

ESTUDOS ESTRUTURAIS COM A IMPORTINA- α DO FUNGO

Neurospora crassa E SEQUÊNCIAS DE LOCALIZAÇÃO

NUCLEAR

NATÁLIA ELISA BERNARDES

Dissertação apresentada ao Instituto de Biociências, Câmpus de Botucatu, UNESP, para a obtenção do título de Mestre no Programa de Pós-Graduação em Biologia Geral e Aplicada, Área de concentração Biomoléculas - estrutura e função.

Orientador: Prof. Tit. Marcos Roberto de Mattos Fontes
Co-orientadora: Dra. Agnes Alessandra Sekijima Takeda

BOTUCATU – SP

2014

UNIVERSIDADE ESTADUAL PAULISTA

“Júlio de Mesquita Filho”

INSTITUTO DE BIOCÊNCIAS DE BOTUCATU

ESTUDOS ESTRUTURAIS COM A IMPORTINA- α DO FUNGO

Neurospora crassa E SEQUÊNCIAS DE LOCALIZAÇÃO

NUCLEAR

NATÁLIA ELISA BERNARDES

ORIENTADOR: PROF. TIT. MARCOS ROBERTO DE MATTOS FONTES

CO-ORIENTADORA: DRA. AGNES ALESSANDRA SEKIJIMA TAKEDA

Dissertação apresentada ao Instituto de Biociências, Câmpus de Botucatu, UNESP, para a obtenção do título de Mestre no Programa de Pós-Graduação em Biologia Geral e Aplicada, Área de concentração Biomoléculas- estrutura e função.

BOTUCATU - SP

Bernardes, Natália Elisa.

Estudos estruturais com a Importina- α ; do fungo *Neurospora crassa* e sequências de localização nuclear / Natália Elisa Bernardes. - Botucatu, 2014

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências de Botucatu

Orientador: Marcos Roberto de Mattos Fontes

Coorientador: Agnes Alessandra Sekijima Takeda

Capes: 20901003

1. *Neurospora crassa*. 2. Peptídeos NLS. 3. Análise cromatográfica. 4. Proteínas - Pesquisa. 5. Cristalografia.

Palavras-chave: Cristalografia de proteínas; Importação nuclear; Importina- α ; NLS; *Neurospora crassa*.

Aos que apoiam e incentivam os jovens pesquisadores brasileiros.

AGRADECIMENTOS

Aos meus pais Lincoln e Maria que me acompanham, apoiam e incentivam em todos os desafios que me arrisco a enfrentar nessa vida.

Às minhas irmãs Carolina e Amália e minha tia Márcia, minhas melhores amigas e sempre ótimas companhias para viagens, conversas e risadas.

Ao meu professor, Marcos Fontes, por me aceitar como aluna e por sua paciência e compreensão ao orientar.

À minha co-orientadora Agnes, sempre solícita e atenciosa em suas sugestões e orientações.

Aos meus amigos do Departamento de Física e Biofísica que tornam o trabalho diário prazeroso, além de proporcionar excelentes momentos de discussões, sejam estas científicas ou cotidianas.

Aos meus amigos da Turma V de Engenharia Biotecnológica, que fazem de cada reencontro uma oportunidade de recarregar as energias.

A todos do Instituto de Biociências de Botucatu, pelo apoio financeiro e burocrático.

A Todos do Programa de Biologia Geral e Aplicada, pelo apoio técnico e burocrático.

À Fundação de Apoio a Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro e pelo estímulo a ciência nacional.

“A ignorância gera mais frequentemente confiança do que o conhecimento: são os que sabem pouco, e não aqueles que sabem muito, que afirmam de uma forma tão categórica que este ou aquele problema nunca será resolvido pela ciência.”

Charles Darwin

RESUMO

Um dos mecanismos de transporte de macromoléculas do citoplasma para o núcleo celular ocorre através da passagem da macromolécula pelo complexo poro nuclear (CPN), presente no envoltório nuclear, e depende de proteínas transportadoras denominadas Importinas. Neste processo, conhecido como via clássica de importação nuclear, a proteína Importina- α (Imp α) reconhece sequências de localização nuclear (NLSs) na proteína a ser transportada para a formação de um complexo junto a Importina- β (Imp β) permitindo o transporte da macromolécula. O objetivo do presente trabalho é o estudo estrutural da Imp α proveniente do fungo filamentososo *Neurospora crassa* (Imp α Nc), a fim de reconhecer as regiões que determinam especificidades da proteína, além de comparar seu comportamento nativo e na presença de um peptídeo NLS. Os primeiros experimentos com a Imp α Nc permitiram uma caracterização inicial da proteína e seu comportamento em solução. Experimentos de cromatografia analítica de exclusão molecular, comparando duas amostras de Imp α Nc: (1) na presença do peptídeo NLS da proteína FEN1 (FEN1 NLS) e (2) sem peptídeo NLS, indicaram a formação de aglomerados na amostra 2 e a conformação, predominantemente, monomodal na amostra 1, sugerindo uma maior estabilidade da proteína na presença de peptídeos NLS. Para aprofundar as informações sobre a Imp α Nc, experimentos de cristalização foram conduzidos com a proteína complexada ao peptídeo de NLS clássica e monopartida do SV40 (SV40 NLS), conforme experimentos anteriores sugeriram a maior estabilidade da proteína na presença de NLSs. Um primeiro cristal (Imp α Nc-1), obtido na condição 0,2mM fosfato de sódio dibásico dihidratado e 20% (w/v) de polietilenoglicol 3350, foi submetido à difração de raios X e apresentou padrão de difração satisfatório para elucidação da estrutura do complexo a uma resolução de 2,05 Å. Um segundo cristal (Imp α Nc-2), obtido na condição 0,2 mM de bicina pH 8.5 e 20% PEG 6000, foi submetido a difração de raios X na presença de 20% de glicerol e permitiu a elucidação da estrutura a 1,75 Å. O conjunto de dados Imp α Nc-2 foi utilizado para as análises estruturais no presente trabalho. Na estrutura da Imp α Nc o peptídeo SV40 NLS pôde ser observado nos sítio principal e secundário da proteína em uma conformação estendida e antiparalela aos motivos ARM da proteína. A estrutura apresentou uma conformação alongada e côncava, conservada em outros organismos. Análises de dicroísmo circular com a Imp α Nc e a Imp α de *Mus musculus* (Imp α Mm) mostraram elevada conservação da estrutura secundária entre as duas proteínas e a sobreposição das estruturas da Imp α Nc e Imp α Mm mostrou que a Imp α Mm é mais côncava do que a Imp α Nc, o que justifica a diferença nos parâmetros de rede entre os cristais das duas proteínas. A ligação do peptídeo SV40 NLS no sítio secundário ocorreu de forma similar à ligação do mesmo peptídeo no sítio secundário da Imp α de arroz, com os aminoácidos K129 e R130, nas posições P1' e P2' respectivamente, enquanto nas Imp α de camundongo e levedura os aminoácidos K128 e K129 que se ligam na posição P1' e P2', respectivamente. Alguns resíduos presentes na região do sítio secundário, conservados entre as Importinas da família α 1, foram também observados na Imp α Nc e sugerem especificidades da proteína na ligação a alguns NLSs monopartidos provenientes de proteínas de fungos. Ensaio de afinidade com Imp α Nc e o peptídeo SV40 NLS, permitiram o cálculo das constantes termodinâmicas envolvidas na interação da proteína com o peptídeo NLS, as quais poderão ser utilizadas como parâmetro para verificar a afinidade de potenciais NLSs específicos de fungos.

Palavras-chave: Importina- α , NLS, *Neurospora crassa*, Cristalografia de proteínas, Importação Nuclear, SV40.

ABSTRACT

*The transport of macromolecules from the cytoplasm to the nucleus occurs by passage through the nuclear pore complex, present in the nuclear envelope. One of that nuclear transport pathway depends on carrier proteins called Importins. In this process, known as classical nuclear import pathway, the Importin- α protein ($Imp\alpha$) recognizes nuclear localization sequences (NLSs) in the protein to be transported to the formation of a complex with Importin- β ($Imp\beta$) allowing the transport of macromolecules. The aim of this work is the structural study of the protein Importin- α , from the filamentous fungus *Neurospora crassa* ($Imp\alpha_{Nc}$) in order to recognize the regions that determine the specificity of the protein and to compare their behavior in its native state and in presence of a NLS peptide. The first experiments with $Imp\alpha_{Nc}$ allowed an initial characterization of the protein and its behavior in solution. Experiments of analytical size exclusion chromatography comparing two samples of $Imp\alpha$: (1) in the presence of the NLS peptide of the protein FEN1 (FEN1 NLS) and, (2) without NLS peptide; indicated the formation of agglomerates in the sample 2 and the conformation predominantly monomodal, in the sample 1, suggesting a greater stability of the protein in the presence of NLS peptides. For further information about the $Imp\alpha_{Nc}$, crystallization experiments were carried out with the peptide complexed to the classic NLS SV40 (SV40 NLS) protein as previous experiments have suggested the increased stability of the protein in the presence of NLSs. A first crystal obtained in the condition 0.2mM dibasic sodium phosphate dihydrate and 20% (w / v) polyethylene glycol 3350 were subjected to x-ray diffraction and showed satisfactory diffraction pattern for the elucidation of the structure of the complex at a resolution of 2.05 Å. A second crystal ($Imp\alpha_{Nc}$ -2) obtained under the condition of 0.2 mM bicine pH 8.5 and 20% PEG 6000, was subjected to x-ray diffraction in the presence of 20% glycerol and allowed the elucidation of the structure 1, 75 Å. The set of $Imp\alpha_{Nc}$ -2 data was used for structural analysis in this work. In the structure of the $Imp\alpha_{Nc}$ the SV40 NLS peptide could be observed in the major and minor sites in an extended conformation and antiparallel to the ARM motifs of the protein. The structure showed an elongated, concave conformation, conserved in other organisms. Analysis of circular dichroism with $Imp\alpha_{Nc}$ and $Imp\alpha$ from *Mus musculus* ($Imp\alpha_{Mm}$) showed high conservation of secondary structure between the two proteins and the overlap of structures of $Imp\alpha_{Nc}$ and $Imp\alpha_{Mm}$ showed that $Imp\alpha_{Mm}$ is more concave than $Imp\alpha_{Nc}$, which explains the difference in network parameters between the two crystals proteins. The SV40 NLS peptide binding occurred at the minor binding site similarly of the same peptide bound in the minor site in rice $Imp\alpha$, with K129 and R130 amino acids at positions P1' and P2' respectively, while the mouse and yeast $Imp\alpha$ are the K128 and K129 amino acids which bind the P1' and P2' position, respectively. Some residues in the minor site, preserved among the family of $Imp\alpha$ were also observed in $Imp\alpha_{Nc}$ and suggest protein specificity in binding to some monopartites NLSs from proteins of fungi. Affinity assays with $Imp\alpha_{Nc}$ and SV40 NLS peptide allowed the calculation of thermodynamic constants involved in the interaction of the protein with the NLS peptide, which may be used as a parameter to verify affinity for potential fungal NLSs.*

Keywords: *Importin- α , NLS, *Neurospora crassa*, Crystallography, Nuclear Import, SV40.*

LISTA DE ABREVIATURAS

- Abs** – absorvância
- IgG** – imonuglobulina G
- ARM** – repetições tipo *armadillo*
- CAS** – Exportina *Cellular apoptosis susceptibility*.
- Cdna** – DNA codificante
- CPN** – Complexo Poro Nuclear
- cNLS** – Sequência de Localização Nuclear clássica
- DLS** – *Dynamic Light Scattering*
- DNase** – Desoxirribonuclease.
- FEN1 NLS** – NLS da Endonuclease 1 flap, nuclease multifuncional, atua no repado de DNA na via de reparo por excisão de base.
- FG** – fenilalanina/glicina.
- GTP** – Guanosina trifosfato.
- GTPaseRan**- proteína nuclear associada a Ras, com atividade GTPase
- H1** – hélice 1 do motivo ARM
- H2** – hélice 2 do motivo ARM
- H3** – hélice 3 do motivo ARM
- IBB**- *Importin-β binding*- domínio de ligação à Importina-β
- Impα** - Importina-α
- ImpαMm** – Importina-α de *Mus musculus*, truncada N-terminalmente
- ImpαNc** – Importina-α de *Neurospora crassa*, truncada N-terminalmente
- ImpαOs** – Importina-α de *Oryza sativa*, truncada N-terminalmente
- ImpαSc** – Importina-α de *Saccharomyces cerevisiae*, truncada N-terminalmente
- Impβ** - Importina-β
- IPTG** – Isopropil-beta-D-tiogalactopiranosídeo.
- Kapα1**- Carioferinas da família α1.
- Kapα2**- Carioferinas da família α2.
- Kapα3**- Carioferinas da família α3.
- LB** – Luria Bertani (meio de cultura).
- LNLS** – Laboratório Nacional de Luz Síncrotron
- NE** – *nuclear envelope*- envelope nuclear
- NTF2**- Nuclear Transport Fator 2

P1-P6 – Posições do sítio de ligação principal da Importina- α .

P1'-P4' – Posições do sítio de ligação secundário da Importina- α .

PDB ID – Código identificador da estrutura depositada no *Protein data bank web site*.

