Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of importin α, the nuclear-import receptor

Trazel Teh, Tony Tiganis and Bostjan Kobe*

St Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia

Correspondence e-mail: b.kobe@medicine_stvin.unimelb.edu.au

Received 8 July 1998
Accepted 6 October 1998

Crystals of recombinant importin α, the nuclear-import receptor, have been obtained at two different pH conditions by vapour diffusion using sodium citrate as precipitant and dithiothreitol as an additive. At pH 4–5, the crystals have the symmetry of the trigonal space group $P3_121$ or $P3_221$ ($a = b = 78.0$, $c = 255.8$ Å, $\gamma = 120^\circ$); at pH 6–7, the crystals have the symmetry of the orthorhombic space group $P2_12_12_1$ (a = 78.5, b = 89.7, c = 100.5 Å). In both cases, there is probably one molecule of importin α in the asymmetric unit. At least one of the crystal forms diffracts to a resolution higher than 3 Å using the laboratory X-ray source; the crystals are suitable for crystal structure determination.

1. Introduction

The nucleus is the defining feature of a eukaryotic cell. All nuclear proteins are synthesized in the cytoplasm and need to be imported into the nucleus through the nuclear-pore complexes (NPCs). While smaller molecules can freely diffuse through the NPCs, particles from 45 kDa to several million Da (25 nm in diameter) can enter the nucleus *via* active import. Such import is directed by special signals, the best characterized being the classical nuclear-localization sequence (NLS).

The classical NLSs are characterized by one or more clusters of basic amino acids, but are too divergent to yield a consensus sequence. The two major classes of NLSs are the single-cluster NLS, exemplified by the NLS of the simian virus 40 large T-antigen PKKKRKV, and the bipartite NLS, exemplified by the NLS of nucleoplasmin KRPAATKKAGQAKKKK (Dingwall & Laskey, 1991). The targeting efficiency of NLSs can be affected by modifications (phosphorylation), flanking sequences (Rihs & Peters, 1989; Rihs *et al.*, 1991; Jans *et al.*, 1991; Jans & Hubner, 1996) and the presence of multiple NLSs within a protein and the distances between them (Lanford *et al.*, 1986; Roberts *et al.*, 1987; Robbins *et al.*, 1991). Despite the variability among NLSs, the two classes compete for import and are therefore recognized by the same receptor (Michaud & Goldfarb, 1991; Gorlich *et al.*, 1994; Weis *et al.*, 1995). The receptor was identified as the cytoplasmic protein importin (also called karyopherin; Gorlich, Kostka *et al.*, 1995; Imamoto *et al.*, 1995), a heterodimer of α and β subunits; the main NLS-binding site is located on importin α (Weis *et al.*, 1995; Adam & Gerace, 1991), but importin β also contributes to the binding (Xiao *et al.*, 1997). Importin β is responsible for the docking of the importin–substrate complex to the cytoplasmic filaments of the NPC and its translocation through the pore (Gorlich, Vogel *et al.*, 1995; Moore & Blobel, 1994; Weis *et al.*, 1996; Gorlich, Henklein *et al.*, 1996). The transfer through the pore requires GTP hydrolysis by Ran (Ras-related nuclear protein; Moore & Blobel, 1993; Melchior *et al.*, 1993) and is facilitated by nuclear-transport factor 2 (NTF2; Moore & Blobel, 1994; Paschal & Gerace, 1995). In the nucleus, the complex disassembles upon binding of Ran-GTP to importin β, and the import substrate is released into the nucleoplasm (Gorlich, Pante *et al.*, 1996). The importin subunits return to the cytoplasm separately and without the substrate (Weis *et al.*, 1996; Gorlich, Henklein *et al.*, 1996). A schematic diagram illustrating the NLS-dependent import pathway is shown in Fig. 1.

