Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-α

Marcos R. M. FONTES*†, Trazel TEH*‡, Gabor TOTH§, Anna JOHN‖, Imre PAVO§, David A. JANS‖¶ and Bostjan KOBE*‡[1]

*Structural Biology Laboratory, St. Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia, †Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista, C.P. 510, 18618-000, Botucatu, SP, Brazil, ‡Department of Biochemistry and Molecular Biology, Institute for Molecular Bioscience, ARC Special Research Centre for Functional and Applied Genomics, and Cooperative Research Centre for Chronic Inflammatory Diseases, University of Queensland, Brisbane, Queensland 4072, Australia, §Institute for Medical Chemistry, Szeged Medical University, Szeged, Hungary, ‖Nuclear Signalling Laboratory, Division for Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia, and ¶Department for Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3168, Australia

The nuclear import of simian-virus-40 large T-antigen (tumour antigen) is enhanced via phosphorylation by the protein kinase CK2 at Ser^{112} in the vicinity of the NLS (nuclear localization sequence). To determine the structural basis of the effect of the sequences flanking the basic cluster KKKRK, and the effect of phosphorylation on the recognition of the NLS by the nuclear import factor importin-α (Impα), we co-crystallized non-autoinhibited Impα with peptides corresponding to the phosphorylated and non-phosphorylated forms of the NLS, and determined the crystal structures of the complexes. The structures show that the amino acids N-terminally flanking the basic cluster make specific contacts with the receptor that are distinct from the interactions between bipartite NLSs and Impα. We confirm the important role of flanking sequences using binding assays. Unexpectedly, the regions of the peptides containing the phosphorylation site do not make specific contacts with the receptor. Binding assays confirm that phosphorylation does not increase the affinity of the T-antigen NLS to Impα. We conclude that the sequences flanking the basic clusters in NLSs play a crucial role in nuclear import by modulating the recognition of the NLS by Impα, whereas phosphorylation of the T-antigen enhances nuclear import by a mechanism that does not involve a direct interaction of the phosphorylated residue with Impα.

Key words: importin-α (karyopherin-α), nuclear localization sequence recognition (NLS recognition), phosphorylation, simian-virus-40 (SV40) large tumour-antigen nuclear localization sequence, X-ray crystal structure.

INTRODUCTION

Nucleocytoplasmic transport occurs through nuclear pore complexes, large supramolecular structures that penetrate the double lipid layer of the nuclear envelope. Most macromolecules require an active, signal-mediated transport process that enables the passage of particles up to 25 nm in diameter (approx. 25 MDa). The first and best-characterized nuclear targeting signals are the conventional NLSs (nuclear localization sequences) that contain one or more clusters of basic amino acids [1]. The NLSs fall into two distinct classes: monopartite NLSs, containing a single cluster of basic amino acids, and bipartite NLSs, containing two basic clusters connected by spacers of variable length.

Despite their variability, the conventional basic NLSs are recognized by the same receptor protein, termed importin or karyopherin, a heterodimer of α and β subunits (for recent reviews see [2–4]). Importin-α (Impα) contains the NLS-binding site, and importin-β (Impβ) is responsible for the translocation of the importin–substrate complex through the nuclear pore complex. Once inside the nucleus, Impβ binds to Ran-GTP (where Ran stands for Ras-related nuclear protein), which causes the dissociation of the import complex and the autoinhibition of Impα. The importin subunits return to the cytoplasm separately without the import complex. The directionality of the nuclear import is conferred by an asymmetric distribution of the GTP- and GDP-bound forms of Ran between the cytoplasm and the nucleus. This distribution is, in turn, controlled by various Ran-binding regulatory proteins.

Impα consists of two structural and functional domains: a short basic N-terminal IBB domain (Impβ-binding domain) [5–7] and a large NLS-binding domain built of armadillo repeats (referred to as Arm repeats hereafter) [8]. The structural basis for the recognition of monopartite and bipartite NLSs by Impα has been studied crystallographically in yeast and mouse Impα proteins [9–11]. Monopartite NLSs bind at a major site located between the first and fourth Arm repeats, and additionally at a minor site spanning repeats 4–8. Bipartite NLSs span the two binding sites, each site recognizing one of the basic clusters [10,11]. The linker sequence between the two basic clusters makes few contacts with Impα, consistent with its tolerance to mutations. The structure of fl-Impα (full-length Impα), in the absence of Impβ, indicated that the IBB domain occupies the major NLS-binding site of Impα, thus autoinhibiting Impα from NLS binding [12]. This is supported by studies indicating significantly higher NLS-binding affinity by Impα/Impβ compared with Impα alone [13–22]. Several studies indicate that the affinity of the importin–targeting sequence interaction is a critical parameter in determining transport efficiency [15,19,23,24].

Abbreviations used: AcN, acetonitrile; Arm repeat, armadillo repeat; T-antigen, tumour antigen; T-Ag, simian-virus-40 large T-antigen; CN peptide, GPGSDDEAAADAQHAAPPKKKRKVG (corresponding to residues 109–133 of T-Ag with modifications); Imp, importin; IBB domain, Impβ-binding domain; fl-Impα, full-length Impα; NLS, nuclear localization sequence; PAH, phenylalanine hydroxylase; TFA, trifluoroacetic acid.

[1] To whom correspondence should be addressed (e-mail b.kobe@uq.edu.au).

The co-ordinates of pCN and CN peptide complexes have been deposited in the Protein Data Bank with the accession codes 1Q1S and 1Q1T respectively.

Table 1 Peptides used in the present study

Peptide	Comprises	Sequence*
CN†	T-Ag residues 110–132 (non-phosphorylated CK2 site + basic cluster)	GPGSDDEAAADAQHAAPPKKKRKVG
Cn†	T-Ag residues 110–132, K128T	GPGSDDEAAADAQHAAPPKtKRKVG
pCN‡	T-Ag residues 110–132, Ser^{112} pre-phosphorylated	GPGpSDDEAAADAQHAAPPKTKRKVG
dCN	T-Ag residues 110–132, S112D	GPGDDDEAAADAQHAAPPKTKRKVG
N	T-Ag amino acids 126–132 (basic cluster)	PKKKRKV
bipN1N2	N1N2 amino acids 535–555 (N1N2 bipartite NLS)	CGRKKRKKTEEESPLKDKAKKSKGY

* Single-letter amino acid code is used; residues with small letters indicate mutations destroying nuclear-targeting activity. Heterologous residues are present at the termini (see the Experimental section).

† The CN and Cn peptides have been described previously [16].

‡ Chemically pre-phosphorylated at Ser^{112} as described in the Experimental section.

Although one basic cluster may be sufficient for nuclear import, as demonstrated by monopartite NLSs, the multitude of bipartite NLSs suggests that it must be beneficial to contain additional interacting regions. Structural analyses of bipartite NLSs bound to Impα show that, whereas the essential contacts are provided by the two basic clusters, the linker sequences also contribute to binding [10,11,24a]. Similarly, the sequences flanking the basic cluster in monopartite NLSs contribute to the binding affinity [16,23,25]. Phosphorylation in the vicinity of the NLS also plays a role in regulating NLS-dependent nuclear protein import [24,26,27]. Phosphorylation events in various phosphorylation-regulated NLSs can have either an enhancing or an inhibitory effect on nuclear import. The best-characterized example is the NLS of T-Ag (simian-virus-40 large tumour antigen) that contains several phosphorylation sites N-terminal to the basic cluster (the basic cluster spans residues 127–131). Phosphorylation by the protein kinase CK2 at Ser^{111}/Ser^{112} accelerates the nuclear import rate 50-fold [16,28], whereas phosphorylation by the double-stranded DNA-dependent protein kinase at Ser^{120} further enhances import [23]. Protein kinase CK2 and double-stranded DNA-dependent protein kinase sites synergize both at the level of phosphorylation and enhancement of T-Ag nuclear import. On the basis of the available data, it has been proposed that the phosphorylated residues may be involved in a direct interaction with Impα [23]. A similar mechanism may operate in the *Drosophila melanogaster* transcription factor Dorsal, whose recognition by Impα/Impβ is enhanced via phosphorylation by cAMP-dependent protein kinase at a site N-terminally flanking the basic cluster [18]. Phosphorylation plays a key regulatory role in most cellular processes, yet the consequences of phosphorylation have been structurally elucidated in very few systems [29].