PEG – polietileno glicol

pET28a – vetor de expressão em bactérias

PVDF – *Polyvinylidene fluoride*

RanBP1 – *Ran-specific binding protein 1* – protein de ligação específica a Ran

RanGAP1 – *RanGTPase-activating protein* – proteína ativadora da RanGTPase

RanGTP - Guanosina trifosfato Ran, associada à GTP

r.m.s.d. – *root mean square deviation* - desvio da raiz quadrática média

RNase – Ribonuclease

rpm – rotações por minuto

SDS-PAGE – Gel de poliacrilamida na presença de dodecil sulfato de sódio

NLS – *Nuclear Localization Sequence*- Sequência de Localização Nuclear

SV40 – Vírus Símio 40

TBS – Tampão tris salina

Tm – Temperatura de *melting* (desnaturação)

LISTA DE FIGURAS

Figura 1. Esquema simplificado da Via Clássica de Importação Nuclear.....	23
Figura 2. Esquema ilustrado da Via Clássica de Importação Nuclear.....	24
Figura 3. Esquema ilustrado do Complexo Poro Nuclear (CPN).....	25
Figura 4. Estrutura da Imp β complexada ao domínio IBB.....	27
Figura 5. Estrutura da Imp α de <i>Mus musculus</i> (Imp α Mm).....	29
Figura 6. Árvore filogenética da proteína Importina- α	30
Figura 7. Cromatograma de eluição e gel SDS-Page 12% das frações obtidas na purificação da Imp α Nc por cromatografia de afinidade.....	40
Figura 8. Verificação do perfil oligomérico da Imp α Nc.....	44
Figura 9. Gráfico de distribuição oligomérica de amostra de Imp α Nc, obtido através de experimento de Espalhamento de Luz Dinâmico (DLS).....	45
Figura 10. Cristais de Imp α Nc crescidos na presença de peptídeos SV40 NLS.....	52
Figura 11. Imagem de difração de raios X dos cristais Imp α Nc-1 e Imp α Nc-2 do complexo Imp α Nc/ SV40 NLS.....	54
Figura 12. Estrutura da proteína Imp α Nc complexada a peptídeos SV40 NLS.....	58
Figura 13. Mapa de densidade eletrônica do peptídeo SV40 NLS ligado aos sítios principal das estruturas Imp α Nc-1 e Imp α Nc-2.....	59
Figura 14. Mapa de densidade eletrônica do peptídeo SV40 NLS ligado ao sítio secundário das estruturas Imp α Nc-1 e Imp α Nc-2.....	60
Figura 15. Diagrama de interações dos peptídeos SV40 NLS com resíduos do sítio principal da Imp α Nc.....	62
Figura 16. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio principal.....	63
Figura 17. Diagrama de interações dos peptídeos SV40 NLS com resíduos do sítio secundário da Imp α Nc.....	65
Figura 18. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio secundário.....	66
Figura 19. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio principal, nas posições P1 a P6.....	68

Figura 20. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio secundário nas posições P1' a P4'	69
Figura 21. Sobreposição do peptídeo SV40 NLS interagindo com os sítios principal e secundário das proteínas Imp α Nc, Imp α Mm e Imp α Os.....	71
Figura 22. Verificação da estrutura secundária e temperatura de desnaturação térmica da Imp α Nc.....	75
Figura 23. Comparação das sequências de aminoácidos α de Imp α de diferentes organismos.....	76
Figura 24. Comparação das estruturas de Imp α de diferentes organismos.....	77
Figura 25. Resíduos não-conservados entre as estruturas da Imp α Mm e Imp α Nc que podem estar relacionados a especificidades na interação com NLSs.....	79
Figura 26. Termograma e titulação isotérmica do complexo Imp α /SV40 NLS.....	82
Figura 27. Representação das curvas de Dicroísmo circular correspondentes a diferentes elementos de estrutura secundária.....	101
Figura 28. Diagrama de fases do processo de cristalização de proteínas.	102
Figura 29. Ilustração das técnicas de Hanging Drop (A) e Sitting Drop (B).....	104

LISTA DE TABELAS

Tabela 1. Sequência de aminoácidos requerida para o reconhecimento de NLSs pela Imp α ..	32
Tabela 2. Estatísticas da coleta de dados de difração de Raios-X e refinamento inicial dos modelos Imp α Nc-1 e Imp α Nc-2 do complexo Imp α Nc /SV40.....	55
Tabela 3. Estatísticas finais de refinamento dos modelos Imp α Nc-1 e Imp α Nc-2 do complexo Imp α Nc /SV40.....	57
Tabela 4. Distribuição do fator de temperatura (β -factor), para cada resíduo do peptídeo SV40 NLS que interage com a proteína Imp α Nc no sítio principal.....	64
Tabela 5. Distribuição do fator de temperatura (β -factor), para cada resíduo do peptídeo SV40 NLS que interage com a proteína Imp α Nc no sítio secundário.....	67
Tabela 6. Ligação do peptídeo SV40 NLS a Imp α de diferentes organismos.....	69
Tabela 7. Valores das constantes termodinâmicas que definem a afinidade da interação da Imp α Nc com o peptídeo SV40 NLS.....	82

SUMÁRIO

LISTA DE ABREVIATURAS	9
LISTA DE FIGURAS	11
LISTA DE TABELAS	13
ESTRUTURA DA DISSERTAÇÃO	18
1. Introdução	19
1.1. Transporte intracelular de macromoléculas.....	20
1.2. Via Clássica de Importação Nuclear.....	22
1.3. Complexo Poro Nuclear (CPN).	25
1.4. Importina- β (Imp β).	26
1.5. Importina- α (Imp α).	27
1.6. Distribuição filogenética de Importina- α homólogas.	29
1.7. Sequências de Localização Nuclear (NLSs).....	31
1.8. Importina- α e potenciais NLSs do fungo <i>Neurospora crassa</i>	32
2. Objetivos	34
3. Expressão e purificação das proteínas ImpαNc e ImpαMm recombinantes.	36
3.1. Material e Métodos	37
3.1.1. Clonagem da proteína recombinante Imp α Nc.	37
3.1.2. Expressão e purificação da proteína recombinante Imp α Nc.	37
3.1.3. .. Clonagem, expressão e purificação da proteína recombinante Imp α de <i>Mus Musculus</i>	39
3.2. Resultados e Discussão	39

4. Verificação do comportamento da proteína ImpαNc na ausência ou presença de NLSs.....	41
4.1. Material e Métodos	42
4.1.1. Síntese de peptídeos NLS.	42
4.1.2. Cromatografia Analítica por Exclusão Molecular.....	42
4.1.3. Western-blot.....	42
4.1.4. Espalhamento de Luz Dinâmico.....	43
4.2. Resultados.....	43
4.3. Discussão.....	45
4.3.1. Mudança no perfil oligomérico da proteína na presença de peptídeos NLSs.....	45
5. Estrutura cristalográfica do complexo ImpαNc/SV40NLS.....	47
5.1. Material e Métodos	48
5.1.1. Síntese de peptídeos NLS.	48
5.1.2. Cristalização.....	48
5.1.3. Coleta e processamento de dados de difração de raios X.....	48
5.1.4. Elucidação da estrutura, modelagem e refinamento.....	49
5.1.5. Análise e comparação das estruturas cristalográficas.....	50
5.2. Resultados.....	50
5.2.1. Cristais do complexo Imp α Nc/SV40NLS.	50
5.2.2. Coleta e processamento dos dados de difração de raios X.....	53
5.2.3. Elucidação da estrutura, modelagem e refinamento.....	55
5.2.4. Estruturas da proteína Imp α Nc complexada ao peptídeo SV40 NLS.....	57
5.2.5. Interação entre o peptídeo SV40 NLS e resíduos presentes no sítio principal da Imp α Nc.....	61

5.2.6. Interação entre o peptídeo SV40 NLS e resíduos presentes no sítio secundário da Imp α Nc.....	64
5.3. Discussão.....	67
5.3.1. Interações significativas entre resíduos do peptídeo SV40 NLS e resíduos presentes nos sítios principal e secundário da Imp α	67
5.3.2. Interação preferencial do peptídeo SV40 NLS ao sítio principal da proteína Imp α Nc.....	70
6. Comparação da estrutura da ImpαNc com Impα de outros organismos.	72
6.1. Material e Métodos.....	73
6.1.1. Dicroísmo circular.	73
6.1.2. Alinhamento das sequências da Imp α de diferentes organismos. ...	73
6.2. Resultados.....	74
6.2.1. Estrutura secundária da Imp α Nc.	74
6.2.2. Alinhamento das sequências de aminoácidos e sobreposição das estruturas das Imp α de diferentes organismos.....	75
6.3. Discussão.....	77
6.3.1. Comparação das estruturas da Imp α de diferentes organismos.....	77
6.3.2. Resíduos conservados apenas entre Imp α da família α 1 relacionados a especificidades na ligação a NLSs.....	78
7. Afinidade do peptídeo SV40 NLS a ImpαNc.	80
7.1. Material e Métodos.....	81
7.1.1. Calorimetria de titulação isotérmica.....	81

7.2. Resultados	81
7.2.1. Curva de titulação e constantes termodinâmicas da interação entre Imp α Nc e SV40NLS.	81
7.3. Discussão	83
7.3.1. Afinidade entre peptídeos SV40 NLS e a proteína Imp α Nc.	83
8. Conclusões	85
9. Perspectivas	87
10. Referências bibliográficas	89
Apêndices	98
Anexos	121

ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está dividida em dez capítulos, além de um conjunto de apêndices e anexos que apresentam o seguinte conteúdo:

Capítulos 1 e 2: Fundamentação teórica e objetivos do trabalho.

Capítulo 3 a 7: São apresentadas as etapas de desenvolvimento do trabalho, sendo que cada capítulo apresenta os materiais e métodos, resultados e discussão, referente a abordagem de cada etapa para atingir os objetivos do trabalho.

Capítulos 8 a 10: Apresentam as conclusões obtidas no decorrer do desenvolvimento do trabalho e as perspectivas para a continuação da pesquisa na área.

Apêndices: Breve resumo das técnicas biofísicas utilizadas e artigos publicados a partir dos dados obtidos no decorrer do trabalho.

Anexos: Tabela de aminoácidos.

1. Introdução

1.1. Transporte intracelular de macromoléculas.

As funções vitais de um organismo eucarioto são coordenadas por um refinado maquinário presente em suas células. A intrincada rede metabólica responsável pela manutenção da vida deve funcionar de forma harmônica, com todos os componentes em seu fluxo regular. As características que definem as funções das células eucarióticas são determinadas pelo material genético nuclear, o qual está envolto em uma dupla membrana, denominada envoltório nuclear, separando as organelas e componentes presentes no nucleoplasma e citoplasma (Chook & Blobel, 1999; Conti & Izaurralde, 2001; Cook *et al.*, 2007; Lange *et al.*, 2007).

Dessa forma, as atividades que acontecem dentro do núcleo, tais como a expressão gênica, replicação e transcrição, ficam separadas da síntese, armazenamento e degradação de proteínas, que ocorrem no citoplasma. Sendo assim, faz-se necessário um eficiente mecanismo de transporte de micro e macromoléculas, através do envoltório nuclear, que conecte as duas regiões intracelulares de forma que as proteínas nucleares sejam importadas do citoplasma para o núcleo, enquanto RNAs transportadores, RNAs mensageiros são exportados ao citoplasma, seu sítio de função (Gorlich, 1996).

As estruturas presentes no envelope nuclear, conhecidas como Complexo Poro Nuclear (CPNs), permitem a passagem de moléculas de diferentes maneiras: íons e pequenas proteínas neutras entram e saem do núcleo por difusão passiva (Gorlich e Kutay, 1999; Kemminer *et al.*, 1999; Lange *et al.*, 2007; Knudsen *et al.*, 2009), proteínas entre 20 a 60KDa, atravessam para o interior do núcleo com o auxílio de proteínas que formam a superfície citoplasmática do CPN, denominadas Nucleoporinas (Gorlich, 1998), sendo que algumas pequenas proteínas são filtradas por essa rede de Nucleoporinas. Por sua vez, proteínas grandes atravessam o envelope nuclear apenas com o auxílio de mediadores, receptores solúveis que interagem com o CPN.

O transporte de proteínas com o auxílio de mediadores foi verificado, inicialmente, através de estudos com proteínas marcadas com elementos radioativos e microinjetadas em oócitos e ovos. Através da autoradiografia de secções desses ovos e oócitos, foi possível visualizar a translocação de proteínas através da membrana nuclear (Arms, 1968; Merriam, 1969; Gurdon, 1970; Bonner, 1975). Também utilizando marcadores radioativos, foi possível identificar a região da proteína Nucleoplasmina que permite sua localização por mediadores e transporte ao núcleo (Dingwall, 1982). Ainda na década de 80, realizou-se uma série de mutações na proteína antígeno grande T do vírus de símio 40 (SV40), levando a produção de

proteínas com ausência de determinados aminoácidos, as quais foram microinjetadas em ovos e, quando analisadas por imunofluorescência, permitiram visualizar alterações no transporte das proteínas mutadas, o que levou a identificação da sequência de aminoácidos necessária para a localização e transporte da proteína para dentro do núcleo (Kalderon, 1984). Essas sequências seriam, mais tarde, identificadas como Sequências de Localização Nuclear (NLSs).

Através de experimentos com fluorescência e microscopia eletrônica, foi possível observar a presença dos CPNs distribuídos por todo o envelope nuclear (Callan and Tamlin, 1950) e a estrutura destes poros, constituídos pelas proteínas Nucleoporinas, que formam uma região membranosa, inserida no envelope nuclear e, na superfície nuclear, emergem estruturas filamentosas que se estendem ao citoplasma (Franke and Scheer, 1970; Kessel, 1969). Logo, o transporte intracelular de macromoléculas para o interior do núcleo pôde ser observado com maior precisão e foi então dividido em dois passos: ligação aos filamentos de Nucleoporinas presentes na face citoplasmática dos poros nucleares e translocação através do poro, em um mecanismo energia-dependente (Newmeyer and Forbes, 1988; Richardson et. al., 1988; Akey and Godfarb, 1989).