Importin α is the protein responsible for the initial recognition of the import substrates by binding their NLSs. It is a ∼60 kDa protein consisting of two functional domains. A short basic N-terminal domain (the IBB domain) is sufficient for binding to importin β (Weis *et al.*, 1996; Gorlich, Henklein *et al.*, 1996), but the majority of the protein consists of eight ∼43-residue repeated motifs termed armadillo (arm) motifs (Gorlich *et al.*, 1994) that constitute the NLS-binding site (Cortes *et al.*, 1994). Importin β is a ∼95 kDa protein that also contains sequence repeats (∼11) (Gorlich, Kostka *et al.*, 1995). Arm motifs are additionally found in several functionally unrelated proteins (Peifer *et al.*, 1996). The structure of a fragment of one of these proteins, β-catenin, which is involved in the morphogenesis and maintenance of tissue integrity in solid tissues, has recently been determined (Huber *et al.*, 1997). Each repeat consists of three α-helices,

Table 1
Diffraction data-collection statistics.

Reservoir solution	0.9 *M* sodium citrate, 10 m*M* DTT, pH 4.0	0.8 *M* sodium citrate, 10 m*M* DTT, pH 6.0
Space group	$P3_121$ or $P3_221$	$P2_12_12_1$
Unit cell (Å, °)		
a	78.0	78.5
b	78.0	89.7
c	255.8	100.5
γ	120	90
Resolution ranges (Å)	∞–3.49 (3.61–3.49)	∞–2.50 (2.59–2.50)
Observations	59742 (1475)	144864 (7045)
Unique reflections	11325 (801)	24615 (2477)
Multiplicity	5.3 (1.8)	5.9 (2.8)
Completeness (%)	93 (67)	97 (100)
$R_{\rm merge}$† (%)	12.1 (66.7)	12.3 (63.1)
Average $I/\sigma(I)$	5.9 (0.6)	9.4 (1.7)
$I > 3\sigma(I)$ (%)	40 (8)	59 (24)

† $R_{\rm merge} = \sum_{hkl}[\sum_i(|I_{hkl,i} - \langle I_{hkl}\rangle|)]/\sum_{hkl,i}\langle I_{hkl}\rangle$, 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 for $I > -3\sigma(I)$.

with the tandemly repeating units forming a superhelical structure. This places proteins with arm repeats in a group of folds termed coiled-folding domains; these proteins contain short repetitive structural units that do not form stable domains individually but arrange in tandem in a superhelical fashion to form stable structures (Kobe, 1996). Many such proteins are involved in protein–protein interactions. It appears that the elongated non-globular structure formed by the repetitive arrangements in these proteins is capable of presenting a large surface that can form many contacts. The recognition event between importin and the NLS is unusual in that NLSs are very diverse and do not conform to a consensus sequence. To understand the structural basis of this recognition process, we set out to determine the crystal structure of importin α and its complexes with NLSs. Here, we report the crystallization and preliminary X-ray diffraction analysis of recombinant importin α as a first step towards this goal.

