4.1 Abstract.....	57
4.2 Introduction.....	57
4.3 Results.....	60
4.4 Discussion.....	76
4.5 Supplementary Figures and Tables.....	80
4.6 Methods .....	87
4.7 Accession codes.....	94
4.8 Acknowledgements to Chapter 4.....	94

<b>CHAPTER 5</b> .....	<b>96</b>
<b>Concluding Remarks and Perspectives</b> .....	<b>96</b>
5.1 How does Ran promote export cargo loading?.....	96
5.2 Dissociation of export cargoes from CRM1 .....	99
5.3 Future perspectives .....	102
<b>CHAPTER 6</b> .....	<b>104</b>
<b>Appendix</b> .....	<b>104</b>
6.1 Establishment of a cost-efficient <i>E. coli</i> culture medium for the perdeuteration of CRM1 .....	105
6.2 Perdeuteration of CRM1 .....	114
6.3 Conclusions.....	116
6.4 Isotope labeling of the PKI $\Phi^0$ Leu NES peptide .....	117
6.5 The previously postulated c-Abl NES is non-functional in the context of the Abl C-terminal domain .....	120
<b>References</b> .....	<b>136</b>
<b>Author's contributions to publications</b> .....	<b>148</b>
<b>Curriculum vitae</b> .....	<b>150</b>

## List of abbreviations used in this work (selection)

A <sub>280</sub>	absorbance at $\lambda = 280$ nm
ATP	<b>a</b> denosine 5'- <b>t</b> ri <b>p</b> hosphate
ADP	<b>a</b> denosine 5'- <b>d</b> i <b>p</b> hosphate
BHK cells	<b>b</b> aby <b>h</b> amster <b>k</b> idney cells
BIB domain	<b>b</b> eta-like <b>i</b> mpor <b>t</b> receptor- <b>b</b> inding domain of ribosomal protein L23
cAMP	3', 5'- <b>c</b> yclic <b>a</b> denosine <b>m</b> ono <b>p</b> hosphate
CAS	<b>c</b> ellular <b>a</b> poptosis <b>s</b> usceptibility (= Exportin 2)
cNLS	classical NLS (Imp $\alpha$ / $\beta$ -dependent), nuclear localization signal
C-terminus	carboxy-terminus
CRINEPT	<b>c</b> ross-correlated <b>r</b> elaxation- <b>e</b> nhanced <b>p</b> olarization
CRM1	<b>c</b> hromosomal <b>r</b> egion <b>m</b> aintenance <b>1</b> (= Exportin 1/ Xpo1p)
Cse1p	<b>c</b> hromosome <b>s</b> egregation 1 (yeast ortholog of CAS)
DFP	<b>d</b> iisopropyl <b>f</b> luoro <b>p</b> hosphate
dsRNA	<b>d</b> ouble- <b>s</b> tranded RNA
DTT	<b>d</b> i <b>t</b> hiot <b>h</b> reitol
eIF	<b>e</b> ukaryotic translation <b>i</b> nitiation <b>f</b> actor
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	<b>e</b> thylened <b>i</b> aminet <b>e</b> tetraacetic <b>a</b> cid
ES cell	<b>e</b> mbryonic <b>s</b> tem cell
Exp	exportin
FG repeat	phenylalanine/glycine repeat
FP	<b>f</b> luorescence <b>p</b> olarization
GAP	<b>G</b> TPase- <b>a</b> ctivating <b>p</b> rotein
Gd(DTPA-BMA)	<b>G</b> adolinium-(III)- <b>d</b> iethylenetriamine <b>p</b> enta <b>a</b> ctic acid <b>b</b> is <b>m</b> ethyl <b>a</b> mid
G protein	see GNBP
GDP	<b>g</b> uanosine 5'- <b>d</b> i <b>p</b> hosphate
GFP	<b>g</b> reen <b>f</b> luorescent <b>p</b> rotein
GNBP	<b>g</b> uanine <b>n</b> ucleotide- <b>b</b> inding <b>p</b> rotein
GppNHp	5'-guanylyl imidodiphosphate (a non-hydrolyzable GTP analog)
GST	<b>g</b> lutathione <b>S</b> - <b>t</b> ransferase
GTP	<b>g</b> uanosine 5'- <b>t</b> ri <b>p</b> hosphate
GTPase	GTP hydrolase
HEAT repeat	class of protein repeats ( <b>h</b> untingtin, <b>e</b> longation factor 3, the PR65/A subunit of protein phosphatase <b>2A</b> (PP2A), lipid kinase <b>T</b> OR)
HEPES	4-(2- <b>h</b> ydroxy <b>e</b> thyl)-1- <b>p</b> iperazine <b>e</b> thanesulfonic acid
HIV	<b>h</b> uman <b>i</b> mmunodeficiency <b>v</b> irus
HMG	<b>h</b> igh- <b>m</b> obility <b>g</b> roup
hnRNP	<b>h</b> eterogeneous <b>n</b> uclear <b>r</b> ibonucleoprotein
HPLC	<b>h</b> igh <b>p</b> erformance <b>l</b> iquid <b>c</b> hromatography
<i>H. sapiens</i>	<i>Homo sapiens</i>
HSQC	<b>h</b> eteronuclear <b>s</b> ingle- <b>q</b> uantum <b>c</b> oherence
HMQC	<b>h</b> eteronuclear <b>m</b> ultiple- <b>q</b> uantum <b>c</b> oherence
IBB domain	<b>I</b> mp <b><math>\beta</math></b> - <b>b</b> inding domain of Imp $\alpha$ (compare to "sIBB")
IgG	<b>i</b> mmunoglobulin <b>G</b>
Imp	importin
ITPG	<b>i</b> sopropyl- $\beta$ -D- <b>t</b> hiogalactopyranoside
Kap	<b>k</b> aryopherin (see also "NTR")
LB	<b>L</b> uria- <b>B</b> ertani (lysogeny broth, medium)
LMB	<b>l</b> eptomycin <b>B</b>
K <sub>a</sub>	dissociation constant
kDa	<b>k</b> ilodalton
MBP	<b>m</b> altose- <b>b</b> inding <b>p</b> rotein
m <sub>3</sub> G	2,2,7-trimethyl guanosine (cap)
m <sup>7</sup> G	7-methyl guanosine ("5' cap")
mCherry	<b>m</b> onomeric <b>C</b> herry (a red-fluorescent protein)
MDa	<b>m</b> egadalton

mRFP	<b>m</b> onomeric <b>r</b> ed <b>f</b> luorescent <b>p</b> rotein
MW	<b>m</b> olecular <b>w</b> eight
NA	<b>n</b> umerical <b>a</b> perture
N-terminus	amino-terminus
ncNLS	<b>n</b> on- <b>c</b> lassical NLS (Imp $\beta$ -dependent)
NE	<b>n</b> uclear <b>e</b> nvelope
NES	"Leucine-rich" <b>n</b> uclear <b>e</b> xport <b>s</b> ignal
NLS	<b>n</b> uclear <b>l</b> ocalization <b>s</b> ignal
NMR	<b>n</b> uclear <b>m</b> agnetic <b>r</b> esonance
NOE	<b>N</b> uclear <b>O</b> verhauser <b>E</b> ffect
NOESY	<b>N</b> OE <b>S</b> pectroscopy
NPC	<b>n</b> uclear <b>p</b> ore <b>c</b> omplex
NTA	<b>n</b> itro <b>t</b> riacetic <b>a</b> cid
NTF2	<b>n</b> uclear <b>t</b> ransport <b>f</b> actor <b>2</b>
NTR	<b>n</b> uclear <b>t</b> ransport <b>r</b> eceptor (see also "Kap")
Nup	<b>n</b> ucleoporin (NPC protein)
OD	<b>o</b> ptical <b>d</b> ensity
PBS	<b>p</b> hosphate- <b>b</b> uffered <b>s</b> aline
PEG	<b>p</b> oly <b>e</b> thylene <b>g</b> lycol
PKA	<b>p</b> rotein <b>k</b> inase <b>A</b> (cAMP-dependent protein kinase)
PKI	<b>p</b> rotein <b>k</b> inase <b>A</b> <b>i</b> nhibitor
P-loop	<b>p</b> hosphate-binding loop
PMSF	<b>p</b> henyl <b>m</b> ethyl <b>s</b> ulfonyl <b>f</b> luoride
p.p.m.	<b>p</b> arts <b>p</b> er <b>m</b> illion
PRE	<b>p</b> aramagnetic <b>r</b> elaxation <b>e</b> nhancement
pre-miRNA	pre-micro RNA
PTHrP	<b>p</b> arathyroid <b>h</b> ormone- <b>r</b> elated <b>p</b> rotein (an Imp $\beta$ cargo)
Ran	<b>R</b> as-related <b>n</b> uclear antigen
RanBD	Ran- <b>b</b> inding <b>d</b> omain
RanBP	Ran- <b>b</b> inding <b>p</b> rotein
RanGAP	Ran <b>G</b> TPase- <b>a</b> ctivating <b>p</b> rotein
RanGEF	Ran <b>g</b> uanine nucleotide <b>e</b> xchange <b>f</b> actor (RCC1)
Ras	<b>R</b> at <b>s</b> arcoma
RCC1	<b>r</b> egulator of <b>c</b> hromosome <b>c</b> ondensation <b>1</b> (see also "RanGEF")
RMSD	<b>r</b> oot <b>m</b> ean <b>s</b> quare <b>d</b> eviation
RNP	<b>r</b> ibonucleoprotein (particle)
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS-PAGE	sodium <b>d</b> odecyl sulfate- <b>p</b> oly <b>a</b> crylamide <b>g</b> el <b>e</b> lectrophoresis
sIBB domain	<b>S</b> PN1 <b>I</b> mp $\beta$ - <b>b</b> inding domain (compare to "IBB")
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
SPN1	Snurportin1
SV40	<b>s</b> imian <b>v</b> irus 40
TEV	<b>t</b> obacco <b>e</b> tch <b>v</b> irus
Tris	Tris(hydroxymethyl)aminomethane, 2-amino-2-hydroxymethyl-1,3-propanediol
TOCSY	<b>T</b> otal <b>C</b> orrelated <b>S</b> pectroscopy
TROSY	<b>T</b> ransverse <b>R</b> elaxation- <b>O</b> ptimized NMR <b>S</b> pectroscopy
UsnRNP	<b>u</b> ridine-rich <b>s</b> mall <b>n</b> uclear <b>r</b> ibonucleoprotein (particle)
v/v	volume per volume
w/v	weight per volume
<i>X. laevis</i>	<i>Xenopus laevis</i>
YT / 2YT	<b>y</b> east extract <b>t</b> ryptone medium
z (in "zz")	IgG-binding domain of the <i>Staphylococcal</i> protein A

*Standard single-letter amino acid codes and the International System of units (SI) were used.*

## Summary

All nucleocytoplasmic exchange proceeds through nuclear pore complexes, which constitute giant aqueous channels in the nuclear envelope. Most of this traffic is mediated by Ran-dependent Importin  $\beta$ -like nuclear transport receptors, which include import mediators (importins) as well as exportins. Importin 13 and Msn5p are known exceptions in that they can carry distinct cargos into different directions. We demonstrated that Exportin 4, in addition to its established function in nuclear export, acts as a *bona fide* nuclear import receptor for Sox-type transcription factors.

CRM1 (also known as Exportin 1) is the cell's most versatile export receptor, recognizing myriads of structurally unrelated proteins. How CRM1 achieves this remarkable multispecificity and how Ran triggers cargo loading was a fundamental but unresolved question. The instability of export complexes was considered to preclude crystallographic studies.

I established constructs and conditions to reconstitute stable and chromatographically homogeneous export complexes that allowed us to determine the crystal structure of the Snurportin1·CRM1·RanGTP complex. Snurportin is an exceptional cargo in that it binds CRM1 with an extraordinarily high affinity. The structure shows that Ran promotes Snurportin binding solely through long-range conformational changes in CRM1 and reveals how the exportin contacts a protein comprising an extensive and complex tripartite export signature that includes a folded domain.

Far simpler CRM1-dependent export determinants are the so-called "Leucine-rich" *nuclear export signals (NESs)* – a diverse family of peptides that contain spaced hydrophobic ( $\Phi$ ) residues. NES-like sequence patterns occur very frequently in proteins, even in those that are not recognized by CRM1. We found that the previously proposed NES of the Abelson (Abl) tyrosine kinase is functional in isolation but not in the context of the Abl C-terminal domain, where  $\Phi$  residues are buried in the hydrophobic core. This emphasizes that structural information is indispensable for the correct prediction of export signals. To elucidate how CRM1 can interact with various NESs, we used a crystallographic approach that relies on NES-Snurportin chimeras. We solved the crystal structures of the RanGTP-CRM1 complex alone and when bound to the prototypic PKI (protein kinase A inhibitor) or HIV-1 Rev NESs. These NESs differ drastically in the spacing of their  $\Phi$  residues. Yet, Ran-bound CRM1

recognizes them with the same set of five  $\Phi$  pockets. While these pockets are rigid, variable  $\Phi$  spacings in the NESs are compensated for by different conformations of the bound NES peptides. Our NMR analysis of the PKI NES in its free state and when bound to CRM1 suggests that CRM1 selects NES conformers that pre-exist in solution. These findings and our observations that individual  $\Phi$  residues are dispensable and that each  $\Phi$  pocket can accept a wide range of hydrophobic residues explain the enormous flexibility in CRM1·NES recognition. Our data lead to a new structure-based NES consensus that provides the basis for predicting the affinities of NESs for CRM1.

---

**Results described in this thesis have been published (or accepted for publication) in the following research articles:**

Hantschel, O., Wiesner, S., **Güttler, T.**, Mackereth, C. D., Rix, L. L., Mikes, Z., Dehne, J., Görlich, D., Sattler, M. & Superti-Furga, G. (2005) Structural basis for the cytoskeletal association of Bcr-Abl/c-Abl. *Mol Cell* **19**, 461-473.

Gontan, C., **Güttler, T.**, Engelen, E., Demmers, J., Fornerod, M., Grosveld, F. G., Tibboel, D., Görlich, D., Poot, R. A. & Rottier, R. J. (2009) Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *J Cell Biol* **185**, 27-34.

*C. Gontan and T. Güttler contributed equally to this work.*

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

*T. Monecke and T. Güttler contributed equally to this work.*

**Güttler, T.**, Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., Görlich, D. (2010) NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol*, **accepted for publication** (on condition of editorial changes in the text).

*T. Güttler, T. Madl and P. Neumann contributed equally to this work.*

Note added for publication of this thesis:

The citation of the article's print version is *Nat Struct Mol Biol* **17**, 1367-1376.

## CHAPTER 1

### Introduction to nucleocytoplasmic transport

#### 1.1 Compartmentation of eukaryotic cells

Eukaryotic cells are subdivided into membrane-bounded compartments. Each of these compartments harbors characteristic sets of structural proteins and enzymes to carry out specific tasks (for examples, see De Duve *et al.*, 1953; Leighton *et al.*, 1968). The advantages of such organization are best witnessed by the fact that only eukaryotes evolved to complex multicellular organisms.

The defining feature of a eukaryotic cell is the cell nucleus, which is delineated by the double membrane of the *nuclear envelope (NE)*. The benefits of this nucleocytoplasmic compartmentation are manifold. Firstly, the NE encloses the genetic material and contributes in various ways to genomic stability, allowing eukaryotes to handle  $\approx 1000$ -fold larger genomes than prokaryotes. Secondly, the confinement of the genome in the nucleus allows the cell to fine-tune gene expression by controlling the access of transcription factors to DNA – an instance of control unavailable to prokaryotes (Kaffman and O'Shea, 1999). Further, the NE separates the nuclear events of transcription and RNA processing from translation in the cytoplasm. It thereby prevents translation of unspliced or incompletely spliced transcripts, which would give rise to non-functional or even dominant-negative proteins. The NE also ensures temporal separation of transcription and translation, as transcripts will normally only be released from the spliceosome when processed to completion (Legrain and Rosbash, 1989; Custódio *et al.*, 1999). This separation provides the basis for the overall fidelity of gene expression.

The division of eukaryotic cells into nucleus and cytoplasm necessitates nucleocytoplasmic exchange. All nuclear proteins, such as histones and components required for DNA replication and transcription, originate in the cytoplasm and need to be imported into the nucleus (e.g. Gurdon, 1970). Conversely, tRNAs, mRNAs and ribosomes are generated or assembled in the nucleus and have to be exported to the cytoplasm, where they function in translation. During the cell cycle of higher eukaryotes, the NE breaks down, causing nuclear and cytoplasmic contents to intermix. With re-assembly of the NE in telophase, nuclear and

cytoplasmic proteins are re-segregated – a task that obviously requires transport into and out of the nucleus (reviewed in Görlich and Kutay, 1999).

## 1.2 Overview of nucleocytoplasmic exchange

Nucleocytoplasmic exchange is a truly impressive activity: in a growing mammalian cell, more than two million macromolecules per second are actively exchanged between nucleus and cytoplasm, amounting to a total mass flow of > 200 GDa (Ribbeck and Görlich, 2001). To accomplish this enormous task, the cell invests a considerable amount of resources, including ≈80-100 distinct, often highly abundant proteins.

All nucleocytoplasmic traffic proceeds through *nuclear pore complexes (NPCs)* - proteinaceous gates that are embedded in circular openings of the NE, where outer and inner nuclear membrane fuse. The term "pore complex" has been proposed about half a century ago (Watson, 1959), based on several electron microscopic studies in the early 1950s. With molecular weight estimates ranging from ≈55-66 MDa for yeast NPCs (Rout and Blobel, 1993; Yang *et al.*, 1998) and ≈125 MDa for vertebrate pore complexes (Reichelt *et al.*, 1990), NPCs are gigantic. (For comparison, the molecular weight of a ribosome is ≈4 MDa.) For a summary on the more recent advances in the structural characterization of the NPC, I would like to refer to a review by Brohawn *et al.*, 2009. The central portion of the NPC constitutes an aqueous channel with an inner diameter as large as ≈40 nm (Panté and Kann, 2002), which can accommodate large objects (of up to 25-50 MDa in molecular weight), such as the giant Balbiani ring particles, intact viral capsids and pre-ribosomes (see Franke and Scheer, 1974 and references therein; Feldherr *et al.*, 1984; Daneholt, 1997; Kiseleva *et al.*, 1998; Panté and Kann, 2002).

Small molecules (i.e. metabolites and small proteins) can freely diffuse through the NPC (Bonner, 1975). In contrast, macromolecules above the passive diffusion limit (i.e. with a molecular weight exceeding 20-40 kDa) depend on an appropriate *nuclear transport receptor (NTR)* for rapid translocation. These two modes of NPC passage are referred to as passive diffusion or facilitated translocation, respectively. Aspects of the latter are in the focus of this work.

### 1.2.1 *Importin $\beta$ -like nuclear transport receptors*

Most nuclear transport pathways are mediated by NTRs of the *Importin  $\beta$*  (*Imp $\beta$* ) superfamily (also referred to as  $\beta$ -karyopherins, Kap- $\beta$ s). These receptors constantly shuttle between nucleus and cytoplasm, bind their cargo (directly or through adapter molecules) on one side of the NE and release it on the other side. Unifying features of *Imp $\beta$* -like NTRs are their rather large sizes (90-150 kDa), their acidic isoelectric points (pI 4.0-5.0), their capacity to bind the small guanine nucleotide-binding protein Ran (Görlich *et al.*, 1997; Fornerod *et al.*, 1997b, see below), and their affinity for phenyl-Sepharose (Ribbeck and Görlich, 2002). Despite their related functions and similar domain organizations, *Imp $\beta$* -like NTRs share surprisingly little overall sequence homology, with identity ranging from only 8 to 15%. The sequence similarity is highest in their N-terminal region, which accounts for most of their Ran-binding activity and negative charge (Görlich *et al.*, 1997).

Ran is a family member of the small (Ras-related) GTPases (Drivas *et al.*, 1990; Bischoff and Ponstingl, 1991a; Melchior *et al.*, 1993; Moore and Blobel, 1993). It can cycle between two states – a GTP-bound (ON) and a GDP-bound (OFF) state. Ran's intrinsic GTPase activity is very low, and thus it strictly depends on its *GTPase-activating protein RanGAP* to hydrolyze GTP (Bischoff *et al.*, 1994; Klebe *et al.*, 1995). Notably however, RanGAP alone is not sufficient to activate Ran bound to *Imp $\beta$* -like NTRs – it requires the assistance of *RanBP1* (*Ran-binding protein 1*) or the homologous Ran-binding domains of RanBP2/Nup358 (see **Chapter 5**). The *Ran guanine nucleotide exchange factor (RanGEF) RCC1* (*regulator of chromosome condensation 1*) catalyzes GDP-to-GTP exchange on Ran (Bischoff and Ponstingl, 1991b; Klebe *et al.*, 1995). RanGAP, RanBP1 and RanBP2 are restricted to the cytoplasm or the cytoplasmic side of NPCs (Yokoyama *et al.*, 1995; Richards *et al.*, 1996; Matunis *et al.*, 1996; Mahajan *et al.*, 1997; Saitoh *et al.*, 1997) where they deplete RanGTP. In contrast, their opponent RCC1 is chromatin-bound, generating RanGTP exclusively in the nuclear compartment (Ohtsubo *et al.*, 1989). This asymmetric distribution of Ran's regulators establishes a steep nucleocytoplasmic gradient of RanGTP, with a high nuclear and a low cytoplasmic concentration (Görlich *et al.*, 1996; Izaurralde *et al.*, 1997).

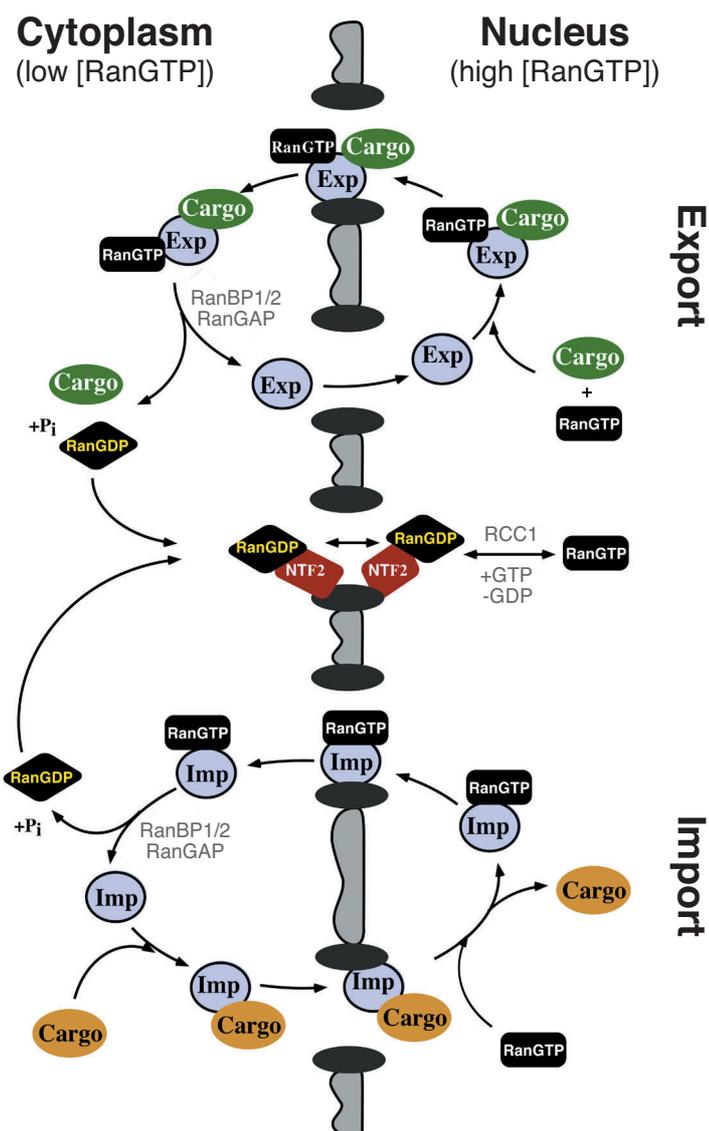
In its ON state, Ran binds to *Imp $\beta$* -like NTRs to trigger cargo loading or cargo release (see also below). Thus, all *Imp $\beta$* -likes exploit the chemical potential of the RanGTP gradient to mediate unidirectional cargo transport. Based on their responses to RanGTP binding, two classes of transport receptors can be distinguished: nuclear import mediators (importins) and exportins (reviewed by Görlich and Kutay, 1999). Their transport cycles are illustrated in

**Figure 1-1.** Importins bind their cargo at a low RanGTP concentration in the cytoplasm and traverse the NPC as dimeric importin·cargo complexes. In the nucleus, binding of RanGTP to the importin dissociates the import complex and thereby renders cargo import irreversible. The resulting importin·RanGTP complex is then recycled back to the cytoplasm, where the GTPase activators cause hydrolysis of Ran-bound GTP to GDP. This ultimately dissociates the importin·Ran complex and allows the importin to mediate another round of import. Exportins operate in exactly the opposite manner, recruiting their cargo at high RanGTP levels in the nucleus. In this process, cargo binding and RanGTP binding to the exportin are coupled by positive cooperativity, i.e., RanGTP increases the affinity of the exportin for its cargo and *vice versa*. Exportins traverse the NPC as ternary cargo·exportin·RanGTP complexes and discharge their cargo and Ran upon GTP hydrolysis in the cytoplasm. The exportin then re-enters the nucleus to act in another export cycle. The just-described processes illustrate how input of metabolic energy into the Ran cycle drives directional nuclear transport. This allows accumulation of transport substrates against gradients of chemical activity. It should be noted that the translocation process *per se* is fully reversible and energy-independent (Kose *et al.*, 1997; Nakielny and Dreyfuss, 1998; Ribbeck *et al.*, 1998; Schwoebel *et al.*, 1998; Ribbeck *et al.*, 1999; Englmeier *et al.*, 1999; Nachury and Weis, 1999; Zeitler and Weis, 2004).

Both import and export cycles steadily carry Ran to the cytoplasm and should eventually deplete the nuclear RanGTP pool. Although Ran is sufficiently small to rapidly re-enter the nucleus by passive diffusion (25 kDa), the cell maintains the predominant nuclear localization of Ran (Bischoff and Ponstingl, 1991a) by employing a dedicated nuclear import receptor specific for RanGDP – *nuclear transport factor 2* (NTF2, Ribbeck *et al.*, 1998; Smith *et al.*, 1998; see **Figure 1-1**). Indeed, the lack of NTF2 would reduce the steepness of the RanGTP gradient (Görlich *et al.*, 2003), which is probably the reason why NTF2 was found to be required for efficient nuclear import of proteins (Moore and Blobel, 1994; Paschal and Gerace, 1995). By the criteria mentioned above and by its structure (Bullock *et al.*, 1996; also see below), NTF2 is not an Imp $\beta$ -like NTR. Efficient disassembly of NTF2·RanGDP complexes and release of Ran into the nucleoplasm appears to require the concerted action of RCC1 and Imp $\beta$ -like NTRs (Ribbeck *et al.*, 1998; Smith *et al.*, 1998). An additional (ATP-dependent) dissociation factor had been postulated but not identified so far (Yamada *et al.*, 2004).

The Importin  $\beta$  superfamily comprises at least 21 members in mammals and 15 members in *Saccharomyces cerevisiae*. **Table 1-1** lists functionally characterized NTRs and a selection of

their respective cargoes. Of the mammalian transport receptors, seven have been found to mediate nuclear export while 13 have been described as importins. Importin 13 (in higher eukaryotes) and Msn5p (in *S. cerevisiae*) are known exceptions in that they can mediate both import and export (Mingot *et al.*, 2001; Yoshida and Blobel, 2001).



**Figure 1-1: Schematic overview of the nuclear export and import cycles and their coordination by the RanGTPase system.**

The figure has been adapted from Görlich and Kutay, 1999. "Exp" = exportin, "Imp" = importin. Nuclear export and import are signal-mediated, i.e. they depend on nuclear export signals or nuclear localization signals, respectively. See text for details.

**Table 1-1: Functionally characterized NTRs of the Imp $\beta$  superfamily and their adapters or co-receptors.**

The mammalian NTRs are listed in black, selected *S. cerevisiae* orthologs in gray. Only a selection of cargoes and references is shown. Exp4 had been re-classified based on this work (see **Chapter 2**).

<b>NTR</b>	<b>adapters or co-receptors</b>	<b>selected cargoes</b>	<b>selected references</b>
<i>Exportins</i>			
CRM1 (Exportin 1) Xpo1p/Kap124p	HIV Rev PHAX	Leu-rich NES cargoes RRE-containing RNAs m <sup>7</sup> G-capped UsnRNAs Snurportin1 (SPN1)	Wen <i>et al.</i> , 1995 Fischer <i>et al.</i> , 1995 Fischer <i>et al.</i> , 1995 Ohno <i>et al.</i> , 2000 Paraskeva <i>et al.</i> , 1999
CAS (Exportin 2) Cse1p/Kap109p		Importin $\alpha$ s	Kutay <i>et al.</i> , 1997
Exp-t (Exportin 3) Los1p/Kap127p		tRNA	Kutay <i>et al.</i> , 1998 Arts <i>et al.</i> , 1998
Exportin 5 (Msn5p/Kap142p, see below)	aa-tRNA dsRNA	eEF1A dsRNA-binding proteins pre-miRNAs	Bohnsack <i>et al.</i> , 2002 Calado <i>et al.</i> , 2002 Brownawell and Macara, 2002 Lund <i>et al.</i> , 2004 Bohnsack <i>et al.</i> , 2004
Exportin 6		actin-profilin complexes	Stüven <i>et al.</i> , 2003
Exportin 7		p50RhoGAP, 14-3-3 $\sigma$	Mingot <i>et al.</i> , 2004
<i>Importins</i>			
Importin $\beta$ (Imp $\beta$ -1) Kap- $\beta$ 1p/Kap95p	Importin $\alpha$ Kap- $\alpha$ /Kap60p Snurportin1 XRIP $\alpha$ Importin 7 Nmd5p/Kap119p	ribosomal proteins HIV Rev, HIV Tat histones classical NLS-cargoes m <sub>3</sub> G-capped UsnRNPs replication protein A histone H1	Jäkel and Görlich, 1998 Truant and Cullen, 1999 Mühlhäusser <i>et al.</i> , 2001 Mosammaparast <i>et al.</i> , 2001 Baake <i>et al.</i> , 2001 Adam and Adam, 1994 Görlich <i>et al.</i> , 1994 Görlich <i>et al.</i> , 1995 Chi <i>et al.</i> , 1995 Imamoto <i>et al.</i> , 1995 Radu <i>et al.</i> , 1995 Huber <i>et al.</i> , 1998 Jullien <i>et al.</i> , 1999 Jäkel <i>et al.</i> , 1999
Transportin 1+2 (Trn, Imp $\beta$ -2) Kap- $\beta$ 2p/Kap104p		hnRNP proteins (M9-NLS) ribosomal proteins TAP/NXF1 histones c-Fos SRP19	Pollard <i>et al.</i> , 1996 Siomi <i>et al.</i> , 1997 Jäkel and Görlich, 1998 Truant and Cullen, 1999 Mühlhäusser <i>et al.</i> , 2001 Baake <i>et al.</i> , 2001 Arnold <i>et al.</i> , 2006 Dean <i>et al.</i> , 2001
Transportin SR 1+2 (TrnSR, Trn 3) Mtr10p/Kap111p		SR proteins tRNA	Kataoka <i>et al.</i> , 1999 Lai <i>et al.</i> , 2000 Shaheen and Hopper, 2005

Importin 4 a+b	ribosomal proteins	Schlenstedt <i>et al.</i> , 1997 Jäkel <i>et al.</i> , 2002
	histones	Mosammaparast <i>et al.</i> , 2001 Mosammaparast <i>et al.</i> , 2002
Importin 5 Pse1p/Kap121p	ribosomal proteins	Deane <i>et al.</i> , 1997 Jäkel and Görlich, 1998
	histones	Mühlhäusser <i>et al.</i> , 2001 Mosammaparast <i>et al.</i> , 2001 Baake <i>et al.</i> , 2001 Mosammaparast <i>et al.</i> , 2002
Importin 7 Nmd5p/Kap119p	ribosomal proteins	Jäkel and Görlich, 1998
	histones	Mühlhäusser <i>et al.</i> , 2001 Baake <i>et al.</i> , 2001
	ERK2, SMAD3, MEK1	Chuderland <i>et al.</i> , 2008
Importin 8	SRP19	Dean <i>et al.</i> , 2001
	Argonaute proteins	Weinmann <i>et al.</i> , 2009
Importin 9 (a+b)	histones	Mühlhäusser <i>et al.</i> , 2001 Mosammaparast <i>et al.</i> , 2001
	ribosomal proteins	Jäkel <i>et al.</i> , 2002
Importin 11	UbcM2	Plafker and Macara, 2000
	rpL12	Plafker and Macara, 2002
<i>Bidirectionally operating NTRs</i>		
Importin 13	hUBC9, MGN/Y14 (import)	Mingot <i>et al.</i> , 2001
	eIF1A (export)	
	TF NF-Y (import)	Kahle <i>et al.</i> , 2005
	histone fold heterodimers (import)	Walker <i>et al.</i> , 2009
Msn5p/Kap142p	Pho4p (export)	Kaffman <i>et al.</i> , 1998
	RPA (import)	Yoshida and Blobel, 2001
Exportin 4	eIF5A (export)	Lipowsky <i>et al.</i> , 2000
	SMAD3 (export)	Kurisaki <i>et al.</i> , 2006
	Sox2, SRY (import)	Gontan <i>et al.</i> , 2009 <b>(this work)</b>

### 1.2.2 The nuclear export receptor CRM1/Exportin 1

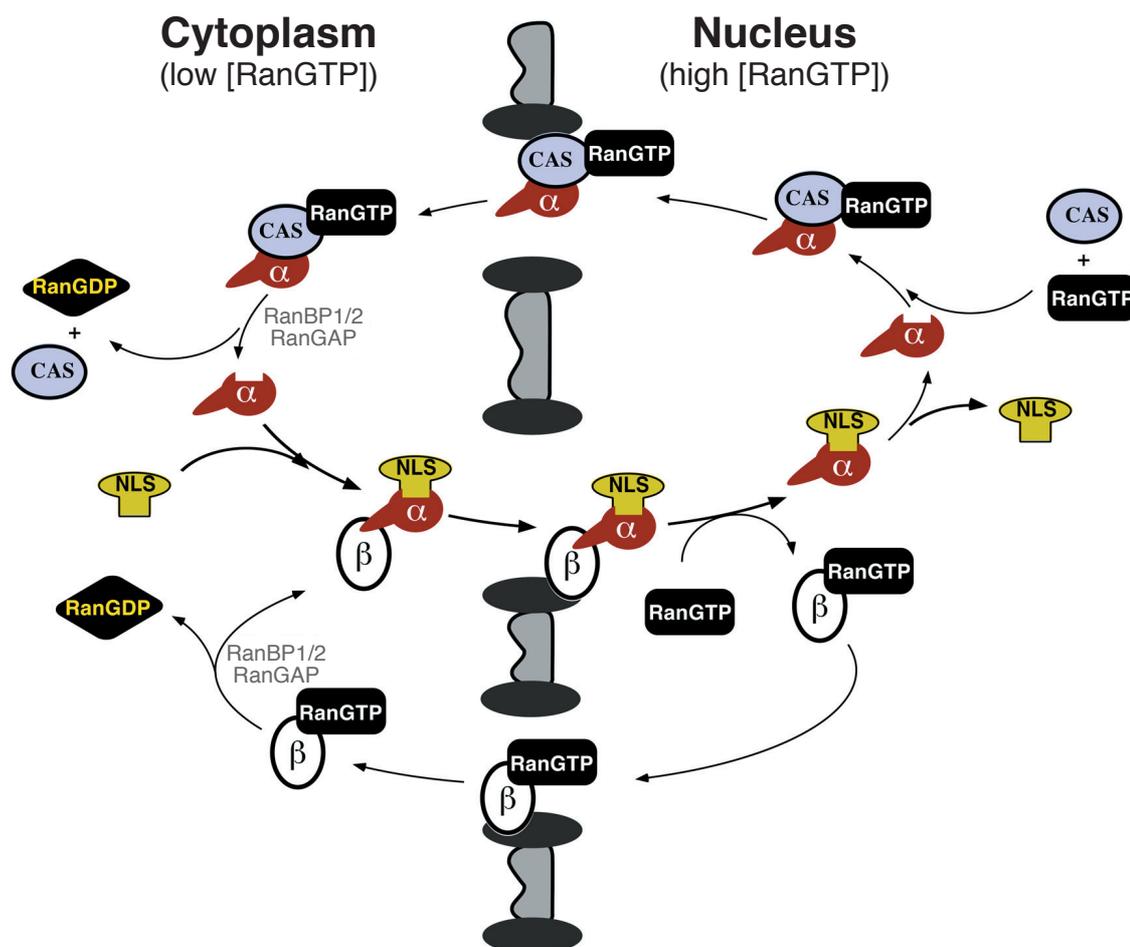
CRM1/Exportin 1 was initially discovered in *Schizosaccharomyces pombe*, based on cold-sensitive mutants with "deformed nuclear chromosome domains" (Adachi and Yanagida, 1989). This pleiotropic phenotype gave rise to the protein's name ("**chromosomal region maintenance 1**"). CRM1/Xpo1p and CAS/Cse1p (**Table 1-1**) are the prototypical exportins (Stade *et al.*, 1997; Fornerod *et al.*, 1997a; Ossareh-Nazari *et al.*, 1997; Fukuda *et al.*, 1997; Kutay *et al.*, 1997; Solsbacher *et al.*, 1998).

CRM1 is essential for the viability of all organisms tested so far. Among the exportins, it clearly is the most versatile receptor, recognizing hundreds of structurally unrelated proteins. Its activities can be grouped into four categories. The first one comprises "biosynthetic

transport", which is required for the assembly of biosynthetic "molecular machines" such as the ribosome and the spliceosome (Ho *et al.*, 2000; Gadal *et al.*, 2001; Moy and Silver, 2002; Thomas and Kutay, 2003; Fornerod *et al.*, 1997a; Ohno *et al.*, 2000; Will and Lührmann, 2001). Secondly, CRM1 counteracts the leakage of cytoplasmic factors into nuclei and thereby contributes to compartment identity. Examples of such cargoes are several translation factors (Bohnsack *et al.*, 2002) or RanBP1 (Richards *et al.*, 1996). Thirdly, CRM1 is a key element in many regulatory networks. For instance, it controls the nuclear activity of cAMP-dependent protein kinase (protein kinase A, PKA) by expelling the PKA inhibitor (PKI)·PKA complex from the nucleus (Fantozzi *et al.*, 1994; Wen *et al.*, 1994; Wen *et al.*, 1995). Here, PKI acts as an adaptor to CRM1. Finally, CRM1 is also required for the infection cycles of numerous viruses. HIV-1, for example, relies on CRM1 to export its genomic RNA from nuclei (Malim *et al.*, 1991; Fischer *et al.*, 1995; Fornerod *et al.*, 1997a). The HIV-1 Rev protein is an essential adaptor in this process. It binds the unspliced viral RNA, recruits CRM1 and thereby promotes HIV RNA export to the cytoplasm, where the next generation of viral particles is assembled. For a comprehensive review on CRM1, see Hutten and Kehlenbach, 2007.

### 1.2.3 Nuclear transport signals

NTR-cargo interaction is highly specific and governed by the cargoes' nuclear transport signals. Import cargoes bind to importins by virtue of their import signatures. These signals are as varied as the import substrates – there is neither a unifying consensus sequence, nor a length to which they are restricted, reflecting that multiple import receptors of different specificities join forces to accomplish the protein supply of the nucleus (**Table 1-1**). Prototypes of the so-called "*classical nuclear localization signals*" (*cNLSs*) are the NLS of the *simian virus 40* (*SV40*) large T-antigen, which comprises a single cluster of basic amino acids (hence called "monopartite") and the "bipartite" NLS of nucleoplasmin, where a short spacer separates two basic clusters (Kalderon *et al.*, 1984; Lanford and Butel, 1984; Robbins *et al.*, 1991). *cNLSs* do not bind to Imp $\beta$  directly. Instead, Imp $\alpha$  serves as the actual NLS receptor, which, upon cargo binding, pairs with Imp $\beta$  for translocation through the NPC (**Table 1-1**). Seven mammalian Imp $\alpha$  homologs have been characterized so far (**Table 1-2**). The transport cycle of the Imp $\alpha/\beta$  system is depicted in **Figure 1-2**. In this process, Imp $\alpha$  interacts with Imp $\beta$  through a sequence stretch termed the *IBB* (*Imp $\beta$ -binding*) domain (Görlich *et al.*, 1996; Weis *et al.*, 1996).



**Figure 1-2: Schematic representation of the Imp $\alpha$  transport cycle.**

For import into the nucleus, the cargo binds to Imp $\alpha$ , which acts as the actual NLS receptor. In turn, cargo-bound Imp $\alpha$  pairs with Imp $\beta$  (via its Imp $\beta$ -binding domain, "IBB domain"). The resulting trimeric complex translocates through the NPC. In the nucleus, RanGTP triggers disassembly of the import complex. The cargo is released into the nucleus and Imp $\alpha$  is re-exported by CAS (see **Table 1-1** and **Figure 1-1**). The resulting Imp $\beta$ ·RanGTP complex directly exits the nucleus. The figure has been adapted from Görlich and Kutay, 1999.

It should be noted that the IBB domain auto-inhibits Imp $\alpha$ 's NLS-binding site in the absence of NLS cargo (not shown, Kobe, 1999). This mechanism and the action of CAS and nuclear pore complex components promote cargo release from Imp $\alpha$  and render the Imp $\alpha$  cycle efficient (Solsbacher *et al.*, 2000; Catimel *et al.*, 2001; Gilchrist *et al.*, 2002; Gilchrist and Rexach, 2003; Matsuura *et al.*, 2003; Harreman *et al.*, 2003; Kutay *et al.*, 1997; Matsuura and Stewart, 2004; Matsuura and Stewart, 2005).

**Table 1-2: (next page) Functionally characterized mammalian Importin  $\alpha$  homologs.**

The Importin  $\alpha$  protein family is diverse and a rational basis for a generalized nomenclature has only been suggested recently (Mason *et al.*, 2009). Most metazoan Importin- $\alpha$ s assort into subtypes  $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 3 (Malik *et al.*, 1997; Tsuji *et al.*, 1997; Köhler *et al.*, 1997; Goldfarb *et al.*, 2004). The table shows human Importin  $\alpha$  isoforms according to the nomenclature used by Köhler *et al.*, 1999. For clarity, NCBI protein accession codes are given. Importin  $\alpha$ 8 has only been found recently (Tejomurtula *et al.*, 2009). See Köhler *et al.*, 1999 for references. "\*" marks those factors that had been employed in the course of this work (**Chapter 2**).

Imp $\alpha$ subtype	nomenclature used	selected synonyms	accession number ( <i>H. sapiens</i> )
$\alpha 1$	Importin $\alpha 5^*$	Karyopherin $\alpha 1$ Rch2 hSRP1- $\beta$ hSRP1	NP_002255
	Importin $\alpha 6$	Karyopherin $\alpha 5$	NP_002260
	Importin $\alpha 7^*$	Karyopherin $\alpha 6$	NP_036448
$\alpha 2$	Importin $\alpha 1^*$	Karyopherin $\alpha 2$ Rch1 hSRP1- $\alpha$	NP_002257
$\alpha 3$	Importin $\alpha 4^*$	Karyopherin $\alpha 3$ hSRP1- $\gamma$	NP_002258
	Importin $\alpha 3^*$	Karyopherin $\alpha 4$ SRP3 Qip1	NP_002259
?	Importin $\alpha 8$	Karyopherin $\alpha 7$	NP_001139187

It should be noted that Imp $\beta$  can also pair with Imp7 to import, for instance, histone H1 (Görlich *et al.*, 1997; Jäkel *et al.*, 1999, see **Table 1-1**). Many cargoes interact directly with Imp $\beta$ , without the need of Imp $\alpha$ . Prominent examples for which the NLSs have been mapped to a defined peptide stretch are the *parathyroid hormone-related protein* (PTHrP, Lam *et al.*, 1999) and the highly basic (pI 12.2) *beta-like import receptor-binding* (BIB) domain of ribosomal protein L23a (Jäkel and Görlich, 1998). To discriminate these NLSs from the Imp $\alpha$ -dependent signals, they have been termed "*non-classical*" NLSs (*ncNLSs*). As shown for ribosomal proteins, recognition of import substrates can also involve very extended and highly basic sequences or three-dimensional patches of great complexity (Jäkel and Görlich, 1998). The best-characterized transportin-dependent NLS is the 38-amino acid M9 signal of hnRNP A1 (Pollard *et al.*, 1996), which, remarkably, only contains two basic residues.

CRM1-dependent export determinants have been studied in great detail. The simplest CRM1-dependent nuclear export determinants are the so-called classical *nuclear export signals* (*NESs*). These are short peptides reported to comprise four spaced hydrophobic residues (denoted  $\Phi^1$  to  $\Phi^4$ ) and to follow the consensus  $\Phi^1-(x)_{2,3}-\Phi^2-(x)_{2,3}-\Phi^3-x-\Phi^4$ , with "x" preferentially being charged, polar or small amino acids (Kutay and Güttinger, 2005). The prototypical representatives are the PKI NES (**LALKLAGLDI**; critical hydrophobics shown in bold), which exemplifies the most common spacing of the hydrophobic positions ( $\Phi^1_{xxx}\Phi^2_{xx}\Phi^3_x\Phi^4$ ), as well as the HIV-1 Rev NES (**LPPLERLTL**) with a reported

$\Phi^1\text{xx}\Phi^2\text{xx}\Phi^3\text{x}\Phi^4$  spacing (Wen *et al.*, 1995; Fischer *et al.*, 1995). Even though classical NESs are also referred to as leucine-rich, analyses of other export cargoes as well as randomization-and-selection-screens revealed that Ile, Val, Met, or Phe are also permitted at the hydrophobic positions (Bogerd *et al.*, 1996; Zhang and Dayton, 1998; Kosugi *et al.*, 2008). *Snurportin 1 (SPN1)*, the nuclear import adaptor for m<sub>3</sub>G-capped spliceosomal *uridine-rich small nuclear ribonucleoproteins (UsnRNPs; Huber et al., 1998)*, exemplifies a CRM1 cargo with a complex export signature (Paraskeva *et al.*, 1999). Deciphering the structural basis for CRM1-cargo and CRM1-RanGTP recognition is the main focus of this work. More details on NESs are given in **Chapter 4**.

### 1.3 The structure of Importin $\beta$ -like nuclear transport receptors

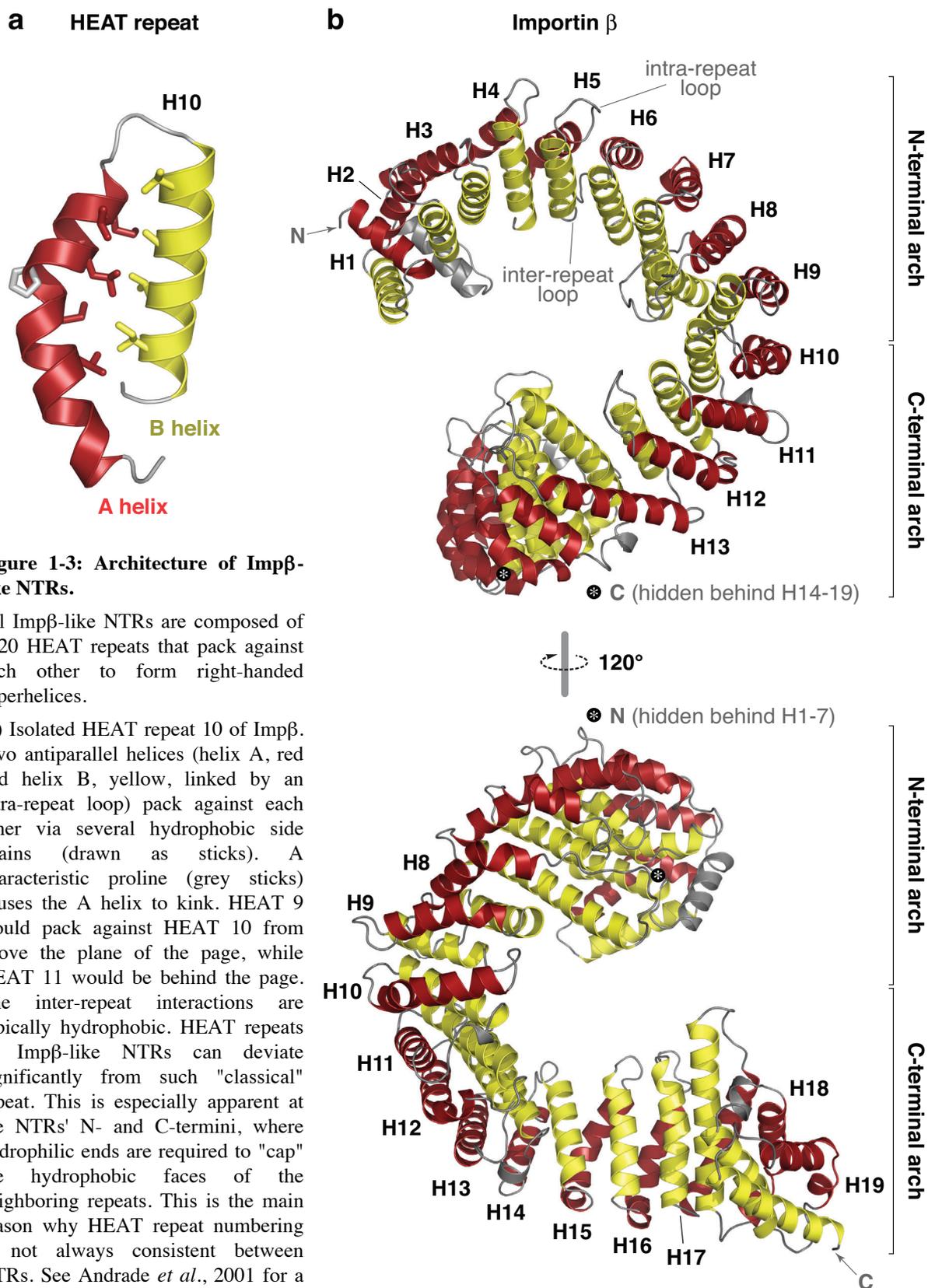
The smallest architectural units of Imp $\beta$ -like NTRs are the so-called HEAT repeats, named after the first members identified for the class of HEAT repeat proteins: **h**untingtin, **e**longation factor 3, the PR65/A subunit of protein phosphatase **2A** (PP2A) and the lipid kinase **T**OR (Andrade and Bork, 1995; Andrade *et al.*, 2001; Perry and Kleckner, 2003). One HEAT repeat is made up of two  $\alpha$ -helices (of 10-20 residues each) that are linked by a short intra-repeat loop (**Figure 1-3a**). The  $\alpha$ -helices (denoted A and B) pack in an antiparallel orientation to form a helical hairpin. In Imp $\beta$ -like NTRs,  $\approx 20$  of these HEAT repeats stack in a tandemly arranged fashion. Since stacking is not perfectly head-to-tail but with a slight clockwise twist, the resulting molecules are not just elongated rods, but form right-handed superhelical structures (**Figure 1-3b**). In these solenoids, the A-helices face the convex back of the transport receptor while the B-helices make up the concave "inner" surface of the protein.

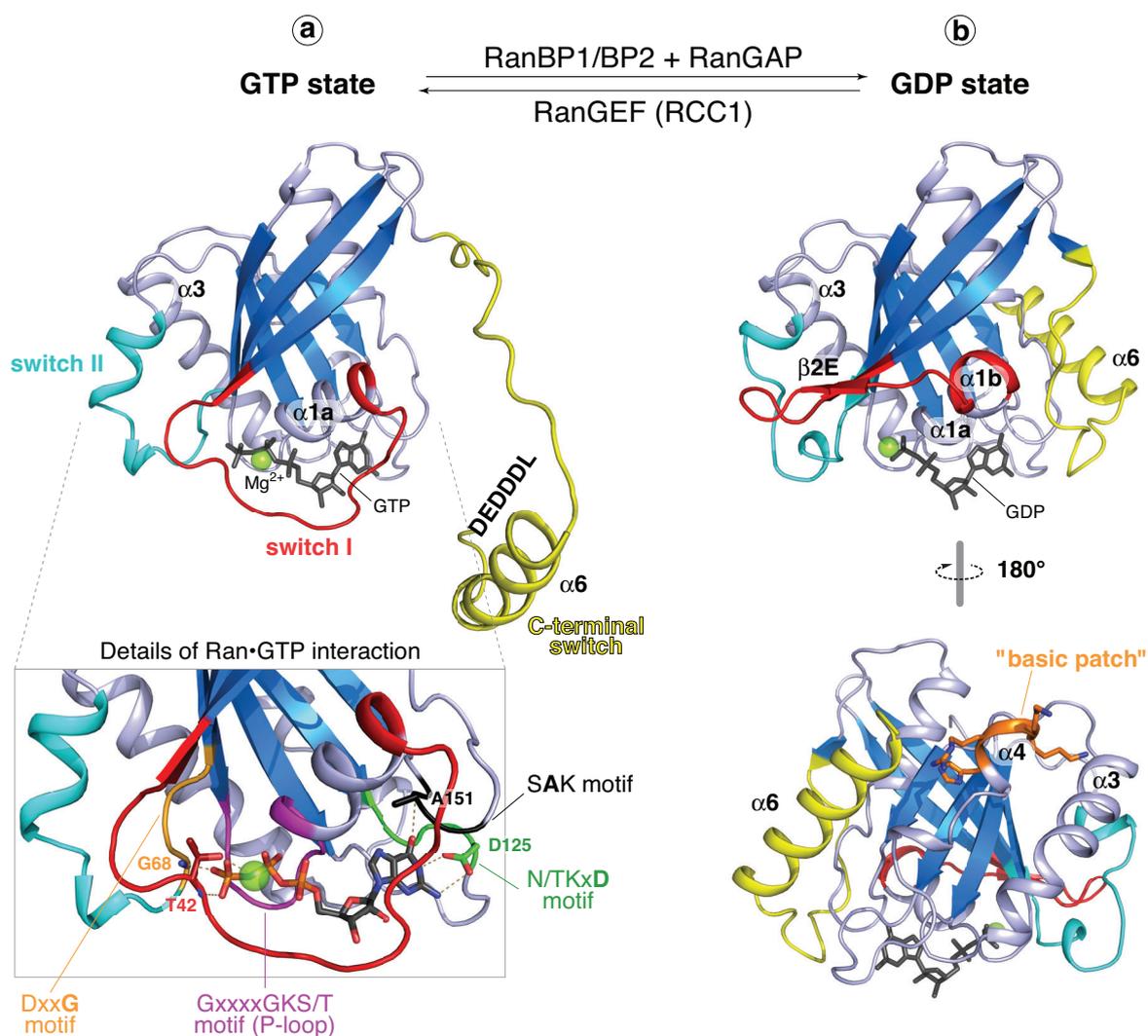
The intra-repeat and inter-repeat contacts establish a continuous hydrophobic core. One can distinguish two arches of the superhelix that are sometimes referred to as "domains" – an N-terminal arch, whose B-helices enwrap Ran, and a C-terminal arch that, generally speaking, recruits the transport cargo, also via its B-helices (**Figure 1-3b**). In contrast, the interaction with NPC components (such as Phe/Gly-rich, so-called "FG repeats") is normally accomplished by hydrophobic pockets formed by the (outer) A-helices (Bayliss *et al.*, 2000). The interactions of Imp $\beta$ -like NTRs with NPC components are not in the focus of this thesis and I would therefore like to refer the reader to recent review articles discussing this topic (Stewart, 2007; Hutten and Kehlenbach, 2007; Wälde and Kehlenbach, 2010). One can think of an NTR-superhelix as of a tightly coiled spring with each one of its turns corresponding to

one HEAT repeat (Stewart, 2003; Stewart, 2007). Such repetitive structure is inherently flexible and this flexibility appears to contribute substantially to cargo recognition, cargo loading and cargo release (reviewed by Conti *et al.*, 2006; see also Cansizoglu and Chook, 2007). It also explains, at least in part, why Imp $\beta$  can pair with a wealth of structurally unrelated cargo proteins (compare Cingolani *et al.*, 1999 and Lee *et al.*, 2003 for a truly striking example). Principles of these processes will be introduced below and in **Chapters 3 to 5**.

### 1.3.1 *The guanine nucleotide-binding protein Ran*

As described above, Ran serves as the master regulator of NTR-cargo interaction. The GTP-driven switch of Ran and Ran-related proteins has been extensively studied (Vetter and Wittinghofer, 2001). **Figure 1-4a** shows the first structure solved for RanGTP (Vetter *et al.*, 1999b). Ran's principal structural feature is the conserved, so-called "G domain". It consists of a central six-stranded  $\beta$ -sheet that is stabilized by packing of five peripheral  $\alpha$ -helices. Nucleotide binding is mainly accomplished by the G domain's loops (**Figure 1-4a**). These loops contain several sequence motifs or residues that are highly conserved among *guanine nucleotide-binding proteins (GNBPs)*. The nucleotide's  $\beta$  and  $\gamma$ -phosphates are coordinated by a network of hydrogen bonds involving a  $Mg^{2+}$  ion. Most importantly, the  $\gamma$ -phosphate is hydrogen-bonded to the backbone of two invariant residues –Thr 42 and Gly 68 (**Figure 1-4a**). These two hydrogen bonds can be imagined holding the G domain in a  $\gamma$ -phosphate-dependent "spring-loaded" conformation.





**Figure 1-4: The guanine nucleotide switch of Ran.**

(a) Ran is shown in its GTP state, as described for the RanGTP-RanBP2/RanBD1 complex (Vetter *et al.*, 1999b; PDB-ID 1RRP). **Upper:** The central  $\beta$ -strands of Ran are shown in dark blue, surrounding  $\alpha$ -helices and loops are depicted in light blue. Those parts of Ran that undergo marked conformational changes during nucleotide switching are highlighted: switch I (residues 30–47) is shown in red, switch II (residues 65–80) is shown in cyan and the C-terminal switch (residues 177–end) is colored in yellow. It should be noted that the conformation of the RanGTP C-terminus appears to depend on its binding partners. Here, it packs against RanBP2/RanBD1, which has been omitted for clarity. The very C-terminus of Ran (the "DEDDDL motif") was not completely resolved in the electron density map and had therefore not been added to the model. GTP is shown as sticks, the Mg<sup>2+</sup> ion is shown as a green sphere. See text for details. Secondary structure elements referred to in the text are labeled. **Lower:** Close-up view of the nucleotide-binding region of Ran. Coloring is as in the upper panel, conserved motifs and residues referred to in the text are colored as indicated. Details shown in stick representation are colored according to atoms. Oxygen is light red, nitrogen is blue and phosphate is orange. Interactions mentioned are indicated by dotted lines. Specificity for guanine nucleotides is conferred by the N/TKxD and SAK motifs that interact with the nucleotide's base (conserved residues shown in bold). The nucleotide's  $\alpha$  and  $\beta$ -phosphates are positioned by the so-called *P-loop* (phosphate-binding loop) that harbors the GxxxxGKS/T motif.

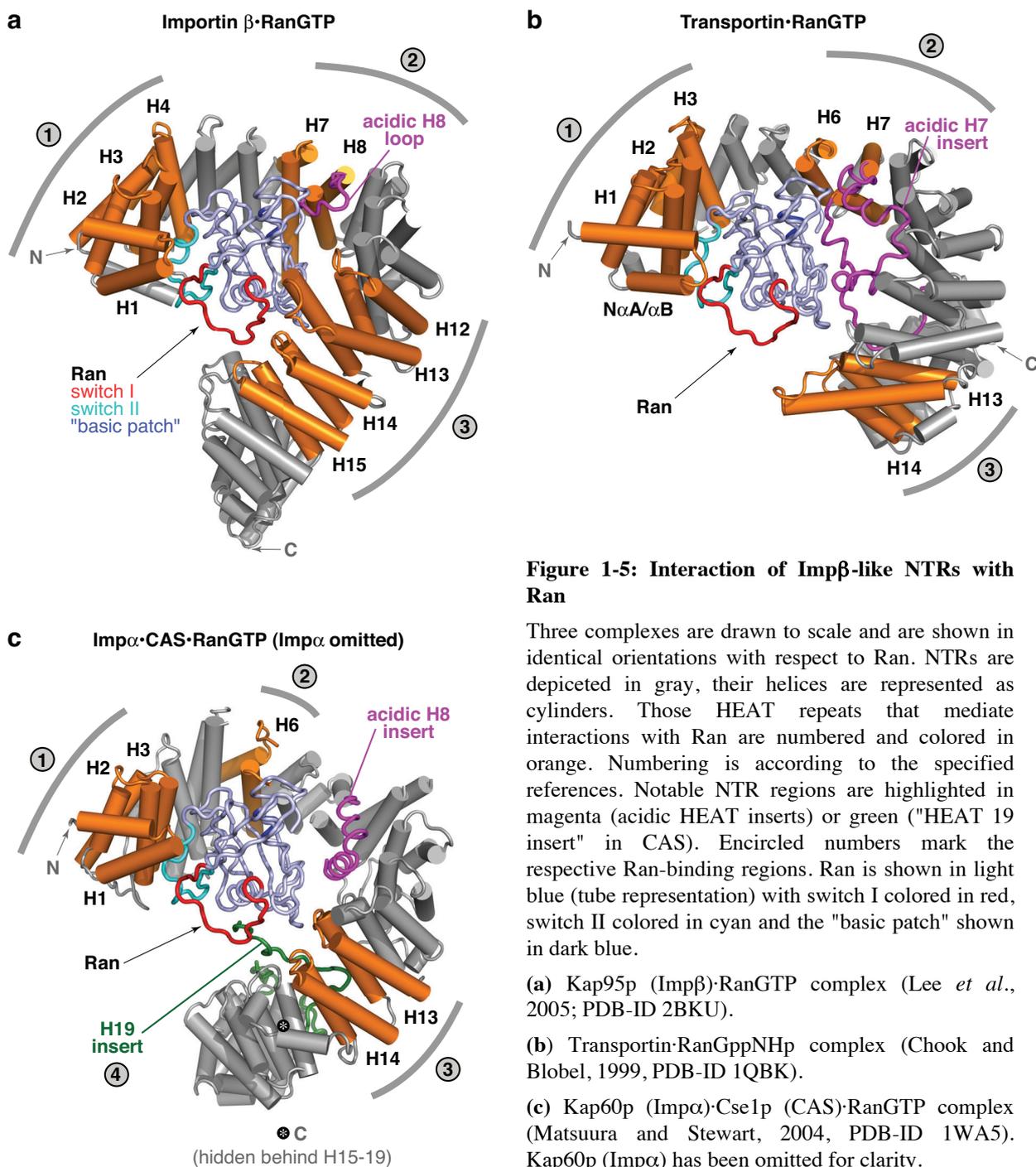
(b) Ran is shown in the GDP state (Scheffzek *et al.*, 1995; Partridge and Schwartz, 2009; PDB-ID 3GJ0). Coloring is as in panel (a) (upper). **Upper:** View as in panel (a). Note the dramatic structural rearrangements of Ran' switch regions that occur upon hydrolysis of a single phosphodiester bond in GTP. See text for details. **Lower:** "Back" view of Ran. The ultimate C-terminus of Ran (the DEDDDL motif) was not resolved in the electron density map, but it likely packs against a conspicuously basic patch of Ran (residues 139–142, shown as orange sticks with nitrogen in blue).

Hydrolysis of a single phosphodiester bond (upon GTPase activation) triggers a dramatic conformational switch, causing Ran to relax to its GDP state (Scheffzek *et al.*, 1995; Partridge and Schwartz, 2009; **Figure 1-4b**). While the core of the G domain remains essentially unaltered, the so-called "switch regions" rearrange substantially (Milburn *et al.*, 1990; Scheffzek *et al.*, 1995; Vetter *et al.*, 1999b). Switch I relocates almost entirely. Helix  $\alpha 1a$  partially melts while a new helical stretch ( $\alpha 1b$ ) and a short  $\beta$ -strand ( $\beta 2E$ ) form further downstream. The conformational changes associated with switch II are more subtle, but (nevertheless) functionally important (see also below): in the GTP state, its Gln<sup>69</sup> side chain (which is crucial for GTP hydrolysis) is in close proximity to the GTP  $\gamma$ -phosphate. There are also changes in the positioning and coordination of the Mg<sup>2+</sup> ion. The hydrolysis of GTP renders these conformational changes essentially irreversible until RanGEF/RCC1 displaces GDP, allowing GTP to re-bind. The so-called "C-terminal switch" (residues 177-216), which, among the Ras superfamily members, is unique to Ran, experiences by far the most drastic rearrangements when Ran changes its nucleotide state (Scheffzek *et al.*, 1995; Vetter *et al.*, 1999b). Given that the C-terminal switch is not part of Ran's primary nucleotide sensor, these molecular acrobatics are particularly intriguing. The C-terminal switch consists of a long, rather extended linker, followed by an  $\alpha$ -helix ( $\alpha 6$ ) and the highly acidic C-terminal motif (DEDDDL, Ran<sup>211-216</sup>). In the GDP state, the extended linker region is close to the G domain and even contacts switch I and the preceding helix ( $\alpha 1a$ , **Figure 1-4b**). Together, these interactions appear to stabilize RanGDP, which is consistent with the observation that Ran's affinity for GDP is  $\approx 10$ -fold higher than that for GTP (Klebe *et al.*, 1995). The C-terminal  $\alpha 6$ -helix packs against the "back" of Ran while the DEDDDL motif probably contacts the so-called "basic patch". Upon GDP-to-GTP exchange, switch I re-shapes into its above-described triphosphate form, which would clash with the linker of Ran's C-terminal switch. Thereby, it can contribute to the displacement of the entire C-terminal switch from the G domain (Richards *et al.*, 1995; Hieda *et al.*, 1999; Vetter *et al.*, 1999b). This is consistent with the observation that monoclonal antibodies directed against the linker region of the C-terminal switch preferentially recognize the triphosphate state of Ran (Richards *et al.*, 1995). As you will see below, this displacement and the resulting "unlocking" of the "basic patch" are crucial for Ran binding to Imp $\beta$ -like NTRs.

### 1.3.2 Interaction of Importin $\beta$ -like NTRs with RanGTP and cargo

To operate as unidirectional cargo transporters, Imp $\beta$ -like NTRs must faithfully discriminate the GDP-bound form of Ran from the GTP-bound species. Indeed, they bind RanGTP with a  $\approx 1000$  times higher affinity than RanGDP, which makes them efficient sensors of the RanGTP gradient. NTRs do not directly sense the nucleotide, but instead, very elegantly probe those regions of Ran that differ most between the nucleotide states – the switch loops I and II. Ran's C-terminal switch does not contribute to NTR binding, but plays a key role in the disassembly of NTR·RanGTP complexes (see **Chapter 5**). NTRs also contact Ran at several "invariant" loops and features of its "back side" (including the "basic patch", see below).

The comparison of the RanGTP complexes of Imp $\beta$ , Transportin and CAS (**Figure 1-5**) reveals that all these NTRs enwrap Ran with their N-terminal arches (Vetter *et al.*, 1999a; Chook and Blobel, 1999; Matsuura and Stewart, 2004; Lee *et al.*, 2005). In all cases, three distinct HEAT repeat regions contribute. The receptors' N-termini (which are most conserved among NTRs; Görlich *et al.*, 1997; Fornerod *et al.*, 1997b; Petosa *et al.*, 2004) constitute the first Ran-binding region. This area interacts with switch II and also contacts  $\alpha 3$  on the "back" of Ran (see **Figure 1-4**). It is also near switch I, and for Transportin and CAS, there are indeed some relevant contacts. Region 2 extends over Ran's back and shields, among others, the "basic patch" of Ran. In RanGDP, large parts of this contacted area are held in check by Ran's C-terminal switch (**Figure 1-4**) and would hence be inaccessible for a transport receptor. The third region binds those loops of Ran that are involved in holding the nucleotide's guanine base (**Figure 1-4a**). Most importantly, however, Imp $\beta$  also contacts switch I via this area. This interaction has not been described for Transportin, but slight conformational changes of the intra-HEAT 13 loop would be sufficient to establish such a contact. In the case of CAS, equivalent interactions are definitely absent, but here a very peculiar loop inserted into HEAT 19 contacts switch I. Another type of HEAT repeat insertion is noticeable from **Figure 1-5** – the acidic insertions into HEATs 7 or 8. In Imp $\beta$  and Transportin, they are part of contact area 2 and contribute to binding the "back" of Ran, including its "basic patch".



**Figure 1-5: Interaction of Imp $\beta$ -like NTRs with Ran**

Three complexes are drawn to scale and are shown in identical orientations with respect to Ran. NTRs are depicted in gray, their helices are represented as cylinders. Those HEAT repeats that mediate interactions with Ran are numbered and colored in orange. Numbering is according to the specified references. Notable NTR regions are highlighted in magenta (acidic HEAT inserts) or green ("HEAT 19 insert" in CAS). Encircled numbers mark the respective Ran-binding regions. Ran is shown in light blue (tube representation) with switch I colored in red, switch II colored in cyan and the "basic patch" shown in dark blue.

(a) Kap95p (Imp $\beta$ )-RanGTP complex (Lee *et al.*, 2005; PDB-ID 2BKU).

(b) Transportin-RanGppNHp complex (Chook and Blobel, 1999, PDB-ID 1QBK).

(c) Kap60p (Imp $\alpha$ )-Cse1p (CAS)-RanGTP complex (Matsuura and Stewart, 2004, PDB-ID 1WA5). Kap60p (Imp $\alpha$ ) has been omitted for clarity.

N- and C-termini of NTRs are labeled. See text for further details.

For Transportin, these interactions are very elaborate and also involve Ran's guanine-binding loops. In CAS, the reported HEAT 8 insert does not contact Ran. In fact, the entire Ran-binding region 2 of CAS is not very pronounced, but here, also the cargo (which has been omitted in **Figure 1-5c**) contributes to the formation of the CAS-RanGTP complex. This cooperativity mechanism is discussed in **Chapter 5**.