In the present study, we investigated the role of flanking sequences and phosphorylation in the recognition of the T-Ag NLS by Impα. We co-crystallized a non-autoinhibited mouse Impα with both the non-phosphorylated and phosphorylated versions of the peptide corresponding to the T-Ag NLS, $G^{109}PG\mathbf{S^{112}}$DDEAAADAQHAAPPKKKRKVG[133] (phosphorylation site is indicated in boldface; see Table 1 for the description of the peptides), and complemented the structural studies using ELISA-based binding assays. The structures show that the sequence N-terminally flanking the basic cluster makes several contacts with the receptor that are distinct from the interaction of bipartite NLSs with Impα [10,11]. Binding assays confirm the important role of this region in binding. Unexpectedly, the regions of the peptides containing the phosphorylation site do not make specific contacts with the receptor. The accompanying binding studies confirm that direct recognition of the phosphorylated residue by Impα is not responsible for enhancement of nuclear import. We conclude that the sequences flanking the basic clusters in NLSs play a crucial role in nuclear import by modulating the recognition of the NLS by Impα, but T-Ag phosphorylation enhances nuclear import by a mechanism that does not involve the direct interaction of the phosphorylated residue with the receptor.

EXPERIMENTAL

Materials

Dichloromethane, dimethylformamide, di-isopropylethylamine, methanol, t-butanol, TFA (trifluoroacetic acid) and HPLC-grade AcN (acetonitrile) were purchased from Merck (Darmstadt, Germany) and were used without further purification. Dicyclohexylcarbodi-imide was purchased from Fluka (Buchs, Switzerland). Protected amino acid derivatives and *p*-alcoxybenzyl alcohol resins were obtained from Bachem (Torrance, CA, U.S.A.). 1-Hydroxybenzotriazole, piperidine, PCl3, *N*,*N*-diethylamine, t-butylperoxide, *p*-cresol and dimethyl sulphide were purchased from Aldrich (Steinheim, Germany). Hydrogen fluoride was obtained from UCAR (Olen, Belgium).

Peptide synthesis

Non-phosphorylated T-Ag derivative peptides were synthesized using an ABI 433 A synthesizer [30,31]. Peptide PKKKRKV (N peptide; see Table 1 for the description of the peptides) corresponds to T-Ag residues 126–132, whereas peptide GPGSDDEAAADAQHAAPPKKKRKVG (referred to as CN peptide hereafter) corresponds to residues 110–132, additionally including the CK2 site (Ser^{112}), with residue 111 replaced by glycine and residues 117, 120, 123 and 124 replaced by alanine (to remove all phosphorylation sites other than the CK2 site Ser^{112}, and thereby ensure that the chemical phosphorylation procedure was unique to the Ser^{112} position); N- and C-terminally flanking glycine residues are included. Substitution of Thr^{117}, Ser^{120}, Ser^{123} and Thr^{124} by alanine residues has no significant effect on importin recognition relative to the wild-type sequence [16]. Either of the residues 111 or 112 can be phosphorylated by protein kinase CK2, but Ser^{112} is the preferred phosphorylation site [28]. The peptide GPGSDDEAAADAQHAAPPKTKRKVG (referred to as Cn peptide hereafter) is identical with the CN peptide, except that it contains Thr in place of Lys^{128}. The peptide GPGDDDEAAADAQHAAPPKTKRKVG (referred to as dCN peptide hereafter) is identical with the CN peptide, except that it contains Asp in place of Ser^{112}. The peptide CGRKKRKKTEEESPLKDKAKKSKGY (bipN1N2 peptide) contains *Xenopus laevis* chromatin assembly factor N1N2 (amino acids 535–555), encompassing the bipartite NLS with two heterologous residues at each terminus.

The chemically phosphorylated peptide GPGpSDDEAAADAQHAAPPKKKRKVG (referred to as pCN peptide hereafter) was synthesized using the SPPS–Fmoc methodology (where SPPS stands for solid-phase peptide synthesis and Fmoc for fluoren-9-ylmethoxycarbonyl) [32]. Side-chain-protecting groups were as follows: Arg(Pmc), Glu(OBut), Asp(OBut), His(Trt) and Lys(Boc) (where Pmc stands for 2,2,5,7,8-pentamethylcroman-6-sulphonyl, OBut for t-butyl ester, Trt for trityl and Boc for t-butoxycarbonyl), with the alcoholic hydroxy groups of the serine residue unprotected. The peptides were prepared on alcoxybenzyl alcohol resin (0.6 mmol/g) using an ABI 433 A synthesizer and a modified synthesis method. Couplings were performed with diphenylcarbamoyl chloride, with the exception of Asn, Gln and

Arg, which were incorporated as their 1-hydroxybenzotriazole esters. Amino acid incorporation was monitored using the ninhydrin test. After the incorporation of the serine residue, these hydroxy groups were phosphitylated using di-t-butyl-*N*,*N*-diethyl-phosphoramidite [33]. The di-t-butyl-*N*,*N*-diethyl-phosphoramidite was prepared from PCl_3, diethylamine and t-butanol [34]. After the oxidation of the phosphite into phosphate using t-butylperoxide, the elongation of the peptide chain was completed. The crude phosphopeptide detached from the resin using a TFA/dichloromethane/anisole mixture (31:15:4, by vol), at 0 °C for 1 h, and the free peptides were precipitated with diethyl ether, filtered, dissolved in water and freeze-dried. The crude peptides were purified by reversed-phase HPLC on a Lichrosorb RP-18 10 μm column (16 mm × 250 mm), using a Knauer HPLC apparatus (Knauer, Berlin, Germany), with the following solvent systems: A, 0.1 % TFA in water; B, 0.1 % TFA + 80 % AcN in water; gradient: 0–60 % B in 1.5 h at a flow rate of 3 ml/min. Appropriate fractions were pooled and freeze-dried. Purity was assessed by analytical reversed-phase HPLC. The analytical HPLC investigations were performed on a Vydac C_{18} column (4 mm × 250 mm, 218TP54; Vydac, Hesperia, CA, U.S.A.), at a flow rate of 0.8 ml/min with detection at 220 nm wavelength and with the following solvent systems: A, 0.1 % TFA in water; B, 0.1 % TFA + 80 % AcN in water; gradient: 0–30 % B in 20 min. MS of the peptide was performed on a Finnigan TSQ 7000 tandem quadrupole mass spectrometer equipped with an electrospray ion source. The pCN peptide indicated a molecular mass of 2.5806 kDa, consistent with the theoretical value of 2.58078 kDa.

Expression and purification of N-terminally truncated Impα

N-terminally truncated (non-autoinhibited) mouse Impα (α2 isoform [35]) lacking 69 N-terminal residues [Impα(70–529)] was expressed recombinantly as a fusion protein containing a His_6 tag in *Escherichia coli* [11]. Full-length and N-terminally truncated mouse Impα2 was expressed in bacteria as glutathione S-transferase fusion proteins [16,23].

Crystallization

For crystallization, Impα(70–529) was concentrated to 18.8 mg/ml using a Centricon-30 (Millipore, Billerica, MA, U.S.A.) and stored at −20 °C. Crystallization conditions were screened by altering systematically various parameters using the crystallization conditions for other peptide complexes [11] as a starting point. The crystals of both complexes (rod-shaped, 0.4 mm × 0.15 mm × 0.1 mm for the pCN peptide complex and 0.3 mm × 0.1 mm × 0.07 mm for the CN peptide complex) were obtained by co-crystallization, combining 1 μl of protein solution, 0.7 μl of peptide solution (1.7 mg/ml, with peptide/protein ratios of 3.5 and 2.5 for the non-phosphorylated and phosphorylated peptides respectively) and 1 μl of reservoir solution. The crystals were suspended in 0.5 ml of reservoir solution containing 0.7–0.8 M sodium citrate as the precipitant in 100 mM Hepes (pH range 6.0–6.5) containing 10 mM dithiothreitol.

Collection of diffraction data

Crystals of native Impα(70–529) and the peptide complexes exhibit orthorhombic symmetry (space group $P2_12_12_1$; Table 2). Diffraction data were collected from single crystals transiently soaked in a solution corresponding to the reservoir solution but supplemented with 23 % (v/v) glycerol and flash-frozen at 100 K in a stream of nitrogen (Oxford Cryosystems, Oxford, U.K.), using an MAR-Research image plate detector (plate diameter 345 mm) and synchrotron radiation ($\lambda = 0.97$ Å; 1 Å = 0.1 nm) at Stanford Synchrotron Radiation Laboratory (Menlo Park, CA, U.S.A.). Data were autoindexed and processed with the HKL suite [36] (Table 2).