Outros componentes citoplasmáticos essenciais para esse mecanismo de transporte foram identificados através da purificação de proteínas do citoplasma, ligadas a NLSs previamente reconhecidos e marcados. Células permeabilizadas com digitonina (glicosídeo que perfura a membrana plasmática, mas não rompe o envelope nuclear nem a membrana de outras estruturas intracelulares), foram utilizadas para simulações do transporte intracelular de macromoléculas *in vitro*, através da inserção de substratos de importação fluorescentes e acompanhamento por microscopia de fluorescência. Também foi possível o acompanhamento individual dos fatores envolvidos na importação nuclear, devido a célula estar ausente de seus componentes citosólicos solúveis (Adam et. al., 1990). Esses ensaios permitiram o reconhecimento da função das principais proteínas citoplasmáticas envolvidas nesse mecanismo de transporte (Adam et. al., 1990; Adam e Gerace, 1991; Moore e Blobel, 1993; Gorlich et. al., 1994).

Quatro fatores principais, responsáveis pela importação nuclear, foram identificados: a GTPase-Ran (Moore e Blobel, 1993; Melchior et. al., 1993) e p10 (ou NTF2 –*nuclear transport fator 2*) (Moore e Blobel, 1994), requeridas para a translocação da macromolécula através do poro, e um complexo heterodimérico composto por duas proteínas, Importina 60 (Importina- α , Imp α) e Importina 90 (Importina- β , Imp β) (Gorlich, 1995) responsáveis pelo

reconhecimento da macromolécula a ser transportada e ligação aos filamentos de Nucleoporinas do poro nuclear, respectivamente, para posterior etapa de translocação da proteína.

A identificação do complexo de importação nuclear de macromoléculas formado por Importinas e proteínas contendo NLS, junto a caracterização da estrutura e função do CPN e outros componentes citoplasmáticas envolvidos no transporte de macromoléculas, permitiram a elucidação das principais vias relacionadas a esse tipo de transporte, sendo que, a mais bem elucidada é a chamada Via Clássica de Importação Nuclear dependente de Importinas.

Atualmente, inúmeros estudos relacionam essa via de importação nuclear a processos e alterações metabólicas que ocorrem em diferentes organismos. O vírus da dengue, por exemplo, utiliza desse maquinário para o transporte de proteínas virais para dentro do núcleo da célula hospedeira, auxiliando na replicação do DNA viral (Hannemann et. al., 2013), já o vírus do ebola induz a produção de proteínas virais na célula hospedeira que bloqueiam o transporte de proteínas da célula, essenciais ao funcionamento adequado do organismo (Shabman et. al., 2011). Alterações na expressão de uma das proteínas Importinas em humanos podem estar relacionadas ao desenvolvimento de câncer gástrico, tornando-a um potencial biomarcador para predição do potencial maligno de tumores (Altan et. al., 2013).

Todos esses estudos corroboram a necessidade de continuar avançando na compreensão detalhada do processo de transporte de macromoléculas em diferentes organismos, destacando suas especificidades e possibilitando o uso dessas informações para o desenvolvimento de ferramentas biotecnológicas, inclusive o desenho de drogas.

1.2. Via Clássica de Importação Nuclear.

A Via Clássica de Importação Nuclear é a via de transporte nuclear melhor caracterizada, com todos seus componentes já identificados. Simplificadamente, o processo compreende a junção das proteínas transportadoras $Imp\alpha$ e $Imp\beta$, seguido do reconhecimento do substrato a ser transportado pela $Imp\alpha$ e a passagem do complexo trimérico pelo CPN para o interior do núcleo onde ocorre a liberação do substrato e a reciclagem das proteínas transportadoras para atuarem em um novo ciclo (**Figura 1**).

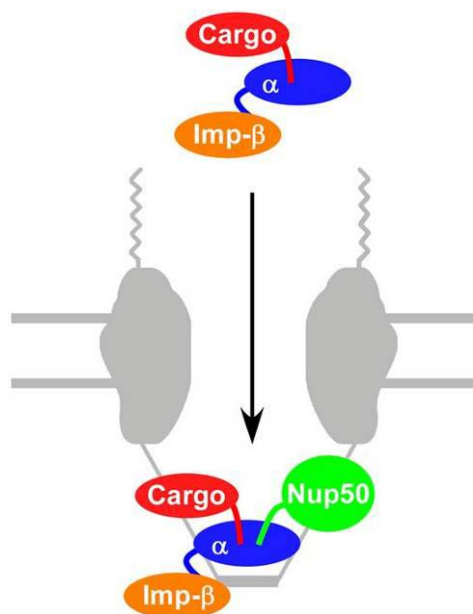


Figura 1: Esquema simplificado da Via Clássica de Importação Nuclear. A Importina- α , representada em azul, forma um complexo com a Importina- β (laranja) e reconhece a proteína a ser transportada (vermelho). O complexo trimérico atravessa o CPN (cinza) e, dentro do núcleo, o complexo se desfaz, com o auxílio do filamento de nucleoporina Nup50 (verde), liberando o substrato (Sun et. al., 2013).

Para a formação do complexo heterodimérico entre as proteínas transportadoras Imp α e Imp β , a proteína Imp β ancora-se à Imp α através do deslocamento de seu domínio autoinibitório, ou domínio IBB (*Importin beta binding* - Domínio de ligação a Imp β ; Kobe, 1999). Em seguida, o reconhecimento da proteína a ser transportada (substrato) ocorre devido à presença de uma NLS (Dingwall *et. al.*, 1991). As NLSs são reconhecidas pela Imp α , formando o complexo de importação nuclear. Após a formação do complexo, a Imp β ancora-se aos filamentos de Nucleoporinas do CPN, com destaque para a Nucleoporina Nup50, fazendo com que o complexo formado pela Imp α/β e a proteína contendo a NLS, atravesse o poro, chegando ao núcleo (Moore & Blobel, 1993; Gorlich & Mattaj, 1996).

No núcleo, o complexo Ran-GTP se liga a Imp β , promovendo o seu desacoplamento do domínio IBB, o qual é responsável pelo mecanismo auto-inibitório da Imp α e passa a competir com o sítio de ligação da NLS do substrato (Kobe, 1999). Juntos, domínio IBB, Nucleoporina (Nup50) e o complexo receptor de exportação da Imp α , formado pela proteína Exportina CAS + RanGTP, trabalham juntos para a liberação do substrato. Após a liberação do substrato, a Imp α ligada ao complexo RanGTP + CAS é transportada ao citoplasma. No citoplasma, a proteína RanBP1 se liga a RanGTP, e estimula a ação da proteína RanGAP que promove a hidrólise da RanGTP, liberando a Imp α do complexo. A Imp β volta ao citoplasma

através da ligação com RanGTP, que ao ser hidrolisada através da ação de RanBP1 e RanGAP, libera a Imp β para a atuação em um novo ciclo de transporte junto a Imp α (Cingolani *et al.*, 2002; Lee *et al.*, 2003; Lange *et al.*, 2007). A proteína Ran ligada a GTP se liga fortemente a Imp β . Como no citoplasma a forma dominante da Ran é ligada a GDP e no núcleo, está ligada ao GTP, no citoplasma a Imp β está livre para se ligar a Imp α . O processo completo, com todos os componentes envolvidos, pode ser observado na **Figura 2**.

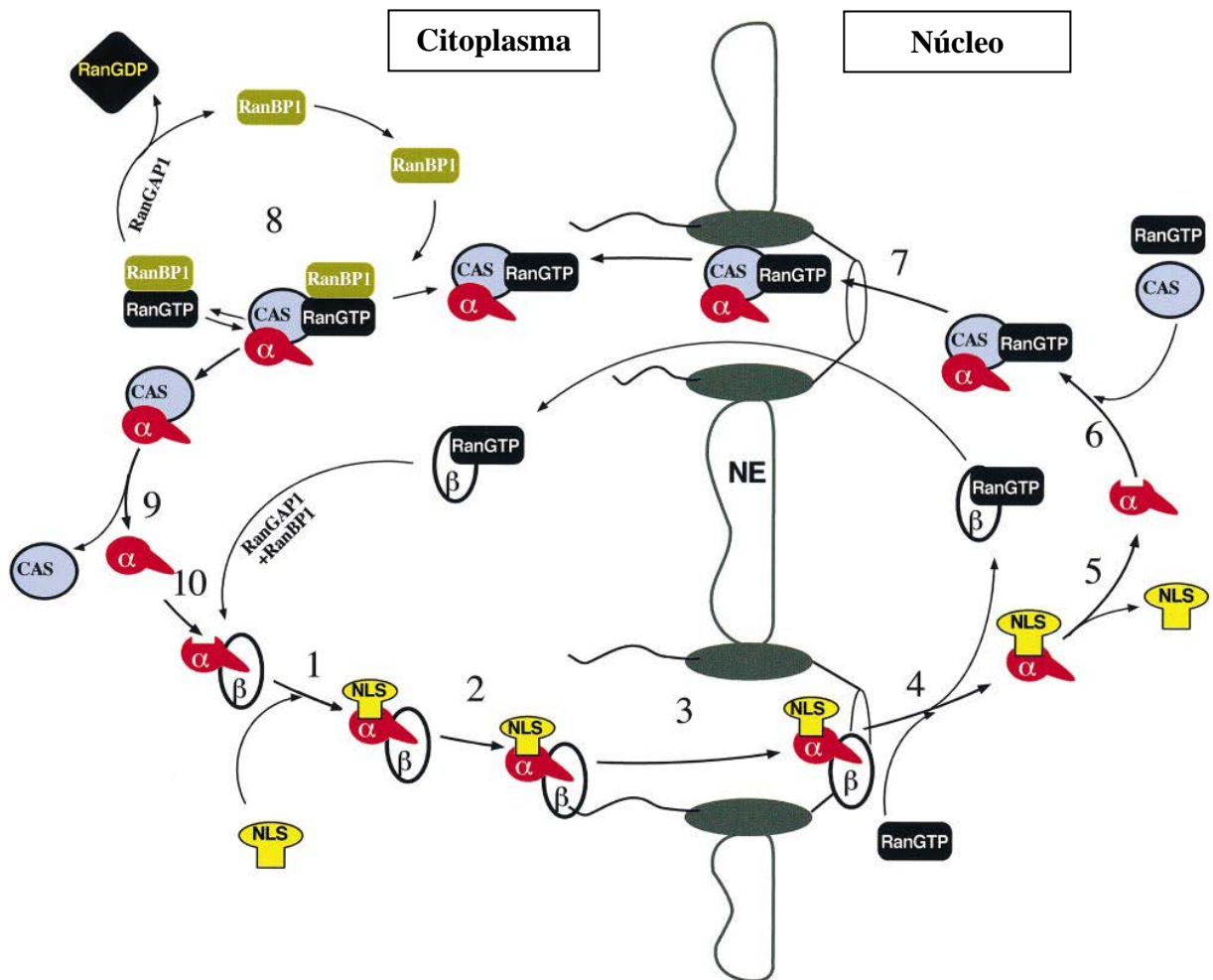


Figura 2: Esquema ilustrado da Via Clássica de Importação Nuclear. 1) Formação do complexo Imp α / Imp β /proteína contendo NLS; 2) Imp β ancora -se ao filamento de Nucleoporina do CPN; 3) Translocação do complexo pelo CPN; 4) No núcleo, Imp β se liga a RanGTP, liberando o complexo Imp α /NLS. A Imp β ligada a RanGTP volta ao citoplasma, onde as proteínas RanGAP1 e RanBP1 hidrolisam a RanGTP e permitem a liberação da proteína para um novo ciclo de transporte. 5) O domínio IBB da Imp α passa a competir com o sítio de ligação de NLSs da proteína, induzindo a liberação do substrato; 6) Imp α se liga ao complexo de exportação nuclear formado pela proteína de exportação CAS junto a RanGTP; 7) O complexo Imp α /CAS/RanGTP atravessa o CPN em direção ao citoplasma; 8) RanBP1 e RanGAP1 atuam juntas na hidrólise da RanGTP e liberação do complexo CAS/Imp α ; 9) A exportina CAS se dissocia da Imp α ; 10) Imp α fica livre para atuar em um novo ciclo de importação nuclear. (Gorlich,1998).

1.3. Complexo Poro Nuclear (CPN).

A dupla membrana que constitui o envoltório nuclear é recoberta por poros que permitem a passagem de micro e macromoléculas. Os poros nucleares são estruturas complexas formadas por diferentes proteínas que, juntas, formam canais seletivos que conectam citoplasma e nucleoplasma e recebem o nome de Complexo Poro Nuclear (CPN).

O CPN possui, aproximadamente, 120MDa e 125nm de diâmetro e apresenta uma estrutura conservada em diferentes organismos (DeGrasse *et. al.*, 2009). Através da técnica de microscopia eletrônica foi possível observar que o CPN possui uma estrutura cilíndrica, octagonal e simétrica (Gall,1967) (**Figura 3**). Essa estrutura é composta por cerca de 450 cópias de 30 diferentes proteínas, denominadas Nucleoporinas, que formam oito unidades simétricas chamadas raios. Os raios são conjugados por anéis coaxiais que revestem a superfície da membrana do poro e formam um canal central. A partir dessa estrutura central, emergem filamentos, tanto na face citoplasmática quanto nucleoplasmática do poro, que se juntam na região circular distal, formando uma “cesta nuclear” (Fernandez-Martinez e Rout, 2012).