With the notable exceptions of the PTHrP ncNLS (see **1.2.3**; Cingolani *et al.*, 2002) and c-Fos (Arnold *et al.*, 2006), Imp $\beta$  and Transportin cargoes mostly rely on the NTRs' C-terminal arches for binding (Cingolani *et al.*, 1999; Lee *et al.*, 2003; Lee *et al.*, 2006; Imasaki *et al.*, 2007; Wohllwend *et al.*, 2007; Mitrousis *et al.*, 2008). This raises the question as to how Ran triggers efficient import cargo release. For Imp $\beta$  two mechanisms have been proposed (Vetter *et al.*, 1999a; Lee *et al.*, 2005). The first mechanism involves a steric clash between Ran and cargo: the cargo- and Ran-binding sites partially overlap in the region that connects the two arches. For instance, both Ran and the IBB domain of Imp $\alpha$  bind Imp $\beta$ 's acidic H8 loop. However, this direct competition appears to be insufficient for productive disassembly of the import complex (Lee *et al.*, 2005). A second, allosteric mechanism is required for complete release: the interactions of Ran with region 3 of Imp $\beta$  (**Figure 1-5a**) increase the helicoidal pitch of the importin, which ultimately expels the IBB domain. Consistent with that (and quite impressively), the disruption of that region 3 interface (by point mutations in Ran) inhibits RanGTP-driven release of the IBB domain from Imp $\beta$  but does not prevent RanGTP binding to the importin. For Transportin, the contacts of Ran with region 3 are less extensive, but here yet another allosteric cargo release strategy is employed, involving Transportin's strikingly long acidic H7 insert. In the cargo-bound state, the H7 insert appears to be disordered (Lee *et al.*, 2006; Imasaki *et al.*, 2007), but upon Ran binding, it is "forced" to a path that blocks Transportin's cargo-binding site in the C-terminal arch (**Figure 1-5b**, Chook and Blobel, 1999). In contrast to Imp $\beta$ , cargo-bound and RanGTP-bound Transportin are virtually indistinguishable by their overall shapes.

How precisely Ran contacts exportins and thereby promotes export cargo loading was not understood at the beginning of my doctoral studies (with the exception of CAS, see **Figure 1-5c** and **Chapter 5**). The work presented in this thesis provides unprecedented insight into Ran-dependent CRM1 export complex formation. For the sake of simplicity, the role of Ran in nuclear export is discussed further in **Chapter 5**.

## 1.4 About this work

This thesis addresses various central aspects of exportin function. The work presented in **Chapter 2** (Gontan *et al.*, 2009) has been performed in collaboration with Dr. Raymond Poot, Cristina Gontan and Dr. Robbert Rottier (Rotterdam, Netherlands). We discovered that Exportin 4 not only carries cargoes from the nucleus to the cytoplasm, but also ferries proteins in the opposite direction (**Table 1-1**). This revealed a fully unanticipated and novel nuclear import pathway. The study focuses on the import of Sox family transcription factors, which constitute an important class of developmental regulators. Sox proteins have been implicated in various human diseases (see **Chapter 2**). This work is also highly relevant in light of the recent proposal that a change in the expression pattern of Imp $\alpha$  isoforms (called "Imp $\alpha$  subtype switching") would trigger stem cell differentiation through regulated import of certain transcription factors, including Sox2 (Yasuhara *et al.*, 2007). We found that the import activity of Importin- $\alpha$ s towards Sox2 is actually negligible compared to that of Exp4. Instead, Sox2 is imported not only by Exp4, but also by the Imp $\beta$ /7 heterodimer and Imp9, which are known to mediate the constitutive nuclear import of essential proteins such as histones and ribosomal components (**Table 1-1**). Mutation of conserved amino acids in the import signal of Sox2 disrupted import by Exp4, Imp $\beta$ /7 and Imp9. The fact that Sox2 is imported by at least three (constitutively active) importins clearly excludes the possibility that Sox2 import can be regulated by activating or deactivating an individual pathway.

The structural characterization of the cell's export workhorse CRM1 had been a major goal since its description as an export factor in 1997 (see **1.2.2**). However, the atomic details of CRM1 had been elusive up to this work (with the exception of a C-terminal fragment; Petosa *et al.*, 2004). The intrinsic flexibility of CRM1 and the instability of its cargo complexes (and hence the apparent impossibility to purify them to homogeneity) were regarded as principal obstacles towards structure determination (Cook *et al.*, 2007). **Chapter 3** is dedicated to the first structural characterization of a CRM1 export complex – the Snurportin1·CRM1·RanGTP complex (Monecke *et al.*, 2009). This has been a joined project with Prof. Ralf Ficner, Thomas Monecke, Dr. Piotr Neumann and Dr. Achim Dickmanns. The structure shows that CRM1 not only has a rather peculiar overall shape as compared to the other characterized family members, but also revealed that CRM1 contacts RanGTP and cargo in an unprecedented manner. Our structural data and biochemical analysis show that Ran promotes cargo binding solely by an allosteric mechanism, i.e. through long-range conformational changes in the exportin. Our study also provides insight into how cargo loading onto the

import adapter Snurportin1 is controlled. Snurportin is indeed a very special cargo in that it contacts CRM1 with an extensive tripartite export signature that also includes features of a folded domain.

Much simpler CRM1-dependent export determinants are the "Leucine-rich" nuclear export signals (NESs). NES-like sequence patterns are also frequently found in proteins that are not to be exported by CRM1. For the Abelson tyrosine kinase we uncovered that its previously proposed NES is functional in isolation, but not in the context of the protein's C-terminal domain, where  $\Phi$  residues contribute to the hydrophobic core (Hantschel *et al.*, 2005). This work, which was part of a collaboration with Prof. Giulio Superti-Furga and Dr. Oliver Hantschel (CeMM, Wien) and the group of Prof. Michael Sattler (Technical University and Helmholtz Zentrum München), highlights the fact that careful considerations on protein structure are instrumental for predicting actual NESs. This will be discussed at the end of **Chapter 4**.

As described above, NESs can vary greatly in their  $\Phi$  residues and the length and sequence of their inter- $\Phi$  spacers. How CRM1 can recognize such a diversity of NES peptides has been a fascinating but unresolved question of molecular recognition. This question will be addressed in **Chapter 4** (Güttler *et al.*, 2010, accepted for publication). We tackled the problem by conceiving a crystallographic approach from the structure of the Snurportin1·CRM1·RanGTP complex. In collaboration with Prof. Ralf Ficner, Dr. Piotr Neumann, Thomas Monecke and Dr. Achim Dickmanns, we solved crystal structures of the Ran·CRM1 complex alone and when paired with the classic PKI or HIV-1 Rev NESs. Despite their drastically different  $\Phi$  spacings, Ran-bound CRM1 recognizes these NESs with the very same set of five distinct  $\Phi$  pockets. These pockets are rigid, i.e. they do not adapt to the NESs. Instead, variable  $\Phi$  spacings in the NESs are compensated for by different conformations of the bound NES peptides. In collaboration with Prof. Michael Sattler, Dr. Tobias Madl and Lorenzo Corsini (Technical University and Helmholtz Zentrum München) we employed NMR spectroscopy to analyze the PKI NES in its free state and when bound to CRM1. This part of the work sheds light on the actual process of NES recognition and suggests that CRM1 selects NES conformers that pre-exist in solution. These structural data and a rigorous biochemical and functional analysis explain the enormous flexibility in CRM1·NES recognition. Certain viral proteins harbor so-called "supraphysiological" export signals - NESs of exceptionally high CRM1 affinity (Engelsma *et al.*, 2004; Engelsma *et al.*, 2008). These proteins strongly compete CRM1-mediated export and might thus represent a viral strategy to exploit or even

disable the host cell's infrastructure. Our data lead to a new structure-based NES consensus that provides the basis for predicting the affinities of NESs for CRM1.

**Chapter 2** and **Chapter 3** are compilations of journal reprints and supplementary information. **Chapter 4** contains our recently completed manuscript on CRM1·NES recognition. In **Chapter 5**, I discuss our findings in light of the recent progress in the structural characterization of exportins. The reconstitution of export complexes with various isotope labeling schemes for NMR spectroscopy had been a major effort of my doctoral work and will therefore be described in a separate appendix to this work (**Chapter 6**). This section also contains a reprint of the above-mentioned work on the Abelson tyrosine kinase (Hantschel *et al.*, 2005) as well as additional data. My contributions to the individual studies are listed on page 148.

## CHAPTER 2

**Exportin 4 mediates a novel nuclear import pathway**

*This work has been originally published in the following article:*

**Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors**

Cristina Gontan\*, Thomas Güttler\*, Erik Engelen, Jeroen Demmers, Maarten Fornerod, Frank G. Grosveld, Dick Tibboel, Dirk Görlich, Raymond A. Poot and Robbert J. Rottier

**Journal of Cell Biology**, 2009, Volume 185, Pages 27-34.

\* These authors contributed equally to this work.

# Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors

Cristina Gontan,<sup>1</sup> Thomas Güttler,<sup>4</sup> Erik Engelen,<sup>2</sup> Jeroen Demmers,<sup>3</sup> Maarten Fornerod,<sup>5</sup> Frank G. Grosveld,<sup>2</sup> Dick Tibboel,<sup>1</sup> Dirk Görlich,<sup>4</sup> Raymond A. Poot,<sup>2</sup> and Robbert J. Rottier<sup>1</sup>

<sup>1</sup>Department of Pediatric Surgery, <sup>2</sup>Department of Cell Biology, and <sup>3</sup>Proteomics Center, Erasmus Medical Center, Rotterdam, Netherlands

<sup>4</sup>Max-Planck-Institut für Biophysikalische Chemie, Göttingen, Germany

<sup>5</sup>Netherlands Cancer Institute, Amsterdam, Netherlands

**S**RY and other Sox-type transcription factors are important developmental regulators with various implications in human disease. In this study, we identified Exp4 (exportin 4) as an interaction partner of Sox2 in mouse embryonic stem cells and neural progenitors. We show that, besides its established function in nuclear export, Exp4 acts as a bona fide nuclear import receptor for Sox2 and SRY. Thus, Exp4 is an example of a nuclear transport receptor carrying distinct cargoes into different

directions. In contrast to a published study, we observed that the import activity of Imp- $\alpha$  (importin- $\alpha$ ) isoforms toward Sox2 is negligible. Instead, we found that Imp9 and the Imp- $\beta$ /7 heterodimer mediate nuclear import of Sox2 in parallel to Exp4. Import signals for the three pathways overlap and include conserved residues in the Sox2 high-mobility group (HMG) box domain that are also critical for DNA binding. This suggests that nuclear import of Sox proteins is facilitated by several parallel import pathways.

## Introduction

Sox proteins are important regulators of different stages in embryonic development such as sex determination, neural development, bone formation, and neural crest development. They share the so-called noncanonical high-mobility group (HMG) box domain, an  $\sim$ 80-amino acid domain that confers DNA-binding activity and sequence specificity (Lefebvre et al., 2007). Sox2 is required in both embryonic and extra-embryonic tissues; the lack of Sox2 results in peri-implantation lethality (Avilion et al., 2003). Sox2, together with Oct4 and Nanog, is required for embryonic stem (ES) cell self-renewal and pluripotency (Masui et al., 2007) and is one of the factors needed for reprogramming somatic cells into induced pluripotent stem cells (Takahashi and Yamanaka, 2006). Later in development, Sox2 is required for neural progenitor expansion and maintenance (Graham et al., 2003; Ferri et al., 2004), eye development (Taranova et al., 2006), development of tastebuds (Lefebvre et al., 2007), inner ear formation (Kiernan et al., 2005), and

development of the trachea and esophagus (Que et al., 2007). Sox2 mutations in humans cause anophthalmia-esophageal-genital syndrome (Williamson et al., 2006). It is believed that Sox proteins, including Sox2, exert their specificity by synergistic binding with other transcription factors to DNA. For Sox2, partners like Oct4, Pax6, and Brn2 have been found (Lefebvre et al., 2007).

In this study, we identified Exp4 (exportin 4) as an interaction partner of Sox2 in mouse ES cells and neural progenitors and show that it serves, besides its established function in nuclear export (Lipowsky et al., 2000; Kurisaki et al., 2006), as a bona fide nuclear import receptor for Sox2 and SRY. In contrast to earlier findings (Yasuhara et al., 2007), we found that the import activity of Imp- $\alpha$  (importin- $\alpha$ ) isoforms toward Sox2 is negligible when compared with Exp4. Instead, we observed that Imp9 and the Imp- $\beta$ /7 heterodimer operate as import mediators of Sox2 in parallel to Exp4, suggesting that regulation of Sox2 function is most likely not dependent on a single nuclear import mechanism. The import signals for the three pathways overlap and include conserved residues in the Sox HMG box domain that are critical for *in vivo* nuclear localization and DNA binding.

C. Gontan and T. Güttler contributed equally to this paper.

R.A. Poot and R.J. Rottier contributed equally to this paper.

Correspondence to Raymond A. Poot: r.poot@erasmusmc.nl; or Robbert J. Rottier: r.rotter@erasmusmc.nl

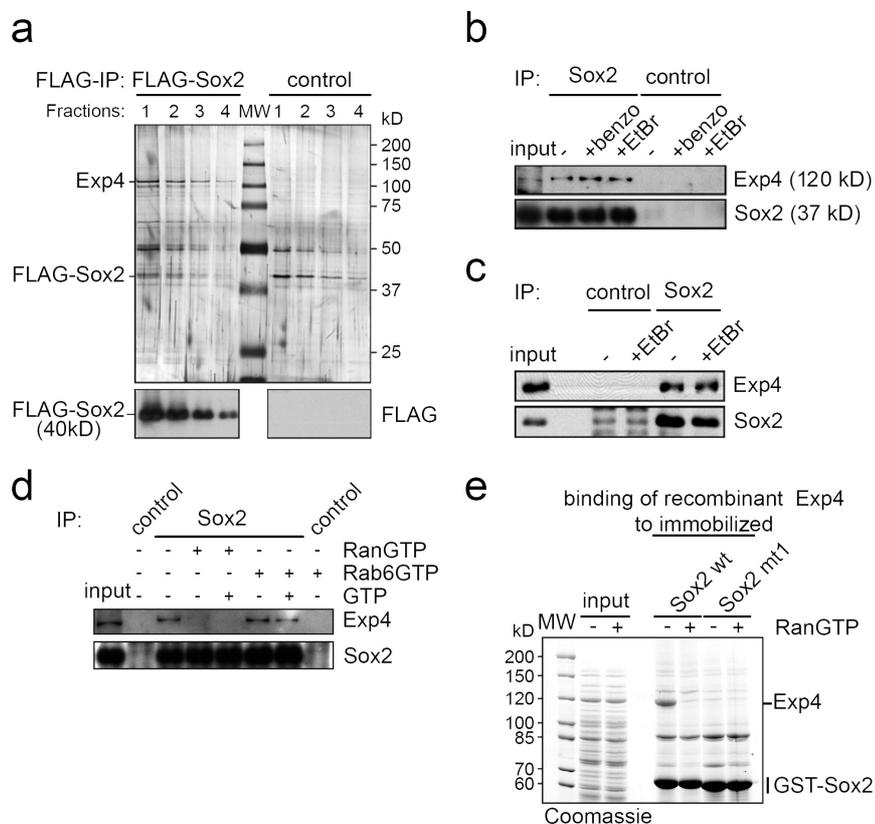
Abbreviations used in this paper: ES, embryonic stem; HMG, high-mobility group; IBB, Imp- $\beta$ -binding domain of Imp- $\alpha$ ; MBP, maltose-binding protein; NPC, nuclear pore complex.

© 2009 Gontan et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.jcb.org/misc/terms.shtml>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

Supplemental Material can be found at:  
<http://jcb.rupress.org/content/suppl/2009/04/06/jcb.200810106.DC1.html>

Published April 6, 2009

Figure 1. **Exp4 binds to Sox2.** (a) Identification of Exp4 as a Sox2 interaction partner by Flag affinity purification. (top) Flag affinity purifications from Flag-Sox2 (left) or control 46C ES cell extract (right). Eluted fractions were analyzed by SDS-PAGE and silver staining. Bands representing Flag-Sox2 and Exp4 are indicated. (bottom) Anti-Flag immunoblot showing Flag-Sox2 in the eluted fractions. (b) Exp4 coimmunoprecipitated with endogenous Sox2 from ES cell extracts. Where indicated, Benzomase (benzo) or ethidium bromide (EtBr) was added to the extract to test for DNA independence of the Sox2–Exp4 interaction. Input and bound fractions were analyzed by SDS-PAGE and immunoblotting. (c) Exp4 coimmunoprecipitated with endogenous Sox2 as in b, but from neural progenitor cell extracts. (d) Exp4 was coimmunoprecipitated with endogenous Sox2 as in b but with indicated additions. RanGTP disrupted the Sox2–Exp4 interaction, whereas the negative control, Rab6GTP, had no effect. (e) Recombinant Sox2 and Exp4 interact directly. Immobilized GST-Sox2, but not GST-Sox2 mt1 (Fig. 4), binds recombinant Exp4 from an *E. coli* extract. RanGTP disrupted the GST-Sox2–Exp4 interaction. Input and bound fractions were analyzed by SDS-PAGE and Coomassie staining. IP, immunoprecipitation; MW, molecular weight; wt, wild type.



## Results and discussion

### Exp4 is a novel interaction partner of Sox2

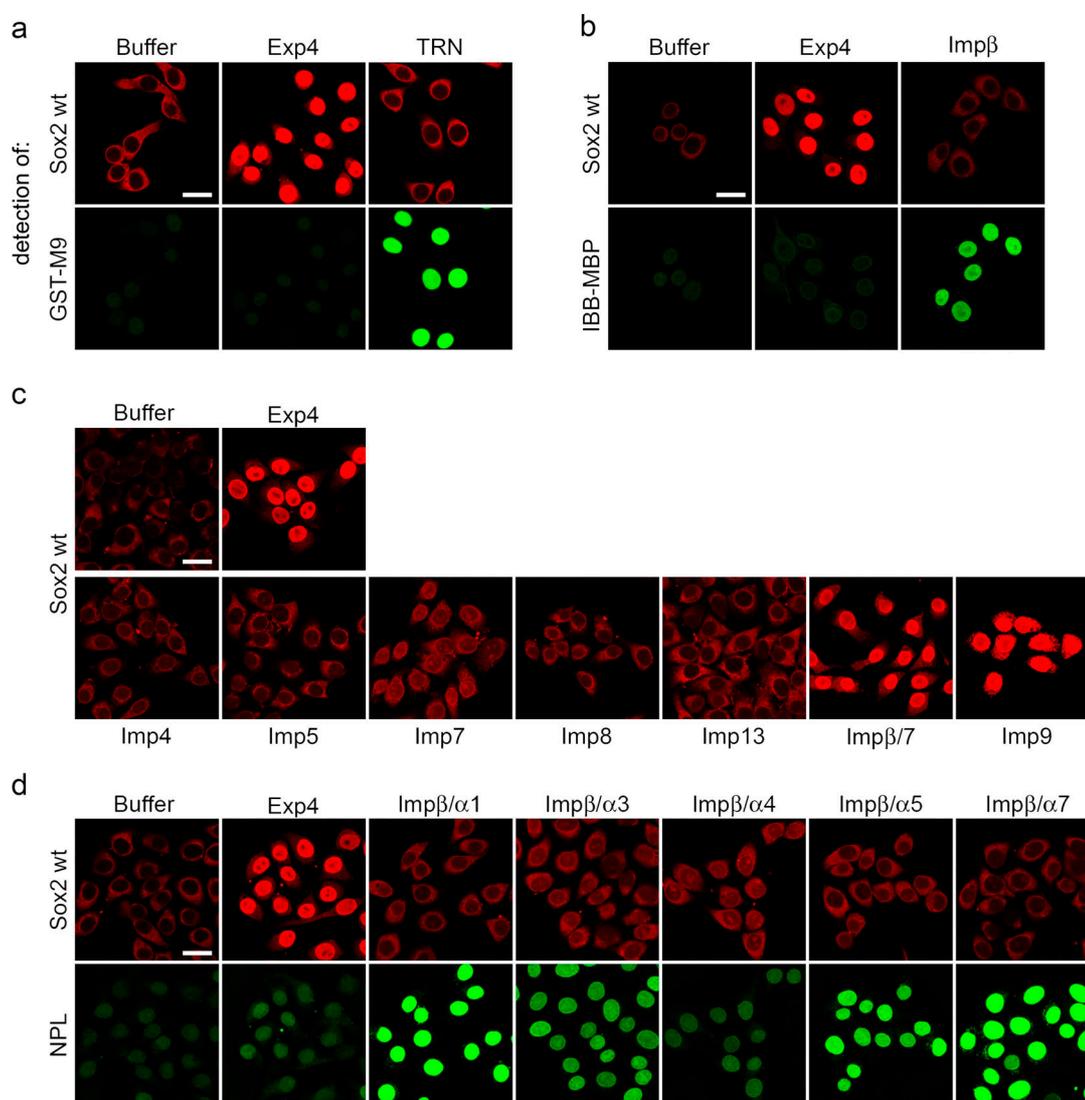
To identify Sox2 interactors by affinity purification, we introduced a Flag-Sox2 transgene into mouse ES cells and isolated individual clones stably expressing the fusion protein. A selected clone expressed Flag-Sox2 protein at ~30% of the endogenous Sox2 level (Fig. S1 a, bottom left). These cells displayed normal ES cell growth behavior and expressed the ES cell marker Oct4 (Fig. S1 a). We then performed anti-Flag immunoprecipitation experiments with extracts of both Flag-Sox2 and control ES cells and analyzed the bound fractions. One of the predominant Sox2-binding partners, migrating at ~120 kD (Fig. 1 a), was identified by mass spectrometry as Exp4 (39 unique peptides). We confirmed the identity of Exp4 by Western blotting (not depicted) and showed that endogenous (untagged) Sox2 also interacts with Exp4 (Fig. 1 b). We also analyzed anti-Flag immunoprecipitates from extracts of Flag-Sox2 ES cell-derived neural progenitors (Fig. S1 b; Conti et al., 2005) by mass spectrometry and again found the prominent presence of Exp4 (27 unique peptides), which was confirmed by coimmunoprecipitation of Exp4 with endogenous Sox2 (Fig. 1 c). The addition of Benzomase or ethidium bromide did not abolish the interaction (Fig. 1, b and c), which suggests that binding is independent of DNA.

Exp4 is a nuclear export receptor for eIF5A and the transcription factor SMAD3 (Lipowsky et al., 2000; Kurisaki et al., 2006). It belongs to the superfamily of Ran-regulated nuclear transport receptors, which includes nuclear import mediators

(importins) and exportins (Mattaj and Englmeier, 1998; Görlich and Kutay, 1999; Pemberton and Paschal, 2005). Importins bind cargoes at low RanGTP levels in the cytoplasm, facilitate translocation through nuclear pore complexes (NPCs), release their load upon RanGTP binding in the nucleus, and, finally, return to the cytoplasm, where GTP hydrolysis dissociates the Ran–importin complex and allows the importin to bind and import another cargo molecule. Exportins operate in the opposite manner; they bind cargo together with RanGTP in the nucleus and release their cargo upon GTP hydrolysis in the cytoplasm. Surprisingly, the addition of RanGTP almost completely abolished binding of Exp4 to Sox2 (Fig. 1 d), suggesting that Exp4 acts as an importin for Sox2. The effect was specific for Ran, as the Rab6 GTPase did not dissociate the Sox2–Exp4 complex (Fig. 1 d). In line with this observation, activation of the RanGTPase by exogenous RanGTPase-activating protein and RanBP1 (Bischoff and Görlich, 1997) did not disrupt the Sox2–Exp4 interaction (Fig. S1 c).

The observed Sox2–Exp4 binding could have been either direct or mediated by some additional component from the cell extract. Therefore, we tested whether recombinant GST-Sox2 and Exp4 interact within an *Escherichia coli* extract lacking other components of the nuclear transport machinery. Indeed, GST-Sox2 bound Exp4 in the absence of recombinant RanGTP but not in its presence (Fig. 1 e). Interestingly, a Sox2 mutant (mutant 1 [mt1]; see Fig. 4 a) previously reported to be compromised in import (Yasuhara et al., 2007) did not interact with Exp4 (Fig. 1 e). Thus, Exp4 specifically and directly interacts with Sox2 in an importin-like manner.

Published April 6, 2009



**Figure 2. Exp4, Imp- $\beta$ /7, and Imp9, but not Imp- $\beta$  alone or Imp- $\beta$ / $\alpha$ , facilitate nuclear import of Sox2.** The panels show nuclear import of fluorophore-labeled GST-Sox2 (red) and control cargo proteins (green) into nuclei of digitonin-permeabilized HeLa cells with either buffer or the indicated import receptors. All import reactions contained the components of the Ran cycle and an ATP/GTP-replenishing system (see Materials and methods for details). (a) Exp4 imported GST-Sox2 into nuclei of permeabilized cells. The transportin (TRN) substrate GST-M9 was used as an internal specificity control. (b) Imp- $\beta$  failed to import GST-Sox2, whereas it efficiently imported its substrate Imp- $\beta$ -binding domain of Imp- $\alpha$  (IBB) fused to the maltose-binding protein (MBP [IBB-MBP]). (c) Apart from Exp4, only Imp- $\beta$ /7 and Imp9 efficiently imported GST-Sox2. Note that only the cognate import receptors suppressed cytoplasmic aggregation of GST-Sox2. (See also a, b, and d as well as Fig. 4 and Fig. S2 b.) (d) The panel shows GST-Sox2 import by Exp4 compared with the import activity of the indicated Imp- $\beta$ / $\alpha$  combinations. The Imp- $\beta$ / $\alpha$  substrate nucleoplasmin (NPL) served as an internal positive control. wt, wild type. Bars, 25  $\mu$ m.

### Exp4 facilitates nuclear import of Sox2

To test whether Exp4 imports Sox2, we prepared fluorophore-labeled GST-Sox2 for import assays with permeabilized HeLa cells (Adam et al., 1990; Jäkel and Görlich, 1998). In the absence of transport receptors, GST-Sox2 failed to accumulate inside nuclei but precipitated at the cytoplasmic remnants of the permeabilized cells (Fig. 2, a and b), which is similar to other nucleic acid-binding proteins such as histone H1 and ribosomal factors (Jäkel et al., 1999, 2002). However, the addition of Exp4 strongly stimulated nucleoplasmic accumulation of GST-Sox2 (Fig. 2, a and b). Remarkably, Exp4 also suppressed cytoplasmic aggregation of Sox2 (Fig. 2, a and b), which is an effect that appeared to be Exp4 dose dependent (Fig. S2, a and b, compare

Exp4 panels) but did not require Ran (Fig. S2 b; see next section). These effects were specific to Exp4, as transportin neither imported GST-Sox2 nor affected Sox2 aggregation (Fig. 2 a). As for all Imp- $\beta$ -like importins, efficient multiround nuclear import by Exp4 depended on both Ran and energy (Fig. S2 a), reflecting the need for recycling Exp4-RanGTP complexes back into the cytoplasm. Thus, Exp4 is a bona fide importin for Sox2 and mediates a novel nuclear import pathway. So far, only Imp13 from higher eukaryotes and Msn5p from *Saccharomyces cerevisiae* have been recognized to transport distinct sets of cargoes into opposite directions through the NPC (Mingot et al., 2001; Yoshida and Blobel, 2001). It is now tempting to speculate that any exportin could operate as an importin, provided its

Published April 6, 2009

affinity for RanGTP in the absence of export cargo is sufficiently high for displacing import substrates into the nucleus. We are currently testing this hypothesis. In contrast to Imp- $\beta$  but like transportin, Exp4 required RanGTP only to release its import cargo into the nucleus but not to detach from the nuclear side of the NPC (Fig. S2 b, middle and right panels; Ribbeck et al., 1999).

#### **Imp- $\beta$ /7 and Imp9 also facilitate nuclear import of Sox2**

Cargo binding by importins is highly specific and governed by NLSs. Imp- $\beta$  either binds its cargoes directly or pairs with Imp- $\alpha$  isoforms that recognize so-called monopartite or bipartite classical NLSs. Previous studies proposed that Sox proteins share a conserved bipartite NLS at the N terminus of the HMG box domain and a monopartite NLS at its C-terminal end (Figs. 3 a and 4 a; Poulat et al., 1995; Südbek and Scherer, 1997) and that Imp- $\beta$ / $\alpha$ 3 and Imp- $\beta$ / $\alpha$ 5 import Sox2 in neural progenitors (Yasuhara et al., 2007). It was also suggested that Imp- $\beta$  alone can import Sox2 in undifferentiated ES cells (Yasuhara et al., 2007). Therefore, it was fully unexpected to find Exp4 as a predominant interaction partner (Fig. 1 a) and efficient nuclear import mediator of Sox2 (Fig. 2 a). This situation prompted us to also test other members of the Imp- $\beta$  superfamily for import activity toward Sox2. Although, in the parallel incubation, Exp4 strongly stimulated nuclear accumulation of GST-Sox2, we found no indication for Imp- $\beta$ , Imp4, Imp5, Imp7, Imp8, and Imp13 mediating import of the protein (Fig. 2, b and c). However, the Imp- $\beta$ /7 heterodimer (Jäkel et al., 1999) and Imp9 imported GST-Sox2 into nuclei as efficiently as Exp4 (Fig. 2 c). We obtained virtually identical results using permeabilized ES cells instead of HeLa cells (Fig. S3; see figure legend for the rationale). Thus, Imp9 and the Imp- $\beta$ /7 heterodimer can facilitate nuclear import of Sox2 in parallel to Exp4. Importantly, just as Exp4, Imp- $\beta$ /7 and Imp9 markedly reduced cytoplasmic precipitation of GST-Sox2 (Fig. 2 c). We found no evidence for any Imp- $\alpha$ / $\beta$  combination stimulating Sox2 import, although they all imported the internal positive control, nucleoplasmin (Fig. 2 d). Only when using Imp- $\alpha$ 3 and Imp- $\alpha$ 4, GST-Sox2 showed slight nucleolar accumulation, which was not seen for endogenous or overexpressed Sox2 (unpublished data) and, thus, probably reflects mislocalized Sox2. In contrast to Exp4, none of the Imp- $\alpha$ / $\beta$  combinations suppressed cytoplasmic aggregation of GST-Sox2 (Fig. 2 d). In summary, our comprehensive survey identified three import pathways for Sox2, all of which meet the criteria for cognate import receptors, namely stimulation of nuclear import and suppression of import substrate aggregation. In contrast, Imp- $\beta$  alone and in conjunction with Imp- $\alpha$  isoforms met none of these criteria.

#### **SRY is imported by Exp4 through conserved residues of its HMG box domain**

Of the three import pathways identified, Exp4 represents a novel mediator of nuclear import, and, therefore, it is of interest to know whether Sox2 represents the sole import cargo of this pathway. Thus, we tested whether Exp4 also imports SRY, the prototypic Sox family member (Sinclair et al., 1990), which shares in its HMG box domain 85% identity with Sox2 (Fig. 3 a; Lefebvre et al., 2007) and whose loss of function leads to XY sex reversal

in humans (Lefebvre et al., 2007). Indeed, fluorophore-labeled GST-SRY behaved identically to GST-Sox2, showing pronounced nuclear accumulation upon Exp4 addition and nuclear exclusion with cytoplasmic precipitation when only buffer or Imp- $\beta$  had been added (Fig. 3 b). Consistent with this result, Flag-SRY co-immunoprecipitated Exp4 from ES cell extracts in a RanGTP-sensitive manner (Fig. 3 c).

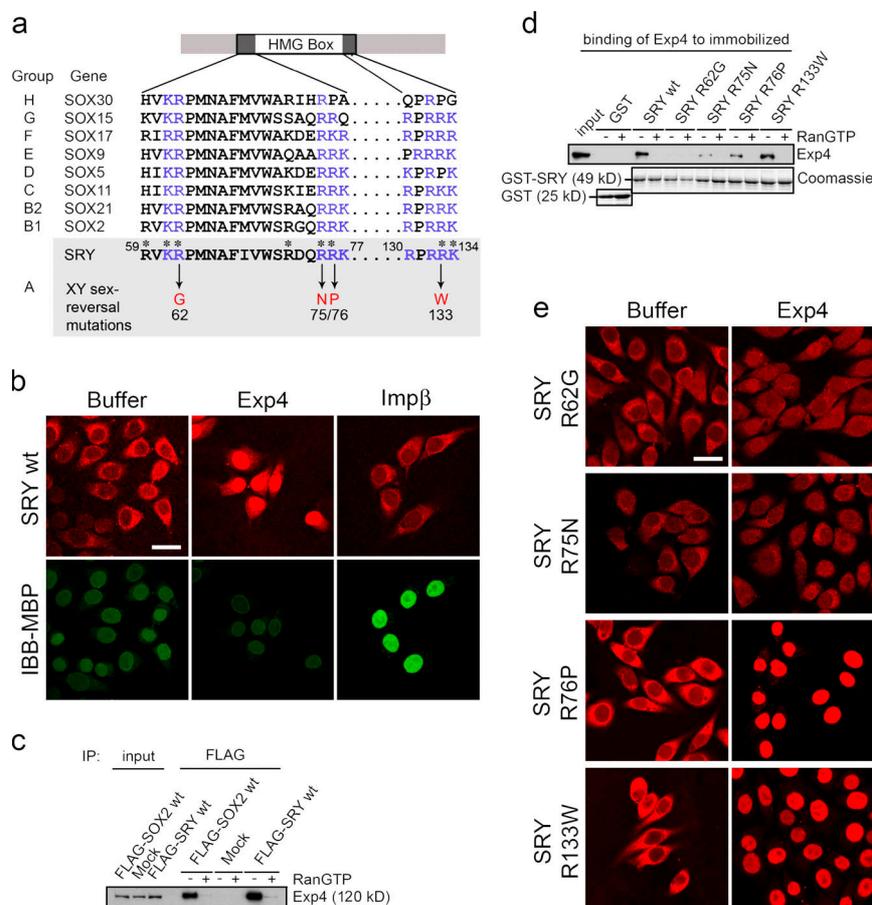
Interestingly, some SRY mutations in XY females have been reported to impair SRY nuclear localization, with R62G and R75N causing the most drastic defect, whereas R76P and R133W only display a minor (if any) change (Fig. 3 a; Harley et al., 2003). These effects were previously allocated to defects in Imp- $\beta$ -dependent import (R62G and R133W) or as yet unidentified import pathways (R75N and R76P; Harley et al., 2003). However, as our data argue against a significant contribution of Imp- $\beta$  to SRY nuclear localization (Fig. 3 b), we tested whether the SRY mutants are compromised in recruiting Exp4 from an ES cell extract. Strikingly, Exp4 binding was completely lost for the R62G mutant and strongly impaired for SRY R75N, whereas SRY R76P retained a reduced Exp4-binding activity (Fig. 3 d). R133W had no detectable effect on Exp4 binding (Fig. 3 d). This correlation with the previously observed localization defects (Harley et al., 2003) was perfectly recapitulated by our nuclear import assays: although Exp4 efficiently facilitated nuclear import of GST-SRY wild type (Fig. 3 b), nuclear import was lost for GST-SRY R62G and GST-SRY R75N but not affected by the R76P or R133W mutation (Fig. 3 e). In summary, our data suggest that Exp4 operates as a nuclear import receptor for Sox family transcription factors and that the Exp4-Sox protein recognition requires residues in the HMG box domain that are conserved in all Sox proteins (Fig. 3 a).

#### **Nuclear import signals for Exp4, Imp- $\beta$ /7, and Imp9 overlap**

Sox2 mutations within the HMG box domain have also been reported to cause localization defects. GST-Sox2 mutants mt1, mt2, and mt1.2 (Fig. 4 a) were compromised in nuclear import when microinjected into the cytoplasm of HeLa cells (Yasuhara et al., 2007). Therefore, we tested whether these mutations affect Sox2 import by Exp4, Imp- $\beta$ /7, or Imp9. Strikingly, mt1 and mt1.2 completely failed to be imported by any of these pathways, whereas for mt2, only Imp- $\beta$ /7- and Imp9-mediated import was reduced (Fig. 4 b). In line with these observations, we detected binding of Exp4 only to wild-type Sox2 and mt2, but not to mt1 or mt1.2 (Fig. 4 c). Therefore, the Exp4-dependent import signal involves parts previously shown to be important for Sox2 nuclear localization (Yasuhara et al., 2007) and overlaps with the signal recognized by Imp- $\beta$ /7 and Imp9. Given that some of the residues necessary for Sox protein import (for example, SRY R75) reside within one of the three  $\alpha$  helices of the HMG box domain (Werner et al., 1995; Murphy et al., 2001), it appears that the import signature is not a linear sequence but instead features a three-dimensional structure.

The aforementioned import determinants coincide conspicuously with DNA-binding residues (Fig. 3 a; Werner et al., 1995; Murphy et al., 2001). Thus, mutations of these residues such as those found in SRY-dependent sex-reversal conditions have at least two consequences. First, they directly inhibit DNA

Published April 6, 2009



**Figure 3. Exp4 facilitates nuclear import of SRY.** (a) Sequence alignment of N- and C-terminal conserved regions of the HMG box domain previously implicated in the nuclear localization of Sox proteins (selected members of all nine subgroups are shown; Poulat et al., 1995; Südbek and Scherer, 1997). Conserved basic residues are indicated in purple, and basic residues binding DNA in the SRY-DNA complex (Werner et al., 1995; Murphy et al., 2001) are marked with asterisks. Amino acid substitutions of conserved arginines (R62G, R75N, R76P, and R133W) in the HMG box domain of SRY, originally identified in human patients with XY sex reversal (Murphy et al., 2001), are shown in red. (b) Exp4-mediated import of SRY but not IBB-MBP. The import experiment was performed as in Fig. 2 b but with fluorophore-labeled GST-SRY (red). Imp-β imported IBB-MBP but not SRY wt, wild type. (c) Exp4 was coprecipitated with Flag-SRY in a RanGTP-sensitive manner. The panel shows Flag affinity purifications from extracts of ES cells, which were transiently transfected with the indicated Flag plasmids. Mock refers to the empty control Flag plasmid. Starting material and bound fractions were analyzed by SDS-PAGE and Western blotting. The addition of RanGTP to the extract disrupted the SRY-Exp4 interaction. IP, immunoprecipitation. (d) SRY sex-reversal mutations affected the SRY interaction with Exp4 to different degrees. The panel shows a GST pull-down from ES cell extracts with GST-SRY or the indicated mutants. (top) Eluted Exp4 was detected by SDS-PAGE and Western blotting. (bottom) Eluted GST and GST-SRY fusions were detected by SDS-PAGE and Coomassie staining. The addition of RanGTP to the extract disrupted the SRY-Exp4 interaction as in c. (e) Sex-reversal mutations affected SRY nuclear import by Exp4 to different degrees. The experiment was performed as in b but with the indicated GST-SRY sex-reversal mutants. Only a subset of the SRY mutations compromised Exp4-mediated nuclear import of SRY. Bars, 25 μm.

binding of the affected Sox protein and thereby reduce Sox-dependent transcriptional activity. Second, they inhibit its import, which aggravates this effect. Residues directly binding to DNA also appear to cause the observed cytoplasmic aggregation of Sox proteins, as the respective Sox2 mutations reduced precipitation in the absence of cognate importins (Fig. 4 b). The mutations also compromised import (Fig. 4 b), and, therefore, we propose that Exp4, Imp-β/7, and Imp9 act as chaperones for the exposed polycationic DNA-binding domain of Sox2 to prevent potentially deleterious aggregation with polyanions such as RNA (Jäkel et al., 2002). Given this important function, it is not surprising that all cognate importins identified in this study recognize similar features of Sox2.

Sox2-dependent transcription is essential at different stages of embryonic development, including the maintenance of the inner cell mass, from which ES cells are derived (Avilion et al., 2003), and neural progenitor expansion (Ferri et al., 2004). A recent study proposed that selective up-regulation of Imp-α3 and Imp-α5 would trigger the differentiation of ES cells into neural progenitors by mediating import of certain transcription factors such as Sox2 (Yasuhara et al., 2007). Our finding that the import

activity of Imp-α isoforms toward Sox2 is negligible compared with that of Exp4, Imp9, or the Imp-β/7 dimer argues against the scenario that up-regulated Imp-α isoforms facilitate nuclear localization of Sox2 during neural differentiation. Instead, our data are in agreement with the observation that Imp-α5-deficient mice undergo normal brain development (Shmidt et al., 2007). It was suggested that up-regulated Imp-α4 expression may compensate for the loss of Imp-α5 (Shmidt et al., 2007); however, we found no supporting evidence for Sox2 import activity of any Imp-α isoform (Fig. 2 d). Sox2 does not rely on a single nuclear import pathway but can use three pathways, all of whose import mediators are expressed in ES cells (Fig. 1 and not depicted). At least Imp-β/7 and Imp9 import essential mass proteins such as histones and ribosomal proteins (Jäkel et al., 1999, 2002; Mühlhäusser et al., 2001); thus, they should represent constitutive import pathways that do not become limiting under physiological conditions. Whether other transcription factors required for neural differentiation can also access multiple, possibly constitutive import pathways remains to be tested. At least for Sox2, it appears very unlikely that import regulation occurs by up- or down-regulating a specific nuclear import pathway. With a single known exception

Published April 6, 2009

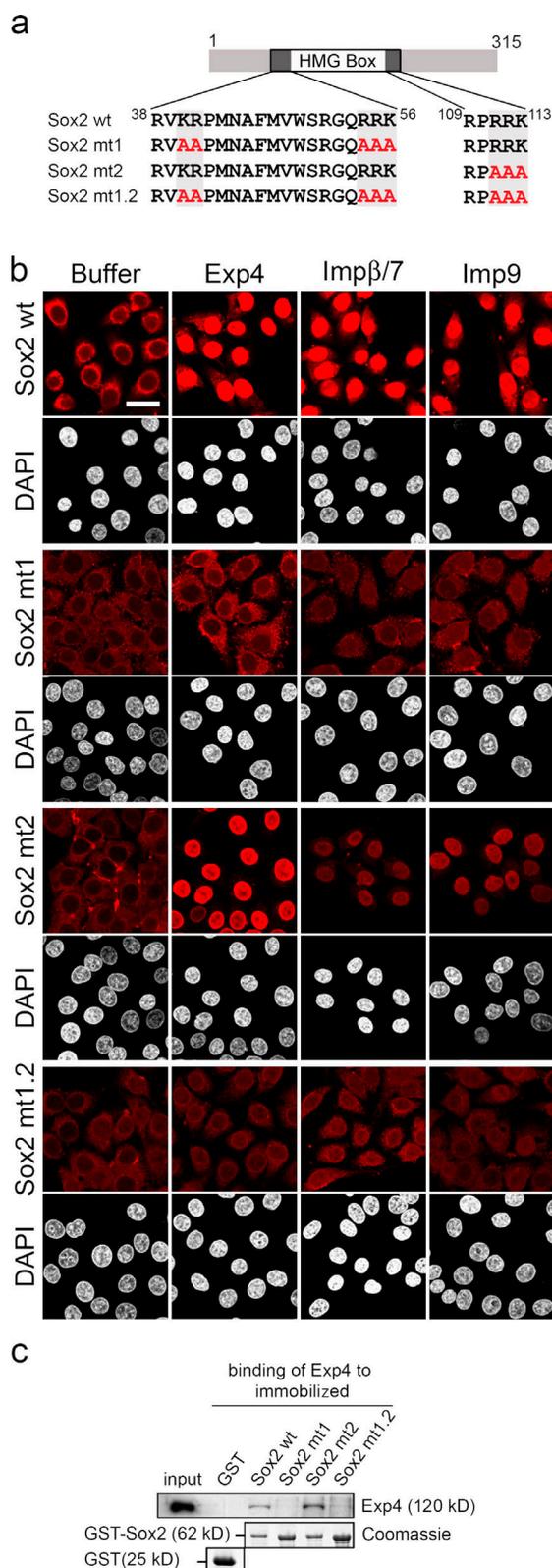


Figure 4. **The Exp4-, Impβ/7-, and Imp9-dependent nuclear import signals overlap.** (a) Schematic representation of N- and C-terminal regions of the HMG box domain that have been previously implicated in nuclear localization of Sox2 (Fig. 3 a; Yasuhara et al., 2007). The gray-shaded

(Exp6, having only one substrate in *Xenopus laevis* oocytes and eggs; Bohnsack et al., 2006), regulation of transport is achieved by modifying the cargo, which is a strategy that does not affect myriads of other cargoes depending on the same pathway (Kaffman and O'Shea, 1999). If regulation of nucleocytoplasmic distribution plays a role in Sox protein function, we would expect it to follow the aforementioned paradigm.

## Materials and methods

### Antibodies and cell culture

Anti-Exp4 antibodies (Figs. 1 c, 3, and 4) against residues 1,015–1,150 of the *Xenopus* protein were raised in rabbits and affinity purified via the antigen. For other figures (Fig. 1, b and d, and Fig. S1 c), commercial anti-Exp4 was used (V-18; Santa Cruz Biotechnology, Inc.). Other antibodies were against Flag (M2; Sigma-Aldrich), Sox2 (Millipore), Oct3/4 (Santa Cruz Biotechnology, Inc.), nestin, and RC2 (Developmental Studies Hybridoma Bank).

Mouse 46C ES cells (provided by A. Smith, Wellcome Trust Centre for Stem Cell Research, Cambridge, England, UK; Ying et al., 2003) and derivatives were grown under standard conditions (Yasuhara et al., 2007). Flag-Sox2 46C ES cells were generated by electroporation of a 2× Flag-tagged murine Sox2 under the control of the CAG [CMV early enhancer/chicken β-actin/globin] promoter followed by neomycin selection (pCBA vector provided by A. Tang, Medical Research Council Clinical Sciences Centre, London, England, UK). Differentiation into neural progenitors and immunostaining were performed as described previously (Conti et al., 2005). HeLa cells were grown under standard conditions (Kurisaki et al., 2006). Lipofectamine 2000 (Invitrogen) was used for transient transfections.

### Preparative Flag-Sox2 purification and mass spectrometry

Flag-Sox2 ES cell or neural progenitor nuclear extract (20 mM Hepes, pH 7.6, 10% [wt/vol] glycerol, 100 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, and 1× complete protease inhibitor [Roche]) was incubated with anti-Flag M2 agarose beads (Sigma-Aldrich). Bound material was eluted with a Flag-tripeptide (Sigma-Aldrich) and analyzed by SDS-PAGE and silver staining. Mass spectrometry analysis was performed on a capillary liquid chromatography system (NanoflowLC-MS/MS 1100 series; Agilent Technologies) coupled to a mass spectrometer (LTQ-Orbitrap; Thermo Fisher Scientific). Data analysis was performed as described previously (Sanchez et al., 2007).

### Bacterial protein expression and purification

GST fusions of human SRY (subcloned into a modified pGEX-6P-1 plasmid; GE Healthcare; DNA construct provided by V. Harley, Human Molecular Genetics Laboratory, Clayton, Australia) and chicken Sox2 (both C-terminally His<sub>6</sub> tagged; DNA construct provided by Y. Yoneda, Osaka University, Osaka, Japan) were expressed in a protease-deficient *E. coli* strain and purified by nickel chelate affinity and gel filtration chromatography. The gel filtration step was essential to keep the background fluorescence in import assays sufficiently low to assess specific import effects. Alexa Fluor 488- and 568-maleimide labeling was performed essentially as previously described (Jäkel and Görlich, 1998). Other recombinant proteins were prepared as previously described (Pollard et al., 1996; Jäkel and Görlich, 1998; Köhler et al., 1999; Lipovsky et al., 2000; Dean et al., 2001; Mingot et al., 2001; Jäkel et al., 2002; Ribbeck and Görlich, 2002).

regions mark basic amino acid residues that were substituted for alanines (marked in red) in mt1, mt2, and mt1.2 (Yasuhara et al., 2007). wt, wild type. (b) Exp4-, Impβ/7-, and Imp9-mediated Sox2 nuclear import is affected by the previously described Sox2 mutations (Yasuhara et al., 2007). The nuclear import assay was performed as in Fig. 2 c. Note that the GST-Sox2 mutants showed reduced cytoplasmic aggregation in the absence of importins. DNA-staining with DAPI highlights nuclei. (c) GST-Sox2 mt2 but not GST-Sox2 mt1 binds Exp4. The figure shows a GST pull-down from ES cell extracts with GST-Sox2 or the indicated Sox2 mutants. (top) Eluted Exp4 was detected by SDS-PAGE and Western blotting. (bottom) Eluted GST and GST-Sox2 fusions were detected by SDS-PAGE and Coomassie staining. Bar, 25 μm.

Published April 6, 2009

Expression plasmids for human Imp- $\alpha$  isoforms were provided by M. Köhler (University of Kiel, Kiel, Germany) and E. Hartmann (University of Lübeck, Lübeck, Germany).

#### Immunoprecipitations and binding assays

ES cell or neural progenitor nuclear extract was incubated with protein A-Sepharose (GE Healthcare) and anti-Sox2 antibody or control rabbit IgG (Santa Cruz Biotechnology, Inc.) or with anti-Flag M2-agarose (Sigma-Aldrich). Elution was performed with SDS sample buffer. Where indicated, 5 U/ml Benzonase (EMD), 25  $\mu$ g/ml ethidium bromide, 3–5  $\mu$ M RanQ69L, 5  $\mu$ M Rab6 (provided by A. Akhmanova, Erasmus Medical Center, Rotterdam, Netherlands), 1 mM GTP, 12  $\mu$ M RanGTPase-activating protein, or 6  $\mu$ M RanBP1 was added.

GST pull-downs were performed by incubating GST-Sox2 bound to glutathione-Sepharose 4B (GE Healthcare) with an ES cell nuclear extract or 1  $\mu$ M mouse Exp4-spiked *E. coli* lysate (50 mM Tris, pH 7.5, 50 mM NaCl, 2 mM Mg-acetate, 1 mM DTT, and 0.005% digitonin) in the presence of an ATP/GTP-regenerating system and, if indicated, 3  $\mu$ M RanGTP. Bound proteins were eluted with SDS sample buffer and analyzed by immunoblotting and/or Coomassie staining.

#### Import assays

Import assays were performed essentially as described previously (Adam et al., 1990; Jäkel and Görlich, 1998). Unless otherwise noted, import mixes contained 1  $\mu$ M of import cargoes and 1  $\mu$ M of transport receptors (2  $\mu$ M for Imp- $\alpha$  isoforms). Where indicated, the energy dependence of nuclear import was tested by replacing the energy-regenerating system with 0.1 U/ $\mu$ l apyrase (Sigma-Aldrich). Fixed samples were mounted in Vectashield (Vector Laboratories); imaging was performed with a confocal laser-scanning microscope system (SP5; Leica) using the 405-, 488-, and 561-nm laser lines and a 63 $\times$  NA 1.4 Plan-Apochromat oil objective (Leica). Confocal imaging was vital to minimize blurring of the strong cytoplasmic signal into nuclear regions. Figures were assembled in Photoshop or Illustrator (Adobe).

#### Online supplemental material

Fig. S1 shows the characterization of Flag-Sox2 ES cells and neural progenitor cells. Fig. S2 shows Ran and energy dependence of Exp4-mediated nuclear import of Sox2. Fig. S3 shows that Exp4, Imp- $\beta$ /7, and Imp9 also mediate nuclear import of Sox2 in permeabilized ES cells. Online supplemental material is available at <http://www.jcb.org/cgi/content/full/jcb.200810106/DC1>.

We thank Drs. Vincent Harley, Yoshihiro Yoneda, Matthias Köhler, Enno Hartmann, Amy Tang, Anna Akhmanova, and Austin Smith for providing reagents. We also thank Dr. Laura Riobos for advice on import assays, Dr. Steffen Frey for helpful comments on the manuscript, and Danilo Deichsel and Heinzjürgen Dehne for excellent technical assistance.

The Rottier laboratory is supported by the Sophia Foundation for Medical Research (SSWO project no. 413), and the Poot laboratory is supported by a Vidi grant from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek. Work in the Görlich laboratory is funded by the Max-Planck-Gesellschaft. T. Güttler was also supported by fellowships from the Boehringer Ingelheim Fonds and the Alfred Krupp Foundation.

Submitted: 16 October 2008

Accepted: 9 March 2009

## References

Adam, S.A., R.S. Marr, and L. Gerace. 1990. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J. Cell Biol.* 111:807–816.

Avilion, A.A., S.K. Nicolis, L.H. Pevny, L. Perez, N. Vivian, and R. Lovell-Badge. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* 17:126–140.

Bischoff, F.R., and D. Görlich. 1997. RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors. *FEBS Lett.* 419:249–254.

Bohnsack, M.T., T. Stuken, C. Kuhn, V.C. Cordes, and D. Görlich. 2006. A selective block of nuclear actin export stabilizes the giant nuclei of *Xenopus* oocytes. *Nat. Cell Biol.* 8:257–263.

Conti, L., S.M. Pollard, T. Gorba, E. Reitano, M. Toselli, G. Biella, Y. Sun, S. Sanzone, Q.L. Ying, E. Cattaneo, and A. Smith. 2005. Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol.* 3:e283.

Dean, K.A., O. von Ahlsen, D. Görlich, and H.M. Fried. 2001. Signal recognition particle protein 19 is imported into the nucleus by importin 8 (RanBP8) and transportin. *J. Cell Sci.* 114:3479–3485.

Ferri, A.L., M. Cavallaro, D. Braidà, A. Di Cristofano, A. Canta, A. Vezzani, S. Ottolenghi, P.P. Pandolfi, M. Sala, S. DeBiasi, and S.K. Nicolis. 2004. Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. *Development.* 131:3805–3819.

Görlich, D., and U. Kutay. 1999. Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* 15:607–660.

Graham, V., J. Khudyakov, P. Ellis, and L. Pevny. 2003. SOX2 functions to maintain neural progenitor identity. *Neuron.* 39:749–765.

Harley, V.R., S. Layfield, C.L. Mitchell, J.K. Forwood, A.P. John, L.J. Briggs, S.G. McDowall, and D.A. Jans. 2003. Defective importin beta recognition and nuclear import of the sex-determining factor SRY are associated with XY sex-reversing mutations. *Proc. Natl. Acad. Sci. USA.* 100:7045–7050.

Jäkel, S., and D. Görlich. 1998. Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO J.* 17:4491–4502.

Jäkel, S., W. Albig, U. Kutay, F.R. Bischoff, K. Schwamborn, D. Doenecke, and D. Görlich. 1999. The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *EMBO J.* 18:2411–2423.

Jäkel, S., J.M. Mingot, P. Schwarzmaier, E. Hartmann, and D. Görlich. 2002. Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J.* 21:377–386.

Kaffman, A., and E.K. O'Shea. 1999. Regulation of nuclear localization: a key to a door. *Annu. Rev. Cell Dev. Biol.* 15:291–339.

Kiernan, A.E., A.L. Pelling, K.K. Leung, A.S. Tang, D.M. Bell, C. Tease, R. Lovell-Badge, K.P. Steel, and K.S. Cheah. 2005. Sox2 is required for sensory organ development in the mammalian inner ear. *Nature.* 434:1031–1035.

Köhler, M., C. Speck, M. Christiansen, F.R. Bischoff, S. Prehn, H. Haller, D. Görlich, and E. Hartmann. 1999. Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. *Mol. Cell Biol.* 19:7782–7791.

Kurisaki, A., K. Kurisaki, M. Kowanzet, H. Sugino, Y. Yoneda, C.H. Heldin, and A. Moustakas. 2006. The mechanism of nuclear export of Smad3 involves exportin 4 and Ran. *Mol. Cell Biol.* 26:1318–1332.

Lefebvre, V., B. Dumitriu, A. Penzo-Mendez, Y. Han, and B. Pallavi. 2007. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. *Int. J. Biochem. Cell Biol.* 39:2195–2214.

Lipowsky, G., F.R. Bischoff, P. Schwarzmaier, R. Kraft, S. Kostka, E. Hartmann, U. Kutay, and D. Görlich. 2000. Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *EMBO J.* 19:4362–4371.

Masui, S., Y. Nakatake, Y. Toyooka, D. Shimosato, R. Yagi, K. Takahashi, H. Okochi, A. Okuda, R. Matoba, A.A. Sharov, et al. 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.* 9:625–635.

Mattaj, I.W., and L. Englmeier. 1998. Nucleocytoplasmic transport: the soluble phase. *Annu. Rev. Biochem.* 67:265–306.

Mingot, J.M., S. Kostka, R. Kraft, E. Hartmann, and D. Görlich. 2001. Importin 13: a novel mediator of nuclear import and export. *EMBO J.* 20:3685–3694.

Mühlhäusser, P., E.C. Muller, A. Otto, and U. Kutay. 2001. Multiple pathways contribute to nuclear import of core histones. *EMBO Rep.* 2:690–696.

Murphy, E.C., V.B. Zhurkin, J.M. Louis, G. Cornilescu, and G.M. Clore. 2001. Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. *J. Mol. Biol.* 312:481–499.

Pemberon, L.F., and B.M. Paschal. 2005. Mechanisms of receptor-mediated nuclear import and nuclear export. *Traffic.* 6:187–198.

Pollard, V.W., W.M. Michael, S. Nakielnny, M.C. Siomi, F. Wang, and G. Dreyfuss. 1996. A novel receptor-mediated nuclear protein import pathway. *Cell.* 86:985–994.

Poulat, F., F. Girard, M.P. Chevron, C. Goze, X. Rebillard, B. Calas, N. Lamb, and P. Berta. 1995. Nuclear localization of the testis determining gene product SRY. *J. Cell Biol.* 128:737–748.

Que, J., T. Okubo, J.R. Goldenring, K.T. Nam, R. Kurotani, E.E. Morrisey, O. Taranova, L.H. Pevny, and B.L. Hogan. 2007. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development.* 134:2521–2531.

Ribbeck, K., and D. Görlich. 2002. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* 21:2664–2671.

Ribbeck, K., U. Kutay, E. Paraskeva, and D. Görlich. 1999. The translocation of transportin-cargo complexes through nuclear pores is independent of both Ran and energy. *Curr. Biol.* 9:47–50.

Sanchez, C., I. Sanchez, J.A. Demmers, P. Rodriguez, J. Strouboulis, and M. Vidal. 2007. Proteomics analysis of Ring1B/Rnf2 interactors identifies a

Published April 6, 2009

- novel complex with the Fbx10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. *Mol. Cell. Proteomics*. 6:820–834.
- Shmidt, T., F. Hampich, M. Ridders, S. Schultrich, V.H. Hans, K. Tenner, L. Vilianovich, F. Qadri, N. Alenina, E. Hartmann, et al. 2007. Normal brain development in importin-alpha5 deficient-mice. *Nat. Cell Biol.* 9:1337–1338.
- Sinclair, A.H., P. Berta, M.S. Palmer, J.R. Hawkins, B.L. Griffiths, M.J. Smith, J.W. Foster, A.M. Frischauf, R. Lovell-Badge, and P.N. Goodfellow. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*. 346:240–244.
- Südbeck, P., and G. Scherer. 1997. Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9. *J. Biol. Chem.* 272:27848–27852.
- Takahashi, K., and S. Yamanaka. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 126:663–676.
- Taranova, O.V., S.T. Magness, B.M. Fagan, Y. Wu, N. Surzenko, S.R. Hutton, and L.H. Pevny. 2006. SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev.* 20:1187–1202.
- Werner, M.H., J.R. Huth, A.M. Gronenborn, and G.M. Clore. 1995. Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell*. 81:705–714.
- Williamson, K.A., A.M. Hever, J. Rainger, R.C. Rogers, A. Magee, Z. Fiedler, W.T. Keng, F.H. Sharkey, N. McGill, C.J. Hill, et al. 2006. Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. *Hum. Mol. Genet.* 15:1413–1422.
- Yasuhara, N., N. Shibazaki, S. Tanaka, M. Nagai, Y. Kamikawa, S. Oe, M. Asally, Y. Kamachi, H. Kondoh, and Y. Yoneda. 2007. Triggering neural differentiation of ES cells by subtype switching of importin-alpha. *Nat. Cell Biol.* 9:72–79.
- Ying, Q.L., M. Stavridis, D. Griffiths, M. Li, and A. Smith. 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.* 21:183–186.
- Yoshida, K., and G. Blobel. 2001. The karyopherin Kap142p/Msn5p mediates nuclear import and nuclear export of different cargo proteins. *J. Cell Biol.* 152:729–740.

Published April 6, 2009

## Supplemental Material

JCB

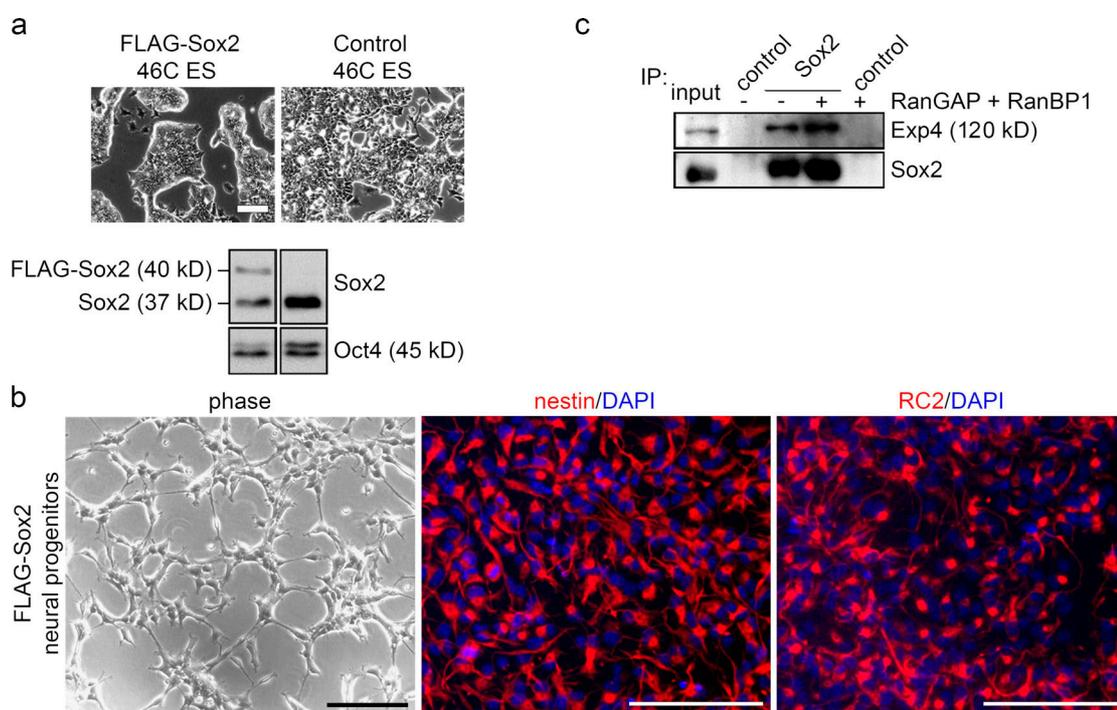
Gontan et al., <http://www.jcb.org/cgi/content/full/jcb.200810106/DC1>

Figure S1. **Characterization of Flag-Sox2 ES cells and neural progenitor cells.** (a) Flag-Sox2 cells display normal ES cell growth behavior and express the ES cell marker Oct4. (top) Phase-contrast images of control 46C ES cells and 46C ES cells stably transfected with Flag-Sox2. (bottom) Anti-Sox2 and Oct4 Western blots using extracts from Flag-Sox2 ES cells (left) and control 46C ES cells (right). (b) Characterization of ES cell-derived Flag-Sox2 neural progenitors. The left panel shows a phase-contrast image of Flag-Sox2 neural progenitors. Immunostaining shows that the cells expressed the neural progenitor markers nestin (middle) and RC2 (right). (c) The Sox2-Exp4 complex is insensitive to RanGTPase activation by RanGTPase-activating protein (RanGAP) and RanBP1. Coimmunoprecipitation of Exp4 with endogenous Sox2 as in Fig. 1 b but with the indicated additions. IP, immunoprecipitation. Bars: (a) 200  $\mu$ m; (b) 100  $\mu$ m.

Published April 6, 2009

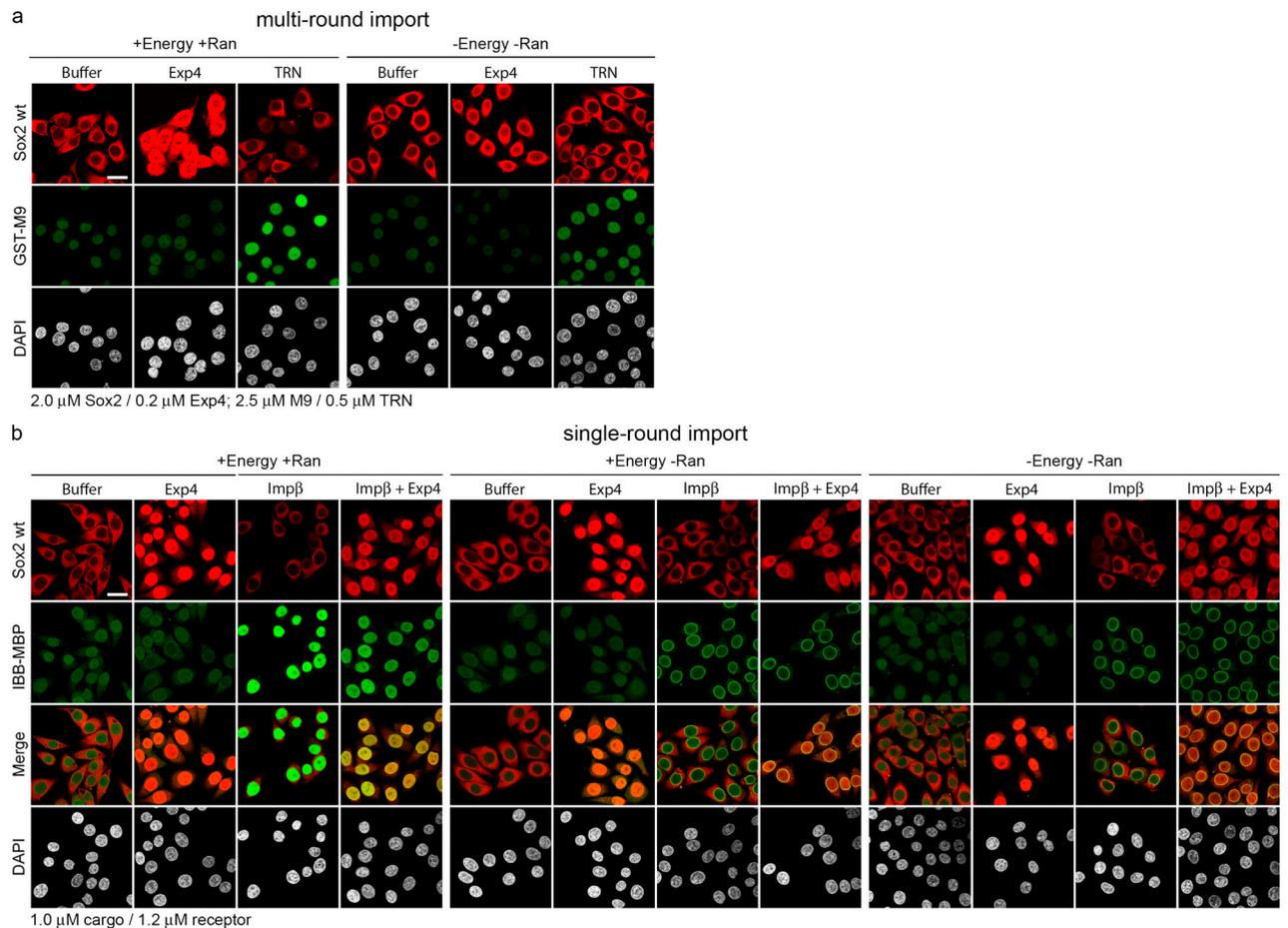


Figure S2. **Ran and energy dependence of Exp4-mediated nuclear import of Sox2.** (a) Efficient multiround nuclear import by Exp4 depends on Ran and metabolic energy. The import experiment was performed as in Fig. 2 a, but import substrates were in 10-fold (Sox2) or fivefold (M9) excess over the receptors, making efficient nuclear accumulation of cargoes dependent on receptor recycling. In the absence of Ran and energy, nuclear accumulation of both GST-M9 and GST-Sox2 was impaired. wt, wild type; TRN, transportin. (b) Single-round nuclear import of Exp4-Sox2 is independent of Ran and metabolic energy. The import experiment was performed as in Fig. 2 b, but transport receptors were in slight excess over the substrates to render nuclear accumulation of the cargoes independent of receptor recycling. Omission of Ran or of Ran and energy did not affect GST-Sox2 import by Exp4 but did prevent the nuclear accumulation of Imp- $\beta$ -IBB-MBP complexes (which instead arrested at nuclear pores). This was also evident when both Imp- $\beta$  and Exp4 had been added simultaneously (right). Bars, 25  $\mu$ m.

Published April 6, 2009

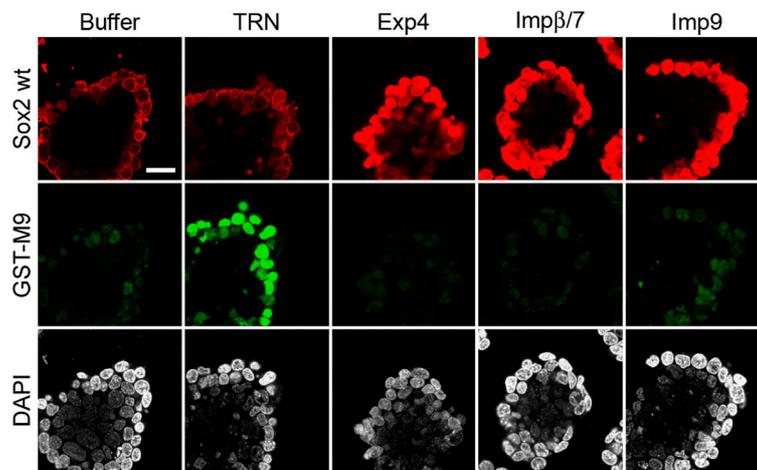


Figure S3. **Exp4, Imp-β/7, and Imp9 also mediate nuclear import of Sox2 in permeabilized ES cells.** NPC composition may vary among different cell types and change during development (Lupu, F., A. Alves, K. Anderson, V. Doye, and E. Lacy. 2008. *Dev. Cell.* 14:831–842). As it cannot be excluded that such changes regulate nuclear import, we tested the import activity of Exp4, Imp-β/7, and Imp9 with permeabilized 46C ES cells instead of HeLa cells. The import experiment was performed as in Fig. 2 a. The transportin (TRN) substrate GST-M9 was used as an internal specificity control. wt, wild type. Bar, 25 μm.