Table 2 Structure determination

Statistics	CN peptide	pCN peptide
Diffraction data		
Unit cell dimensions (Å)		
a	79.2	78.8
b	89.3	89.7
c	100.3	99.7
Resolution (Å)	99–2.5 (2.59–2.5)*	99–2.3 (2.38–2.3)*
Observations	443 936	272 363
Unique reflections	25 902	32 147
Completeness (%)	97.2 (88.6)*	99.4 (95.8)*
R_{merge}† (%)	7.8 (25.5)*	5.9 (23.4)*
Average $I/\sigma(I)$	19.8 (3.9)*	25.7 (5.3)*
Refinement		
Resolution (Å)	30–2.5	30–2.3
Number of reflections ($F > 0$)	24 804	31 890
Completeness (%)	98.2	99.4
R_{cryst}‡ (%)	20.2	19.8
R_{free}§ (%)	23.3	22.1
Number of non-hydrogen atoms		
Protein	3256	3300
Peptide	154	161
Water	302	305
Mean B-factor (Å^2)	39.0	39.9
R.m.s deviations from ideal values‖		
Bond lengths (Å)	0.0067	0.008
Bond angles (°)	1.6	1.7
Ramachandran plot¶: residues in most favoured (disallowed) regions (%)	94.3 (0.3)	91.7 (0.3)
Co-ordinate error (Å)‖:		
Luzzati plot (cross-validated Luzzati plot)	0.28 (0.33)	0.25 (0.29)
SIGMAA (cross-validated SIGMAA)	0.24 (0.31)	0.20 (0.26)

* Numbers in parentheses are for the highest-resolution shell.

† $R_{merge} = \Sigma_{hkl}(\Sigma_i(|I_{hkl,i} - \langle I_{hkl}\rangle|))/\Sigma_{hkl,i}\langle I_{hkl}\rangle$, 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 $I > -3\sigma(I)$.

‡ $R_{cryst} = \Sigma_{hkl}(||F_{obs,hkl}|-|F_{calc,hkl}||)/|F_{obs,hkl}|$, where $|F_{obs,hkl}|$ and $|F_{calc,hkl}|$ are the observed and calculated structure factor amplitudes.

§ R_{free} is equivalent to R_{cryst}, but calculated with reflections (10 %) omitted from the refinement process.

‖ Calculated with the program CNS [37].

¶ Calculated with the program PROCHECK [40].

Structure determination and refinement

Crystals of peptide complexes were highly isomorphic with the crystals of fl-Impα [12] and peptide complexes obtained previously [11]. Therefore the structure of mouse Impα [12] (Protein Data Bank entry 1IAL) with residues 44–54 omitted was used as a starting model for crystallographic refinement. Electron density maps were inspected for the presence of the peptide after rigid-body refinement using the program CNS [37] [pCN peptide–Impα(70–529) complex, $R_{cryst} = 32.3$ %, $R_{free} = 32.4$ %, 10–4 Å resolution; CN peptide–Impα(70–529) complex, $R_{cryst} = 28.2$ %, $R_{free} = 31.4$ %, 10–4 Å resolution; see Table 2 for an explanation of *R*-factors]. Electron density maps calculated with coefficients $3|F_{obs}| - 2|F_{calc}|$ and simulated annealing omit maps (Figure 1) calculated with analogous coefficients were generally used. The model was improved, as judged by the free *R*-factor [38], through rounds of crystallographic refinement (positional and

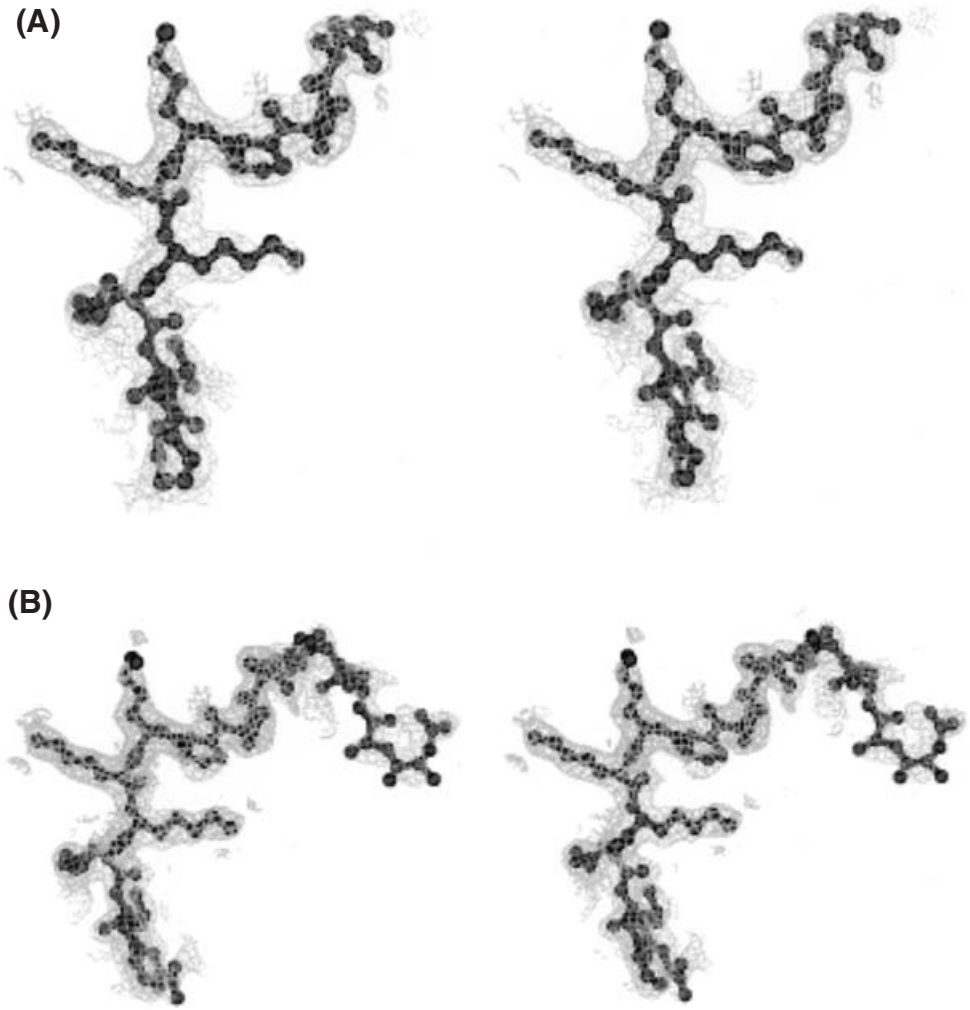

Figure 1 Structure determination

(**A**) Stereo view of the electron density (drawn with the program BOBSCRIPT [48]) in the region of the pCN peptide bound to the major binding site of Impα(70–529). All peptide residues were omitted from the model and simulated annealing was run with the starting temperature of 1000 K. The electron density map was calculated with coefficients $3|F_{obs}| - 2|F_{calc}|$ and for resolution between 40 and 2.3 Å, and contoured at 1.2 S.D. The refined model of the peptide is superimposed. (**B**) Stereoview of the electron density (contoured at 1.5 S.D.) in the region of the CN peptide bound to the major binding site of Impα(70–529), shown as in (**A**).

restrained isotropic individual *B*-factor refinement, with an overall anisotropic temperature factor and bulk solvent correction), and manual rebuilding (the program O [39]). Solvent molecules were added using the program CNS [37]. Asn^{239} is an outlier in the Ramachandran plot, as also observed in all other structures of mouse Impα [11,12]. Pro^{242} is a *cis*-proline residue. The final models comprise 434 Impα residues (residues 64–497) and 20 peptide residues for the pCN peptide–Impα(70–529) complex, and 427 Impα residues (residues 70–496) and 19 peptide residues for the CN peptide–Impα(70–529) complex. Residues 64–69 of Impα in the pCN peptide–Impα(70–529) complex are part of the His tag and associated linker. The results of the structural determination are shown in Table 2. It should be noted that the crystals were grown in the presence of sodium citrate at high ionic strength, which may affect polar interactions between the peptides and Impα. The effect does not appear to be significant based on our binding studies and the similarity of the structures of NLS peptides bound to mouse and yeast Impα crystallized in different conditions [9–11]. The co-ordinates have been deposited in the Protein Data Bank (codes 1Q1S and 1Q1T for the pCN and CN peptide complexes respectively).

Structural analysis

The quality of the models was checked with the program PROCHECK [40]. The contacts were analysed with the program CONTACT, and the buried surface areas were calculated using the program CNS [37].

ELISA-based binding assay

Peptide binding by importin was quantified using an ELISA-based assay [16,23]. Peptides were coated on to microtitre plates and then incubated with increasing concentrations of Impα(70–529), fl-Impα or pre-formed fl-Impα/Impβ heterodimer. Importin binding was then measured using a glutathione S-transferase-specific primary antibody and a secondary antibody conjugated with alkaline phosphatase. The colorimetric reaction after the addition of *p*-nitrophenyl phosphate substrate (Sigma) was followed at 405 nm for 90 min on a microtitre plate reader (Molecular Devices, Sunnyvale, CA, U.S.A.). Values were corrected against background absorbance at 0 min and against wells incubated without importin. In some cases, peptides were coated in the presence of 0.8 M citrate (crystallization solution).

