doi:10.1016/j.jmb.2011.07.038 J. Mol. Biol. (2011) 412, 226–234

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

# Structural Basis of Importin-α-Mediated Nuclear Transport for Ku70 and Ku80

Agnes A. S. Takeda[1], Andrea C. de Barros[1], Chiung-Wen Chang[2], Boštjan Kobe[2] and Marcos R. M. Fontes[1]*

[1]*Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, SP 18618-970, Brazil*

[2]*School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Queensland 4072, Australia*

*Received 12 May 2011;*
*received in revised form*
*15 July 2011;*
*accepted 19 July 2011*
Available online
23 July 2011

***Edited by K. Morikawa***

***Keywords:***
importin-α;
nuclear import pathway;
DNA repair proteins;
Ku70 and Ku80 proteins;
X-ray crystallography

Ku70 and Ku80 form a heterodimeric complex involved in multiple nuclear processes. This complex plays a key role in DNA repair due to its ability to bind DNA double-strand breaks and facilitate repair by the nonhomologous end-joining pathway. Ku70 and Ku80 have been proposed to contain bipartite and monopartite nuclear localization sequences (NLSs), respectively, that allow them to be translocated to the nucleus independently of each other via the classical importin-α (Impα)/importin-β-mediated nuclear import pathway. To determine the structural basis of the recognition of Ku70 and Ku80 proteins by Impα, we solved the crystal structures of the complexes of Impα with the peptides corresponding to the Ku70 and Ku80 NLSs. Our structural studies confirm the binding of the Ku80 NLS as a classical monopartite NLS but reveal an unexpected binding mode for Ku70 NLS with only one basic cluster bound to the receptor. Both Ku70 and Ku80 therefore contain monopartite NLSs, and sequences outside the basic cluster make favorable interactions with Impα, suggesting that this may be a general feature in monopartite NLSs. We show that the Ku70 NLS has a higher affinity for Impα than the Ku80 NLS, consistent with more extensive interactions in its N-terminal region. The prospect of nuclear import of Ku70 and Ku80 independently of each other provides a powerful regulatory mechanism for the function of the Ku70/Ku80 heterodimer and independent functions of the two proteins.



## Introduction

Ku70 and Ku80 form a heterodimeric complex involved in multiple nuclear processes. This complex plays a key role in DNA repair due to its ability to bind DNA double-strand breaks and facilitate repair by the nonhomologous end-joining pathway.[1] Other functions are related to chromosome and telomere maintenance,[2] regulation of gene-specific transcription[3] and protection of developing neurons against apoptosis.[4] The crystal structure of the Ku heterodimer has been elucidated alone and bound to DNA.[5] Ku70 and Ku80 have a common topology, and the Ku complex forms an asymmetric ring to encircle duplex DNA. The channel that is formed by the Ku ring interacts with the sugar-phosphate DNA backbone in a sequence-independent interaction.[5] Because both subunits of the Ku complex in the crystallographic structure have their C-terminal

*Corresponding author. E-mail address: fontes@ibb.unesp.br.

Abbreviations used: ARM, armadillo; Impα, importin-α; Impβ, importin-β; NLS, nuclear localization sequence; SV40, simian virus 40; TAg, large tumor antigen; cNLS, classical NLS; AR, androgen receptor; GST, glutathione *S*-transferase; PDB, Protein Data Bank.

domains truncated or not modeled due to weak electron density, the C-terminal domains for Ku70 and Ku80 have been elucidated independently by nuclear magnetic resonance techniques.[6,7]

Because the majority of the Ku complex functions take place in the cell nucleus, it is important to understand how its nuclear transport occurs. Nuclear localization sequences (NLSs) were identified in both Ku70 and Ku80,[8,9] which allow them to be translocated to the nucleus independently of each other using their own NLS by the classical importin-α (Impα)/importin-β (Impβ)-mediated nuclear import pathway.[8,10,11]

The classical NLSs (cNLSs) contain one cluster or two clusters of positively charged amino acids and are therefore usually divided into monopartite cNLS—containing a single cluster of basic residues—and bipartite cNLS—containing two clusters of basic residues separated by 10–12 variant residues.[12] Structural studies have shown that both classes are recognized by the nuclear import receptor Impα. This nuclear protein receptor has two NLS binding sites formed by conserved residues in its armadillo (ARM) repeat domain, the major and the minor NLS binding sites. The N- and C-terminal clusters of a bipartite NLS interact with the minor and the major NLS binding sites, which correspond to ARM repeats 4–8 and 1–4, respectively. On the other hand, a monopartite NLS interacts primarily with the major binding site.[13–15]

While the Ku80 NLS (amino acids 561–569) was proposed to be a monopartite cNLS similar to the simian virus 40 (SV40) large tumor antigen (TAg) NLS,[8] the Ku70 NLS (amino acids 539–556) was classified as a bipartite NLS similar to nucleoplasmin and N1N2 NLSs.[9]

Nuclear localization is an important control mechanism that allows the cell to regulate DNA replication, DNA repair and many other biological functions. Although many NLSs have been identified, the understanding of the specificity of NLS recognition remains limited.[16] Because nuclear localization appears to play a key role in regulating the physiological function of the Ku heterodimer by regulating the nuclear localization of Ku monomers independently *in vivo*,[8] we set out to study the structural basis of Ku70 and Ku80 NLS binding to Impα. In particular, we were interested in the binding mechanism of the unusual bipartite NLS proposed in Ku70. We determined the crystal structures of the peptides corresponding to Ku70 and Ku80 NLSs bound to Impα and found that,