## CHAPTER 3

**Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP**

*This work has been originally published in the following article:*

**Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP**

Thomas Monecke<sup>\*</sup>, Thomas Güttler<sup>\*</sup>, Piotr Neumann, Achim Dickmanns, Dirk Görlich and Ralf Ficner

**Science**, 2009, Volume 324, Pages 1087-91.

\* These authors contributed equally to this work.

10. I. Ulbert, E. Halgren, G. Heit, G. Karmos, *J. Neurosci. Methods* **106**, 69 (2001).
11. Materials and methods are available as supporting material on Science Online.
12. M. Massimini, R. Huber, F. Ferrarelli, S. Hill, G. Tononi, *J. Neurosci.* **24**, 6862 (2004).
13. J. Numminen, J. P. Makela, R. Hari, *Electroencephalogr. Clin. Neurophysiol.* **99**, 544 (1996).
14. M. Velasco *et al.*, *Clin. Neurophysiol.* **113**, 25 (2002).
15. R. A. Wennberg, A. M. Lozano, *Clin. Neurophysiol.* **114**, 1403 (2003).
16. M. Volgushev, S. Chauvette, M. Mukovski, I. Timofeev, *J. Neurosci.* **26**, 5665 (2006).
17. M. Mukovski, S. Chauvette, I. Timofeev, M. Volgushev, *Cereb. Cortex* **17**, 400 (2007).
18. F. Amzica, M. Steriade, *Neurology* **49**, 952 (1997).
19. M. Bazhenov, I. Timofeev, M. Steriade, T. J. Sejnowski, *J. Neurosci.* **22**, 8691 (2002).
20. A. Compte, M. V. Sanchez-Vives, D. A. McCormick, X. J. Wang, *J. Neurophysiol.* **89**, 2707 (2003).
21. M. O. Cunningham *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5597 (2006).
22. D. Contreras, A. Destexhe, T. J. Sejnowski, M. Steriade, *Science* **274**, 771 (1996).
23. J. A. Hobson, E. F. Pace-Schott, *Nat. Rev. Neurosci.* **3**, 679 (2002).
24. R. Stickgold, *Nature* **437**, 1272 (2005).
25. G. Tononi, C. Cirelli, *Sleep Med. Rev.* **10**, 49 (2006).
26. S. Ribeiro *et al.*, *PLoS Biol.* **2**, E24 (2004).
27. A. Rechtschaffen, A. Kales, *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. (U.S. Government Printing Office, Washington, DC, 1968).
28. Supported by National Institute of Neurological Disorders and Stroke (NIH) grants NS18741 and NS44623 to E.H. and Neuroprobes IST-027017, OTKA49122, ETT135/2006, János Szentágothai Knowledge Center RET 05/2004 grants to I.U. We thank L. Papp for technical support; and L. Grand, A. Magony, B. Dombóvári, and D. Fabó for support in data analysis and recording. The authors dedicate this work to the memory of Edward Bromfield, who passed away on 10 May 2009.

## Supporting Online Material

www.sciencemag.org/cgi/content/full/324/5930/1084/DC1

Materials and Methods

Fig. S1

Table S1

References

11 December 2008; accepted 20 March 2009

10.1126/science.1169626

# Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP

Thomas Monecke,<sup>1\*</sup> Thomas Güttler,<sup>2\*</sup> Piotr Neumann,<sup>1</sup> Achim Dickmanns,<sup>1</sup> Dirk Görlich,<sup>2†</sup> Ralf Ficner<sup>1</sup>

CRM1 mediates nuclear export of numerous unrelated cargoes, which may carry a short leucine-rich nuclear export signal or export signatures that include folded domains. How CRM1 recognizes such a variety of cargoes has been unknown up to this point. Here we present the crystal structure of the SPN1-CRM1-RanGTP export complex at 2.5 angstrom resolution (where SPN1 is snurportin1 and RanGTP is guanosine 5' triphosphate-bound Ran). SPN1 is a nuclear import adapter for cytoplasmically assembled, m<sub>3</sub>G-capped spliceosomal U snRNPs (small nuclear ribonucleoproteins). The structure shows how CRM1 can specifically return the cargo-free form of SPN1 to the cytoplasm. The extensive contact area includes five hydrophobic residues at the SPN1 amino terminus that dock into a hydrophobic cleft of CRM1, as well as numerous hydrophilic contacts of CRM1 to m<sub>3</sub>G cap-binding domain and carboxyl-terminal residues of SPN1. The structure suggests that RanGTP promotes cargo-binding to CRM1 solely through long-range conformational changes in the exportin.

Nuclear transport proceeds through nuclear pore complexes (NPCs) and supplies cell nuclei with proteins and the cytoplasm with nuclear products such as ribosomes and tRNAs. Most nuclear transport pathways are mediated by importin  $\beta$ -type nuclear transport receptors, which include nuclear export receptors (exportins), as well as importins (1, 2). These receptors bind cargoes directly or through adapter molecules, shuttle constantly between the nucleus and cytoplasm, and use the chemical potential of the nucleocytoplasmic RanGTP-gradient to act as unidirectional cargo pumps (where GTP is guanosine 5' triphosphate and RanGTP is GTP-bound Ran) (3).

Exportins recruit cargo at high RanGTP levels in the nucleus, traverse NPCs as ternary cargo-exportin-RanGTP complexes, and release their cargo upon GTP hydrolysis into the cytoplasm. CRM1 (exportin1/Xpo1p) (4, 5) and CAS (Cse1p/exportin2) (6) are the prototypical exportins. Whereas CAS is specialized to retrieve the nuclear import adapter importin  $\alpha$  back to the cytoplasm (6), CRM1 exports a very broad range of substrates from nuclei (4, 5, 7–11), including ribosomes and many regulatory proteins. It also depletes translation factors from nuclei and is essential for the replication of viruses such as HIV.

CRM1 has a dual function during biogenesis of spliceosomal U small nuclear ribonucleoproteins (snRNPs). It exports m<sup>7</sup>G-capped U small nuclear RNAs to the cytoplasm (4, 12), where they recruit Sm-core proteins and receive a 2,2,7-trimethyl (m<sub>3</sub>G) cap structure. The import adapter snurportin 1 (SPN1) and importin  $\beta$  then transport the mature m<sub>3</sub>G-capped U snRNPs into nuclei (13). To mediate another import cycle, SPN1 is returned to the cytoplasm by CRM1 (14).

Many CRM1 cargoes harbor a leucine-rich nuclear export signal (NES) that typically includes four characteristically spaced hydrophobic residues (7). Examples are the HIV-Rev protein (15) or the protein kinase A inhibitor (PKI) (16). In other cases, however, CRM1 recognizes not just a short peptide, but instead a large portion of the export cargo; here, SPN1 is the prototypical example (14). CRM1 binds SPN1 tighter than other export substrates, apparently because CRM1 must displace the imported U snRNP from SPN1 before export may occur.

The cytoplasmic dissociation of CRM1 from SPN1 is essential for multi-round import of U snRNPs. Hydrolysis of the Ran-bound GTP alone is insufficient to fully disrupt the interaction (Fig. 1, A to C) (14), but importin  $\beta$  can displace CRM1 from SPN1 (Fig. 1A). Thus, either the binding sites of SPN1 for CRM1 and importin  $\beta$  overlap, or importin  $\beta$  forces SPN1 into a conformation that is incompatible with CRM1 binding.

Two functional domains in SPN1 have been described: (i) the m<sub>3</sub>G cap-binding domain (SPN<sup>97–300</sup>) (17) and (ii) the N-terminal importin  $\beta$ -binding (IBB) domain (SPN<sup>40–65</sup>) (14, 18, 19), which confers binding to and import by importin  $\beta$  (20). A multiple alignment of SPN1 from various species revealed another conserved region that precedes the IBB domain and includes the hydrophobic residues Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup>, and Val<sup>14</sup>. Mutating any of those residues to serine or deleting Met<sup>1</sup> strongly impaired the interaction with CRM1, in particular at higher salt concentrations (Fig. 1B and fig. S1). Even though the SPN1 N terminus (with its conserved hydrophobic residues) resembles a classical NES, there are clear differences: foremost that CRM1 binds the isolated SPN1 N terminus (SPN1<sup>1–21</sup>) considerably weaker than, for instance, the PKI-NES (Fig. 1C). In the context of full-length SPN1, however, this difference is more than compensated by the contribution of the m<sub>3</sub>G cap-binding domain to the CRM1 interaction.

We then assembled, purified, and crystallized an export complex containing full-length human SPN1<sup>1–360</sup>, full-length mouse CRM1<sup>1–1071</sup>, and GTP-RanQ69L<sup>1–180</sup>, a C-terminally truncated and GTPase-deficient form of human Ran (21).

<sup>1</sup>Abteilung für Molekulare Strukturbiologie, Institut für Mikrobiologie und Genetik, GZMB, Georg-August-Universität Göttingen, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany. <sup>2</sup>Abteilung Zelluläre Logistik, Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg 11, 37077 Göttingen, Germany.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: goerlich@mpibpc.mpg.de

## REPORTS

The resulting crystals contained two complexes per asymmetric unit. The structure was solved by molecular replacement with the use of known structures of GTP-bound Ran<sup>7-176</sup> (22), SPN1<sup>97-300</sup> (17), and a short human CRM1<sup>707-1027</sup> fragment (23). The final model, refined at a resolution of 2.5 Å, includes residues 12 to 1055 of CRM1 and Ran<sup>9-179</sup>, as well as SPN1<sup>1-360</sup>. CRM1<sup>67-69</sup>

and four regions of SPN1 appear to be disordered (table S1) (21).

As expected from previous sequence analysis (23, 24), CRM1 is built from so-called HEAT repeats (Fig. 2, fig. S2, and table S2), which include two consecutive helices (A and B) that pack in antiparallel orientation against each other and against the adjacent repeat (25). However,

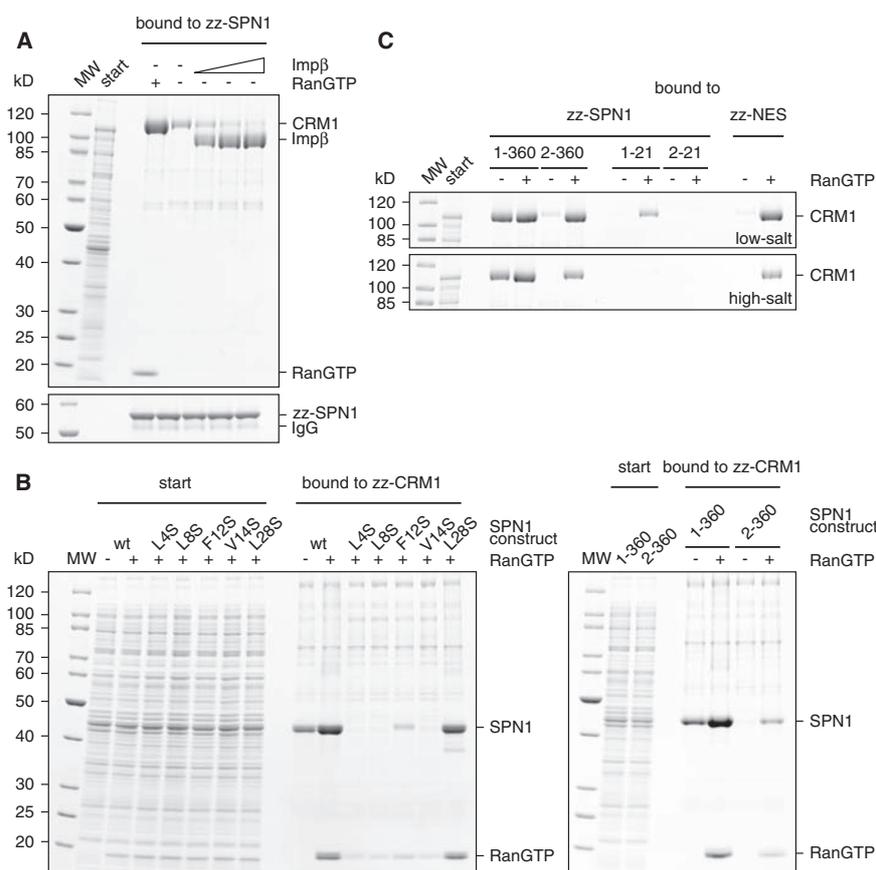
previous structure prediction (23) failed to predict the correct number and exact positions of the 21 repeats. This reflects the highly degenerate nature of some of the repeats, which even leads to an inverted topology of helices at the C terminus of CRM1 (fig. S2).

In contrast to importin β (26), transportin (27), and CAS/Cse1p (22, 28), the overall CRM1 structure shows remarkably little superhelical twist (Fig. 2). However, it is bent to a distorted toroid structure, with HEAT 21 touching helices 2B and 5A, as well as the loop between HEATs 4 and 5 (Fig. 2 and fig. S2). Ran is enclosed into this toroid and stabilizes the ring closure by extensive contacts. In contrast to the IBB–importin β interaction (18, 19, 26), the cargo SPN1 is not enveloped by CRM1 but instead rests on the outside of the CRM1 toroid (Fig. 2). This different binding topology might reflect the fact that CRM1 carries cargoes, such as ribosomal subunits, that are too large to be engulfed by an exportin. In addition, the outside of the torus provides a larger surface area and possibly also a greater variety of binding sites for cargo recognition than does the inner face that already accommodates the Ran molecule.

The structure of m<sub>3</sub>G cap-bound SPN1<sup>97-300</sup> was previously solved (17) and remained essentially unaltered in the SPN1-CRM1-RanGTP complex (root mean square deviation = 0.67 Å). However, several residues of SPN1 as well as of CRM1 HEATs 12 and 13 protrude into the m<sub>3</sub>G cap-binding pocket (fig. S3). With the physiological import cargo of SPN1 (fully assembled U snRNPs), the clashes would be even more severe, because the RNA would run into the CRM1 molecule. Thus, SPN1 cannot simultaneously bind its import cargo and its export receptor, which agrees with previous data (14). This ensures that only cargo-free SPN1 is returned to the cytoplasm and allows SPN1 to mediate unidirectional transport of m<sub>3</sub>G-capped U snRNPs into nuclei.

SPN1 binds CRM1 through an elaborate contact area (2330 Å<sup>2</sup>), which includes three parts: (i) the N terminus (SPN1<sup>1-35</sup>), (ii) the m<sub>3</sub>G cap-binding domain (SPN1<sup>97-300</sup>), and (iii) a C-terminal region, SPN1<sup>349-360</sup> (Fig. 3A). This is consistent with biochemical data that revealed strong contributions of SPN1<sup>1-21</sup> and the cap-binding domain to CRM1 binding (Fig. 1, B and C, and fig. S1) (14) and a weaker contribution of SPN1<sup>286-360</sup> (14).

All N-terminal residues that were found to be critical for CRM1-binding (SPN1<sup>Met1, Leu4, Leu8, Phe12, Val14</sup>) (Fig. 1B and fig. S1) dock into a hydrophobic cleft that is formed by helices 11A and 12A and the intervening helical linker between 11B and 12A of CRM1 (Fig. 3B and fig. S4). The side chain of CRM1<sup>Lys534</sup>, which is positioned by a salt bridge to CRM1<sup>Glu575</sup>, closes the cleft and introduces a sharp kink into the SPN1 chain between SPN1<sup>Val14</sup> and SPN1<sup>Ser15</sup>. There are several additional contacts in this area, such as electrostatic attraction between the negatively charged N-terminal helix of SPN1 and basic regions on the CRM1 surface, as well as hydrogen bonds



**Fig. 1.** (A) Effects of RanGTP and importin β on the SPN1-CRM1 interaction. 1 μM SPN1 was mixed with an *Escherichia coli* lysate containing 200 mM NaCl, 1 μM CRM1, and the indicated combinations of 3 μM RanGTP and 1, 2, or 3 μM importin β (Impβ) (21). Complexes were retrieved by immunoglobulin G (IgG)–Sephacrose via the zz-tag of SPN1. SPN1-ligands were eluted with 1.5 M MgCl<sub>2</sub> (upper panel); the remaining baits (zz-SPN1; lower panel) were eluted with SDS. Analysis was by SDS–polyacrylamide gel electrophoresis/Coomassie staining. Note that RanGTP enhanced CRM1-binding to SPN1; however, this interaction was also detectable in the absence of Ran. This residual CRM1-SPN1 interaction could be suppressed by importin β that binds the IBB domain of SPN1. MW, molecular weight standard. (B) Met<sup>1</sup>, Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup>, and Val<sup>14</sup> of SPN1 are all required for high-affinity binding to the CRM1-RanGTP complex. zz-tagged CRM1 immobilized on IgG-Sephacrose was incubated with an *E. coli* lysate containing 200 mM NaCl and indicated combinations of 3 μM RanGTP and 1 μM untagged wild-type SPN1 or the specified mutants. CRM1-ligands were eluted with MgCl<sub>2</sub> and analyzed as described in (A). Note that mutating Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup>, or Val<sup>14</sup> to Ser (left) or deleting Met<sup>1</sup> (right) abolished or substantially impaired SPN1 binding to CRM1-RanGTP, whereas mutating Leu<sup>28</sup> did not (left). (C) The N terminus of SPN1 contains export determinants that allow autonomous, RanGTP-stimulated binding to CRM1. Indicated zz-tagged SPN1 derivatives or the PKI-NES immobilized on IgG-Sephacrose were incubated with an *E. coli* lysate containing 1 μM CRM1 and 3 μM RanGTP as specified. Bound ligands were eluted with MgCl<sub>2</sub> and analyzed as described in (A). At low NaCl concentration (50 mM, upper panel), SPN1<sup>2-360</sup> bound CRM1-RanGTP nearly as efficiently as full-length SPN1<sup>1-360</sup>, however, a clear decrease in binding was observed without Ran. SPN1<sup>1-21</sup> recruited CRM1 in a strictly RanGTP-dependent manner. This CRM1 binding was lost when SPN1<sup>Met1</sup> was deleted. Even though SPN1<sup>1-21</sup> contains five hydrophobic residues, it bound CRM1 considerably weaker than the classical PKI-NES with only four hydrophobic residues. This difference was particularly apparent at 200 mM NaCl (lower panel).

between SPN1<sup>Ser15</sup> and CRM1<sup>Glu575</sup> and between SPN1<sup>Tyr35</sup> and CRM1<sup>Glu529</sup> (fig. S4). The CRM1 inhibitor leptomycin B (LMB) covalently modifies CRM1<sup>Cys528</sup> (29). Cys<sup>528</sup> is located within the hydrophobic cleft (Fig. 3B), which explains plausibly why LMB-modified CRM1 cannot bind export cargoes that rely on this cleft.

The N-terminal part of snurportin's export signature (with its five critical hydrophobic residues) resembles a classical NES and binds CRM1 in a conformation where residues Met<sup>1</sup> to Ser<sup>11</sup> form an  $\alpha$  helix (Fig. 3A and fig. S4). The classical NES from the HIV Rev protein (15) must be recognized differently for three reasons: (i) The spacing of the

hydrophobic residues is different, (ii) the intervening prolines would not allow such a helix to form, and (iii) this classical NES contains only four critical hydrophobic residues (15). Nevertheless, we cannot exclude the possibility that the same hydrophobic cleft also accommodates some or all of the key hydrophobic residues from classic NESs.

The interaction between the SPN1 m<sub>3</sub>G cap-binding domain and CRM1 is dominated by polar contacts. SPN1<sup>349–360</sup>, the third part of the export signature, binds to helices 14A, 15A, and 16A of CRM1 (Figs. 2 and 3A).

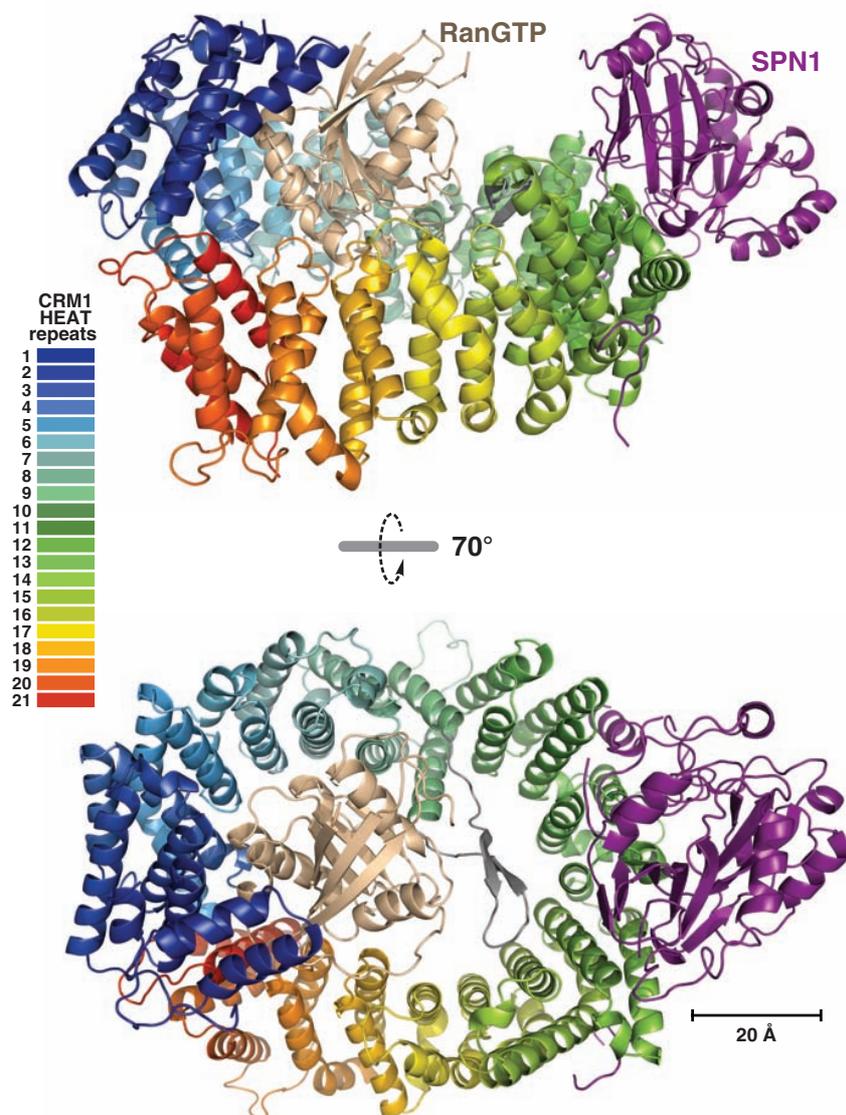
Importin  $\beta$  can displace CRM1 from the rather stable Ran-free SPN1·CRM1 complex

(Fig. 1A) and therefore restore m<sub>3</sub>G cap-binding of SPN1 in the cytoplasm (fig. S3) (14). This antagonism between CRM1 and importin  $\beta$  is not caused by an overlap of the respective binding sites; rather it appears to be caused by a combination of conformational changes in SPN1 and volume extrusion. The IBB domain binds importin  $\beta$  as a straight helix (18, 19). However, within the CRM1 complex, the central part of this IBB helix is unwound, and the remaining helix-fragments are kinked by  $\approx 80^\circ$  (shown in green in Fig. 3A). This distortion of the IBB helix appears to be enforced by contacts of the 35 N-terminal residues of SPN1 with CRM1. Thus, straightening of the IBB helix by importin  $\beta$  is likely to break crucial contacts between CRM1 and SPN1.

The structure of Ran in the SPN1·CRM1·RanGTP complex is virtually identical to that in other transport receptor·RanGTP complexes (22, 30, 31). Ran is almost completely engulfed by the CRM1 toroid and contacts four distinct areas of CRM1 (Figs. 2 and 4, A to C, and movie S1). The first area is located within the region that is most conserved between nuclear transport receptors (24, 32). HEATs 1 to 3 bind switch II of Ran, whereas HEATs 4 and 5 pack against Ran helix 3 and the so-called “basic patch” (30), respectively (Fig. 4, B and C). The second Ran-binding region (HEATs 7 to 9) also contacts the basic patch and extends to  $\beta$  strand 6 of Ran. Analogous interactions occur in RanGTP complexes with CAS, transportin, and importin  $\beta$  (22, 30, 31). In contrast, the long “acidic loop” within HEAT 9 (region 3) binds Ran in an unprecedented manner. It forms a  $\beta$  hairpin, touches HEAT helices B12 to B15, and reaches through the entire central “hole” of the CRM1 toroid (Figs. 2 and 4, A and B, and figs. S2 and S5). The acidic loop locks RanGTP closely to the N- and C-terminal HEAT repeats and binds Ran<sup>37</sup> from switch I, as well as Ran<sup>127,129,155</sup> from the loops involved in guanine recognition. The fourth C-terminal Ran-binding region (HEATs 17 and 19) was not anticipated by sequence similarity or previous structures. It contacts both switch regions of Ran.

To function as an effective, unidirectional cargo-pump, CRM1 must strongly discriminate between GTP- and guanosine diphosphate (GDP)-bound Ran. CRM1 can sense the nucleotide state of Ran because it contacts switches I and II; i.e., the regions that differ most between GDP- and GTP-Ran (Fig. 4, B and C, and fig. S5). The structure of RanGDP (33, 34) is incompatible with CRM1 binding, because switch I of RanGDP would clash with CRM1 HEAT 1 and switch II would collide with HEAT 19.

Ran switches the affinity of importin  $\beta$ -type transport receptors for their cargoes and thereby provides energy for the transport cycles. In the case of Cse1p, RanGTP increases the affinity of the exportin for its cargo importin  $\alpha$  by directly interacting with both Cse1p and importin  $\alpha$  (22). There are, however, no direct contacts between Ran and cargo in the SPN1·CRM1·RanGTP com-

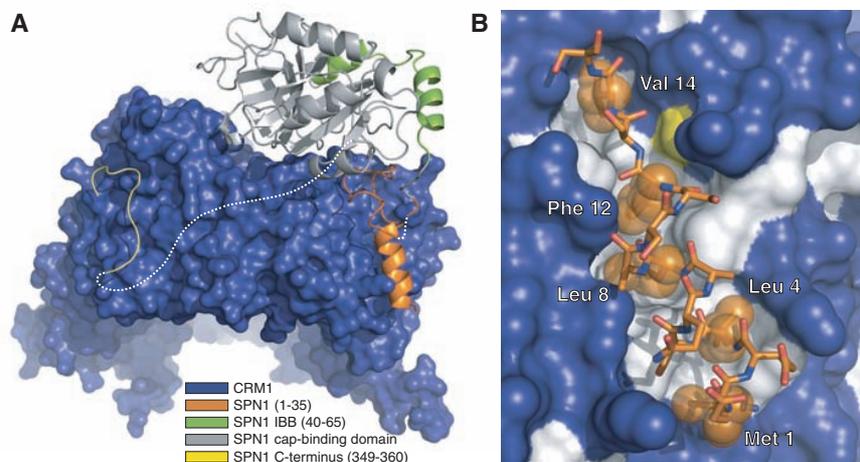


**Fig. 2.** Structure of the SPN1·CRM1·RanGTP nuclear export complex. Two views of the complex are depicted. Color-codes for Ran, SPN1, and the 21 consecutive HEAT repeats of CRM1 are indicated. Except for HEAT 21, A helices of the HEAT repeats are located at the outer surface of the CRM1 toroid and B helices at the inner surface (see also fig. S2). RanGTP is engulfed by the toroid-shaped structure of CRM1 and fixed by the so-called acidic loop (shown in the lower panel in gray), which is part of HEAT repeat 9. SPN1 is bound on the outer surface of CRM1, far away from the Ran molecule.

## REPORTS

plex (Fig. 2). The  $\approx 1000$ -fold increase in the affinity of CRM1 for RanGTP by SPN1 and the equally large strengthening of the CRM1·SPN1

interaction by RanGTP (14) must therefore be caused solely by conformational changes in the CRM1 molecule.

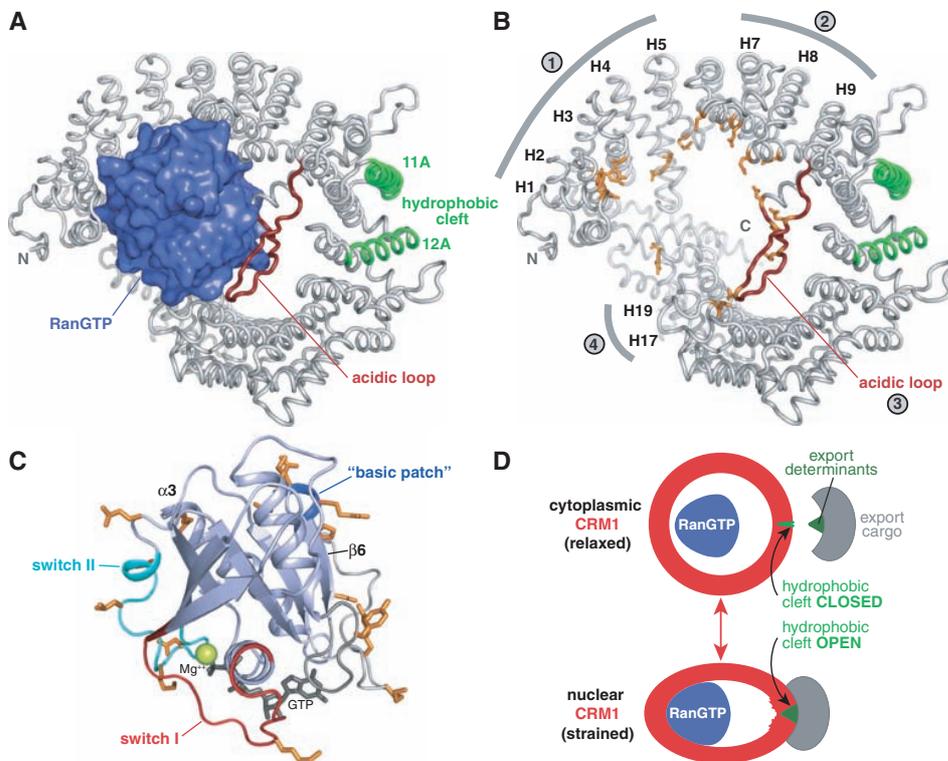


**Fig. 3.** The nuclear export signature of SPN1 involves a large interface formed by residues from all three domains of SPN1. **(A)** Three domains of SPN1 contact CRM1. These include N-terminal residues of SPN1 (orange), the cap-binding domain (gray), and the C-terminal residues (SPN1<sup>349–360</sup>, yellow). The IBB domain of SPN1, which forms a straight helix within the importin  $\beta$  complex (18, 19), is here partially unwound and bent (green). White dashed lines mark unresolved stretches. **(B)** The N-terminal hydrophobic residues of SPN1 (Met<sup>1</sup>, Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup>, and Val<sup>14</sup>) dock into a hydrophobic cleft of CRM1. Carbons of SPN1 are shown as orange, oxygens as red, and nitrogens as blue sticks. The side chains of the hydrophobic residues are depicted as spheres. CRM1 is shown as a surface representation (blue indicates hydrophilic areas, white denotes hydrophobic areas). The yellow patch marks the sulfur of Cys<sup>528</sup>, which is covalently modified by the CRM1-specific inhibitor LMB (29).

Both RanGTP and SPN1 obviously select the same conformation of CRM1 for high-affinity binding (here referred to as the nuclear conformation), whereas the free cytoplasmic conformation of CRM1 has only a low affinity for the two ligands (Fig. 1 and fig. S1). To explain this cooperativity, one must assume that the nuclear conformation is under considerable tension and that this tension is compensated for by the released binding energies of the Ran·CRM1 and SPN1·CRM1 interactions. In other words, Ran apparently promotes SPN1-binding by stabilizing the strained nuclear conformation of CRM1 and vice versa.

Ran and SPN1 are  $\approx 55$  Å apart in the export complex (Fig. 2). The conformational changes in CRM1 that coordinate their binding must therefore be transmitted over a considerable distance, probably through rigid body movements along the HEAT repeats. The splitting of the Ran-binding site on CRM1 into distinct regions is probably crucial for driving these movements, because even small changes in their distances will greatly affect the binding of Ran. Ran-binding regions 2 and 4, for example, are located on opposite sides of the CRM1 toroid (Fig. 4B), and it is quite possible that they are too far apart in the relaxed CRM1 conformation to contact Ran simultaneously (Fig. 4D). Hence, a low affinity for Ran would result. The transition to the nuclear conformation would bring these inter-

**Fig. 4.** Molecular details of the CRM1·RanGTP interaction (see also table S3). **(A)** CRM1 is shown as a gray backbone tube, and Ran is depicted as a surface representation. HEAT helices 11A and 12A, forming the cargo-binding hydrophobic cleft, are shown in green, and the acidic loop is shown in red. **(B)** RanGTP contact areas on CRM1. Orientation of CRM1 is as in (A), but Ran has been removed, and Ran-binding residues of CRM1 (distance  $< 3.6$  Å) are shown as orange sticks. Note that the Ran-binding site includes four distinct areas (labeled 1 to 4). See also movie S1. **(C)** Contacts of RanGTP to CRM1. Ran is depicted as a ribbon diagram. Orientation is as in (A). CRM1-binding residues are shown in orange, switch I (Ran<sup>30–47</sup>) is shown in red, and switch II (Ran<sup>65–80</sup>) is depicted in blue. CRM1 contacts both switches. The basic patch (Ran<sup>139–142</sup>, dark blue) shows extensive contacts to CRM1 regions 1 and 2 [see panel (B)]. Secondary structure elements are numbered as in (33). GTP is depicted as gray sticks; the Mg<sup>2+</sup> ion is shown as a green sphere. **(D)** Model for conformational states of CRM1. CRM1 switches between a relaxed cytoplasmic (top) and a strained nuclear conformation (bottom). In the hypothetical cytoplasmic conformation, the contact sites for RanGTP inside the CRM1 toroid are too far apart to bind Ran with high affinity. Also, the hydrophobic cleft on the outer side of the toroid is closed. Rigid body movements allow transition to the nuclear conformation. Here, the Ran-binding sites are close enough to bind Ran



simultaneously and thus with high affinity. The conformational change also alters the curvature of the toroid near the cargo-binding site, opens the hydrophobic cleft, and allows the export cargo to dock. For details, see main text.

faces closer together and allow high-affinity binding of Ran.

Conformational changes in CRM1, which favor the CRM1·Ran interaction, must also activate the cargo binding site(s). Thus, we suggest that the hydrophobic cleft is also controlled by these transitions. The cleft might be closed in the cytoplasmic, relaxed conformation of CRM1 (Fig. 4D). Putting the CRM1 toroid under tension to bind Ran with all interfaces should also change the curvature of the CRM1 molecule around the cargo-binding site. This might stretch the contacts between the outer A helices of HEATs 11 and 12 and consequently open the hydrophobic cleft. The observation that CRM1 binding of the isolated SPN1 N terminus and, thus, docking into the hydrophobic cleft is efficient only in the presence of RanGTP (Fig. 1C) strongly supports this model.

#### References and Notes

- D. Görlich, U. Kutay, *Annu. Rev. Cell Dev. Biol.* **15**, 607 (1999).
- A. Cook, F. Bono, M. Jinek, E. Conti, *Annu. Rev. Biochem.* **76**, 647 (2007).
- D. Görlich, N. Pante, U. Kutay, U. Aebi, F. R. Bischoff, *EMBO J.* **15**, 5584 (1996).
- M. Fornerod, M. Ohno, M. Yoshida, I. W. Mattaj, *Cell* **90**, 1051 (1997).
- K. Stade, C. S. Ford, C. Guthrie, K. Weis, *Cell* **90**, 1041 (1997).
- U. Kutay, F. R. Bischoff, S. Kostka, R. Kraft, D. Görlich, *Cell* **90**, 1061 (1997).
- U. Kutay, S. Güttinger, *Trends Cell Biol.* **15**, 121 (2005).
- O. Gadal *et al.*, *Mol. Cell. Biol.* **21**, 3405 (2001).
- T. I. Moy, P. A. Silver, *J. Cell Sci.* **115**, 2985 (2002).
- A. W. Johnson, E. Lund, J. Dahlberg, *Trends Biochem. Sci.* **27**, 580 (2002).
- M. T. Bohnsack *et al.*, *EMBO J.* **21**, 6205 (2002).
- M. Ohno, A. Segref, A. Bachi, M. Wilm, I. W. Mattaj, *Cell* **101**, 187 (2000).
- J. Huber *et al.*, *EMBO J.* **17**, 4114 (1998).
- E. Paraskeva *et al.*, *J. Cell Biol.* **145**, 255 (1999).
- U. Fischer, J. Huber, W. C. Boelens, I. W. Mattaj, R. Lührmann, *Cell* **82**, 475 (1995).
- W. Wen, J. Meinkoth, R. Tsien, S. Taylor, *Cell* **82**, 463 (1995).
- A. Strasser, A. Dickmanns, R. Lührmann, R. Ficner, *EMBO J.* **24**, 2235 (2005).
- D. Wohlwend, A. Strasser, A. Dickmanns, R. Ficner, *J. Mol. Biol.* **374**, 1129 (2007).
- G. Mitrousis, A. S. Olia, N. Walker-Kopp, G. Cingolani, *J. Biol. Chem.* **283**, 7877 (2008).
- D. Görlich, P. Henklein, R. Laskey, E. Hartmann, *EMBO J.* **15**, 1810 (1996).
- Materials and methods are available as supporting material on Science Online.
- Y. Matsuura, M. Stewart, *Nature* **432**, 872 (2004).
- C. Petosa *et al.*, *Mol. Cell* **16**, 761 (2004).
- D. Görlich *et al.*, *J. Cell Biol.* **138**, 65 (1997).
- M. A. Andrade, C. Petosa, S. I. O'Donoghue, C. W. Müller, P. Bork, *J. Mol. Biol.* **309**, 1 (2001).
- G. Cingolani, C. Petosa, K. Weis, C. W. Müller, *Nature* **399**, 221 (1999).
- B. J. Lee *et al.*, *Cell* **126**, 543 (2006).
- A. Cook *et al.*, *Mol. Cell* **18**, 355 (2005).
- N. Kudo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9112 (1999).
- I. R. Vetter, A. Arndt, U. Kutay, D. Görlich, A. Wittinghofer, *Cell* **97**, 635 (1999).
- Y. M. Chook, G. Blobel, *Nature* **399**, 230 (1999).
- M. Fornerod *et al.*, *EMBO J.* **16**, 807 (1997).
- K. Scheffzek, C. Klebe, K. Fritz-Wolf, W. Kabsch, A. Wittinghofer, *Nature* **374**, 378 (1995).
- M. Stewart, H. M. Kent, A. J. McCoy, *J. Mol. Biol.* **284**, 1517 (1998).
- We thank the staff of synchrotron beamlines at BESSY (Berlin), the European Molecular Biology Laboratory/Deutsches Elektronen-Synchrotron (Hamburg), and the European Synchrotron Radiation Facility (Grenoble) for assistance during data collection; D. Deichsel for excellent technical help; and S. Frey and S. Güttler for comments on the manuscript. We received funding from the Deutsche Forschungsgemeinschaft (grant SFB523 to R.F. and T.M.), the Max-Planck-Gesellschaft (to D.G. and T.G.), the Alfred Krupp-Foundation (to D.G. and T.G.), and the Boehringer Ingelheim Fonds (to T.G.). Coordinates and structure factors have been deposited in the Protein Data Bank with accession code 3GJX.

#### Supporting Online Material

[www.sciencemag.org/cgi/content/full/1173388/DC1](http://www.sciencemag.org/cgi/content/full/1173388/DC1)

Materials and Methods

Figs. S1 to S5

Tables S1 to S3

References

Movie S1

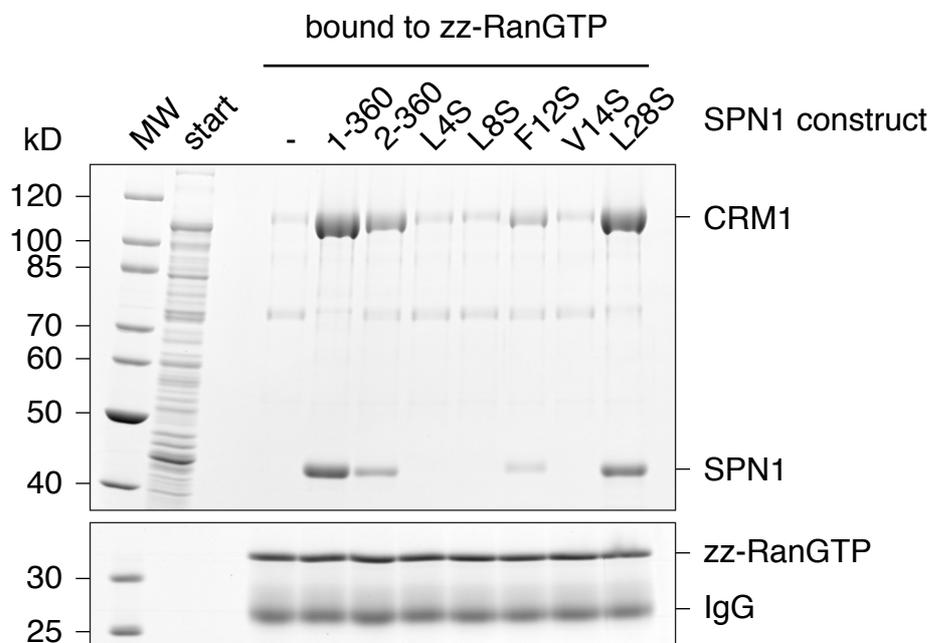
11 March 2009; accepted 9 April 2009

Published online 23 April 2009;

10.1126/science.1173388

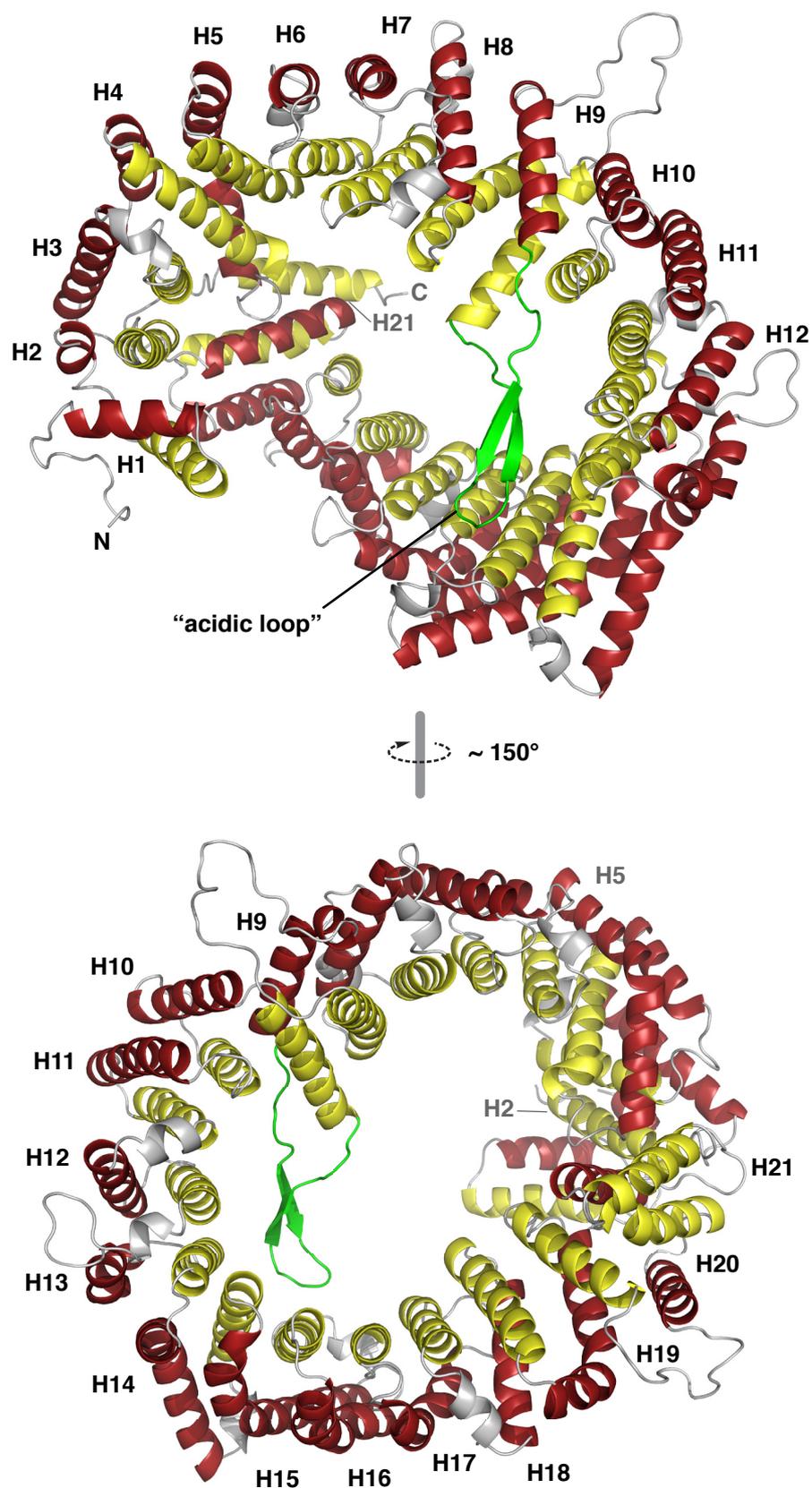
Include this information when citing this paper.

## 3.1 Supplementary Figures and Tables



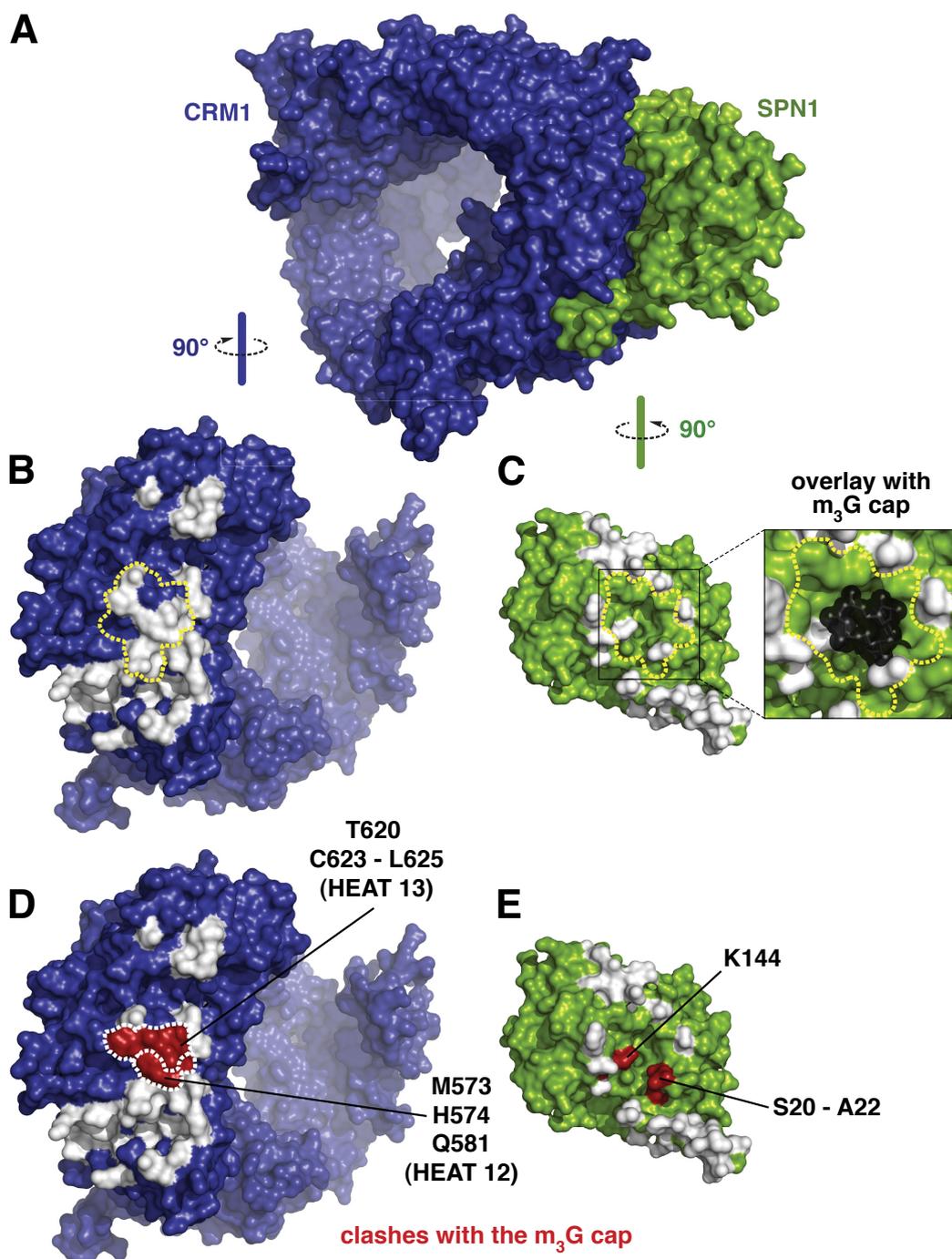
**Fig. S1. Met<sup>1</sup>, Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup> and Val<sup>14</sup> of SPN1 are critical for cooperative export complex formation.**

The experimental setup was essentially identical to that described in Fig. 1B with the difference that here, RanGTP instead of CRM1 was zz-tagged and immobilized on IgG-Sepharose. The beads were incubated at 200 mM NaCl with an *E. coli* extract that contained 1  $\mu$ M CRM1 and 1  $\mu$ M of wild-type or mutant SPN1 as indicated. RanGTP-bound ligands were eluted with MgCl<sub>2</sub> and analyzed by SDS-PAGE and Coomassie-staining. As the affinity of CRM1 for RanGTP is low in the absence of export cargo, RanGTP recruited CRM1 only very inefficiently without SPN1. However, addition of wild-type SPN1 (1-360) promoted CRM1-binding to RanGTP. Note that the deletion of Met<sup>1</sup> of SPN1 or the change of either Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup> or Val<sup>14</sup> to Ser abolished or strongly weakened the CRM1·RanGTP interaction. Mutating Leu<sup>28</sup> to Ser had no effect. See also Fig. 1B.



**Fig. S2. HEAT repeat organization of CRM1.**

The panels illustrate the organization of CRM1 from consecutive HEAT repeats, numbered H1-H21. Two views are depicted. A-helices (colored in red) are located on the outside of the torus, while B-helices (yellow) face the inside. However, this topology is inverted for HEAT 21. HEAT 21 also circularizes CRM1 by contacting HEATs 2 and 5. The long intra-repeat "acidic loop" within HEAT 9 is shown in green. See also Table S2.



**Fig. S3.  $m_3G$  cap binding is excluded in the export complex.**

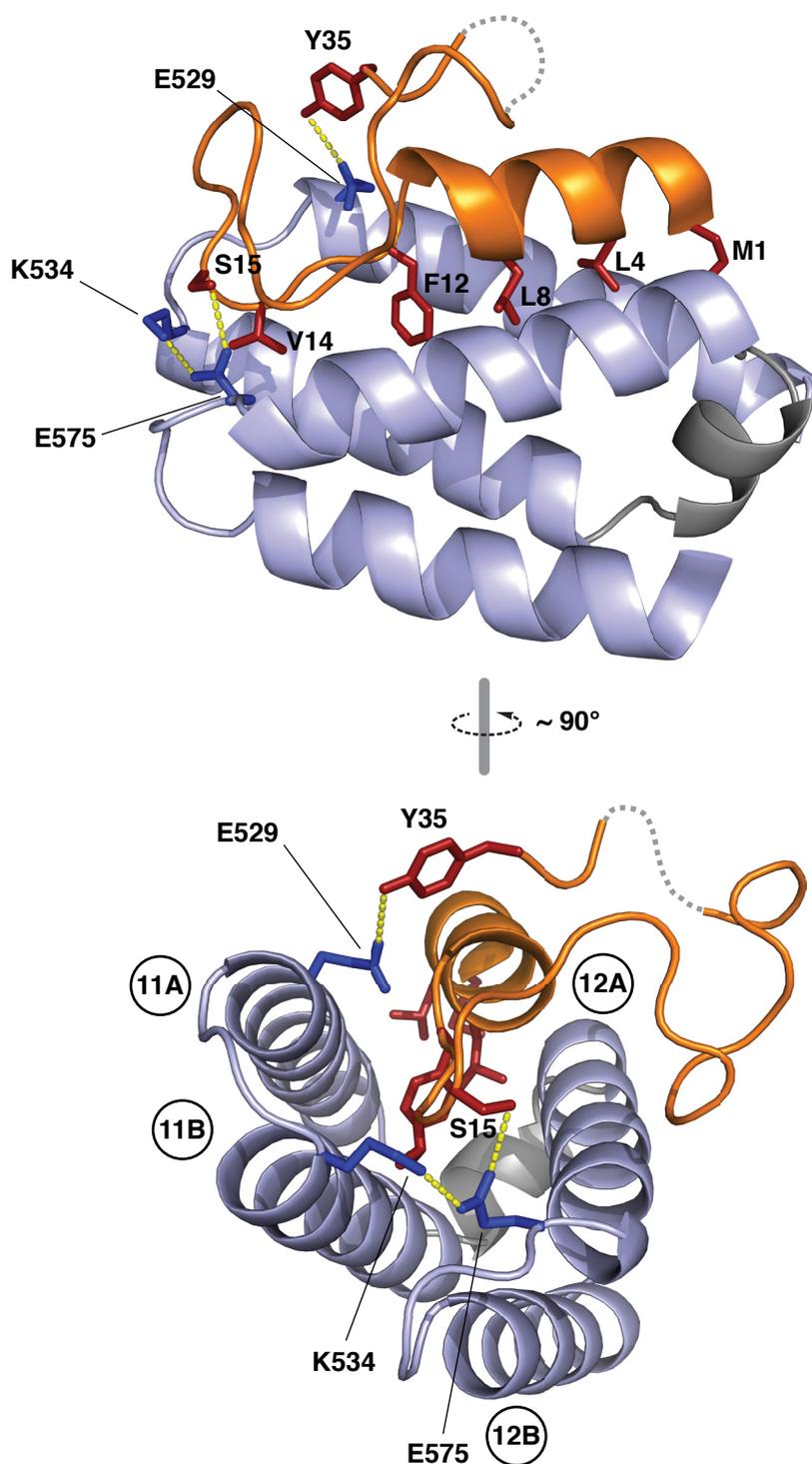
(A) The complex of CRM1 (blue) and SPN1 (green) is shown as a surface representation. In the following panels, it is flipped open to illustrate the contact sites (shown in white). Deviation from exact complementarity is due to coloring interacting residues and not interacting atoms.

(B) The CRM1 molecule from (A) is shown alone and rotated as indicated. Areas that contact SPN1 are colored in white. The region that covers the  $m_3G$  cap-binding site of SPN1 is demarcated by a yellow dashed line.

(C) The SPN1 molecule from (A) is shown alone and rotated as indicated. Residues contacting CRM1 are shown in white. The  $m_3G$  cap-binding site is bordered by a yellow dashed line. The insert shows superposition with an  $m_3G$  cap (in black) from the SPN1· $m_3GpppG$  complex (PDB accession code 1XK5).

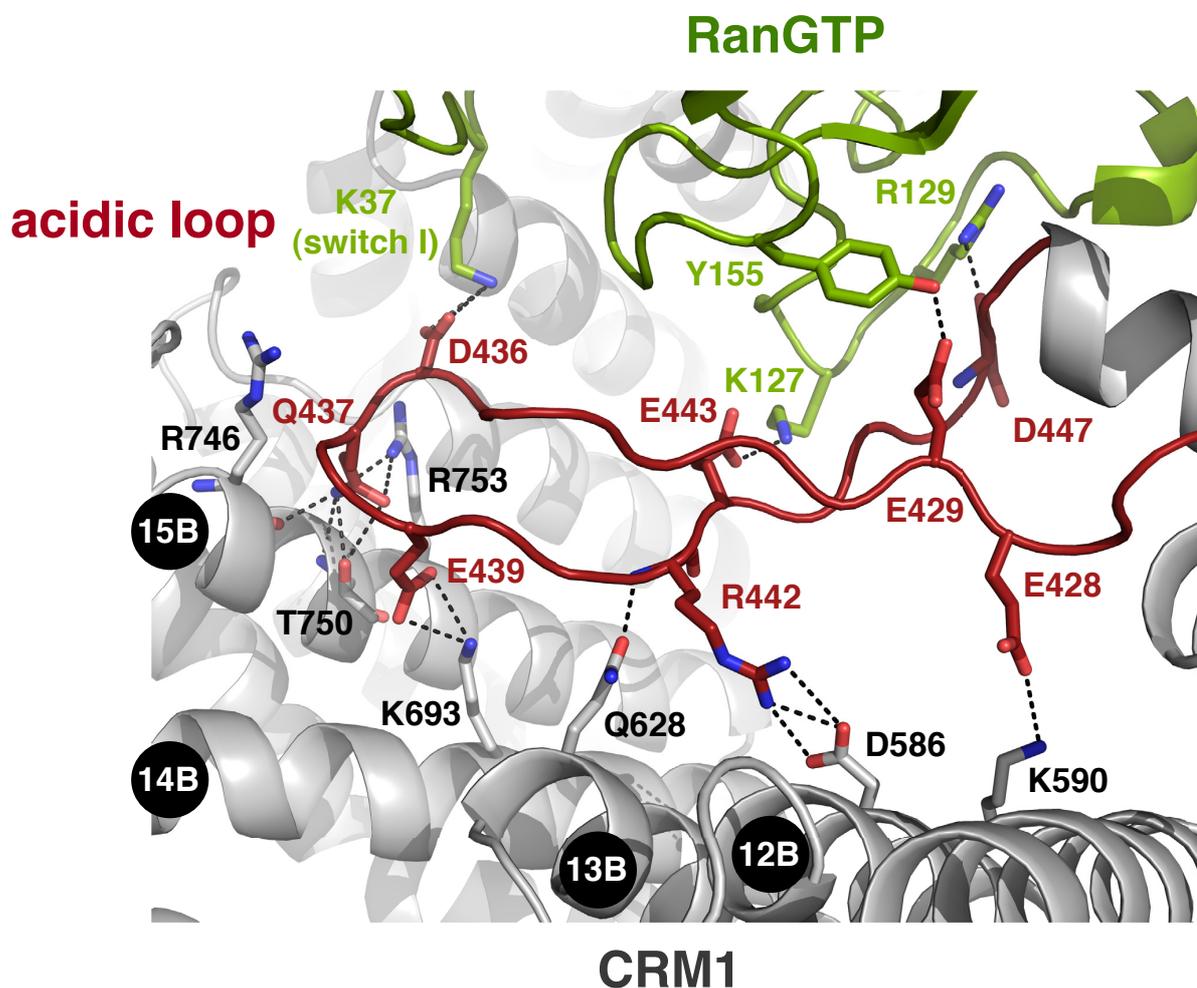
(D) The CRM1 molecule is shown as in (B), but those residues are shown in red that would clash with an  $m_3G$  cap bound to SPN1.

(E) The SPN1 molecule is shown as in (C), but residues that would clash with a bound  $m_3G$  cap are shown in red.



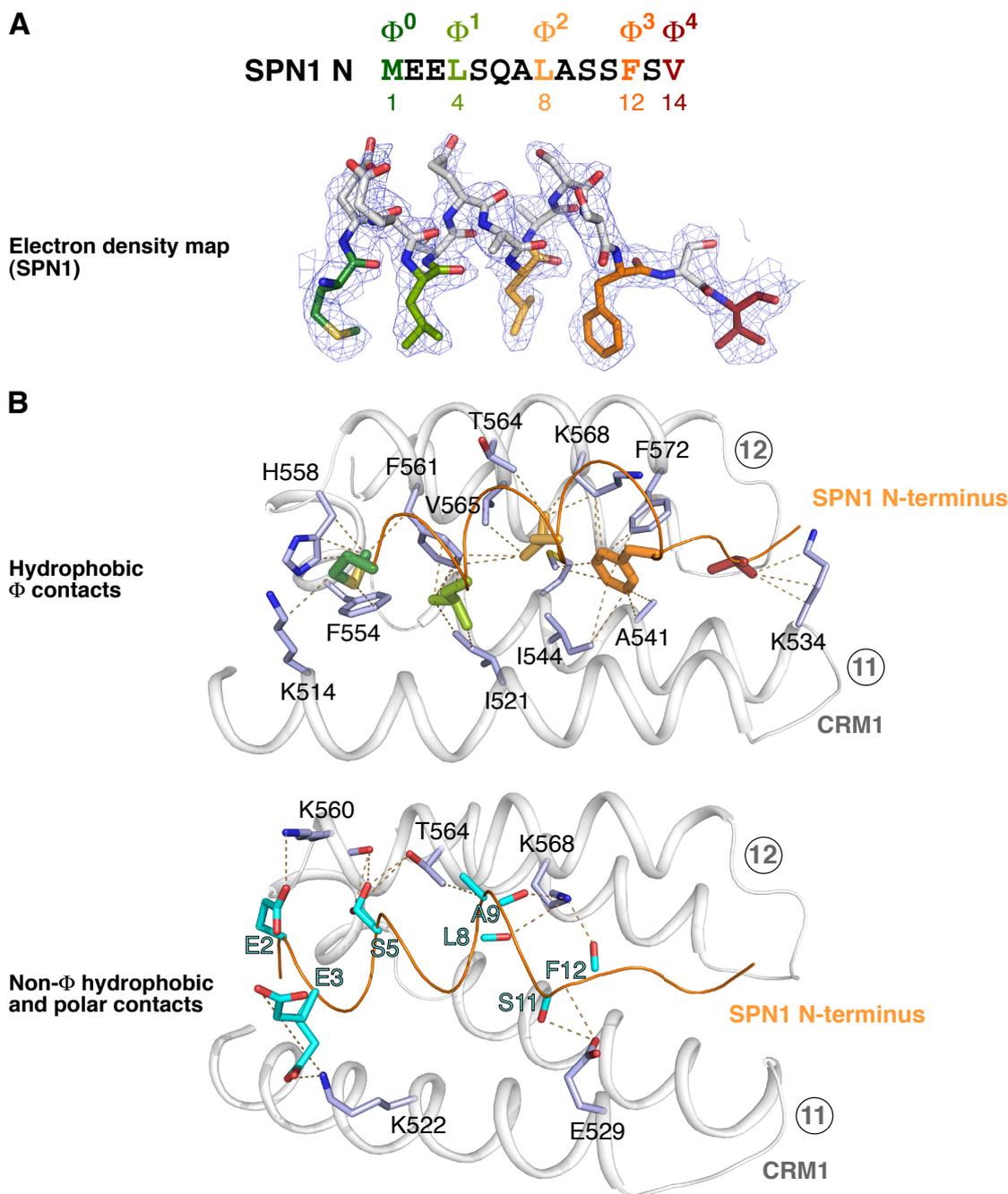
**Fig. S4. Molecular details of the CRM1·SPN1<sup>1-35</sup> interaction (overview).**

HEAT repeats 11-12 (CRM1<sup>510-595</sup>) are shown in blue, SPN1<sup>1-35</sup> in orange. SPN1 residues Met<sup>1</sup>, Leu<sup>4</sup>, Leu<sup>8</sup>, Phe<sup>12</sup> and Val<sup>14</sup> (shown as red sticks) wedge into the hydrophobic cleft that is formed by CRM1 helices 11A and 12A and the helical linker (gray) connecting helices 11B and 12A. Hydrogen bonds (CRM1<sup>E529</sup>-SPN1<sup>Y35</sup>; CRM1<sup>E575</sup>-SPN1<sup>S15</sup>) and the salt bridge (CRM1<sup>K534</sup>-CRM1<sup>E575</sup>) are illustrated as yellow dashed lines. CRM1 residues engaged in polar contacts are shown as blue sticks. SPN1<sup>30-32</sup> remained unresolved and are shown as a gray dashed line. See main text for further details. A complete analysis is given in Fig. S6.



**Fig. S5. Detailed view on the interactions of the acidic loop with Ran and HEAT repeats of CRM1.**

The acidic loop is colored in dark red, other parts of CRM1 in gray, and Ran in green. Atoms involved in polar interactions (black dashed lines) are colored in blue (nitrogen) or in light red (oxygen). Numbers in black circles denote those B-helices of CRM1 HEATs that contact the acidic loop with a distance of less than 3.6 Å. The extensive contacts immediately suggest that the acidic loop helps to couple Ran binding to those conformational changes that activate the cargo-binding site at HEATs 11 and 12. See also Fig. 4 and Table S3.



**Fig. S 6. Details of the interaction between the NES-like SPN1 N-terminus and CRM1.**

By sequence, the SPN1 N-terminus displays striking similarity to "Leu-rich" NES peptides. Residues 4; 8; 12 and 14 are equivalent to the four  $\Phi$  residues of the PKI NES ( $\Phi^1$ - $\Phi^4$ ). In addition, SPN1 uses its start-Methionine (termed " $\Phi^0$ ") to wedge into the hydrophobic cleft of CRM1. See also Fig. S4.

**(A) Upper:** Sequence of the SPN1 N-terminus.  $\Phi$  residues are color-coded. **Lower:** Panel displays the 2Fo-Fc electron density map (blue mesh, contoured at 1.0  $\sigma$ ) for the SPN1 N-terminus (shown as sticks) in the RanGTP·CRM1·SPN1 complex.  $\Phi$  residues are colored according to the shown sequence. Note that all  $\Phi$  residues are very well defined in the map. In all panels, dark blue marks nitrogen, oxygen is shown in light red and sulfur is colored in yellow.

**(B) CRM1 HEAT repeats 11-12** (gray cartoon) are shown with the SPN1 peptide bound (backbone traced in orange). SPN1-binding residues of CRM1 are depicted as blue sticks. Dashed lines link interacting atoms. Lines pointing onto backbones indicate contacts to carbonyl-carbons or amide groups. **Upper:** Panel shows the hydrophobic contacts of the  $\Phi$  residues (distance  $\leq 4.0$  Å). The respective  $\Phi$  residues are shown as sticks, the color code is explained in (A). **Lower:** Panel shows the non- $\Phi$  hydrophobic (distance  $\leq 4.0$  Å) as well as polar (distance  $\leq 3.8$  Å) contacts of SPN1 residues (cyan sticks).

**Table S1:** Crystallographic data statistics

Crystal	SPN1·CRM1·RanGTP
<b>Data collection</b>	
Space group	P 2 <sub>1</sub>
<i>Cell dimensions</i>	
a, b, c (Å)	72.17, 225.72, 163.41
α, β, γ (°)	90.00, 100.56, 90.00
Wavelength (Å)	0.9
X-ray source	BL14.1, BESSY (Berlin)
Resolution range (Å)	38.84-2.50 (2.60-2.50)
No. of reflections	749341 (64346)
Completeness (%)	97.3 (96.1)
R <sub>merge</sub> <sup>a</sup> (%)	10.8 (44.3)
Average I/σ	10.8 (2.9)
Redundancy	4.4 (3.4)
<b>Refinement</b>	
Resolution (Å)	38.8-2.5
Complexes per AU	2
<i>No. of atoms</i>	
Protein	24247
Ligand	66
Waters	959
R <sub>work</sub> <sup>b</sup> (%)	24.4 (28.1)
R <sub>free</sub> <sup>c</sup> (%)	28.1 (32.9)
Figure of merit	0.80
<i>Average B factors (Å<sup>2</sup>)</i>	
Protein	48.9
Ligand	29
Waters	40
<i>RMS deviations</i>	
Bond lengths (Å)	0.007
Bond angles (°)	1.079

Values in parentheses indicate the specific values in the particular highest resolution shell.

<sup>(a)</sup>  $R_{\text{merge}} = \frac{\sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I_i(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_i I_i(\text{hkl})}$ , where the sum  $i$  is over all separate measurements of the unique reflection hkl.

<sup>(b)</sup>  $R_{\text{work}} = \frac{\sum_{\text{hkl}} |F_{\text{obs}} - F_{\text{calc}}|}{\sum_{\text{hkl}} F_{\text{obs}}}$ .

<sup>(c)</sup> R<sub>free</sub> as R<sub>work</sub>, but summed over a 5 % test set of reflections.

**Table S2:** HEAT repeat helices in CRM1. See Fig. S2 for further details.

<b>HEAT repeat</b>	<b>Helix A residues</b>	<b>Helix B residues</b>
<b>1</b>	25 – 35	38 – 52
<b>2</b>	59 – 64	69 – 90
<b>3</b>	96 – 115	124 – 140
<b>4</b>	148 – 159	161 – 181
<b>5</b>	188 – 215	219– 234
<b>6</b>	246 – 254	261 – 273
<b>7</b>	280 – 297	314 – 339
<b>8</b>	344 – 359	363 – 383
<b>9</b>	405 – 423	
<b>Acidic loop</b>	424 – 448	
<b>9</b>		449 – 467
<b>10</b>	469 – 485	491 – 504
<b>11</b>	510 – 530	534 – 550
<b>12</b>	559 – 574	580 – 595
<b>13</b>	610 – 623	627 – 643
<b>14</b>	647 – 674	682 – 702
<b>15</b>	704 – 735	746 – 765
<b>16</b>	769 – 790	798 – 811
<b>17</b>	815 – 834	842 – 858
<b>18</b>	868 – 883	887 – 906
<b>19</b>	908 – 931	939 – 954
<b>20</b>	970 – 985	991 – 1005
<b>21</b>	1008 – 1023	1037 – 1052

**Table S3:** List of CRM1 residues that contact the RanGTP molecule with a distance of less than 3.6 Å. Corresponding HEAT repeats are numbered. “A”: A helix; “B”: B helix; “A-B”: Loop between A and B helix; “AL”: acidic loop within HEAT 9. Contact areas of CRM1 for Ran are numbered as in Fig. 4B.