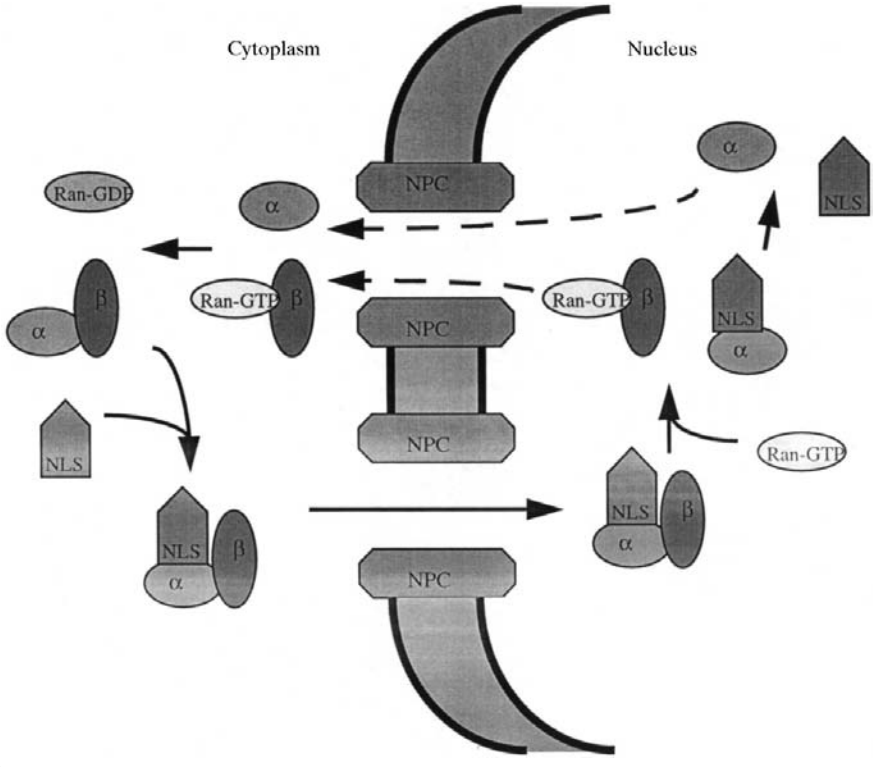

Figure 1
Schematic diagram illustrating the NLS-dependent import pathway. The importin α–β heterodimer (oval objects labelled α and β) binds the cargo protein containing an NLS (pentagonal object labelled NLS) in the cytoplasm and transports it through the nuclear-pore complex (NPC) into the nucleus. Ran-GTP (oval object labelled Ran-GTP) binding to importin β causes the release of the cargo into the nucleoplasm. The importin subunits return to the cytoplasm separately and without the substrate. For simplicity, other factors involved in the pathway such as NTF2, the nuclear-export receptor for importin α and Ran-binding proteins have been omitted from the diagram.

2. Materials and methods

2.1. Expression and purification

Full-length mouse importin α was expressed as a fusion protein containing a hexahistidine tag using the pET30a (Novagen) expression vector (Chi *et al.*, 1996; the molecular mass of the expressed protein is 63 kDa); the recombinant protein is biologically active in *in vitro* nuclear-import assays (Chi *et al.*, 1996; Tiganis *et al.*, 1997). The plasmid was transformed into BL21 (DE3) *E. coli* cells. For expression, the cells were grown at 310 K to an OD (600 nm) of 1.0, induced with 1 m*M* isopropyl thio-β-D-galactoside (IPTG) and grown for 3 h at 303 K. All subsequent purification steps were performed at 277 K. Bacteria were lysed with 1 mg ml^{-1} lysozyme in buffer *A* [20 m*M* HEPES pH 7.0, 500 m*M* NaCl, 1 m*M* $MgCl_2$, 5 m*M* imidazole, 0.1 m*M* tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 1 µg ml^{-1} leupeptin, 1 µg ml^{-1} aprotinin, 1 µg ml^{-1} pepstatin, 1 m*M* phenylmethylsulfonyl fluoride (PMSF)] and cell debris was pelleted by centrifugation at 100000*g* for 30 min. Importin α was affinity-purified from the soluble fraction using Ni^{2+}-agarose (Qiagen). After incubation with the resin for 1 h on the rotating wheel, the resin was washed with buffer *A*, followed by buffer *A* containing 1 *M* NaCl and finally eluted with buffer *A* containing 150 m*M* imidazole. The protein was dialyzed against 20 m*M* Tris–HCl (pH 8.0), 100 m*M* NaCl and 2 m*M* dithiothreitol (DTT). The protein was >95% pure as assessed by SDS–PAGE.