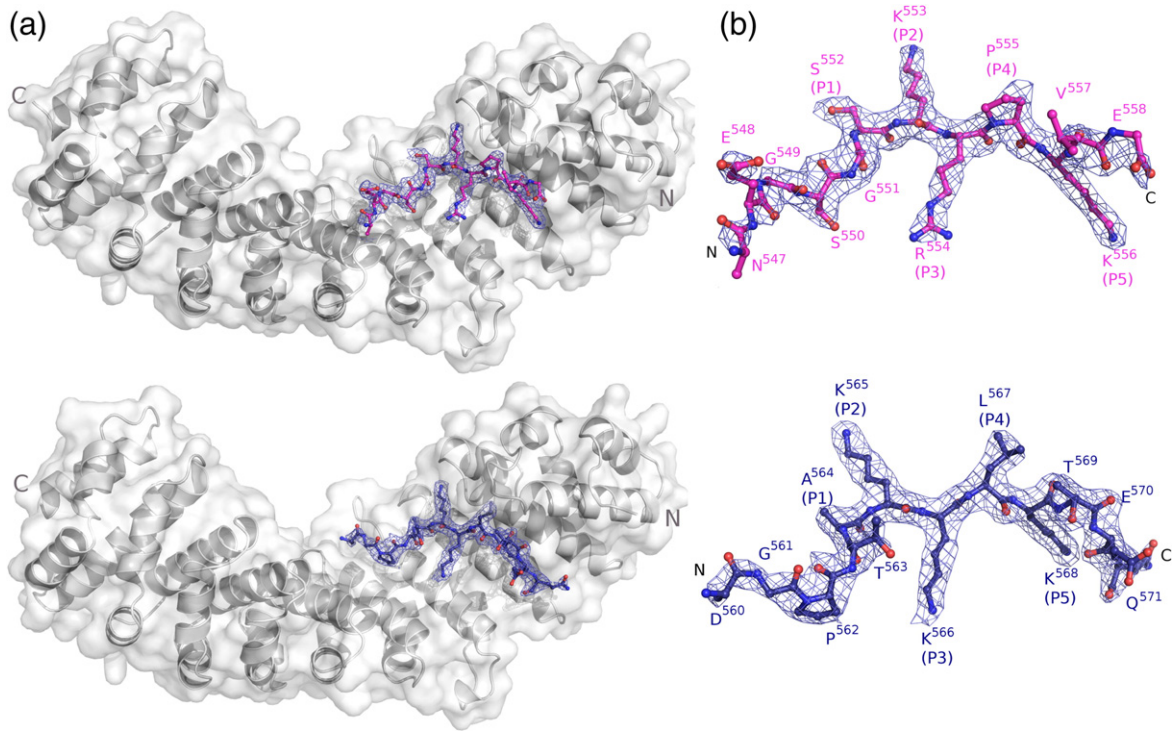


**Fig. 1.** Crystal structures of Impα:Ku70NLS and Impα:Ku80NLS complexes. (a) Overall structure of the Impα:Ku70NLS complex. Impα is shown as a ribbon diagram. Ku70 NLS (pink) is shown in a stick representation. (b) As in (a), but for the Impα:Ku80NLS (blue) complex. (c) Electron density map (coefficients $3|F_{obs}| - 2|F_{calc}|$) of the Ku70NLS:Impα complex crystals in the area corresponding to the peptide (contoured at 1.5 SD). All peptide residues were omitted from the model and simulated annealing run with the starting temperature of 1000 K. (d) As in (c), but for the Ku80 NLS peptide (contoured at 1.2 SD).

while Ku80 NLS binds as a typical monopartite NLS, contrary to what was proposed, only one basic cluster is bound to Impα in Ku70 NLS. In both Ku70 and Ku80 NLSs, residues outside the basic cluster make favorable interactions with Impα similar to SV40 TAg NLS. The Ku70 NLS shows a higher affinity for Impα than the Ku80 NLS as measured by a solid-phase binding assay, consistent with more extensive interactions in its N-terminal region. Our study provides insights into the regulation of the Ku protein function.

## Results

### Structures of Impα in complex with peptides corresponding to Ku70 and Ku80 NLSs

The peptides corresponding to Ku70 and Ku80 NLSs were co-crystallized with an N-terminally truncated mouse Impα lacking residues 1–69 (mImpαΔIBB); the truncated residues are responsible for autoinhibition.[17] The co-crystals with both peptides were grown under similar conditions and isomorphously to other mouse Impα crystals.[12,14] Electron density maps based on the Impα model, following rigid-body refinement, clearly showed electron density corresponding to the peptides (Fig. 1). The structures were refined at 2.6 and 2.3 Å resolutions, respectively, for Ku70NLS:Impα and Ku80NLS:Impα complexes (Table 1).

Impα is an elongated protein composed of 10 ARM motifs displayed in tandem,[17,18] each containing three α-helices (H1, H2 and H3).[14] The NLS binding sites are located in a concave groove on the surface of Impα. The major binding site is formed mainly by the H3 helices of ARM repeats 1–4, while the minor site is located at ARM repeats 6–8. The Impα structure in both Ku70 NLS and Ku80 NLS complexes is essentially identical with that of full-length Impα and other complexes with monopartite NLS-like peptides reported previously. The root-mean-square deviations (RMSDs) of $C^{\alpha}$ atoms of Impα residues 72–496 are 0.34 and 0.55 Å between the full-length Impα and the Impα:Ku70NLS and Impα:Ku80NLS complexes, respectively, and 0.36 and 0.69 Å between the Impα:CN-SV40TAg complex [extended SV40 TAg NLS peptide (comprising TAg amino acids 110–132) complex][19] and the Impα:Ku70NLS and Impα:Ku80NLS complexes, respectively.

### Binding of Ku70 NLS to Impα

The peptide corresponding to Ku70 NLS, $^{537}$EGKVTKRKHDNEGSGSKRPKVG$^{558}$, binds only to the Impα major binding site with the main chain positioned in antiparallel configuration when compared to the direction of the ARM repeats.