RESULTS AND DISCUSSION

Structural determination

The NLS peptides were co-crystallized with the N-terminally truncated Impα [Impα(70–529)], which lacks residues 1–69 that are responsible for autoinhibition. The co-crystals with both peptides grew under similar conditions and isomorphic to other mouse Impα crystals [11,12]. Electron density maps based on the Impα model, following rigid-body refinement, clearly showed electron density corresponding to the peptides. The structures were refined at 2.3 and 2.5 Å resolution for the phosphorylated and non-phosphorylated peptide complexes respectively (Table 2). Residues 119–133 and 123–133 of the pCN and CN peptides respectively could be interpreted unambiguously in the electron density map at the major site, and additionally residues 128–132 and 127–133 have been modelled for the minor site of the pCN and CN peptides respectively (Figure 1).

Binding of NLS peptides to Impα

Impα(70–529) forms a single elongated domain built from ten Arm structural repeats, each containing three α-helices (H1, H2 and H3) connected by loops (Figure 2). The structure of Impα is comparable with the crystal structures of fl-Impα [r.m.s. deviations of Cα atoms of residues 7–496 are 0.27 Å between the pCN and CN peptide–Impα(70–529) complexes, and 0.43 and 0.30 Å between fl-Impα and the pCN and CN peptide–Impα(70–529) complexes respectively].

Both the pCN and CN peptides bind at two sites on the surface of Impα(70–529) molecule (Figure 2). At the major site, the electron density is present for 15 peptide residues (residues 119–133; average *B* factor, 50.8 Å^2) and 11 peptide residues (residues 123–133; average *B* factor, 36.6 Å^2) of the pCN and CN peptides respectively (Figure 1) [the average *B* factor for Impα(70–529) is 37.9 and 37.4 Å^2 for the pCN and CN peptide complexes respectively]. At the minor site, the electron density is present for five peptide residues (residues 128–132; average *B* factor, 72.3 Å^2) and seven peptide residues (residues 127–133; average *B* factor, 60.7 Å^2) for the pCN and CN peptides respectively (Figure 1). The parts of the pCN and CN peptides modelled in both complexes have similar structures. After the superposition of the equivalent Cα atoms of the pCN and CN peptides, the r.m.s. deviation of residues 123–133 at the major binding site was 0.91 Å, and that for residues 128–132 at the minor binding site was 0.36 Å (Figure 3).

There is a surface area of 2014 Å^2 buried between the pCN peptide and Impα(70–529) at the major site (239 contacts <4 Å), and 1009 Å^2 at the minor site (74 contacts <4 Å). The corresponding numbers for the CN peptide are 1583 Å^2 of surface area buried at the major site (171 contacts <4 Å), and 1213 Å^2 at the minor site (103 contacts <4 Å). In both complexes, residues 125–131 at the major site have *B* factors below the average number for the entire structure, but the mobility starts increasing towards both termini.

The general features of peptide binding are not different from those observed in other available structures of Impα–NLS peptide complexes [9–11]. As discussed previously [9–11], only the binding at the major site is likely to be physiologically relevant. The peptides bind in an extended conformation, with the chain running antiparallel to the direction of the Arm repeat superhelix. The base of the groove that contains the binding sites is formed mainly by the H3 helices of the Arm repeats, which carry some residues conserved among the repeats, including the tryptophan and asparagine residues at the 3rd and 4th turns respectively in H3 helices of the Arm repeats [10,11]. Residues 126–132 superimpose closely with the corresponding residues of the shorter T-Ag peptide (N peptide; residues 126–132; r.m.s. deviation of Cα atoms, 0.26 Å) [11]. Following the chain from Pro^{126} towards the N-terminus, the peptide turns initially in the same direction as for the peptide corresponding to the c-Myc NLS bound to yeast Impα [10], and continues further towards the edge of the groove defined by the N-terminal end of the H3 helices. The protein residues involved in the interaction with this region of the peptide come mainly from Arm repeats 5 and 6, with additional interactions of Impα Arg^{238} (Arm 4) with the peptide residue Pro^{125} and Impα Asn^{350} (Arm 7) with the peptide residue His^{122}. The peptide residues 119–122 interact exclusively with the sequence 305–308 at the N-terminal end of helix H3 of Arm 6. There are some close interactions with the protein formed at distances <3.5 Å in this region of the peptide, with most of the interactions in this region being hydrophobic (Figure 4).

Recognition of the NLS by autoinhibited and non-autoinhibited forms of Impα is enhanced by the N-terminal flanking residues

To confirm the role of the flanking region, N-terminal to the basic cluster, in mediating the binding to Impα as indicated by the structural results, we used an ELISA-based binding assay to compare the T-Ag NLS peptide-binding properties of autoinhibited (fl-Impα) and non-autoinhibited Impα [Impα(70–529)] in the absence or presence of Impβ (Figure 5A, Table 3). The binding to T-Ag NLS peptides with (CN peptide) and without (N peptide; Table 1) T-Ag residues 112–125 were compared directly. The CN peptide bound with approx. 4-fold higher affinity than peptide N to fl-Impα in either the presence or absence of Impβ (Table 3, Figure 5A). For Impα(70–529), the longer peptide showed almost 2-fold higher binding affinity, whereas the maximal binding level (B_{max}) was approx. 5-fold higher. None of the importins recognized the derivative peptide with a mutation in the basic cluster (Cn peptide; Table 1) with high affinity, exhibiting 7–13-fold higher apparent dissociation constant (K_d) values when compared with the CN peptide and >2-fold decrease in B_{max} value. Using a peptide corresponding to the bipartite NLS from the phosphoprotein N1N2 (bipN1N2; Table 1), fl-Impα again exhibited the lowest binding affinity, with K_d values at least 4-fold higher than those of Impα(70–529) or fl-Impα/Impβ. Impβ had no effect on the NLS-binding properties of Impα(70–529), consistent with its inability to bind Impα(70–529) which lacks the IBB domain. Our results show that the non-autoinhibited Impα(70–529) resembles fl-Impα/Impβ in its NLS-binding properties, whereas fl-Impα shows decreased NLS-binding abilities consistent with its autoinhibited state [12].

The crystallization solution used for structure determination contained 0.8 M sodium citrate. To assess the possible effects of this solution on the recognition of NLSs by importin, we repeated the binding assays in the presence of 0.8 M citrate (Figure 5B, Table 3). For Impα(70–529) and fl-Impα/Impβ, the presence of citrate resulted in a significant, but not dramatic, increase in K_d values, and a decrease in B_{max} values for binding to most of the peptides. One exception was the N-peptide–fl-Impα/Impβ combination, where the K_d value decreased slightly in the presence of citrate. In contrast, the presence of citrate increased the NLS-binding abilities of fl-Impα, whereas the K_d values decreased by approx. 2-fold. These results suggest that in the presence of citrate, fl-Impα exhibits NLS-binding properties more similar to those of Impα(70–529) and fl-Impα/Impβ. Importantly, the binding properties are not altered dramatically by citrate, thus validating the conclusions drawn from the structural studies.