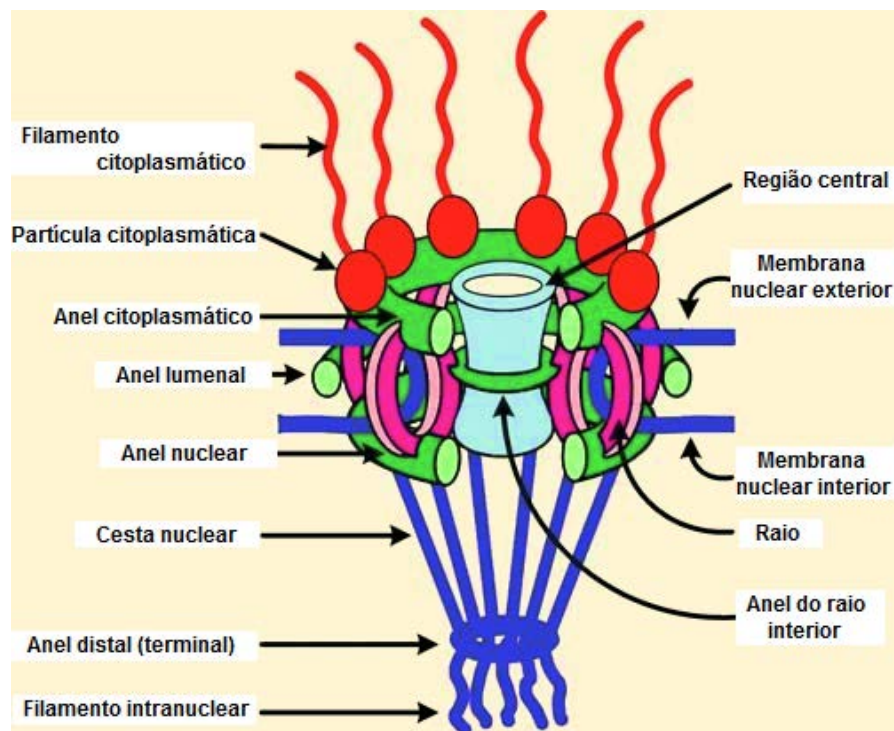


Figura 3: Esquema ilustrado do Complexo Poro Nuclear (CPN). Em verde, os anéis formados pelas proteínas Nucleoporinas, ancorados no envelope nuclear (azul). Os filamentos das Nucleoporinas, voltados ao citoplasma, são mostrados em vermelho e os filamentos voltados ao nucleoplasma, formando a “cesta nuclear”, em azul. As unidades simétricas chamadas raios (rosa), sustentam a estrutura e formam o canal central (azul claro) (Sorokin *et. al.*, 2007).

Por sua vez, as Nucleoporinas, arranjadas repetidamente por todo o CPN, são divididas em três principais grupos: grupo transmembrana, formado por Nucleoporinas que ancoram o CPN ao envoltório nuclear; grupo de Nucleoporinas com conformação α -solenóides e β -propulsores, voltadas para o interior do CPN; e Nucleoporinas que apresentam sequências repetidas e conservadas dos aminoácidos fenilalanina e glicina (motivos FG), as quais formam a armação central e estrutura periférica do CPN (Lim *et. al.*, 2008; Rout *et. al.*, 2000). As repetições dos motivos FG, nesse último grupo, formam uma estrutura coordenada que funciona como um “portão virtual”, que através de movimentos rápidos, afastam moléculas não-específicas, enquanto proteínas transportadoras, como a Importina- β , ancoram-se nas regiões dos motivos FG, através de interações hidrofóbicas com o anel da fenilalanina, e atravessam o NPC junto a suas “cargas” (Krull *et. al.*, 2004; Rout *et. al.*, 2000; Isgro e Schulten, 2005, 2007 a,b).

1.4. Importina- β (Imp β).

A proteína Imp β , com cerca de 97 kDa é um dos constituintes do complexo trimérico que atravessa o CPN na Via Clássica de Importação Nuclear. Inicialmente identificada como Importina 90 (Gorlich, 1995), a Imp β é a proteína responsável pelo contato com proteínas do CPN e translocação do complexo de importação para o interior do núcleo. Para a formação do complexo de importação, a Imp β interage com a Imp α acontece através do chamado domínio IBB, que é formado por resíduos básicos presentes na região N-terminal da Imp α , os quais interagem com resíduos ácidos da Imp β (Gorlich *et. al.*, 1996; Kobe, 1999).

A Imp β apresenta uma estrutura formada pela repetição de, cerca de, 19 motivos HEATs, sendo cada um destes formados por duas hélices- α conectadas por um curto *turn* (Andrade & Bork, 1995) (**Figura 4**). A região côncava da Imp β entre os HEATs 7 a 19 é responsável pelo reconhecimento do domínio de interação com a Imp α (IBB). Durante essa interação, a proteína adota uma conformação globular com as regiões N- e C- terminal quase se tocando (Cingolani *et al.*, 1999; Stewart, 2006; Cook *et al.*, 2007; **Figura 4**). Já a região côncava entre os HEATs 1 a 6 está envolvida na interação com a proteína Ran ligada a GTP. Nesse caso, a Imp β adota uma conformação alongada que pode ser responsável pela dissociação do complexo de importação (Chook & Blobel, 1999). A região de interação da Imp β com Nucleoporinas do CPN ocorre através da porção convexa da Imp β (Bayliss *et al.*, 2000).

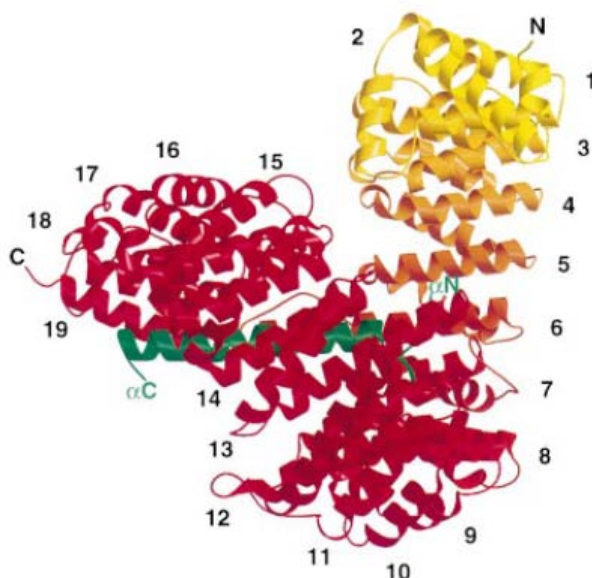


Figura 4: Estrutura da Imp β complexada ao domínio IBB. A proteína Imp β , representada em fitas, está colorida, progressivamente, de vermelho a amarelo, conforme a cadeia procede do C- ao N-terminal. O domínio IBB está representando em verde. Os domínios HEATS, estão numerados de 1 a 16 (Cingolani,1999).

1.5. Importina- α (Imp α).

A proteína Imp , inicialmente identificada como Importina 60 (Gorlich, 1995), pertence á família das proteínas chamadas carioferinas. Apresenta cerca de 55kDa e participa ativamente do mecanismo de transporte de macromoléculas para o interior do núcleo celular, através do reconhecimento de NLSs, presentes nas moléculas a serem transportadas. A Imp α funciona como um “adaptador”, ligando-se à NLS da macromolécula a ser transportada e depois se ancorando a Imp β , que media a passagem do complexo através do CPN.

A Imp α apresenta uma estrutura conservada em diferentes organismos, sendo que alguns destes organismos apresentam mais de uma isoforma da Imp α . Através da análise das sequências dessas proteínas, foi possível classifica-las em três famílias: Imp α 1, Imp α 2, e Imp α 3 (Kohler *et al.*, 1999). As diferenças entre as proteínas de cada família podem estar relacionadas com especificidades da Imp α em cada organismo.

Através da elucidação da estrutura molecular da Imp α de organismos-modelo como *Saccharomyces cerevisiae* (Conti *et. al.*, 1998; Imp α Sc) e *Mus musculus* (Kobe,1999; Imp α Mm; **Figura 5**), foi possível observar dois principais domínios funcionais responsáveis pela atuação da proteína no mecanismo de transporte celular via reconhecimento de NLSs: o primeiro consiste em uma curta sequência N-terminal envolvida na ligação com a Imp β e

formação do complexo de transporte (domínio IBB) (Gorlich & Mattaj, 1996; Moore & Blobel, 1993); o outro domínio é formado por dez repetições em *armadilho* (ARM) que formam os dois sítios de ligação para sequências NLSs (Gorlich *et. al.*, 1994; Peifer & Reynolds, 1994): o sítio principal que, usualmente, compreende a região entre os ARMs 2-4 e o secundário, entre os ARMs 7-8 (**Figura 5**). Cada motivo ARM é constituído, normalmente, por três hélices- α (chamadas de hélices H1, H2 e H3) que, devido à angulação que formam entre si, determinam a estrutura curva e alongada da proteína (Kobe, 1999; Andrade *et al.*, 2001; **Figura 5**). Nas hélices H3 encontram-se resíduos de triptofano e asparagina que formam uma superfície ideal para a ligação de NLSs.

A elucidação da Imp α Sc nativa e complexada ao peptídeo SV40 NLS permitiu caracterizar a ligação da proteína com NLSs e sugeriu sua participação no mecanismo de transporte nuclear na forma homodimérica (Conti *et. al.*, 1998). Posteriormente, com os estudos da proteína Imp α Mm, verificou-se que, para efetuar a importação de proteínas para o núcleo, a Imp α forma um complexo com a proteína Imp β , resultando em um heterodímero (Kobe, 1999). Nesse mesmo trabalho, reconheceu-se um mecanismo auto-inibitório da Imp α , quando nativa, onde a sequência N-terminal (domínio IBB) interage com o domínio formado pelas repetições dos motivos ARM.

As estruturas da Imp α Mm (Kobe, 1999) e Imp α de *Homos sapiens* (Imp α Hs; Pumroy *et. al.*, 2012; Tarendau *et. al.*, 2007) já elucidadas foram obtidas a partir da cristalização da proteína junto a NLSs, devido a dificuldade de cristalizá-las na forma nativa. Suspeita-se, portanto, que a Imp α apresenta uma menor estabilidade na ausência de ligantes nos sítios de ligação de NLSs. Mudanças na estabilidade da proteína estão relacionadas com variações no perfil oligomérico da mesma, o que afeta diretamente o processo de cristalização. Portanto, a verificação de variações na oligomerização da Imp α na presença ou ausência de ligantes é parte importante para a compreensão e delineamento de experimentos de cristalização da proteína.

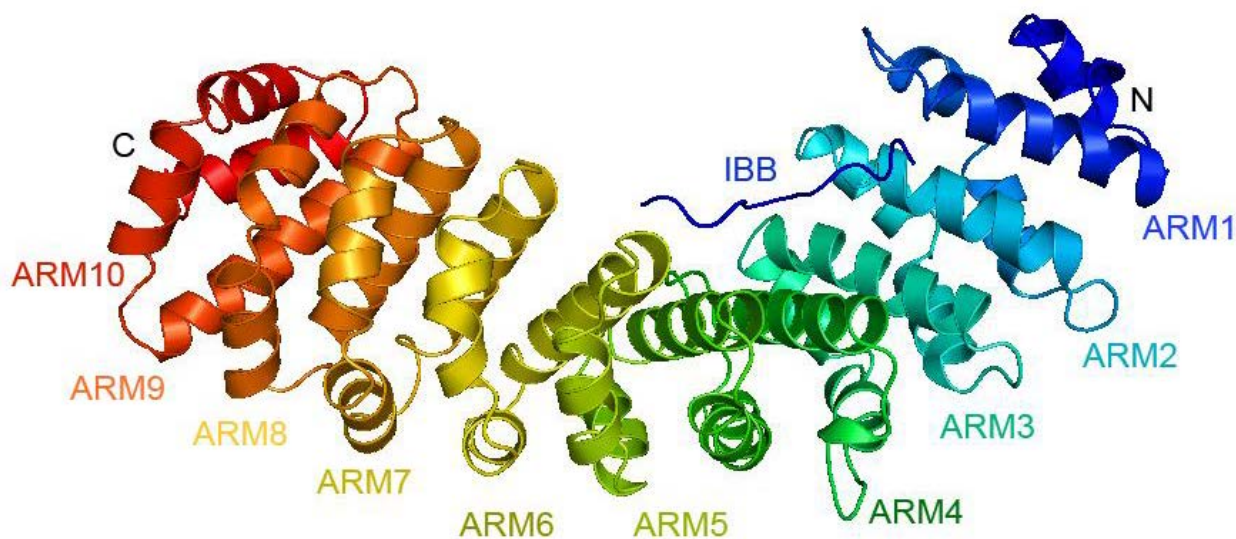


Figura 5: Estrutura da Impα de *Mus musculus* (ImpαMm). Estrutura da ImpαMm, destacando os dez motivos ARMs e o domínio IBB (Kobe, 1999). Cada motivo ARM está representado por uma cor, destacando o conjunto de três hélices-α que formam cada motivo.