**Table 1.** X-ray data collection and refinement statistics

| | Ku70NLS: mImpαΔIBB | Ku80NLS: mImpαΔIBB |
|---|---|---|
| *Diffraction data statistics* | | |
| Unit cell (Å) | $a=78.5$, $b=90.0$, $c=100.1$ | $a=77.4$, $b=88.3$, $c=98.3$ |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ |
| Resolution (Å) | 40.0–2.6 (2.69–2.6)[a] | 40.0–2.29 (2.34–2.29)[a] |
| Unique reflections | 22,152 | 29,010 |
| Multiplicity | 4.3 (4.2)[a] | 5.3 (4.5)[a] |
| Completeness (%) | 99.3 (99.9)[a] | 98.7 (91.8)[a] |
| $R_{merge}$[b] (%) | 11.8 (61.5)[a] | 8.1 (52.3)[a] |
| Average $I/\sigma(I)$ | 14.36 (2.73)[a] | 20.77 (2.45)[a] |
| *Refinement statistics* | | |
| Resolution (Å) | 40.0–2.6 (2.69–2.6)[a] | 40.0–2.3 (2.34–2.29)[a] |
| Number of reflections | 23,640 | 29,010 |
| $R_{cryst}$ (%)[c] | 16.76 | 17.77 |
| $R_{free}$ (%)[d] | 22.29 | 21.79 |
| Number of non-hydrogen atoms | | |
| Protein | 3237 | 3207 |
| Peptide | 79 | 87 |
| Solvent | 98 | 216 |
| Mean *B*-factor (Å$^2$) | | |
| Protein | 53.8 | 41.2 |
| Peptide | 53.4 | 47.2 |
| Water | 55.0 | 42.3 |
| Coordinate error (Å)[e] | 0.27 | 0.26 |
| RMSDs from ideal values[e] | | |
| Bond lengths (Å) | 0.021 | 0.022 |
| Bond angles (°) | 2.132 | 1.984 |
| Ramachandran plot (%) | | |
| Residues in most favored (disallowed) regions[f] | 94.8 (0.3) | 98 (0.2) |

[a] Values in parentheses are for the highest-resolution shell.
[b] $R_{merge}=\sum_{hkl}(\sum_i(|I_{hkl,i}-\langle I_{hkl}\rangle|))/\sum_{hkl,i}\langle I_{hkl}\rangle$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices $h$, $k$ and $l$, and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. Calculated for $I>-3\ \sigma(I)$.[34]
[c] $R_{cryst}=\sum_{hkl}(||F_{obs,hkl}|-|F_{calc,hkl}||)/|F_{obs,hkl}|$, where $|F_{obs,hkl}|$ and $|F_{calc,hkl}|$ are the observed and calculated structure factor amplitudes, respectively.
[d] $R_{free}$ is equivalent to $R_{cryst}$ but calculated with reflections (5%) omitted from the refinement process.
[e] Calculated based on Luzzati plot with the program SFCHECK.[37]
[f] Calculated with the program PROCHECK.[37]

Residues 547–558 of Ku70 NLS peptide (residues 537–546 had no interpretable electron density) could be identified unambiguously in the electron density maps (Fig. 1a). The buried surface between the protein and the peptide is 893.8 Å$^2$.

Protein–peptide interactions (Fig. 2a) are consistent with the distribution of *B*-factors along the peptide chain. The average *B*-factor of Ku70 NLS in the major site (55.0 Å$^2$) has *B*-factors below the average *B*-factors of the entire structure (53.8 Å$^2$). The major site residues (residues 551–556; positions $P_1$ to $P_5$) have lower average *B*-factors (47.4 Å). The residues K553 and K556 (positions $P_2$ and $P_5$) have the lowest *B*-factors (35.9 and 42.0 Å$^2$, respectively) and the largest number of interactions between NLS