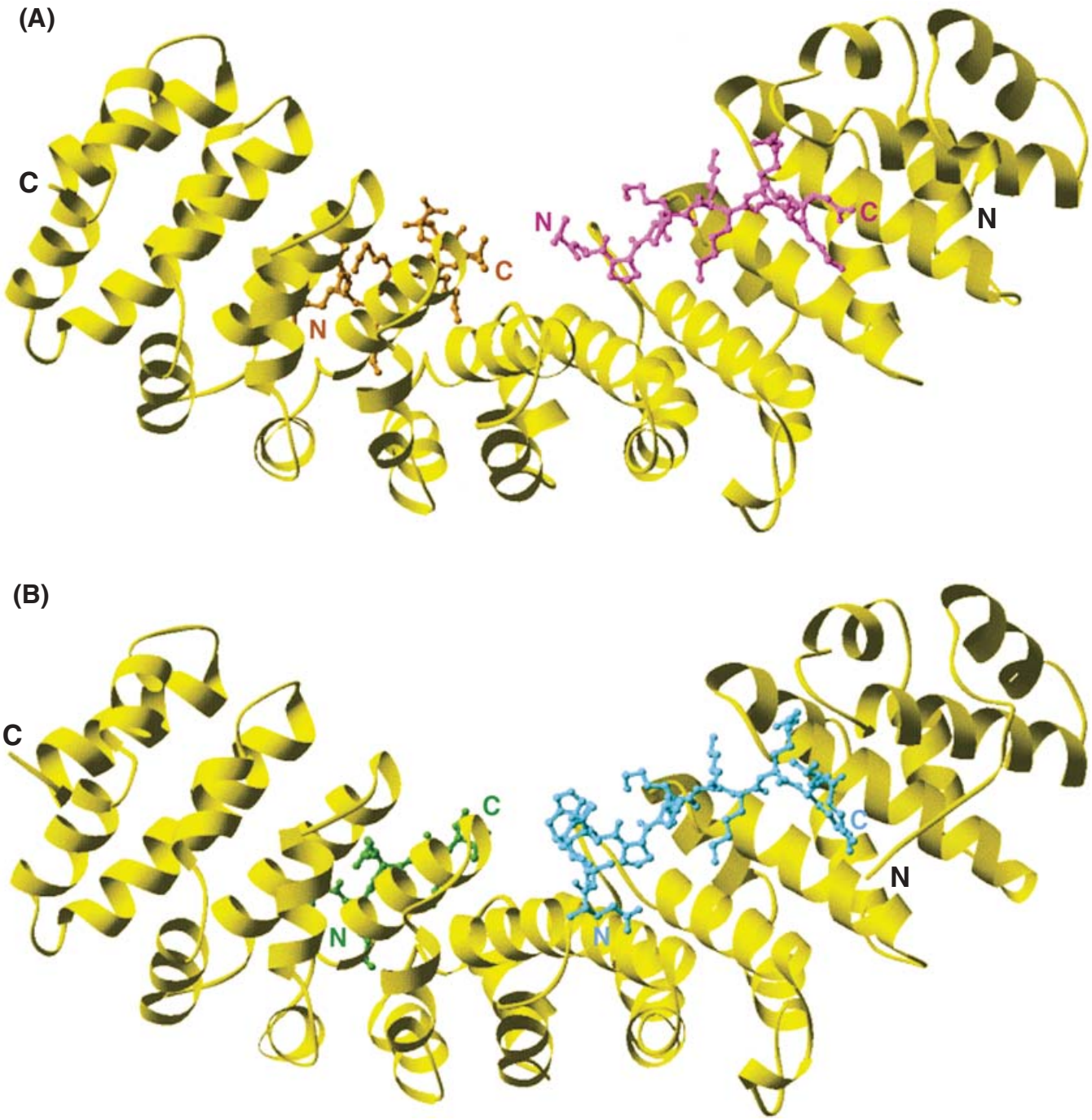

Figure 2 Structures of Impα–peptide complexes

(**A**) Structure of the pCN peptide–Impα(70–529) complex. Impα is shown as a ribbon (coloured yellow; drawn with the program RIBBONS [49]). The superhelical axis of the repetitive part of the molecule is approximately horizontal. The two NLS peptides are shown in a ball-and-stick representation; the peptide bound to the major site is coloured cyan, and the peptide bound to the minor site is coloured green. (**B**) Structure of the CN peptide–Impα(70–529) complex, shown as in (**A**). The peptide bound to the major site is coloured magenta, and the peptide bound to the minor site is coloured gold.

Phosphorylation does not enhance the recognition of the NLS by importin

To assess the effect of phosphorylation on the binding of T-Ag NLS peptides to Impα, we compared directly the binding of the phosphorylated pCN peptide and the non-phosphorylated CN peptide to the autoinhibited and non-autoinhibited forms of Impα (Figure 5C, Table 3). Surprisingly, the pCN peptide bound with approx. 3–4-fold lower affinity to fl-Impα/Impβ, Impα(70–529) and fl-Impα, and the B_{max} value was decreased by approx. 20%. We also performed analyses using the dCN peptide, which is identical with the CN peptide, except that it contains the negatively charged aspartate residue in place of Ser^{112}. This peptide was also less well recognized by importins than the CN peptide (Figure 5C,

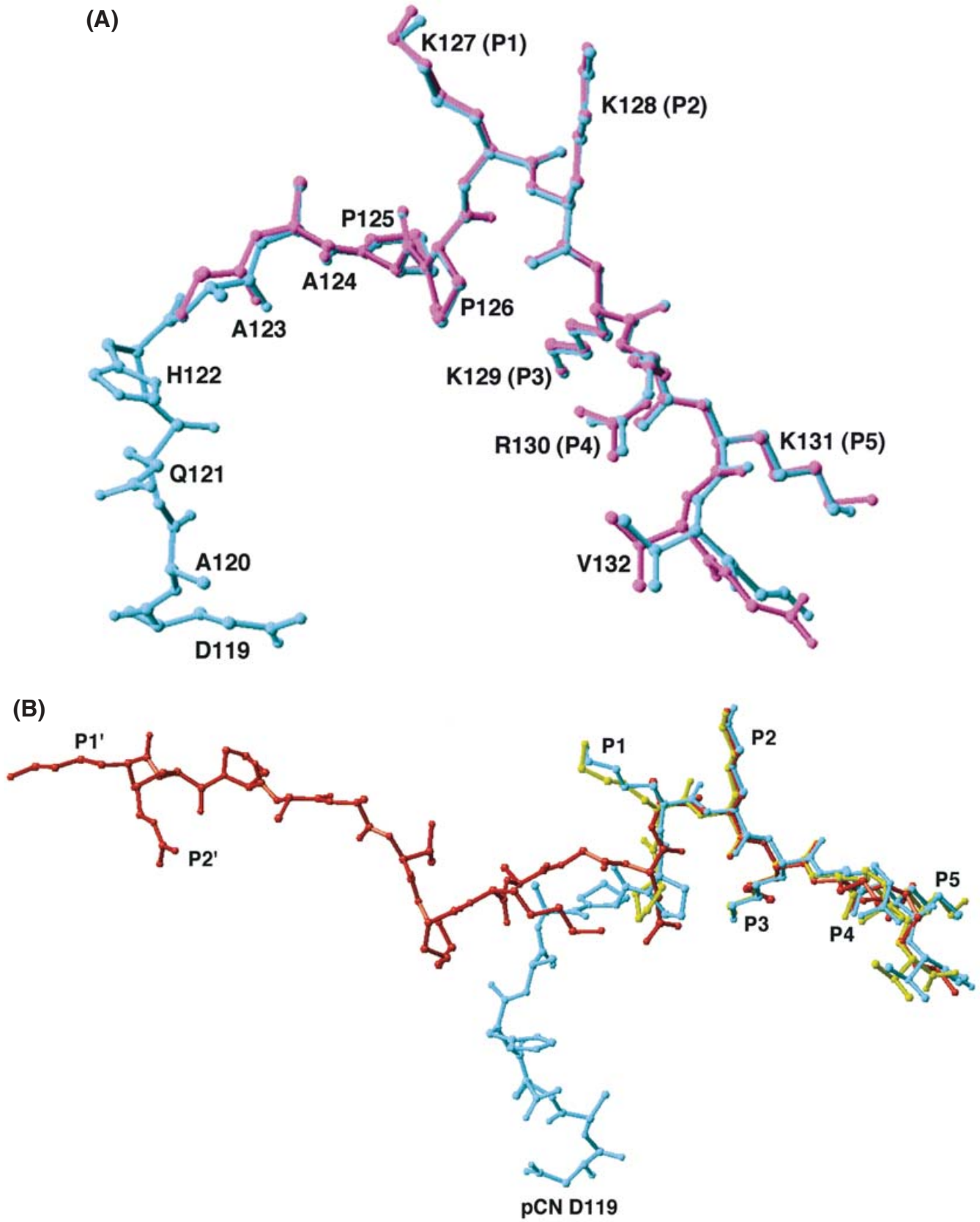

Figure 3 Comparison of the NLS peptides bound to Impα

(**A**) Superposition of the pCN (cyan) and CN (magenta) peptides bound to the major site of Impα(70–529). The Cα atoms of Impα(70–529) in the structures of the two complexes were used in the superposition (drawn with the program RIBBONS [49]). (**B**) Superposition of the pCN peptide (cyan), the N peptide (yellow) [11] and the nucleoplasmin NLS peptide (red) [11] bound to the major site of Impα(70–529), superimposed and drawn as in (**A**).