1.6. Distribuição filogenética de Importina-α homólogas.

A similaridade entre os motivos ARMs presentes na Impα e os motivos HEATs da Impβ, sugerem que essas proteínas surgiram de um ancestral comum (Goldfarb *et. al.*, 2004). Como os motivos ARMs estão presentes apenas em organismos eucariotos e os motivos HEATs são encontrados tanto em eucariotos quanto procariotos, é provável que esse ancestral comum é proveniente de procariotos e a Impα evoluiu para estender e refinar as funções de transporte de uma Impβ progenitora (Goldfarb *et. al.*, 2004). Enquanto o genoma de *S. cerevisiae* apresenta o gene para apenas uma Importina-α, o genoma de *H. sapiens* codifica seis diferentes tipos de Importina-α (Goldfarb *et. al.*, 2004). Essas variedades de Impα foram agrupadas em três diferentes famílias α1, α2 e α3 (Köhler *et. al.*, 1997; 1999; **Figura 6**), sendo que membros de diferentes subfamílias tem cerca de 50% de identidade, enquanto membros da mesma família apresentam cerca de 80% de identidade (Köhler *et. al.*, 1999). As proteínas da família α1 são encontradas em todos os organismos eucariotos (animais, plantas, fungos), enquanto as proteínas das famílias α2 e α3 estão presentes apenas em animais metazoários (Köhler *et. al.*, 1997; **Figura 6**). Essa distribuição levanta a possibilidade de que o surgimento das proteínas das famílias α2 e α3 está relacionado ao desenvolvimento e diferenciação de células e tecidos. Essa hipótese é evidenciada quando se observa que Impα1, α2 e α3 apresentam padrões de expressões distintos em tecidos e células tipo-específicas e em

diferentes estágios do desenvolvimento de animais (Köhler *et. al.*, 1997; Goldfarb *et. al.*, 2004). Além disso, as proteínas das diferentes famílias podem apresentar preferências por NLSs específicos (Köhler *et. al.*, 1999; Sekimoto *et. al.*, 1997; Fang *et. al.*, 2001; Mason *et. al.*, 2002). Estudos com *Drosophila melanogaster* que possui um exemplar de Imp α de cada subfamília, mostraram que a Imp α 2 é essencial no processo de gametogênese das fêmeas da espécie (Mason *et. al.*, 2002), já a Imp α 3 é a única que promove o transporte de um fator de transcrição que está relacionado a produção de uma proteína de choque térmico (*heat shock protein*) durante um estágio específico de desenvolvimento de *D. melanogaster* (Fang. *et. al.*, 2001). Um estudo funcional comparativo entre as proteínas Imp expressas em *H. sapiens*, Imp α 2 de *Xenopus* e Imp α 1 de *S. cerevisiae*, mostrou diferenças, entre as proteínas das diferentes famílias, na eficiência de importação de determinados substratos com NLSs, especialmente quando eram oferecidos dois ou mais substratos ao mesmo tempo (Köhler *et. al.*, 1999).

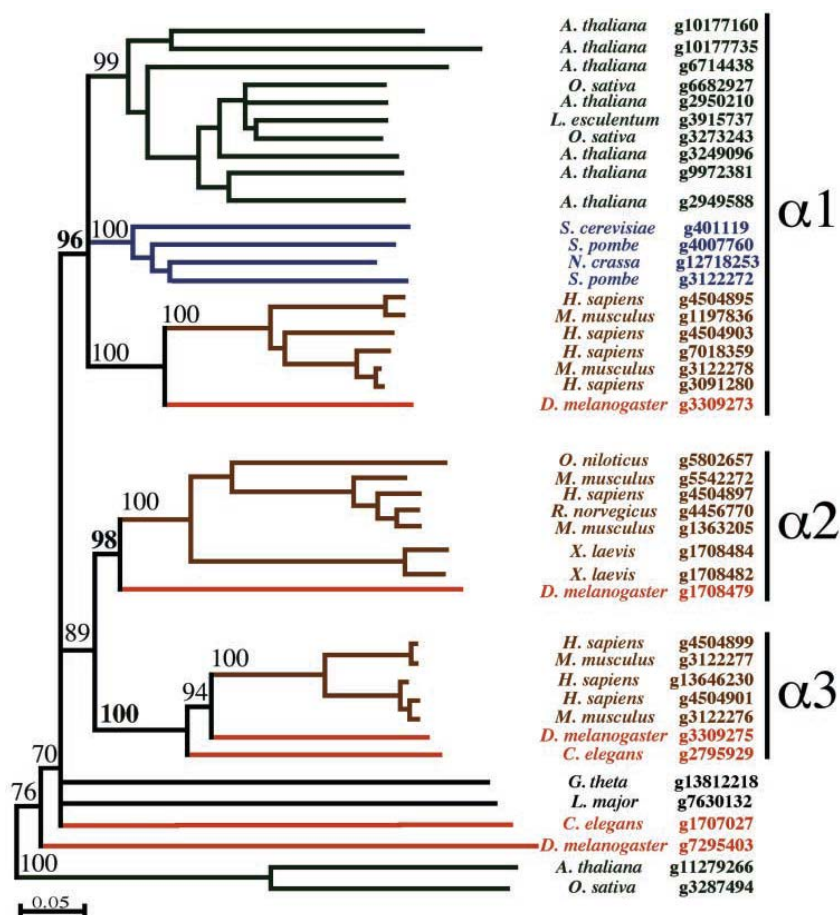


Figura 6: Árvore filogenética da proteína Importina- α . Distribuição filogenética da proteína Imp α , destacando a divisão das proteínas homólogas em três famílias α 1, α 2, α 3. Os genes provenientes de organismos vertebrados são mostrados em marrom, invertebrados em vermelho, plantas em verde e fungos em azul (Mason *et. al.*, 2002).

1.7. Sequências de Localização Nuclear (NLSs).

As Sequências de Localização Nuclear (NLSs) são parte essencial do mecanismo de transporte nuclear mediado por Importinas. Na via clássica de importação nuclear, as proteínas que apresentam as chamadas NLSs clássicas (cNLSs) são reconhecidas pela Imp α , que funciona como um adaptador para a ligação com a Imp β qual promove o transporte para o interior do núcleo. As cNLSs são caracterizadas pela presença de um ou mais grupos de aminoácidos básicos (Kobe, 1999). As cadeias laterais básicas das cNLSs interagem especificamente com resíduos dos aminoácidos triptofano e asparagina, presentes na porção côncava da Imp α (Conti *et al.*, 1998; Conti & Kuriyan, 2000; Fontes *et al.*, 2000; Stewart, 2006; Stewart, 2007).

As cNLSs são divididas em dois grupos: as monopartidas, que apresentam um grupo de aminoácidos básicos e são exemplificadas pela cNLS do antígeno T do vírus de símio 40 (SV40 NLS; ¹²⁶PKKKR¹³²KV), e as bipartidas, exemplificadas pela cNLS da Nucleoplasmina (¹⁵⁵KRPAATKKAGQAK¹⁷⁰KKK) ou da fap-endonuclease 1 (FEN1 NLS, ³⁵¹SSAKRKEPEPKGSTKKK³⁷²AKT), que apresentam mais de um grupo de aminoácidos básicos. As cNLSs bipartidas ligam-se a dois sítios de ligação da proteína Imp α (sítios principal e secundário) ao mesmo tempo, enquanto as monopartidas se ligam, preferencialmente, ao sítio de ligação principal. Nesses sítios, posições específicas são importantes para o reconhecimento de NLSs. No sítio principal, essas posições são denominadas P1-P6, enquanto no sítio secundário há as posições P1'-P4'.

Ensaio de afinidade mostraram que a sequência de aminoácidos requerida para o reconhecimento das cNLSs monopartidas pela Imp α deve apresentar uma lisina na posição P1, seguida por resíduos básicos nas posições P2 e P4, seguindo o seguinte padrão: K(K/R)X(K/R) (Hodel, 2001; revisão em Lange *et al.*, 2007; **Tabela 1**). Já as cNLSs bipartidas são caracterizadas pelo seguinte padrão: KRX10-12KRXXK, em que os dois grupos de resíduos básicos (lisinas e argininas) são separados por uma região de ligação, formada por 10 a 12 aminoácidos (Dingwall & Laskey, 1991; Fanara *et al.*, 2000; Lange *et al.*, 2007; **Tabela 1**).

A interação da NLS monopartida SV40 NLS com a Imp α foi apresentada através da elucidação das estruturas de Imp de mamífero e levedura (Kobe, 1999; Conti, 1998) complexadas ao peptídeo. Desde então essa sequência é utilizada como referência para estudos envolvendo a Imp α , tanto análises estruturais quanto ensaios de importação nuclear e interação proteína-proteína.

Alguns substratos apresentam NLSs que não seguem o padrão das cNLS, são as chamadas NLS não-clássicas. As NLS não-clássicas podem apresentar resíduos positivos e maior hidrofobicidade do que as cNLSs e mesmo assim interagem com a Imp α , algumas vezes em regiões diferentes das cNLS (Chen *et. al.*, 2005), ou interagem diretamente com a Imp β (Cingolani *et. al.*, 2002). Um estudo de Kosugi e colaboradores (Kosugi *et. al.*, 2009) sugeriu a identificação de NLSs não clássicos, no entanto, pouco ainda se sabe sobre a identificação dessas sequências e o reconhecimento das mesmas pela Imp α .

Tabela 1. Sequência de aminoácidos requerida para o reconhecimento de NLSs pela Imp α . As posições dos aminoácidos no sítio principal são representadas por P1-P6 e no sítio secundário por P1'-P4'. A região de ligação dos NLSs bipartidos requer a presença de 10 a 12 aminoácidos entre os dois sítios (região conhecida como *linker*).

	Sítio secundário				Região de ligação	Sítio principal						
	P1'	P2'	P3'	P4'		P1	P2	P3	P4	P5	P6	
NLS: Impα												
Monopartidos						K	K/R	X	(K/R)			
Bipartidos		K	R	X	10 - 12	K	R	X	K			

1.8. Importina- α e potenciais NLSs do fungo *Neurospora crassa*.

O fungo *Neurospora crassa* pertence ao filo Ascomyceta e é um fungo filamentosamente empregado como modelo para estudos de biologia celular e molecular. No século 19, estudos de biologia celular demonstraram que os fungos filamentosos podem causar doenças em animais e plantas, tornando-os alvo de pesquisas clínicas (Bristowe, 1854; Kühn, 1858). Em 1940, o trabalho de Beadle e Tatum (Beadle e Tatum, 1941) apresentou a relação entre genes e proteínas, através de experimentos com o fungo filamentosamente *Neurospora crassa*. A partir desta década, inúmeros trabalhos subsequentes com o fungo apresentaram importantes aspectos acerca da metilação do DNA, silenciamento gênico pós-transcricional, reparo de DNA e transporte de proteínas (Davis, 2000). No século 20, consolidou-se o interesse das pesquisas voltadas para questões fundamentais de desenvolvimento e crescimento dos fungos e o fato de *Neurospora crassa* ser um fungo filamentosamente multicelular, permitiu que se consolidasse como um organismo- modelo, caracterizado como um sistema que permite estudar diferenciação celular e desenvolvimento além de outros aspectos das células eucarióticas (Davis, 2002). A descoberta do genoma de *N. crassa* (Galagan, 2003), acelerou ainda mais a investigação de muitos aspectos da biologia desse organismo.

Dentro desse panorama, pesquisas envolvendo o mecanismo de regulação da expressão gênica e o controle pré e pós-transcricional em fungos despontam como os novos desafios de pesquisas nessa área. O transporte intracelular de macromoléculas, que compreende a via clássica de importação nuclear, ainda não é bem elucidada em fungos, mas, na literatura, estudos apontam a interação da Imp α de fungos com potenciais NLSs presentes em fatores de transcrição para que estes sejam transportados ao interior do núcleo (Stinnett *et. al.*, 2007; Araújo-Bazán *et. al.*, 2008).

Alguns fatores de transcrição relacionados à regulação da concentração de glicogênio em diferentes condições fisiológicas e ambientais foram identificados através de experimentos que empregaram mutações sítio-dirigidas, onde foi possível verificar variações no acúmulo de glicogênio, de acordo com os genes silenciados (Gonçalves, 2011). Os fatores de transcrição identificados apresentam sequências de aminoácidos característicos de potenciais NLSs e que ainda não foram caracterizados. Essas sequências não apresentam características de um NLS clássico, o que sugere que em *Neurospora* o mecanismo de reconhecimento de proteínas a ser transportadas para o núcleo seja caracterizado por sequências específicas. Portanto, a caracterização estrutural da Imp do fungo *Neurospora crassa* (Imp α Nc) é de grande importância para posterior caracterização de peptídeos NLSs específicos de fungos, o que viabilizará o reconhecimento de especificidades no transporte desses fatores de transcrição, assim como o mecanismo de reconhecimento de NLSs não-clássicos.

2. Objetivos

O objetivo principal deste trabalho é identificar características específicas da Imp α Nc no reconhecimento de sequências de localização nuclear. Para este fim, os objetivos específicos são:

- Caracterização biofísica da proteína Imp α Nc para avaliar a viabilidade e comportamento oligomérico da proteína recombinante e auxiliar na cristalização.
- Cristalização e elucidação da estrutura da proteína Imp α Nc recombinante.
- Quantificar a afinidade da Imp α Nc e o peptídeo SV40 NLS.

3. *Expressão e purificação das proteínas Imp α Nc e Imp α Mm recombinantes.*

3.1. Material e Métodos

3.1.1. Clonagem da proteína recombinante Imp α Nc.

A proteína Importina- α de *Neurospora crassa* (Imp α Nc) foi clonada pela Dra. Fernanda Zanolli Freitas, do grupo da Prof^a Maria Célia Bertolini do Departamento de Química da UNESP de Araraquara. A clonagem foi realizada através da inserção do cDNA da proteína truncada em pET28A (Novagen) e a expressão em *Escherichia coli*, linhagem Rosetta (DE3) pLysS (Novagen). A proteína recombinante foi produzida fusionada a uma cauda de histidina e truncada, através da exclusão da porção N-terminal correspondente ao domínio auto-inibitório:

¹MADRYIPEHRRTQFKAKSAFKPDELRRRREEQQVEIRKAKREENLAKRRGIGAGDSR
PGASLGAAPDSDDENPP⁷⁴.