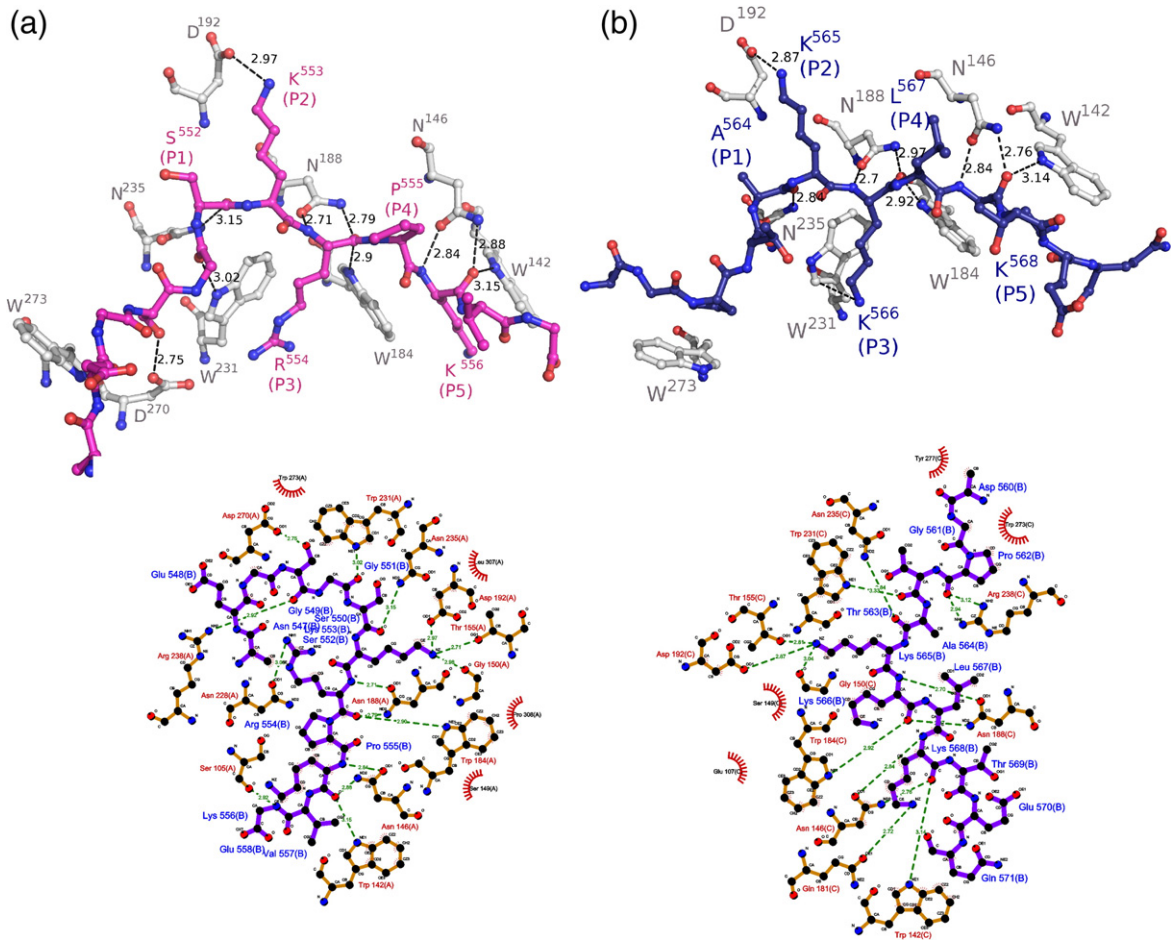


**Fig. 2.** Schematic diagram of the interactions between the Ku70 and Ku80 NLS peptides and the major binding site of Impa. Polar contacts are shown with broken lines, and hydrophobic contacts are indicated by arcs with radiating spokes. The NLS peptide residues are labeled with "R". Carbon, nitrogen and oxygen atoms are shown in black, white and gray, respectively. Generated with the program LIGPLOT.[38] (a) Ku70 NLS; (b) Ku80 NLS.

side chains and side chains of conserved residues of Impα.

The comparison of Ku70 NLS to CN-SV40TAg and the bipartite nucleoplasmin NLS[14] showed a high structural similarity between Ku70 NLS and CN-SV40TAg NLS (RMSDs of $C^{\alpha}$ atoms are 1.42 and 2.06 Å for Ku70/CN-SV40TAg and Ku70/nucleoplasmin, respectively), as can be observed in Fig. 3.

Very weak electron density was found in the minor binding site of Impα; however, it was not possible to model unambiguously any amino acid residue at this site.

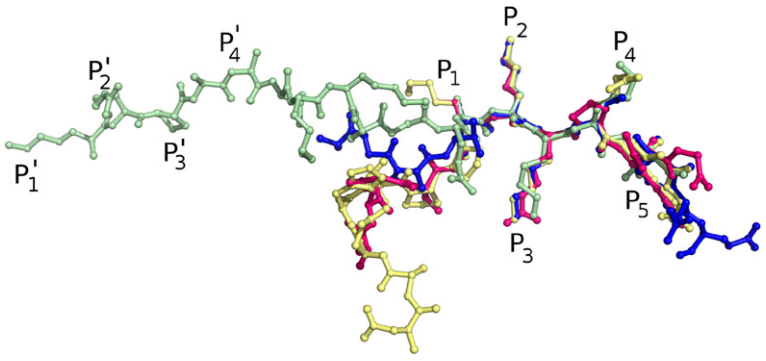


**Fig. 3.** Comparison of NLS peptides in the major NLS binding site. Ku70 (pink), Ku80 (blue), nucleoplasmin (green) (PDB ID 1EJY) and CN-SV40TAg (yellow) (PDB ID 1EJL) NLSs were superimposed using $C^{\alpha}$ atoms of the peptides. Positions binding to the major ($P_1$–$P_5$) and minor binding sites ($P_1'$–$P_4'$) are identified along the chains.

### Binding of Ku80 NLS to Impα

The peptide corresponding to the Ku80 NLS, $^{559}$EDGPTAKKLKTEQ$^{571}$, binds only to the Impα major binding site with the main chain positioned in antiparallel configuration when compared to the direction of the ARM repeats. Residues 560–571 of Ku80 NLS peptide (residue 559 had no interpretable electron density) could be identified unambiguously in the electron density maps (Fig. 1b). The buried surface between the protein and the peptide is 891.0 Å$^2$.

The average *B*-factor of Ku80 NLS at the major site (47.2 Å$^2$) is higher than the average *B*-factor of the entire structure (41.4 Å$^2$), although the average *B*-factor of residues 564–568 positioned at the major binding site (positions $P_1$–$P_5$) is 33.5 Å$^2$. The residues K565 and K568 (positions $P_2$ and $P_5$) have the lowest *B*-factors (30.3 and 32.5 Å$^2$, respectively).

The superposition of $C^{\alpha}$ atoms between Ku80 and SV40 TAg NLS peptides yields an RMSD of 0.43 Å.

### Affinity of Ku70 and Ku80 NLS binding to Impα

Solid-phase binding assays were used to compare the affinities of Ku70 and Ku80 for Impα (Fig. 4). The apparent dissociation constants ($K_d$) for Ku70 and Ku80 NLS peptides for mImpαΔIBB were measured as 29 ± 4 and 120 ± 9 nM, respectively. The $K_d$ value of the Ku70 peptide for Impα is compatible with an optimal monopartite NLS peptide binding to mImpαΔIBB (53 ± 15 nM),[20] which was obtained using a similar technique. These values are also comparable to the values obtained for CN-SV40TAg NLS peptide binding to Impα.[19]

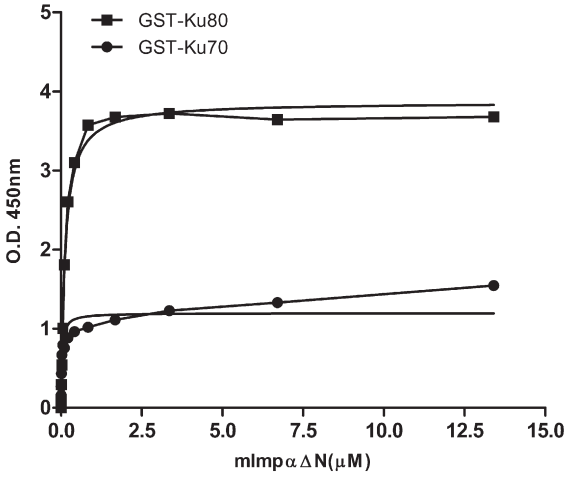


**Fig. 4.** Quantitative binding assays. The affinities of Impα for Ku70 and Ku80 NLS peptides were determined using solid-phase binding assay.[33] Binding of Impα to GST-Ku70 NLS peptide (●) and to GST-Ku80 NLS peptide (■).