Residue of CRM1	CRM1 HEAT	Contact area of CRM1	Distance [Å]	Contacted residue of Ran
<b>Tyr36</b>	1A-1B	1	3.28	<b>Gln82</b>
<b>Tyr77</b>	2B	1	3.54	<b>Asp77</b>
<b>Gln81</b>	2B	1	2.66	<b>Leu75</b>
<b>Gln81</b>	2B	1	2.56	<b>Asp77</b>
<b>Gln81</b>	2B	1	3.27	<b>Gly78</b>
<b>Lys129</b>	3B	1	3.51	<b>Asp77</b>
<b>Glu176</b>	4B	1	3.54	<b>Arg110</b>
<b>Glu176</b>	4B	1	2.58	<b>Arg110</b>
<b>Glu177</b>	4B	1	3.08	<b>Arg110</b>
<b>Glu177</b>	4B	1	2.80	<b>Arg110</b>
<b>Arg231</b>	5B	1	2.99	<b>Lys142</b>
<b>Asp313</b>	7A-7B	2	3.16	<b>Lys167</b>
<b>Asn317</b>	7B	2	3.20	<b>Asn143</b>
<b>Gln320</b>	7B	2	3.59	<b>Arg140</b>
<b>Asn321</b>	7B	2	3.54	<b>Asn143</b>
<b>Glu364</b>	8B	2	2.78	<b>His139</b>
<b>Glu364</b>	8B	2	3.41	<b>Arg140</b>
<b>Glu371</b>	8B	2	3.58	<b>Arg140</b>
<b>Glu429</b>	AL	3	2.53	<b>Tyr155</b>
<b>Asp436</b>	AL	3	3.35	<b>Lys37</b>
<b>Asp436</b>	AL	3	3.26	<b>Lys37</b>
<b>Glu443</b>	AL	3	3.19	<b>Lys127</b>
<b>Asp 447</b>	AL	3	2.47	<b>Arg129</b>
<b>Asp449</b>	9B	4	2.87	<b>Asp148</b>
<b>Glu843</b>	17B	4	3.38	<b>Lys37</b>
<b>Asp932</b>	19A-19B	4	2.88	<b>Lys71</b>
<b>Thr933</b>	19A-19B	4	2.62	<b>Glu70</b>
<b>Thr933</b>	19A-19B	4	2.78	<b>Lys71</b>
<b>Thr933</b>	19A-19B	4	3.07	<b>Lys71</b>

**Movie S1.** (available at <http://www.sciencemag.org/cgi/content/full/1173388/DC1>) The movie illustrates the multi-partite nature of the Ran-binding surface of CRM1. Coloring is according to Fig 4B: The Ran-binding residues are shown in orange, while the helices that form the hydrophobic cleft (11A and 12A) are shown in green. Orientation in the first frame is approximately identical to the orientation in Fig. 4B.

## 3.2 Materials and Methods

### 3.2.1 Protein expression and purification

Full-length mouse CRM1<sup>1-1071</sup> was expressed at 16 °C in *E. coli* BLR as an N-terminal His-zz-[TEV] fusion protein from a pQE80-derived plasmid (QIAGEN, Hilden, Germany). Cells were lysed in 50 mM Tris (pH 7.5), 500 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 2 mM imidazole, 5 mM DTT. The protein was purified via Ni<sup>2+</sup>-chelate affinity chromatography with elution in lysis buffer containing 200 mM imidazole. After cleavage of the His-zz-tag during dialysis to the lysis buffer, tag and the bulk of contaminants were removed via another Ni<sup>2+</sup>-chelate column. The flow-through was further purified by gel filtration on a Superdex 200 column (equilibrated in 50 mM Tris (pH 7.5), 50 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 5 mM DTT).

We truncated the C-terminus of Ran (residues 181-216), because it is disordered in other transport receptor complexes (Vetter *et al.*, 1999a; Cook *et al.*, 2007), it destabilizes the GTP-bound form of Ran and weakens the interactions with transport receptors (Richards *et al.*, 1995). The Q69L mutation blocks the GTPase-activity (Klebe *et al.*, 1995). Human RanQ69L<sup>1-180</sup> was expressed as an N-terminal His-zz-[TEV] fusion at 20 °C in *E. coli* BL21 (DE3). Cells were lysed in 50 mM K-Phosphate (pH 7.0), 500 mM NaCl, 5 mM Mg(OAc)<sub>2</sub>, 2 mM imidazole, 2 mM DTT. The purification of Ran was analogous to that of CRM1 with the difference that all steps performed after cell lysis were in the presence of 30 μM GTP. The nucleotide state of Ran was validated by extracting the nucleotide in de-ionized urea, followed by analysis on MonoQ (Amersham Biosciences).

Expression and purification of human SPN1 for crystallization was as described (Strasser *et al.*, 2004). Wild-type and mutant SPN1 constructs for binding assays were expressed as N-terminal His-[TEV] or His-zz-[TEV] fusions in *E. coli* BLR at 20-25 °C and purified via Ni<sup>2+</sup>-chelate affinity chromatography (lysis buffer: 50 mM Tris (pH 7.5), 200 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 2 mM imidazole, 2 mM DTT). Where indicated, the His-tag had been cleaved off by TEV-protease and removed as described above. Untagged and His-zz-tagged SPN1<sup>1-360/2-360</sup> were further purified on a Superdex 200 gel filtration column equilibrated to 50 mM Tris (pH 7.5), 200 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 2 mM DTT. His-zz-[TEV]-Hs SPN1<sup>1-21/2-21</sup> were expressed in *E. coli* BLR at 37 °C, purified under denaturing conditions via the His-tag (lysis in 50 mM Tris (pH 8.0), 6 M Guanidinium hydrochloride, 1 mM EDTA, 2 mM DTT; elution in 50 mM Tris (pH 7.5), 8 M urea, 50 mM NaCl, 1 mM EDTA, 200 mM imidazole, 2 mM DTT) and re-folded by dialysis to 50 mM Tris (pH 7.5), 50 mM NaCl, 1 mM EDTA, 2 mM

DTT. Expression and purification of human importin  $\beta$  was as described (Jäkel and Görlich, 1998).

### 3.2.2 *Binding assays*

Binding assays were performed by incubating the specified components (see Fig. 1 and fig. S1) for 3 h with 20  $\mu$ l IgG-Sepharose 6 FastFlow (Amersham Biosciences) at 50 mM Tris (pH 7.5), 200 mM NaCl (if not denoted differently), 2 mM Mg(OAc)<sub>2</sub>, 1 mM DTT, 0.005% (w/v) digitonin in the presence of an ATP/GTP-regenerating system (Jäkel and Görlich, 1998). Where indicated, 3  $\mu$ M RanQ69L<sup>1-180</sup> (GTP-form) was added. The binding volume was 500  $\mu$ l. Beads were washed 3x with 500  $\mu$ l of the respective binding buffer; bound material was eluted with 50 mM Tris (pH 7.5), 1.5 M MgCl<sub>2</sub> and precipitated with isopropanol (95% v/v final). Baits immobilized on IgG-Sepharose were then eluted with SDS.

### 3.2.3 *Reconstitution and structure determination of the SPN1·CRM1·RanGTP complex*

The complex was prepared by mixing purified CRM1<sup>1-1071</sup>, GTP-RanQ69L<sup>1-180</sup> and SPN1<sup>1-360</sup> and further purified by gel filtration on a Superdex 200 column (equilibrated in 50 mM Tris (pH 7.5), 50 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 5 mM DTT). For crystallization, the protein solution was concentrated to 4 mg·ml<sup>-1</sup> and stored on ice for a maximum of 2 weeks.

The complex was crystallized by the vapor diffusion method in sitting drop 24-well ChrysChem-plates (Hampton Research, CA, USA). 1  $\mu$ l of a reservoir solution containing 11% (w/v) PEG1000 and 100 mM Tris pH 8.05 was mixed with 1  $\mu$ l of the prepared protein complex solution. Single crystals with dimensions of 50  $\mu$ m x 50  $\mu$ m x 300  $\mu$ m grew within 4 days at 293 K and belonged to the space group P2<sub>1</sub> with cell dimensions of a=72.17 Å, b=225.74 Å and c=163.45 Å and angles of  $\alpha$ =90.0°,  $\beta$ =100.6° and  $\gamma$ =90.0°. The crystals were soaked in 15.7 % (v/v) propanediol for 10 seconds and flash-frozen.

For structure determination more than 500 crystals had to be tested. The X-ray diffraction data of two crystals, diffracting to a maximum resolution of 2.5 Å, were integrated, scaled, reduced and merged using XDS (Kabsch, 1993). The structure was solved by means of molecular replacement using PHASER (McCoy, 2007) with the crystal structures of GTP-Ran<sup>7-176</sup> (PDB-ID 1WA5; Matsuura and Stewart, 2004), the m<sub>3</sub>G-cap-binding domain of SPN1

(residues 97-300; PDB-ID 1XK5; Strasser *et al.*, 2005) and CRM1<sup>707-1027</sup> (PDB ID 1W9C, Petosa *et al.*, 2004) as search models. The resulting electron density map was used to complete the initial search model. The structure was improved by iterative cycles of refinement using CNS (Brunger, 2007) and PHENIX (Adams *et al.*, 2002), as well as model building in COOT (Emsley and Cowtan, 2004). Waters were built manually in COOT.

In the final model of SPN1, residues 31-33, 72-91, 162-165 and 288-348 of one complex in the asymmetric unit, and residues 30-32, 74-92, 162-165 and 291-348 of the other one were not defined in the electron density map and thus they were not built. The electron density corresponding to the C-terminal residues of SPN1 could not be interpreted unambiguously. The model containing residues 349-360 yielded the best R-factor and  $R_{\text{free}}$ -value, and strong electron density peaks correlated with the positions of the sulfur atoms of Cys<sup>356</sup> and Met<sup>358</sup>. Ran comprises the residues 9-179, the bound GTP molecule as well as a coordinated magnesium ion. The polypeptide chain of CRM1 could be traced for residues 12-1055 with exception of the flexible region encompassing residues 67-69.

The structure was refined at a resolution of 2.5 Å to an  $R_{\text{work}}$  of 24.4 % and an  $R_{\text{free}}$ -value of 28.1 % (see Table S1). In the final model, 88 % of the residues are located within the most favored regions of the Ramachandran plot, 11 % in the additionally allowed ones, 1% in the generously allowed regions and none in the disallowed regions. Contact surfaces were calculated with the program AREAIMOL as implemented in the CCP4 suite (version 6.0.2; Collaborative Computational Project, 1994). Figures were prepared with PyMOL (DeLano, W.L. The PyMOL Molecular Graphics System (2002), DeLanoScientific, USA).

## CHAPTER 4

**The NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1**

*This work has been accepted for publication in the following article:*

**NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1**

Thomas Güttler<sup>\*</sup>, Tobias Madl<sup>\*</sup>, Piotr Neumann<sup>\*</sup>, Danilo Deichsel, Lorenzo Corsini, Thomas Monecke, Ralf Ficner, Michael Sattler and Dirk Görlich

**Nature Structural & Molecular Biology**, 2010, accepted for publication (on condition of editorial changes in the text).

\* These authors contributed equally to this work.

Note added for publication of this thesis:

The citation of the article's print version is **Nature Structural & Molecular Biology**, 2010, Volume 17, Pages 1367-76.

## 4.1 Abstract

Classic nuclear export signals (NESs) confer CRM1-dependent nuclear export. Here we present crystal structures of the RanGTP·CRM1 complex alone and bound to the prototypic PKI or HIV-1 Rev NESs. These NESs differ drastically in the spacing of their key hydrophobic ( $\Phi$ ) residues. Yet, CRM1 recognizes them with the same rigid set of five  $\Phi$  pockets. The different  $\Phi$  spacings are compensated for by different conformations of the bound NESs: an  $\alpha$ -helical conformation in the case of PKI and, in the case of Rev, an extended conformation with a critical proline docking into a  $\Phi$  pocket. NMR analyses of CRM1-bound and free PKI NES suggest that CRM1 selects NES conformers that pre-exist in solution. Our data lead to a new structure-based NES consensus, explain why NESs differ in their affinities for CRM1, and why supraphysiological NESs bind the exportin so tightly.

## 4.2 Introduction

Nuclear export is essential for eukaryotic life. It proceeds through nuclear pore complexes (NPCs) and is typically accomplished by exportins (Görlich and Kutay, 1999; Cook and Conti, 2010). The most versatile exportin is CRM1, which is also called exportin 1 or Xpo1p (Adachi and Yanagida, 1989; Fornerod *et al.*, 1997a; Stade *et al.*, 1997; Nishi *et al.*, 1994; Wolff *et al.*, 1997). Exportins shuttle between nucleus and cytoplasm, bind cargo molecules at high RanGTP levels inside the nucleus, traverse NPCs as ternary cargo·exportin·RanGTP complexes, and release their cargo upon hydrolysis of the Ran-bound GTP into the cytoplasm (Kutay *et al.*, 1997; Fornerod *et al.*, 1997a).

RanGTP greatly increases the affinity of exportins for their cargoes. For the exportins CAS, exportin-t, and exportin 5, it is striking that RanGTP contacts not only the exportin but also the cargo (Matsuura and Stewart, 2004; Cook *et al.*, 2009; Okada *et al.*, 2009; see also **Chapter 5**), i.e., here, the binding energy released at the cargo·RanGTP interface is likely to make a substantial contribution to the observed cooperative effect. This mechanism is unavailable for CRM1, where RanGTP is enclosed within the toroid-like exportin molecule (Monecke *et al.*, 2009), far away from the cargo-binding site on the outside of the toroid (Dong *et al.*, 2009b; Monecke *et al.*, 2009). It was therefore suggested that CRM1 switches between a (nuclear) high-affinity state for RanGTP and cargo, and a (cytoplasmic) low-affinity conformation (Monecke *et al.*, 2009). Cooperativity should arise if the nuclear conformation is strained and if this strain is counterbalanced by the released binding energies

of the RanGTP·CRM1 and cargo·CRM1 interactions. In this scenario, RanGTP promotes cargo binding solely by stabilizing the strained nuclear conformation of CRM1. Direct evidence for such an allosteric mechanism comes from the very recently solved crystal structure of a RanBP1-stabilized disassembly intermediate of a RanGTP·CRM1 complex (Koyama and Matsuura, 2010; see also **Chapter 5**).

CRM1 serves a very broad range of export substrates and mediates, for instance, the nuclear export of small and large ribosomal subunits (Ho *et al.*, 2000; Gadal *et al.*, 2001; Moy and Silver, 2002; Thomas and Kutay, 2003). It is also required for numerous viral infection cycles. HIV-1, for example, uses CRM1 to export its genomic RNA from nuclei (Malim *et al.*, 1991; Fischer *et al.*, 1995; Fornerod *et al.*, 1997a). The HIV-1 Rev protein is an adaptor in this process. It binds the unspliced viral RNA, recruits CRM1 and thereby triggers export of the RNA into the cytoplasm, where the RNA is assembled into the next generation of viral particles. CRM1 is also a key element in many regulatory networks. For instance, it controls the nuclear activity of protein kinase A (PKA) by expelling the PKA·PKI (PKA inhibitor) complex from nuclei (Wen *et al.*, 1995). In this process, PKI acts as an adaptor to CRM1. Finally, CRM1 counteracts the leakage of cytoplasmic factors into nuclei. Examples are certain translation factors (Bohnsack *et al.*, 2002) as well as RanBP1 (Richards *et al.*, 1996), whose exclusive cytoplasmic localization is crucial for RanGTPase-driven nucleocytoplasmic transport (Izaurrealde *et al.*, 1997).

The simplest CRM1-dependent nuclear export determinants are the so-called classic nuclear export signals (NESs). These are short peptides reported to comprise four spaced hydrophobic residues (denoted  $\Phi^1$  to  $\Phi^4$ ) and to follow the consensus  $\Phi^1$ -(x)<sub>2,3</sub>- $\Phi^2$ -(x)<sub>2,3</sub>- $\Phi^3$ -x- $\Phi^4$ , with "x" preferentially being charged, polar or small amino acids (Kutay and Güttinger, 2005). The prototypical representatives are the PKI NES (**LALKLAGLDI**; critical hydrophobics shown in bold), which exemplifies the most common spacing of the hydrophobic positions ( $\Phi^1$ xxx $\Phi^2$ xx $\Phi^3$ x $\Phi^4$ ), as well as the HIV-1 Rev NES (**LPPLERLTL**) with a reported  $\Phi^1$ xx $\Phi^2$ xx $\Phi^3$ x $\Phi^4$  spacing (Wen *et al.*, 1995; Fischer *et al.*, 1995). Even though classic NESs are also referred to as leucine-rich NESs, analyses of other export cargoes as well as randomization-and-selection-screens revealed that Ile, Val, Met, or Phe are also permitted at the hydrophobic positions (Bogerd *et al.*, 1996; Zhang and Dayton, 1998; Kosugi *et al.*, 2008).

Snurportin 1 (SPN1), the nuclear import adaptor for m<sub>3</sub>G-capped spliceosomal U snRNPs (Huber *et al.*, 1998), exemplifies a CRM1 cargo with a complex export signature. Mapping

experiments (Paraskeva *et al.*, 1999) and the solved crystal structures of the CRM1·SPN1 (Dong *et al.*, 2009b) and the RanGTP·CRM1·SPN1 complexes (Monecke *et al.*, 2009) revealed that three regions of SPN1, namely the N-terminus, the m<sub>3</sub>G cap-binding domain and a C-terminal region contact the exportin. The SPN1 N-terminus (residues 1-14) docks into the so-called hydrophobic cleft of CRM1. In sequence, the SPN1 N-terminus resembles a leucine-rich NES, but differs from previously described classic NESs in that it contains not four but five critical hydrophobic positions ( $\Phi^0$ - $\Phi^4$ ) and binds CRM1 much more weakly than, for example, the PKI NES.

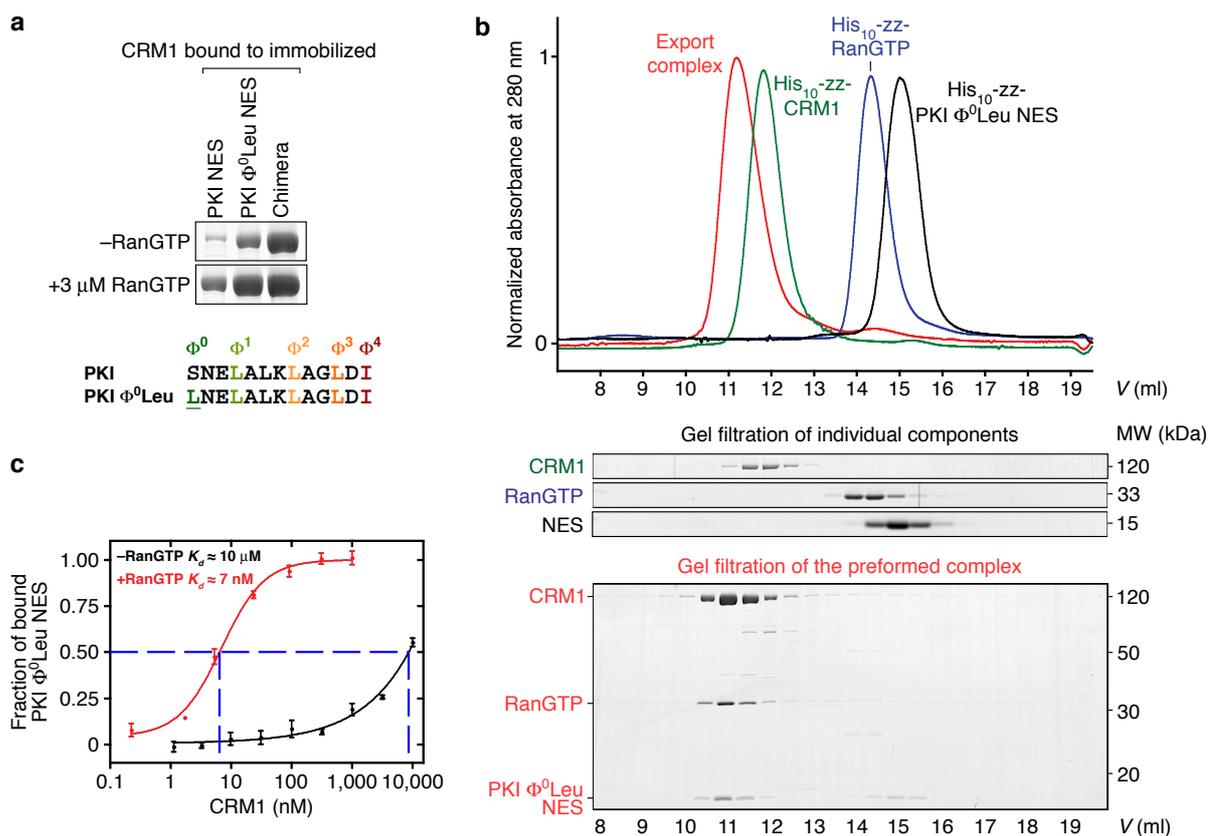
We now introduce a unifying structure-based NES consensus comprising five  $\Phi$  positions ( $\Phi^0$ ,  $\Phi^1$ - $\Phi^4$ ), which explains differences in binding strength that were previously not understood. The general importance of a hydrophobic  $\Phi^0$  residue so far escaped detection. However, we observed that it not only causes exceptionally strong CRM1 binding of the so-called supraphysiological S1 and NS2 NESs (Engelsma *et al.*, 2004; Engelsma *et al.*, 2008), but also that a  $\Phi^0$  residue is critically important for CRM1 binding of the classic Rev NES. In fact, our data revise the Rev NES from the previously suggested  $\Phi^1$ xx $\Phi^2$ xx $\Phi^3$ x $\Phi^4$  spacing to a fully unexpected  $\Phi^0$  $\Phi^{1Pro}$ x $\Phi^2$ xx $\Phi^3$ x $\Phi^4$  pattern. It has been an intriguing, unresolved question of how CRM1 can recognize diverse NESs that differ not only in their hydrophobic  $\Phi$  positions, but also in the length and sequence of their inter- $\Phi$ -spacers. One possibility was that the NES-binding site flexibly adapts its conformation to match individual NES sequences. Alternatively, different  $\Phi$  spacings might be compensated for by the use of additional  $\Phi$  pockets (Dong *et al.*, 2009b) or by allowing  $\Phi$  residues to "slide" within the hydrophobic cleft. We now solved the structures of RanGTP·CRM1 complexes without cargo, with a PKI NES or with an HIV-1 Rev NES bound; and we found that yet another scenario holds true: The NES-binding site of RanGTP-bound CRM1 is rigid, it does not adapt to individual NES sequences and uses the very same five binding pockets to capture the  $\Phi$  residues of either SPN1 N-terminus, PKI NES or Rev NES. These ligands bind to CRM1, however, in drastically different conformations. The CRM1-bound PKI  $\Phi^0$ Leu NES is  $\alpha$ -helical from  $\Phi^0$  to  $\Phi^2$ , whereas the backbone of the shorter-spaced Rev NES is extended. Thus, the NES ligands themselves compensate different  $\Phi$  spacings by a conformational adaptation to the rigid NES-binding site.

## 4.3 Results

### 4.3.1 Crystal structure of the classic PKI NES bound to CRM1

The PKI NES (LALKLAGLDI) contains four critical hydrophobic positions ( $\Phi^1$ - $\Phi^4$ ). By means of a linker, we had serendipitously introduced an additional upstream leucine into a fusion construct. We observed that this  $\Phi^0$ Leu improved CRM1 binding such that it became clearly detectable even in the absence of RanGTP, i.e. when CRM1 is in its low-affinity state for cargo binding (**Figure 4-1a**). A successful crystallization of classic NES·CRM1·RanGTP complexes was so far hampered by their apparent instability (Cook *et al.*, 2007). The PKI  $\Phi^0$ Leu NES was therefore a logical candidate for further crystallization trials. Indeed, we obtained a stable PKI  $\Phi^0$ Leu NES·CRM1·RanGTP complex (**Figure 4-1b, c**). However, even exhaustive screens failed to identify suitable crystallization conditions for the intact complex.

Because of this, we switched strategies and proceeded on the assumption that the PKI  $\Phi^0$ Leu NES docks into a binding site on CRM1 that is similar to that used by the SPN1 N terminus (Dong *et al.*, 2009b; Monecke *et al.*, 2009). We exploited the fact that the SPN1·CRM1·RanGTP complex readily crystallizes, facilitated by the m<sub>3</sub>G cap-binding domain of SPN1 that engages in critical crystal contacts (Monecke *et al.*, 2009). To implement this strategy, we exchanged the N-terminus of SPN1 for the PKI  $\Phi^0$ Leu NES, yielding a PKI  $\Phi^0$ Leu-SPN1 chimera. The addition of the SPN1<sup>15-360</sup> module to the NES not only stabilized the NES·CRM1 interaction (**Figure 4-1a**), but also allowed the NES·CRM1·RanGTP complex to form crystals that were isomorphous to the crystals of the previously solved SPN1·CRM1·RanGTP export complex (PDB-ID 3GJX, Monecke *et al.*, 2009, **Table 4-1**). As already experienced with this previous structure, we had to screen many crystals of the chimeric NES·CRM1·Ran complex to identify one that diffracted beyond 4 Å resolution. Also here, data collection and structure solving were complicated by high mosaicity and anisotropic diffraction, leading to low completeness and redundancy, especially for the highest resolution shell (**Table 4-1**). However, refinement strategies optimized for datasets with low completeness (see Methods and Rice and Brünger, 1994) yielded a reliable model at 3.42 Å resolution with good geometry for the PKI NES bound to CRM1 (**Figure 4-2; Supplementary Figure 4-1; Table 4-1**). As detailed below, the NES part of the model is supported by solution NMR, while the CRM1·RanGTP part is supported by a total of five independently refined crystal structures (PDB-ID 3GJX, Monecke *et al.*, 2009; and 3NBY, 3NBZ, 3NC0 as well as 3NC1 from this study).



**Figure 4-1: Enhancement of the PKI NES·CRM1 interaction by  $\Phi^0$ Leu and the SPN1<sup>15-360</sup> fusion module.**

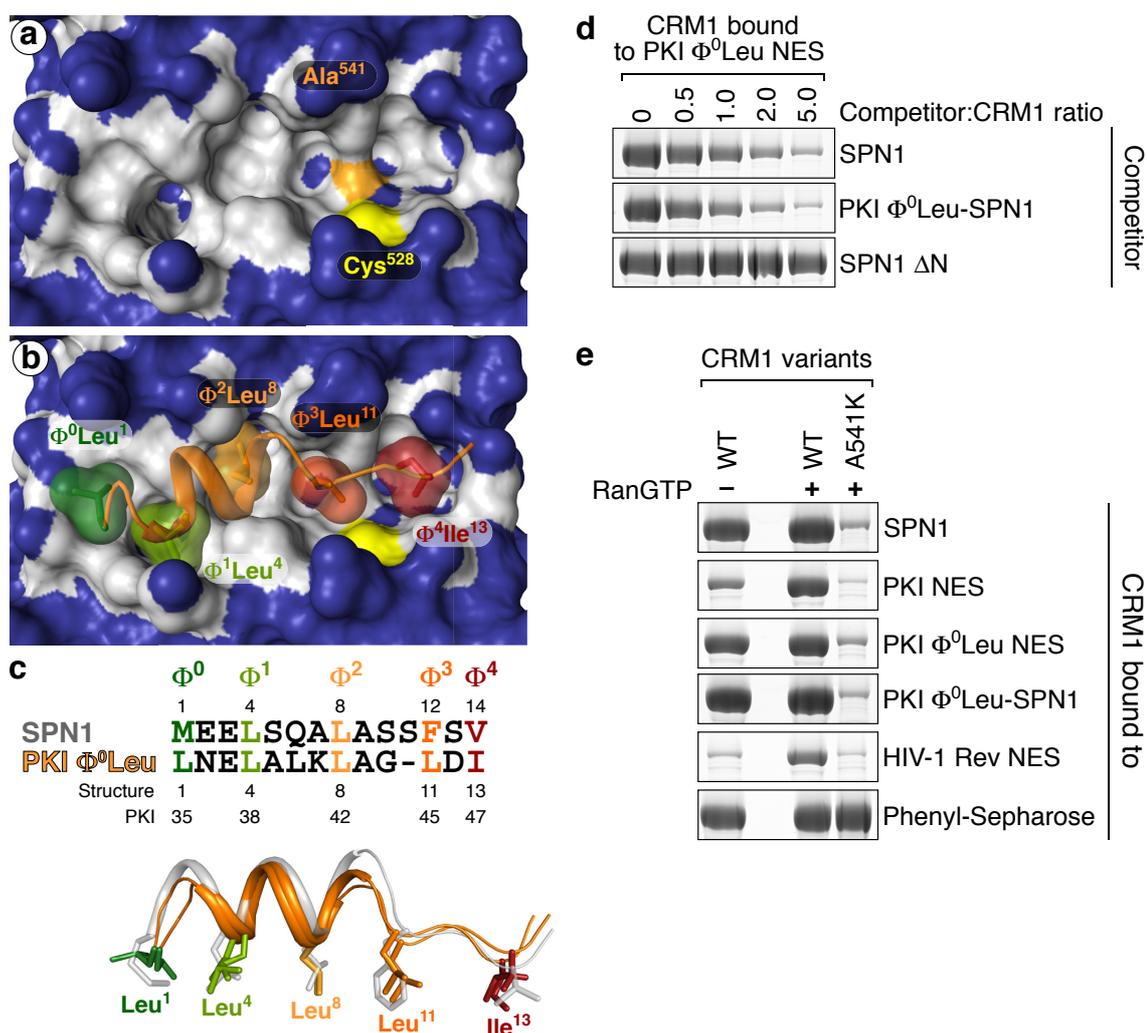
(a) The zz-tagged CRM1 cargoes, immobilized to IgG Sepharose, were incubated at 200 mM NaCl with an *E. coli* lysate containing 1  $\mu$ M CRM1. Beads were washed and bound fractions analyzed by SDS-PAGE/Coomassie staining. CRM1 binding to the PKI NES was clearly enhanced by the  $\Phi^0$ Ser→Leu mutation and stabilized further by fusion of the SPN1<sup>15-360</sup> module to the NES ("chimera"). This was most obvious when CRM1 was supplied in its Ran-free, low-affinity form for cargo binding. The NaCl concentration was chosen according to the strength of the CRM1-cargo interaction: strong interactions tolerate 200 mM NaCl whereas weak ones are evident only at 50 mM. This and subsequent figures also show relevant sequences with color-coded  $\Phi$  residues.

(b) Size exclusion chromatography of indicated export complex constituents on a Superdex 200 column. **Upper:** Elution profiles recorded at 280 nm **Lower:** Fractions analyzed by SDS-PAGE/Coomassie-staining. CRM1, Ran and NES co-eluted as a complex when mixed prior to gel filtration. (MW, molecular weight standard)

(c) CRM1 binding to the fluorescently labeled PKI  $\Phi^0$ Leu NES was quantitated at equilibrium by fluorescence polarization. The assay exploits the fact that the rotational diffusion of the 2.7 kDa NES peptide is slowed down when assembled into an > 120 kDa CRM1 complex. Apparent dissociation constants ( $K_d$ ) for the CRM1·NES interaction were determined by non-linear regression. Error bars illustrate standard deviations for three independent experiments. RanGTP (1  $\mu$ M) increased the affinity of CRM1 for the NES by a factor of  $\approx$  1400.

As expected, the structures of RanGTP, CRM1 and of the retained SPN1 parts showed only minor deviations from our reported structure with wild type SPN1. Crucially however, we could now see that all experimentally confirmed  $\Phi$  side chains of the PKI  $\Phi^0$ Leu NES (Wen *et al.*, 1995 and see below) docked into the exportin's hydrophobic cleft (**Figure 4-2a, b**), which represents also the docking site for the authentic SPN1 N-terminus (Monecke *et al.*, 2009; Dong *et al.*, 2009b; **Supplementary Figure 4-1; Figure S6 in Chapter 3**). The  $\Phi^2$ - $\Phi^3$

spacers of PKI NES and SPN1 N-terminus differ in length. Yet, equivalent  $\Phi$  side chains of the two NESs dock at identical positions of the CRM1 molecule (**Figure 4-2c**). The shorter  $\Phi^2$ - $\Phi^3$  spacer of the PKI  $\Phi^0$ Leu NES is compensated for by an earlier break in the  $\alpha$ -helix, which allows  $\Phi^3$ Leu<sup>11</sup> and  $\Phi^4$ Ile<sup>13</sup> to reach into their corresponding pockets of the hydrophobic cleft (**Figure 4-2b, c**). The docking of  $\Phi^0$ Leu<sup>1</sup> of the PKI NES into the  $\Phi^0$  pocket explains why this additional leucine confers a stronger CRM1 binding to the PKI NES (**Figure 4-1a**).



**Figure 4-2: Crystal structure of a classic PKI-type NES bound to the CRM1·RanGTP complex.**

(a) A complex of GTP-Ran<sup>5-180</sup>, mouse CRM1 and the PKI  $\Phi^0$ Leu NES-SPN<sup>115-360</sup> fusion was assembled, purified and crystallized, yielding crystals that diffracted to 3.42 Å. Picture shows the NES-binding site of CRM1 (hydrophobic cleft) as a surface representation. Blue denotes hydrophilic and white hydrophobic areas. Cys<sup>528</sup> is colored yellow, Ala<sup>541</sup> orange. Deep pockets are visible that accommodate the  $\Phi$  residues of the NES.

(b) (*previous page*) Same structure as in (a), but here the bound PKI NES (chain E) is also shown. The NES backbone is colored in orange. Helical regions were defined by PROMOTIF (Hutchinson and Thornton, 1996). Side chains of  $\Phi^0$ - $\Phi^4$  (color-coded as in **Figure 4-1a**) are presented as sticks and surfaces. Note that the  $\Phi$  residues perfectly fit into their pockets. NES·CRM1 interactions are further detailed in **Supplementary Figure 4-1**.

(c) **Upper:** sequence alignment of SPN1 N-terminus and PKI  $\Phi^0$ Leu NES. **Lower:** structural overlay of NES and SPN1 N-terminus in their CRM1-bound forms. The NES is colored as above, the SPN1 N-terminus (PDB-ID 3GJX, Monecke *et al.*, 2009) is colored gray.  $\Phi$  side chains of the two ligands are superimposable, despite their different  $\Phi^2$ - $\Phi^3$  spacings. Overlay derives from a C $\alpha$  alignment of the corresponding hydrophobic clefts. Both NES chains of the asymmetric unit (B & E) are shown to illustrate allowed flexibility.

(d) CRM1 was bound to immobilized PKI  $\Phi^0$ Leu NES (with 3  $\mu$ M RanGTP and 100 mM NaCl). This CRM1 binding was competitively inhibited by wild type SPN1 or the PKI NES-SPN<sup>115-360</sup> fusion. This is consistent with the assumption that PKI NES and SPN1 N-terminus are recognized by the same binding site. The SPN1 module lacking SPN1 residues 1-14 ( $\Delta$ N) alone did not compete.

(e) The A541K CRM1 mutation disables binding to the indicated export ligands, probably because the mutant residue clashes in all cases with the respective  $\Phi^3$  side chains (see panels a and b). Binding of CRM1 to Phenyl-Sepharose (a mimic of FG repeats, Ribbeck and Görlich, 2002) remained unaffected, indicating that the mutant's effects are restricted to the hydrophobic cleft. Binding to NES peptides was at 100 mM NaCl, binding to SPN1 chimeras and Phenyl-Sepharose at 200 mM. See also **Supplementary Figure 4-5**.

### 4.3.2 Point mutations in CRM1 that block NES binding to the hydrophobic cleft

One might argue that the addition of the SPN1<sup>15-360</sup> module forced the NES into the hydrophobic cleft and that an NES without the SPN1 fusion partner would bind to a different site and/or in a different mode. We therefore probed the atomic structure of a non-fused NES in its CRM1-bound state as well as the location of the NES on the exportin molecule by independent experimental strategies.

**Figure 4-2d** shows that the PKI NES-SPN1 chimera competitively blocks CRM1 binding of the non-fused PKI NES, supporting the assumption that fused and non-fused NES contact identical sites on CRM1.

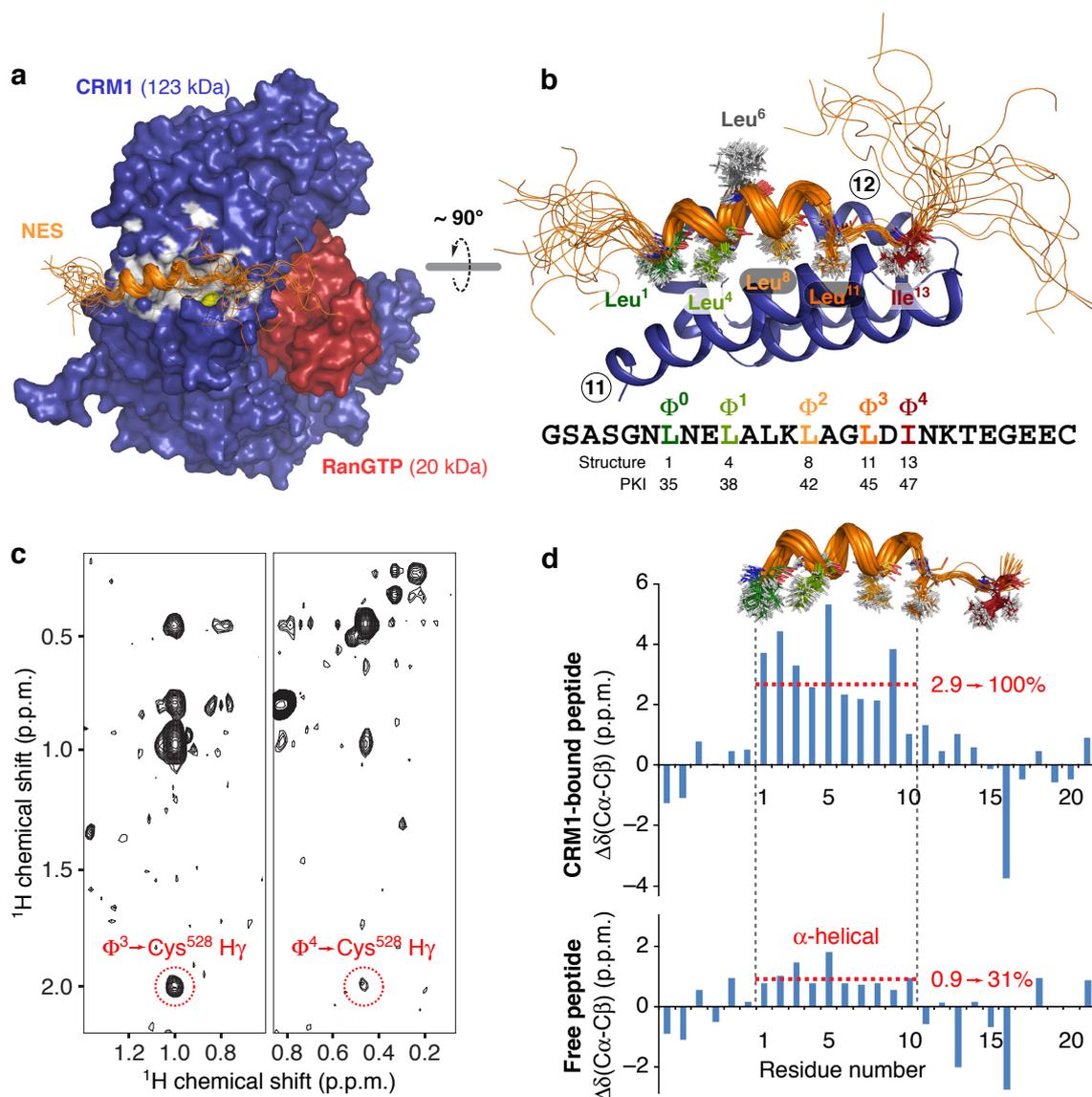
Furthermore, we mutated CRM1 Ala<sup>541</sup> (located at the base of the  $\Phi^3$ -binding pocket, Dong *et al.*, 2009b; Monecke *et al.*, 2009; **Figure 4-2a**) to a lysine that we expected to clash with the CRM1-bound PKI NES (**Supplementary Figure 4-1b**). Indeed, the A541K mutation excluded binding of all tested PKI NES variants (**Figure 4-2e**). Likewise, NES binding was impaired when Cys<sup>528</sup>, the residue covalently modified by the CRM1 inhibitor leptomycin B (Kudo *et al.*, 1999), was mutated. While Ala, Thr and Val were still well tolerated at this position, NES binding was diminished in the C528S and clearly reduced in the C528W mutant (**Supplementary Figure 4-5d**).

Considered together, these experiments provided, independently of the above-described crystal structure, evidence for the hydrophobic cleft being the binding site not only for the SPN1 N-terminus, but also for the classic PKI NES.

#### 4.3.3 NMR structure of the CRM1-bound PKI $\Phi^0$ Leu NES

To elucidate the structure of a CRM1-bound PKI  $\Phi^0$ Leu NES without the SPN1 fusion module, we performed a series of NMR experiments. We prepared several isotope-labeled variants of the NES peptide and assembled them into NES·CRM1·RanGTP complexes. (See **Chapter 6** for more details on the sample preparation.) Initial attempts were hampered by poor signal-to-noise in the NMR spectra, probably due to the size of this complex ( $\approx 150$  kDa) and extensive broadening of the NES NMR signals by relaxation with protons of the exportin. Complexes with uniformly deuterated CRM1, however, yielded high-quality NMR spectra and allowed structural analysis of the bound NES and its docking to the exportin (**Supplementary Figure 4-4** and **Figure 4-3**) in a two-step protocol.

First, we determined the structure of the bound NES from NOE-derived intramolecular (intrapeptide) distance restraints and chemical shift-derived restraints for the backbone torsion angles (**Table 4-2**). In the second step, the NES was docked onto the CRM1·RanGTP complex, which we found to undergo surprisingly little structural change upon export ligand binding (see below, **Supplementary Figure 4-3**). For docking, the following additional restraints were used: **i**) NOE crosspeaks between the sulfhydryl proton of CRM1-Cys<sup>528</sup> and NES protons from the methyl groups of  $\Phi^3$ Leu<sup>11</sup> and  $\Phi^4$ Ile<sup>13</sup> as well as from the backbone amide of NES-Asp<sup>12</sup> (**Figure 4-3c** and **Supplementary Figure 4-5c**); and **ii**) ambiguous distance restraints (Nilges, 1993; Nilges and O'Donoghue, 1998) between CRM1 amide protons and buried PKI NES methyl protons, which were derived from NOEs detected in <sup>13</sup>C-edited NOESY-HMQC spectra. The obtained docking model was further refined based on the solvent accessibility of methyl groups from the NES side chains as probed by paramagnetic relaxation enhancement experiments (**Supplementary Figure 4-4b**, Madl *et al.*, 2009).



**Figure 4-3: Solution NMR structure of the CRM1-bound PKI  $\Phi^0$ Leu NES.**

(a) Overview of the NES·CRM1·RanGTP complex. RanGTP (red) and CRM1 (blue) are shown as a surface representation. The NES-binding site is colored as in **Figure 4-2**. The ensemble of NMR structures of the PKI  $\Phi^0$ Leu NES is shown in orange, docked onto the CRM1·RanGTP complex.

(b) Close-up of the CRM1-PKI  $\Phi^0$ Leu NES interaction.  $\Phi$  side chain carbons are colored according to **Figure 4-1a**, the non-interacting, solvent-accessible Leu<sup>6</sup> is shown in gray. The hydrophobic cleft, formed by HEAT repeats 11 and 12, is depicted as a blue cartoon.

(c) Strips from a  $^{13}\text{C}$ -edited NOESY-HMQC spectrum displaying intermolecular NOE crosspeaks between CRM1-Cys<sup>528</sup> H $\gamma$  and methyl protons of  $\Phi^3$ Leu<sup>11</sup>/  $\Phi^4$ Ile<sup>13</sup> from the NES. The  $^{13}\text{C}$  chemical shifts for the  $\Phi^3$  and  $\Phi^4$   $\delta$  methyl groups are 26.38 and 16.32 p.p.m., respectively. See also **Supplementary Figure 4-5**.

(d) **Upper:** Cartoon representation of the CRM1-bound NES as defined by the NMR secondary chemical shifts ( $\Delta\delta$ ). **Middle:** Difference of  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  secondary chemical shifts for the residues of the CRM1-bound NES. The average over residues 1-10 (red dashed line) serves as a reference point to define the 100% bound population. **Lower:** The secondary chemical shift analysis indicates that free NES molecules already populate the CRM1-bound  $\alpha$ -helical conformation to  $\approx 31\%$ . See also **Supplementary Figure 4-4**.

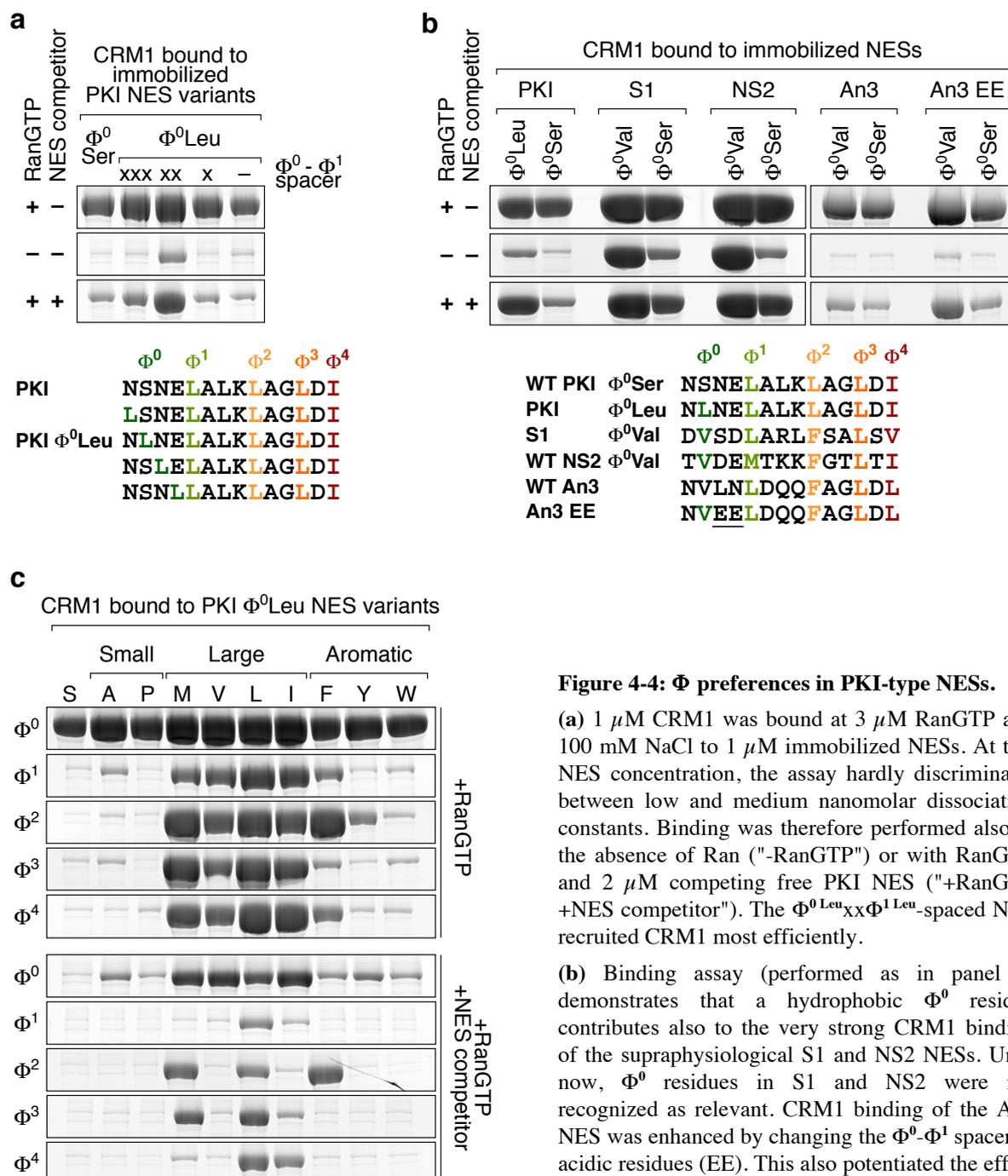
The final ensemble of lowest-energy structures is well converged (pair-wise backbone coordinate RMSD of 0.9 Å) and does not exhibit any restraint violations. Residues  $\Phi^0$ Leu<sup>1</sup>-Gly<sup>10</sup> of the NES form an amphipathic  $\alpha$ -helix, whereas  $\Phi^3$ Leu<sup>11</sup>- $\Phi^4$ Ile<sup>13</sup> adopt an extended conformation (for helix definition, see **Figure 4-3d**). Residues N-terminal of  $\Phi^0$  and C-terminal of  $\Phi^4$  are highly flexible in solution and do not interact with CRM1. This is also indicated by similar chemical shifts of the corresponding residues in the free and bound forms of the NES peptide (**Supplementary Figure 4-4a**). Importantly, the solution structure of the CRM1-bound PKI NES is fully consistent with the crystal structure of the SPN1<sup>15-360</sup>-fused PKI NES. This validation suggests that our crystallographic approach using chimeras is generally applicable for elucidating the atomic details of CRM1 binding to different NESs.

#### 4.3.4 *Solution conformation of the unbound NES peptide*

NMR experiments also allowed us probe the conformation of the NES peptide prior to exportin binding (**Figure 4-3d**). The spectra revealed that peptide molecules with an  $\alpha$ -helical conformation between residues  $\Phi^0$ Leu<sup>1</sup>-Gly<sup>10</sup>, corresponding to the CRM1-bound conformation, pre-exist in solution and that this conformation is already populated to  $\approx 31\%$ . This suggests that recruitment of cargo to CRM1 may not solely rely on a (CRM1-) induced fit mechanism. Instead, it appears that CRM1 selects those NES molecules from a pre-existing equilibrium that are already in a "bound conformation". We assume that a highly populated "bound conformation" of the free peptide is a prerequisite for high-affinity binding to CRM1.

#### 4.3.5 *A consensus for optimal NES binding to CRM1*

NESs vary considerably in their affinities for CRM1, but the sequence features underlying this variability are still poorly understood. We therefore addressed this issue systematically, starting with a characterization of the  $\Phi^0$  position of the PKI NES. We found that a  $\Phi^0$ xx $\Phi^1$  spacing is optimal (**Figure 4-4a**), and that  $\Phi^0$ Ile, Val or Met confer an even stronger CRM1 binding than a  $\Phi^0$ Leu (see below, **Figure 4-4c**). In the light of previously considered  $\Phi$  residues (Bogerd *et al.*, 1996; Zhang and Dayton, 1998; Kosugi *et al.*, 2008), it was quite surprising that even  $\Phi^0$ Ala, Pro or Tyr enhanced CRM1 binding as compared to the PKI wild type  $\Phi^0$ Ser variant.



**Figure 4-4:  $\Phi$  preferences in PKI-type NESs.**

(a) 1  $\mu$ M CRM1 was bound at 3  $\mu$ M RanGTP and 100 mM NaCl to 1  $\mu$ M immobilized NESs. At this NES concentration, the assay hardly discriminates between low and medium nanomolar dissociation constants. Binding was therefore performed also in the absence of Ran ("-RanGTP") or with RanGTP and 2  $\mu$ M competing free PKI NES ("RanGTP +NES competitor"). The  $\Phi^0$ Leu<sub>xxx</sub> $\Phi^1$ Leu-spaced NES recruited CRM1 most efficiently.

(b) Binding assay (performed as in panel a) demonstrates that a hydrophobic  $\Phi^0$  residue contributes also to the very strong CRM1 binding of the supraphysiological S1 and NS2 NESs. Until now,  $\Phi^0$  residues in S1 and NS2 were not recognized as relevant. CRM1 binding of the An3 NES was enhanced by changing the  $\Phi^0$ - $\Phi^1$  spacer to acidic residues (EE). This also potentiated the effect of the hydrophobic  $\Phi^0$  residue.

(c) Starting with the PKI  $\Phi^0$ Leu NES, all five  $\Phi$  positions were systematically mutated to indicated residues. These mutants were tested for CRM1 binding as in a, using the conditions "RanGTP" and "RanGTP+NES competitor". The control without Ran is shown in **Supplementary Figure 4-6**. Note that the individual  $\Phi$  pockets differ in their preferences. This is most obvious in the "RanGTP+NES competitor" panels.

It is curious that hydrophobic  $\Phi^0$  positions have not yet been described for high-affinity NESs such as the supraphysiological S1 and NS2 NESs. We therefore re-evaluated this issue. S1 was identified through a phage display selection for 15-mer peptides that show strong CRM1

binding even in the absence of RanGTP (Engelsma *et al.*, 2004). With the previously considered sequence (LARLFSALSV), which lacks a  $\Phi^0$  position, we could not reproduce the high affinity reported. However, we noticed a seemingly perfect  $\Phi^0$ Val in the sequence stretch N-terminal to the reported 4- $\Phi$  NES. Indeed, including this Val dramatically improved CRM1 binding, while a  $\Phi^0$ Val $\rightarrow$ Ser mutation abolished the effect (**Figure 4-4b**). The NS2 NES from Minute Virus of Mice was also published without considering a hydrophobic  $\Phi^0$  position (Engelsma *et al.*, 2008), and again, we found an upstream Val from the original sequence context to be required for very tight CRM1 binding (**Figure 4-4b**). We thus conclude that the NES consensus comprises not four but five hydrophobic positions ( $\Phi^0$ ,  $\Phi^1$ - $\Phi^4$ ) and that high-affinity NESs are characterized by a hydrophobic  $\Phi^0$  residue.

Our analysis of the An3 NES (Askjaer *et al.*, 1999) documents another critical feature of NESs with high affinity for CRM1. This NES contains an optimally spaced  $\Phi^0$ Val. Here, however, a change to Ser had surprisingly little effect on CRM1 binding (**Figure 4-4b**). This correlated with another difference: whereas the  $\Phi^0$ - $\Phi^1$  spacer is acidic in the SPN1 N-terminus as well as in the PKI  $\Phi^0$ Leu, NS2 and S1 NESs, it comprises two neutral residues in the An3 NES. Indeed, changing these neutral residues to acidic ones greatly improved the CRM1 interaction. In such context, also the  $\Phi^0$ Val $\rightarrow$ Ser mutant had a striking effect (**Figure 4-4b**). Acidic residues from the  $\Phi^0$ - $\Phi^1$  spacer engage in electrostatic interactions with K<sup>522</sup> and K<sup>560</sup> of CRM1 (Dong *et al.*, 2009b; Monecke *et al.*, 2009). This explains their positive effect on exportin binding. Apparently, they also position the  $\Phi^0$  residue within the hydrophobic cleft.

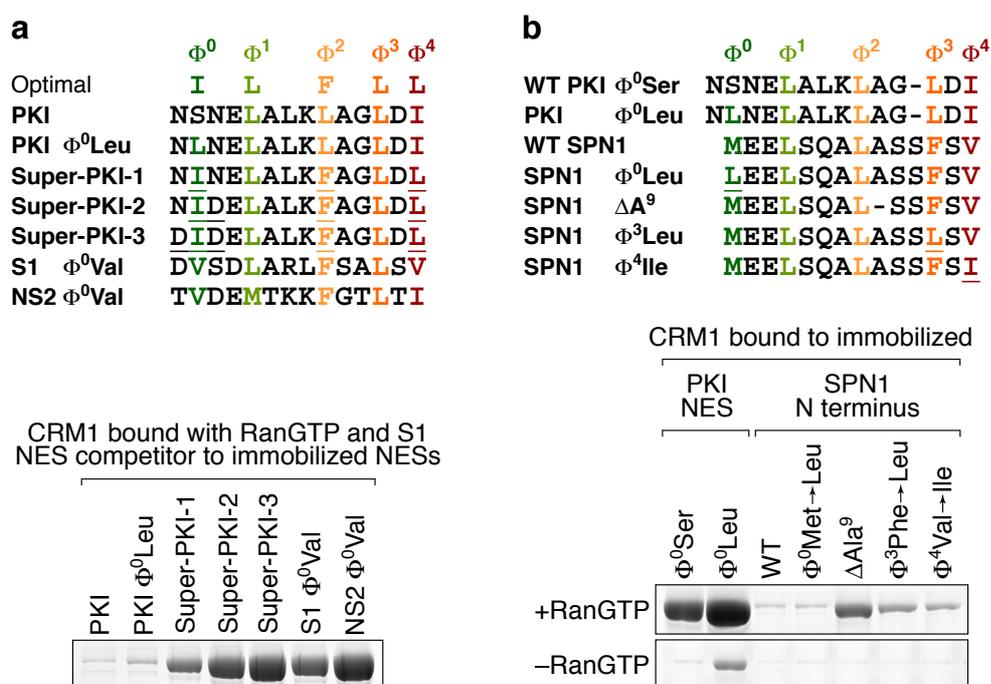
To complete our systematic analysis, we probed, in the context of the PKI NES, all five  $\Phi$ -binding pockets for their side chain preferences. The result of our survey is summarized in **Figure 4-4c** (see also **Supplementary Figure 4-6** and below **Figure 4-8a**). Leu is well accepted by all five  $\Phi$  pockets, but only in  $\Phi^1$  clearly preferred over all other hydrophobics. Notably, the  $\Phi^2$  pocket has the preference Phe  $\approx$  Met > Leu and it even accepts Tyr. These data suggest that the five  $\Phi$ -binding pockets are non-equivalent and that the strength of NES binding to CRM1 can be adjusted at each  $\Phi$  position by the choice of the hydrophobic residue.

#### 4.3.6 Rational design of a supraphysiological NES

With respect to maximum CRM1 binding, our PKI  $\Phi^0$ Leu NES was still suboptimal at  $\Phi^0$  (Leu instead of Ile),  $\Phi^2$  (Leu instead of Phe) and  $\Phi^4$  (Ile instead of Leu). We therefore

constructed a fully  $\Phi$ -optimized version and found that this Super-PKI-1 NES bound CRM1 nearly as strongly as the supraphysiological S1 NES (**Figure 4-5a**).

Around their  $\Phi^0$  positions, the NS2 and S1 NESs are more negatively charged than the PKI NES (**Figure 4-5a**). This is interesting, because a negatively charged  $\Phi^0$ - $\Phi^1$  spacer confers stronger CRM1 binding to the An3 NES (**Figure 4-4b**). We therefore gradually adjusted the original  $N\Phi^0NE\Phi^1$  sequence of the Super-PKI-1 NES to more acidic patterns. The effect was striking. The final Super-PKI-3 NES had a  $D\Phi^0DE\Phi^1$  pattern and a CRM1 affinity that clearly exceeded the affinity of S1 and matched that of the NS2 NES (**Figure 4-5a**). Structurally, this effect can be explained by additional electrostatic interactions of the acidic  $\Phi^0$  neighbors and Lys514 and Lys560 of CRM1 (**Supplementary Figure 4-1b**, **Supplementary Figure 4-2b**).



**Figure 4-5: Supraphysiological and attenuated NESs.**

(a) Conversion of the PKI NES into a supraphysiological NES. To maximize CRM1 affinity of the PKI  $\Phi^0$ Leu NES, all  $\Phi$  positions were changed to preferred residues (yielding Super-PKI-1 NES). Subsequently, Asp residues were introduced next to  $\Phi^0$ . The final Super-PKI-3 NES matched the CRM1 affinity of NS2 and outperformed the S1 NES. Binding assay was performed as in **Figure 4-4a**, using the supraphysiological S1 NES as a competitor.

(b) Mutants, which render the weakly CRM1-binding SPN1 N-terminus more similar to the PKI NES, were tested for CRM1 binding at 50 mM NaCl. The greatest gain in CRM1 binding was observed for the  $\Delta A^9$  mutant that adjusted the  $\Phi^2xxx\Phi^3$  to the PKI-like  $\Phi^2xx\Phi^3$  spacing. The  $\Phi^3F\rightarrow L$  and  $\Phi^4V\rightarrow I$  mutations that introduced more optimal  $\Phi$  residues also facilitated the interaction with CRM1.

Supraphysiological NESs naturally occur in certain viral proteins (Engelsma *et al.*, 2008). They strongly compete CRM1-mediated nuclear export and thus might even represent a viral strategy to disable a critical element of the host's cellular infrastructure. Our data suggest that such high CRM1 affinity results if an NES sequence closely matches our consensus for PKI-type NESs with 5  $\Phi$  positions and an acidic  $\Phi^0$  context (see below, **Figure 4-8a**).

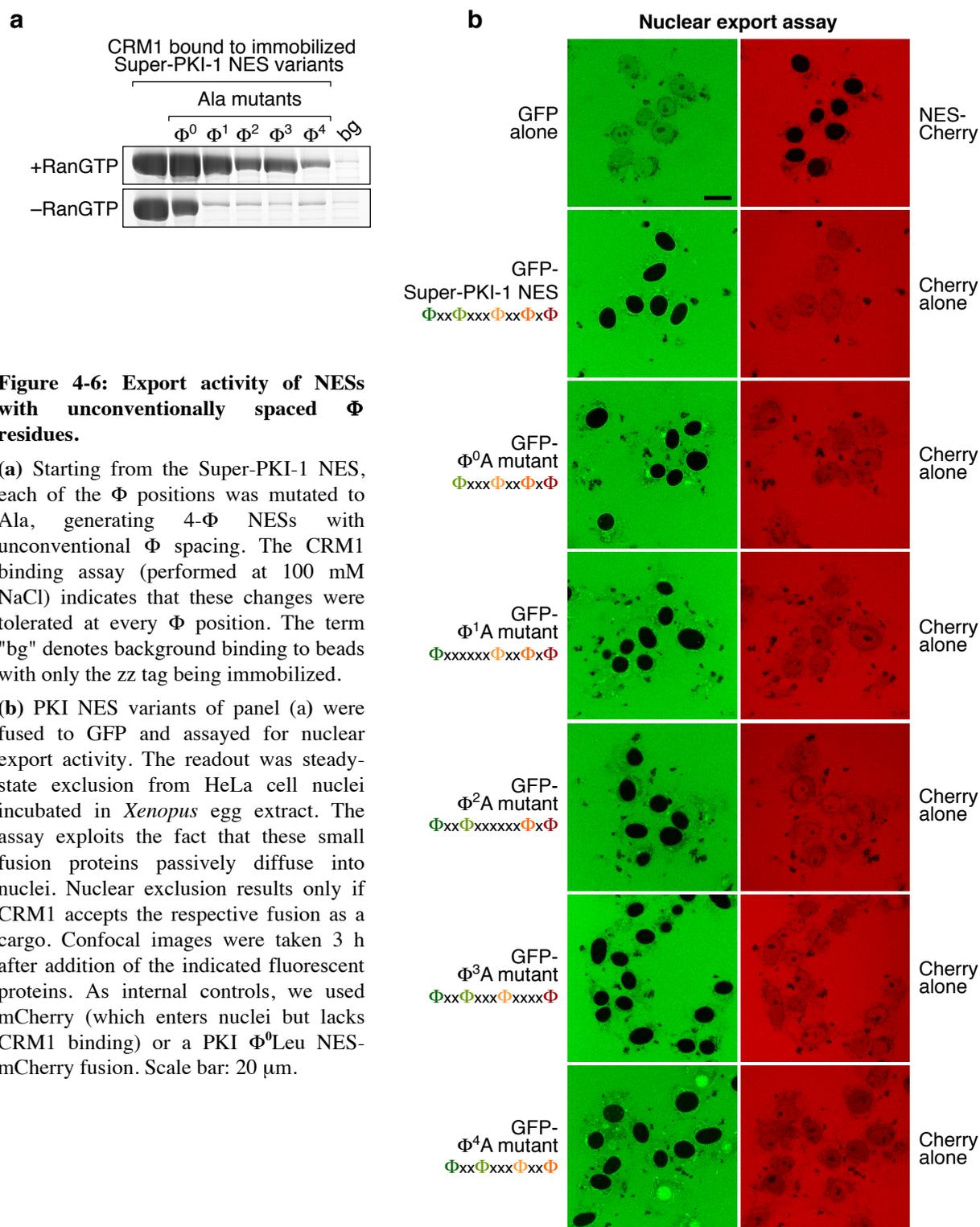
#### 4.3.7 NES attenuation

Cellular NESs are apparently not optimized to maximum CRM1 binding (Engelsma *et al.*, 2008; Kutay and Güttinger, 2005), because that too strong an NES·CRM1 interaction would cause the Ran-free form of CRM1 to stably bind the NES in the cytoplasm and thus to re-import such cargo into the nucleus. A typical cellular NES will therefore be attenuated as compared to the strongest possible ligand of the hydrophobic cleft. An extreme example of such attenuation is the NES-like N-terminus of SPN1. Its  $\Phi$  residues occupy virtually the same positions within the hydrophobic cleft as the equivalent residues of the PKI  $\Phi^0$ Leu NES (**Figure 4-2c**). Yet, the SPN1 N-terminus binds CRM1  $\approx 100$  times more weakly. Our analysis (**Figure 4-5b**) revealed that the SPN1 N-terminus is down-tuned in at least three regards as compared to the PKI NES, namely by containing suboptimal  $\Phi^3$ Phe and  $\Phi^4$ Val residues and by having a  $\Phi^2$ xxx $\Phi^3$  spacing instead of the optimal  $\Phi^2$ xx $\Phi^3$  spacing. Possibly, the  $\Phi^2$ xxx $\Phi^3$  spacing results in tension within the hydrophobic cleft and /or the NES peptide.

#### 4.3.8 Unconventionally-spaced NESs derived from the 5- $\Phi$ consensus

The allowed variability of the  $\Phi^2$ - $\Phi^3$  spacing is also interesting in that it creates a greater diversity of functional NESs. An even greater diversity arises from the observation that not all hydrophobic positions are necessarily essential, especially if the other  $\Phi$  positions are sufficiently strong. In the context of the  $\Phi$ -optimized Super-PKI-1 NES (**Figure 4-5a**), for example, each of the  $\Phi$  residues can be exchanged for Ala without abolishing CRM1 binding and export activity (**Figure 4-6a, b**). These  $\Phi \rightarrow$ Ala mutations lead to 4- $\Phi$  NESs with very unusual hydrophobic spacings, such as  $\Phi$ xx $\Phi$ xxx $\Phi$ xxxx $\Phi$  in the case of a  $\Phi^3 \rightarrow$ Ala mutant. Likewise, the  $\Phi^4 \rightarrow$ Ala mutation generates a peptide with a  $\Phi$ xx $\Phi$ xxx $\Phi$ xx $\Phi$  spacing that corresponds to an exotic "class 3 NES" described by Kosugi *et al.* (Kosugi *et al.*, 2008). Taken together, this suggests that the number of peptide sequences with CRM1-dependent NES activity is far larger than was previously thought. The data also show that an optimized

5- $\Phi$  NES tolerates losses at hydrophobic positions considerably better than the wild type 4- $\Phi$  version of the PKI NES (Wen *et al.*, 1995). This tolerance certainly reaches its limits, however, when the peptide contains residues that are incompatible with the required NES conformation (e.g. proline in the  $\Phi^1$ - $\Phi^3$  helical region) or cause steric exclusion from the central  $\Phi^1$ - $\Phi^3$  pockets.



**Figure 4-6: Export activity of NESs with unconventionally spaced  $\Phi$  residues.**

(a) Starting from the Super-PKI-1 NES, each of the  $\Phi$  positions was mutated to Ala, generating 4- $\Phi$  NESs with unconventional  $\Phi$  spacing. The CRM1 binding assay (performed at 100 mM NaCl) indicates that these changes were tolerated at every  $\Phi$  position. The term "bg" denotes background binding to beads with only the zz tag being immobilized.

(b) PKI NES variants of panel (a) were fused to GFP and assayed for nuclear export activity. The readout was steady-state exclusion from HeLa cell nuclei incubated in *Xenopus* egg extract. The assay exploits the fact that these small fusion proteins passively diffuse into nuclei. Nuclear exclusion results only if CRM1 accepts the respective fusion as a cargo. Confocal images were taken 3 h after addition of the indicated fluorescent proteins. As internal controls, we used mCherry (which enters nuclei but lacks CRM1 binding) or a PKI  $\Phi^0$ Leu NES-mCherry fusion. Scale bar: 20  $\mu$ m.

### 4.3.9 The "4- $\Phi$ " Rev NES docks in an unanticipated manner into all 5 $\Phi$ pockets

The HIV-1 Rev NES was initially identified as the Rev-activation domain (Malim *et al.*, 1991) and later shown to mediate nuclear export of the genomic HIV-1 RNA (Fischer *et al.*, 1995) in a CRM1-dependent manner (Fornerod *et al.*, 1997a). Its recognition by the exportin is particularly interesting, because here, the suggested " $\Phi^1$ " is separated from  $\Phi^2$  by a double-Pro spacer that should break any helicity. Its mode of CRM1 binding can therefore not be explained by the SPN1/PKI paradigm described above.

To investigate this problem, we used the chimera strategy, crystallized the corresponding Rev NES·CRM1·RanGTP complex and solved its structure. Considering that the Rev NES had so far been assumed to contain four  $\Phi$  residues, all being leucines, the structure revealed a fully unexpected mode of binding, namely that the Rev NES docks with five residues into all five  $\Phi$  pockets of CRM1 (**Figure 4-7a, c; Supplementary Figure 4-2**). The structure shows that the traditional alignment between the Rev and PKI-type NESs was correct only for the stretch from  $\Phi^2$ Leu<sup>9</sup> to  $\Phi^4$ Leu<sup>14</sup>, i.e. for the region that is indeed analogous between the two NES types (**Figure 4-7c**). The critical Leu<sup>6</sup>, however, is not the  $\Phi^1$  residue, but docks into the  $\Phi^0$  pocket. Surprisingly, the  $\Phi^1$  pocket accommodates the following Pro<sup>7</sup>. The spacer to  $\Phi^2$ Leu<sup>9</sup> does not comprise three residues as in the PKI NES or the SPN1 N-terminus, but only a single Pro. These shorter spacers between  $\Phi^0$  and  $\Phi^2$  are compensated for by an extended instead of an  $\alpha$ -helical structure. Strikingly, this allows the five newly defined  $\Phi$  residues of the Rev NES to occupy positions in 3D space that are nearly identical to the equivalent hydrophobics of SPN1 or PKI (**Figure 4-7c; compare Figure 4-2b, c; Figure 4-7a-c, Supplementary Figure 4-1; Supplementary Figure 4-2**).