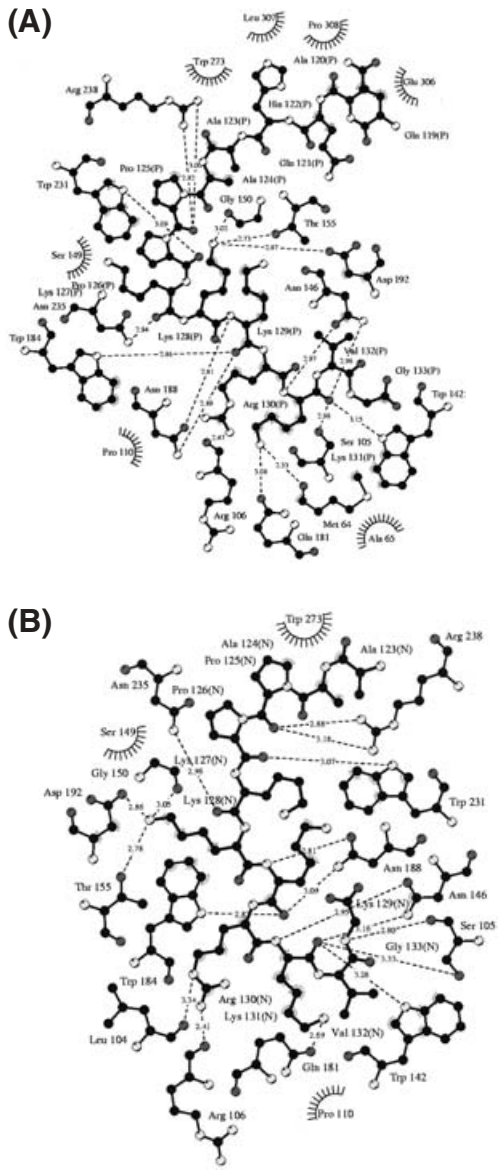

Figure 4 Interactions of the peptides with Impα

(**A**) Schematic diagram of the interactions between the pCN peptide and the major binding site of Impα(70–529). Polar contacts are indicated by broken lines, and hydrophobic contacts are indicated by arcs with radiating spokes. The NLS peptide residues are labelled with 'P'. Carbon, nitrogen and oxygen atoms are shown in black, white and grey respectively. Prepared with the program LIGPLOT [50]. (**B**) Schematic diagram of the interactions between the CN peptide and the major binding site of Impα(70–529), shown as in (**A**), except that the NLS peptide is labelled with 'N'.

Table 3). Hence, the results indicate that negative charge at position 112, in the form of either phospho-Ser or Asp, does not enhance the binding of the T-Ag NLS to the import receptor.

Role of flanking sequences in the recognition of T-Ag NLS by Impα

Structures of the complexes of peptides comprising the basic cluster of the T-Ag NLS, i.e. ^{125}SPKKKRKVE^{133} bound to the N-terminally truncated yeast Impα homologue, Kapα50 [9], and ^{126}PKKKRKV^{132} bound to Impα(70–529) [11], have been reported previously. Whereas the binding of the basic cluster is essentially identical in the structures reported previously and in the structures presented here, several additional N-terminal residues present in the extended CN and pCN peptides are well ordered and clearly make contacts with Impα. This is consistent with the results of ELISA-based binding assays presented here (Table 3), indicating that Impα(70–529), fl-Impα and the Impα/Impβ heterodimer all bind to the extended peptides corresponding to T-Ag residues 112–132 with higher affinity than to the N peptide, comprising only residues 126–132. Similarly, the binding of residues 111–132 of T-Ag fused to β-galactosidase by Impα/Impβ has been shown to be 10-fold higher than that of the fusion protein containing only the basic cluster (residues 126–132), whereas residues 120–132 showed 5-fold higher binding than the basic NLS cluster alone [16]. Clearly, the structural and binding affinity results suggest that the residues flanking the basic cluster play an important role in NLS recognition. Analogously, flanking sequences have been reported to be essential in other NLSs (e.g. DNA helicase Q1 [41,42]).

Previous studies have shown that fl-Impα exhibits high affinity to NLSs only when bound to Impβ [13–22]. Results of the present study (Figure 5, Table 3) show a higher affinity of T-Ag NLS peptides to the non-autoinhibited Impα when compared with the autoinhibited Impα, confirming that the NLS-binding properties of the non-autoinhibited Impα(70–529) closely resemble those of the Impα/Impβ heterodimer. In conclusion, the structural and binding results presented here demonstrate that the residues flanking the basic cluster in a conventional basic monopartite NLS make a significant contribution to NLS recognition by Impα.

The interaction of the T-Ag residues 119–125 with Impα is distinct from the interaction of the residues N-terminal to the C-terminal basic cluster in bipartite NLSs with the receptor. This sequence in the peptide corresponding to the bipartite NLS from nucleoplasmin binds along the major groove on the surface of Impα, where the binding sites for both basic clusters are located. In contrast, the N-terminal sequences of the CN and pCN peptides climb towards the side of the groove, interacting with a completely different set of Impα residues. There appears to be a plethora of possibilities in terms of surface features of Impα that the NLSs exploit for binding.

Role of phosphorylation in the nuclear import of T-Ag

Previously, phosphorylation at the CK2 site has been shown to increase the affinity of binding of β-galactosidase fusion proteins containing the T-Ag NLS to Impα/Impβ complex by approx. 2-fold [16]. A 3-fold increase in binding was observed for a pre-phosphorylated fusion protein containing an inactivated T-Ag NLS, relative to the non-phosphorylated protein, implying that the phosphate may participate directly in binding to Impα [16]. Surprisingly, our structural analysis reveals no direct interaction of the phosphate group with Impα.

One interesting observation is that despite increasing *B* factors, continuous electron density is present in the N-terminal region of the pCN peptide extending to the residue 119. In contrast, the electron density becomes discontinuous N-terminal to residue 123 in the CN peptide, and no clear electron density is present in the region 119–122 of this non-phosphorylated peptide. This is despite the fact that a higher concentration of the CN peptide compared with that of the pCN peptide was used in the co-crystallization experiments. Phosphorylation at residue 112 may therefore influence the conformational properties of the surrounding sequence, facilitating additional interactions in the region 119–122 of the peptide. Previously, the importance of conformation in this region of T-Ag has been suggested on the basis of the effect of proline substitutions in the region 109–111 [25]. However, our binding studies show a slightly lower affinity of the phosphorylated peptide to Impα. We conclude that the phosphate

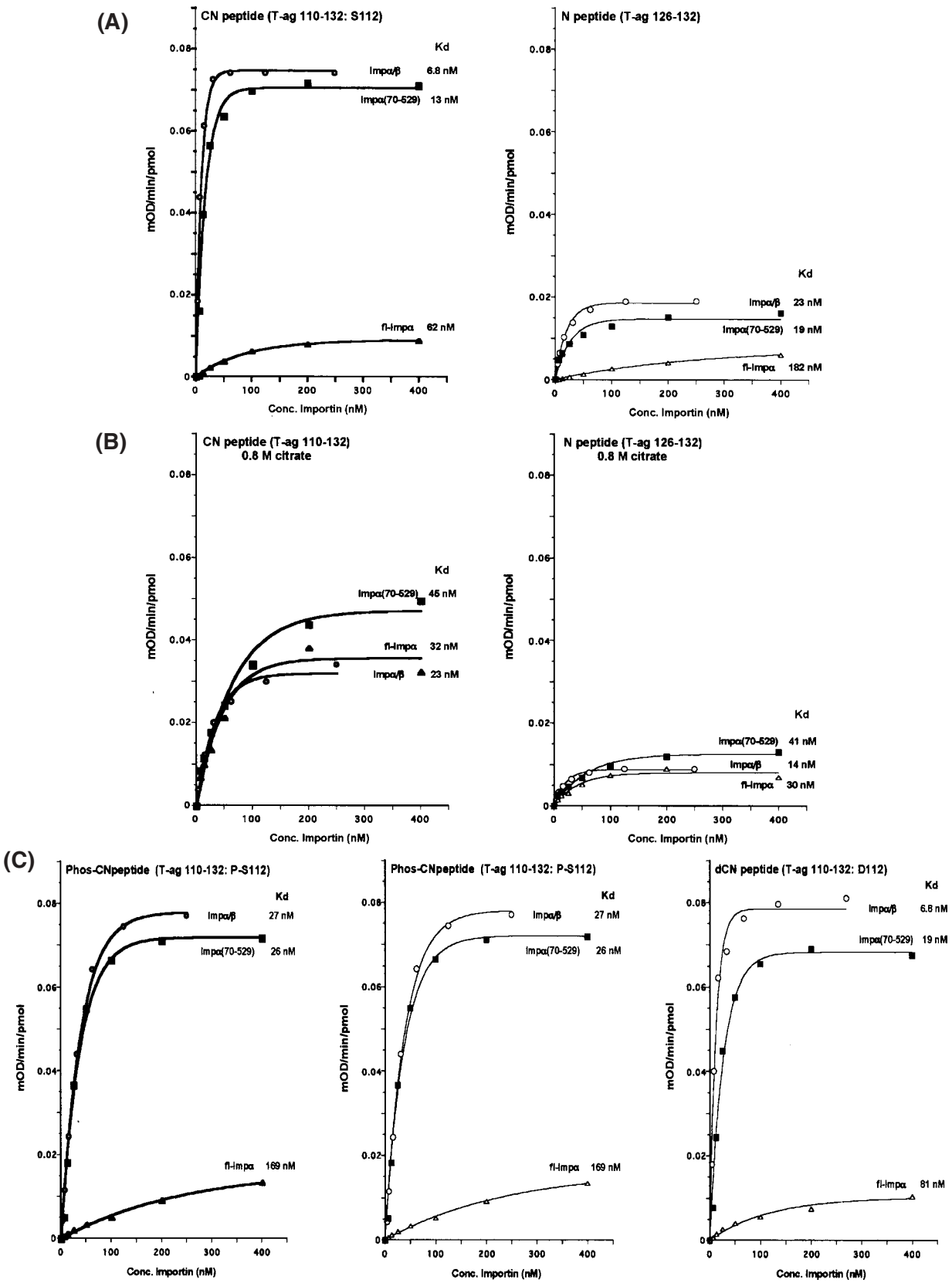

Figure 5 ELISA-based binding assays

Binding of importin to T-Ag synthetic NLS peptides as determined using an ELISA-based binding assay [16,23]. Data were fitted to the function $B(x) = B_{max}(1 - e^{-kt})$, where x is the concentration of importin and B the level of importin binding. The results are from a single typical experiment, performed in triplicate, for incubations with Impα(70–529), fl-Impα and fl-Impα/Impβ in the absence (**A**, **C**) and presence (**B**) of 0.8 M citrate. See Table 3 for pooled data.