3.1.2. Expressão e purificação da proteína recombinante Imp α Nc.

A bactéria *E. coli* contendo o plasmídeo de expressão pET28a, armazenada como cultura permanente em nitrogênio líquido (-196°C), foi pré-inoculada em tubos tipo falcon contendo 15 mL de meio LB (Luria Bertani) líquido e 15 μ L dos agentes de seleção cloranfenicol e canamicina (50mg/ml). O pré-inóculo foi mantido sob agitação a 140 rpm, 37°C, por 12 h para crescimento da cultura. Em seguida, o conteúdo dos falcons foi dispensado em erlenmeyers contendo 500 mL de meio LB líquido e 500 mL dos antibióticos cloranfenicol e canamicina (a 50 mg/ml) e mantidos sob agitação a 140 rpm, 37 °C até atingir a densidade óptica Abs₆₀₀= 0,6, medida em espectrofotômetro. Após essa etapa, foi adicionado aos erlenmeyers 5 mL de IPTG (Isopropil β - d- tiogalactopiranosídeo) a 100 mM, para indução da produção da proteína recombinante. Os erlenmeyers foram submetidos a agitação a 140 rpm, 28 °C, por 4 h. Ao fim desse período, o conteúdo dos frascos foi centrifugado a 11.5 g por 30 min para sedimentação das células bacterianas, permitindo a separação destas (*pellets*) do meio de cultura.

Os *pellets* obtidos após a centrifugação foram submetidos à lise celular para ruptura das células bacterianas e obtenção da proteína solúvel, através da resuspensão das células em tampão [50 mM de Hepes, pH 7, 150 mM de Imidazole, 500 mM de NaCl contendo lisozima, DNase, RNase e 1 μ g/ml de inibidores de proteases (Aprotinina, Leupeptina e Pepstatina)]. A solução foi mantida em gelo, por cerca de 1 h e agitada, periodicamente. Ao fim desse

período, a solução foi centrifugada a 11.5 g, por 30 min, 4 °C e a porção sobrenadante, onde se esperava encontrar a proteína solúvel, foi armazenada para a etapa posterior de purificação por cromatografia líquida.

A purificação da proteína (contida no extrato obtido após a lise celular), foi realizada por cromatografia de afinidade, utilizando uma coluna de níquel His-TrapHP (GE Healthcare), acoplada ao aparelho ÄKTA Purifier System (GE Healthcare). O extrato foi injetado na coluna a um fluxo de 0,5 mL/min e pressão máxima de 0,3 MPa. Após a injeção, foi realizada a eluição de moléculas com ligação inespecífica através da passagem de tampão A contendo 20 mM de Hepes pH 7,0, 1 M de NaCl e 15 mM de Imidazol. Posteriormente, a proteína Imp α Nc foi eluída e coletada em frações de 1,0 mL, através do fluxo de tampão A junto a um gradiente de 0 a 100% de tampão B contendo 20 mM de Hepes pH7,0, 500 mM de NaCl e 300 mM de Imidazol.

Os picos de eluição observados através do cromatograma obtido foram coletados separadamente e amostras de cada pico foram submetidas á análise por gel SDS-PAGE 12% para verificação das frações mais puras. Em seguida, as frações mais puras foram submetidas à diálise para remoção do tampão com Imidazol para tampão contendo 20 mM de Tris HCl pH 8 e 100 mM de NaCl, para a manutenção da estabilidade da proteína. Para a diálise, as amostras foram armazenadas em membranas e mergulhadas em tampão, sob agitação constante, a 8 °C, por cerca de 12 h. Em seguida, essas amostras foram concentradas, utilizando o dispositivo de ultrafiltração *Amicon* (Millipore), através de centrifugação a 11,5 g. As amostras concentradas foram utilizadas em experimentos de cristalização e análises biofísicas.

3.1.3. Clonagem, expressão e purificação da proteína recombinante Importina- α de *Mus musculus*.

A proteína Imp α Mm, clonada em pET28a e truncada em sua porção N-terminal, correspondente ao domínio IBB, foi cedida pelo colaborador Prof. Kobe, da Universidade de Queensland, Austrália. A proteína foi transformada em *E. coli*, linhagem BL21(DE3) pLysS. A metodologia empregada está descrita em Teh *et al.*, 1999. As etapas de expressão e purificação da Imp α Mm seguem o mesmo protocolo já descrito para a expressão e purificação da Imp α Nc (item 3.1.2) e foram realizadas pela Msc. Andrea Coelho de Barros.

3.2. Resultados e Discussão

A proteína recombinante Imp α Nc foi produzida fusionada a uma cauda de histidina e truncada em sua porção N-terminal (resíduos 1-74), correspondente ao domínio auto-inibitório. A proteína foi produzida dessa maneira para simular sua ligação com a Importina- β , tornando assim disponível os sítios de ligação de NLSs. Após a etapa de indução da expressão da proteína recombinante e posterior lise celular e centrifugação do extrato desse inóculo, obteve-se a proteína solúvel no material sobrenadante proveniente da centrifugação. A parte solúvel do extrato de *E. coli*, foi submetida a cromatografia de afinidade (**Figura 7A**). A Imp α Nc teve eluição a partir de 30 mM de imizadol (pico 1), no entanto houve eluição simultânea de contaminantes, assim como observado no pico 2 (**Figuras 7A e 7B**). Na concentração de 90 mM de imidazol (pico 3) foi possível a eluição da Imp α Nc com maior grau de pureza (**Figuras 7A e 7B**). Frações correspondentes a esse pico foram concentradas para a utilização em experimentos em virtude da pureza das amostras.

Após a concentração das frações eluídas, obteve-se cerca de 6mg de proteína para 1 L de indução, que foi concentrada em frações a 20 mg/ml, as quais foram submetidas a diálise para a troca do tampão com Imidazol para tampão contendo 20 mM de Tris HCl pH 8 e 100 mM de NaCl. Em seguida, as amostras foram congeladas e armazenadas a -20 °C até a realização de outros experimentos.

A proteína recombinante Imp α Mm também foi eluída com alto grau de pureza na concentração de 50 mM de Imidazol, resultando em cerca de 6 mg para 1 L de indução. A proteína concentrada a 18 mg/ml foi submetida a diálise com o mesmo tampão da Imp α Nc e , em seguida, as amostras foram congeladas e armazenadas a -20 °C até a realização de outros

experimentos (dados não mostrados).

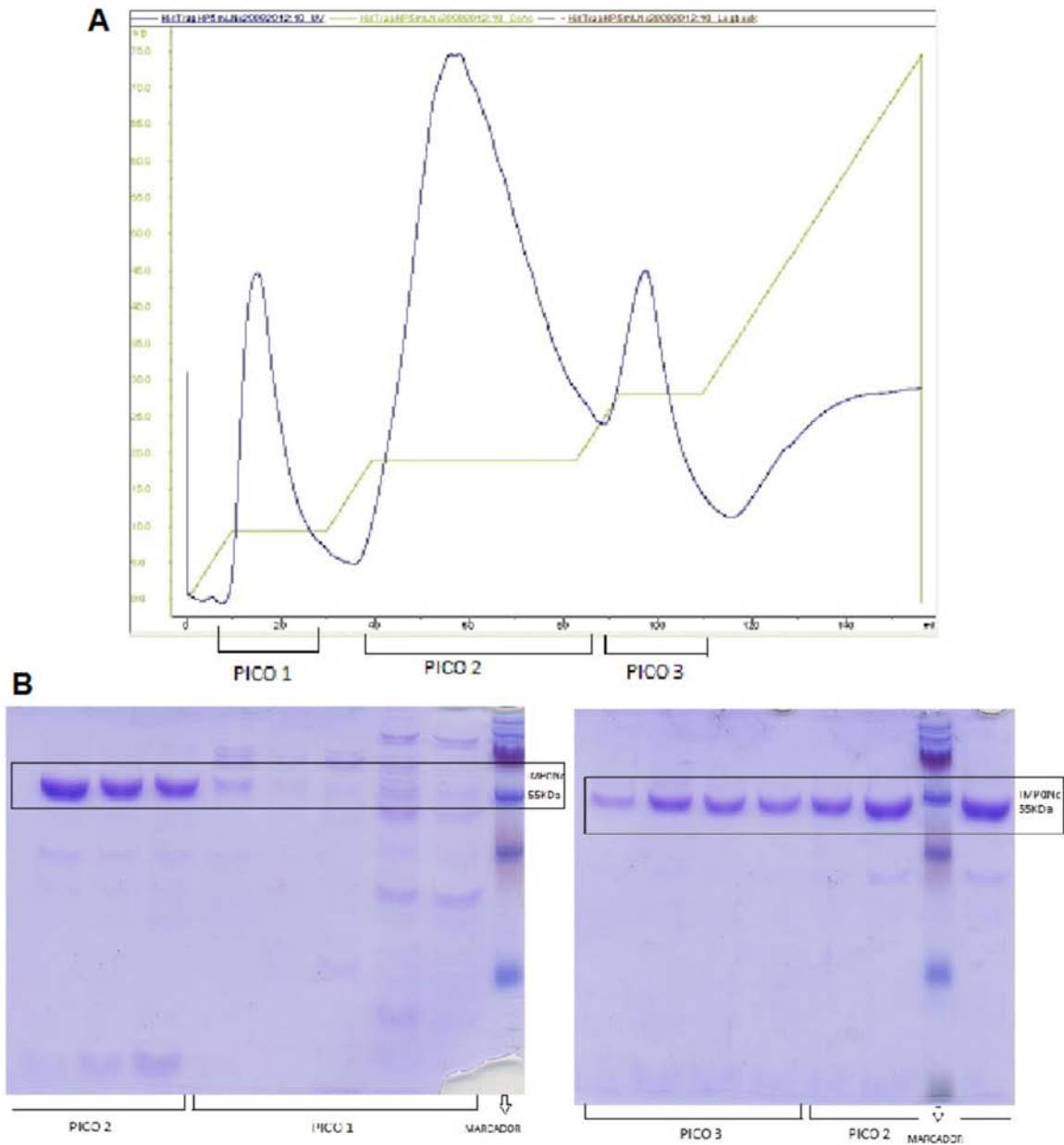


Figura 7. Cromatograma de eluição e gel SDS-Page 12% das frações obtidas na purificação da ImpαNc por cromatografia de afinidade. A) Cromatograma de eluição da ImpαNc através de cromatografia de afinidade. A linha verde apresenta as fases do gradiente do tampão de eluição. Os picos 1 e 2 correspondem a eluição da proteína junto a outras proteínas contaminantes e o pico 3 apresenta a proteína pura, conforme observado em gel SDS-PAGE 12% (B).

***4. Verificação do comportamento da
proteína ImpaNc na ausência ou
presença de NLSs.***

4.1. Material e Métodos

4.1.1. Síntese de peptídeos NLS.

O peptídeo correspondente a FEN1 NLS (³⁵¹SSAKRKEPEPKGSTKKKAKT³⁷²) foi sintetizado pela empresa Proteimax com pureza maior do que 98%. O peptídeo foi sintetizado com aminoácidos adicionais nas porções N e C- terminal para garantir a ligação eficiente nessas regiões.

4.1.2. Cromatografia Analítica por Exclusão Molecular.

O perfil oligomérico da Imp α Nc foi avaliado por cromatografia analítica por exclusão molecular, utilizando uma coluna Superdex-200 (GE Helthcare), acoplada ao equipamento ÄKTA Purifier System (GE Healthcare). A coluna foi equilibrada com tampão contendo 20 mM de Tris HCl pH 8,0 e 100 mM de NaCl e as amostras da proteína purificada foram aplicadas a um fluxo de 0,5ml/min, sendo as amostras: (1) Imp α Nc nativa e (2) Imp α Nc junto ao peptídeo FEN1 NLS na proporção de 1:8, incubadas por 20 minutos antes da cromatografia. A calibração da coluna foi realizada com a utilização de padrões de massa molecular conhecida, provenientes do *kit* de calibração (Sigma-Aldrich) composto por: Blue dextran (2.000 kDa), Beta-amilase (200 kDa), Álcool desidrogenase (150 kDa), Albumina bovina (66 kDa), Anidrase carbônica (29 kDa) e Citocromo C (12,4 kDa), sendo que essas amostras foram diluídas no mesmo tampão utilizado para equilibrar a coluna. A partir das medidas do volume de retenção das amostras da proteína e dos padrões, calculou-se o coeficiente de partição (K_{av}) através da equação $K_{av} = (V_r - V_o) / (V_c - V_o)$, onde V_r = volume de retenção, V_o = volume inicial (baseado no tempo de retenção do padrão *blue dextran*), e V_c = volume da coluna. O coeficiente K_{av} obtido para cada padrão foi representado graficamente contra o *log* do peso molecular para gerar a curva padrão. As massas moleculares estimadas para as amostras 1 e 2 referentes aos picos obtidos no cromatograma, foram calculados a partir da curva padrão.

4.1.3. Western-blot.

As amostras referentes aos picos obtidos na cromatografia analítica por exclusão

molecular foram diluídas 0,25 mg/mL e submetidas ao experimento de *western- blot* para confirmação da presença da Imp α Nc nas amostras. Inicialmente, as amostras contidas em gel de poliacrilamida SDS-PAGE 15% foram transferidas, para uma membrana de PVDF que, em seguida, foi saturada em solução de bloqueio contendo TBS 1X , 3 g de leite em pó desnatado; 0,05% Tween® 20, *overnight*. O próximo passo foi a lavagem da membrana com TBS 1X, 3 g de leite em pó desnatado, 0,25% Tween® 20, por 3 vezes, em intervalos de 5 minutos. Nessa mesma solução, a membrana foi incubada, por 1 h, junto ao anticorpo monoclonal anti-histidina produzido em rato, diluído a 1:2000 (Sigma-Aldrich) e lavada, três vezes, com solução de bloqueio por 5 minutos. Posteriormente, a membrana foi incubada em solução de bloqueio com o segundo anticorpo, anti-IgG de camundongo conjugado com peroxidase (Sigma-Aldrich), em uma diluição 1:10000, seguido de 3 lavagens com solução de bloqueio, com duração de 5 minutos cada. A revelação foi feita em câmara escura, utilizando a solução substrato *ECL plus Western Blotting Detection System* (Amersham Biosciences) em filme de raios X (Kodak).