Our Rev NES·CRM1 structure is supported by biochemical data. The notion that the Rev NES docks into the hydrophobic cleft, for example, is fully consistent with the inhibition of the Rev NES·CRM1 interaction by competing PKI NES or SPN1 (not shown), by leptomycin B (Fornerod *et al.*, 1997a), by the A541K and C528W mutations in CRM1 (**Fig. 2e and Supplementary Fig. 5d**), and by the I<sup>521</sup>L<sup>525</sup>F<sup>561</sup>F<sup>572</sup>→A quadruple mutation in CRM1 (Dong *et al.*, 2009b).



(d) (*previous page*) 1  $\mu\text{M}$  CRM1 was bound at 3  $\mu\text{M}$  RanGTP and 75 mM NaCl to indicated Rev NES variants. CRM1 binding was drastically reduced by mutations of leucines 6, 9, 12, or 14 to Ser, or change of the critical proline 7 to Gly, Ala or Leu. Pro<sup>8</sup>→Gly also impaired CRM1 binding. Changing Leu<sup>4</sup> also had some effect, probably because Leu<sup>4</sup> shields Leu<sup>6</sup> from the solvent (panel a).

(e) Overlay of the structures of the hydrophobic clefts from RanGTP·CRM1 complexes in their SPN1-, PKI NES- and Rev NES-bound forms (light blue backbones) or cargo-free form (orange). The side chains shown are those that contact the  $\Phi$  residues of the NESs. Note that the backbones show no major deviation and that structural differences are restricted to varying rotational states of the side chains and minor alterations in the intra-HEAT loops.

Mutation of the critical  $\Phi$  leucines of the Rev NES to serine reduced CRM1 binding to background levels (**Figure 4-7d**). The very strong effect of the  $\Phi^0\text{Leu}^6\rightarrow\text{Ser}$  mutation is particularly interesting, because it classifies the Rev NES as an export signal where a hydrophobic  $\Phi^0$  position not just tightens the interaction with CRM1, but appears essential for CRM1 binding. The critical  $\Phi^1\text{Pro}^7$  turned out to be very intolerant towards mutations (**Figure 4-7d**). Changes to Gly or Ala abolished CRM1 binding. Even the  $\Phi^1\text{Pro}\rightarrow\text{Leu}$  mutant showed strongly reduced CRM1 binding, emphasising that Pro<sup>7</sup> is not just a hydrophobic ligand, but also favors the special backbone conformation that positions other  $\Phi$  residues within the hydrophobic cleft. Indeed, this mutant reveals an unexpected strong interdependence of  $\Phi$  preference and backbone conformation:  $\Phi^1\text{Pro}$  is optimal in the Rev NES context but kills the PKI NES, whereas  $\Phi^1\text{Leu}$  disables the Rev NES but is optimal for PKI (summarized in **Figure 4-8**). Interestingly, also the change of the spacer-Pro<sup>8</sup> to Gly impaired the interaction with CRM1, while a change to Ala had no negative effect (**Figure 4-7d**). This probably reflects a negative entropic impact of the very flexible Gly residue, and it emphasizes once again the great influence of the NES backbone conformation on recognition by CRM1.

#### 4.3.10 *The NES-binding site is rigid and does not adapt to individual cargoes*

We also obtained a crystal structure of the binary RanGTP·CRM1 complex. It lacked a cargo molecule and crystallized in a different space group than the RanGTP·CRM1·cargo complexes (**Supplementary Figure 4-3** and **Table 4-1**). Nevertheless, Ran and CRM1 showed only negligible deviations from the structures containing cargo. In particular, the backbones of the NES-binding sites of all four RanGTP·CRM1 complexes (containing SPN1, the PKI NES or Rev NES fusions, or lacking cargo) are essentially indistinguishable (**Figure 4-7e**). Differences appear merely restricted to conformations of the side chains. Thus, the NES-binding site in RanGTP-bound CRM1 is rigid, i.e. it does not adapt to individual NES

sequences. Instead, the NES ligands themselves compensate different  $\Phi$  spacings by their conformational adaptation to the rigid NES-binding site.

The fact that RanGTP alone suffices to force the CRM1 molecule into the nuclear conformation emphasizes that Ran is the master regulator of the cargo-binding site. It further predicts that the conformation of CRM1 in export complexes with other cargoes will be very similar to the structures we observed.

a	PKI-class NESs		b	Rev-class NESs
x	Acidic residues preferred			
$\Phi^0$	I $\approx$ V $\approx$ M > L > A $\approx$ Y > F $\approx$ W > P		$\Phi^0$	
x <sub>2</sub>	Acidic residues preferred			
$\Phi^1$	L > I > V $\approx$ M > F > A > W	$\Phi^1$ (P) $\Phi^1$ (Y)	$\Phi^1$	= P >> L
x <sub>3</sub>			x	
$\Phi^2$	F $\approx$ M > L > I $\approx$ V > Y > W	$\Phi^2$ (P)	$\Phi^2$	
x <sub>2</sub> > x <sub>3</sub>			x <sub>2</sub>	
$\Phi^3$	L $\approx$ M > I > V > F > W $\approx$ A	$\Phi^3$ (P) $\Phi^3$ (Y)	$\Phi^3$	
x			x	
$\Phi^4$	L > I > M > V > F	$\Phi^4$ (P) $\Phi^4$ (A)	$\Phi^4$	

**Figure 4-8: Redefinition of the NES consensus.**

Our structural and biochemical analyses revealed that five and not just four  $\Phi$  positions need to be considered. In addition, the rules for  $\Phi^0$ - $\Phi^2$  spacing differ strikingly between PKI-class NESs (a) and Rev-class NESs (b).  $\Phi$  preferences for PKI-class NESs were elucidated in **Figure 4-4c** and **Supplementary Figure 4-6**. A given  $\Phi$  residue was scored as active if the tested NES bound CRM1 in a RanGTP-stimulated manner and stronger than the corresponding Ser variant. It is remarkable that while Pro is the only well accepted residue at  $\Phi^1$  of the Rev-type NES, it is forbidden at PKI  $\Phi^1$ .

## 4.4 Discussion

The great diversity of functional leucine-rich NESs poses an intriguing question, namely, how a single receptor, CRM1, can recognize all of these different peptides. In particular, it was difficult to comprehend how the exportin can cope with different spacings of the key  $\Phi$  residues. To address this recognition problem, we developed and validated a general crystallographic approach (based on NES-SPN1 chimeras) that allowed us to elucidate the structures of NES·CRM1·RanGTP complexes. We applied this approach to the export signals of the protein kinase A inhibitor PKI and the HIV-1 Rev protein, which not only represent the two prototypical NESs, but also mark (together with the SPN1 N-terminus) the known extremes of inter- $\Phi$  spacing.

Strikingly, we found that CRM1 uses the very same set of  $\Phi$  pockets for the recognition of each of these diverse NES peptides. The high-affinity state of the hydrophobic cleft has a very defined conformation that does not adapt to different NESs. We also found no indication for a sliding of  $\Phi$  residues within the hydrophobic cleft or for alternative  $\Phi$ -binding sites. Instead, different  $\Phi$  spacings are compensated for by allowing NESs to dock their  $\Phi$  residues with an  $\alpha$ -helical, near  $\alpha$ -helical or an extended backbone conformation into the  $\Phi$  pockets. This, along with the facts that single  $\Phi$  residues are dispensable (provided the others are close to optimal) and that each  $\Phi$ -binding pocket accepts various hydrophobic residues, account for the observed very wide range of specifically recognized NESs.

NESs of the PKI class are characterized by a  $\Phi^1\text{xxx}\Phi^2$  spacing. The traditional NES consensus, however, also included a deviating  $\Phi^1\text{xx}\Phi^2$  spacing. We now have reason to assume that this latter  $\Phi$  spacing is highly disfavored, because it is too short to bridge the distance between the  $\Phi^1$  and  $\Phi^2$  pockets in a PKI-typical  $\alpha$ -helical conformation, and too long for a favorable extended conformation. The  $\Phi^1\text{xx}\Phi^2$  pattern was historically derived from NESs containing a Rev-like  $\Phi$ -Pro-x- $\Phi$  motif (see Kutay and Güttinger, 2005 and citations therein). Our crystal structure now revealed that Pro is not part of an inter- $\Phi$  spacer, but the essential  $\Phi^1$  residue of the NES, whereas the preceding Leu is  $\Phi^0$ . In other words, Rev-class NESs are characterized by a  $\Phi^0\Phi^{1\text{Pro}}\text{x}\Phi^2$  and not a  $\Phi^1\text{xx}\Phi^2$  spacing (**Figure 4-8b**).

NES-like sequence patterns occur rather frequently in proteins, even in those that are not recognized by CRM1. This phenomenon has been studied in detail for the Abl tyrosine kinase (Hantschel *et al.*, 2005), whose actin-binding domain contains a perfect 4- $\Phi$  PKI-type NES stretch (LENNLRELQI, Taagepera *et al.*, 1998). When taken out of its original protein

context, this NES indeed confers efficient CRM1-dependent nuclear export. However, the complete domain cannot be recruited as a CRM1 cargo, because the three leucines of the "NES" are buried in the hydrophobic core and are thus not accessible for the exportin (Hantschel *et al.*, 2005; see also **Chapter 6**; Rittinger *et al.*, 1999; Kadlec *et al.*, 2004; la Cour *et al.*, 2004). Functional NESs should hence lack an appropriate "packing partner" in their sequence context. Operational classic NESs will therefore not occur within compactly folded domains, but rather at the N-terminus, C-terminus, or within an unstructured region of an export cargo. This consideration is obviously crucial for a precise prediction of export signals. From a protein folding perspective, the docking of an NES into CRM1's hydrophobic cleft can be seen as a *trans*-complementation that creates a new hydrophobic core.

The most conserved part of CRM1 (and probably of all nuclear transport components) comprises residues 495-595 and includes the hydrophobic cleft as well as flanking residues (**Supplementary Figure 4-7**). This part from human is 100% identical to that of fish (*Danio rerio*), 94% to *Drosophila melanogaster*, 91% to *Penicillium chrysogenum*, 82% to rice (*Oryza sativa*) and still very similar to that of distant protozoa such as *Toxoplasma gondii* (76% identity) or *Tetrahymena thermophila* (53% identity). This extreme conservation readily explains why the Rev and PKI NESs are functional from yeast to human. Remarkably, however, the conservation is not restricted to the residues that contact the  $\Phi$  side chains. Instead, strict evolutionary constraints were evidently imposed onto the entire HEAT repeats 11 and 12, which harbor the hydrophobic cleft, as well as onto flanking regions, in particular the acidic loop that contacts RanGTP. This probably reflects that the NES-binding site and the acidic loop cooperate in a highly optimized manner when CRM1 is switched between its cytoplasmic state (which is the form of low affinity for Ran and cargo) and its nuclear high-affinity state (Monecke *et al.*, 2009; Koyama and Matsuura, 2010). The necessary rigid coupling between RanGTP binding and export cargo loading probably requires the hydrophobic cleft to be rigid in its nuclear state. As a consequence, the NES-binding site cannot adapt to different  $\Phi$  spacings. Therefore, peptides will display CRM1-dependent NES activity only if a favorable conformation of their backbone can place a sufficient number of  $\Phi$  side chains into the rigid arrangement of  $\Phi$  pockets.

Table 4-1: Data collection and refinement statistics.

	PKI Φ <sup>0</sup> Leu NES export complex	HIV-1 Rev NES export complex I	HIV-1 Rev NES export complex II	RanGTP-CRM1 complex
<b>Data collection</b>				
Space group	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	C222 <sub>1</sub>
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	72.09, 223.73, 163.06	72.66, 224.62, 164.02	73.26, 225.90, 163.98	156.80, 216.16, 123.82
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90.00, 100.63, 90.00	90.00, 100.83, 90.00	90.00, 100.75, 90.00	90.00, 90.00, 90.00
Resolution (Å)	38.60-3.42 (3.58-3.42)*	39.00-2.80 (2.90-2.80)*	39.00-2.90 (3.00-2.90)*	36.50-3.35 (3.40-3.35)*
<i>R</i> <sub>rim</sub> <sup>a</sup>	23.30 (51.90)	15.20 (64.90)	16.20 (62.90)	7.2 (53.90)
<i>R</i> <sub>pim</sub> <sup>b</sup>	18.35 (34.35)	12.17 (43.13)	13.64 (39.04)	5.2 (48.70)
<i>I</i> / <i>σI</i>	5.68 (1.96)	7.05 (1.99)	8.97 (1.96)	18.0 (2.1)
Completeness (%)	83.9 (56.3)	90.0 (92.9)	95.4 (80.9)	94.1 (66.6)
Mosaicity (°)	1.86	0.90	0.85	0.38
Redundancy	3.0 (1.4)	2.9 (2.8)	4.2 (1.8)	4.1 (2.7)
<b>Refinement</b>				
Resolution (Å)	38.63-3.42 (3.47-3.42)	38.90-2.80 (2.83-2.80)	39.00-2.90 (2.93-2.90)	36.52-3.35 (3.40-3.35)
No. reflections	57482	114008	110353	28804
<i>R</i> <sub>work</sub> (%)	25.78 (29.91)	22.62 (28.28)	24.21 (31.12)	21.87 (38.11)
<i>R</i> <sub>free</sub> (%)	31.45 (36.81)	28.50 (34.89)	29.47 (37.78)	25.39 (41.97)
No. atoms	24087	25146	25190	9754
Protein	24021	24293	24325	9721
Ligand/ion	66	68	236	33
Water		785	629	
<i>B</i> -factors (Å <sup>2</sup> )				
Protein	152.2/140.2 <sup>#</sup>	52.2/34.3 <sup>#</sup>	50.3/33.9 <sup>#</sup>	151.3/149.7 <sup>#</sup>
Ligand/ion	137.2	32.1	51.0	129.6
Water		33.8	30.3	
R.m.s. deviations				
Bond lengths (Å)	0.013	0.006	0.003	0.002
Bond angles (°)	1.545	0.938	0.747	0.538
Corr. Coeff.	0.69 / 0.73 <sup>#</sup>	0.89 / 0.91 <sup>#</sup>	0.88 / 0.90 <sup>#</sup>	0.91 / 0.91 <sup>#</sup>

\*Values in parentheses are for highest-resolution shell.

<sup>a</sup>  $R_{rim} = \sum_{hkl} \{N/(N-1)\}^{1/2} \sum_i |I_i(hkl) - I(hkl)| / \sum_{hkl} \sum_i I_i(hkl)$ , with *N* being the number of times a given reflection has been observed. (Weiss and Hilgenfeld, 1997)

<sup>b</sup>  $R_{pim} = \sum_{hkl} \{1/(N-1)\}^{1/2} \sum_i |I_i(hkl) - I(hkl)| / \sum_{hkl} \sum_i I_i(hkl)$ , with *N* being the number of times a given reflection has been observed. (Weiss and Hilgenfeld, 1997)

<sup>#</sup>*B*-factor or map correlation coefficient for NES-binding site comprising CRM1 residues 511-574.

Note that the quality of maps and models for the NES regions of the Snurportin chimeras is above average of the data, as judged by *B*-factors and map correlation coefficients. See Online Methods for further details.

**Table 4-2: NMR and refinement statistics for the CRM1-bound PKI  $\Phi^0$ Leu NES.**

	PKI $\Phi^0$ Leu NES bound to CRM1·Ran complex
<b>NMR restraints</b>	
Distance restraints <sup>a</sup>	
Total NOE	37
Intra-residue	0
Inter-residue (NES)	
Sequential ( $ i - j  = 1$ )	0
Medium-range ( $ i - j  \leq 4$ )	9
Long-range ( $ i - j  > 4$ )	10
Intermolecular (NES-CRM1)	
unambiguous	3
ambiguous	15
Dihedral angle restraints (NES) <sup>b</sup>	
$\phi$	9
$\psi$	10
Solvent PREs (NES)	11
<b>Structure statistics</b>	
Violations (mean $\pm$ s.d.)	
Distance restraints (Å)	0.331 $\pm$ 0.106
Dihedral angle restraints (°)	0.000 $\pm$ 0.000
Max. dihedral angle violation (°)	0.00
Max. distance restraint violation (Å)	0.13
Deviations from idealized geometry	
Bond lengths (Å)	0.00344
Bond angles (°)	0.57480
Impropers (°)	1.46510
Average pairwise r.m.s. deviation (Å)	
Backbone	0.89 $\pm$ 0.25
Heavy	1.62 $\pm$ 0.22
Ramachandran plot (PROCHECK) <sup>c</sup>	
Most favored regions	88.3 %
Allowed regions	8.3 %
Generously allowed regions	3.3 %
Disallowed regions	0.0 %

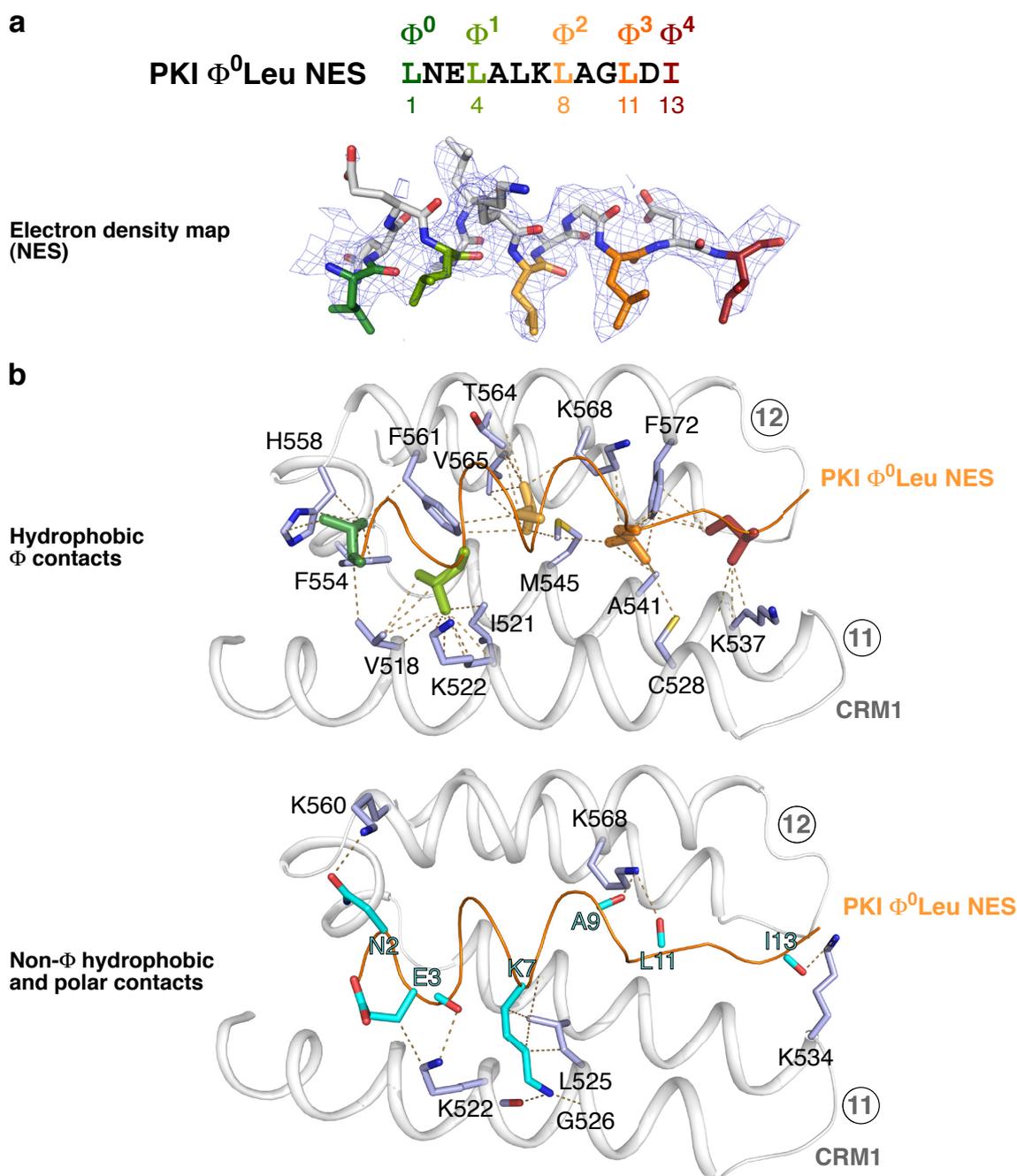
Statistics are given for the 10 lowest energy structures out of 1000 calculated for the PKI NES docking to CRM1. These structures were obtained after scoring against the solvent PRE data as described in the Methods. The CNS  $E_{\text{repe1}}$  function was used to simulate van der Waals interactions with an energy constant of 25 kcal mol<sup>-1</sup> Å<sup>-4</sup> using "PROLSQ" van der Waals radii (Linge *et al.*, 2003). RMSD and PROCHECK values apply for residues 1-11 and 1-13, for the PKI NES peptide and the docking to CRM1, respectively.

<sup>a</sup> Distance restraints were employed with a soft square-well potential using an energy constant of 50 kcal mol<sup>-1</sup> Å<sup>-2</sup>.

<sup>b</sup> Torsion angle restraints derived from TALOS (Shen *et al.*, 2009) were applied to  $\phi, \psi$  backbone angles using energy constants of 200 kcal mol<sup>-1</sup> rad<sup>-2</sup>.

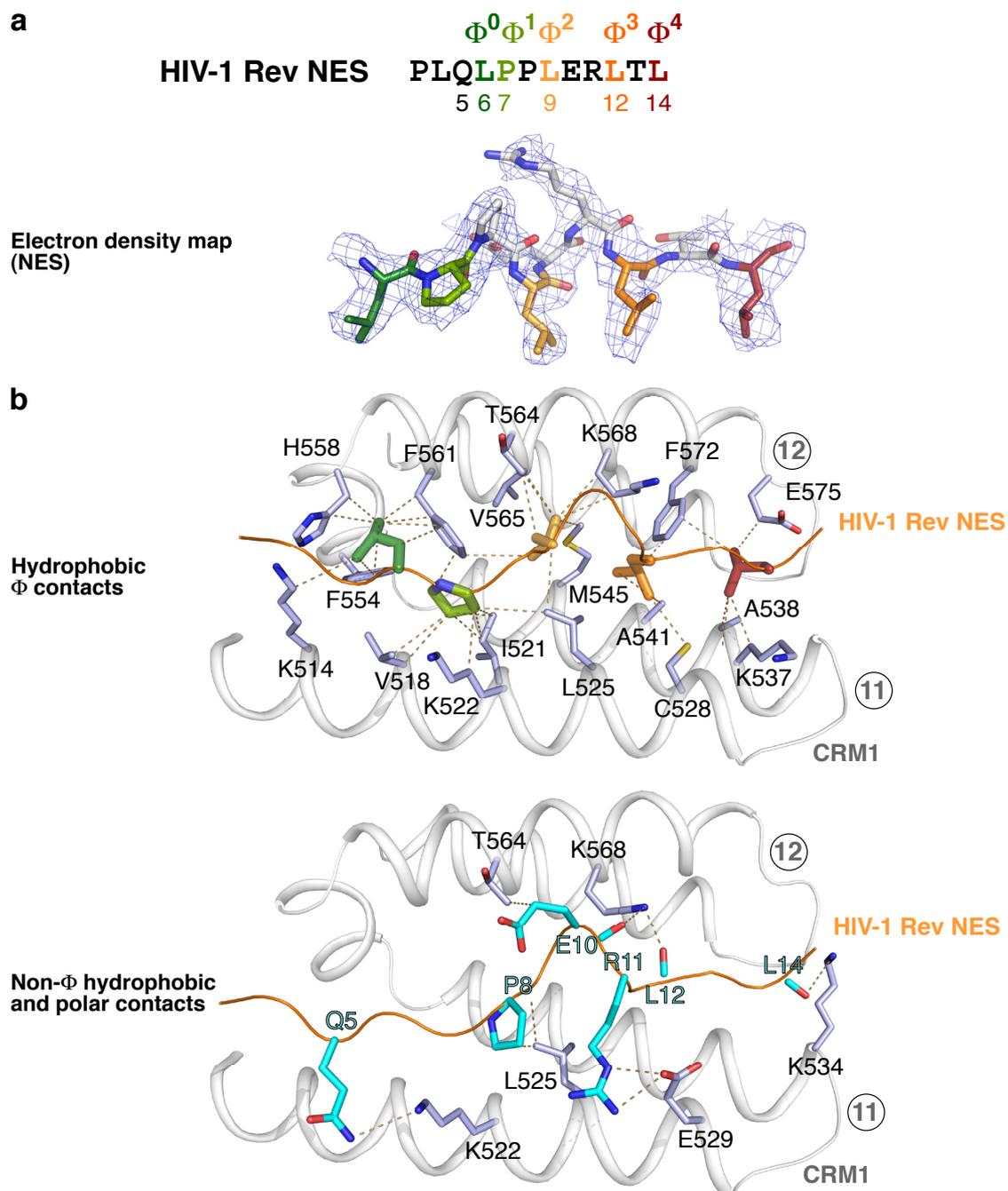
<sup>c</sup> PROCHECK (Laskowski *et al.*, 1996) was used to determine the quality of the structure.

## 4.5 Supplementary Figures and Tables

**Supplementary Figure 4-1: Details of the PKI  $\Phi^0$ Leu NES-CRM1 interaction.**

(a) **Upper:** Sequence of the PKI  $\Phi^0$ Leu NES.  $\Phi$  residues are colored according to **Figure 4-1a**. **Lower:** Panel displays the 2Fo-Fc electron density map (blue mesh, contoured at 1.0  $\sigma$ ) for the PKI  $\Phi^0$ Leu NES (shown as sticks) in the chimeric RanGTP-CRM1-NES-SPN<sup>115-360</sup> complex.  $\Phi$  residues are colored according to the shown sequence. Note that all  $\Phi$  residues are well defined in the map. In all panels, dark blue marks nitrogen, oxygen is shown in light red and sulfur is colored in yellow.

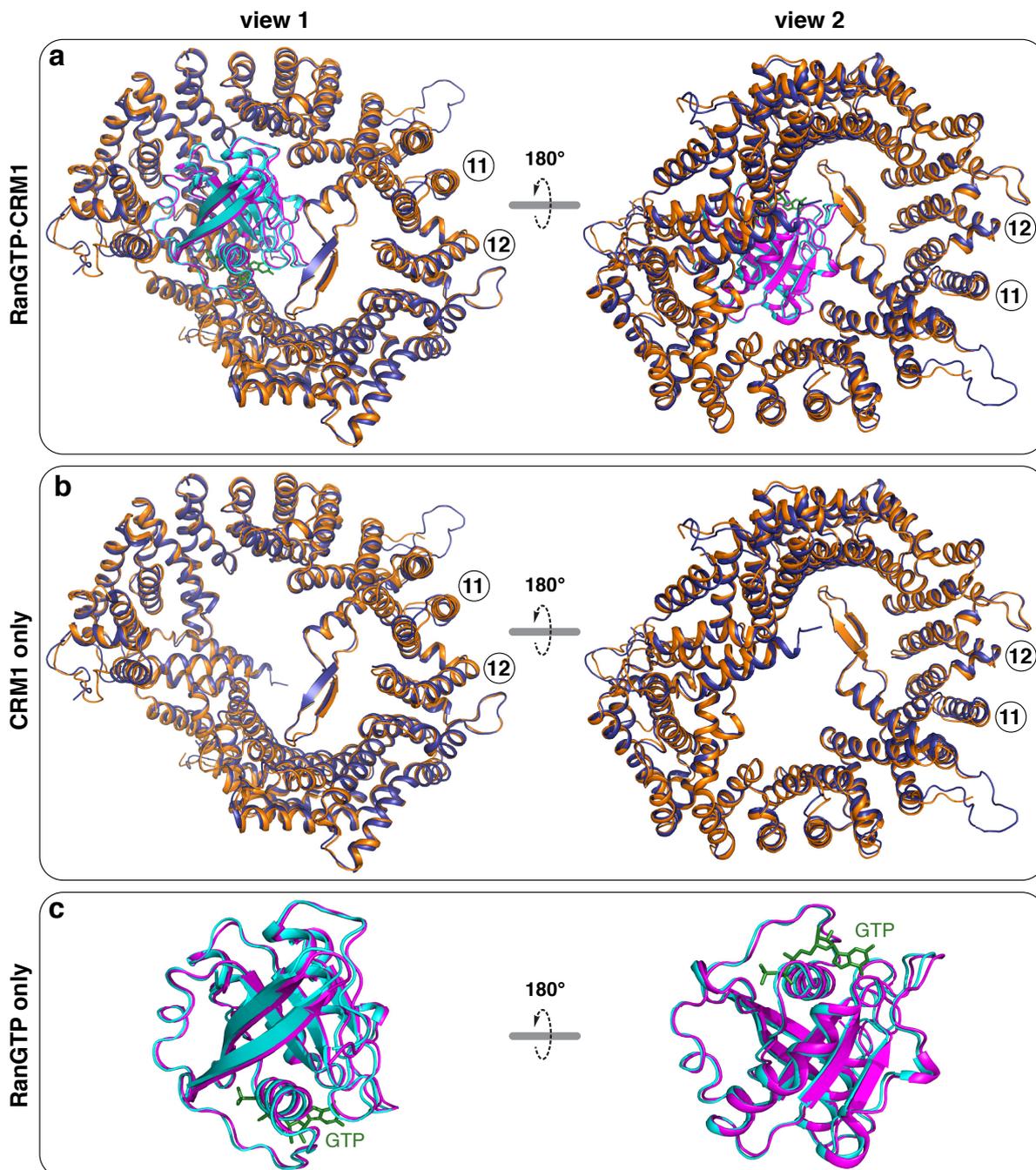
(b) CRM1 HEAT repeats 11-12 (gray cartoon) are shown with the NES peptide bound (backbone traced in orange). NES-binding residues of CRM1 are depicted as blue sticks. Dashed lines link interacting atoms. Lines pointing onto backbones indicate contacts to carbonyl-carbons or amide groups. **Upper:** Panel shows the hydrophobic contacts of the  $\Phi$  residues (distance  $\leq 4.0$  Å). The respective  $\Phi$  residues are shown as sticks, the color code is explained in a. **Lower:** Panel shows the non- $\Phi$  hydrophobic (distance  $\leq 4.0$  Å) as well as polar (distance  $\leq 3.8$  Å) contacts of NES residues (cyan sticks).



**Supplementary Figure 4-2: Details of the HIV-1 Rev NES·CRM1 interaction.**

For explanation, see **Supplementary Figure 4-1**.

## OVERLAY OF:

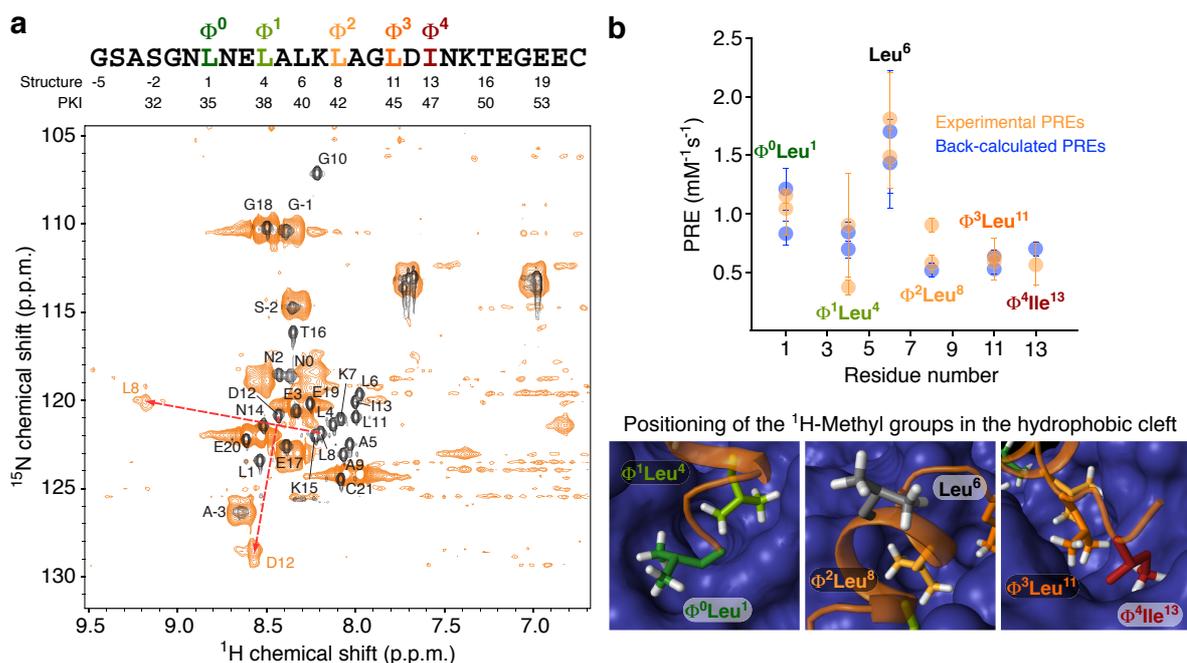
Cargo-free **RanGTP·CRM1** (binary RanGTP·CRM1 complex)Cargo-bound **RanGTP·CRM1** (ternary RanGTP·CRM1·SPN1 complex, SPN1 being omitted)

**Supplementary Figure 4-3: Comparison of the overall structures of RanGTP·CRM1 from the binary (cargo-free) RanGTP·CRM1 complex and the ternary RanGTP·CRM1·SPN1 complex (PDB-ID 3GJX, chains F and D, Monecke *et al.*, 2009).**

**a)** Pictures show an overlay of RanGTP·CRM1 from the indicated complexes (in cartoon representation). The color code is explained on top of the figure. The overlay is based on a C $\alpha$  alignment of the CRM1 molecules (RMSD = 0.843 Å). HEAT repeats forming the hydrophobic cleft (11 and 12) are labeled.

**(b)** As in (a), but here RanGTP was omitted for clarity.

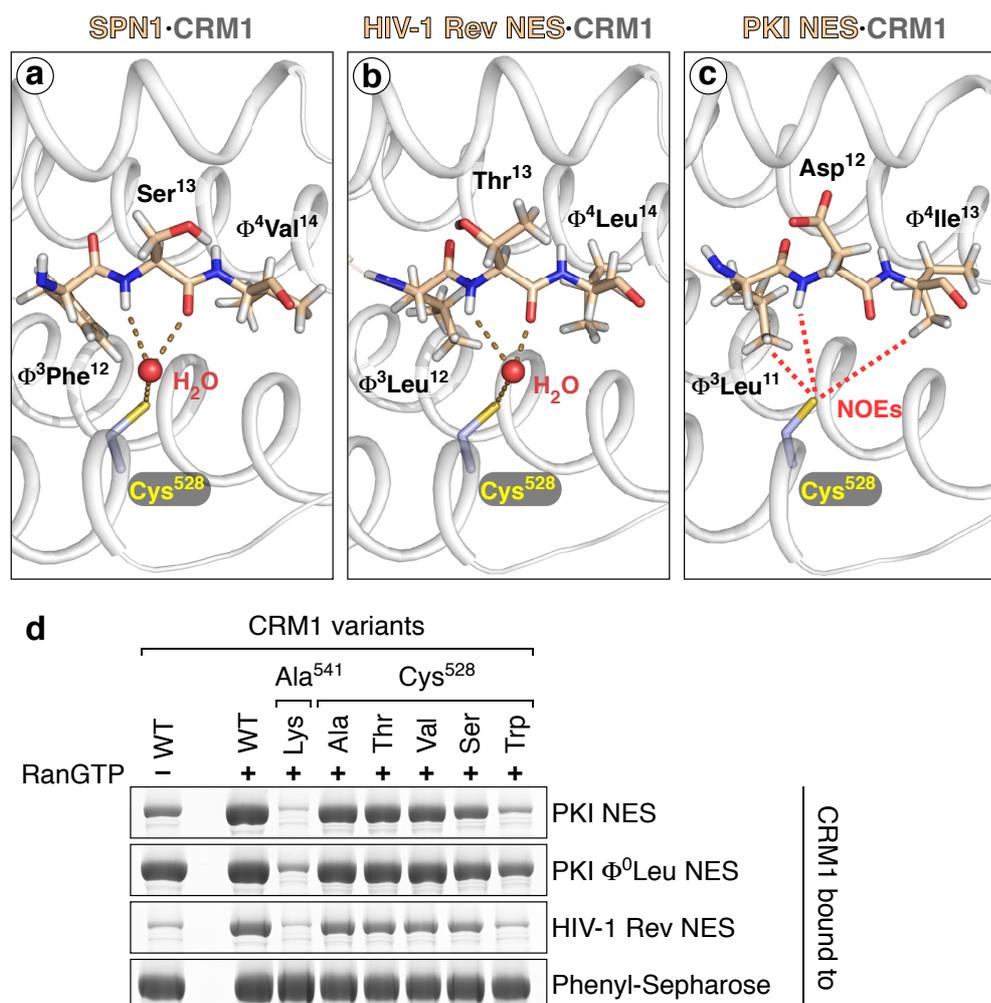
**(c)** As in (a), but here CRM1 was omitted and Ran was enlarged. GTP is shown for orientation (green sticks).



**Supplementary Figure 4-4: NMR-spectroscopic analysis of the free and CRM1·RanGTP-bound PKI  $\Phi^0$ Leu NES.**

(a) Overlay of the  $^1\text{H}$ , $^{15}\text{N}$ -HSQC NMR spectrum of the unbound PKI NES peptide (black) and the  $^1\text{H}$ , $^{15}\text{N}$ -CRINEPT-HMQC spectrum of the PKI NES peptide in the export complex (orange). Signals are labeled according to the shown residue numbers. Arrows indicate changes in the chemical shift of selected residues that occur when the NES is incorporated in the export complex.

(b) **Upper:** Diagram shows solvent PRE (paramagnetic relaxation enhancement) data for the CRM1-bound NES. PRE values positively correlate with the solvent-accessibility of methyl groups. Experimental (orange) and back-calculated  $^1\text{H}$  PREs (blue) for methyl groups are displayed. **Lower:** The panels show how the  $^1\text{H}$ -methyl groups of the indicated residues are positioned in the hydrophobic cleft of CRM1. The backbone of the PKI NES is shown in orange, side chains are color-coded as in **Figure 4-3b**, protons are colored in light gray. CRM1 is shown as a surface representation (blue).



**Supplementary Figure 4-5: Evidence for a hydrogen bonding network involving CRM1<sup>Cys528</sup> and NES peptide backbones.**

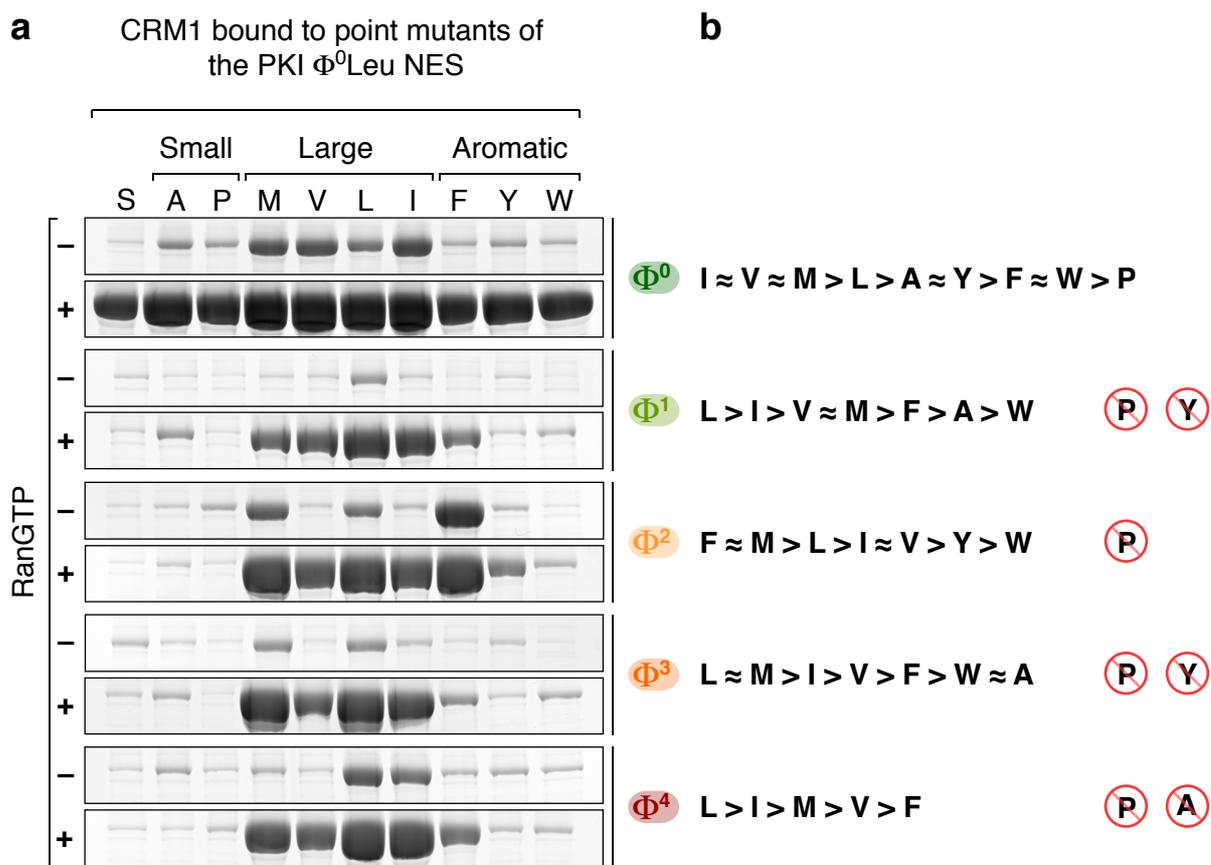
Panels (a)-(c) show cartoon representations of CRM1 (gray) and the indicated ligand (light orange), focussing on the region around CRM1<sup>Cys528</sup>. Selected residues are depicted as sticks (with oxygen in red, nitrogen in blue, sulfur in yellow and protons in gray). The higher-resolution structures of the SPN1 (PDB-ID 3GJX, Monecke *et al.*, 2009) and HIV-1 Rev NES complexes revealed a conspicuous water molecule (red sphere) in the vicinity of CRM1<sup>Cys528</sup>.

(a) The dashed lines illustrate the hydrogen bonding network that involves this water molecule, CRM1<sup>Cys528</sup> and the backbone of SPN1<sup>Ser13</sup>.

(b) The panel illustrates the analogous hydrogen bonding network for the HIV-1 Rev NES·CRM1·RanGTP complex.

(c) The PKI  $\Phi^0$ Leu NES·CRM1·RanGTP electron density map could not resolve water molecules. However, we observed NOE cross peaks for a cysteine sulfhydryl protected against solvent exchange (Figure 4-3c), typical for stable hydrogen bonding interactions. This cysteine can be assigned to CRM1-Cys<sup>528</sup>, which is located in the vicinity of the PKI  $\Phi^0$ Leu NES peptide (panel d) and is the only cysteine within the hydrophobic cleft. Thus, the NOE pattern (cross peaks between CRM1-Cys<sup>528</sup> H $\gamma$  and side chain methyl protons of NES-Leu<sup>11</sup>/Ile<sup>13</sup> as well as the backbone amide of NES-Asp<sup>12</sup>, illustrated by red dashed lines) is consistent with an equivalent hydrogen bonding network in the PKI  $\Phi^0$ Leu NES·CRM1·RanGTP complex.

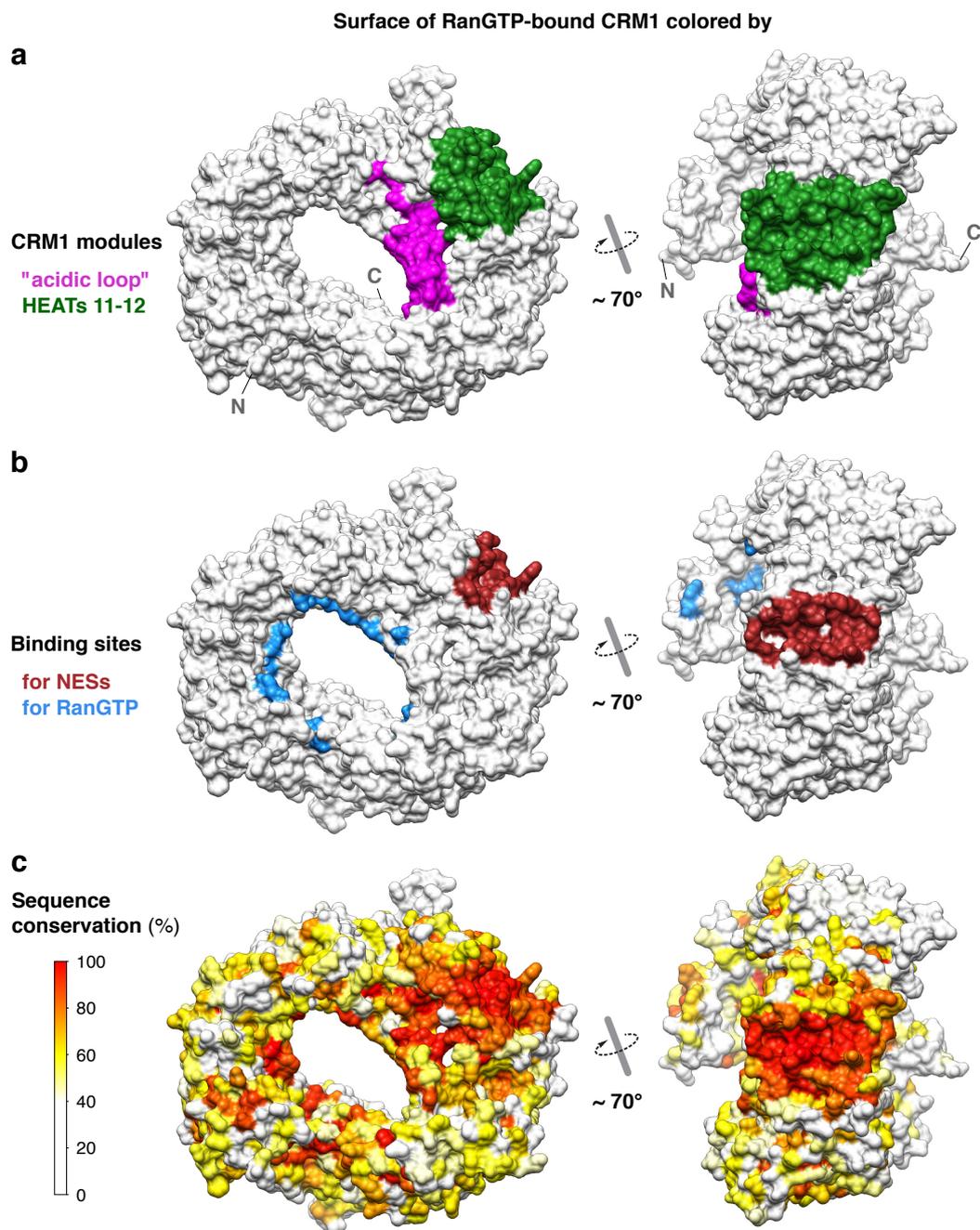
(d) Binding of the indicated CRM1 variants to the specified export ligands. Changing CRM1<sup>Cys528</sup> to other small residues with hydrophobic potential (Ala, Thr, Val) did not reduce cargo binding detectably. A change to the more hydrophilic Ser, however, caused some reduction and a change to the bulky residue Trp resulted in a clear decrease in cargo binding. See Figure 4-2e for further details. These results confirm that PKI and Rev NES bind in close vicinity of CRM1<sup>Cys528</sup> and CRM1<sup>Ala541</sup>. See also Supplementary Figure 4-1 and Supplementary Figure 4-2.



**Supplementary Figure 4-6: Assessment of the amino acid specificities of the  $\Phi$  pockets.**

(a) Each  $\Phi$  residue of the zz-tagged PKI  $\Phi^0$ Leu NES was systematically mutated to the indicated hydrophobic residues (single letter codes) and tested for CRM1 binding in the absence or presence of RanGTP as described for **Figure 4-4a**. Ser mutants were used as negative controls, because Ser should not engage in hydrophobic interactions with  $\Phi$  pockets. We considered residues as clearly  $\Phi$ -active if (I) CRM1 recruitment to the mutant peptide was stimulated by RanGTP and (II) if CRM1 binding was more pronounced than for the corresponding Ser mutant. Substitutions that reduced bound CRM1 to background levels can be regarded as disallowed.

(b) Panel shows the NES consensus derived from a and **Figure 4-4c** (see also **Figure 4-8**).



#### Supplementary Figure 4-7: Sequence conservation of CRM1.

CRM1 (in export complex conformation) is shown as a surface representation, ligands have been omitted for clarity.

(a) The "acidic loop" (residues 423-448) is colored in magenta, HEAT repeats 11 and 12 (residues 510-595) in green. The locations of N- and C-terminus are shown for orientation.

(b) The Ran-binding surface is depicted in blue and the NES-binding surface is shown in brown.

(c) CRM1 surface is colored according to sequence conservation. Note that the NES-binding site is the most conserved part of CRM1. Striking conservation further extends to the acidic loop that links Ran- and NES-binding. The surface representation was generated with UCSF Chimera (Pettersen *et al.*, 2004) from Clustal W-aligned CRM1 sequences (Larkin *et al.*, 2007). The alignment was essentially based on all full-length CRM1 sequences that were identifiable in the non-redundant NCBI protein sequence database, the most distant being CRM1 from *Trichomonas vaginalis* (22 % identity to mouse CRM1). For sets of highly similar sequences however, only one orthologue was included in order to avoid bias by overrepresenting individual clades, leaving 58 sequences.

**Supplementary Table 4-1: Crystallization and cryoprotection conditions.**

Complex	Crystallization		Cryoprotection
PKI $\Phi^0$ Leu NES-SPN1·CRM1·Ran(Q69L) <sup>5-180</sup>	100 mM Tris/HCl pH 8.08-8.30, 10-14% (w/v) PEG 1000 (20 °C)	sitting drop 1 $\mu$ l prot. (4 mg ml <sup>-1</sup> ) + 1 $\mu$ l reservoir	reservoir + 20% (v/v) 1,2-propanediol
HIV-1 Rev NES-SPN1·CRM1·Ran(Q69L) <sup>5-180</sup> (crystal I)	100 mM Tris/HCl pH 7.8, 16% (w/v) PEG 1000 (16 °C)	hanging drop 2 $\mu$ l prot. (8 mg ml <sup>-1</sup> ) + 2 $\mu$ l reservoir	reservoir + 5% (w/v) PEG 1000 + 12% (v/v) glycerol
HIV-1 Rev NES-SPN1·CRM1·Ran(Q69L) <sup>5-180</sup> (crystal II)	100 mM Tris/HCl pH 7.8, 16% (w/v) PEG 1000, 2 mM phenol (20 °C)	hanging drop 2 $\mu$ l prot. (8 mg ml <sup>-1</sup> ) + 2 $\mu$ l reservoir	reservoir + 5% (w/v) PEG 1000 + 12% (v/v) glycerol
CRM1·Ran(Q69L) <sup>1-180</sup>	100 mM Tris/HCl pH 8.5, 12% (w/v) PEG 4000 (20→4 °C)	sitting drop 1 $\mu$ l prot. (12.2 mg ml <sup>-1</sup> ) + 1 $\mu$ l reservoir	reservoir + 11% (v/v) 2,3-butanediol

## 4.6 Methods

### 4.6.1 Protein Expression and Purification

2YT medium supplemented with 2% (w/v) glycerol and 30 mM K<sub>2</sub>HPO<sub>4</sub> was used for standard expression cultures. Mouse CRM1<sup>1-1071</sup> (full-length) was expressed as an N-terminal His<sub>10</sub>-zz fusion in *E. coli* BLR using a fermenter (Labfors 3, Infors AG). Expression was performed at 16 °C for  $\approx$ 20 hours (0.1 mM IPTG). A total of 1 mM PMSF and 1 mM EDTA were added to the culture prior to centrifugation. Cells were resuspended in lysis buffer (50 mM Tris/HCl pH 7.5, 500 mM NaCl, 1 mM EDTA, 2 mM imidazole) supplemented with 100  $\mu$ M Amidino-PMSF and 100  $\mu$ M DFP and lysed in the presence of 5 mM DTT and 20 U ml<sup>-1</sup> Benzonase (Novagen). The lysate was cleared by centrifugation and the protein bound to a Ni<sup>2+</sup>-chelate affinity column. To remove CRM1-associated chaperones, the column was washed with lysis buffer supplemented with 100 mM KCl, 10 mM Mg(OAc)<sub>2</sub> and 1 mM ATP. CRM1 was eluted with lysis buffer/200 mM imidazole. The eluate (diluted to 100 mM NaCl) was passed through a Heparin Sepharose column (GE Healthcare), applied to a Q-Sepharose column (GE Healthcare) and eluted in a concentration gradient of NaCl. After cleavage of the His<sub>10</sub>-zz-tag by His-tagged TEV protease, the tag, protease and residual contaminants were removed via another Ni<sup>2+</sup>-column. The flow-through was then subjected to a Superdex 200 gel filtration column (HiLoad 26/60, GE Healthcare, equilibrated in 50 mM Tris/HCl pH 7.5, 500 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 5 mM DTT).

We truncated the C-terminus of Ran, since it is disordered in other transport receptor complexes (Vetter *et al.*, 1999a; Cook *et al.*, 2007), it destabilizes the GTP-bound form of

Ran and counteracts the interactions with transport receptors (Richards *et al.*, 1995). Human Ran<sup>1-180</sup>(Q69L) and Ran<sup>5-180</sup>(Q69L) were expressed as an N-terminal His<sub>10</sub>-zz fusion in *E. coli* BLR (0.1 mM IPTG for  $\approx$  16 hours at 20 °C). Protease inhibitors were applied as described for CRM1<sup>1-1071</sup>. Cells were lysed in 50 mM K-Phosphate pH 7.0, 500 mM NaCl, 5 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 2 mM imidazole, 2 mM DTT (lysis buffer). The protein was purified by Ni<sup>2+</sup>-chelate affinity chromatography. 20  $\mu$ M GTP were included in all buffers subsequent to cell lysis. The His<sub>10</sub>-zz tag was cleaved off by TEV protease during dialysis of the Ni<sup>2+</sup>-eluate to lysis buffer. The flow-through of a second Ni<sup>2+</sup>-column was further purified by gel filtration chromatography (HiLoad 26/60 Superdex 200, GE Healthcare, equilibrated in 50 mM K-Phosphate pH 7.0, 500 mM NaCl, 5 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 2 mM DTT). To confirm the nucleotide state of Ran, the nucleotide was dissociated from the protein by addition of de-ionized urea and the nucleotide pattern analyzed by anion exchange chromatography on a MonoQ 5/50 GL column (GE Healthcare).

All human SPN1 and SPN1-derived constructs (SPN1<sup>1-360</sup>, the NES-SPN1<sup>15-360</sup> chimera and SPN1  $\Delta$ N) were expressed as N-terminal His<sub>10</sub>-zz fusions in *E. coli* BLR (0.1 mM IPTG for  $\approx$  16 hours at 18 °C). Protease inhibitors were applied as described for CRM1<sup>1-1071</sup>. Cells were lysed in 50 mM Tris/HCl pH 7.5, 200 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 2 mM imidazole, 5 mM DTT (lysis buffer). The protein was purified by Ni<sup>2+</sup>-chelate affinity chromatography. The eluate (diluted to 50 mM NaCl) was bound to a Q-Sepharose column (GE Healthcare) and eluted in a concentration gradient of NaCl. After cleavage of the His<sub>10</sub>-zz tag by His-tagged TEV protease, the protein was passed over a second Ni<sup>2+</sup>-chelate affinity column and polished by gel filtration (HiLoad 26/60 Superdex 200, GE Healthcare, equilibrated in 50 mM Tris/HCl pH 7.5, 200 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 5 mM DTT).

His<sub>10</sub>-zz-tagged NES peptides (PKI NES: *Homo sapiens*, Rev NES: HIV-1, NS2 NES: Minute Virus of Mice, An3 NES: *Xenopus laevis*, S1: synthetic) were expressed in *E. coli* BLR (1 mM IPTG for  $\approx$  5 hours at 37 °C) and purified under denaturing conditions (lysis buffer: 50 mM Tris/HCl pH 8.0, 6 M guanidinium hydrochloride, 1 mM EDTA, 5 mM DTT) by Ni<sup>2+</sup>-chelate affinity chromatography (elution in 50 mM Tris/HCl pH 7.5, 8 M urea, 50 mM NaCl, 200 mM imidazole, 1 mM EDTA, 5 mM DTT), followed by dialysis to 50 mM Tris/HCl pH 7.5, 50 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 2 mM DTT for re-folding. The His<sub>10</sub>-zz tag used as a control in binding assays (**Figure 4-6a** and **Figure 4-7d**) was obtained by TEV protease cleavage of an NES fusion construct. His<sub>10</sub>-zz-tagged human Ran<sup>1-180</sup>(Q69L) (**Figure 4-1b**) was prepared according to the procedure described for Ran<sup>1-180</sup>(Q69L) and Ran<sup>5-180</sup>(Q69L). All

CRM1 forms used for **Figure 4-2e** and **Supplementary Figure 4-5d** were expressed as N-terminal His fusions (as described for CRM1<sup>1-1071</sup>) and purified by Ni<sup>2+</sup>-chelate affinity chromatography.

eGFP-spacer-NES fusions (N-terminally His<sub>10</sub>-tagged, C-terminal Cys) were expressed in *E. coli* TOP10F' (0.2 mM IPTG for  $\approx$ 16 hours at 18 °C) and bound to a Ni<sup>2+</sup>-chelate affinity column (lysis buffer: 50 mM K-Phosphate pH 7.0, 200 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 2 mM imidazole, 5 mM DTT). The proteins were directly eluted onto a home-made thiopyridine-activated, SH-reactive Sepharose matrix (pH adjusted to 7.5) to select for full-length NES species. Elution from SH-Sepharose was performed with lysis buffer (pH adjusted to 7.5, supplemented with 5 mM DTT). The eluate was dialyzed to 50 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 2 mM DTT. mCherry and PKI  $\Phi^0$ Leu NES-mCherry (TEV protease-cleavable His<sub>14</sub>-fusions) were expressed and purified following the procedure described for mCherry (Frey and Görlich, 2009).

To obtain NMR spectra of good quality, uniform deuteration of CRM1 was essential. For this, we optimized the fully deuterated minimal medium to support robust expression of CRM1 and cell growth to high optical densities (OD<sub>600</sub> 6-10, manuscript in preparation). CRM1 expression was induced with 0.05 mM IPTG for nine (!) days at 16 °C (see **Chapter 6**). Purification of deuterated CRM1 followed the procedure described for the non-deuterated protein. The His<sub>10</sub>-zz-tagged PKI  $\Phi^0$ Leu NES was produced with various methyl-protonation schemes (Gardner and Kay, 1998; Tugarinov *et al.*, 2006) and purified as the unlabeled peptide fusion constructs. The tag was cleaved off by TEV protease and removed via a Ni<sup>2+</sup>-column. The peptide was then loaded onto a reversed-phase HPLC column (C-18, 218TP1022, GraceVydac), eluted by increasing the concentration of acetonitrile in the presence of 0.5% (v/v) TFA and lyophilized.

#### 4.6.2 Protein sample preparation for crystallization and NMR analysis

For crystallization, complexes were reconstituted in 20 mM Tris/HCl pH 7.5, 50 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 5 mM DTT as described (Monecke *et al.*, 2009).

Samples for NMR spectroscopy were prepared by mixing the NES peptide, CRM1 and Ran<sup>5-180</sup>(Q69L) in a 1:1.4:1.4 ratio, followed by dialysis against 20 mM Na-Phosphate pH 6.8, 50 mM NaCl, 2 mM MgCl<sub>2</sub>, 5 mM DTT, 30  $\mu$ M GTP and concentration to 0.1-0.2 mM

(export complex). NMR spectra of the free NES were recorded at 1 mM peptide concentration. Isotope labeling is described in **Chapter 6**.

### 4.6.3 Crystallization and structure determination.

All complexes were crystallized by vapor diffusion (**Supplementary Table 4-1**).

The crystallization of the binary CRM1·RanGTP complex was initially started with a ternary PKI  $\Phi$  <sup>0</sup>Leu NES·CRM1·RanGTP complex. However, crystal growth was most probably initiated only after a contaminating bacterial protease had removed the NES peptide as well as 21 residues from the C-terminus of CRM1. This interpretation is supported by the following observations: **i)** Crystals were observed only after more than five months. **ii)** After this time, CRM1 in the used protein preparation was found to be quantitatively cleaved. **iii)** More highly purified, protease-free preparations of the intact ternary complex were fully refractory to crystallization. **iv)** The structure of CRM1 in the obtained crystals is clearly defined up to His<sup>1050</sup>, whereas the complete C-terminus (ending at residue 1071) can probably not be accommodated into the observed crystal lattice. And **v)** no electron density for an NES peptide was traceable. See also **Table 4-1** and **Supplementary Table 4-1**.

X-ray oscillation photographs were integrated and scaled using XDS (Kabsch, 1993).

Crystals containing the NES-SPN1 chimera were isomorphous with the described crystals of the SPN1·CRM1·RanGTP complex (Monecke *et al.*, 2009, PDB-ID 3GJX). Iterative cycles of refinement were done using CNS (Brunger, 2007). To avoid overfitting, refinement was based on slow-cooling simulated annealing (SA) torsion angle dynamics (TAD) combined with standard minimization and individually restrained B-factor refinement. The structures were refined using strong non-crystallographic symmetry (NCS) restraints between equivalent residues, which were later relaxed for regions showing different conformations. The final refinement steps were performed with PHENIX (Adams *et al.*, 2002) using the NCS restraints and the set of reflections for calculating  $R_{\text{free}}$  that had also been used for CNS refinement. *TLS* (*translation/libration/screw*) refinement was applied to derive the anisotropic motions of the domains. Water molecules for the HIV-1 Rev models had been added manually, based on stringent criteria, i.e. only if both difference electron density maps ( $2F_o-F_c$  contoured at  $1\sigma$  and  $F_o-F_c$  contoured at  $3\sigma$ ) showed a peak within a distance ranging between 2.2 and 3.5 Å to O or N atoms of the protein molecules. Deletion of these water molecules increased both R

and  $R_{\text{free}}$  factors by a similar value, indicating that the water molecules included in the current model are justified and are not simply masking errors.

The structure of the binary RanGTP·CRM1 complex was solved by molecular replacement using PHASER (McCoy, 2007) with the crystal structures of Ran<sup>9-179</sup> and CRM1<sup>12-1055</sup> (PDB-ID 3GJX, Monecke *et al.*, 2009) as separated search models. Refinement was conducted using the same strategy as was used for the cargo complexes (refinement with CNS, followed by use of PHENIX with grouped B-factor refinement and TLS).

Models were built in COOT (Emsley and Cowtan, 2004) based on  $\text{SIGMA}_A$ -weighted SA omit maps and/or difference electron density maps. Molecular contacts were analyzed with the CONTACTS program of the CCP4 suite (version 6.1.3, Collaborative Computational Project, 1994). Map correlation coefficients in **Table 4-1** were calculated with PHENIX (against 2Fo-Fc map). All main figures were prepared with PyMOL (<http://www.pymol.org>).

#### 4.6.4 NMR Spectroscopy and NMR structure calculation.

Protein sample preparation is described in the supplementary information online. NMR spectra were recorded at 298 K using Avance 900 and Avance III 600 Bruker NMR spectrometers, equipped with cryogenic triple resonance gradient probes. Assignments were carried out using a combination of CRINEPT (Riek *et al.*, 1999), Methyl-TROSY (Tugarinov *et al.*, 2003), NOE-based (mixing times 30-140 ms) and <sup>13</sup>C direct-detected (Bermel *et al.*, 2006; Oh *et al.*, 1988) NMR experiments (see supplementary information online). <sup>13</sup>C direct-detected spectra were recorded at 298 K using Avance III 600 (equipped with a TCI probehead) and Avance III 700 (equipped with a CP-TXO probehead) spectrometers. Spectra were processed with NMRPipe/Draw (Delaglio *et al.*, 1995) and analyzed with Sparky 3 (T. D. Goddard & D. G. Kneller, University of California, San Francisco, USA) and NMRView (Johnson and Blevins, 1994).

For measurement of solvent PREs, Gd(DTPA-BMA) was added to the protein sample to final concentrations from 0.5 to 4.5 mM. Proton  $T_1$  relaxation times were obtained from a series of saturation-recovery 2D <sup>1</sup>H,<sup>13</sup>C-HMQC and <sup>1</sup>H,<sup>15</sup>N-HMQC spectra. Typically, eight such 2D data sets were acquired with recovery delays of 0.2, 0.5, 1.0, 1.5, 2.0, 4.0, 6.0 and 8.0 s. For two delays, duplicates were acquired for statistical purposes. To obtain the relaxation times, the peak intensities were fitted to

$$I = I_0 \left( 1 - e^{-\frac{t}{T_1}} \right)$$

using the NMRViewJ (version 8.0) rate analysis tool. PREs were calculated as reported (Madl *et al.*, 2009; Pintacuda and Otting, 2002) and are represented by the slope of the relaxation rate as function of the concentration of the paramagnetic agent. PREs were back-calculated and averaged for the ten lowest-energy structures by numerical integration of the paramagnetic environment (Madl *et al.*, 2009).

Structures were calculated with the standard ARIA/CNS setup (Linge *et al.*, 2003; Linge *et al.*, 2001) using NOEs and backbone dihedral angles derived from  $^{13}\text{C}$  secondary chemical shifts (TALOS+) (Shen *et al.*, 2009). The peptide was docked to the CRM1·RanGTP crystal structure (PDB-ID 3GJX, Monecke *et al.*, 2009) using all intra-peptide restraints, ambiguous PKI NES - CRM1 NOEs (Nilges, 1993; Nilges and O'Donoghue, 1998) and unambiguous PKI NES - CRM1 Cys<sup>528</sup> NOEs. The protein backbone of CRM1 was fixed while CRM1 side chains were left flexible for optimal accommodation of the PKI NES in the hydrophobic cleft. From the 1000 structures generated, the 100 structures of lowest energy were evaluated against solvent PRE data (**Supplementary Figure 4-4b**), and ranked according to the correlation between measured and back-calculated PREs. The 10 best structures in terms of the correlations are displayed and used to calculate the structural statistics (**Table 4-2**).

#### 4.6.5 NMR assignment

Conventional TROSY versions of triple resonance experiments failed for the bound PKI NES residues in the complex. This is due to the relatively low concentrations around 0.1-0.2 mM (limited by the solubility of the protein complex) and the adverse relaxation properties that reflect the size of the complex (calculated correlation time: 90 ns (HYDRONMR, García de la Torre *et al.*, 2000). Sample stability further limited NMR measurements to ambient temperature. Resonance assignment was carried out using several specifically isotope labeled samples, namely [U- $^2\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ , I/L/V  $^1\text{H}$ ,  $^{13}\text{C}$ -methyl] and [U- $^2\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ , Ala- $^1\text{H}$ , Leu/Val  $^1\text{H}$ ,  $^{13}\text{C}$ -methyl] PKI NES, respectively. Unique amino acids (e.g. Ile13) were used as starting points for the resonance assignment in  $^{13}\text{C}/^{15}\text{N}$  edited NOESY-HMQC experiments acquired at different mixing times. Two continuous methyl-walks involved residues Leu1, Leu4 and Leu8, Leu11, Ile13, respectively. For one leucine methyl group (Leu6) no NOE cross peaks to any other methyl groups were observed.

As the standard route towards backbone dihedral angle restraints from  $^{13}\text{C}$  chemical shifts (Shen *et al.*, 2009) via conventional TROSY versions of triple resonance experiments failed, we used  $^{13}\text{C}$  direct-detected experiments to assign backbone  $^{13}\text{C}\alpha$  and side chain  $^{13}\text{C}$  chemical shifts. A combination of  $^{13}\text{C}$ - $^{13}\text{C}$  TOCSY and  $^{13}\text{C}$ - $^{13}\text{C}$  NOESY experiments provided complementary chemical shift data. Whereas cross peaks for residues located within the flexible termini of the PKI NES peptide are observed in the  $^{13}\text{C}$ - $^{13}\text{C}$  TOCSY experiment, the  $^{13}\text{C}$ - $^{13}\text{C}$  NOESY provides chemical shifts for the CRM1-bound residues that tumble with the correlation time of the complex. The efficiency of the dipolar transfer between  $^{13}\text{C}$  spins increases with the correlation time (Fischer *et al.*, 1996), and residues located within the flexible termini did not give rise to detectable cross peaks at the used mixing times of 500 ms. The  $^{13}\text{C}$ - $^{13}\text{C}$  NOESY yields cross peaks exclusively for the directly bonded carbons. This is due to the generally much shorter distances between directly bonded carbon and the strong ( $r^{-6}$ ) distance dependency of the dipolar transfer.  $^{13}\text{C}$  chemical shifts of leucine and isoleucine methyl groups were used as starting point for the  $^{13}\text{C}$  resonance assignment. The remaining amino acids were assigned based on their characteristic  $^{13}\text{C}$  chemical shifts and spin systems.  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  chemical shifts were then used as input to derive backbone dihedral angle restraints (Shen *et al.*, 2009). Further details of the assignments and structure calculation will be described in a manuscript by T. Madl *et al.* (in preparation).