## Discussion

### Ku80 NLS is a monopartite cNLS

The nine-amino-acid sequence corresponding to residues 561–569 of Ku80 was previously identified as the minimum region required for optimal transport of this protein into the nucleus.[8] It was also demonstrated that Ku80 NLS was recognized and transported to the nucleus by the Impα/Impβ complex but could not be translocated to the nucleus with either Impα or Impβ alone. In our crystal structure of the Ku80NLS:mImpαΔIBB complex, the Ku80 NLS peptide binds to Impα with Lys residues at positions $P_2$, $P_3$ and $P_5$ (KKxK), which is in accord with the monopartite NLS consensus: K R/K × R/K.[12,21] Moreover, Ku80 NLS has a sequence exactly the same as that of the well-studied monopartite SV40 TAg NLS (KKxK) at positions $P_2$, $P_3$ and $P_5$, respectively. We conclude that Ku80 binds to Impα using a monopartite NLS and that Ku80 can be transported alone (independently of Ku70) into the nucleus by the classical Impα/Impβ-mediated pathway.

### Ku70 NLS is not a bipartite NLS

It has been proposed that Ku70 NLS is bipartite because it contains two basic amino acid clusters.[9] The deletion of the N-terminal region of Ku70 NLS (K539–K544), which corresponds to the putative minor site-binding region, led to the loss of detection of the protein in the nucleus. Additionally, the deletion of only the first residue of Ku70 NLS (K539) led to a decrease of protein localization to the nucleus.[9] Thus, the authors suggested that this putative bipartite NLS has a 12-residue linker region (the sequence that connects the minor and the major site-binding basic clusters), with Lys and Val binding at the $P_1'$ and $P_2'$ positions, respectively (Table 2). Another possibility is that the NLS may have a nine-residue linker region, with the more conventional Lys and Arg residues at the $P_1'$ and $P_2'$ positions, respectively. The typical bipartite NLS consensus sequence corresponds to KR$X_{10–12}$KRRK, with a 10- to 12-residue linker region.[12] Clearly, both proposed possibilities are not entirely compatible with the available structural, functional and mutagenesis studies.[12–14,21–24] Firstly, a linker region shorter than 10 residues would not be able to span the distance between the minor site and the major site. Secondly, the N-terminal basic cluster requires basic residues (Lys or Arg) at both $P_1'$ and $P_2'$ position[19] and typically contains Lys and Arg residues at $P_1'$ and $P_2'$ positions, respectively.[12] Considering these points, Ku70 NLS is unlikely to be bipartite. The structure of the Ku70NLS:mImpα-ΔIBB complex confirms that Ku70 NLS uses only

**Table 2.** Binding to specific binding pockets of Impα based on structural data

| | Minor NLS binding site | | | | | | | Linker | | | | | | | | Major NLS binding site | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NLS source protein[a] | | | $P'_1$ | $P'_2$ | $P'_3$ | $P'_4$ | | | | | | | | | | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | | |
| Ku80 | | | | | | | | | | | | D | G | P | T | A | K | K | L | K | T | E |
| Ku70 | | | | | | | | | | N | E | G | S | G | S | K | R | P | K | | | |
| CN-SV40TAg | | | | | | | D | A | Q | H | A | A | P | P | K | K | K | R | K | | | |
| AR | | | | | | | E | A | G | M | T | L | G | A | R | K | L | K | K | L | | |
| PLSCR1 | | | | | | | | | | | | | | | G | K | I | S | K | H | W | |
| Nucleop1 | A | V | K | R | P | A | | A | T | K | K | A | G | Q | A | K | K | K | K | L | | |
| RB | | | K | R | S | A | E | G | S | N | P | P | K | P | L | K | K | L | R | G | | |
| N1N2 | R | K | K | R | K | T | E | E | S | P | L | K | D | K | A | K | K | S | K | G | | |
| c-Myc | | | K | R | V | K | L | | | | | | P | A | A | K | R | V | K | L | D | |
| Consensus | | | K | R | | | | $X_{10-12}$ | | | | | | | | K | R | x | K | | | |

NLSs are aligned as observed to bind to the NLS binding sites ($P'_1$–$P'_4$, minor binding site; $P_1$–$P_6$, major binding site as defined in Ref. 12). The consensus sequences for monopartite and bipartite NLSs have been defined in Refs. 21 and 22, respectively.

[a] CN-SV40TAg, simian virus antigen T phosphorylated on residue S112;[19] PLSCR1, phospholipid scramblase 1;[39] Nucleop1, nucleoplasmin;[14] RB, retinoblastoma protein;[12] N1N2, *Xenopus laevis* phosphoprotein N1N2;[12] c-Myc, proto-oncogene protein.[13]

one basic cluster to bind to Impα and is therefore a monopartite NLS.

## Comparison of Ku70NLS:Impα and Ku80NLS:Impα complex structures with other available NLS:Impα complex structures

In the monopartite SV40 TAg NLS, seven residues N-terminal to the basic cluster form favorable interactions with Impα, as revealed by the crystal structure of the Impα:CN-SV40TAg complex.[19] Interestingly, Ku70 NLS closely resembles the CN-SV40TAg NLS structure in that the N-terminal flanking sequences follow the same path in relation to Impα and make favorable interactions with the receptor (Fig. 3). In the case of CN-SV40TAg NLS, the flanking sequences increase its affinity to the receptor compared to a shorter sequence comprising only the basic cluster.[19,25] Analogously, the affinity assays performed with Ku70 and Ku80 NLS peptides for Impα are also in agreement with the importance of flanking sequence. Ku70 peptide bound with approximately 4-fold higher affinity than Ku80 peptide, highlighting the importance of Ku70 extended N-terminal regions.