4.1.4. Espalhamento de Luz Dinâmico.

Amostras da proteína Imp α Nc nativa, purificadas e concentradas a 3,5 mg/mL foram submetidas ao experimento de espalhamento de luz dinâmico, no equipamento DynaPro TITAN (Wyatt Technology), a fim de verificar o perfil oligomérico e polidispersidade da proteína em solução. A amostra analisada foi previamente filtrada em poro-membrana de 0,22 μ m (Millipore) e os dados foram coletados em cem medições a um ângulo fixo de 90 °C, temperatura de 4 °C e analisados com o programa *Dynamics v.6.10*.

4.2. Resultados

Foram realizados experimentos de Cromatografia Analítica de Exclusão Molecular tanto com a proteína Imp α Nc nativa (Imp α Nc, **Figura 8A**), quanto na presença do peptídeo NLS bipartido FEN1 (Imp α Nc-FEN1, **Figura 8A**). O cromatograma obtido a partir da eluição da amostra contendo a proteína nativa indicou a separação da amostra em dois picos: o primeiro (pico 1) mais largo e composto de partículas com uma massa estimada superior ao padrão de maior massa presente na curva de calibração, e o segundo (pico 2) com partículas de massa estimada de 83 kDa, valor diferente do esperado para a Imp α Nc (55 kDa), porém

muito próximo ao obtido através do experimento de DLS (90 kDa) (**Figuras 8A e 8B**). Já o cromatograma da proteína no complexo Imp α Nc + FEN1 NLS, apresentou apenas um pico (pico 3), com a massa inferida de 81,1 kDa (**Figuras 8A e 8B**).

Análises em gel de SDS-PAGE mostraram que a amostra da fração eluída da proteína complexada ao peptídeo FEN1 NLS (pico 3) apresentou apenas uma banda em torno dos 55 kDa, massa referente a Imp α Nc, enquanto que a amostra da eluição da proteína nativa (pico 2) apresentou, além da banda próxima aos 55 kDa, uma banda em torno de 105 kDa (**Figura 8C**). Através de *Western blot* foi possível confirmar que todas as bandas correspondiam a Imp α Nc (**Figura 8C**).

Os resultados obtidos no experimento de espalhamento de luz dinâmico, com a Imp α Nc nativa a 4 °C, mostraram a massa da proteína com cerca de 90 kDa, em concordância com a massa obtida na gel-filtração. A proteína também apresentou uma polidispersidade de 20,4%, mostrando uma conformação, predominantemente, monodispersa, a essa temperatura, e raio hidrodinâmico de 4,1 nm.

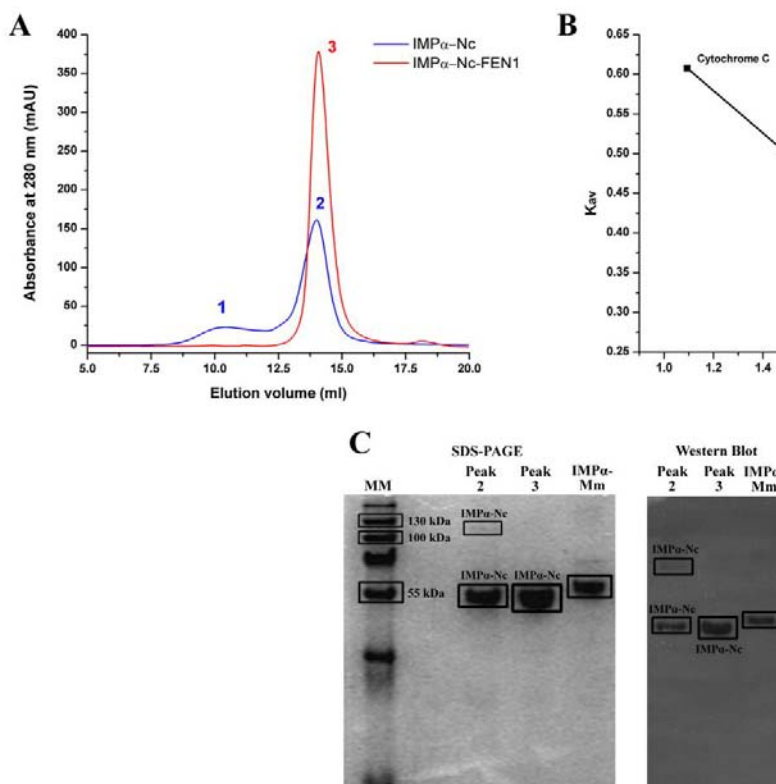


Figura 8: Verificação do perfil oligomérico da Imp α Nc. (A) Cromatografia analítica de exclusão molecular da proteína Imp α Nc com e sem a presença do peptídeo FEN1 NLS em vermelho e azul, respectivamente. (B) Curva de calibração obtida utilizando substâncias com massa conhecida. Utilizando a curva padrão e a fórmula $K_{av} = (v_e - v_o) / (v_t - v_o)$, foi calculado a massa das moléculas presentes nos picos 2 e 3 obtidos na gel-filtração. (C) Gel SDS-PAGE e *Western blot* referente às amostras obtidas nos picos 2 e 3 e à Imp α de *Mus musculus*.

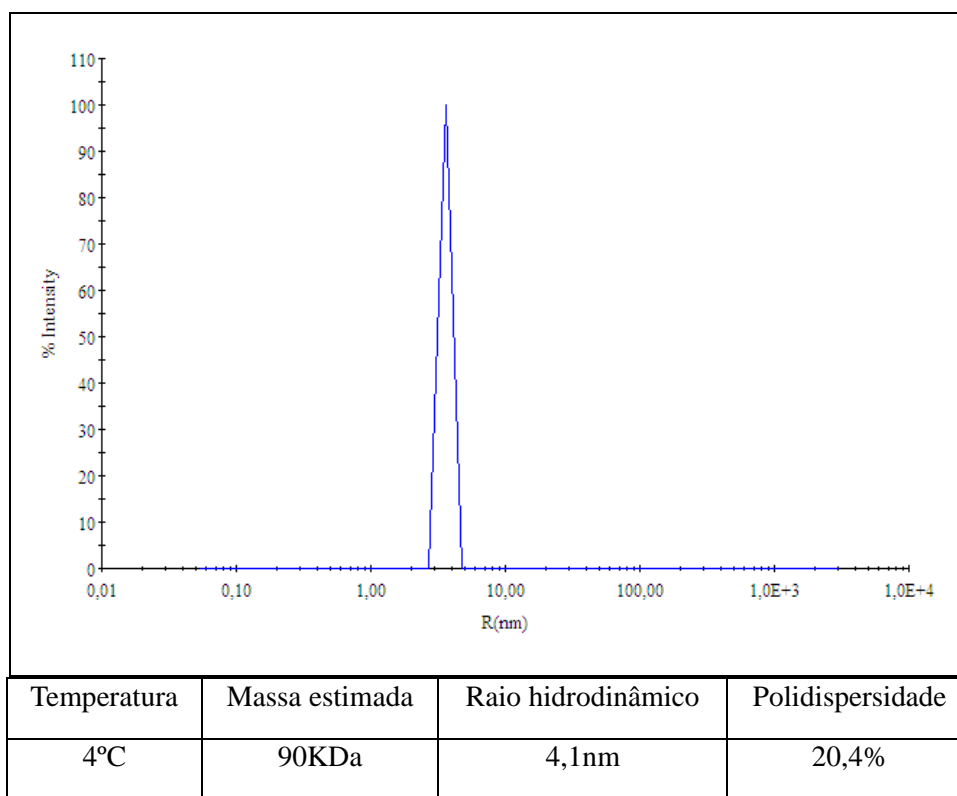


Figura 9: Gráfico de distribuição oligomérica de amostra de *Imp* obtido através de experimento de Espalhamento de Luz Dinâmico (DLS). Gráfico obtido a 4 °C, mostrando a distribuição, predominantemente, monomodal da proteína a essa temperatura, levando ao aparecimento de apenas um pico de proteína com raio hidrodinâmico estimado de 4,1 nm.

4.3. Discussão

4.3.1. Mudança no perfil oligomérico da proteína na presença de peptídeos NLSs.

Estruturas da proteína *Imp* de diferentes organismos já foram elucidadas através da cristalização da proteína junto à peptídeos NLSs ou o domínio auto-inibitório (domínio IBB) (Kobe, 1999; Hubner *et. al.*, 1999). Uma exceção foi a *Imp α* de levedura, truncada na sua porção N-terminal (*Imp α Sc*). Entretanto, essa estrutura apresentou-se na forma homodimérica (Conti *et. al.*, 1998), cuja interface inclui a região de ligação de NLSs bipartidos. A dificuldade de cristalizar a proteína nativa sugere uma menor estabilidade da proteína na ausência de uma sequência peptídica, ligada a seus sítios principal e secundário.

Experimentos de cromatografia analítica de exclusão molecular, aliado a análises por DLS e Western blot, indicaram a tendência da proteína a formar aglomerados na ausência de ligantes em seus sítios de ligação. No cromatograma da proteína Imp α Nc nativa, pôde ser observado a formação de dois picos durante a eluição, sendo que um deles apresentou um valor de massa estimado muito maior do que o esperado, formado por aglomerados da proteína, como foi corroborado pelo gel -de eletroforese e *western blot* onde, inclusive, foi possível observar o aparecimento de uma banda acima do valor esperado para a amostra do pico correspondente a proteína predominantemente monomodal, mesmo em condições desnaturantes.

Já a amostra da proteína complexada com um peptídeo NLS (Imp α Nc-FEN1 NLS) levou ao aparecimento de apenas um pico de eluição, com valor de massa estimado condizente com a proteína em sua forma monomodal. A massa estimada para o complexo Imp α Nc-FEN1 NLS na cromatografia analítica de exclusão molecular é condizente com a massa estimada da proteína em sua forma monomodal no DLS (90 kDa). Esse valor de massa maior do que o esperado (55 kDa) se deve, provavelmente, a forma mais alongada da proteína, que leva ao cálculo de um raio hidrodinâmico alto nos experimentos de DLS, uma vez que o programa de análises realiza as estimativas partindo de uma proteína globular como referência. A técnica de cromatografia analítica também parte do princípio de que as proteínas são na sua grande maioria globulares. Uma proteína alongada tem maior dificuldade de deslocamento ao longo da coluna cromatográfica, o que leva a uma eluição mais lenta da mesma.

Esses resultados indicam que a proteína tem a tendência a agregação na ausência de um ligante associado ao sítio de ligação de NLS, seja o seu domínio autonibitório ou um peptídeo NLS. Além disso, destaca a importância da realização de experimentos de co-cristalização da Imp α Nc com peptídeos NLS para sucesso na obtenção de cristais e elucidação da estrutura dessa proteína. Isso é corroborado com os cristais de Imp α Nc obtidos na presença do peptídeo SV40NLS. Somente com a utilização do peptídeo nos experimentos de cristalização foi possível o crescimento de cristais adequados para a coleta de dados e elucidação da estrutura da Imp α Nc.

***5. Estrutura cristalográfica do
complexo Imp α Nc/SV40NLS***

5.1. Material e Métodos

5.1.1. Síntese de peptídeos NLS.

O peptídeo correspondente a SV40 NLS ($^{125}\text{PPKKKRKV}^{132}$) foi sintetizado pela empresa Proteimax com pureza maior do que 98%. O peptídeo foi sintetizado com aminoácidos adicionais nas porções N e C- terminal para garantir a ligação eficiente nessas regiões.

5.1.2. Cristalização.

As amostras da proteína Imp α Nc purificadas e concentra das foram submetidas a experimentos de cristalização para a obtenção de cristais adequados a difração de raios X. Inicialmente, buscou-se uma condição de cristalização com a técnica de difusão de vapor (Mcpherson, 1982) por *hanging-drop* ou *sitting-drop* e o método da matriz esparsa (Jancarik e Kim, 1991) sendo que algumas placas foram montadas utilizando o robô de cristalização de proteínas Oryx4 (*Douglas Instruments*). Testes foram realizados tanto com a proteína Imp α Nc nativa, quanto complexada ao peptídeo SV40 NLS na proporção de 1:8, em número de moléculas. Foram utilizados *Kits* de soluções pré-estabelecidas, contendo condições de cristalização com as maiores taxas de sucesso em triagens iniciais. Os *Kits* utilizados na triagem de condições iniciais da proteína Imp α Nc foram *Screen I e II (Hampton Research)*, *JCSG + suíte (Qiagen)*, *JBScreen Classic (Jena Bioscience)* e *Screen PEG/ION da Hampton Research*. As gotas foram compostas pela mistura de 1 μL de proteína, 0,5 μL de peptídeo e 0,5 μL de solução do reservatório. Após a identificação das condições de cristalização com melhores resultados, foram realizados experimentos para a otimização dos cristais obtidos, utilizando outras técnicas de cristalização, como *microbath* e *microseeding*, além de *Kits* de aditivos e variação das proporções na solução inicial. Os melhores cristais foram montados em *loops* de *nylon* e mantidos em fluxo de nitrogênio gasoso a 100K (Oxford Cryosystems[®]) para serem submetidos à coleta de dados de difração de raios X.