2.2. Crystallization

For crystallization, the protein was concentrated to 12 mg ml^{-1} using Centricon (Amicon) and stored at 253 K. Crystallization conditions were screened by the sparse-matrix approach using the hanging-drop vapour-diffusion technique (McPherson, 1982; Jancarik & Kim, 1991). 1 µl of protein solution was combined with 1 µl of reservoir solution and suspended over 0.5 ml reservoir solution. Small crystals were initially observed in 0.8 *M* sodium citrate and 100 m*M* HEPES (pH 7.5). Subsequent attempts to reproduce these crystals were unsuccessful, until fresh DTT was supplemented in the well solution. Crystals could be grown in the pH range 4–9, with the best crystals obtained using 0.6–0.8 *M* sodium citrate as the precipitant and 100 m*M* citrate buffer (pH 4–6) or HEPES (pH 7) and 5–40 m*M* DTT. DTT appears to be an essential additive and its concentra-

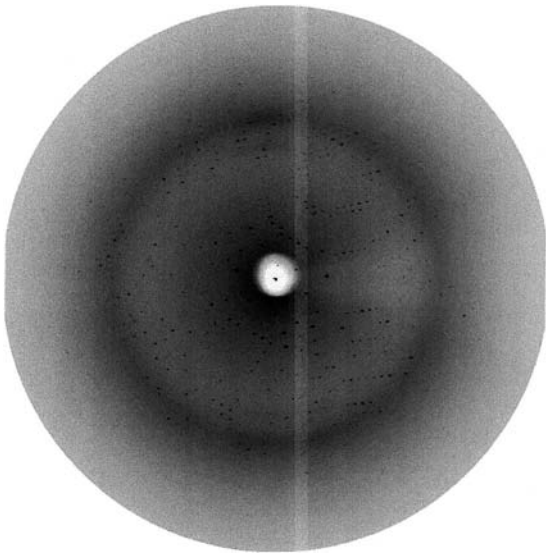

Figure 2
Diffraction pattern from a 1° oscillation image of the orthorhombic crystal described in Table 1. The crystal-to-detector distance was 160 mm; the resolution at the edge of the image is 2.5 Å.

tion has important effects on crystal nucleation, suppressing nucleation both at low and high concentrations. SDS–PAGE of dissolved crystals confirmed that they consisted of importin α.

2.3. Data collection

For X-ray diffraction experiments, crystals were transiently soaked in a solution corresponding to the well solution but supplemented with 20% glycerol and were flash frozen at 100 K in a nitrogen stream (Oxford Cryosystems Cryostream). Data were collected from single crystals using an MAR Research image-plate detector and Cu $K\alpha$ radiation from a Rigaku RU-200 rotating-anode generator. Data were autoindexed and processed with the programs *DENZO* and *SCALEPACK* (Otwinowski, 1993). An oscillation image of an orthorhombic importin α crystal is shown in Fig. 2.

3. Results and discussion

We obtained two crystal forms of importin α, both using sodium citrate as the precipitant and DTT as an additive; at pH 4–5, the crystals are diamond-shaped (dimensions 0.2 × 0.2 × 0.1 mm) and have the symmetry of the trigonal space group $P3_121$ or $P3_221$, while at pH 6–7, the crystals are rod-shaped (dimensions 0.3 × 0.1 × 0.1 mm) and have the symmetry of the orthorhombic space group $P2_12_12_1$ (Table 1). Crystals of both forms appear after a few days and grow to their maximum dimensions within three weeks. In both cases, there is probably one molecule of importin α in the asymmetric unit [the Matthews coefficient V_M (Matthews, 1968) and the solvent content are 3.9 Å^3 Da^{-1} and 68%, respectively, for the trigonal crystal form, and 3.0 Å^3 Da^{-1} and 59%, respectively, for the orthorhombic crystal form]. The orthorhombic crystal form diffracts X-rays using the laboratory source at a resolution higher than 3 Å. The large unit-cell dimensions of the trigonal crystal form have currently precluded data collection beyond 3.5 Å resolution. Data-collection statistics are given in Table 1.

Determination of the crystal structure of importin α will give us a structural reference for understanding the interactions of NLSs and importin β with this protein, and will have implications for understanding the structures and functions of other proteins containing arm repeats. Screening of heavy-atom derivatives is in progress in order to solve the structure of importin α by the multiple isomorphous replacement method. The availability of two crystal forms should facilitate structure determination through density modification by multiple-crystal averaging. Co-crystallization of importin α and peptides corresponding to NLS is also under way in order to determine the structural basis of their interaction.