The presence of proline residues in this region appears favorable, as observed in CN-SV40TAg NLS and in optimized NLSs, presumably providing rigidity to the NLS.[19,25] Ku70 NLS displays a shorter flanking sequence than CN-SV40TAg NLS, possibly due to the presence of flexible glycine residues in this region.

Flanking sequences have been reported to be important in other NLSs.[25,26] An extended N-terminal sequence of the peptide makes favorable interactions with Impα also in the structure of the complex between Impα and the peptide corresponding to the NLS from the human androgen receptor (AR).[27] However, in this case, the flanking sequence is unusually curved compared to CN-SV40TAg or nucleoplasmin NLSs, probably due to the presence of a rigid zinc finger motif at the N-terminus of this ligand.

Despite Ku70 NLS having only one extra residue at its N-terminus ordered in the structure compared to Ku80 NLS, and both peptides adopt basically the same conformation (RMSD of $C^{\alpha}$ atoms is 0.59 Å), only the N-termini of Ku70 and CN-SV40TAg NLS peptides make contacts with Impα W273 and P308 residues, which appears to be a common hydrophobic interaction of monopartite NLSs with extended N-termini.

As discussed previously, it has been shown that the deletion of the N-terminal region of Ku70 NLS peptide (K539–K544) leads to the loss of detection of the peptide in the nucleus.[31] In light of the previous results with CN-SV40TAg NLS and other studies with sequences flanking the major NLS site, we suggest that the N-terminally truncated Ku70 peptide would have a lower affinity for the receptor, which may lead to the lack of detection of this peptide in the nucleus.

## The functional role of Ku70 and Ku80 nuclear transport

Although Ku70 and Ku80 have been generally considered to function as a heterodimeric complex, these proteins have unique functions that are independent of each other.[28] Nuclear transport could therefore represent an important regulatory mechanism of the physiological function of Ku proteins.[28] Although the crystal structure of the heterodimeric complex Ku70/Ku80 has been solved, the C-terminal regions of the two molecules that contain the NLSs could not be modeled.[5] The C-terminal regions of both molecules have therefore been solved by nuclear magnetic resonance

techniques, but again, the regions containing the NLSs could not be modeled due to their high flexibility. These results are in complete agreement with the expectations that NLSs are found in flexible regions of the protein, facilitating NLS binding to Impα.

Some studies suggested that Ku70 and Ku80 are transported to the nucleus as a heterodimeric complex.[28] This complex has an asymmetrical topology[29] with the C-terminal region of Ku80 longer than that of Ku70; therefore, the complex could attach to Impα using only the Ku80 NLS. However, the proteins can also be transported into the nucleus independently, providing a powerful means of regulation for their respective functions as a heterodimer or on their own.[10,11]

In conclusion, we confirm that both Ku70 and Ku80 NLSs can bind to the Impα. Moreover, we demonstrate that both NLSs are monopartite cNLSs and that Ku70 NLS is not bipartite as reported previously. Finally, we show that the N-terminal flanking sequence is important for binding to the receptor in a manner analogous to SV40 TAg NLS.

## Materials and Methods

### Synthesis of NLS peptides

The peptide corresponding to Ku70 NLS ($^{537}$EGKVTKRKHDNEGSGSKRPKVE$^{558}$; Ku70 NLS) was synthesized by Auspep (Australia) with purity higher than 95%, and the peptide corresponding to the Ku80 NLS ($^{559}$EDGPTAKKLKTEQ$^{571}$; Ku80 NLS) was synthesized by Proteimax (Brazil) with purity higher than 98%. The peptides contain N- and C-terminal residues additional to the minimal identified NLSs[8,9] to avoid artifactual binding at the termini.[14]

### Protein expression and purification

Hexa-His-tagged truncated *Mus musculus* Impα2, comprising amino acids 70–529 (mImpαΔIBB), was expressed and purified by nickel affinity chromatography as described previously.[30] Additionally, cation-exchange chromatography using a Resource Q (GE Healthcare) column was also performed to increase purity. The protein was eluted using a gradient of sodium chloride followed by dialysis to remove excess salt. Then, the Impα sample was stored in a buffer composed of 20 mM Tris–HCl (pH 8.0), 100 mM sodium chloride and 10 mM DTT at −20 °C. The purity was estimated to be 98% by SDS-PAGE.

### Quantitative binding assay

The glutathione *S*-transferase (GST)-Ku70 and GST-Ku80 NLS fusion proteins were expressed in *Escherichia coli* BL21 (DE3) for 6 h at 37 °C followed by 18 h at 20 °C in an auto-induction medium.[31] The bacterial pellet was resuspended with GST-A buffer (50 mM Tris–HCl, 1 mM DTT and 125 mM NaCl at pH 7.8) and frozen–thawed in three cycles before purification. We added 1 mg/ml lysozyme, 1 mg DNase and 1 mM PMSF (Sigma) to the cell crude extract and mixed them thoroughly. After centrifugation, the soluble fraction was loaded into a GSTrap column (5 ml; GE Healthcare) and then washed with 100-ml GST-A buffer. After washing, the GST-Ku70/Ku80 NLSs were eluted with GST-B buffer (50 mM Tris–HCl, 125 mM NaCl, 1 mM DTT and 10 mM reduced glutathione at pH 7.8). The eluted fractions were pooled and concentrated to perform gel-filtration purification by using a Superdex 200 (20/60) gel-filtration column (GE Healthcare). Pure GST-Ku70/Ku80 NLSs were concentrated by Amicon Ultra centrifugal filter devices (10-kDa cutoff).