Table 3 NLS peptide–importin binding parameters* as determined using an ELISA-based binding assay

	Impα(70–529)		fl-Impα		fl-Impα/β	
Peptide	K_d (nM)	B_{max}	K_d (nM)	B_{max}	K_d (nM)	B_{max}
CN peptide	10 ± 1.8 (8)†	100	43 ± 9 (10)	15 ± 3	5.8 ± 0.7 (9)	98 ± 3
+ 0.8 M citrate	39 ± 5 (2)	74 ± 8	27 ± 4 (3)	66 ± 10	23 ± 8 (3)	49 ± 11
pCN peptide	24 ± 8 (4)	81 ± 2	166 ± 36 (4)	12 ± 1.1	24 ± 6 (4)	88 ± 4
Cn peptide	74 ± 1.4 (1)	20 ± 1.5	315 ± 46 (2)	6.5 ± 1.2	76 ± 7 (2)	21 ± 1
dCN Peptide	14 ± 1.5 (7)‡	97 ± 0.9	63 ± 11 (7)	12.5 ± 1.4	9.4 ± 0.7 (5)	92 ± 7
+ 0.8 M citrate	48 ± 5.3 (2)	83 ± 3.0	25 ± 5 (2)	43 ± 2	30 ± 19 (3)	50 ± 2
N peptide	20 ± 1.1 (2)	18 ± 1.8	168 ± 14 (2)	7.8 ± 3.0	20 ± 3 (2)	25 ± 1
+ 0.8 M citrate	77 ± 37 (2)	15 ± 2.6	117 ± 15	21 ± 1.2	14 ± 3.6 (2)	12 ± 2.7
bipN1N2 peptide	10 ± 1.0 (3)§	98 ± 2	40 ± 6.5 (3)	17 ± 1.8	5.2 ± 0.6 (4)	95 ± 11
+ 0.8 M citrate	40 ± 4.9 (1)	104 ± 4	257 ± 36 (1)	37 ± 6.7	30 ± 6 (2)	65 ± 2

* Data represent the means ± S.E.M. (*n* in parentheses) for the apparent dissociation constant (K_d) and maximal binding B_{max}* [% relative to that of Impα(70–529) to peptide CN], in the absence of 0.8 M citrate; where $n = 1$, the S.E.M. was estimated from the curve fit. Measurements were made as described in the Experimental section [16,17,23,25].

† In the presence of Impβ, which does not bind truncated Impα(70–529), the binding affinity was 14.8 ± 0.9 nM (2), with B_{max} 107 ± 7 %.

‡ In the presence of Impβ, the binding affinity was 21.2 ± 1.3 nM (2), with B_{max} 100 ± 3 %.

§ In the presence of Impβ, the binding affinity was 13.7 ± 1.2 nM (1), with B_{max} 104 ± 27 %.

group is not directly involved in binding to Impα, and that the observed increase in the efficiency of nuclear import cannot be explained by the increase in the affinity of the NLS to Impα, as hypothesized previously [23]. However, the observed increase in the binding of the phosphorylated T-Ag NLS–β-galactosidase fusion protein to Impα/Impβ heterodimer [16] suggests that phosphorylation may improve the presentation or accessibility of the NLS in the context of a larger protein, as would be the case in the native protein. Phosphorylation may also alter the kinetics of binding, decreasing the off-rate at the expense of decreasing the on-rate, without significantly affecting the equilibrium constant of binding. Such effects could be investigated in the future by binding assays capable of measuring the kinetic parameters (e.g. surface plasmon resonance biosensor). Furthermore, other mechanisms may contribute to the observed increase in the efficiency of nuclear import [16,28]. These may include the effects on the NLS release in the nucleus, interaction with NLS-masking factors or interactions with the nuclear pore during transport.

An intriguing similarity exists between the effects of phosphorylation on the T-Ag NLS and the metabolic enzyme PAH (phenylalanine hydroxylase). The enzyme activity of PAH is stimulated by phenylalanine, and phosphorylation by cAMP-dependent protein kinase at Ser^{16} further modulates the activity by decreasing the concentration of phenylalanine required for activation [43,44]. The crystal structures of PAH have been determined in both phosphorylated and non-phosphorylated states. The N-terminal 18-residue sequence containing the phosphorylation site contains no interpretable electron density in either form, and there are no major differences in the structures of the two forms [45]. For PAH, it is quite probable that phosphorylation only has a structural effect concurrently with regulatory phenylalanine binding. Examples of similar concerted action of regulatory events include the combination of cyclin binding and phosphorylation in cyclin-dependent protein kinase 2 [29,46] and dual phosphorylation in the mitogen-activated protein kinase extracellular-signal-regulated kinase-2 [47]. Our results suggest that phosphorylation may act in concert with another regulatory mechanism during the nuclear import of T-Ag.

Conclusion

The structures of the complexes of Impα with the peptides corresponding to the phosphorylated and non-phosphorylated forms of the T-Ag phosphorylation-regulated NLS provide new insights into the mechanism of regulation of nuclear import. Both structural and binding results demonstrate an important role of residues, N-terminal to the basic cluster in the T-Ag monopartite NLS, in the recognition by Impα. Surprisingly, the residue phosphorylated by protein kinase CK2 is not involved in a specific interaction with Impα, consistent with the results from direct binding assays, indicating that phosphorylation does not increase the affinity of the T-Ag NLS to Impα. We conclude that the sequences flanking the basic clusters in NLSs play a crucial role in nuclear import by modulating the recognition of the NLS, whereas phosphorylation of T-Ag enhances nuclear import by a mechanism that does not appear to involve direct interaction of the phosphorylated residue with Impα.

We thank Ian Jennings (St. Vincent's Institute of Medical Research, Melbourne, Australia) for helpful discussions and Chenoa Barton (Monash University, Melbourne, Australia) for assistance with the ELISA-based binding assays. M.R.M.F. was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (Brazil). D.A.J. and B.K. are recipients of a NHMRC (National Health and Medical Research Council, Australia) Senior Research Fellowship and B.K. was a Wellcome Senior Research Fellow in Medical Science in Australia.

REFERENCES

1 Dingwall, C. and Laskey, R. A. (1991) Nuclear targeting sequences – a consensus? Trends Biochem. Sci. **16**, 478–481
2 Weis, K. (2002) Nucleocytoplasmic transport: cargo trafficking across the border. Curr. Opin. Cell Biol. **14**, 328–335
3 Damelin, M., Silver, P. A. and Corbett, A. H. (2002) Nuclear protein transport. Methods Enzymol. **351**, 587–607
4 Conti, E. (2002) Structures of importins. Results Probl. Cell Differ. **35**, 93–113
5 Gorlich, D., Henklein, P., Laskey, R. A. and Hartmann, E. (1996) A 41 amino acid motif in importin-α confers binding to importin-β and hence transit into the nucleus. EMBO J. **15**, 1810–1817
6 Weis, K., Ryder, U. and Lamond, A. I. (1996) The conserved amino-terminal domain of hSRP1α is essential for nuclear protein import. EMBO J. **15**, 1818–1825
7 Moroianu, J., Blobel, G. and Radu, A. (1996) The binding site of karyopherin α for karyopherin β overlaps with a nuclear localization sequence. Proc. Natl. Acad. Sci. U.S.A. **93**, 6572–6576
8 Peifer, M., Berg, S. and Reynolds, A. B. (1996) A repeating amino acid motif shared by proteins with diverse cellular roles. Cell (Cambridge, Mass.) **76**, 789–791
9 Conti, E., Uy, M., Leighton, L., Blobel, G. and Kuriyan, J. (1998) Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin α. Cell (Cambridge, Mass.) **94**, 193–204