#### 4.6.6 Nuclear export assay

HeLa cells were prepared as described (Adam *et al.*, 1990; Stüven *et al.*, 2003), with some changes of the protocol. Briefly, cells were grown to  $\approx 95\%$  confluency, washed with PBS, detached with a Citrate/EDTA solution, and washed 3x with ice-cold transport buffer (20 mM HEPES/KOH pH 7.5, 110 mM KOAc, 5 mM  $\text{Mg}(\text{OAc})_2$ , 0.5 mM EGTA, 250 mM sucrose). Cell membranes were permeabilized with 25  $\mu\text{g ml}^{-1}$  digitonin (in ice-cold transport buffer) and nuclei recovered by centrifugation. After two washes, nuclei were resuspended in transport buffer (without sucrose) containing 5% (w/v) glycerol and 0.5 M trehalose, and were slowly frozen at  $-80^\circ\text{C}$ . We used *Xenopus laevis* egg extract (Leno and Laskey, 1991) for nuclear export assays, as it stabilizes the nuclei over a period of hours (Stüven *et al.*, 2003). Nuclei were mixed with egg extract (supplemented with 0.5  $\mu\text{M}$  CRM1 and an ATP/GTP-regenerating system) and incubated at  $22^\circ\text{C}$  for 30 min before addition of 2  $\mu\text{M}$  of the indicated fluorescent proteins. Imaging was performed with a confocal laser-scanning

microscope system (SP5; Leica) using the 488-, and 561-nm laser lines and a 63x-NA 1.3 Plan-Apochromat glycerol objective (Leica).

#### 4.6.7 Fluorescence polarization (FP) assay

10 nM of the PKI  $\Phi^0$ Leu NES (Fluorescein maleimide-labeled via its C-terminal Cys) were incubated for 4 h with the indicated concentrations of CRM1 and GTP-Ran (Q69L). The assay buffer contained 20 mM Tris/HCl pH 7.5, 130 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM DTT, 0.005% (w/v) digitonin, 0.1 mg ml<sup>-1</sup> BSA and a RanGTP-regenerating system (0.5 mM GTP, 10 mM PEP, 0.5  $\mu$ M pyruvate kinase, 0.5  $\mu$ M RanGEF/RCC1). FP measurements were performed with a "Victor 3" multilabel plate reader (PerkinElmer). Apparent equilibrium dissociation constants were derived by non-linear regression using GraphPad Prism (GraphPad Software, Inc.).

#### 4.6.8 Software used for preparation of figures

Structural representations have been generated with PyMOL (<http://www.pymol.org>) and, where indicated, with UCSF Chimera (Pettersen *et al.*, 2004). Figures were prepared using Photoshop and Illustrator (Adobe Systems Inc.).

### 4.7 Accession codes

Protein Data Bank: Coordinates and structure factors have been deposited with accession codes 3NBY (PKI  $\Phi^0$ Leu NES·CRM1·RanGTP), 3NBZ (HIV-1 Rev NES·CRM1·RanGTP, crystal I), 3NC0 (HIV-1 Rev NES·CRM1·RanGTP, crystal II) and 3NC1 (CRM1·RanGTP). The NMR data of the CRM1·RanGTP-bound PKI  $\Phi^0$ Leu NES have been deposited with accession code 2L1L.

### 4.8 Acknowledgements to Chapter 4

We thank A. Berndt and A. Dickmanns for performing the final purification step and the crystallization of the RanGTP·CRM1 complex, T.A. Rapoport, S. Frey, S. Güttler, C. Siebenhaar and B. Hülsmann for critical reading of the manuscript, B. Hülsmann for preparing the *Xenopus* egg extract, the Bavarian NMR Centre (BNMRZ) for NMR

measurement time, the staff of synchrotron beamlines at the SLS (Villigen, X10SA, PXII) and the ESRF (Grenoble, ID 23-2) for assistance during data collection and M. Wahl, V. Pena and J. Wawrzinek for setting up the crystallization facility at the MPI. This study was funded by the MPG, the Boehringer Ingelheim Fonds and the Alfried Krupp von Bohlen und Halbach Foundation (fellowships to T. Güttler), EMBO (fellowship to T. Madl), the Austrian Science Fund (FWF, Schrödinger fellowship to T. Madl), the European Commission contract 3D Repertoire (LSHG-CT-2005-512028 to M. Sattler) and the contract EU-NMR (No. RII3-026145).

## CHAPTER 5

**Concluding Remarks and Perspectives**

The past couple of years have seen tremendous progress in the structural characterization of Ran-dependent nuclear export factors (Cook and Conti, 2010). The crystal structure of the RanGTP·CRM1·Snurportin1 complex (**Chapter 3**) revealed how CRM1 recognizes a cargo that carries a very complicated export signature, which also includes features of a folded domain (see also Dong *et al.*, 2009b). Moreover, it suggests a mechanism by which RanGTP promotes cargo binding to CRM1. The work presented in **Chapter 4** explains how the same exportin can recognize extremely divergent NESs. This study redefines the NES consensus and explains why NESs differ in their affinities for CRM1. Together with the structures presented in this work, crystal structures of Cse1p/CAS, Exportin-t, Exportin 5 as well as those of the bidirectional transport receptor Importin 13 (Matsuura and Stewart, 2004; Cook *et al.*, 2005; Cook *et al.*, 2009; Okada *et al.*, 2009; Bono *et al.*, 2010) begin to reveal principles of exportin-cargo recognition and the specific roles of Ran in cargo loading. In the following, I will highlight some of these concepts and briefly sketch the major future goals in the characterization of the nuclear transport receptors that have been in the focus of this work (i.e. CRM1 and Exp4).

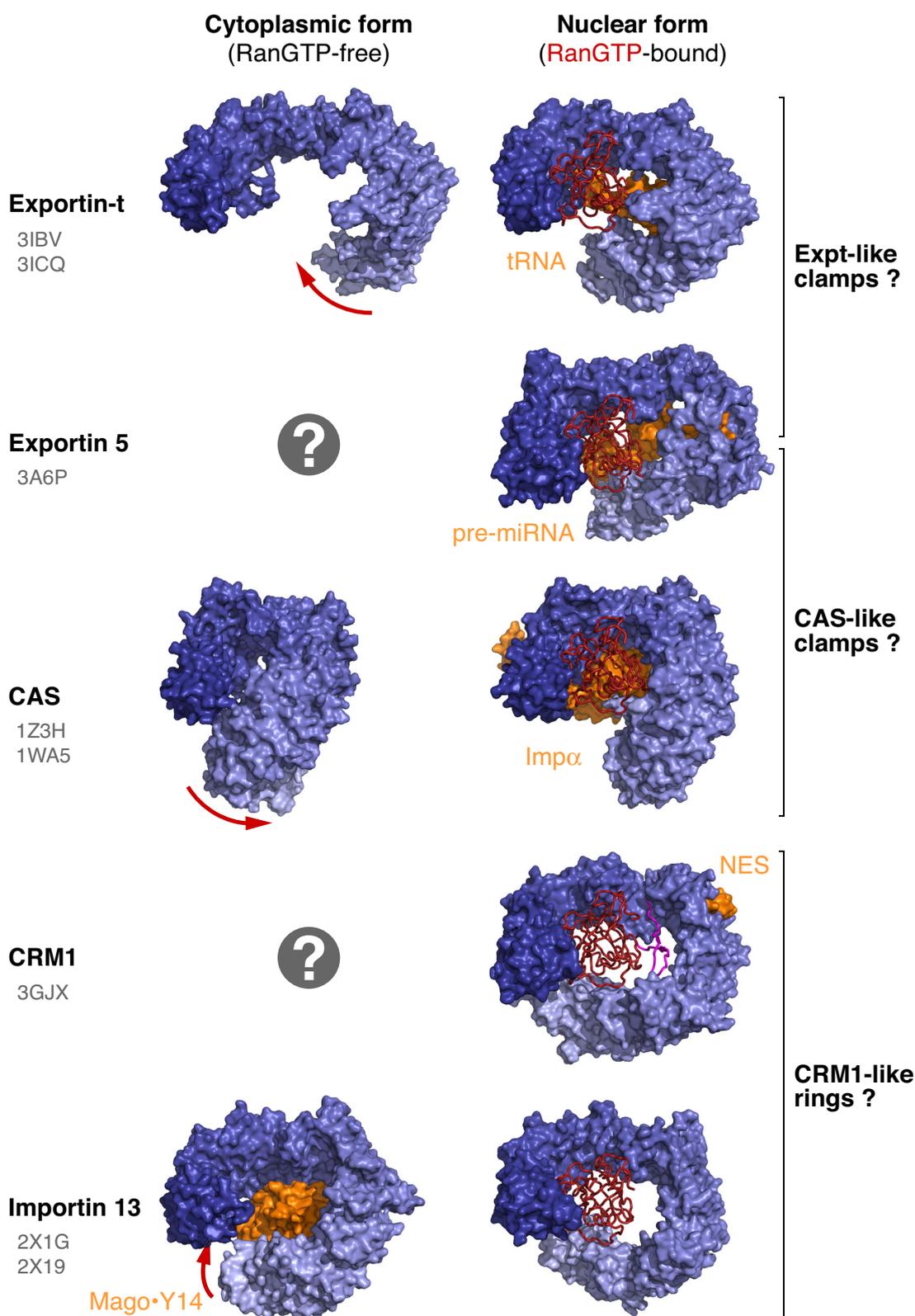
**5.1 How does Ran promote export cargo loading?**

In all exportin·RanGTP complexes crystallized so far, GTP-bound Ran confers high cargo affinity by interacting simultaneously with both the N-terminal as well as the C-terminal arches of the export receptors. However, the associated structural changes in the exportins appear to vary considerably. In general, these rearrangements are either the result of discrete hinge movements (e.g. in the case of CAS), the consequence of cumulative changes over the entire molecule (e.g. in Exp-t) or a combination of both (such as in Imp13). Among the exportins shown in **Figure 5-1**, Exp-t probably experiences the most dramatic Ran-driven structural changes (Cook *et al.*, 2009). In its cytosolic state, Exp-t is open and rather elongated, with its N- and C-termini being distant in space. Ran stabilizes a closed conformation of Exp-t that can recognize tRNA.

Although the CAS/Cse1p export complex (Matsuura and Stewart, 2004) resembles that of Exp-t, Ran-promoted cargo binding by CAS deviates markedly from the just described principle. While cytoplasmic Exp-t is open, unliganded CAS/Cse1p is tightly closed by the C-terminal region clamping onto the N-terminal arch (Cook *et al.*, 2005, **Figure 5-1**). This conformation clearly precludes cargo loading. However, when RanGTP intercalates between the two arches, it opens the transport receptor to a horseshoe-like conformation, which can readily accommodate the export cargo (Matsuura and Stewart, 2004). Thus, both Exp-t and CAS act as Ran-driven "clamps", but the ways of how Ran triggers cargo clamping are diametrically opposite.

How does Ran control cargo loading onto Exp5? Architecturally, the RanGTP·Exp5·pre-miRNA complex is very similar to the Exp-t export complex (Okada *et al.*, 2009). The interaction with Ran appears to bend the exportin into a U-shaped structure that can sandwich the pre-micro RNA (**Figure 5-1**). The atomic details of cytoplasmic, unliganded Exp5 are elusive, but it most likely assumes a more open conformation similar to that of cytosolic Exp-t. Alternatively, Exp5 might follow the paradigm described for CAS.

In their RanGTP-bound states, Exp-t, Exp5 and CAS use their inner HEAT helices to contact their load, i.e. they wrap around their cargoes. Ran facilitates loading of these exportins not only by stabilizing their cargo-competent states, but also by directly binding to the cargoes. As discussed in **Chapters 3 and 4**, this mechanism of cooperative export complex formation is unavailable for the CRM1 ring (**Figure 5-1**), because here the cargo is recruited to the outer HEAT helices – far away from the inner Ran-binding site. In other words, Ran appears to promote CRM1 loading by an exclusively allosteric mechanism. The atomic structure of unliganded CRM1 is elusive, but it is likely that this "relaxed" form assumes a more open conformation that is incompatible with cargo and Ran binding. Data obtained by small angle X-ray scattering and electron microscopy (Fukuhara *et al.*, 2004; Petosa *et al.*, 2004) and the observation that CRM1's ring conformation is stabilized by several strong salt bridges (Dong *et al.*, 2009b; Monecke *et al.*, 2009) suggest that an overall toroid-like shape will be preserved in cytoplasmic CRM1. The recent crystal structure of a RanBP1-stabilized export complex disassembly intermediate (Koyama and Matsuura, 2010, see below) provides proof for a RanGTP-driven allosteric mode of CRM1 cargo loading and unloading. This disassembly mechanism will be discussed in the next section.



**Figure 5-1: Comparison of structurally characterized Imp $\beta$ -like export mediators.**

The indicated transport receptors are shown in a surface representation, gradient-colored from dark blue (N-terminus) to light blue (C-terminus). The left column shows the NTRs in their (Ran-free) cytosolic states, while the right column depicts their nuclear (RanGTP-bound) forms. The orientations are the same with respect to HEATs 2-7. RanGTP is shown in red, the indicated transport cargoes are depicted in orange. The acidic loop of CRM1 is shown in magenta. Red arrows sketch the apparent global conformational changes that occur upon RanGTP binding. At the right, the exportins are classified by the Ran-triggered conformational changes and/or their shape. The PDB-IDs are indicated. See text for details.

## 5.2 Dissociation of export cargoes from CRM1

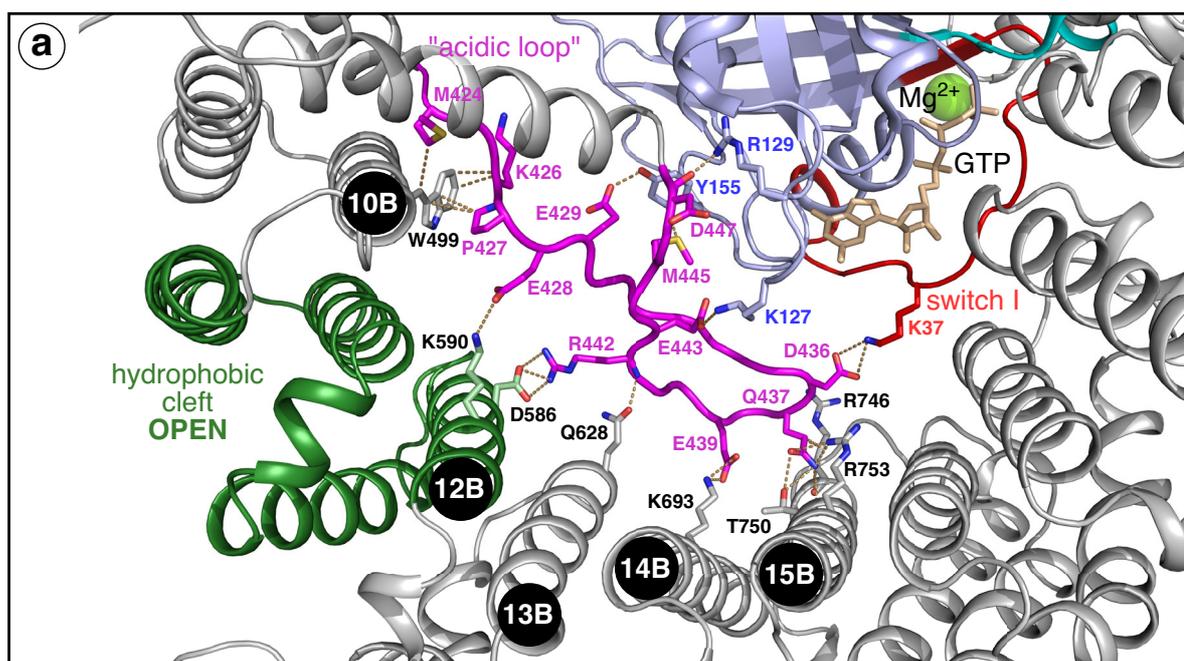
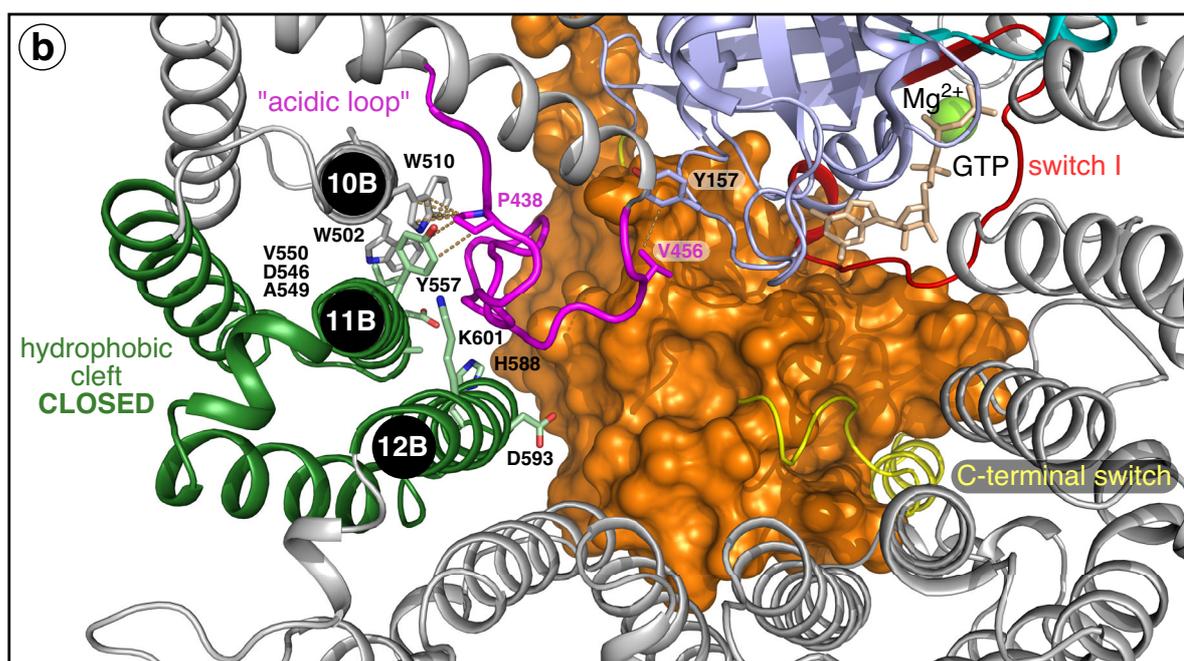
In a CRM1 export complex, Ran-bound GTP is protected from RanGAP-stimulated hydrolysis (Paraskeva *et al.*, 1999), similarly to what had earlier been described for Imp $\beta$  (Floer and Blobel, 1996; Görlich *et al.*, 1996). RanBP1 (Bischoff *et al.*, 1995) or RanBP2/Nup358 (Yokoyama *et al.*, 1995) release this inhibition (Bischoff and Görlich, 1997; Lounsbury and Macara, 1997; Paraskeva *et al.*, 1999; Askjaer *et al.*, 1999; Maurer *et al.*, 2001; Koyama and Matsuura, 2010), priming Ran for GTPase activation by RanGAP. The crystal structure of the RanGTP·CRM1·RanBP1 complex from *S. cerevisiae* (Koyama and Matsuura, 2010) has revealed that RanBP1 can actively displace an NES cargo from CRM1.

In all CRM1 cargo complexes crystallized so far (see **Chapters 3 and 4**), the exportin's "acidic loop" (within HEAT 9) engages in extensive interactions with RanGTP and CRM1 HEAT repeats 10 and 12-15 (**Figure 5-2a**). When an export complex reaches the cytoplasmic face of the NPC, it encounters RanBP1 and RanBP2/Nup358. Their *Ran-binding domains* (*RanBDs*) interact with Ran by a tight "molecular embrace": the C-terminal switch region of Ran (see **Figure 1-4**) wraps around the RanBD, while the N-terminus of the RanBD clasps around Ran (Vetter *et al.*, 1999b). As revealed by the RanGTP·CRM1·RanBP1 complex (**Figure 5-2b**), the very same embrace in the context of an export complex causes the export cargo to be released. How is this accomplished? Both RanBP1 and the C-terminal switch region of Ran would not only clash with the acidic loop of "nuclear" CRM1, but also with its HEAT repeats 14 and 15. As a consequence, CRM1 performs a series of structural rearrangements. Most strikingly, the acidic loop assumes a highly compacted conformation, packing against the "back" of the NES-binding HEAT repeats 11 and 12 and helix 10B by a very elaborate network of mostly hydrophobic but also polar interactions (**Figure 5-2b**).

The incompatibility of the RanBP1·RanGTP component with "nuclear" CRM1 and the binding of RanBP1 to HEAT 15 also trigger some more global remodeling of the exportin (**Figure 5-3a**). The RanBP1-induced rearrangements of the acidic loop and the changes in the superhelical path of the CRM1 HEAT repeats are probably coupled. Most crucially, these changes close the NES-binding site (the "hydrophobic cleft") of CRM1 and thereby cause the exportin to release its cargo (**Figure 5-3b, c**). This mechanism is in perfect agreement with the cooperativity model proposed in **Chapters 3 and 4** of this work.

The described RanBP1-triggered changes in CRM1 also break the exportin's contacts to the switch I loop and the guanine-binding loops of Ran, preparing the complex for RanGAP-stimulated disassembly.

## RanGTP·CRM1·SPN1 complex (SPN1 being omitted)

RanGTP·CRM1·RanBP1 complex (*S. cerevisiae*, Gsp1p·Xpo1p·Yrb1p)

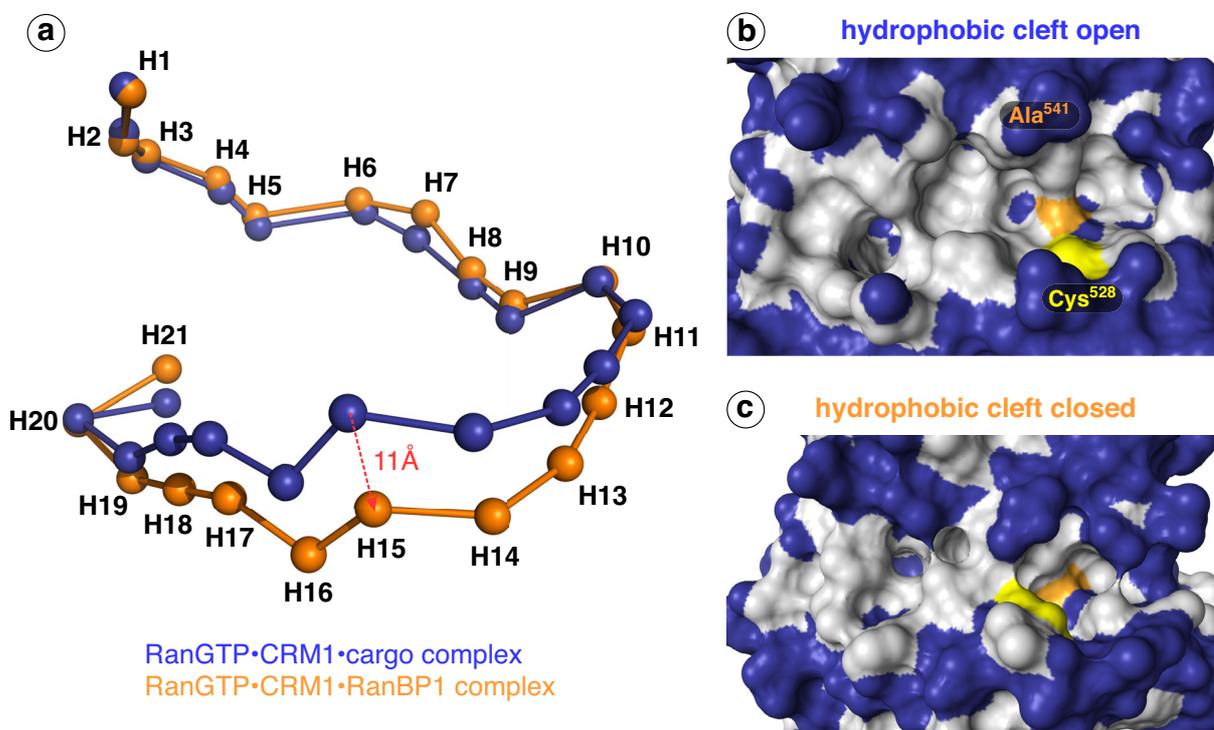
**Figure 5-2: Interactions of the acidic loop with RanGTP and the HEAT repeats of CRM1.**

CRM1 is shown in gray (with hydrophobic cleft HEATs 11 and 12 in green and the acidic loop in magenta), RanGTP is depicted in light blue (with switch I in red, switch II in cyan, the C-terminal switch in yellow, GTP in ochre, and the magnesium ion as a green sphere). Numbers in black circles denote those B helices of CRM1 that contact the acidic loop. Residues involved in interactions are shown as sticks with dashed lines linking interacting atoms (hydrophobic contacts  $\leq 3.8$  Å, polar contacts  $\leq 3.6$  Å). Oxygens are red, nitrogens blue and sulfur atoms yellow.

(a) Interactions of the acidic loop in the RanGTP·CRM1·SPN1 complex (PDB-ID 3GJX, Monecke *et al.*, 2009). Note that the acidic loop extends over HEAT repeats 10-15 and that it also interacts with Ran loops (including switch I). CRM1 is from mouse, RanGTP is human.

(b) (*previous page*) Interactions of the acidic loop in the RanGTP·CRM1·RanBP1 complex (PDB-ID 3M1I, Koyama and Matsuura, 2010). RanBP1 is depicted in a surface representation (orange). Its binding to an export complex triggers major conformational changes in CRM1, mainly in the acidic loop, but also in the arrangement of HEAT repeats. Collectively, these changes close the hydrophobic cleft and thereby release the NES cargo. See text and **Figure 5-3** for details. The "compacted" acidic loop tightly packs against HEAT helices 10B, 11B and 12B. For clarity, only the B helix residues are shown as sticks, interacting acidic loop residues (Met<sup>435</sup>, Val<sup>436</sup>, Arg<sup>437</sup>, Glu<sup>440</sup>, Leu<sup>442</sup>, Val<sup>443</sup>, Ile<sup>451</sup> and Arg<sup>453</sup>, yeast numbering) are not depicted, with the exception of the invariant Xpo1p<sup>Pro438</sup> (Pro<sup>427</sup> in mouse CRM1). Note that practically all interactions of the acidic loop with Ran are lost. (The depicted bond is only weak, distance 3.7 Å).

The view in the panels is the same with respect to RanGTP, whose structure is virtually identical in both complexes.



**Figure 5-3: Global and local RanBP1-triggered conformational changes in CRM1 HEAT repeats.**

(a) Change in superhelical paths of CRM1 HEAT repeats accompanying RanBP1 binding (Figure adapted from Koyama and Matsuura, 2010). The indicated CRM1 complexes were aligned via their RanGTP components as in **Figure 5-2**. The geometric centers of individual HEAT repeats are represented by spheres (blue: RanGTP·CRM1·cargo complex, orange: RanGTP·CRM1·RanBP1 complex). Note that CRM1 in the RanBP1 complex assumes a more "open" conformation than cargo-bound CRM1. The most pronounced deviations are seen for HEAT repeats 12-19 with HEAT 15 (which represents the binding site for RanBP1) re-locating its center most drastically. Similar results are obtained with the CRM1 molecules aligned via HEAT repeats 1-8, whose interactions with RanGTP remain essentially unchanged upon RanBP1 binding (Koyama and Matsuura, 2010). This representation merely shows the relocation of HEAT repeats, but cannot properly reflect rotations that drive hinge-like motions, as they do not shift the repeats' centers.

(b) The picture shows the NES-binding site ("hydrophobic cleft") of Ran- and cargo-bound CRM1 in a surface representation (see also **Figure 4-2a**). The cargo has been omitted. Hydrophilic regions are shown in blue, hydrophobic surface is depicted in gray. For orientation, Cys<sup>528</sup> is colored yellow, while Ala<sup>541</sup> is depicted in orange. Note the deep pockets, which can accommodate the  $\Phi$  residues of NESs (see also **Figure 4-2b**).

(c) RanBP1 binding to an export complex causes the hydrophobic cleft to close. Approximately the same view as in (a) is depicted (PDB-ID 3M1I, Koyama and Matsuura, 2010).

### 5.3 Future perspectives

Deciphering the precise choreography of conformational changes that drive CRM1 cargo loading and unloading is probably one of the most exciting future perspectives of this work. Besides the invariant Pro<sup>438</sup> in the CRM1 acidic loop (yeast numbering, **Figure 5-2b**), 19 further residues are absolutely conserved from human to species as distant as *Trichomonas vaginalis* (see also **Supplementary Figure 4-7**). It is likely that most of these residues play key roles in export complex formation and dissociation. In addition, the C-terminus of CRM1 had been suggested to stabilize the cytosolic form of CRM1 (Dong *et al.*, 2009b; Dong *et al.*, 2009a; Cook and Conti, 2010), however, its role in cooperative export complex formation is somewhat controversial. The crystal structure of unliganded CRM1 and an understanding of how intramolecular interactions in CRM1 can stabilize either the nuclear or the cytoplasmic state of the exportin, will provide insight not only into mechanisms of nuclear export but also into general principles of allosteric regulation and conformational dynamics of proteins.

The above-described RanGTP·CRM1·RanBP1 complex raises the question of how RanBP1 can be both a disassembly factor of CRM1 export complexes and an NES-dependent CRM1 cargo (Richards *et al.*, 1996).

Which features enable CRM1 to recognize hundreds of structurally unrelated cargoes? The structural basis for Snurportin and NES recognition provides part of the answer (see **Chapters 3 and 4**). Our observation that virtually the entire central "hole" of the CRM1 ring is occupied by RanGTP and the acidic loop (see **Figure 5-1**) strongly suggests that all CRM1 cargoes dock onto the outer, convex surface of the exportin. Such binding topology would indeed put the least constraints onto the cargoes' export determinants and could thereby greatly expand the spectrum of cargoes. The outer CRM1 surface is large and at the same time very diverse in charge distribution. Any surface patch could serve as a cargo-binding site, provided that its shape and/or charge complementarity to the cargo is regulated by Ran. Such general CRM1-binding mode would not impose any size limitations onto the cargoes (see also **Chapter 3**) as opposed to the wrapping or sandwiching seen for Exp-t, Exp5 and CAS (**Figure 5-1**). Curiously, the cell indeed employs CRM1 for the export of "bulk goods" such as ribosomal subunits (Ho *et al.*, 2000; Gadal *et al.*, 2001; Moy and Silver, 2002; Thomas and Kutay, 2003). The lack of obligatory cargo-Ran contacts (see above) should further reduce the restraints on CRM1-dependent export signals. In fact, in the case of CRM1, Ran might even add to the exportin's versatility by directly contributing to cargo binding, though no such example has been encountered so far. Whether all CRM1 cargoes rely on binding to the

hydrophobic cleft is still unknown. Further assessment of the allowed variability of functional NES peptides, in particular the elucidation of the still ill-defined Rev-type NES consensus will help to clarify this fundamental question.

Imp13 and Msn5p have been known as bidirectional transport receptors. The characterization of Exp4 as an import mediator (**Chapter 2**) adds yet a third member to this unusual class of Imp $\beta$ -likes. The question of how these factors can act both in nuclear import and export is largely unresolved, and the structural description of their nuclear and cytosolic cargo complexes is a major future goal. Crystal structures of Imp13 bound to import cargo or Ran have been determined recently (Bono *et al.*, 2010, **Figure 5-1**). Imp13 tightly encircles its import cargo (the Mago-Y14 complex). Although the Mago-Y14-binding sites on Imp13 do not overlap with the Ran-binding sites, recruitment of the ligands is mutually exclusive, because they would clash when binding simultaneously. This nicely explains as to how Imp13 can perform unidirectional cargo import. The modes by which Imp13 and CRM1 interact with Ran are very similar and thus, their nuclear forms are virtually indistinguishable by their overall shapes (**Figure 5-1**). How Imp13 binds export cargoes remains to be determined. Bono *et al.*, 2010 suggested that the import cargo-binding site of Imp13 might also recruit proteins destined for export, provided that they do not interfere with RanGTP binding. Such binding would probably involve direct contacts between Ran and the export cargo. However, given the architectural similarity of CRM1 and Imp13, it is tempting to speculate that also in the case of Imp13, Ran can allosterically promote export cargo binding to the outer face of the ring. Interestingly, nuclear Imp13 assumes a more closed conformation as compared to the cytosolic form. "Compaction" of the ring indeed coincides with considerable shape changes in the outer Imp13 surface (**Figure 5-1**). Taken together, future studies on Imp13, Exp4 and CRM1 will certainly reveal exciting surprises.

## CHAPTER 6

**Appendix**

*This chapter serves as an appendix to **Chapter 4**. The first part contains a description of how I optimized a synthetic *E. coli* culture medium for the cost-efficient large-scale perdeuteration of CRM1. This section also includes details on the isotope labeling of the PKI NES peptide. In the second part, I append a reprint of our structure-function study on the Abl C-terminal domain:*

**Structural Basis for the Cytoskeletal Association of Bcr-Abl/c-Abl**

Oliver Hantschel, Silke Wiesner, Thomas Güttler, Cameron D. Mackereth, Lily L. Remsing Rix, Zsuzsanna Mikes, Jana Dehne, Dirk Görlich, Michael Sattler and Giulio Superti-Furga

**Molecular Cell**, 2005, Volume 19, Pages 461-73.

*Additional data are presented at the end of this chapter. See **Chapter 4** for further details.*

## 6.1 Establishment of a cost-efficient *E. coli* culture medium for the perdeuteration of CRM1

Virtually complete deuteration of proteins (> 95% D, also referred to as "perdeuteration") requires growth of bacteria in D<sub>2</sub>O-based defined or rich media that are practically devoid of <sup>1</sup>H and provide fully deuterated carbon and nitrogen sources. The most commonly employed defined bacterial culture media are derived from the classical M9 minimal medium (Sambrook *et al.*, 1989). This medium had been originally conceived to control stationary phase cell densities for synchronization of bacterial cells (Nagata, 1963), and was later refined as a standard medium for physiological studies on enterobacteria (Neidhardt *et al.*, 1974).

Perdeutering proteins is a challenging task because biological systems respond sensitively to high concentrations of D<sub>2</sub>O. H<sub>2</sub>O and D<sub>2</sub>O differ in their solvent properties since hydrogen bonds involving deuterium instead of hydrogen are generally stronger (see Katz, 1960). Proteins dissolved in D<sub>2</sub>O therefore tend to be more "rigid" than the equivalent hydrogenated forms, which often reduces the rates of enzymatic reactions (see Katz, 1960; Flaumenhaft *et al.*, 1965). This phenomenon is commonly referred to as the "kinetic isotope effect". It is not possible to reliably predict the effects of perdeuteration on the stability or "rigidity" of a given protein. In any case, a change from H<sub>2</sub>O to D<sub>2</sub>O imposes a major challenge on the cell's metabolism. While a high degree of deuteration is toxic to animals and plants (mice and rats die when their body water approaches ≈ 30% deuteration, see Katz, 1960), bacteria and unicellular green algae can be propagated in fully deuterated media, allowing highly efficient <sup>2</sup>H-labeling of biomolecules (Crespi *et al.*, 1959). Bacterial growth in D<sub>2</sub>O is drastically reduced (by at least a factor of five) and normally requires prior adaptation of the cells to high levels of D<sub>2</sub>O – either via serial sub-cultures with increasing D<sub>2</sub>O concentrations or by direct accommodation over extended periods of time.

With the classical M9 medium, however, this adaptation is often further complicated because the medium does not support optimal bacterial growth (even in an H<sub>2</sub>O-based setting). Bacterial cultures grown on classical M9 medium not only fail to reach high population densities but often also display decreased levels of protein overexpression. These problems have been partially overcome with the commercial advent of rich culture media - hydrolysates from chemolithoautotrophic bacteria or autotrophic algae grown on inorganic, isotopically labeled substrates (Crespi *et al.*, 1959; <http://www.silantes.com/>). While these media indeed allow for more robust protein expression, the cultures do not reach the saturation densities achieved with yeast extract-based media such as 2YT or TB. Moreover, they are often costly,

especially when needed for large-scale expression of high-molecular weight proteins for NMR studies.

The expression and purification procedure for CRM1 (described in **Chapter 4**) yields  $\approx 16$  mg of pure protein from 1 liter of a culture grown in a fermenter that allowed for a saturation density of 14 OD<sub>600</sub> in 2YT medium. With the unjustified optimistic assumption that expression level and protein solubility remain unchanged in a deuteration culture grown in a standard medium (see below, **Table 6-1**, maximum saturation density of  $\approx 3$  OD<sub>600</sub>), 18 liters of medium would be required to obtain 60 mg of perdeuterated CRM1 (which is the protein amount that was needed for our NMR study). The costs for the culture medium would amount to at least 24000 € (with D<sub>12</sub>-glucose as a carbon source) or 8000 € if deuterated glycerol is used instead (see **Table 6-2b** for the underlying pricing information). Obviously, D<sub>2</sub>O becomes a significant cost factor with these medium requirements. In the following, I will describe the establishment of a cost-efficient *E. coli* culture medium for the efficient and reliable deuteration of CRM1. The main focus will be on the reasonable exploitation of D<sub>2</sub>O and the deuterated carbon and nitrogen sources.

### 6.1.1 Formulation of a high-density *E. coli* culture medium

Starting point for the optimization of the culture conditions was a defined medium adapted from the classical M9 recipe (Sambrook *et al.*, 1989). Its composition is described in **Table 6-1**. All optimization steps had been performed with <sup>1</sup>H-based media unless stated otherwise.

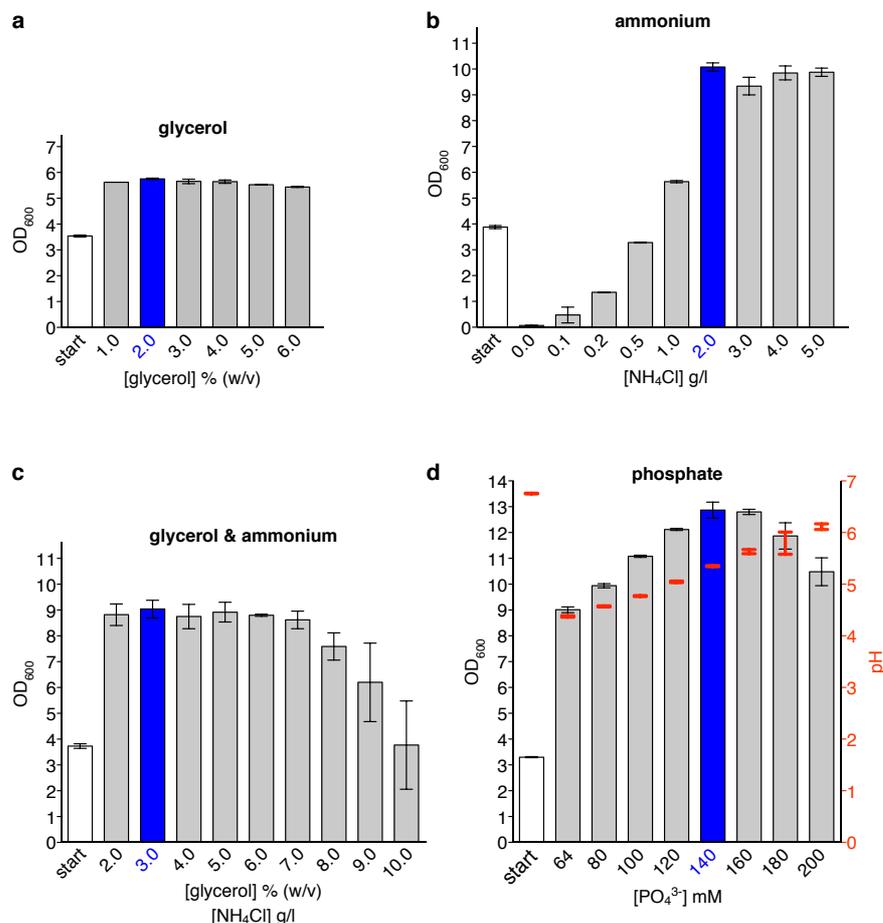
**Table 6-1: Composition of the *E. coli* culture medium that served as the starting point for the formulation of an optimized deuteration medium.**

(\*) Instead of glucose, glycerol was chosen as a carbon source (2 g/l, 108.6 mM) since D<sub>8</sub>-glycerol is ≈ 10-fold less expensive than D<sub>12</sub>-glucose. See text for details.

HD-M9 salts		
compound	concentration (g/l)	concentration (mM)
Na <sub>2</sub> HPO <sub>4</sub>	13.06	92.1
KH <sub>2</sub> PO <sub>4</sub>	6.53	47.9
NaCl	0.50	8.6
MgSO <sub>4</sub>	0.12	1.0
CaCl <sub>2</sub>	0.03	0.3
carbon and nitrogen source		
compound	concentration (g/l)	concentration (mM)
NH <sub>4</sub> Cl	2.00	37.4
Glucose*	2.00	11.1
Trace elements		
149 μM EDTA, 30 μM FeCl <sub>3</sub> , 6 μM ZnCl <sub>2</sub> , 0.8 μM CuSO <sub>4</sub> , 1.6 μM H <sub>3</sub> BO <sub>3</sub> , 0.1 μM MnCl <sub>2</sub> , 0.4 μM CoCl <sub>2</sub>		
Other cofactors		
100 μg/ml Biotin. 100 μg/ml Thiamine		

The concentrations of carbon and nitrogen source are the most obvious determinants of the maximally possible cell density of a bacterial culture. **Figure 6-1a** shows that the final population density increased markedly when the glycerol concentration was doubled. Addition of extra ammonium chloride led to a saturation density more than twice as high than that achieved with the initial medium (**Figure 6-1b**). Co-titration of glycerol and ammonium chloride confirmed that neither the carbon source nor the nitrogen source were limiting under the culture conditions used (**Figure 6-1c**). Analysis of the culture supernatant of a stationary culture grown at 3% (w/v) glycerol and 3 g/l ammonium, however, revealed that the pH had dropped below 5 whereas the final pH under the starting conditions usually ranges between 6.5 and 7.0 (see **Figure 6-1d**). This observation can be readily explained by the phosphate levels being insufficient to buffer the metabolic products of the cells. Indeed, increasing the

phosphate concentration stabilized the pH and also raised the saturation density of the culture (**Figure 6-1d**).



**Figure 6-1: Optimization of a defined *E. coli* culture medium for high-density growth.**

*E. coli* BL21 (DE3) was grown in an M9-derived medium (see **Table 6-1**) to an OD<sub>600</sub> < 2.0. This starter culture was used to inoculate 5 ml of the indicated media (in 100 ml-Erlenmeyer flasks) to an OD<sub>600</sub> of 0.02. These cultures were then grown at 37 °C for 15 hours. The final OD<sub>600</sub> was recorded and plotted. "Start" refers to the medium described in **Table 6-1**. Error bars show the minimum and maximum OD<sub>600</sub> readings from two independent experiments. The blue bars indicate the respective condition chosen for the subsequent optimization steps.

(a) Titration of the glycerol concentration in the medium shown in **Table 6-1**. 2.0% (w/v) glycerol were selected for the subsequent optimization step.

(b) The medium described in **Table 6-1** (but with 2.0% (w/v) glycerol) was prepared without ammonium chloride. The NH<sub>4</sub>Cl concentration was varied as indicated. 2.0 g/l were sufficient to confer the maximum saturation density to the culture.

(c) To test whether glycerol or ammonium chloride are limiting under the culture conditions used, their concentration was co-titrated over the indicated range. An increase of the glycerol and ammonium chloride concentrations beyond the levels established in (a) and (b) did not lead to higher saturation densities and very high concentrations even inhibited growth, probably because of ammonium toxicity. The indicated concentrations (blue) were chosen for the experiment shown in (d).

(d) An increased phosphate concentration (molar ratio of phosphate compounds as specified in **Table 6-1**) not only allowed for clearly higher saturation densities but also stabilized the pH of the cultures (red, the minimum and maximum readings are shown).

The MgSO<sub>4</sub> concentration was not adjusted in these optimization experiments. However, Studier, 2005 reported that a change from 1 to 2 mM MgSO<sub>4</sub> can increase the saturation density of a culture.

Trace metal addition is required for optimal growth of *E. coli* in defined media. When the medium is prepared with drinking water, the trace metal amounts present are probably sufficient to support growth to moderate or even high cell densities. However, as the D<sub>2</sub>O supply is virtually free of trace elements, we used double-distilled water for the optimization of the culture medium. The composition of our trace element solution (see **Table 6-3**) is similar to that described in Neidhardt *et al.*, 1974, but there are four notable differences. (1) Some additional trace elements had been included (Ni, Se and W). (2) The relative concentrations of the trace elements were roughly adjusted to their abundance in *E. coli* (Rouf, 1964). (3) The Fe<sup>2+</sup> concentration in our culture medium is 5 times higher. Studier, 2005 reported iron to be decisive for robust bacterial growth and for protein overexpression. Including 100 µM FeCl<sub>3</sub> was shown to partially compensate for the complete omission of a trace element solution. It should be noted that Fe<sup>2+</sup> and Fe<sup>3+</sup> performed equally well in our experiments (not shown). (4) Instead of EDTA (**Table 6-1**) we added NTA as a chelator to buffer the free metal ions. The final formulation of our trace element solution (see **Table 6-3**) is similar to what had been used for other synthetic media (e.g., Studier, 2005).

### 6.1.2 Cost-efficiency

Although a higher glycerol and ammonium concentration raises the saturation density of a bacterial culture (**Figure 6-1a-c**), the increase is not necessarily cost-efficient since the perdeuterated compounds are expensive. To take cost-efficiency into consideration, I adjusted the concentrations of carbon and nitrogen source based on the experiment shown in **Table 6-2**.

---

**Table 6-2:** (next page) **Adaptation of carbon and nitrogen source to cost-efficient concentrations.**

To determine the most cost-efficient concentrations of carbon and nitrogen sources, *E. coli* BL21 (DE3) was grown in media containing the indicated concentrations of glycerol and ammonium chloride. The experiment was performed as described in **Figure 6-1** with the notable difference that here, the cells had been transformed with the CRM1 expression plasmid (see main text for rationale). 1.0% (w/v) glycerol and 2.0 g/l NH<sub>4</sub>Cl (blue) were chosen for the subsequent experiments.

(a) Saturation densities recorded.

(b) The costs per 1000 OD<sub>600</sub> equivalents were calculated based on the saturation densities reached (see panel a) and the following pricing information (June 2007): D<sub>2</sub>O 327.00 €/l (CU Chemie Uetikon AG), ND<sub>4</sub>Cl 16.90 €/g (Aldrich), D<sub>8</sub>-glycerol 50.00 €/g (Spectra Stable Isotopes). For comparison, the costs for D<sub>12</sub>-glucose would amount to 500.00 €/g (Spectra Stable Isotopes).

**a** final OD<sub>600</sub>

		[glycerol] % (w/v)							
		0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
[NH <sub>4</sub> C] g/l	1.00	2.14	4.36	4.56	4.86	5.18	5.54	6.16	5.46
	1.25	2.18	4.68	6.36	6.58	7.04	7.32	7.28	7.54
	1.50	2.22	4.72	6.30	7.22	7.14	7.96	7.96	8.32
	1.75	2.24	4.36	6.06	8.14	8.70	8.46	10.04	8.60
	2.00	2.36	4.12	5.54	7.6	8.82	8.7	9.66	10.08

**b** costs / 1000 OD<sub>600</sub> (€)

		[glycerol] % (w/v)							
		0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
[NH <sub>4</sub> C] g/l	1.00	207.43	124.75	141.21	153.07	162.92	170.38	169.46	209.51
	1.25	205.56	117.12	101.91	113.70	120.47	129.53	143.97	152.27
	1.50	203.76	117.02	103.55	104.20	119.38	119.64	132.20	138.50
	1.75	203.83	127.65	108.35	92.95	98.46	113.07	105.24	134.49
	2.00	195.25	136.12	119.28	100.11	97.60	110.44	109.81	115.16

Even though the final medium is about twice as expensive as the initial setup (not shown), the cost-efficiency is clearly better, because the adjustment raised the saturation density by a factor of more than four (**Table 6-2**). The adaptation of the glycerol and ammonium concentrations as shown in **Table 6-2** was not just pivotal for cost-efficiency but also for growth of *E. coli* transformed with the CRM1 expression plasmid (with Amp<sup>R</sup> as a selection marker). These cells grew only poorly on medium with carbon and nitrogen concentrations higher than those highlighted in **Table 6-2**, possibly because of a greater sensitivity of plasmid-transformed cells to high ammonium concentrations. The composition of the final "high-density (HD)-M9" medium is described in **Table 6-3**. Cultures grown in HD-M9 reached cell densities equivalent to those achieved with rich media such as 2YT (**Figure 6-2**). Spiking the medium with 2% (v/v) Silantes *E. coli*-OD2 medium (a hydrolysate of chemolithoautotrophic bacteria) increased the robustness of bacterial growth (especially in the initial growth phase) but did not affect the final culture density. The supply of the medium with growth-promoting cofactors is a likely explanation of this effect.

**Table 6-3: Composition of the optimized M9 medium ("high density (HD)-M9").**

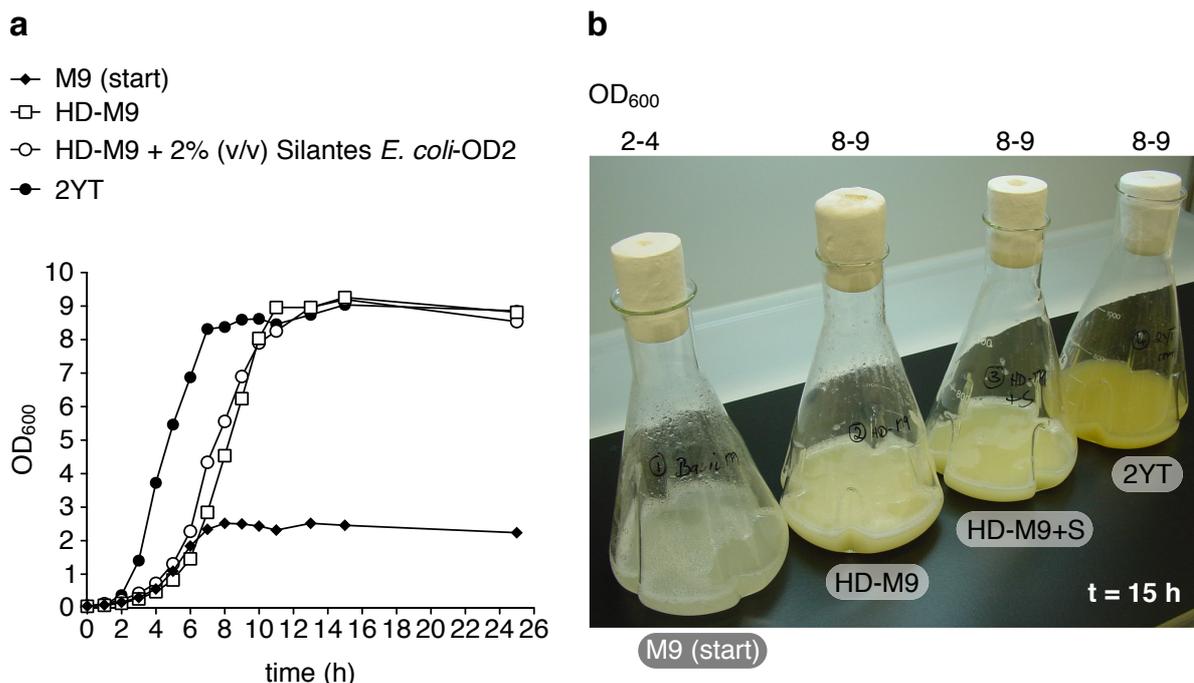
The recipe for the  $^1\text{H}$ -based medium is shown in (a), the details for the equivalent deuteration medium are given in (b). Note that for the  $^2\text{H}$ -based medium,  $^1\text{H}$ -free chemicals were used (with the exception of the cofactors). Supplementation of the HD-M9 medium with 2% (v/v) Silantes *E. coli*-OD2 medium shortened the lag phase and increased the initial growth rate to some extent (see **Figure 6-2**).

**a HD-M9 medium ( $\text{H}_2\text{O}$ )**

<b>HD-M9 salts</b>		
compound	concentration (g/l)	concentration (mM)
$\text{Na}_2\text{HPO}_4$	13.06	92.1
$\text{KH}_2\text{PO}_4$	6.53	47.9
$\text{NaCl}$	0.50	8.6
$\text{MgSO}_4$	0.12	1.0
$\text{CaCl}_2$	0.03	0.3
<b>carbon and nitrogen source</b>		
compound	concentration (g/l)	concentration (mM)
$\text{NH}_4\text{Cl}$	2.00	37.4
Glycerol	10.00	108.6
<b>Trace elements</b>		
150 $\mu\text{M}$ NTA, 50 $\mu\text{M}$ $\text{FeSO}_4$ , 10 $\mu\text{M}$ $\text{ZnCl}_2$ , 5 $\mu\text{M}$ $\text{CuSO}_4$ , 5 $\mu\text{M}$ $\text{H}_3\text{BO}_3$ , 5 $\mu\text{M}$ $\text{MnCl}_2$ , 1 $\mu\text{M}$ $\text{NiCl}_2$ , 1 $\mu\text{M}$ $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}$ , 1 $\mu\text{M}$ $\text{Na}_2\text{SeO}_3$ , 1 $\mu\text{M}$ $\text{Na}_2\text{WO}_4$ , 0.5 $\mu\text{M}$ $\text{CoSO}_4$		
<b>Other cofactors</b>		
100 $\mu\text{g/ml}$ Biotin. 100 $\mu\text{g/ml}$ Thiamine		

**b HD-M9 medium ( $\text{D}_2\text{O}$ )**

<b>HD-M9 salts</b>		
compound	concentration (g/l)	concentration (mM)
$\text{Na}_3\text{PO}_4$	10.07	61.4
$\text{K}_3\text{PO}_4$	3.39	16.0
$\text{D}_3\text{PO}_4$	-	62.6
$\text{NaCl}$	0.50	8.6
$\text{MgSO}_4$	0.12	1.0
$\text{CaCl}_2$	0.03	0.3
<b>carbon and nitrogen source</b>		
compound	concentration (g/l)	concentration (mM)
$\text{ND}_4\text{Cl}$	2.15	37.4
$\text{D}_8$ -Glycerol	11.09	108.6
<b>Trace elements</b>		
150 $\mu\text{M}$ NTA, 50 $\mu\text{M}$ $\text{FeSO}_4$ , 10 $\mu\text{M}$ $\text{ZnCl}_2$ , 5 $\mu\text{M}$ $\text{CuSO}_4$ , 5 $\mu\text{M}$ $\text{H}_3\text{BO}_3$ , 5 $\mu\text{M}$ $\text{MnCl}_2$ , 1 $\mu\text{M}$ $\text{NiCl}_2$ , 1 $\mu\text{M}$ $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}$ , 1 $\mu\text{M}$ $\text{Na}_2\text{SeO}_3$ , 1 $\mu\text{M}$ $\text{Na}_2\text{WO}_4$ , 0.5 $\mu\text{M}$ $\text{CoSO}_4$		
<b>Other cofactors</b>		
100 $\mu\text{g/ml}$ Biotin. 100 $\mu\text{g/ml}$ Thiamine		



**Figure 6-2: Growth of *E. coli* BL21 (DE3) in the specified media at 37 °C.**

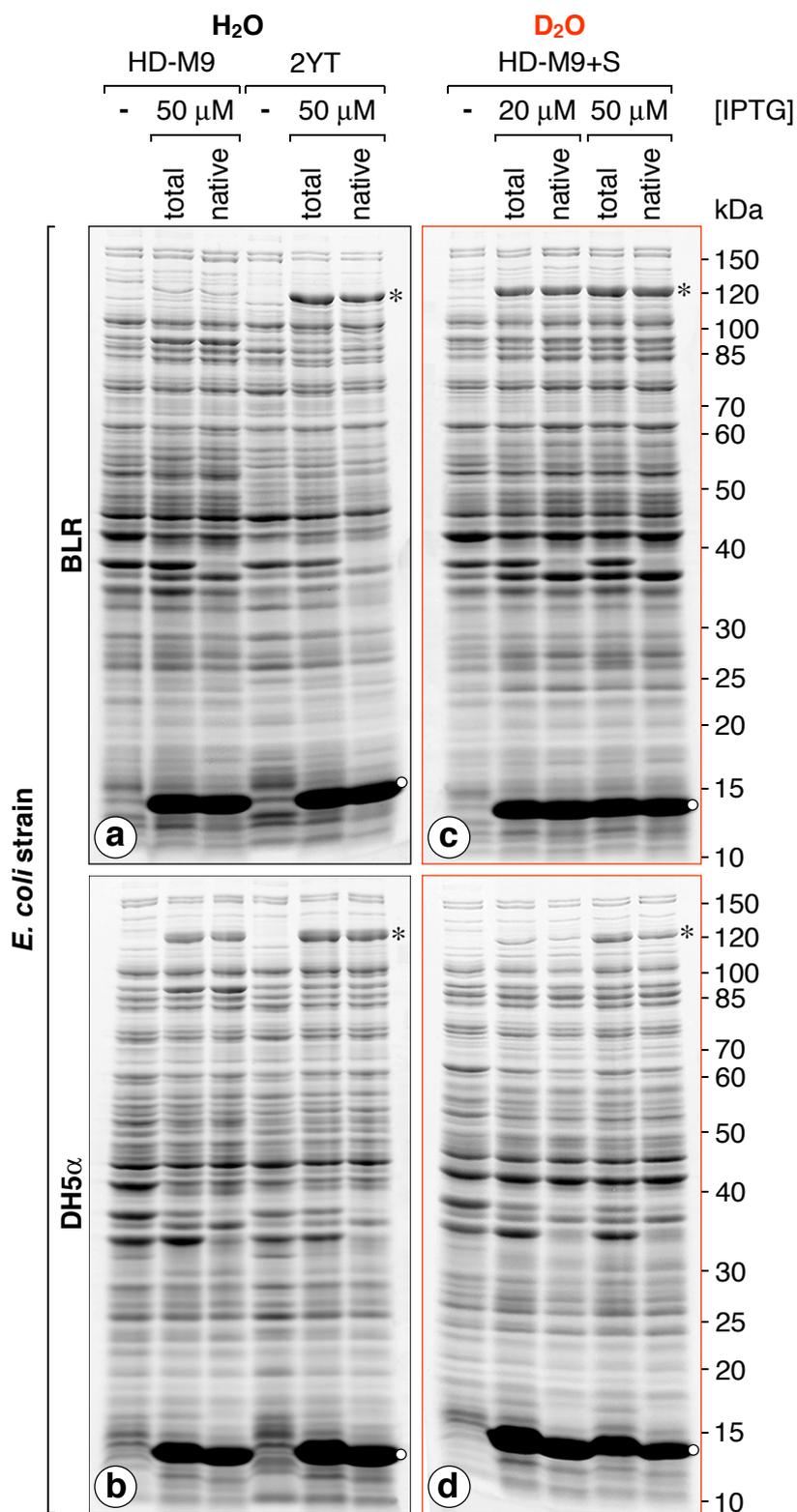
(a) Bacterial growth curves. Note that the HD-M9-based media (**Table 6-3a**) allow for saturation densities of 8-9 OD<sub>600</sub>, which is common for rich media such as 2YT and 2-4 times higher than what can be reached with "traditional" M9 media (**Table 6-1**). "HD-M9+S" refers to HD-M9 medium spiked with 2% (v/v) Silantes *E. coli*-OD2 medium. See text for details.

(b) Photograph shows the bacterial cultures from (a) at t=15h. The maximally achievable saturation densities (OD<sub>600</sub>) are indicated on top of the image.

### 6.1.3 Choice of the *E. coli* strain for deuteration

Large amounts of soluble CRM1 can be obtained by expressing the protein in *E. coli* BLR grown in 2YT medium at 16 °C (Monecke *et al.*, 2009; **Chapter 4**). I therefore chose these conditions for a first CRM1 expression test with HD-M9 medium (**Table 6-3a**). In contrast to the routinely obtained results with 2YT broth, the expression levels were drastically reduced in the defined medium (**Figure 6-3a**). To determine whether this effect was simply related to the choice of the expression strain, I tested an *E. coli* K derivative (DH5 $\alpha$ ) and found that it indeed supported good expression of soluble CRM1 in both 2YT and HD-M9 medium (**Figure 6-3b**). This suggested that DH5 $\alpha$  would also be the best choice for the perdeuteration of CRM1. Unexpectedly, however, the expression test with the D<sub>2</sub>O-based HD-M9 medium (**Table 6-3b**) revealed that the expression patterns under deuteration conditions opposed those observed under H<sub>2</sub>O: while DH5 $\alpha$  showed clear CRM1 expression (with a solubility of just  $\approx$ 50%) only at an elevated IPTG concentration, expression levels and solubility turned out to be robustly high with *E. coli* BLR (compare **Figure 6-3d and c**). There is no obvious

explanation for this notable observation. Taken together, this shows that the expression strain for perdeuteration should be carefully chosen based on pre-trials. For the deuteration of the exportin, described in the following section, I chose *E. coli* BLR.



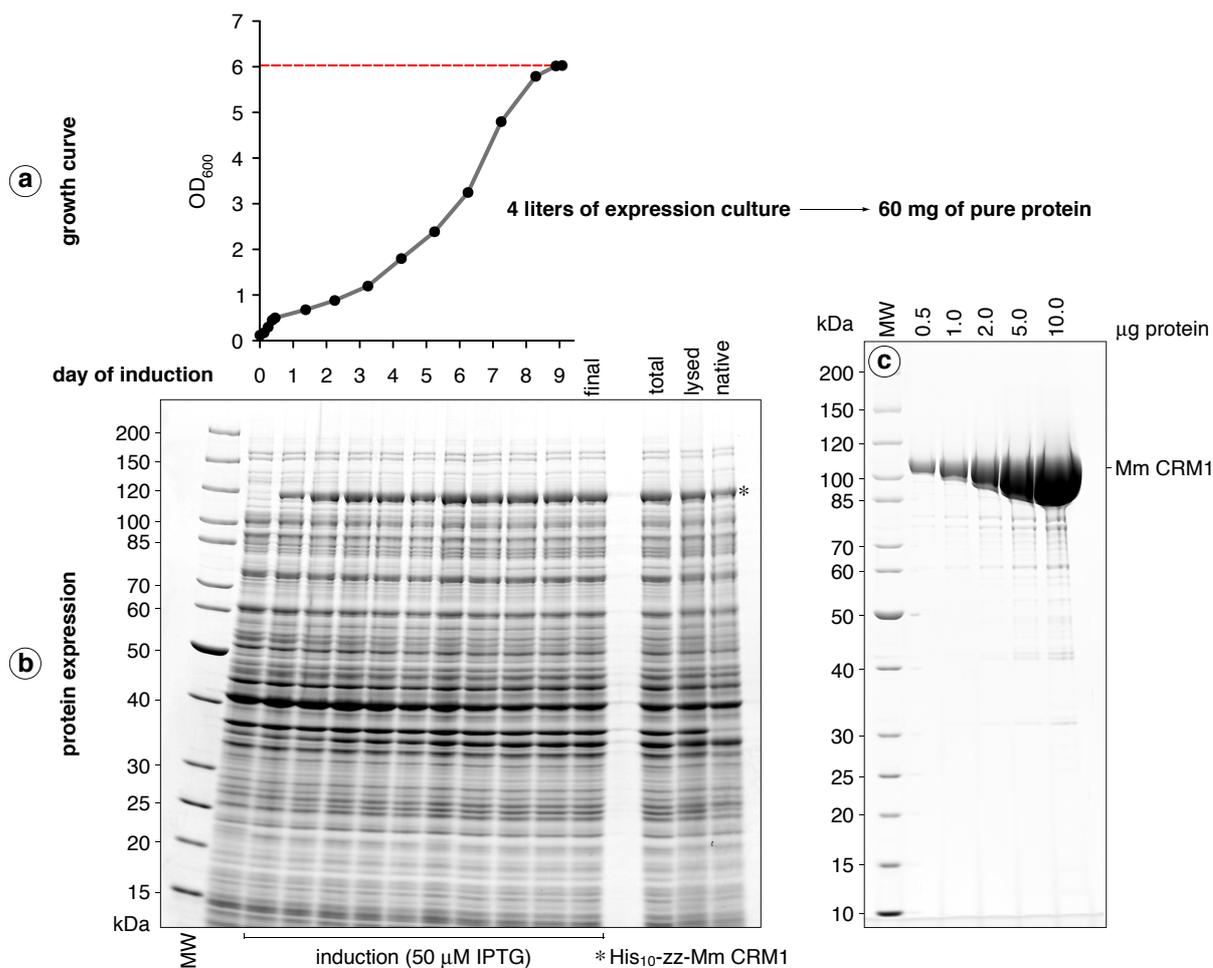
**Figure 6-3: (previous page) Choice of the *E. coli* strain for perdeuteration of CRM1.**

The performance of a given *E. coli* strain in D<sub>2</sub>O cannot be extrapolated from the outcome of experiments conducted with H<sub>2</sub>O-based media. The panels show the result of a CRM1 expression and solubility test with the *E. coli* strains BLR and DH5 $\alpha$  grown in the specified media (see **Figure 6-2**). Cultures were inoculated to an OD<sub>600</sub> of 0.08 and grown to a density of  $\approx 0.4$  at 37 °C before cooling to 16 °C. At an OD<sub>600</sub> of  $\approx 0.5$ , expression was induced with the indicated concentration of IPTG. Prior to induction, the D<sub>2</sub>O cultures were split in two to test two different IPTG levels. The cells were harvested when they reached the stationary growth phase and resuspended to equal densities in 50 mM Tris/HCl pH 7.5, 500 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 5 mM DTT and lysed by sonication in the presence of lysozyme (O). The lysates were then cleared by ultracentrifugation and the samples analyzed by SDS-PAGE followed by Coomassie staining. His<sub>10</sub>-zz-tagged CRM1 is marked with "\*". See text for further explanation. *E. coli* BLR grew to the high HD-M9 culture densities reported in **Figure 6-2** (both for H<sub>2</sub>O and D<sub>2</sub>O), while DH5 $\alpha$  only reached optical densities of  $\approx 6$  (H<sub>2</sub>O) or  $\approx 2$  (D<sub>2</sub>O). This suggests that the optima for the culture media are strain-specific.

---

## 6.2 Perdeuteration of CRM1

For deuteration of CRM1, *E. coli* BLR was transformed with the CRM1 expression plasmid and grown at 37 °C in 2YT medium for about 6 hours. 1 ml of this culture (OD<sub>600</sub>  $\approx 0.8$ ) was then centrifuged and the pellet resuspended directly in 20 ml of Silantes *E. coli*-OD2 D medium (> 98% D). After over-night incubation at 37 °C, this starter culture reached the expected OD<sub>600</sub> of  $\approx 2$ . Four liters of HD-M9 medium (**Table 6-3b**, pre-warmed to 37 °C, supplemented with Silantes *E. coli*-OD2 D medium to 2% (v/v), one liter in each of four baffled five-liter Erlenmeyer flasks) were inoculated to a start-OD<sub>600</sub> of  $\approx 0.1$ . A smaller inoculum is not recommended because it would extend the lag phase (not shown). The culture grew to an OD<sub>600</sub> of  $\approx 0.5$  within 10 hours, showing that the procedure followed can be used to bypass any time-consuming adaptation to D<sub>2</sub>O. The cultures were then cooled to about 20 °C on ice and the cultivation was continued at 16 °C. Two hours after the temperature shift, CRM1 expression was induced with 50  $\mu$ M IPTG. The expression plasmid was selected for with 100  $\mu$ g/ml ampicillin. The Amp<sup>R</sup> gene (encoding  $\beta$ -lactamase) is not a preferred resistance marker, especially in slowly growing cultures for perdeuteration. Since  $\beta$ -lactamase is secreted into the culture medium, the antibiotic is rapidly inactivated at higher cell densities (with a half-life of less than an hour). This increases the danger of reduced expression levels since even the slightest growth advantage of cells lacking the plasmid would give rise to loss of the plasmid from the culture. Attempts to express CRM1 from a construct conferring kanamycin resistance failed because the protein appeared to be toxic under these conditions. To minimize the risk of plasmid loss in the course of expression, the cultures were supplemented with 100  $\mu$ g/ml ampicillin once or twice a day.



**Figure 6-4: Perdeuteration of CRM1.**

Four liters of <sup>2</sup>H-based HD-M9 medium (Table 6-3b) supplemented with 2% (v/v) Silantes *E. coli*-OD2 D medium yielded 60 mg of pure CRM1 (15 mg/liter culture). For comparison, the yield from 1 liter of a 2YT-based fermenter culture is just slightly higher (16 mg).

(a) Growth of the *E. coli* BLR CRM1 expression culture. Although the culture grew slowly, growth was logarithmic up to day 7.

(b) At indicated time points, samples were removed from the expression culture and analyzed by SDS-PAGE followed by Coomassie staining. Note that CRM1 was not degraded over the course of its expression and that most of the protein was soluble in 50 mM Tris/HCl pH 7.5, 500 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 1 mM EDTA, 5 mM DTT (right, see also Figure 6-3).

(c) The panel shows the purity of the final preparation. Note that there is no significant degradation of the protein.

The culture followed a typical bacterial growth curve, indicative of ordered cell division (Figure 6-4a). The observed slower growth is expected for a D<sub>2</sub>O-based defined culture at low temperature. The cell density increased logarithmically up to day 7. At day 9, the culture reached the stationary phase. Figure 6-4b displays the expression of CRM1 during the course of induction. There is no indication for degradation of CRM1 over time (see also Figure 6-3c and Figure 6-4c). The expression was stopped as soon as the cells reached the stationary

phase. CRM1 was purified as described in **Chapter 4**. Impressively, the relative protein yield of the deuteration cultures was nearly as high as that of a fermenter-based CRM1 culture (see **Figure 6-4a**). Mass-spectrometrical analysis confirmed  $^2\text{H}$  incorporation into CRM1 to close to 100% (not shown). I did not notice any solubility differences between perdeuterated and unlabeled CRM1 during the purification.