The solid-phase binding assays were performed essentially as previously described.[32,33] The assay was carried out on an Immuno MaxiSorp 96-well plate (Thermo Fisher Scientific). The plates were coated with 50 nM GST-Ku70/Ku80 NLS or GST in each well. Binding reactions were carried out for 2 h at 4 °C with 100 μl/well of the indicated amount of S-tagged mImpαΔIBB (0–20 μM, in 2× serial dilution) in binding buffer. After binding and few steps of washing, we incubated the plates in an S-protein horseradish peroxidase conjugate (Novagen). Horseradish peroxidase substrate (100 μg/ml 3,3′,5,5′-tetramethylbenzidine; Sigma) was added for 20 min at room temperature, and then the reaction was stopped by adding an equal volume of 0.5 M $H_2SO_4$. The signal was determined at 450 nm with a Molecular Devices plate reader (SpectraMax 250). The average absorbance values at $OD_{450}$ were determined for GST, GST-Ku70 and GST-Ku80 at each S-tagged mImpαΔIBB concentration. Background absorbance values (without mImpαΔIBB added) were subtracted. The average absorbance values were used to generate binding curves by nonlinear regression using the GraphPad Prism software. The apparent dissociation constants ($K_d$) of mouse ImpαΔIBB binding to GST and GST-Ku70/Ku80 were also calculated using the GraphPad Prism software. The curves were plotted against the serial concentration of mImpαΔIBB and fitted using the following equation:

$$Y = B_{max} \times X / (K_d + X)$$

where $X$ is the concentration of the protein, $B_{max}$ is the maximum specific binding, and $K_d$ is the apparent dissociation constant representing the concentration of the protein yielding half maximal binding.

### Crystallization and crystal structure determination

mImpαΔIBB was concentrated to 20 mg/ml using a Centricon-30 (Millipore) and stored at −20 °C. Crystallization conditions were screened by systematically altering various parameters using, as a starting point, the crystallization conditions that had been successful for other peptide complexes.[12,14] The crystals were obtained using co-crystallization by combining 1 μl of protein solution, 0.5 μl of peptide solution (peptide/protein molar ratios of 4 for Ku70 NLS peptide and 8 for Ku80 NLS peptide) and 1 μl of reservoir solution on a coverslip and suspending the mixture over 0.5 ml of reservoir solution. Single crystals were obtained with a reservoir solution

containing 0.65–0.70 M sodium citrate (pH 6.0) and 10 mM DTT after 15–20 days for both complexes.

X-ray diffraction data were collected using a wavelength of 1.46 Å at a synchrotron radiation source (Laboratório Nacional de Luz Sincrotron, Campinas, Brazil) with a MarMosaic 225 imaging plate detector (Marresearch). The crystals were mounted in nylon loops, transiently soaked in reservoir solution supplemented with 25% glycerol and flash cooled at 100 K in a nitrogen stream (Oxford Nitrogen Cryojet XL; Oxford Cryosystems). Data were processed using the HKL2000 package.[34] The crystals had the symmetry of the space group $P2_12_12_1$ and were isomorphous to other mImpαΔIBB:NLS peptide complexes (Table 1). The structure of the complex with N1N2 NLS [Protein Data Bank (PDB) ID 1PJN[12]], with the NLS peptide omitted, was employed as the starting model for crystallographic refinement. After a rigid-body refinement using the program REFMAC5,[35] inspection of the electron density maps confirmed the presence of the peptides at the major binding site of both models. Rounds of crystallographic refinement (positional and restrained isotropic individual *B*-factor with an overall anisotropic temperature factor and bulk solvent correction) and manual modeling using the program Coot[36] were used to improve the models, considering free *R*-factors. The final model of the mImpαΔIBB:Ku70NLS complex consists of 427 residues of Impα (71–497), 1 peptide ligand (12 residues could be modeled in the major site), 96 water molecules and 1 citrate ion, while the final model of the mImpαΔIBB:Ku80NLS complex consists of 426 residues of Impα (72–496), 1 peptide ligand (11 residues could be modeled at the major site) and 116 water molecules (Table 1). Asn239 is an outlier in the Ramachandran plot as also observed in all other structures of mouse Impα.[14,19] Structure quality was checked with the program PROCHECK,[37] and the contacts were analyzed by the program LIGPLOT.[38]

### PDB accession codes

Coordinates and structure factors from both structures have been deposited in the PDB under accession codes 3RZX (Impα:Ku70NLS) and 3RZ9 (Impα:Ku80NLS).

## Acknowledgements

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil, and Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil. B.K. is a National Health and Medical Research Council Research Fellow, and M. R.M.F. is a Conselho Nacional de Desenvolvimento Científico e Tecnológico Research Fellow. We acknowledge the use of the Laboratório Nacional de Luz Sincrotron (Brazil).