This work was supported by the Wellcome Trust; B. Kobe is a Wellcome Senior Research Fellow in Medical Science in Australia. T. Tiganis is a National Health and Medical Research Council of Australia C. J. Martin Fellow.

References

Adam, S. A. & Gerace, L. (1991). *Cell*, **66**, 837–847.
Chi, N. C., Adam, E. J., Visser, G. D. & Adam, S. A. (1996). *J. Cell Biol.* **135**, 559–569.
Cortes, P., Ye, Z. S. & Baltimore, D. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 7633–7637.
Dingwall, C. & Laskey, R. A. (1991). *Trends Biochem. Sci.* **16**, 478–481.
Gorlich, D., Henklein, P., Laskey, R. A. & Hartmann, E. (1996). *EMBO J.* **15**, 1810–1817.
Gorlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R. A., Hartmann, E. & Prehn, S. (1995). *Curr. Biol.* **5**, 383–392.
Gorlich, D., Pante, N., Kutay, U., Aebi, U. & Bischoff, F. R. (1996). *EMBO J.* **15**, 5584–5594.
Gorlich, D., Prehn, S., Laskey, R. A. & Hartmann, E. (1994). *Cell*, **79**, 767–778.
Gorlich, D., Vogel, F., Mills, A. D., Hartmann, E. & Laskey, R. A. (1995). *Nature (London)*, **377**, 246–248.
Huber, A. H., Nelson, W. J. & Weis, W. I. (1997). *Cell*, **90**, 871–882.
Imamoto, N., Tachibana, T., Matsubae, M. & Yoneda, Y. (1995). *J. Biol. Chem.* **270**, 8559–8565.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jans, D. A., Ackermann, M. J., Bischoff, J. R., Beach, D. H. & Peters, R. (1991). *J. Cell Biol.* **115**, 1203–1212.
Jans, D. A. & Hubner, S. (1996). *Physiol. Rev.* **76**, 651–685.
Kobe, B. (1996). *Nature Struct. Biol.* **3**, 977–980.
Lanford, R. E., Kanda, P. & Kennedy, R. C. (1986). *Cell*, **46**, 575–582.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: John Wiley and Sons.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Melchior, F., Paschal, B., Evans, J. & Gerace, L. (1993). *J. Cell Biol.* **123**, 1649–1659.
Michaud, N. & Goldfarb, D. S. (1991). *J. Cell Biol.* **112**, 215–223.
Moore, M. S. & Blobel, G. (1993). *Nature (London)*, **365**, 661–663.
Moore, M. S. & Blobel, G. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 10212–10216.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Paschal, B. M. & Gerace, L. (1995). *J. Cell Biol.* **129**, 925–937.
Peifer, M., Berg, S. & Reynolds, A. B. (1996). *Cell*, **76**, 789–791.
Rihs, H. P., Jans, D. A., Fan, H. & Peters, R. (1991). *EMBO J.* **10**, 633–639.
Rihs, H. P. & Peters, R. (1989). *EMBO J.* **8**, 1479–1484.
Robbins, J., Dilworth, S. M., Laskey, R. A. & Dingwall, C. (1991). *Cell*, **64**, 615–623.
Roberts, B. L., Richardson, W. D. & Smith, A. E. (1987). *Cell*, **50**, 465–475.
Tiganis, T., Flint, A. J., Adam, S. A. & Tonks, N. K. (1997). *J. Biol. Chem.* **272**, 21548–21557.
Weis, K., Mattaj, I. W. & Lamond, A. I. (1995). *Science*, **268**, 1049–1053.
Weis, K., Ryder, U. & Lamond, A. I. (1996). *EMBO J.* **15**, 1818–1825.
Xiao, C. Y., Hubner, S. & Jans, D. A. (1997). *J. Biol. Chem.* **272**, 22191–22198.