### 6.3 Conclusions

This work presents the establishment of a highly cost-efficient and robust procedure for the perdeuteration of CRM1. The total costs for 60 mg of very pure protein (**Figure 6-4c**) amount to  $\approx 3500$  € (instead of at least 8000 or 24000 € needed with conventional deuteration protocols, see above). The yield of the CRM1 deuteration culture was unexpectedly high, obviously because expression levels under perdeuteration conditions had even been superior over those achieved with 2YT medium (**Figure 6-3** and data not shown). It should be mentioned that the observed expression levels could be reproduced in two independent experiments. However, cost-efficiency is not the only advantage of our optimized setup. (1) No time-consuming adaptation to  $\text{D}_2\text{O}$  was required even though the growth conditions (16 °C for induction) had been far from optimal for *E. coli*. And (2) the use of anhydrous and  $^1\text{H}$ -free chemicals completely eliminated the need for the laborious (multi-round) recrystallization and drying of hydrogenated chemicals (Meilleur *et al.*, 2009). The procedure should be readily applicable to the perdeuteration of proteins other than CRM1. Below, I describe the use of the HD-M9 medium for a complex labeling scheme of the PKI  $\Phi^0\text{Leu NES}$  peptide. The protocol will prove valuable not only for NMR spectroscopy but any application relying on highly deuterated proteins (such as neutron scattering and neutron diffraction) or robust growth in defined media (e.g. metabolic labeling with selenomethionine).

## 6.4 Isotope labeling of the PKI $\Phi^0$ Leu NES peptide

In this section I will briefly summarize the expression of PKI  $\Phi^0$ Leu NES peptides of various isotope labeling schemes. The expression was performed with a plasmid that confers kanamycin resistance. For  $^{15}\text{N}$  labeling, an Amp<sup>R</sup> plasmid was used, but the culture was repeatedly replenished with ampicillin to minimize the risk of plasmid loss (see above).

### 6.4.1 [ $^{15}\text{N}$ ]-PKI $\Phi^0$ Leu NES

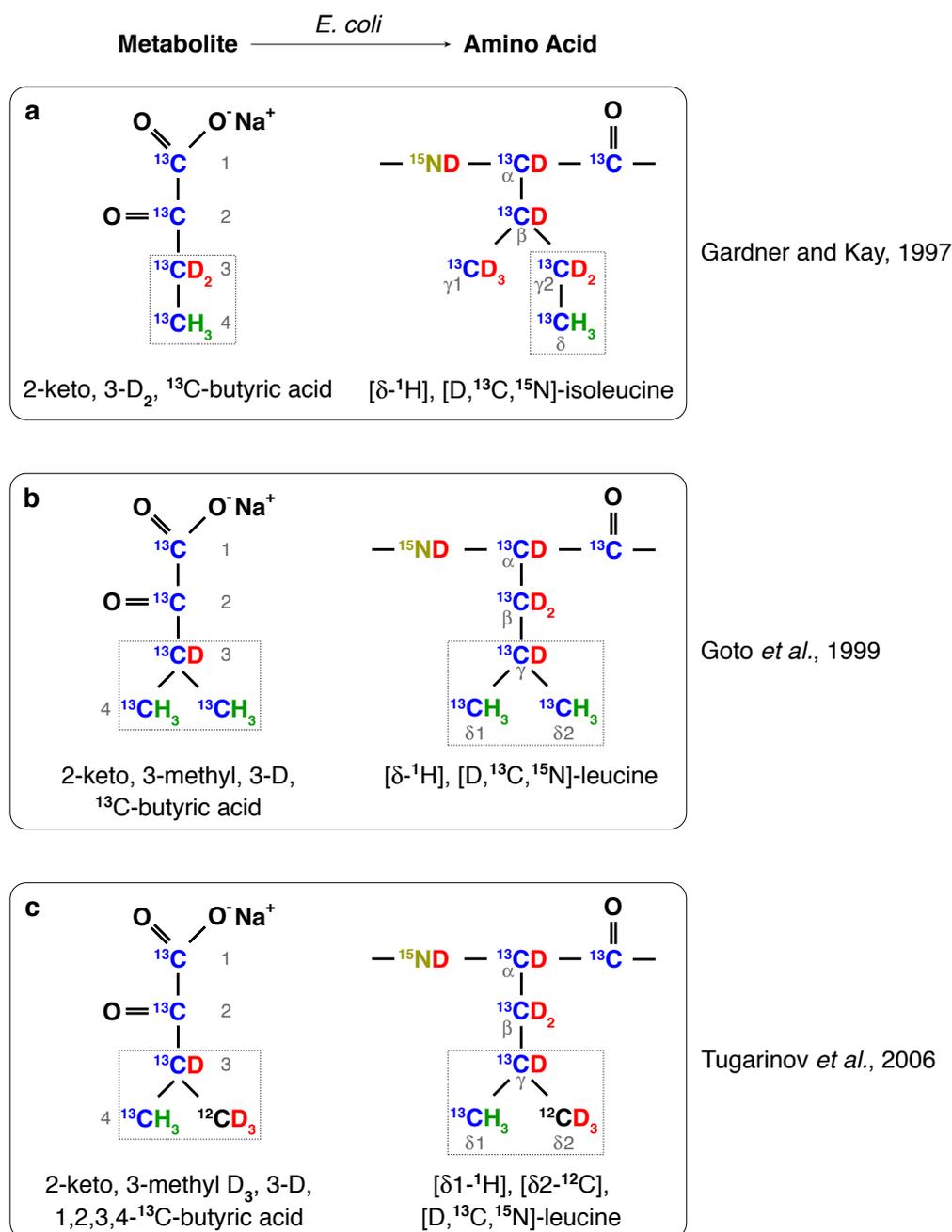
$^{15}\text{N}$  labeling of the NES peptide had been performed with 1.8 liters of the "conventional" M9-based medium described in **Table 6-1**, using glycerol as a carbon source and  $^{15}\text{NH}_4\text{Cl}$  for nitrogen supply. *E. coli* BLR was grown at 37 °C to an  $\text{OD}_{600}$  of  $\approx 0.6$  and the expression of the His<sub>10</sub>-zz-tagged peptide induced with 500  $\mu\text{M}$  IPTG for 3.5 h (i.e. until the culture reached the stationary growth phase,  $\text{OD}_{600} = 2.2$ ). The peptide was purified as described in **Chapter 4**. The final yield was  $\approx 7.5$  mg.

### 6.4.2 [ $\text{D},^{13}\text{C},^{15}\text{N}$ ]-PKI $\Phi^0$ Leu NES

For triple-labeling of the NES peptide, we employed the Silantes *E. coli* CDN (D (>95%),  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) medium. *E. coli* BLR was transformed with the expression plasmid and plated on a selective [ $^2\text{H}$ ]-plate for adaptation to the isotopes. The resulting colonies were washed off the plate and used to inoculate the medium (1 liter) to an  $\text{OD}_{600}$  of 0.02. The culture was grown at 37 °C to an  $\text{OD}_{600}$  of  $\approx 1.0$  and the expression was induced with 1 mM IPTG until the stationary growth phase was reached ( $\text{OD}_{600} = 2.1$ ). The peptide was purified as described in **Chapter 4**. The final yield was  $\approx 4.5$  mg.

### 6.4.3 [ $\delta$ $^{-1}\text{H}$ -Ile/Leu], [ $\text{D},^{13}\text{C},^{15}\text{N}$ ]-PKI $\Phi^0$ Leu NES ( $\delta$ -methyl-specific protonation at Leu and Ile)

Labeling with this isotope pattern required the use of a fully defined medium, because amino acid precursors for Ile and Leu had to be added (**Figure 6-5**; Gardner and Kay, 1997; Goto *et al.*, 1999).



**Figure 6-5: Isotope-labeled amino acid precursors used in this study.**

The resulting amino acid labeling scheme (growth in D<sub>2</sub>O-based medium) is shown.

The medium used was essentially identical to that described in **Table 6-1** (D<sub>2</sub>O-based). <sup>15</sup>ND<sub>4</sub>Cl was taken as a nitrogen source and <sup>13</sup>C, D<sub>12</sub>-glucose served for carbon supply. *E. coli* BLR was transformed with the expression plasmid and plated on a selective [D]-plate for adaptation to the isotope. The resulting colonies were washed off the plate and used to inoculate the medium (1 liter) to an OD<sub>600</sub> of 0.3. When the culture reached the mid-log growth phase (OD<sub>600</sub> = 0.8), 50 mg/l 2-keto, 3-D<sub>2</sub>, <sup>13</sup>C-butyric acid (**Figure 6-5a**) and 85 mg/l

2-keto, 3-methyl, 3-D,  $^{13}\text{C}$ -butyric acid (**Figure 6-5b**) were added. One hour later ( $\text{OD}_{600} = 0.88$ ), expression was induced with 1 mM IPTG for 3 hours (final  $\text{OD}_{600} = 2.0$ ). A longer induction time is not recommended as it might result in inhomogeneity ("scrambling") of the labeling pattern. The peptide was purified as described in **Chapter 4**. The final yield was 1.5 mg.

As the use of precursors limits the time frame for induction to only a few hours, overexpression of the protein should be preferably started at a high culture density but also in mid-log phase of the expression culture, i.e. when the generation time of *E. coli* is shortest. However, with "traditional" M9 media, the culture density at mid-log phase is typically low. Moreover, short generation times are not achieved when nutrients in the medium are limiting. Therefore, the HD-M9 medium should be best suited to generate isotope labeling schemes that involve the use of expensive precursors. In the following section, I describe the application of the HD-M9 medium to isotope labeling entailing amino acid precursors.

#### 6.4.4 [ $\delta^1\text{-}^1\text{H}$ , $\delta^2\text{-}^{12}\text{C}$ -Leu], [ $^1\text{H}$ , $^{12}\text{C}$ , $^{14}\text{N}$ -Ala], [ $\text{D}$ , $^{13}\text{C}$ , $^{15}\text{N}$ ]-PKI $\Phi^0\text{Leu NES}$ ( $\delta$ -methyl-specific labeling of Leu, protonated Ala)

The starter culture for this expression setup was prepared essentially as described for the deuteration of CRM1. To determine the best time point for the induction of NES expression, I recorded a growth curve in HD-M9 medium containing  $\text{D}_8$ -glycerol as a carbon source and  $^{15}\text{ND}_4\text{Cl}$  for nitrogen supply. Whereas for the deuteration of CRM1 no adaptation to the medium was required, this culture entered the logarithmic growth phase only after  $\approx 4$  days at  $37^\circ\text{C}$ . It remains to be tested if supplementing the HD-M9 medium with Silantes medium is decisive for more rapid adaptation. The longer adaptation time might also be a result of the different isotope composition ( $^{15}\text{N}$  instead of  $^{14}\text{N}$ ) or the antibiotic employed (kanamycin instead of ampicillin). Once the culture had entered the logarithmic phase, it proceeded into stationary phase within only  $\approx 12$  hours, reaching a saturation density of 9.8 ( $\text{OD}_{600}$ ).

The final 1-liter culture was inoculated from the logarithmically growing test setup (start- $\text{OD}_{600} = 0.1$ ). The expression culture entered the logarithmic growth phase without any delay. At  $\text{OD}_{600} = 4.2$ , 300 mg of the leucine precursor (**Figure 6-5c**, Tugarinov *et al.*, 2006) and 500 mg alanine were added to the medium. One hour later ( $\text{OD}_{600} = 4.7$ ), expression of the NES peptide was induced with 1 mM IPTG for 3 hours (final  $\text{OD}_{600} = 6.4$ ). This short induction time was sufficient to reach the maximum expression level (not shown). The purification of

the peptide was performed as described (**Chapter 4**). Notably, the yield was  $\approx 10$  times higher than that for the [ $\delta$ - $^1\text{H}$ -Ile/Leu], [D, $^{13}\text{C}$ , $^{15}\text{N}$ ]-PKI  $\Phi^0$ Leu NES (19 mg vs. 1.5 mg, see above).

---

## **6.5 The previously postulated c-Abl NES is non-functional in the context of the Abl C-terminal domain**

### *6.5.1 Reprint of Hantschel et al., 2005*

Supplemental data are available online:

<http://www.cell.com/molecular-cell/retrieve/pii/S1097276505014334>

Additional data are presented after this section.

## Structural Basis for the Cytoskeletal Association of Bcr-Abl/c-Abl

Oliver Hantschel,<sup>1,2,5</sup> Silke Wiesner,<sup>3,5,7</sup>  
Thomas Güttler,<sup>4,6</sup> Cameron D. Mackereth,<sup>3,6</sup>  
Lily L. Remsing Rix,<sup>1,2,6</sup> Zsuzsanna Mikes,<sup>2</sup>  
Jana Dehne,<sup>2</sup> Dirk Görlich,<sup>4</sup> Michael Sattler,<sup>3,\*</sup>  
and Giulio Superti-Furga<sup>1,2,\*</sup>

<sup>1</sup>Center for Molecular Medicine of the Austrian  
Academy of Sciences

Lazarettgasse 19/3

1090 Vienna

Austria

<sup>2</sup>Developmental Biology Programme

<sup>3</sup>Structural and Computational Biology Programme

European Molecular Biology Laboratory

Meyerhofstrasse 1

69117 Heidelberg

Germany

<sup>4</sup>ZMBH

University of Heidelberg

Im Neuenheimer Feld 282

69120 Heidelberg

Germany

### Summary

The Bcr-Abl tyrosine kinase causes different forms of leukemia in humans. Depending on its position within the cell, Bcr-Abl differentially affects cellular growth. However, no structural and molecular details for the anticipated localization determinants are available. We present the NMR structure of the F-actin binding domain (FABD) of Bcr-Abl and its cellular counterpart c-Abl. The FABD forms a compact left-handed four-helix bundle in solution. We show that the nuclear export signal (NES) previously reported in this region is part of the hydrophobic core and nonfunctional in the intact protein. In contrast, we could identify the critical residues of helix  $\alpha$ III that are responsible for F-actin binding and cytoskeletal association. We propose that these interactions represent a major determinant for both Bcr-Abl and c-Abl localization.

### Introduction

Expression of Bcr-Abl, the oncogenic counterpart of the tyrosine kinase c-Abl and the outcome of the Philadelphia chromosome translocation (t[9;22]), is the basis for all cases of chronic myelogenous leukemia (CML) and a subset of acute lymphocytic leukemia (ALL) (Wong and Witte, 2004). The predominant distinguishing feature of Bcr-Abl is its high level of tyrosine kinase

activity compared to the kinase activity of c-Abl, which is effectively autoinhibited by different intramolecular interactions (Raitano et al., 1997; Hantschel and Superti-Furga, 2004). Inhibition of the tyrosine kinase activity of Bcr-Abl by the small-molecule inhibitor Imatinib/Gleevec has become a paradigm for modern targeted cancer therapy, but emerging cases of drug resistance create a demand for additional therapeutic intervention strategies (Druker, 2004; Sawyers, 2004). Bcr-Abl is predominantly localized to the cytoplasm of cells, where it activates different signal transduction pathways, leading to cell proliferation (McWhirter and Wang, 1991, 1993; Steelman et al., 2004). However, relocalization to the nucleus can be achieved by either combination of pharmacological treatment or genotoxic stress resulting in alteration of its growth-promoting properties (Dierov et al., 2004; Vigneri and Wang, 2001). The cellular form c-Abl localizes to varying degrees to both the cytoplasm and the nucleus depending on tissue type and environmental cues and is thought to shuttle between these two compartments (Van Etten et al., 1989; Wetzler et al., 1993; Renshaw et al., 1988; Taagepera et al., 1998).

Bcr-Abl spans more than 2000 amino acid residues and comprises a complex array of multiple protein-protein interaction and signaling domains, including a centrally located SH3-SH2-tyrosine kinase domain module (Figure 1A). The proposed localization determinants are all found in the C-terminal last exon region that is common to both Bcr-Abl and c-Abl. These include three nuclear localization signals (NLSs) (Wen et al., 1996), a leucine-rich NES (Taagepera et al., 1998), and an FABD (McWhirter and Wang, 1993; Van Etten et al., 1994) (Figure 1A). Although major insight into the regulatory mechanisms of c-Abl autoinhibition could be provided by solving the structure of the SH3-SH2-kinase domain fragment, still only about one-third of Bcr-Abl's protein structure is known to date (Hantschel and Superti-Furga, 2004; Harrison, 2003; Nagar et al., 2003; Hantschel et al., 2003). To explore the molecular basis of Bcr-Abl/c-Abl localization, we solved the structure of the FABD and studied its molecular function. Our data provide structural insights into a localization determinant of Bcr-Abl/c-Abl.

### Results and Discussion

#### NMR Structure of the FABD

We determined the three-dimensional (3D) solution structure of the FABD (corresponding to residues 1026–1149 of human c-Abl spliceform 1b) by heteronuclear NMR spectroscopy (Figures 1B and 1C and Table 1). This domain also includes the previously proposed NES (residues 1109–1118) (Taagepera et al., 1998).

The ordered region of the FABD, which encompasses residues 1047–1149, is a monomer in solution and folds as a compact bundle of four antiparallel  $\alpha$  helices, which are arranged in a left-handed topology (Figures 1B and 1C). A predominantly aliphatic core formed by

\*Correspondence: gsuperti@cemm.oeaw.ac.at (G.S.-F.); sattler@embl.de (M.S.).

<sup>5</sup>These authors contributed equally to this work.

<sup>6</sup>These authors contributed equally to this work.

<sup>7</sup>Present address: Structural Biology and Biochemistry, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada.

Molecular Cell  
462

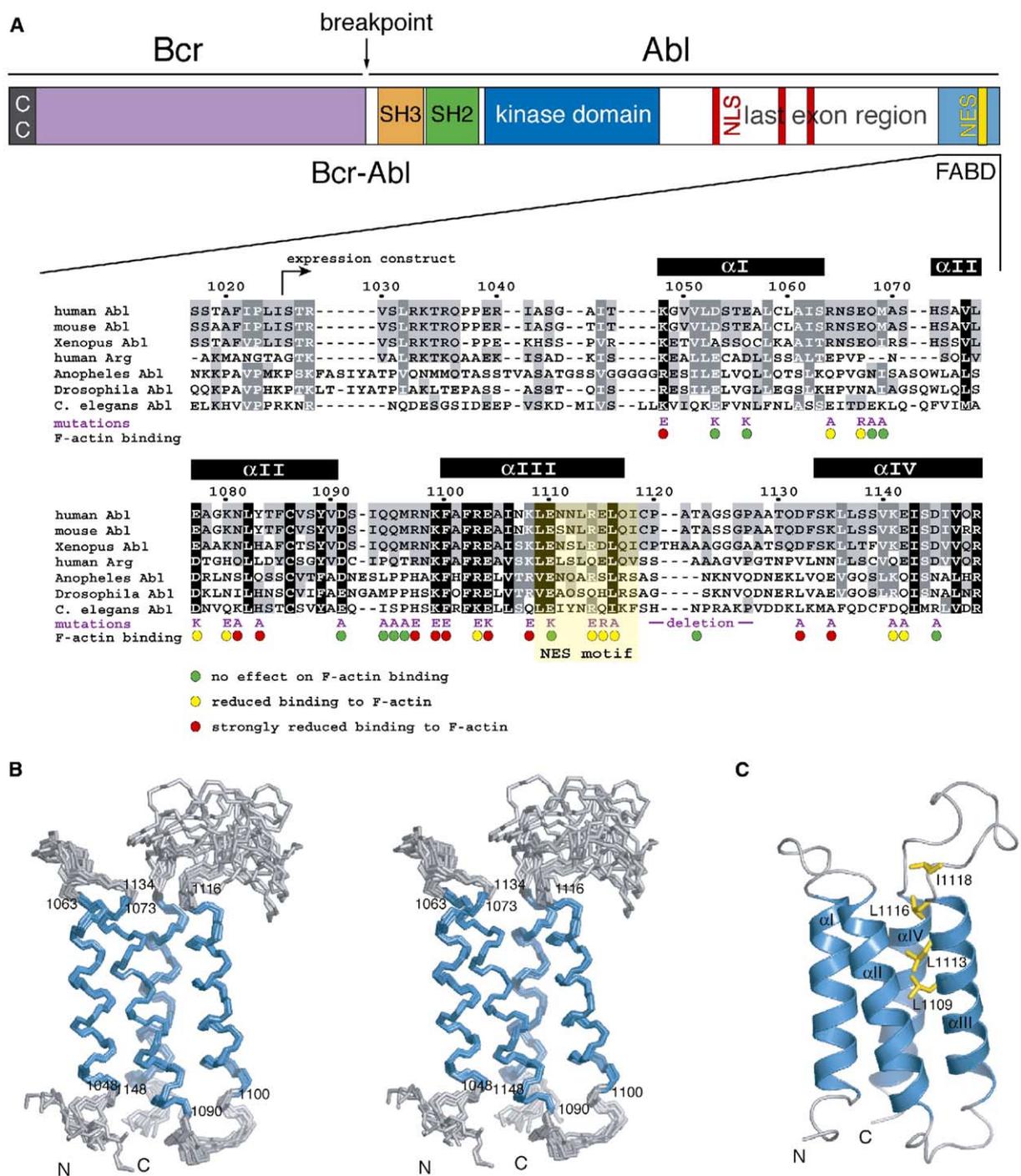


Figure 1. Structure of the Bcr-Abl/c-Abl F-actin binding domain

(A) Domain architecture of the Bcr-Abl fusion protein (top). Bottom, multiple sequence alignments of Abl F-actin binding domains highlighting partial (gray), high (white letters on gray background), and complete (white letters on black background) residue conservation. Sequences of Abl in *Homo sapiens*, *Mus musculus*, *Xenopus laevis*, *Anopheles gambiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and of the Abl paralogue Arg (ABL2) in *Homo sapiens* are shown. Residue numbers and secondary structure elements (boxes) are shown on top. The NES and introduced mutations together with results from the F-actin cosedimentation assay are shown on the bottom. Residue numbers correspond to human c-Abl 1b numbering. Abbreviations: CC, coiled-coil domain; SH3, Src-homology 3 domain; SH2, Src-homology 2 domain; NLS, nuclear localization signal; and NES, nuclear export signal.

(B) Stereoview of the ensemble of the ten lowest-energy NMR structures. The backbone trace is shown in gray, whereas secondary structure elements are shown in blue and labeled by residue numbers. Unstructured, N-terminal residues (1026–1039) are omitted for clarity.

(C) Ribbon representation of the lowest energy structure. Secondary structure elements are labeled and shown in blue. The side chains of residues defining the previously proposed NES are shown in yellow.

F-Actin Binding Domain of Bcr-Abl  
463

Table 1. Structural Statistics of the Ten Lowest Energy Structures of the Bcr-Abl/c-Abl FABD

Number of Structural Restraints		
All	2693	
Sequential ( $ i - j  = 1$ )	497	
Medium range ( $2 \leq  i - j  \leq 4$ )	460	
Long range ( $ i - j  > 4$ )	369	
Intraresidual	1265	
Unambiguous	2591	
Ambiguous	0	
Hydrogen bonds	102	
Dihedral angles	71 $\phi$ , 71 $\psi$ , 14 $\chi_1$	
$^1\text{H}^{\text{N}}$ - $^{15}\text{N}$ residual dipolar couplings	64	
Rmsd from Experimental Restraints <sup>a</sup>	<SA> <sup>b</sup>	<SA> <sup>water-refined</sup>
All distance restraints [ $\text{\AA}$ ]	$0.027 \pm 0.002$	$0.037 \pm 0.002$
Unambiguous NOEs [ $\text{\AA}$ ]	$0.027 \pm 0.002$	$0.037 \pm 0.002$
Hydrogen bonds [ $\text{\AA}$ ]	$0.024 \pm 0.002$	$0.025 \pm 0.003$
Dihedral angles [ $^\circ$ ]	$0.510 \pm 0.101$	$0.518 \pm 0.101$
RDC Q factor <sup>c</sup>	$0.046 \pm 0.003$	$0.058 \pm 0.004$
Rmsd from Idealized Covalent Geometry		
Bond lengths [ $\text{\AA}$ ]	$(2.9 \pm 0.2) \times 10^{-3}$	$(4.9 \pm 0.2) \times 10^{-3}$
Bond angles [ $^\circ$ ]	$0.468 \pm 0.014$	$0.655 \pm 0.017$
Improper dihedral angles [ $^\circ$ ]	$0.415 \pm 0.028$	$1.668 \pm 0.102$
Coordinate Precision <sup>d</sup> [ $\text{\AA}$ ]		
Secondary structure elements (N, C $^\alpha$ , C')	$0.33 \pm 0.06$	$0.42 \pm 0.08$
Secondary structure elements (all heavy atoms)	$0.94 \pm 0.10$	$1.03 \pm 0.11$
Residues 1043–1149 (N, C $^\alpha$ , C')	$1.73 \pm 0.45$	$1.83 \pm 0.85$
Residues 1043–1149 (all heavy atoms)	$2.04 \pm 0.35$	$2.15 \pm 0.67$
Ramachandran Plot <sup>e</sup> [%]		
Most favored regions (secondary structure elements)	$99.7 \pm 0.7$	$99.8 \pm 0.5$
Additionally allowed regions (secondary structure elements)	$0.3 \pm 0.7$	$0.2 \pm 0.5$
Most favored regions (residues 1043–1149)	$77.6 \pm 2.5$	$80.3 \pm 2.3$
Additionally allowed region (residues 1043–1149)	$22.4 \pm 2.5$	$19.7 \pm 2.3$

<sup>a</sup>In any of the final <SA> structures, no distance restraint was violated by more than 0.4  $\text{\AA}$ , no dihedral angle restraint was violated by more than 5 $^\circ$ , and no residual dipolar coupling restraint was violated by more than 2.5 Hz.

<sup>b</sup><SA> refers to the ensemble of the ten structures with the lowest energy.

<sup>c</sup>Reference: Cornilescu et al., 1998.

<sup>d</sup>Coordinate precision is given as the pair-wise Cartesian coordinate rmsd of the ten lowest-energy structures.

<sup>e</sup>Excluding glycine and proline residues.

Val1051, Leu1058 (in helix  $\alpha\text{II}$ ), Val1075, Leu1082, Cys1086 ( $\alpha\text{II}$ ), Phe1102, Ala1105, Leu1109, Leu1113 ( $\alpha\text{III}$ ), Leu1136, Val1140, Ile1143, and Val1147 ( $\alpha\text{IV}$ ) stabilizes this fold. The high degree of conservation of these residues suggests a similar fold for the FABDs of all c-Abl orthologs as well as the c-Abl paralogue Arg (*ABL2*) (Figure 1A). The  $\alpha$  helices are connected by long loops that are poorly conserved in length and sequence.  $^{15}\text{N}$  relaxation data shows that the  $\alpha\text{III}$ - $\alpha\text{IV}$  loop displays high internal mobility and that the 20 N-terminal residues of the expression construct are unstructured (Figure S1 available in the Supplemental Data with this article online).

#### Structural Homologs of the FABD

The structure of the Bcr-Abl/c-Abl FABD is clearly distinct from calponin-homology domains that were recently proposed for the Arg equivalent of the FABD but that bind F-actin only in a tandem configuration (Galkin et al., 2005). Although the Bcr-Abl/c-Abl FABD has no significant sequence similarity to other established F-actin binding domains, the 3D structure exhibits strong homology to domains of other cytoskeletal pro-

teins of which some bind F-actin directly or indirectly: the vinculin head and tail domain (with DALI Z scores [Z] of 9.9 and 8.2, respectively), the dimerization and adhesion modulating domain of  $\alpha$ -catenin (Z scores of 9.3 and 9.5, respectively), the vinculin binding domain of talin (Z = 7.9), and the focal adhesion targeting domain of focal adhesion kinase (Z = 7.5) (Figure 2). Sequence homology of these four-helix bundles is limited only to their helical amphipathicity and could not be revealed by simple sequence comparison. Although many proteins include four-helix bundles, the similarity in the interhelical geometry and the unusual left-handed topology between these domains and the Bcr-Abl/c-Abl FABD is intriguing and, besides F-actin binding (see below), may suggest a role in protein complex assembly that is currently not understood in molecular detail (Fillingham et al., 2005). The helix bundle architecture is emerging as a common interaction module in a number of proteins involved in cytoskeletal regulation (Hayashi et al., 2002), pointing at a possible function of Bcr-Abl/c-Abl in the multiprotein complexes anchoring actin filaments. Indeed, cytoplasmic c-Abl colocalizes with many cytoskeletal structures and effectors, thereby

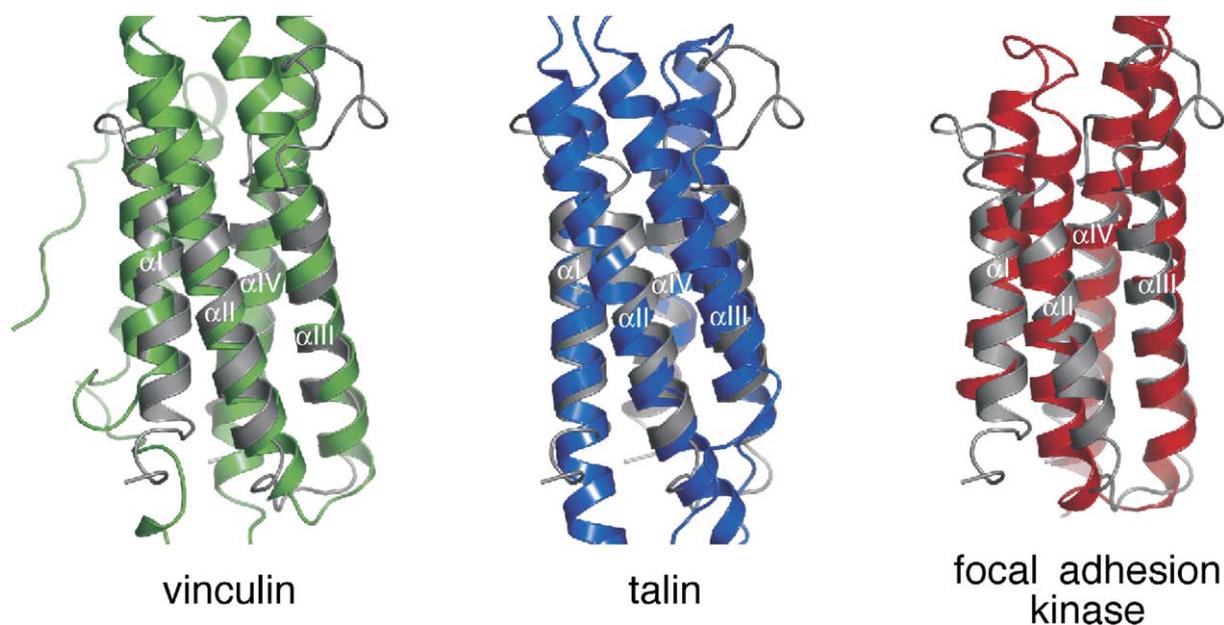


Figure 2. Structural Homologs of the Bcr-Abl/c-Abl FABD

Despite complete absence of sequence similarity, the Bcr-Abl/c-Abl FABD is structurally homologous to domains present in cell adhesion proteins. In all panels, the Bcr-Abl/c-Abl FABD is shown in gray, whereas the vinculin tail domain is displayed in green (left, PDB-entry 1QKR), the vinculin binding domain of talin in blue (middle, PDB-entry 1S7J), and the focal adhesion targeting domain of focal adhesion kinase in red (right, PDB-entry 1K40).

integrating multiple signals to coordinate F-actin dynamics (Woodring et al., 2003). Previous studies have shown that F-actin binding inhibits c-Abl kinase activity (Woodring et al., 2001). The mechanism by which this occurs is currently unclear but may involve intramolecular interactions between the FABD and other regions of the protein, e.g., the kinase domain. Further structure-function analysis of the FABD should be instrumental in elucidating this regulatory function.

#### The NES Is Buried and Nonfunctional

Residues comprising the proposed NES (residues 1109–1118) (Taagepera et al., 1998) are located on the  $\alpha$ III helix and the  $\alpha$ III- $\alpha$ IV loop (Figure 1C). Whereas Ile1118 is solvent accessible, Leu1116 is only partially exposed. Both Leu1109 and Leu1113 are buried in the hydrophobic core of the domain. All three leucines (Leu1109, Leu1113, and Leu1116) are involved in an extensive network of hydrophobic interactions with Leu1136 ( $\alpha$ IV) and with Leu1076, Gly1079, and Leu1082 ( $\alpha$ II).

The fact that most residues of the previously proposed NES (residues 1109–1118) are not solvent exposed raised questions about its physiological relevance. We first investigated whether the isolated NES peptide can form an export complex with the nuclear export receptor CRM1 (exportin 1). The immobilized peptide efficiently recruited CRM1 from a HeLa cell extract, provided RanGTP had been added to mimic a nuclear environment (Figure 3A, lane 7). Leu1116 is essential for NES function of the isolated peptide (Taagepera et al., 1998), and indeed, the L1116A mutation abolishes CRM1 binding (Figure 3A, lane 9). Thus, residues 1109–1118 behave like a leucine-rich NES when presented in isolation. If this NES were also func-

tional in the original domain context, then the FABD should also specifically interact with CRM1. However, CRM1 bound only inefficiently (Figure 3A, lane 3), and crucially, this weak binding was not further reduced by the L1116A mutation (Figure 3A, lane 5). This indicates that the NES is not accessible to CRM1 and that the FABD binds CRM1 only nonspecifically.

Furthermore, immobilized CRM1 specifically bound a green fluorescent protein (GFP)-NES fusion, but not the untagged recombinant FABD, thus excluding the possibility that tagging of the domain or a protein present in the HeLa cell extract had compromised export complex formation (Figure 3B).

To test for nuclear export of the FABD, we performed experiments with HeLa cell nuclei in *Xenopus laevis* egg extract depleted of endogenous nuclear transport receptors (Ribbeck and Gorlich, 2002). The nucleocytoplasmic distribution of added recombinant GFP-FABD and GFP-NES was monitored by confocal microscopy after initial equilibration by diffusion. Upon CRM1 addition, we observed rapid export of GFP-NES, but not of GFP-FABD (Figure 3C). No export occurred when exportin 6 or buffer were added instead of CRM1 (Figure 3C). In agreement with these results, we did not detect export of GFP-FABD or  $^{35}\text{S}$ -methionine-labeled full-length c-Abl upon microinjection into the nuclei of frog oocytes (data not shown). Finally, treatment of HeLa or COS cells with the specific CRM1 inhibitor leptomycin B (LMB) did not lead to nuclear accumulation of either endogenous or overexpressed c-Abl (Figure 4 and data not shown). Overall, our results indicate that residues 1109–1118 of Bcr-Abl/c-Abl do not constitute a functional NES in the context of the folded FABD.

Furthermore, as judged from  $^{15}\text{N}$  relaxation studies,

the NES does not exhibit internal mobility (Figure S1). Together with the functional data above, this suggests that the NES may not become exposed by a flipping or unfolding mechanism. However, we cannot completely exclude the possibility that under conditions different than the ones tested here, mechanisms may exist that trigger unfolding of the FABD and a release of another functionality, such as the hidden NES motif.

#### FABD-Dependent Cytoplasmic Localization Pattern

To analyze the precise role of the FABD in the cellular localization of Bcr-Abl and c-Abl, we performed extensive structure-based mutagenesis, careful to minimize general structure perturbation and maximize specific functional effects. We introduced a set of 21 mutants covering 39 solvent-exposed residues (Figure 1A and Table 2) into full-length p210 Bcr-Abl and human c-Abl 1b to perform confocal immunofluorescence microscopy with transiently transfected COS cells. Under the settings used, the low levels of endogenous c-Abl could hardly be detected (data not shown).

Wild-type (wt) Bcr-Abl colocalized with actin filaments, whereas the FABD-deletion mutant failed to do so. Instead, the mutant displayed a diffused localization throughout the cytoplasm with additional punctuated staining, which had previously been observed with other deletion mutants encompassing the same region (McWhirter and Wang, 1993) (Figure 5A and data not shown). Localization patterns of the different Bcr-Abl FABD point mutants fell between that of the wt protein and that of the FABD-deletion mutant (Table 2). Three mutations, R1097E, K1099E, and F1100E, showed the Bcr-Abl localization phenotype most similar to the FABD-deletion mutant (F1100E shown in Figure 5A). The punctuate staining seemed to depend on kinase activity, as cells treated with the kinase inhibitor Imatinib/Gleevec did not show this phenotype (Figure S2).

Differences in actin colocalization between wt and mutant proteins of c-Abl were less evident but still readily detectable (Figure 5B). Mutants that showed the strongest defect in association with actin filaments were the same for both Bcr-Abl and c-Abl. Interestingly, some of these c-Abl FABD mutants showed a partial or strong nuclear localization (Table 2 and Figure 5B), such as F1100E and the L1116A mutation of the previously proposed NES.

#### The F-Actin Binding Site

To establish the precise molecular nature of the FABD-dependent localization properties of Bcr-Abl/c-Abl described above, we performed a quantitative analysis of the ability of the recombinant FABD wt and mutants to bind purified F-actin *in vitro*. All FABD proteins were expressed as GST-fusions, purified to homogeneity, and tested in an F-actin cosedimentation assay (Figure S3A). Structural integrity was assessed by using gel filtration chromatography, NMR, and far-ultraviolet circular dichroism spectroscopy (for selected mutants, see Figures S3B-S3D and data not shown). Although GST-FABD wt was enriched in the pellet in the presence of F-actin, only a minor proportion was found in the pellet fraction in the absence of F-actin (Figure 6A, top). This suggests specific binding to F-actin *in vitro*. The FABD mutants were tested accordingly and compared

to the wt protein (Table 2). Mutation of Arg1097, Lys1099, Phe1100, and Glu1104/Lys1108 impaired F-actin binding strongest (Table 2 and Figures 1A, 6A, and 6B). Importantly, mutation of residues within the previously proposed NES (Leu1116, Arg1114, and Glu1115) led to an intermediate reduction in F-actin binding (Table 2). Independently, NMR titration experiments with <sup>15</sup>N-labeled FABD and F-actin showed a general line-broadening effect depending on the F-actin concentration consistent with the formation of a high-molecular weight complex (Figure S4). Although for some domains structurally related to the FABD a potential to undergo dramatic conformational changes upon ligand binding has been observed, we have not observed any major chemical shift changes in our NMR titrations that might indicate such a behavior for the Bcr-Abl/c-Abl FABD.

Color coding of the surface of the FABD according to the contribution of the residues to F-actin binding revealed a cluster of residues along helix  $\alpha$ III, including the previously proposed NES, as well as adjacent portions of helix  $\alpha$ II and helix  $\alpha$ IV (Figure 6B). Independently, analysis of evolutionary conservation among Abl FABDs showed that, in contrast to most of the solvent-exposed residues, residues in the N-terminal half of  $\alpha$ III are highly conserved (Figure 6C). The fact that the only highly conserved region of the domain coincides with the F-actin binding epitope suggests that F-actin binding is the key conserved function among the various Abl FABDs.

To visualize how the evolutionarily conserved helix  $\alpha$ III could possibly interact with F-actin, we generated models of the Bcr-Abl/c-Abl FABD-actin complex with the program HADDOCK (Dominguez et al., 2003) using an ensemble of seven NMR structures of the Abl FABD and actin coordinates from four different actin complexes. Although the residues in actin involved in binding the FABD could not be determined experimentally, crystal structures of actin complexes reveal a hydrophobic cleft between actin subdomains 1 and 3 (Figure S5A) as the primary interaction site for F-actin binding proteins (Dominguez, 2004; Aguda et al., 2005). By superimposing structures of four actin complexes, 19 residues in and surrounding this hydrophobic cleft were chosen to generate docking restraints. Indeed, the conserved helix  $\alpha$ III could be docked reliably onto a hydrophobic cleft in the actin structure. It is worth noting that this binding mode would be compatible with the simultaneous actin-actin interactions within F-actin as predicted in the Holmes model (Holmes et al., 1990) (Figure S5).

#### Molecular Basis for the NES Phenotype

Leucine-rich NESs consist of a stretch of regularly spaced hydrophobic residues, an abundant sequence motif, and an intrinsic feature of amphipathic helices. Previous studies have led to the widely acknowledged view that c-Abl and Bcr-Abl possess a functional NES at the C terminus (Taagepera et al., 1998; Wang, 2000; Henderson and Eleftheriou, 2000; Zhu and Wang, 2004). Several key experiments that led to this perception, however, were performed with the NES as a peptide or with constructs that inadvertently affected the FABD structural integrity. Unfolding of the FABD had two consequences. On one hand, normally buried elements

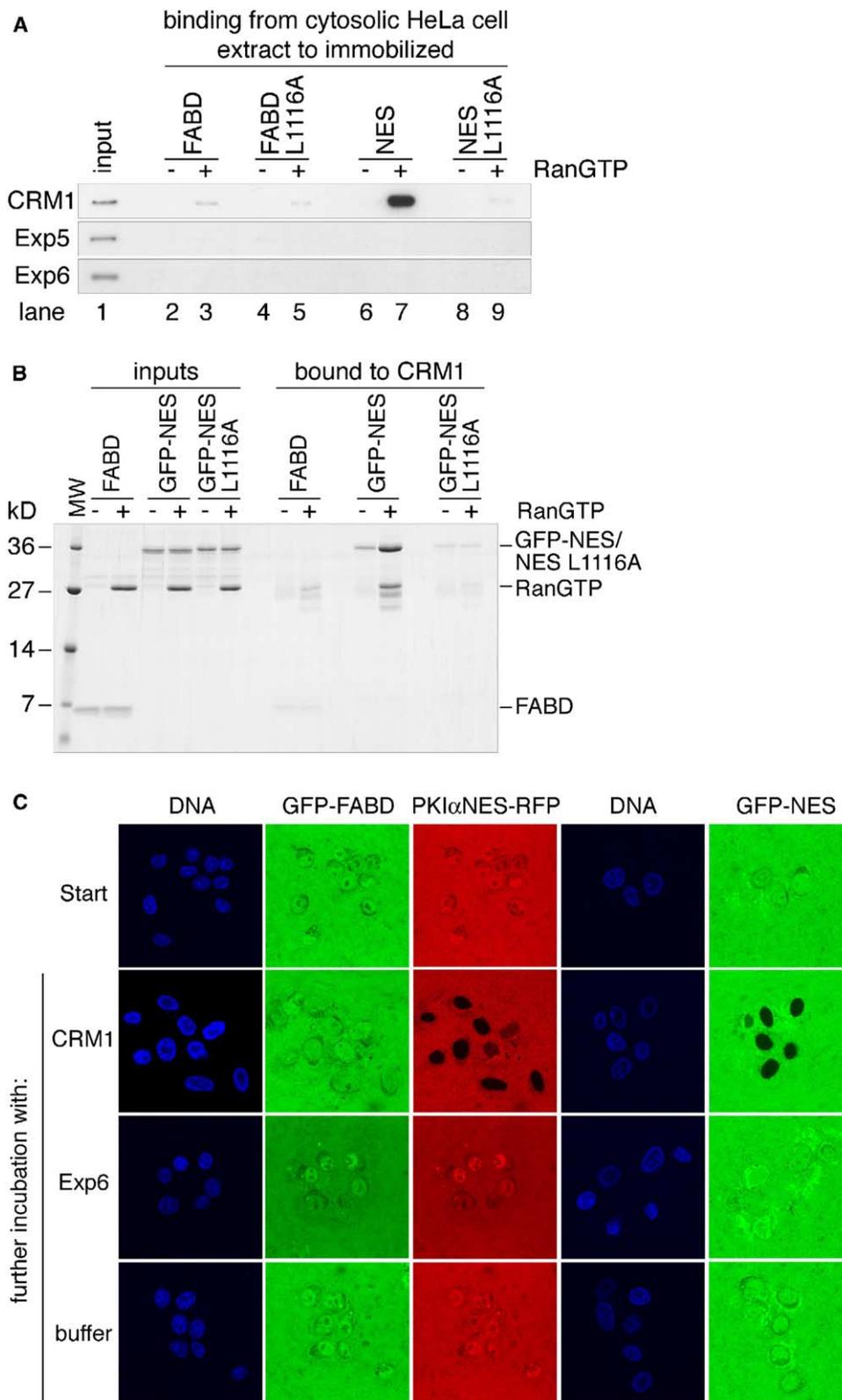
Molecular Cell  
466

Figure 3. The Folded FABD Does Not Bind to CRM1 and Is Not Exported from the Nucleus

(A) FABD, NES<sup>1109-1118</sup>, and the respective L1116A mutants were immobilized on a Sepharose resin in equimolar amounts and allowed to bind exportins from a HeLa cell extract. RanGTP was used to mimic a nuclear environment. Starting material and bound fractions were analyzed by SDS-PAGE and immunoblotting. Exportin 5 (Exp5) and exportin 6 (Exp6) are included as specificity controls. Load in the bound fractions corresponds to 60 times the starting material.

F-Actin Binding Domain of Bcr-Abl  
467

the hydrophobic core became solvent accessible and thereby established a functional, but nevertheless non-physiological CRM1 binding site. On the other hand, the disrupted domain could no longer bind F-actin. In all these cases, subcellular localization is affected in consequence. As we have shown above, the classic Abl “NES-” mutation L1116A causes impaired F-actin binding, thereby leading to nuclear accumulation indirectly. This case is consistent with the general notion that mutation of linear motifs of proteins in the absence of 3D structural information can be difficult to interpret (Puntervoll et al., 2003). As shown for the nonsense-mediated mRNA decay factor UPF3b (Kadlec et al., 2004), we predict that once structural information becomes available, several other NESs that have been proposed based only on the amino acid sequence will also be revealed to be just nonfunctional, buried hydrophobic core residues.

### Conclusion

Taken together, our data strongly suggest that an equilibrium between nuclear import and FABD-dependent cytoplasmic retention determines the subcellular distribution of c-Abl. In support of this, caspase cleavage of c-Abl causes removal of the FABD and leads to nuclear localization and contributes to the induction of apoptosis (Barilá et al., 2003).

The c-Abl and Bcr-Abl proteins share most of their sequences, and indeed, evidence that many of the intramolecular regulatory mechanisms that inhibit c-Abl kinase activity are also operational in Bcr-Abl has emerged recently (Azam et al., 2003; Smith et al., 2003; Hantschel and Superti-Furga, 2004). The Bcr moiety, unique to Bcr-Abl, contributes a coiled-coil oligomerization function that results in proximity-induced *trans*-phosphorylation of the activation loop and consequent high levels of kinase activity (McWhirter et al., 1993; Zhang et al., 2001). We also propose here that the subcellular localization of these two proteins is subject to the same basic mechanisms. But why is mutation of the FABD of Bcr-Abl, in contrast to c-Abl, not sufficient to cause relocalization of the protein from the actin cytoskeleton to the nucleus? In our hands, treatment with the kinase inhibitor Imatinib/Gleevec, although changing the pattern of cytoplasmic staining, did not lead to nuclear localization of FABD-deficient forms of Bcr-Abl, thus excluding a predominant role of kinase activity in cytoplasmic retention (Figure S2). However, preliminary experiments with forms of Bcr-Abl mutated in the coiled-coil domain, in addition to a disrupting mutation in the FABD domain, showed a weak but reproducible nuclear staining (data not shown). This suggests that interaction of Bcr-Abl with the cytoskeleton through the

FABD as well as oligomerization contribute to offset any nuclear localization force acting on the protein. Moreover, recent evidence suggests that nuclear import of c-Abl can be inhibited by phosphorylation of a threonine residue (Yoshida et al., 2005). This mechanism might also account for inhibition of nuclear entry of Bcr-Abl.

Irrespective of what precise subcellular distribution may result from actin binding deficiency of Bcr-Abl under different conditions, there appear to be important functional consequences. It has been reported that deletions disrupting FABD integrity impaired both the transforming potential of Bcr-Abl in rat fibroblasts as well as oncogenicity in transgenic mice (McWhirter and Wang, 1993; Heisterkamp et al., 2000). Because we could confine the critical elements of association with the actin cytoskeleton to the N-terminal half of the helix  $\alpha$ III on Bcr-Abl, it might be possible to interfere with this interaction pharmacologically, thus modulating the biological activity of Bcr-Abl. Such intervention may be complementary to targeting of the ATP binding or substrate binding pocket.

### Experimental Procedures

#### Bacterial Expression Constructs

A pETM30 vector (modified pET24d [Novagen] provides an N-hexa-His-GST-tag, followed by a TEV-protease cleavage site) was used for NMR-sample preparation and F-actin cosedimentation assays. For binding and nuclear export assays, two derivatives of the pQE vectors (Qiagen) were used, one providing an N-decaHis-zz-TEV-tag and the other an N-decaHis-EGFP-tag. The coding region of the F-actin binding domain (residues 1026–1149, human c-Abl spliceform 1b numbering) of human Bcr-Abl/c-Abl was cloned into the vectors described above by using standard molecular cloning techniques.

#### Protein Expression and Purification

For NMR studies, uniformly  $^{15}\text{N}$ - and  $^{15}\text{N},^{13}\text{C}$ -labeled hexaHis-GST-FABD was expressed as described (Wiesner et al., 2005).

All other proteins were expressed in *E. coli* BLR and purified on Ni-NTA-agarose. Untagged FABD was generated by TEV-protease cleavage of decaHis-zz-FABD, and removal of the tag was via Ni-NTA-agarose. Expression and purification of zz-CRM1-His, His-CRM1, His-Ran(Q69L)GTP, PKI $\alpha$ -NES-mRFP-His, and His-Exp6 has been described previously (Stuven et al., 2003).

#### Structure Determination

NMR spectra were recorded at 295 K on Bruker DRX 500, 600, and 900 spectrometers, processed with the NMRPipe/NMRDraw package, and analyzed with XEASY.  $^{15}\text{N}$  relaxation data were analyzed by using NMRView.  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  chemical shifts were assigned by standard triple-resonance experiments as described (Wiesner et al., 2005).

Proton-proton distance restraints were derived from 3D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOE experiments. Dihedral angle restraints for the backbone angles  $\phi$  were determined from quantitative  $^3\text{J}(\text{H}^N, \text{H}^\alpha)$ -correlation experiments (Kubinowa et al., 1994), whereas additional

(B) CRM1 was immobilized as above, and recombinant FABD, GFP-NES $^{1109-1118}$ , and GFP-NES L1116A were allowed to bind to CRM1 in the absence or presence of RanGTP. Analysis of starting material and bound fractions was by SDS-PAGE and Coomassie staining. Load in the bound fractions corresponds to 40 times the starting material.

(C) HeLa cell nuclei were incubated in a *Xenopus laevis* egg extract, which had been depleted of endogenous nuclear transport receptors. GFP-FABD or GFP-NES was added and allowed to enter the nuclei (“start”). The NES of the inhibitor of protein kinase A fused to red fluorescent protein (PKI $\alpha$ -NES-RFP) was used as an internal positive control. The samples were then split and either CRM1, exportin 6, or buffer added. 20 min later, the distribution of the fluorescent proteins was imaged by confocal laser scanning microscopy. DNA staining with DAPI highlights nuclei.

Molecular Cell  
468

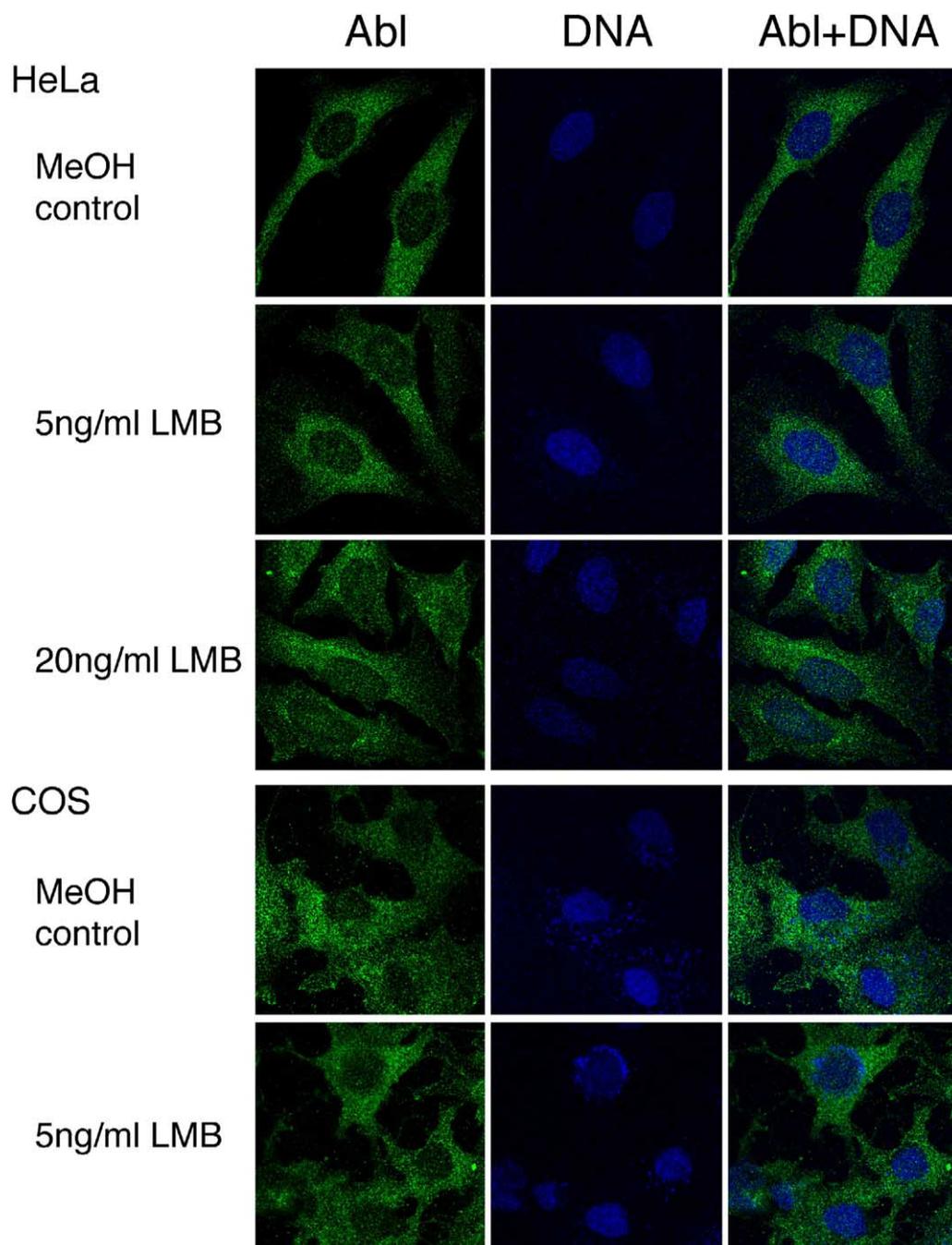


Figure 4. LMB Treatment Does Not Increase Nuclear Levels of c-Abl

Endogenous c-Abl-expressing HeLa and COS cells were treated with 5 ng/mL and/or 20 ng/mL LMB for 4 hr. Cells were fixed and immunostained with anti-Abl antibody (Abl). Nuclei (DNA) were stained with DRAQ5.

$\phi$  and  $\psi$  restraints were derived from TALOS (Cornilescu et al., 1999).  $^3J(N,C')$  coupling constants were measured by using spin-echo difference experiments to restrain the side-chain angle  $\chi_1$  to  $180^\circ \pm 40^\circ$  and  $0^\circ \pm 90^\circ$ , respectively (Hu and Bax, 1997).  $^1H,^{15}N$ -residual dipolar couplings were measured in ether/alcohol liquid crystalline phase (Ruckert and Otting, 2000) by using a 2D spin-state-selective coherence transfer HSQC experiment (Lerche et al., 1999). Axial and rhombic components of the alignment tensor were  $-8.5$  Hz and  $-3.6$  Hz, respectively. Hydrogen bond restraints were applied as indicated by  $^3J(H^N, H^\alpha)$  constants,  $^{13}C^\alpha$  and  $^{13}C^\beta$  second-

ary chemical shifts, H/D amide exchange rates, and NOE patterns. For structure calculation, the experimentally determined restraints were applied in a mixed torsion, and Cartesian angle dynamics simulated annealing protocol by using CNS (Brunger et al., 1998) and ARIA 1.2 (Nilges and O'Donoghue, 1998). Structures were calculated in eight iterations producing 20 structures in each of the first seven iterations and 50 structures in the final iteration. The quality of the ten lowest energy structures was analyzed by using the programs CNS and PROCHECK-NMR (Laskowski et al., 1996). Table 1 summarizes the structural statistics for the ensemble of the

F-Actin Binding Domain of Bcr-Abl  
469

Table 2. Mapping of the F-Actin Binding Site of the FABD

		F-Action Binding	Localization		
		GST-FABD <sup>a</sup>	Bcr-Abl <sup>b</sup>	c-Abl <sup>c</sup>	
Structural Location	Mutation	In Vitro Binding	Actin Colocalization	Cytoplasm-Nucleus	Actin Colocalization
	wt	1.00	wt	C	A
	ΔFABD	–	p	C/N	a
αI	K1048E	0.61	wt	C	a
αI	D1053K/E1056K	1.05	wt	C	A
αI	R1064A/E1067R	0.92	wt	C	A
L I-II	E1067A/Q1068A/M1069A	1.20	wt	C	A
αII	E1077K/K1080E	0.85	wt	C/N	a
αII	N1081A/Y1083A	0.59	wt	C	A
αII	D1091A	1.03	wt	Not tested	Not tested
L II-III	Q1094A/Q1095A/M1096A	1.21	Not tested	C	A
L II-III	R1097E	0.41	p	C/N	a
αIII	K1099E	0.51	p	C	a
αIII	F1100E	0.41	p	N	a
αIII	R1103E	0.70	p	C	a
αIII	E1104K/K1108E	0.55	p	C/N	A
αIII	E1110K	1.17	i	C/N	a
αIII	R1114E/E1115R	0.91	wt	C	A
αIII	L1116A	0.73	wt	N	A
L III-IV	ΔP1120-P1127	1.17	wt	C	A
αIV	D1132A/K1135A	0.66	wt	C	A
αIV	K1141A/E1142A	0.76	wt	C/N	A
αIV	D1145A	1.00	wt	C/N	a

<sup>a</sup>GST-FABD: relative values for F-actin binding in cosedimentation assays.

<sup>b</sup>Bcr-Abl: wt, F-actin colocalization; i, intermediate; and p, punctuate staining (reduced binding).

<sup>c</sup>c-Abl: C, cytoplasmic; C/N, intermediate localization; N, nuclear; A, F-actin colocalization; and a, reduced F-actin colocalization.

ten lowest energy structures. Figures of 3D structures and molecular surface representations were prepared by using PyMol (DeLano, 2002).

#### NMR Titration and <sup>15</sup>N Relaxation Studies

For NMR titrations, <sup>1</sup>H,<sup>15</sup>N-HSQC experiments were recorded on a 500 MHz spectrometer fitted with a cryoprobe. <sup>15</sup>N-labeled FABD was dissolved in 5 mM Tris-HCl (pH 8.0), 50 mM KCl, 0.2 mM CaCl<sub>2</sub>, 1 mM ATP, 0.5 mM DTT, and 10% <sup>2</sup>H<sub>2</sub>O at an initial protein concentration of 30 μM. The sample was titrated at 295 K with freshly polymerized F-actin (Cytoskeleton, Inc.) prepared in the same buffer as described above, toward a final sample containing 1.2 molar equivalents F-actin (based on the monomeric actin concentration) equal to 15 μM and 18 μM of FABD and F-actin, respectively. Backbone amide <sup>15</sup>N relaxation parameters were acquired as described (Wiesner et al., 2005).

#### Protein-Protein Docking Simulations

Structural models of the Bcr-Abl/c-Abl FABD-actin complex were generated with the program HADDOCK by using standard protocols and parameters (Dominguez et al., 2003). A set of 20 ambiguous interaction restraints (AIRs) was defined for the Bcr-Abl/c-Abl FABD based on the results of the F-actin cosedimentation assays. Although we could not experimentally determine the actin residues involved in binding the FABD, analysis of crystal structures of actin complexes reveals a hydrophobic cleft between actin subdomains 1 and 3 as the primary interaction site for F-actin binding proteins (Dominguez, 2004; Aguda et al., 2005). Residues in and surrounding this hydrophobic cleft were chosen for docking from a superposition of known structures of actin complexes. This resulted in 19 AIRs for the actin monomer. An ensemble of seven NMR structures of the Bcr-Abl/c-Abl FABD and actin coordinates from four actin complexes (PDB-entries: 1EQY, 1SQK, 1QZ5, and 1KXP) were used as input structures for initial rigid-body docking. From the resulting 500 energy-minimized structures, 50 were chosen based on AIR restraint energy and subjected to semiflexible simulated annealing followed by a final refinement in explicit water. Among the final water-refined ten lowest energy structures, eight models adopted

conformations as shown in Figure S5 with a pair-wise backbone rmsd of 1.8 ± 0.8 Å, whereas the orientation of the FABD was inverted with respect to the orientation shown in Figure S5 in the remaining two models.

#### Binding Assays

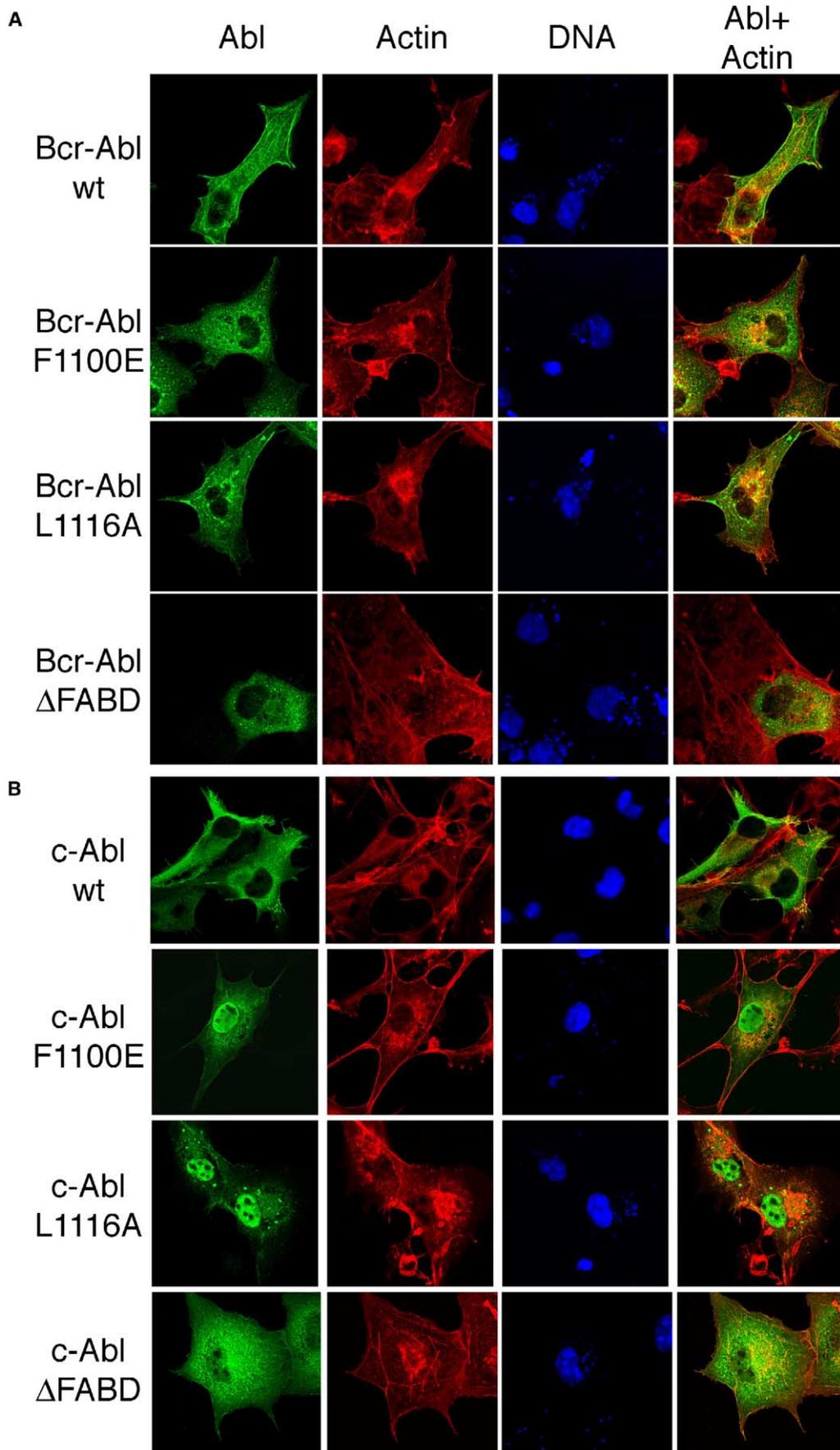
For the binding experiment of Figure 3A, zz-tagged FABD, NES, and the respective L1116A mutants were immobilized on IgG-Sepharose. The beads were incubated with HeLa cell extract in the presence of an energy-regenerating system (Kutay et al., 1997) and 30 μM cytochalasin B. A nuclear environment was mimicked by addition of GTPase-deficient Ran mutant (Ran[Q69L]GTP) (Klebe et al., 1995). Unbound material was removed, and the beads were washed with 10 ml of binding buffer. Bound proteins were eluted with SDS-sample buffer and analyzed by SDS-PAGE followed by Western blotting. Antibodies against CRM1, Exp5, and Exp6 have been described previously (Mingot et al., 2004; Stuken et al., 2003). For Figure 3B, recombinant FABD (3 μM), GFP-NES<sup>1109–1118</sup>, or GFP-NES L1116A (both 1 μM) were allowed to bind to immobilized human CRM1 in the presence or absence of Ran(Q69L)GTP. Bound proteins were eluted with 50 mM Tris (pH 7.5), 1.5 M MgCl<sub>2</sub>, isopropanol precipitated, separated by SDS-PAGE, and analyzed by Coomassie staining.

#### In Vitro Nuclear Export

*Xenopus laevis* egg extract and HeLa cell nuclei were prepared as described (Stuken et al., 2003). Nuclei were mixed with egg extract depleted of nuclear transport receptors and replenished with Ran, NTF2, and an energy-regenerating system (Ribbeck and Gorlich, 2002). GFP-FABD (residues 1039–1149), GFP-NES<sup>1109–1118</sup>, and mRFP-PK1αNES were used at 2 μM. Exportins were added after 20 min of nucleocytoplasmic equilibration at 1 μM.

#### Mammalian Expression Constructs

p210 Bcr-Abl was kindly provided by Martin Ruthardt (Johann Wolfgang Goethe-University, Frankfurt, Germany). pSGT vector and pSGT-Abl constructs were previously described (Barilá and Superti-Furga, 1998). The ΔFABD mutation was performed as pre-

Molecular Cell  
470

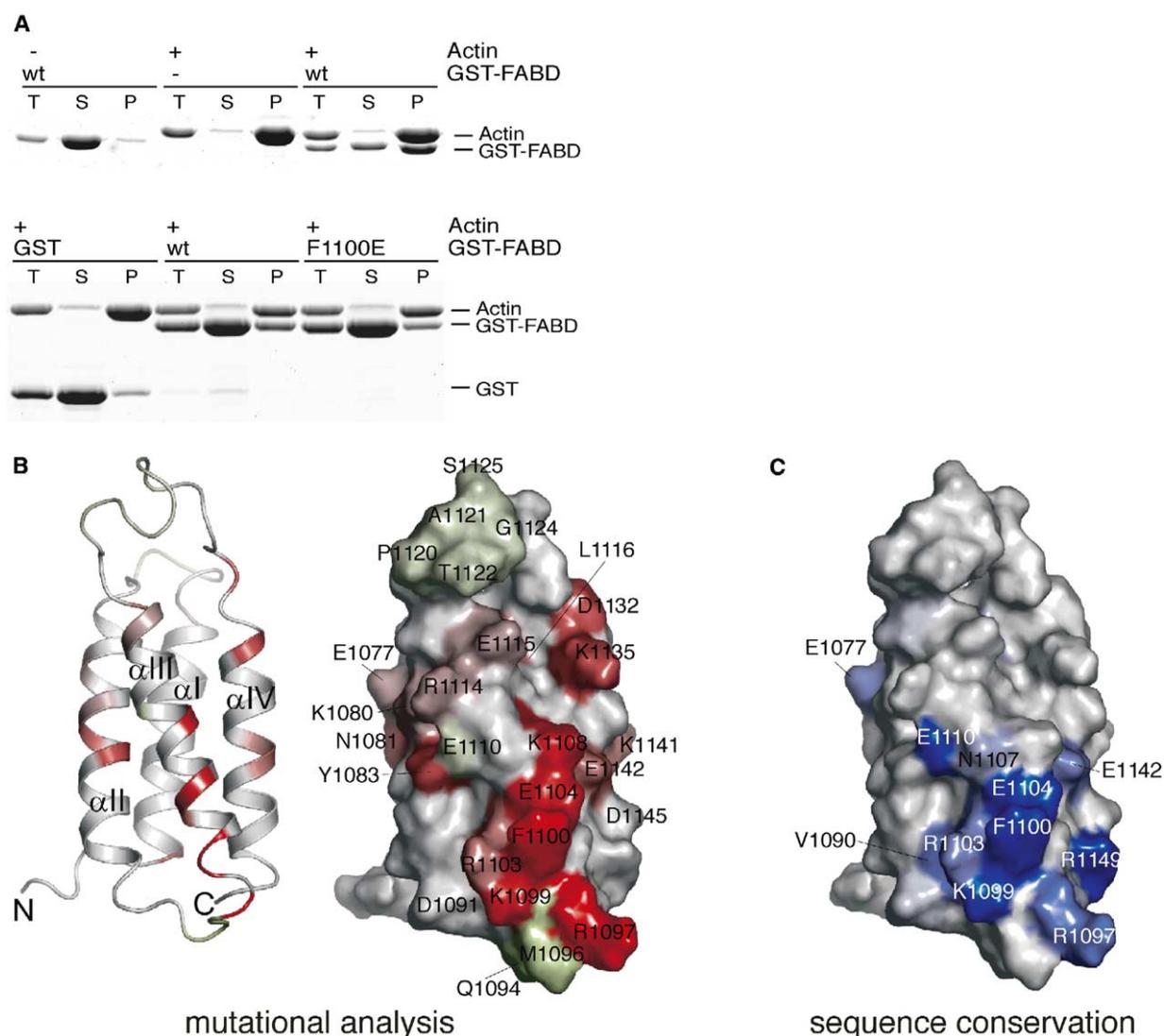
F-Actin Binding Domain of Bcr-Abl  
471

Figure 6. Mutational Analysis of the Bcr-Abl/c-Abl F-Actin Binding Domain

(A) F-actin cosedimentation assay: the indicated proteins were incubated for 30 min and spun at 135,000  $\times$  g, 10% of the total reaction before centrifugation (T) and equal amounts of the supernatant (S) and pellet (P) were resolved by SDS-PAGE, stained with Coomassie, and quantified by using the LI-COR Odyssey system.