## References

1. Ochi, T., Sibanda, B. L., Wu, Q., Chirgadze, D. Y., Bolanos-Garcia, V. M. & Blundell, T. L. (2010). Structural biology of DNA repair: spatial organisation of the multicomponent complexes of nonhomologous end joining. *J. Nucleic Acids*, doi:10.4061/2010/621695. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938450/?tool=pubmed.
2. Bailey, S. M., Meyne, J., Chen, D. J., Kurimasa, A., Li, G. C., Lehnert, B. E. & Goodwin, E. H. (1999). DNA double-strand break repair proteins are required to cap the ends of mammalian chromosomes. *Proc. Natl Acad. Sci. USA*, **96**, 14899–14904.
3. Giffin, W., Torrance, H., Rodda, D. J., Prefontaine, G. G., Pope, L. & Hache, R. J. (1996). Sequence-specific DNA binding by Ku autoantigen and its effects on transcription. *Nature*, **380**, 265–268.
4. Gu, Y., Sekiguchi, J., Gao, Y., Dikkes, P., Frank, K., Ferguson, D. *et al.* (2000). Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proc. Natl Acad. Sci. USA*, **97**, 2668–2673.
5. Walker, J. R., Corpina, R. A. & Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, **412**, 607–614.
6. Zhang, Z., Zhu, L., Lin, D., Chen, F., Chen, D. J. & Chen, Y. (2001). The three-dimensional structure of the C-terminal DNA-binding domain of human Ku70. *J. Biol. Chem.* **276**, 38231–38236.
7. Zhang, Z., Hu, W., Cano, L., Lee, T. D., Chen, D. J. & Chen, Y. (2004). Solution structure of the C-terminal domain of Ku80 suggests important sites for protein–protein interactions. *Structure*, **12**, 495–502.
8. Koike, M., Ikuta, T., Miyasaka, T. & Shiomi, T. (1999). The nuclear localization signal of the human Ku70 is a variant bipartite type recognized by the two components of nuclear pore-targeting complex. *Exp. Cell Res.* **250**, 401–413.
9. Koike, M., Ikuta, T., Miyasaka, T. & Shiomi, T. (1999). Ku80 can translocate to the nucleus independent of the translocation of Ku70 using its own nuclear localization signal. *Oncogene*, **18**, 7495–7505.
10. Bertinato, J., Schild-Poulter, C. & Hache, R. J. (2001). Nuclear localization of Ku antigen is promoted independently by basic motifs in the Ku70 and Ku80 subunits. *J. Cell Sci.* **114**, 89–99.
11. Koike, M., Shiomi, T. & Koike, A. (2001). Dimerization and nuclear localization of Ku proteins. *J. Biol. Chem.* **276**, 11167–11173.
12. Fontes, M. R., Teh, T., Jans, D., Brinkworth, R. I. & Kobe, B. (2003). Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-α. *J. Biol. Chem.* **278**, 27981–27987.
13. Conti, E. & Kuriyan, J. (2000). Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin α. *Structure*, **8**, 329–338.
14. Fontes, M. R., Teh, T. & Kobe, B. (2000). Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-α. *J. Mol. Biol.* **297**, 1183–1194.
15. Rexach, M. & Blobel, G. (1995). Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell*, **83**, 683–692.

16. Marfori, M., Mynott, A., Ellis, J. J., Mehdi, A. M., Saunders, N. F., Curmi, P. M. *et al.* (2011). Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim. Biophys. Acta*, **1813**, 1562–1577.
17. Kobe, B. (1999). Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin α. *Nat. Struct. Biol.* **6**, 388–397.
18. Catimel, B., Teh, T., Fontes, M. R., Jennings, I. G., Jans, D. A., Howlett, G. J. *et al.* (2001). Biophysical characterization of interactions involving importin-α during nuclear import. *J. Biol. Chem.* **276**, 34189–34198.
19. Fontes, M. R., Teh, T., Toth, G., John, A., Pavo, I., Jans, D. A. & Kobe, B. (2003). Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-α. *Biochem. J.* **375**, 339–349.
20. Yang, S. N., Takeda, A. A., Fontes, M. R., Harris, J. M., Jans, D. A. & Kobe, B. (2010). Probing the specificity of binding to the major nuclear localization sequence-binding site of importin-α using oriented peptide library screening. *J. Biol. Chem.* **285**, 19935–19946.
21. Gorlich, 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.
22. Chelsky, D., Ralph, R. & Jonak, G. (1989). Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol. Cell. Biol.* **9**, 2487–2492.
23. Dingwall, C. & Laskey, R. A. (1991). Nuclear targeting sequences—a consensus? *Trends Biochem. Sci.* **16**, 478–481.
24. Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M. & Yanagawa, H. (2009). Six classes of nuclear localization signals specific to different binding grooves of importin α. *J. Biol. Chem.* **284**, 478–485.
25. Hubner, S., Xiao, C. Y. & Jans, D. A. (1997). The protein kinase CK2 site (Ser111/112) enhances recognition of the simian virus 40 large T-antigen nuclear localization sequence by importin. *J. Biol. Chem.* **272**, 17191–17195.
26. Miyamoto, Y., Imamoto, N., Sekimoto, T., Tachibana, T., Seki, T., Tada, S. *et al.* (1997). Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors. *J. Biol. Chem.* **272**, 26375–26381.
27. Cutress, M. L., Whitaker, H. C., Mills, I. G., Stewart, M. & Neal, D. E. (2008). Structural basis for the nuclear import of the human androgen receptor. *J. Cell Sci.* **121**, 957–968.
28. Koike, M. (2002). Dimerization, translocation and localization of Ku70 and Ku80 proteins. *J. Radiat. Res.* **43**, 223–236.
29. Koike, M. & Koike, A. (2005). The Ku70-binding site of Ku80 is required for the stabilization of Ku70 in the cytoplasm, for the nuclear translocation of Ku80, and for Ku80-dependent DNA repair. *Exp. Cell Res.* **305**, 266–276.
30. Teh, T., Tiganis, T. & Kobe, B. (1999). Crystallization of importin α, the nuclear-import receptor. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 561–563.
31. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif.* **41**, 207–234.
32. Bayliss, R., Littlewood, T., Strawn, L. A., Wente, S. R. & Stewart, M. (2002). GLFG and FxFG nucleoporins bind to overlapping sites on importin-β. *J. Biol. Chem.* **277**, 50597–50606.
33. Ben-Efraim, I. & Gerace, L. (2001). Gradient of increasing affinity of importin β for nucleoporins along the pathway of nuclear import. *J. Cell Biol.* **152**, 411–417.
34. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
35. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
36. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
37. Laskowski, R. A., Rullmannn, 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.
38. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127–134.
39. Chen, M. H., Ben-Efraim, I., Mitrousis, G., Walker-Kopp, N., Sims, P. J. & Cingolani, G. (2005). Phospholipid scramblase 1 contains a nonclassical nuclear localization signal with unique binding site in importin α. *J. Biol. Chem.* **280**, 10599–10606.