10 Conti, E. and Kuriyan, J. (2000) Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin α. Structure **8**, 329–338
11 Fontes, M. R. M., Teh, T. and Kobe, B. (2000) Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-α. J. Mol. Biol. **297**, 1183–1194
12 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
13 Rexach, M. and Blobel, G. (1995) Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. Cell (Cambridge, Mass.) **83**, 683–692
14 Gorlich, D., Pante, N., Kutay, U., Aebi, U. and Bischoff, F. R. (1996) Identification of different roles for RanGDP and RanGTP in nuclear protein import. EMBO J. **15**, 5584–5594
15 Efthymiadis, A., Shao, H., Hübner, S. and Jans, D. A. (1997) Kinetic characterization of the human retinoblastoma protein bipartite nuclear localization sequence (NLS) *in vivo* and *in vitro*. A comparison with the SV40 large T-antigen NLS. J. Biol. Chem. **272**, 22134–22139
16 Hübner, S., Xiao, C. Y. and Jans, D. A. (1997) The protein kinase CK2 site (Ser^{111}/Ser^{112}) enhances recognition of the simian virus 40 large T-antigen nuclear localization sequence by importin. J. Biol. Chem. **272**, 17191–17195
17 Hübner, S., Smith, H. M. S., Hu, W., Chen, C. K., Rihs, H. P., Paschal, B. M., Raikhel, N. V. and Jans, D. A. (1999) Plant importin-α binds nuclear localization sequences with high affinity and can mediate nuclear import independent of importin-β. J. Biol. Chem. **274**, 22610–22617
18 Briggs, L. J., Stein, D., Goltz, J., Corrigan, V. C., Efthymiadis, A., Hübner, S. and Jans, D. A. (1998) The cAMP-dependent protein kinase site (Ser^{312}) enhances dorsal nuclear import through facilitating nuclear localization sequence–importin interaction. J. Biol. Chem. **273**, 22745–22752
19 Hu, W. and Jans, D. A. (1999) Efficiency of importin-α/β-mediated nuclear localization sequence recognition and nuclear import. Differential role of NTF2. J. Biol. Chem. **274**, 15820–15827
20 Fanara, P., Hodel, M. R., Corbett, A. H. and Hodel, A. E. (2000) Quantitative analysis of nuclear localization signal (NLS)–importin-α interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding. J. Biol. Chem. **275**, 21218–21223
21 Hodel, M. R., Corbett, A. H. and Hodel, A. E. (2001) Dissection of a nuclear localization signal. J. Biol. Chem. **276**, 1317–1325
22 Catimel, B., Teh, T., Fontes, M. R., Jennings, I. G., Jans, D. A., Howlett, G. J., Nice, E. C. and Kobe, B. (2001) Biophysical characterization of interactions involving importin-α during nuclear import. J. Biol. Chem. **276**, 34189–34198
23 Xiao, C. Y., Hübner, S. and Jans, D. A. (1997) SV40 large tumor antigen nuclear import is regulated by the double-stranded DNA-dependent protein kinase site (serine 120) flanking the nuclear localization sequence. J. Biol. Chem. **272**, 22191–22198
24 Jans, D. A., Xiao, C. Y. and Lam, M. H. (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? Bioessays **22**, 532–544
24a Fontes, M. R., Teh, T., Jans, D. A., Brinkworth, R. I. and Kobe, B. (2003) Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-α. J. Biol. Chem. **278**, 27981–27987
25 Xiao, C. Y., Jans, P. and Jans, D. A. (1998) Negative charge at the protein kinase CK2 site enhances recognition of the SV40 large T-antigen NLS by importin: effect of conformation. FEBS Lett. **440**, 297–301
26 Jans, D. A. (1995) The regulation of protein transport to the nucleus by phosphorylation. Biochem. J. **311**, 705–716
27 Jans, D. A. and Hubner, S. (1996) Regulation of protein transport to the nucleus: central role of phosphorylation. Physiol. Rev. **76**, 651–685
28 Jans, D. A. and Jans, P. (1994) Negative charge at the casein kinase II site flanking the nuclear localization signal of the SV40 large T-antigen is mechanistically important for enhanced nuclear import. Oncogene **9**, 2961–2968
29 Johnson, L. N. and O'Reilly, M. (1996) Control by phosphorylation. Curr. Opin. Struct. Biol. **6**, 762–769
30 Akhlynina, T. V., Jans, D. A., Rosenkranz, A. A., Statsyuk, N. V., Balashova, I. Y., Toth, G., Pavo, I., Rubin, A. B. and Sobolev, A. S. (1997) Nuclear targeting of chlorin e6 enhances its photosensitizing activity. J. Biol. Chem. **272**, 20328–20331
31 Akhlynina, T. V., Jans, D. A., Statsyuk, N. V., Balashova, I. Y., Toth, G., Pavo, I., Rosenkranz, A. A., Naroditsky, B. S. and Sobolev, A. S. (1999) Adenoviruses synergize with nuclear localization signals to enhance nuclear delivery and photodynamic action of internalizable conjugates containing chlorin e6. Int. J. Cancer **81**, 734–740
32 Fields, F. B. and Noble, R. L. (1990) SPPS utilizing 9-fluorenylmethoxycarbonyl amino acids. Int. J. Peptide Protein Res. **35**, 161–214
33 Toth, G. K., Varadi, G., Janaky, Y., Penke, B., Mak, M., Lang, E., Otvos, L., Hegedus, Z. and Monostori, E. (1994) Synthesis of phosphotyrosine-containing peptides – comparison of different methods and their use to investigate cell signalling via the T-cell receptor. In Peptides. Proceedings of the 23rd European Peptide Symposium (Maia, H. L. S., ed.), pp. 743–744, ESCOM, Leiden
34 Perich, J. W. and Johns, R. B. (1988) Di-tert-butyl *N*,*N*-diethylphosphoramidite. A new phosphitylating agent for the efficient phosphorylation of alcohols. Synthesis **2**, 142–144
35 Kussel, P. and Frasch, M. (1995) Yeast Srp1, a nuclear protein related to *Drosophila* and mouse pendulin, is required for normal migration, division, and integrity of nuclei during mitosis. Mol. Gen. Genet. **248**, 351–363
36 Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276**, 307–326
37 Brünger, 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 and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D **54**, 905–921
38 Brünger, A. T. (1992) Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature (London) **355**, 472–475
39 Jones, T. A., Bergdoll, M. and Kjeldgaard, M. (1990) O: a macromolecule modeling environment. In Crystallographic and Modeling Methods in Molecular Design (Bugg, C. E. and Ealick, S. E., eds.), pp. 189–195, Springer-Verlag, New York
40 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) *PROCHECK*: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291
41 Miyamoto, Y., Imamoto, N., Sekimoto, T., Tachibana, T., Seki, T., Tada, S., Enomoto, T. and Yoneda, Y. (1997) Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors. J. Biol. Chem. **272**, 26375–26381
42 Seki, T., Tada, S., Katada, T. and Enomoto, T. (1997) Cloning of a cDNA encoding a novel importin-α homologue, Qip1: discrimination of Qip1 and Rch1 from hSrp1 by their ability to interact with DNA helicase Q1/RecQL. Biochem. Biophys. Res. Commun. **234**, 48–53
43 Hufton, S. E., Jennings, I. G. and Cotton, R. G. H. (1995) Structure and function of the aromatic amino acid hydroxylases. Biochem. J. **311**, 353–366
44 Fitzpatrick, P. F. (2000) The aromatic amino acid hydroxylases. Adv. Enzymol. Relat. Areas Mol. Biol. **74**, 235–294
45 Kobe, B., Jennings, I. G., House, C. M., Michell, B. J., Goodwill, K. E., Santarsiero, B. D., Stevens, R. C., Cotton, R. G. H. and Kemp, B. E. (1999) Structural basis of autoregulation of phenylalanine hydroxylase. Nat. Struct. Biol. **6**, 442–448
46 Pavletich, N. P. (1999) Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. J. Mol. Biol. **287**, 821–828
47 Canagarajah, B. J., Khokhlatchev, A., Cobb, M. H. and Goldsmith, E. J. (1997) Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. Cell (Cambridge, Mass.) **90**, 859–869
48 Esnouf, R. M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J. Mol. Graph. **15**, 133–138
49 Carson, M. (1997) Ribbons. Methods Enzymol. **277**, 493–505
50 Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. Protein Eng. **8**, 127–134

Received 4 April 2003/4 June 2003; accepted 10 July 2003
Published as BJ Immediate Publication 10 July 2003, DOI 10.1042/BJ20030510