(B) Ribbon and surface representation of the FABD highlighting effects of individual mutations on F-actin binding activity (see Table 2). Residues are colored with a linear gradient from red to white to green (red, 60% relative inhibition; green, 20% relative enhanced binding activity). All other residues are shown in white.

(C) Surface representation of the FABD highlighting sequence conservation. High to low sequence conservation is colored from blue to white, respectively. In all panels, the FABD is shown in the same orientation.

viously described (Woodring et al., 2001). All mutations were obtained by using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). All mutations were confirmed by sequencing.

**Immunofluorescence**

COS cells were transfected with either pSG-p210, pSGT-Abl, or mutants thereof by using PolyFect Transfection Reagent (Qiagen). Af-

ter 24–48 hr of incubation, cells were washed with PBS and fixed with 3% para-formaldehyde. Cells were permeabilized with 0.5% Triton X-100 and blocked with 2.5% normal goat serum (Jackson ImmunoResearch). To visualize Bcr-Abl/c-Abl, primary (rabbit anti-Abl, K12; Santa Cruz) and secondary (anti-rabbit AlexaFluor488; Molecular Probes) antibodies were diluted in blocking solution and incubated with the cells at room temperature. F-actin and nuclei

Figure 5. Localization of Bcr-Abl/c-Abl FABD Mutant Proteins

COS cells were transiently transfected with the indicated (A) Bcr-Abl or (B) c-Abl expression constructs. Cells were fixed and immunostained with anti-Abl antibody (Abl). F-actin (Actin) and nuclei (DNA) were stained with rhodamine-conjugated phalloidin and DRAQ5, respectively.

Molecular Cell  
472

were stained with rhodamine-conjugated phalloidin (Molecular Probes) and DRAQ5 (Biostatus Limited), respectively. Treatment with STI571/Gleevec was done for 4 hr at 20  $\mu$ M. Images were obtained on a Zeiss LSM510 confocal microscope by using a 40 $\times$ /1.3 Plan Neofluar objective.

#### F-Actin Cosedimentation Assays

GST-FABD fusion proteins were expressed in pETM30 and purified as described (Wiesner et al., 2005) with the exception that the TEV cleavage step was omitted. The fusion proteins were concentrated (15–25 mg/ml). Rabbit muscle actin (Cytoskeleton Inc.) was polymerized according to the manufacturer's instructions. Centrifugations were performed in a Beckman TLS-55 rotor at 135,000  $\times$  g. GST-FABD proteins were incubated with equimolar amounts of F-actin in a final volume of 50  $\mu$ l. 5  $\mu$ l of the reaction mixture was removed before ultracentrifugation ("total" sample). After ultracentrifugation, the supernatant was removed and the pellet resuspended in SDS-PAGE sample buffer. The "total" sample and equivalent amounts of supernatant and pellet (1/3 of the total reaction) were separated by SDS-PAGE followed by Coomassie staining and quantification using the LI-COR Odyssey system. Levels of F-actin binding in Table 2 were expressed relative to the wt protein.

#### Supplemental Data

Supplemental Data include Supplemental References and five figures and are available with this article online at <http://www.moleculer.org/cgi/content/full/19/4/461/DC1/>.

#### Acknowledgments

We thank the Frankfurt DFG/NMR 900 MHz center; the European Molecular Biology Laboratory (EMBL)/Heidelberg and Institute of Molecular Pathology/Vienna confocal microscopy facilities; G. Stier for advice; M. Way, K. Djinic-Carugo, and V. Small for reading of the manuscript; and K. Dorey for the *Xenopus* sequence. This work was supported by fellowships from the EMBL (O.H.), European Molecular Biology Organization (C.D.M.), German Research Foundation DFG (S.W.), Aventis (O.H.), Boehringer Ingelheim Fonds (T.G.), and the Leukemia and Lymphoma Society (L.L.R.R.). Work in the participating labs is funded by the EMBL, the Austrian Academy of Sciences, and the DFG.

Received: April 15, 2005

Revised: June 2, 2005

Accepted: June 22, 2005

Published: August 18, 2005

#### References

- Aguda, A.H., Burtnick, L.D., and Robinson, R.C. (2005). The state of the filament. *EMBO Rep.* 6, 220–226.
- Azam, M., Latek, R.R., and Daley, G.Q. (2003). Mechanisms of autoinhibition and STI-571/imatinib resistance revealed by mutagenesis of BCR-ABL. *Cell* 112, 831–843.
- Barilá, D., and Superti-Furga, G. (1998). An intramolecular SH3-domain interaction regulates c-Abl activity. *Nat. Genet.* 18, 280–282.
- Barilá, D., Rufini, A., Condo, I., Ventura, N., Dorey, K., Superti-Furga, G., and Testi, R. (2003). Caspase-dependent cleavage of c-Abl contributes to apoptosis. *Mol. Cell. Biol.* 23, 2790–2799.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 54, 905–921.
- Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* 120, 6836–6837.
- Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone

angle restraints from searching a database for protein chemical shifts and sequence homology. *J. Biomol. NMR* 13, 289–302.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System. <http://pymol.sourceforge.net>.

Dierov, J., Dierova, R., and Carroll, M. (2004). BCR/ABL translocates to the nucleus and disrupts an ATR-dependent intra-S phase checkpoint. *Cancer Cell* 5, 275–285.

Dominguez, C., Boelens, R., and Bonvin, A.M.J.J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125, 1731–1737.

Dominguez, R. (2004). Actin-binding proteins—a unifying hypothesis. *Trends Biochem. Sci.* 29, 572–578.

Druker, B.J. (2004). Imatinib as a paradigm of targeted therapies. *Adv. Cancer Res.* 91, 1–30.

Fillingham, I., Gingras, A.R., Papagrigoriou, E., Patel, B., Emsley, J., Critchley, D.R., Roberts, G.C., and Barsukov, I.L. (2005). A vinculin binding domain from the talin rod unfolds to form a complex with the vinculin head. *Structure (Camb)* 13, 65–74.

Galkin, V.E., Orlova, A., Koleske, A.J., and Egelman, E.H. (2005). The Arg non-receptor tyrosine kinase modifies F-actin structure. *J. Mol. Biol.* 346, 565–575.

Hantschel, O., and Superti-Furga, G. (2004). Regulation of the c-Abl and Bcr-Abl tyrosine kinases. *Nat. Rev. Mol. Cell Biol.* 5, 33–44.

Hantschel, O., Nagar, B., Guettler, S., Kretzschmar, J., Dorey, K., Kuriyan, J., and Superti-Furga, G. (2003). A myristoyl/phosphotyrosine switch regulates c-Abl. *Cell* 112, 845–857.

Harrison, S.C. (2003). Variation on an Src-like Theme. *Cell* 112, 737–740.

Hayashi, I., Vuori, K., and Liddington, R.C. (2002). The focal adhesion targeting (FAT) region of focal adhesion kinase is a four-helix bundle that binds paxillin. *Nat. Struct. Biol.* 9, 101–106.

Heisterkamp, N., Voncken, J.W., Senadheera, D., Gonzalez-Gomez, I., Reichert, A., Haataja, L., Reinikainen, A., Pattengale, P.K., and Groffen, J. (2000). Reduced oncogenicity of p190 Bcr/Abl F-actin-binding domain mutants. *Blood* 96, 2226–2232.

Henderson, B.R., and Eleftheriou, A. (2000). A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Exp. Cell Res.* 256, 213–224.

Holmes, K.C., Popp, D., Gebhard, W., and Kabsch, W. (1990). Atomic model of the actin filament. *Nature* 347, 44–49.

Hu, J., and Bax, A. (1997).  $\chi_1$  angle information from a simple two-dimensional NMR experiment that identifies trans  $^3J_{\text{NCY}}$  couplings in isotopically enriched proteins. *J. Biomol. NMR* 9, 323–328.

Kadlec, J., Izaurralde, E., and Cusack, S. (2004). The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat. Struct. Mol. Biol.* 11, 330–337.

Klebe, C., Bischoff, F.R., Ponstingl, H., and Wittinghofer, A. (1995). Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. *Biochemistry* 34, 639–647.

Kubinowa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994). Measurement of HN-H $\alpha$  J-couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J. Biomol. NMR* 4, 871–878.

Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R., and Gorlich, D. (1997). Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* 90, 1061–1071.

Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.

Lerche, M.H., Meissner, A., Poulsen, F.M., and Sorensen, O.W. (1999). Pulse sequences for measurement of one-bond (15N)-(1H) coupling constants in the protein backbone. *J. Magn. Reson.* 140, 259–263.

McWhirter, J.R., and Wang, J.Y. (1991). Activation of tyrosine kinase and microfilament-binding functions of c-abl by bcr sequences in bcr/abl fusion proteins. *Mol. Cell. Biol.* 11, 1553–1565.

McWhirter, J.R., and Wang, J.Y. (1993). An actin-binding function

F-Actin Binding Domain of Bcr-Abl  
473

- contributes to transformation by the Bcr-Abl oncoprotein of Philadelphia chromosome-positive human leukemias. *EMBO J.* 12, 1533–1546.
- McWhirter, J.R., Galasso, D.L., and Wang, J.Y. (1993). A coiled-coil oligomerization domain of Bcr is essential for the transforming function of Bcr-Abl oncoproteins. *Mol. Cell. Biol.* 13, 7587–7595.
- Mingot, J.M., Bohnsack, M.T., Jakle, U., and Gorlich, D. (2004). Exportin 7 defines a novel general nuclear export pathway. *EMBO J.* 23, 3227–3236.
- Nagar, B., Hantschel, O., Young, M.A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G., and Kuriyan, J. (2003). Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* 112, 859–871.
- Nilges, M., and O'Donoghue, S.I. (1998). Ambiguous NOEs and automated NOE assignment. *Progr. NMR Spectr.* 32, 107–139.
- Puntervoll, P., Linding, R., Gemund, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D.M., Ausiello, G., Brannetti, B., Costantini, A., et al. (2003). ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res.* 31, 3625–3630.
- Raitano, A.B., Whang, Y.E., and Sawyers, C.L. (1997). Signal transduction by wild-type and leukemogenic Abl proteins. *Biochim. Biophys. Acta* 1333, F201–F216.
- Renshaw, M.W., Capozza, M.A., and Wang, J.Y. (1988). Differential expression of type-specific c-abl mRNAs in mouse tissues and cell lines. *Mol. Cell. Biol.* 8, 4547–4551.
- Ribbeck, K., and Gorlich, D. (2002). The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* 21, 2664–2671.
- Ruckert, M., and Otting, G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* 122, 7793–7797.
- Sawyers, C. (2004). Targeted cancer therapy. *Nature* 432, 294–297.
- Smith, K.M., Yacobi, R., and Van Etten, R.A. (2003). Autoinhibition of Bcr-Abl through its SH3 domain. *Mol. Cell* 12, 27–37.
- Steelman, L.S., Pohnert, S.C., Shelton, J.G., Franklin, R.A., Bertrand, F.E., and McCubrey, J.A. (2004). JAK/STAT, Raf/MEK/ERK, PI3K/Akt and BCR-ABL in cell cycle progression and leukemogenesis. *Leukemia* 18, 189–218.
- Stuven, T., Hartmann, E., and Gorlich, D. (2003). Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. *EMBO J.* 22, 5928–5940.
- Taagepera, S., McDonald, D., Loeb, J.E., Whitaker, L.L., McElroy, A.K., Wang, J.Y., and Hope, T.J. (1998). Nuclear-cytoplasmic shuttling of C-ABL tyrosine kinase. *Proc. Natl. Acad. Sci. USA* 95, 7457–7462.
- Van Etten, R.A., Jackson, P., and Baltimore, D. (1989). The mouse type IV c-abl gene product is a nuclear protein, and activation of transforming ability is associated with cytoplasmic localization. *Cell* 58, 669–678.
- Van Etten, R.A., Jackson, P.K., Baltimore, D., Sanders, M.C., Matsudaira, P.T., and Janney, P.A. (1994). The COOH terminus of the c-Abl tyrosine kinase contains distinct F- and G-actin binding domains with bundling activity. *J. Cell Biol.* 124, 325–340.
- Vigneri, P., and Wang, J.Y. (2001). Induction of apoptosis in chronic myelogenous leukemia cells through nuclear entrapment of BCR-ABL tyrosine kinase. *Nat. Med.* 7, 228–234.
- Wang, J.Y. (2000). Regulation of cell death by the abl tyrosine kinase. *Oncogene* 19, 5643–5650.
- Wen, S.T., Jackson, P.K., and Van Etten, R.A. (1996). The cytostatic function of c-Abl is controlled by multiple nuclear localization signals and requires the p53 and Rb tumor suppressor gene products. *EMBO J.* 15, 1583–1595.
- Wetzler, M., Talpaz, M., Van Etten, R.A., Hirsh-Ginsberg, C., Beran, M., and Kurzrock, R. (1993). Subcellular localization of Bcr, Abl, and Bcr-Abl proteins in normal and leukemic cells and correlation of expression with myeloid differentiation. *J. Clin. Invest.* 92, 1925–1939.
- Wiesner, S., Hantschel, O., Mackereth, C.D., Superti-Furga, G., and Sattler, M. (2005). NMR assignment reveals an all  $\alpha$ -helical fold for the F-actin binding domain of human Bcr-Abl/c-Abl. *J. Biomol. NMR*, in press.
- Wong, S., and Witte, O.N. (2004). The BCR-ABL story: bench to bedside and back. *Annu. Rev. Immunol.* 22, 247–306.
- Woodring, P.J., Hunter, T., and Wang, J.Y. (2001). Inhibition of c-Abl tyrosine kinase activity by filamentous actin. *J. Biol. Chem.* 276, 27104–27110.
- Woodring, P.J., Hunter, T., and Wang, J.Y. (2003). Regulation of F-actin-dependent processes by the Abl family of tyrosine kinases. *J. Cell Sci.* 116, 2613–2626.
- Yoshida, K., Yamaguchi, T., Natsume, T., Kufe, D., and Miki, Y. (2005). JNK phosphorylation of 14–3–3 proteins regulates nuclear targeting of c-Abl in the apoptotic response to DNA damage. *Nat. Cell Biol.* 7, 278–285.
- Zhang, X., Subrahmanyam, R., Wong, R., Gross, A.W., and Ren, R. (2001). The NH(2)-terminal coiled-coil domain and tyrosine 177 play important roles in induction of a myeloproliferative disease in mice by Bcr-Abl. *Mol. Cell. Biol.* 21, 840–853.
- Zhu, J., and Wang, J.Y. (2004). Death by Abl: a matter of location. *Curr. Top. Dev. Biol.* 59, 165–192.

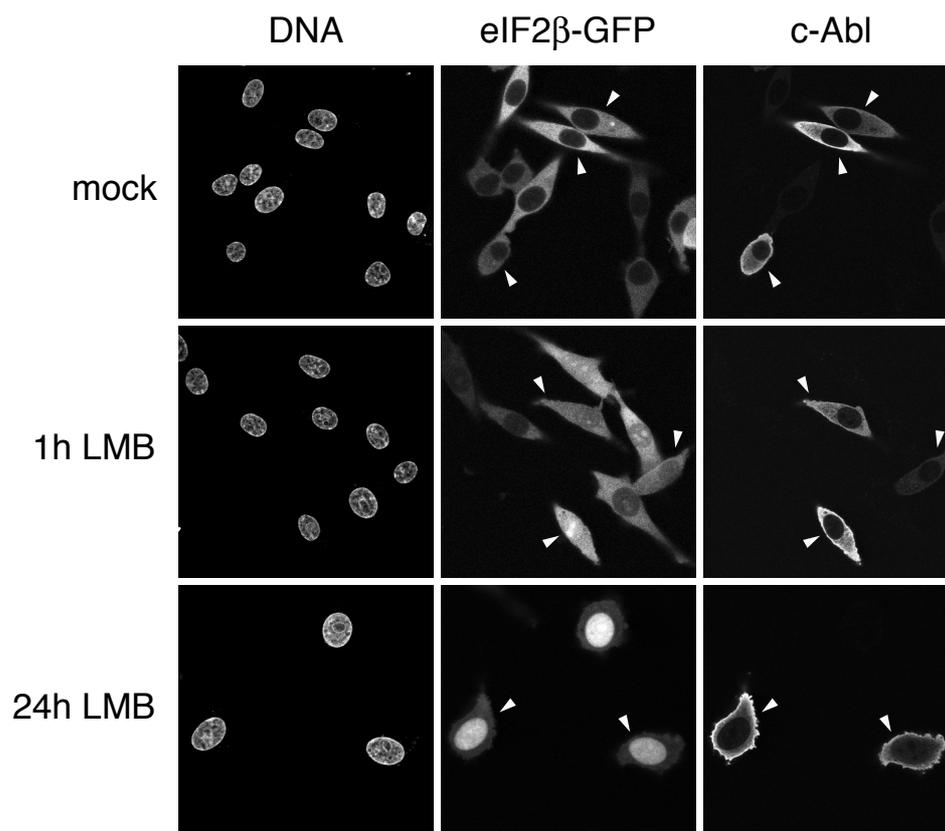
## Accession Numbers

Atomic coordinates have been deposited in the Protein Data Bank under accession number 1ZZP.

### 6.5.2 Additional data to Hantschel *et al.*, 2005

The following experiment was performed to further test if full-length c-Abl is a CRM1 substrate. We used *BHK* (*baby hamster kidney*) cells stably expressing eIF2 $\beta$ -GFP – a known substrate of CRM1 (Bohnsack *et al.*, 2002). We transfected cells with an expression construct for c-Abl, treated them with the CRM1 inhibitor *leptomycin B* (*LMB*) and recorded the nucleocytoplasmic distribution of eIF2 $\beta$ -GFP and c-Abl by microscopy of GFP-fluorescence or indirect immunofluorescence, respectively. As seen with mock-treated cells, both c-Abl and eIF2 $\beta$ -GFP are predominantly cytoplasmic (**Figure 6-6**). Treatment with LMB for one hour did not change the nucleocytoplasmic distribution of c-Abl, whereas the CRM1 cargo eIF2 $\beta$ -GFP clearly shifted to the nucleus. Even extended incubation with LMB did not lead to prominent nuclear accumulation of c-Abl and only a faint nuclear signal was evident after 24 hours. This faint nuclear signal, however, is likely to represent a secondary effect of LMB and is possibly due to the general cytotoxicity of LMB after such prolonged incubation times. LMB selectively inactivates CRM1 but the consequences of this inactivation are ultimately lethal (Nishi *et al.*, 1994) and thus pleiotropic (discussed in Mingot *et al.*, 2004). The morphological changes of cells treated with LMB for 24 hours exemplify such pleiotropic effect.

These results are in perfect agreement with our findings that: (1) the previously postulated leucine-rich nuclear export signal is buried in the hydrophobic core of the F-actin binding domain (FABD), (2) this region does not display mobility in solution, excluding flipping or unfolding that would expose hydrophobic residues (**Figure S1** online), (3) the FABD does not bind to CRM1 (**Figure 3A, B**), (4) CRM1 does not export the FABD from the nucleus (**Figure 3C** and data not shown). In summary, we found no evidence for CRM1-mediated nuclear export of c-Abl.



**Figure 6-6: Block of CRM1 does not increase nuclear levels of c-Abl.**

Baby Hamster Kidney (BHK) cells stably expressing eIF2 $\beta$ -GFP were transfected with a c-Abl expression construct and incubated with leptomycin B (LMB, 5 ng/ml) for the time indicated. Mock refers to cells treated with ethanol only. Nucleocytoplasmic distribution was analyzed by laser scanning confocal microscopy. Arrowheads mark transfected cells.

### 6.5.3 Additional methods

BHK cells stably transduced with an eIF2 $\beta$ -GFP construct (Bohnsack *et al.*, 2002) were transfected with pSGT-Abl using FuGENE6 reagent (Roche). 10h post-transfection, eIF2 $\beta$ -GFP expression was induced with 50 ng/ml doxycycline. After incubation for 12 more hours, cells were treated with 5 ng/ml LMB or ethanol only (the solvent of LMB) for 1h. For a subset of cells, LMB treatment was continued to 24 h total. Cells were washed with PBS and fixed with 3% paraformaldehyde (PFA) for 5 min at 37 °C. PFA was quenched with 50 mM NH<sub>4</sub>Cl. The cells were then permeabilized with 0.5% Triton-X100 for 5 min and blocked with 2.5% normal goat serum. To visualize c-Abl, primary (rabbit anti-Abl, K12; Santa Cruz) and secondary (anti-rabbit AlexaFluor647; Molecular Probes) antibodies were diluted in blocking solution and incubated with the cells at room temperature. Nuclei were stained with DAPI. We ensured that the immunostaining protocol allowed for antibody penetration of the nucleus (data not shown).

## References

- Adachi, Y. & Yanagida, M. (1989) Higher Order Chromosome Structure Is Affected by Cold-Sensitive Mutations in a *Schizosaccharomyces pombe* Gene *crml1sup+/sup* Which Encodes a 115-kD Protein Preferentially Localized in the Nucleus and at Its Periphery. *J Cell Biol* **108**, 1195-1207.
- Adam, E. J. & Adam, S. A. (1994) Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope. *J Cell Biol* **125**, 547-555.
- Adam, S. A., Marr, R. S. & Gerace, L. (1990) Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J Cell Biol* **111**, 807-816.
- Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954.
- Andrade, M. A. & Bork, P. (1995) HEAT repeats in the Huntington's disease protein. *Nat Genet* **11**, 115-116.
- Andrade, M. A., Petosa, C., O'Donoghue, S. I., Müller, C. W. & Bork, P. (2001) Comparison of ARM and HEAT protein repeats. *J Mol Biol* **309**, 1-18.
- Arnold, M., Nath, A., Wohlwend, D. & Kehlenbach, R. H. (2006) Transportin is a major nuclear import receptor for c-Fos: a novel mode of cargo interaction. *J Biol Chem* **281**, 5492-5499.
- Arts, G. J., Fornerod, M. & Mattaj, I. W. (1998) Identification of a nuclear export receptor for tRNA. *Curr Biol* **8**, 305-314.
- Askjaer, P., Bachi, A., Wilm, M., Bischoff, F. R., Weeks, D. L., Ogniewski, V., Ohno, M., Niehrs, C., Kjems, J., Mattaj, I. W. & Fornerod, M. (1999) RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Mol Cell Biol* **19**, 6276-6285.
- Baake, M., Doenecke, D. & Albig, W. (2001) Characterisation of nuclear localisation signals of the four human core histones. *J Cell Biochem* **81**, 333-346.
- Bayliss, R., Littlewood, T. & Stewart, M. (2000) Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* **102**, 99-108.
- Bermel, W., Bertini, I., Felli, I. C., Piccioli, M. & Pierattelli, R. (2006) C-13-detected protonless NMR spectroscopy of proteins in solution. *Prog Nucl Mag Res Sp* **48**, 25-45.
- Bischoff, F. R. & Görlich, D. (1997) RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors. *FEBS Lett* **419**, 249-254.
- Bischoff, F. R., Klebe, C., Kretschmer, J., Wittinghofer, A. & Ponstingl, H. (1994) RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proc Natl Acad Sci USA* **91**, 2587-2591.
- Bischoff, F. R., Krebber, H., Smirnova, E., Dong, W. & Ponstingl, H. (1995) Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J* **14**, 705-715.
- Bischoff, F. R. & Ponstingl, H. (1991a) Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide. *Proc Natl Acad Sci USA* **88**, 10830-10834.
- Bischoff, F. R. & Ponstingl, H. (1991b) Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* **354**, 80-82.
- Bogerd, H. P., Fridell, R. A., Benson, R. E., Hua, J. & Cullen, B. R. (1996) Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vivo randomization-selection assay. *Mol Cell Biol* **16**, 4207-4214.
- Bohnsack, M. T., Czaplinski, K. & Gorlich, D. (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185-191.

- Bohnsack, M. T., Regener, K., Schwappach, B., Saffrich, R., Paraskeva, E., Hartmann, E. & Görlich, D. (2002) Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO J* **21**, 6205-6215.
- Bonner, W. M. (1975) Protein migration into nuclei. I. Frog oocyte nuclei in vivo accumulate microinjected histones, allow entry to small proteins, and exclude large proteins. *J Cell Biol* **64**, 421-430.
- Bono, F., Cook, A. G., Grünwald, M., Ebert, J. & Conti, E. (2010) Nuclear import mechanism of the EJC component Mago-Y14 revealed by structural studies of importin 13. *Mol Cell* **37**, 211-222.
- Brohawn, S. G., Partridge, J. R., Whittle, J. R. & Schwartz, T. U. (2009) The nuclear pore complex has entered the atomic age. *Structure* **17**, 1156-1168.
- Brownawell, A. M. & Macara, I. G. (2002) Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J Cell Biol* **156**, 53-64.
- Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**, 2728-2733.
- Bullock, T. L., Clarkson, W. D., Kent, H. M. & Stewart, M. (1996) The 1.6 angstroms resolution crystal structure of nuclear transport factor 2 (NTF2). *J Mol Biol* **260**, 422-431.
- Calado, A., Treichel, N., Müller, E. C., Otto, A. & Kutay, U. (2002) Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *EMBO J* **21**, 6216-6224.
- Cansizoglu, A. E. & Chook, Y. M. (2007) Conformational heterogeneity of karyopherin beta2 is segmental. *Structure* **15**, 1431-1441.
- Catimel, B., Teh, T., Fontes, M. R., Jennings, I. G., Jans, D. A., Howlett, G. J., Nice, E. C. & Kobe, B. (2001) Biophysical characterization of interactions involving importin-alpha during nuclear import. *J Biol Chem* **276**, 34189-34198.
- Chi, N. C., Adam, E. J. & Adam, S. A. (1995) Sequence and characterization of cytoplasmic nuclear protein import factor p97. *J Cell Biol* **130**, 265-274.
- Chook, Y. M. & Blobel, G. (1999) Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. *Nature* **399**, 230-237.
- Chuderland, D., Konson, A. & Seger, R. (2008) Identification and characterization of a general nuclear translocation signal in signaling proteins. *Mol Cell* **31**, 850-861.
- Cingolani, G., Bednenko, J., Gillespie, M. T. & Gerace, L. (2002) Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta. *Mol Cell* **10**, 1345-1353.
- Cingolani, G., Petosa, C., Weis, K. & Müller, C. W. (1999) Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* **399**, 221-229.
- Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763.
- Conti, E., Müller, C. W. & Stewart, M. (2006) Karyopherin flexibility in nucleocytoplasmic transport. *Curr Opin Struct Biol* **16**, 237-244.
- Cook, A., Bono, F., Jinek, M. & Conti, E. (2007) Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* **76**, 647-671.
- Cook, A., Fernandez, E., Lindner, D., Ebert, J., Schlenstedt, G. & Conti, E. (2005) The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding. *Mol Cell* **18**, 355-367.
- Cook, A. G. & Conti, E. (2010) Nuclear export complexes in the frame. *Curr Opin Struct Biol* **20**, 247-252.
- Cook, A. G., Fukuhara, N., Jinek, M. & Conti, E. (2009) Structures of the tRNA export factor in the nuclear and cytosolic states. *Nature* **461**, 60-65.
- Crespi, H. L., Archer, S. M. & Katz, J. J. (1959) Culture of algae and other micro-organisms in deuterium oxide. *Nature* **184**, 729-730.

- Custódio, N., Carmo-Fonseca, M., Geraghty, F., Pereira, H. S., Grosveld, F. & Antoniou, M. (1999) Inefficient processing impairs release of RNA from the site of transcription. *EMBO J* **18**, 2855-2866.
- Daneholt, B. (1997) A look at messenger RNP moving through the nuclear pore. *Cell* **88**, 585-588.
- De Duve, C., Gianetto, R., Appelmans, F. & Wattiaux, R. (1953) Enzymic content of the mitochondria fraction. *Nature* **172**, 1143-1144.
- Dean, K. A., von Ahsen, O., Görlich, D. & Fried, H. M. (2001) Signal recognition particle protein 19 is imported into the nucleus by importin 8 (RanBP8) and transportin. *J Cell Sci* **114**, 3479-3485.
- Deane, R., Schäfer, W., Zimmermann, H. P., Mueller, L., Görlich, D., Prehn, S., Ponstingl, H. & Bischoff, F. R. (1997) Ran-binding protein 5 (RanBP5) is related to the nuclear transport factor importin-beta but interacts differently with RanBP1. *Mol Cell Biol* **17**, 5087-5096.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293.
- Dong, X., Biswas, A. & Chook, Y. M. (2009a) Structural basis for assembly and disassembly of the CRM1 nuclear export complex. *Nat Struct Mol Biol* **16**, 558-560.
- Dong, X., Biswas, A., Süel, K. E., Jackson, L. K., Martinez, R., Gu, H. & Chook, Y. M. (2009b) Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* **458**, 1136-1141.
- Drivas, G. T., Shih, A., Coutavas, E., Rush, M. G. & D'Eustachio, P. (1990) Characterization of four novel ras-like genes expressed in a human teratocarcinoma cell line. *Mol Cell Biol* **10**, 1793-1798.
- Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.
- Engelsma, D., Bernad, R., Calafat, J. & Fornerod, M. (2004) Supraphysiological nuclear export signals bind CRM1 independently of RanGTP and arrest at Nup358. *EMBO J* **23**, 3643-3652.
- Engelsma, D., Valle, N., Fish, A., Salomé, N., Almendral, J. M. & Fornerod, M. (2008) A Supraphysiological Nuclear Export Signal is Required for Parvovirus Nuclear Export. *Mol Biol Cell* **19**, 2544-2552.
- Englmeier, L., Olivo, J. C. & Mattaj, I. W. (1999) Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis. *Curr Biol* **9**, 30-41.
- Fantozzi, D. A., Harootunian, A. T., Wen, W., Taylor, S. S., Feramisco, J. R., Tsien, R. Y. & Meinkoth, J. L. (1994) Thermostable inhibitor of cAMP-dependent protein kinase enhances the rate of export of the kinase catalytic subunit from the nucleus. *J Biol Chem* **269**, 2676-2686.
- Feldherr, C. M., Kallenbach, E. & Schultz, N. (1984) Movement of a karyophilic protein through the nuclear pores of oocytes. *J Cell Biol* **99**, 2216-2222.
- Fischer, M. W. F., Zeng, L. & Zuiderweg, E. R. P. (1996) Use of C-13-C-13 NOE for the assignment of NMR lines of larger labeled proteins at larger magnetic fields. *J Am Chem Soc* **118**, 12457-12458.
- Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. & Lührmann, R. (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475-483.
- Flaumenhaft, E., Bose, S., Crespi, H. L. & Katz, J. J. (1965) Deuterium isotope effects in cytology. *Int Rev Cytol* **18**, 313-361.
- Floer, M. & Blobel, G. (1996) The nuclear transport factor karyopherin beta binds stoichiometrically to Ran-GTP and inhibits the Ran GTPase activating protein. *J Biol Chem* **271**, 5313-5316.
- Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I. W. (1997a) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051-1060.
- Fornerod, M., van Deursen, J., van Baal, S., Reynolds, A., Davis, D., Murti, K. G., Fransen, J. & Grosveld, G. (1997b) The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88. *EMBO J* **16**, 807-816.
- Franke, W. W. & Scheer, U. (1974) Pathways of nucleocytoplasmic translocation of ribonucleoproteins. *Symp Soc Exp Biol* **28**, 249-282.

- Frey, S. & Görlich, D. (2009) FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties. *EMBO J* **28**, 2554-2567.
- Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M. & Nishida, E. (1997) CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390**, 308-311.
- Fukuhara, N., Fernandez, E., Ebert, J., Conti, E. & Svergun, D. (2004) Conformational variability of nucleocytoplasmic transport factors. *J Biol Chem* **279**, 2176-2181.
- Gadal, O., Strauss, D., Kessel, J., Trumppower, B., Tollervey, D. & Hurt, E. (2001) Nuclear export of 60s ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. *Mol Cell Biol* **21**, 3405-3415.
- García de la Torre, J., Huertas, M. L. & Carrasco, B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson* **147**, 138-146.
- Gardner, K. H. & Kay, L. E. (1997) Production and incorporation of N-15, C-13, H-2 (H-1-delta 1 methyl) isoleucine into proteins for multidimensional NMR studies. *J Am Chem Soc* **119**, 7599-7600.
- Gardner, K. H. & Kay, L. E. (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* **27**, 357-406.
- Gilchrist, D., Mykytka, B. & Rexach, M. (2002) Accelerating the rate of disassembly of karyopherin.cargo complexes. *J Biol Chem* **277**, 18161-18172.
- Gilchrist, D. & Rexach, M. (2003) Molecular basis for the rapid dissociation of nuclear localization signals from karyopherin alpha in the nucleoplasm. *J Biol Chem* **278**, 51937-51949.
- Goldfarb, D. S., Corbett, A. H., Mason, D. A., Harreman, M. T. & Adam, S. A. (2004) Importin alpha: a multipurpose nuclear-transport receptor. *Trends Cell Biol* **14**, 505-514.
- Gontan, C., Güttler, T., Engelen, E., Demmers, J., Fornerod, M., Grosveld, F. G., Tibboel, D., Görlich, D., Poot, R. A. & Rottier, R. J. (2009) Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *J Cell Biol* **185**, 27-34.
- Görlich, D., Dabrowski, M., Bischoff, F. R., Kutay, U., Bork, P., Hartmann, E., Prehn, S. & Izaurralde, E. (1997) A novel class of RanGTP binding proteins. *J Cell Biol* **138**, 65-80.
- Görlich, D., Henklein, P., Laskey, R. A. & Hartmann, E. (1996) A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus. *EMBO J* **15**, 1810-1817.
- Görlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R. A., Hartmann, E. & Prehn, S. (1995) Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Curr Biol* **5**, 383-392.
- Görlich, D. & Kutay, U. (1999) Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* **15**, 607-660.
- Görlich, D., Panté, N., Kutay, U., Aebi, U. & Bischoff, F. R. (1996) Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J* **15**, 5584-5594.
- Görlich, D., Prehn, S., Laskey, R. A. & Hartmann, E. (1994) Isolation of a protein that is essential for the first step of nuclear protein import. *Cell* **79**, 767-778.
- Görlich, D., Seewald, M. J. & Ribbeck, K. (2003) Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J* **22**, 1088-1100.
- Gurdon, J. B. (1970) Nuclear Transplantation and the Control of Gene Activity in Animal Development. *Proceedings of the Royal Society of London. Series B Biological Sciences* **176**, 303-314.
- Güttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M. & Görlich, D. (2010) NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* **17**, 1367-1376.
- Hantschel, O., Wiesner, S., Güttler, T., Mackereth, C. D., Rix, L. L., Mikes, Z., Dehne, J., Görlich, D., Sattler, M. & Superti-Furga, G. (2005) Structural basis for the cytoskeletal association of Bcr-Abl/c-Abl. *Mol Cell* **19**, 461-473.

- Harreman, M. T., Hodel, M. R., Fanara, P., Hodel, A. E. & Corbett, A. H. (2003) The auto-inhibitory function of importin alpha is essential in vivo. *J Biol Chem* **278**, 5854-5863.
- Hieda, M., Tachibana, T., Yokoya, F., Kose, S., Imamoto, N. & Yoneda, Y. (1999) A monoclonal antibody to the COOH-terminal acidic portion of Ran inhibits both the recycling of Ran and nuclear protein import in living cells. *J Cell Biol* **144**, 645-655.
- Ho, J. H., Kallstrom, G. & Johnson, A. W. (2000) Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *J Cell Biol* **151**, 1057-1066.
- Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M. & Lührmann, R. (1998) Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J* **17**, 4114-4126.
- Hutchinson, E. G. & Thornton, J. M. (1996) PROMOTIF--a program to identify and analyze structural motifs in proteins. *Protein Sci* **5**, 212-220.
- Hutten, S. & Kehlenbach, R. H. (2007) CRM1-mediated nuclear export: to the pore and beyond. *Trends in Cell Biology* **17**, 193-201.
- Imamoto, N., Shimamoto, T., Kose, S., Takao, T., Tachibana, T., Matsubae, M., Sekimoto, T., Shimonishi, Y. & Yoneda, Y. (1995) The nuclear pore-targeting complex binds to nuclear pores after association with a karyophile. *FEBS Lett* **368**, 415-419.
- Imasaki, T., Shimizu, T., Hashimoto, H., Hidaka, Y., Kose, S., Imamoto, N., Yamada, M. & Sato, M. (2007) Structural basis for substrate recognition and dissociation by human transportin 1. *Mol Cell* **28**, 57-67.
- Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I. W. & Görlich, D. (1997) The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J* **16**, 6535-6547.
- Jäkel, S., Albig, W., Kutay, U., Bischoff, F. R., Schwamborn, K., Doenecke, D. & Görlich, D. (1999) The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *EMBO J* **18**, 2411-2423.
- Jäkel, S. & Görlich, D. (1998) Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO J* **17**, 4491-4502.
- Jäkel, S., Mingot, J. M., Schwarzmaier, P., Hartmann, E. & Görlich, D. (2002) Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J* **21**, 377-386.
- Johnson, B. A. & Blevins, R. A. (1994) NMR view - a computer-program for the visualization and analysis of NMR data. *J Biomol NMR* **4**, 603-614.
- Jullien, D., Görlich, D., Laemmli, U. K. & Adachi, Y. (1999) Nuclear import of RPA in *Xenopus* egg extracts requires a novel protein XRIPalpha but not importin alpha. *EMBO J* **18**, 4348-4358.
- Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* **26**, 795-800.
- Kadlec, J., Izaurralde, E. & Cusack, S. (2004) The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat Struct Mol Biol* **11**, 330-337.
- Kaffman, A. & O'Shea, E. K. (1999) Regulation of nuclear localization: a key to a door. *Annu Rev Cell Dev Biol* **15**, 291-339.
- Kaffman, A., Rank, N. M., O'Neill, E. M., Huang, L. S. & O'Shea, E. K. (1998) The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature* **396**, 482-486.
- Kahle, J., Baake, M., Doenecke, D. & Albig, W. (2005) Subunits of the heterotrimeric transcription factor NF-Y are imported into the nucleus by distinct pathways involving importin beta and importin 13. *Mol Cell Biol* **25**, 5339-5354.
- Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. (1984) A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499-509.
- Kataoka, N., Bachorik, J. L. & Dreyfuss, G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. *J Cell Biol* **145**, 1145-1152.

- Katz, J. J. (1960) Chemical and biological studies with Deuterium. *American Scientist* **48**, 544-580.
- Kiseleva, E., Goldberg, M. W., Allen, T. D. & Akey, C. W. (1998) Active nuclear pore complexes in *Chironomus*: visualization of transporter configurations related to mRNP export. *J Cell Sci* **111**, 223-236.
- Klebe, C., Bischoff, F. R., Ponstingl, H. & Wittinghofer, A. (1995) Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. *Biochemistry* **34**, 639-647.
- Kobe, B. (1999) Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nat Struct Biol* **6**, 388-397.
- Köhler, M., Ansieau, S., Prehn, S., Leutz, A., Haller, H. & Hartmann, E. (1997) Cloning of two novel human importin-alpha subunits and analysis of the expression pattern of the importin-alpha protein family. *FEBS Lett* **417**, 104-108.
- Köhler, M., Speck, C., Christiansen, M., Bischoff, F. R., Prehn, S., Haller, H., Görlich, D. & Hartmann, E. (1999) Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. *Mol Cell Biol* **19**, 7782-7791.
- Kose, S., Imamoto, N., Tachibana, T., Shimamoto, T. & Yoneda, Y. (1997) Ran-unassisted nuclear migration of a 97-kD component of nuclear pore-targeting complex. *J Cell Biol* **139**, 841-849.
- Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. (2008) Nuclear Export Signal Consensus Sequences Defined Using a Localization-Based Yeast Selection System. *Traffic* **9**, 2053-2062.
- Koyama, M. & Matsuura, Y. (2010) An allosteric mechanism to displace nuclear export cargo from CRM1 and RanGTP by RanBP1. *EMBO J* **29**, 2002-2013.
- Kudo, N., Matsumori, N., Taoka, H., Fujiwara, D., Schreiner, E. P., Wolff, B., Yoshida, M. & Horinouchi, S. (1999) Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc Natl Acad Sci USA* **96**, 9112-9117.
- Kurisaki, A., Kurisaki, K., Kowanez, M., Sugino, H., Yoneda, Y., Heldin, C. H. & Moustakas, A. (2006) The mechanism of nuclear export of Smad3 involves exportin 4 and Ran. *Mol Cell Biol* **26**, 1318-1332.
- Kutay, U., Bischoff, F. R., Kostka, S., Kraft, R. & Görlich, D. (1997) Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* **90**, 1061-1071.
- Kutay, U. & Güttinger, S. (2005) Leucine-rich nuclear-export signals: born to be weak. *Trends in Cell Biology* **15**, 121-124.
- Kutay, U., Lipowsky, G., Izaurralde, E., Bischoff, F. R., Schwarzmaier, P., Hartmann, E. & Görlich, D. (1998) Identification of a tRNA-specific nuclear export receptor. *Mol Cell* **1**, 359-369.
- la Cour, T., Kiemer, L., Mølgaard, A., Gupta, R., Skriver, K. & Brunak, S. (2004) Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel* **17**, 527-536.
- Lai, M. C., Lin, R. I., Huang, S. Y., Tsai, C. W. & Tarn, W. Y. (2000) A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *J Biol Chem* **275**, 7950-7957.
- Lam, M. H., Briggs, L. J., Hu, W., Martin, T. J., Gillespie, M. T. & Jans, D. A. (1999) Importin beta recognizes parathyroid hormone-related protein with high affinity and mediates its nuclear import in the absence of importin alpha. *J Biol Chem* **274**, 7391-7398.
- Lanford, R. E. & Butel, J. S. (1984) Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37**, 801-813.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486.
- Lee, B. J., Cansizoglu, A. E., Süel, K. E., Louis, T. H., Zhang, Z. & Chook, Y. M. (2006) Rules for nuclear localization sequence recognition by karyopherin beta 2. *Cell* **126**, 543-558.

- Lee, S. J., Matsuura, Y., Liu, S. M. & Stewart, M. (2005) Structural basis for nuclear import complex dissociation by RanGTP. *Nature* **435**, 693-696.
- Lee, S. J., Sekimoto, T., Yamashita, E., Nagoshi, E., Nakagawa, A., Imamoto, N., Yoshimura, M., Sakai, H., Chong, K. T., Tsukihara, T. & Yoneda, Y. (2003) The structure of importin-beta bound to SREBP-2: nuclear import of a transcription factor. *Science* **302**, 1571-1575.
- Legrain, P. & Rosbash, M. (1989) Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell* **57**, 573-583.
- Leighton, F., Poole, B., Beaufay, H., Baudhuin, P., Coffey, J. W., Fowler, S. & De Duve, C. (1968) The large-scale separation of peroxisomes, mitochondria, and lysosomes from the livers of rats injected with triton WR-1339. Improved isolation procedures, automated analysis, biochemical and morphological properties of fractions. *J Cell Biol* **37**, 482-513.
- Leno, G. H. & Laskey, R. A. (1991) DNA replication in cell-free extracts from *Xenopus laevis*. *Methods Cell Biol* **36**, 561-579.
- Linge, J. P., Habeck, M., Rieping, W. & Nilges, M. (2003) ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315-316.
- Linge, J. P., O'Donoghue, S. I. & Nilges, M. (2001) Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Meth Enzymol* **339**, 71-90.
- Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M. & Nilges, M. (2003) Refinement of protein structures in explicit solvent. *Proteins* **50**, 496-506.
- Lipowsky, G., Bischoff, F. R., Schwarzmaier, P., Kraft, R., Kostka, S., Hartmann, E., Kutay, U. & Görlich, D. (2000) Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *EMBO J* **19**, 4362-4371.
- Lounsbury, K. M. & Macara, I. G. (1997) Ran-binding protein 1 (RanBP1) forms a ternary complex with Ran and karyopherin beta and reduces Ran GTPase-activating protein (RanGAP) inhibition by karyopherin beta. *J Biol Chem* **272**, 551-555.
- Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. (2004) Nuclear export of microRNA precursors. *Science* **303**, 95-98.
- Madl, T., Bermel, W. & Zangger, K. (2009) Use of relaxation enhancements in a paramagnetic environment for the structure determination of proteins using NMR spectroscopy. *Angew Chem Int Ed Engl* **48**, 8259-8262.
- Mahajan, R., Delphin, C., Guan, T., Gerace, L. & Melchior, F. (1997) A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97-107.
- Malik, H. S., Eickbush, T. H. & Goldfarb, D. S. (1997) Evolutionary specialization of the nuclear targeting apparatus. *Proc Natl Acad Sci USA* **94**, 13738-13742.
- Malim, M. H., McCarn, D. F., Tiley, L. S. & Cullen, B. R. (1991) Mutational definition of the human immunodeficiency virus type 1 Rev activation domain. *J Virol* **65**, 4248-4254.
- Mason, D. A., Stage, D. E. & Goldfarb, D. S. (2009) Evolution of the metazoan-specific importin alpha gene family. *J Mol Evol* **68**, 351-365.
- Matsuura, Y., Lange, A., Harreman, M. T., Corbett, A. H. & Stewart, M. (2003) Structural basis for Nup2p function in cargo release and karyopherin recycling in nuclear import. *EMBO J* **22**, 5358-5369.
- Matsuura, Y. & Stewart, M. (2004) Structural basis for the assembly of a nuclear export complex. *Nature* **432**, 872-877.
- Matsuura, Y. & Stewart, M. (2005) Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. *EMBO J* **24**, 3681-3689.
- Matunis, M. J., Coutavas, E. & Blobel, G. (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* **135**, 1457-1470.
- Maurer, P., Redd, M., Solsbacher, J., Bischoff, F. R., Greiner, M., Podtelejnikov, A. V., Mann, M., Stade, K., Weis, K. & Schlenstedt, G. (2001) The nuclear export receptor Xpo1p forms distinct complexes with NES transport substrates and the yeast Ran binding protein 1 (Yrb1p). *Mol Biol Cell* **12**, 539-549.

- McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32-41.
- Meilleur, F., Weiss, K. L. & Myles, D. A. (2009) Deuterium labeling for neutron structure-function-dynamics analysis. *Methods Mol Biol* **544**, 281-292.
- Melchior, F., Paschal, B., Evans, J. & Gerace, L. (1993) Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. *J Cell Biol* **123**, 1649-1659.
- Milburn, M. V., Tong, L., deVos, A. M., Brünger, A., Yamaizumi, Z., Nishimura, S. & Kim, S. H. (1990) Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247**, 939-945.
- Mingot, J. M., Bohnsack, M. T., Jäkle, U. & Görlich, D. (2004) Exportin 7 defines a novel general nuclear export pathway. *EMBO J* **23**, 3227-3236.
- Mingot, J. M., Kostka, S., Kraft, R., Hartmann, E. & Görlich, D. (2001) Importin 13: a novel mediator of nuclear import and export. *EMBO J* **20**, 3685-3694.
- Mitrousis, G., Olia, A. S., Walker-Kopp, N. & Cingolani, G. (2008) Molecular basis for the recognition of snurportin 1 by importin beta. *J Biol Chem* **283**, 7877-7884.
- Monecke, T., Güttler, T., Neumann, P., Dickmanns, A., Görlich, D. & Ficner, R. (2009) Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* **324**, 1087-1091.
- Moore, M. S. & Blobel, G. (1993) The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature* **365**, 661-663.
- Moore, M. S. & Blobel, G. (1994) Purification of a Ran-interacting protein that is required for protein import into the nucleus. *Proc Natl Acad Sci USA* **91**, 10212-10216.
- Mosammamarast, N., Guo, Y., Shabanowitz, J., Hunt, D. F. & Pemberton, L. F. (2002) Pathways mediating the nuclear import of histones H3 and H4 in yeast. *J Biol Chem* **277**, 862-868.
- Mosammamarast, N., Jackson, K. R., Guo, Y., Brame, C. J., Shabanowitz, J., Hunt, D. F. & Pemberton, L. F. (2001) Nuclear import of histone H2A and H2B is mediated by a network of karyopherins. *J Cell Biol* **153**, 251-262.
- Moy, T. I. & Silver, P. A. (2002) Requirements for the nuclear export of the small ribosomal subunit. *J Cell Sci* **115**, 2985-2995.
- Mühlhäusser, P., Müller, E. C., Otto, A. & Kutay, U. (2001) Multiple pathways contribute to nuclear import of core histones. *EMBO Rep* **2**, 690-696.
- Nachury, M. V. & Weis, K. (1999) The direction of transport through the nuclear pore can be inverted. *Proc Natl Acad Sci USA* **96**, 9622-9627.
- Nagata, T. (1963) The molecular synchrony and sequential replication of DNA in Escherichia coli. *Proc Natl Acad Sci USA* **49**, 551-559.
- Nakielnny, S. & Dreyfuss, G. (1998) Import and export of the nuclear protein import receptor transportin by a mechanism independent of GTP hydrolysis. *Curr Biol* **8**, 89-95.
- Neidhardt, F. C., Bloch, P. L. & Smith, D. F. (1974) Culture medium for enterobacteria. *J Bacteriol* **119**, 736-747.
- Nilges, M. (1993) A calculation strategy for the structure determination of symmetric dimers by 1H NMR. *Proteins* **17**, 297-309.
- Nilges, M. & O'Donoghue, S. I. (1998) Ambiguous NOEs and automated NOE assignment. *Prog Nucl Mag Res Sp* **32**, 107-139.
- Nishi, K., Yoshida, M., Fujiwara, D., Nishikawa, M., Horinouchi, S. & Beppu, T. (1994) Leptomycin B targets a regulatory cascade of crm1, a fission yeast nuclear protein, involved in control of higher order chromosome structure and gene expression. *J Biol Chem* **269**, 6320-6324.

- Oh, B. H., Westler, W. M., Darba, P. & Markley, J. L. (1988) Protein carbon-13 spin systems by a single two-dimensional nuclear magnetic resonance experiment. *Science* **240**, 908-911.
- Ohno, M., Segref, A., Bachi, A., Wilm, M. & Mattaj, I. W. (2000) PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* **101**, 187-198.
- Ohtsubo, M., Okazaki, H. & Nishimoto, T. (1989) The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. *J Cell Biol* **109**, 1389-1397.
- Okada, C., Yamashita, E., Lee, S. J., Shibata, S., Katahira, J., Nakagawa, A., Yoneda, Y. & Tsukihara, T. (2009) A high-resolution structure of the pre-microRNA nuclear export machinery. *Science* **326**, 1275-1279.
- Ossareh-Nazari, B., Bachelier, F. & Dargemont, C. (1997) Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* **278**, 141-144.
- Panté, N. & Kann, M. (2002) Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol Biol Cell* **13**, 425-434.
- Paraskeva, E., Izaurralde, E., Bischoff, F. R., Huber, J., Kutay, U., Hartmann, E., Lührmann, R. & Görlich, D. (1999) CRM1-mediated recycling of snurportin 1 to the cytoplasm. *J Cell Biol* **145**, 255-264.
- Partridge, J. R. & Schwartz, T. U. (2009) Crystallographic and biochemical analysis of the Ran-binding zinc finger domain. *J Mol Biol* **391**, 375-389.
- Paschal, B. M. & Gerace, L. (1995) Identification of NTF2, a cytosolic factor for nuclear import that interacts with nuclear pore complex protein p62. *J Cell Biol* **129**, 925-937.
- Perry, J. & Kleckner, N. (2003) The ATRs, ATMs, and TORs are giant HEAT repeat proteins. *Cell* **112**, 151-155.
- Petosa, C., Schoehn, G., Askjaer, P., Bauer, U., Moulin, M., Steuerwald, U., Soler-López, M., Baudin, F., Mattaj, I. W. & Müller, C. W. (2004) Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex. *Mol Cell* **16**, 761-775.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612.
- Pintacuda, G. & Otting, G. (2002) Identification of protein surfaces by NMR measurements with a paramagnetic Gd(III) chelate. *J Am Chem Soc* **124**, 372-373.
- Plafker, S. M. & Macara, I. G. (2000) Importin-11, a nuclear import receptor for the ubiquitin-conjugating enzyme, UbcM2. *EMBO J* **19**, 5502-5513.
- Plafker, S. M. & Macara, I. G. (2002) Ribosomal protein L12 uses a distinct nuclear import pathway mediated by importin 11. *Mol Cell Biol* **22**, 1266-1275.
- Pollard, V. W., Michael, W. M., Nakielny, S., Siomi, M. C., Wang, F. & Dreyfuss, G. (1996) A novel receptor-mediated nuclear protein import pathway. *Cell* **86**, 985-994.
- Radu, A., Blobel, G. & Moore, M. S. (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 USA* **92**, 1769-1773.
- Reichelt, R., Holzenburg, A., Buhle, E. L., Jarnik, M., Engel, A. & Aebi, U. (1990) Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J Cell Biol* **110**, 883-894.
- Ribbeck, K. & Görlich, D. (2001) Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* **20**, 1320-1330.
- Ribbeck, K. & Görlich, D. (2002) The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J* **21**, 2664-2671.
- Ribbeck, K., Kutay, U., Paraskeva, E. & Görlich, D. (1999) The translocation of transportin-cargo complexes through nuclear pores is independent of both Ran and energy. *Curr Biol* **9**, 47-50.

- Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M. & Görlich, D. (1998) NTF2 mediates nuclear import of Ran. *EMBO J* **17**, 6587-6598.
- Rice, L. M. & Brünger, A. T. (1994) Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277-290.
- Richards, S. A., Lounsbury, K. M., Carey, K. L. & Macara, I. G. (1996) A nuclear export signal is essential for the cytosolic localization of the Ran binding protein, RanBP1. *J Cell Biol* **134**, 1157-1168.
- Richards, S. A., Lounsbury, K. M. & Macara, I. G. (1995) The C terminus of the nuclear RAN/TC4 GTPase stabilizes the GDP-bound state and mediates interactions with RCC1, RAN-GAP, and HTF9A/RANBP1. *J Biol Chem* **270**, 14405-14411.
- Riek, R., Wider, G., Pervushin, K. & Wüthrich, K. (1999) Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proc Natl Acad Sci USA* **96**, 4918-4923.
- Rittinger, K., Budman, J., Xu, J., Volinia, S., Cantley, L. C., Smerdon, S. J., Gamblin, S. J. & Yaffe, M. B. (1999) Structural analysis of 14-3-3 phosphopeptide complexes identifies a dual role for the nuclear export signal of 14-3-3 in ligand binding. *Mol Cell* **4**, 153-166.
- Robbins, J., Dilworth, S. M., Laskey, R. A. & Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615-623.
- Rouf, M. A. (1964) Spectrochemical analysis of inorganic elements in bacteria. *J Bacteriol* **88**, 1545-1549.
- Rout, M. P. & Blobel, G. (1993) Isolation of the yeast nuclear pore complex. *J Cell Biol* **123**, 771-783.
- Saitoh, H., Pu, R., Cavenagh, M. & Dasso, M. (1997) RanBP2 associates with Ubc9p and a modified form of RanGAP1. *Proc Natl Acad Sci USA* **94**, 3736-3741.
- Sambrook, J., Fritsch, E. & Maniatis, T. (1989) Molecular cloning: a laboratory manual, Volume 1. *Plainview, N.Y., Cold Spring Harbor Laboratory Press* 1659.
- Scheffzek, K., Klebe, C., Fritz-Wolf, K., Kabsch, W. & Wittinghofer, A. (1995) Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. *Nature* **374**, 378-381.
- Schlenstedt, G., Smirnova, E., Deane, R., Solsbacher, J., Kutay, U., Görlich, D., Ponstingl, H. & Bischoff, F. R. (1997) Yrb4p, a yeast ran-GTP-binding protein involved in import of ribosomal protein L25 into the nucleus. *EMBO J* **16**, 6237-6249.
- Schwoebel, E. D., Talcott, B., Cushman, I. & Moore, M. S. (1998) Ran-dependent signal-mediated nuclear import does not require GTP hydrolysis by Ran. *J Biol Chem* **273**, 35170-35175.
- Shaheen, H. H. & Hopper, A. K. (2005) Retrograde movement of tRNAs from the cytoplasm to the nucleus in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* **102**, 11290-11295.
- Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223.
- Siomi, M. C., Eder, P. S., Kataoka, N., Wan, L., Liu, Q. & Dreyfuss, G. (1997) Transportin-mediated nuclear import of heterogeneous nuclear RNP proteins. *J Cell Biol* **138**, 1181-1192.
- Smith, A., Brownawell, A. & Macara, I. G. (1998) Nuclear import of Ran is mediated by the transport factor NTF2. *Curr Biol* **8**, 1403-1406.
- Solsbacher, J., Maurer, P., Bischoff, F. R. & Schlenstedt, G. (1998) Cse1p is involved in export of yeast importin alpha from the nucleus. *Mol Cell Biol* **18**, 6805-6815.
- Solsbacher, J., Maurer, P., Vogel, F. & Schlenstedt, G. (2000) Nup2p, a yeast nucleoporin, functions in bidirectional transport of importin alpha. *Mol Cell Biol* **20**, 8468-8479.
- Stade, K., Ford, C. S., Guthrie, C. & Weis, K. (1997) Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* **90**, 1041-1050.
- Stewart, M. (2003) Structural biology. Nuclear trafficking. *Science* **302**, 1513-1514.
- Stewart, M. (2007) Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol* **8**, 195-208.

- Strasser, A., Dickmanns, A., Lührmann, R. & Ficner, R. (2005) Structural basis for m3G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1. *EMBO J* **24**, 2235-2243.
- Strasser, A., Dickmanns, A., Schmidt, U., Penka, E., Urlaub, H., Sekine, M., Lührmann, R. & Ficner, R. (2004) Purification, crystallization and preliminary crystallographic data of the m3G cap-binding domain of human snRNP import factor snurportin 1. *Acta Crystallogr D Biol Crystallogr* **60**, 1628-1631.
- Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234.
- Stüven, T., Hartmann, E. & Görlich, D. (2003) Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. *EMBO J* **22**, 5928-5940.
- Taagepera, S., McDonald, D., Loeb, J. E., Whitaker, L. L., McElroy, A. K., Wang, J. Y. & Hope, T. J. (1998) Nuclear-cytoplasmic shuttling of C-ABL tyrosine kinase. *Proc Natl Acad Sci USA* **95**, 7457-7462.
- Tejomurtula, J., Lee, K. B., Tripurani, S. K., Smith, G. W. & Yao, J. (2009) Role of importin alpha8, a new member of the importin alpha family of nuclear transport proteins, in early embryonic development in cattle. *Biol Reprod* **81**, 333-342.
- Thomas, F. & Kutay, U. (2003) Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *J Cell Sci* **116**, 2409-2419.
- Truant, R. & Cullen, B. R. (1999) The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. *Mol Cell Biol* **19**, 1210-1217.
- Tsuji, L., Takumi, T., Imamoto, N. & Yoneda, Y. (1997) Identification of novel homologues of mouse importin alpha, the alpha subunit of the nuclear pore-targeting complex, and their tissue-specific expression. *FEBS Lett* **416**, 30-34.
- Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. & Kay, L. E. (2003) Cross-correlated relaxation enhanced <sup>1</sup>H-<sup>13</sup>C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* **125**, 10420-10428.
- Tugarinov, V., Kanelis, V. & Kay, L. E. (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc* **1**, 749-754.
- Vetter, I. R., Arndt, A., Kutay, U., Görlich, D. & Wittinghofer, A. (1999a) Structural view of the Ran-Importin beta interaction at 2.3 Å resolution. *Cell* **97**, 635-646.
- Vetter, I. R., Nowak, C., Nishimoto, T., Kuhlmann, J. & Wittinghofer, A. (1999b) Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. *Nature* **398**, 39-46.
- Vetter, I. R. & Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304.
- Wälde, S. & Kehlenbach, R. H. (2010) The Part and the Whole: functions of nucleoporins in nucleocytoplasmic transport. *Trends Cell Biol* **20**, 461-469.
- Walker, P., Doenecke, D. & Kahle, J. (2009) Importin 13 mediates nuclear import of histone fold-containing chromatin accessibility complex heterodimers. *J Biol Chem* **284**, 11652-11662.
- Watson, M. L. (1959) Further observations on the nuclear envelope of the animal cell. *J Biophys Biochem Cytol* **6**, 147-156.
- Weinmann, L., Höck, J., Ivacevic, T., Ohrt, T., Mütze, J., Schwill, P., Kremmer, E., Benes, V., Urlaub, H. & Meister, G. (2009) Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. *Cell* **136**, 496-507.
- Weis, K., Ryder, U. & Lamond, A. I. (1996) The conserved amino-terminal domain of hSRP1 alpha is essential for nuclear protein import. *EMBO J* **15**, 1818-1825.
- Weiss, M. S. & Hilgenfeld, R. (1997) On the use of the merging R factor as a quality indicator for X-ray data. *J Appl Crystallogr* **30**, 203-205.

- Wen, W., Harootunian, A. T., Adams, S. R., Feramisco, J., Tsien, R. Y., Meinkoth, J. L. & Taylor, S. S. (1994) Heat-stable inhibitors of cAMP-dependent protein kinase carry a nuclear export signal. *J Biol Chem* **269**, 32214-32220.
- Wen, W., Meinkoth, J. L., Tsien, R. Y. & Taylor, S. S. (1995) Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463-473.
- Will, C. L. & Lührmann, R. (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13**, 290-301.
- Wohlwend, D., Strasser, A., Dickmanns, A. & Ficner, R. (2007) Structural Basis for RanGTP Independent Entry of Spliceosomal U snRNPs into the Nucleus. *J Mol Biol* **374**, 1129-1138.
- Wolff, B., Sanglier, J. J. & Wang, Y. (1997) Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. *Chem Biol* **4**, 139-147.
- Yamada, M., Mattaj, I. W. & Yoneda, Y. (2004) An ATP-dependent activity that releases RanGDP from NTF2. *J Biol Chem* **279**, 36228-36234.
- Yang, Q., Rout, M. P. & Akey, C. W. (1998) Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Mol Cell* **1**, 223-234.
- Yasuhara, N., Shibasaki, N., Tanaka, S., Nagai, M., Kamikawa, Y., Oe, S., Asally, M., Kamachi, Y., Kondoh, H. & Yoneda, Y. (2007) Triggering neural differentiation of ES cells by subtype switching of importin-alpha. *Nat Cell Biol* **9**, 72-79.
- Yokoyama, N., Hayashi, N., Seki, T., Panté, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, T. & Aebi, U. (1995) A giant nucleopore protein that binds Ran/TC4. *Nature* **376**, 184-188.
- Yoshida, K. & Blobel, G. (2001) The karyopherin Kap142p/Msn5p mediates nuclear import and nuclear export of different cargo proteins. *J Cell Biol* **152**, 729-740.
- Zeitler, B. & Weis, K. (2004) The FG-repeat asymmetry of the nuclear pore complex is dispensable for bulk nucleocytoplasmic transport in vivo. *J Cell Biol* **167**, 583-590.
- Zhang, M. J. & Dayton, A. I. (1998) Tolerance of diverse amino acid substitutions at conserved positions in the nuclear export signal (NES) of HIV-1 Rev. *Biochem Biophys Res Commun* **243**, 113-116.

## Author's contributions to publications

The work presented in this thesis has been a joint endeavor of the laboratories of Prof. Dirk Görlich (MPI for Biophysical Chemistry, Göttingen), Prof. Ralf Ficner (Georg-August-Universität Göttingen), Prof. Michael Sattler (Technische Universität und Helmholtz Zentrum München), Dr. Raymond Poot (Erasmus Medical Center, Rotterdam, Netherlands), Dr. Robbert Rottier (Erasmus Medical Center, Rotterdam, Netherlands) and Prof. Giulio Superti-Furga (CeMM, Vienna, Austria). All contributors to the presented studies are listed on the chapters' title pages or in the acknowledgment sections. For the sake of brevity, the following statement only summarizes the contributions by the thesis' author.

### CHAPTER 2

Gontan, C., **Güttler, T.**, Engelen, E., Demmers, J., Fornerod, M., Grosveld, F. G., Tibboel, D., Görlich, D., Poot, R. A. & Rottier, R. J. (2009) Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *J Cell Biol* **185**, 27-34.

*C. Gontan and T. Güttler contributed equally to this work.*

T. Güttler conceived central nuclear import experiments and prepared reagents (including the  $\alpha$ -X. laevis Exp4 antibody, many fluorescent probes and transport receptors). He performed part of the binding and import assays. T. Güttler contributed to the interpretation of the data, wrote the first draft of the manuscript and helped to complete the paper for publication (including the preparation of figures).

### CHAPTER 3

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

*T. Monecke and T. Güttler contributed equally to this work.*

T. Güttler designed part of the biochemical experiments. He prepared all proteins (with the exception of Snurportin1 used for the crystallization experiments), some with the help of D. Deichsel. T. Güttler established constructs and conditions for the assembly of stable, chromatographically homogeneous and crystallizable export complexes. T. Güttler contributed to screening for crystallization conditions. He performed all biochemical assays described in the main body and in the supplements of the paper. He contributed to the interpretation of the data and to writing of the manuscript (including the preparation of figures).

**CHAPTERS 4 & 6**

**Güttler, T.**, Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., Görlich, D. (2010) NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol*, **accepted for publication** (on condition of editorial changes in the text).

*T. Güttler, T. Madl and P. Neumann contributed equally to this work.*

Note added for publication of this thesis:

The citation of the article's print version is *Nat Struct Mol Biol* **17**, 1367-1376.

T. Güttler conceived part of this study. He designed expression constructs and optimized the recombinant protein expression and purification. All proteins have been prepared by T. Güttler with the help of D. Deichsel. T. Güttler carried out all binding and export experiments, established the deuteration conditions for CRM1, performed isotope labeling and prepared all samples for the NMR-spectroscopic experiments. T. Güttler reconstituted and crystallized all export complexes and contributed to data collection and structure solving. T. Güttler prepared all figures shown in the manuscript (Fig. 3 and Supplementary Fig. 4+5 together with T. Madl). T. Güttler contributed to data interpretation and to writing of the manuscript.

Hantschel, O., Wiesner, S., **Güttler, T.**, Mackereth, C. D., Rix, L. L., Mikes, Z., Dehne, J., Görlich, D., Sattler, M. & Superti-Furga, G. (2005) Structural basis for the cytoskeletal association of Bcr-Abl/c-Abl. *Mol Cell* **19**, 461-473.

T. Güttler conceived part of the nuclear transport experiments. T. Güttler prepared all reagents for the nuclear transport part of the paper and performed all pulldown and export assays (Fig. 3). To verify the data, he carried out numerous experiments not shown in the paper, including microinjection assays with *X. laevis* oocytes and immunofluorescence experiments. T. Güttler contributed to interpretation of the data and to writing of the paper.

Göttingen, August 2010

Thomas Güttler

Prof. Dr. Dirk Görlich

---

## Curriculum vitae

**Name** Thomas Güttler  
**Day and place of birth** July 12, 1978; Löbau (Germany)  
**Nationality** German

### Education

**2007-2010** **Max Planck Institute for Biophysical Chemistry, Göttingen** (Germany) and **Georg-August-Universität Göttingen**  
GGNB Program "Biomolecules: Structure-Function-Dynamics"  
Doctoral studies in the laboratory of Prof. Dirk Görlich

**2004-2007** **Ruprecht-Karls-Universität Heidelberg, ZMBH** (Germany)  
Doctoral studies in the laboratory of Prof. Dirk Görlich

**2001-2003** **Ruprecht-Karls-Universität Heidelberg** (Germany)  
Diploma studies in the International Master's Program "Molecular and Cellular Biology"  
Diploma/Master Thesis in the laboratory of Prof. Dirk Görlich (ZMBH)  
*"Nuclear export of translation factors as a general mechanism to confine translation to the cytoplasm"*

**2001** **University of Copenhagen** (Denmark)  
Workshops on biomembranes, membrane proteins and signal transduction

**2000-2001** **University of Mississippi** (Oxford/MS, U.S.A.)  
Studies of biochemistry and microbiology  
Student research project in the laboratory of Dr. Lidija Halda-Alija  
*"Quorum sensing by wetland rhizobacteria"*

**1998-2000** **Friedrich-Schiller-Universität Jena** (Germany)  
Undergraduate studies in biology  
Pre-degree ("Vordiplom") in biology

**1997** **Geschwister-Scholl-Gymnasium Löbau** (Germany)  
Abitur

**Publications** (research articles)

("\*" denotes equal contribution)

Güttler, T. \*, Madl, T. \*, Neumann, P. \*, Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., Görlich, D. (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. **Nat. Struct. Mol. Biol.** accepted for publication (on condition of editorial changes in the text).

Note added for publication of this thesis:

The citation of the article's print version is *Nat Struct Mol Biol* **17**, 1367-1376.

Mohr, D., Frey, S., Fischer, T., Güttler, T., and Görlich, D. (2009). Characterisation of the passive permeability barrier of nuclear pore complexes. **EMBO J.** **28**, 2541-53.

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

Gontan, C.\* , Güttler, T.\* , Engelen, E., Demmers, J., Fornerod, M., Grosveld, F., Tibboel, D., Görlich, D., Poot, R.A., and Rottier, R. (2009). Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. **J. Cell Biol.**, **185**, 27-34.

Hantschel, O., Wiesner, S., Güttler, T., Mackereth, C. D., Rix, L. L., Mikes, Z., Dehne, J., Görlich, D., Sattler, M., and Superti-Furga, G. (2005). Structural basis for the cytoskeletal association of Bcr-Abl/c-Abl. **Mol. Cell** **19**, 461-73.

Bono, F., Ebert, J., Unterholzner, L., Güttler, T., Izaurrealde, E., and Conti, E. (2004). Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. **EMBO Rep.** **5**, 304-10.

Kruse, C., Jaedicke, A., Beaudouin, J., Böhl, F., Ferring, D., Güttler, T., Ellenberg, J., and Jansen, R. P. (2002). Ribonucleoprotein-dependent localization of the yeast class V myosin Myo4p. **J. Cell Biol.** **159**, 971-82.