5.1.3. Coleta e processamento de dados de difração de raios X.

Os melhores cristais obtidos nos experimentos de cristalização foram submetidos à difração de raios X no Laboratório Nacional de Luz Síncrotron (LNLS), em Campinas -SP,

Brasil e no National Synchrotron Light Source (NSLS) do Brookhaven National Laboratory, de Upton-NY, Estados Unidos da América. Os dados de difração de raios X de cristais do complexo Imp α Nc/SV40NLS foram coletados em dois laboratórios diferentes. Um primeiro conjunto (Imp α Nc-1) foi coletado em um comprimento de onda de 1,58 Å a 100 K na Linha D03B-MX1 do LNLS, utilizando detector MAR CCD (MAR Research). A distância entre o cristal do detector foi de 70 mm e as imagens foram adquiridas a cada 1,5° de oscilação, resultando em um total de 241 imagens. Foi coletado um segundo conjunto (Imp α Nc-2) na linha X25 do NSLS, em um comprimento de onda de 1,0 Å a 100 K, utilizando detector Pilatus. Foram coletadas imagens a cada 0,5° de oscilação, com o detector distante 270 nm do cristal, resultando em 720 imagens. O processamento dos dados dos dois conjuntos foi efetuado com programas do pacote HKL2000 (Otwinowski e Minor, 1997).

5.1.4. Elucidação da estrutura, modelagem e refinamento.

Para obter a estrutura tridimensional de uma proteína, uma das técnicas utilizadas é o método da substituição molecular (Mcree, 1993). O primeiro modelo da Imp α Nc (Imp α Nc-1) foi gerado pela utilização do programa AutoMR e o seu resultado foi enviado ao programa AutoBuild (Terwilliger *et al.*, 2008) para reconstrução automatizada do modelo. Ambos programas são parte do Phenix (Adams *et al.*, 2010).

Posteriormente, foram efetuadas etapas de modelagem e refinamento de posição de átomos e fator de temperatura. Essas etapas utilizaram os programas Coot (Emsley & Cowtan, 2004) e phenix.refine, respectivamente. Durante a etapa de modelagem foi realizado o ajuste manual dos resíduos de acordo com o mapa de densidade eletrônica. Na etapa de refinamento, o primeiro passo foi a leitura e processamento dos arquivos contendo as coordenadas da proteína e as reflexões. Em seguida foram empregados protocolos de refinamento (volume de solvente, refinamento de coordenadas e fator de temperatura, inserção de moléculas de água, etc.) seguidos da produção do arquivo contendo coordenadas do modelo refinado e estatísticas de refinamento (Afonine *et al.*, 2012).

O modelo foi refinado até a máxima concordância entre dados experimentais e estereoquímica da molécula pelos índices R e R_{free} (ver definição: página 54 - tabela 2). Durante esse processo também foram adicionadas moléculas de água e o peptídeo SV40 NLS. A qualidade estereoquímica dos modelos obtidos foi verificada a cada ciclo com os resultados apresentados no programa Phenix.

Para a segunda estrutura (Imp α Nc-2), o modelo obtido Imp α Nc-1 foi utilizado para a

substituição molecular utilizando o programa AutoMR. Posteriormente a estrutura foi submetida a ciclos de refinamento de corpo rígido, seguidos de etapas de modelagem e refinamento de posição de átomos e fator de temperatura. Assim como no modelo Imp α Nc -1, essas etapas utilizaram os programas Coot (Emsley & Cowtan, 2004) e phenix.refine, respectivamente. Além do peptídeo SV40 NLS, foram adicionadas moléculas de água e o modelo foi refinado até sua máxima concordância entre dados experimentais e estereoquímica das moléculas do modelo final.

5.1.5. Análise e comparação das estruturas cristalográficas.

As análises estruturais da proteína Imp α Nc e comparação com as estruturas das Imp α de outros organismos foram realizadas com os programas Coot (Emsley & Cowtan, 2004) e PyMol (Delano, 2002). Todas as figuras das estruturas foram geradas no PyMol (Delano, 2002). As interações entre os resíduos da proteína Imp α Nc e o peptídeo SV40 NLS foram verificadas através da ferramenta *online* PISA (Krissinel & Henrick, 2007; disponível em http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) e do programa LIGPLOT (Wallace et al., 1995).

5.2. Resultados

5.2.1. Cristais do complexo Imp α Nc/SV40NLS.

Os primeiros experimentos de cristalização realizados com a proteína purificada tinham como propósito a busca por condições de cristalização promissoras ao aparecimento de cristais adequados a serem submetidos a difração de raios X. Foram realizados testes tanto com a proteína Imp α Nc nativa, quanto complexada ao peptídeo SV40 NLS, considerando que experimentos anteriores mostraram a maior estabilidade da proteína na presença de peptídeos de NLSs.

Nos experimentos iniciais, com a proteína diluída a 12mg/ml, utilizando as técnicas de *hanging-drop* e *sitting-drop*, foram testadas as condições de cristalização dos *Kits: Screen I e II (Hampton Research), JCSG + suíte (Qiagen)* e *JBScreen Classic (Jena Bioscience)*. As placas foram armazenadas a 18 °C, assim como em experimentos anteriores com a proteína Imp α Mm (Barros *et. al.*, 2012). Após duas semanas, observou-se o aparecimento de pequenos

cristais em formato de agulha e outros cristais maiores porém, sem morfologia externa definida, ou seja, com arestas e vértices irregulares ou cristais sobrepostos (**Figuras 10A, 10B, 10C, 10D, 10E**). Como tais cristais eram inadequados para serem submetidos à difração de raios X, ora pelo tamanho muito pequeno e formato irregular, ora pela fragilidade ao serem coletados, prosseguiu-se com novos experimentos de triagem e otimização de condições de cristalização.

Nessa nova etapa, além das soluções dos *Kits* já utilizados, foram testadas as soluções do *Screen PEG/ION da Hampton Research*, sendo que, dessa vez, os experimentos foram armazenados a temperatura de 4°C, seguindo as indicações do perfil monomodal da proteína a essa temperatura, obtidos nos experimentos de espalhamento de luz dinâmico. As placas de *sitting-drop* foram montadas pelo robô de cristalização de proteínas Oryx4 da *Douglas Instruments*. Após uma semana, foram observados cristais com dimensões adequadas (em torno de 0,1 mm x 0,34 mm) e formações regulares, com arestas e vértices bem definidos, nas condições: (1) 0,2 mM fosfato de sódio dibásico dihidratado e 20% (w/v) de polietilenoglicol (PEG) 3350 e (2) 0,2 mM de bicina pH 8.5 e 20% PEG 6000 [condições do kit de cristalização PEG/ION (Hampton Research) e *The JSCG suite* (QIAGEN), respectivamente] (**Figuras 10F, 10G e 10H**), os quais foram armazenados na presença de crioprotetor para serem submetidos à difração de raios X. Os cristais que apresentaram melhores dados de difração foram os cristais 10H e 10F (**Figura 10**), obtidos nas condições (1) e (2), respectivamente.

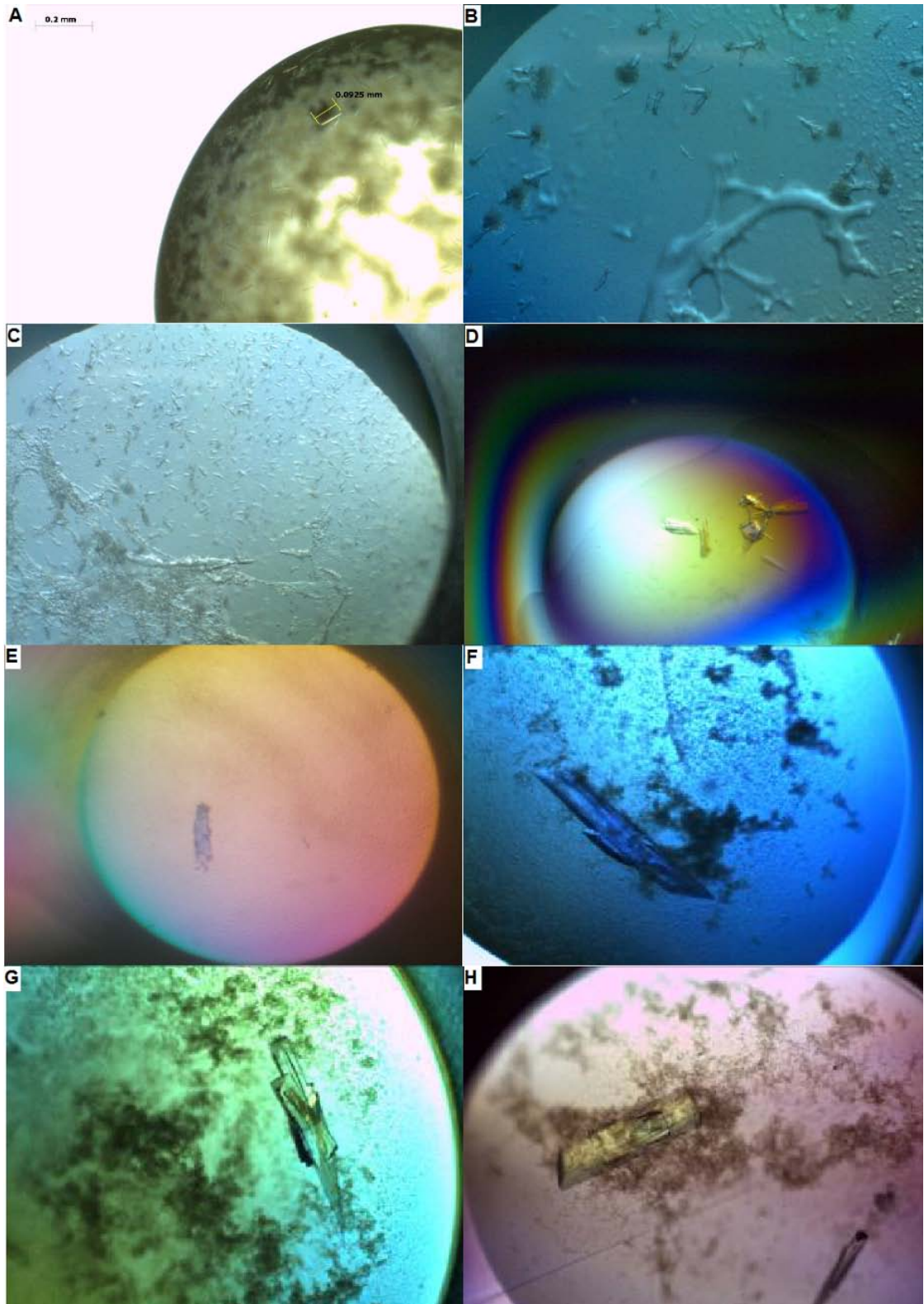


Figura 10: Cristais de Imp α Nc crescidos na presença de peptídeos SV40 NLS. A) critais obtidos a partir de condicao do *Screen The JCSG + Suite* (QIAGEN), utilizando a tecnica de *hanging-drop*. B e C) Cristais pequenos, em formato de agulhas, obtidos em condição do *Screen PACT+HTS* (Jena Bioscience). D e E) Cristais obtidos a partir de diferentes condições do *Screen JCSG+Suite*, utilizando a técnica de *microbatch*. F) Cristal obtidos a partir de condição do *Screen JCSG+Suite*, utilizando a técnica de *sitting-drop*. G e H) Cristais obtidos a partir de condições do Kit de *Screen PEG/ION* da *Hampton Research* com o peptídeo SV40 NLS.

5.2.2. Coleta e processamento dos dados de difração de raios X.

O primeiro conjunto de dados de difração de raios X, referente ao cristal do complexo Imp α Nc/SV40 NLS obtido na condição de cristalização (1) com a adição de 10% de PEG 400 (Figura 10H), foi coletado no LNLS, em Campinas (Imp α Nc-1). Os dados indicaram um padrão de difração em que há um cristal predominante, cercado por cristais satélite, além da presença de anéis de gelo (**Figura 11A**). Por esse motivo, houve a necessidade de coleta de um conjunto com 360 ° para garantir quantidade de dados completos para elucidação da estrutura. Esse primeiro conjunto de dados mostrou que o cristal do complexo Imp α Nc/SV40 NLS (Imp α Nc-1) pertence ao grupo espacial primitivo ortorrômbico (P2₁2₁2₁), com parâmetros de cela unitária a=44,7, b=63,6, c=180,4 Å. Isso indica que o cristal pertence ao mesmo grupo espacial de outros complexos Imp α de camundongo/peptídeos NLSs (Fontes *et al.*, 2003; Takeda *et al.*, 2011), porém não é isomorfo por apresentar diferentes dimensões de cela unitária (**Tabela 2**). O conjunto de dados apresentou 91,4% de dados completos a 2,0 Å de resolução, com um R_{merge} de 9,8 %. Cálculos com base na massa molecular indicaram a presença de uma molécula na unidade assimétrica. Isso corresponde a um coeficiente de Matthews (Matthews, 1968) V_M de 2,42 Å³Da⁻¹ com um conteúdo de solvente calculado de 49,1 %. Esses valores encontram-se no intervalo típico para cristais de proteína, assumindo um valor de 0,74 cm³g⁻¹ para o volume da proteína. Estatísticas de coleta de dados e processamento encontram-se na Tabela 2.

Um segundo cristal do complexo Imp α Nc/ SV40 NLS, obtido na condição de cristalização (2), foi submetido à difração de raios X no NSLS, em Upton- NY, EUA (Imp α Nc-2; **Figura 10F**). Dessa vez, foi adicionado 20 % de glicerol ao cristal, antes do resfriamento em nitrogênio líquido, para atuar como crioprotetor. O conjunto de dados coletado apresentou um padrão de difração significativamente melhor do que o primeiro, com ausência de anéis de gelo (**Figura 11B**). Os dados processados desse conjunto também mostraram significativa melhora com 99,8 % de dados completos a 1,75 Å de resolução, sendo que esta é a resolução mais alta obtida dentre as estruturas cristalográficas de Imp α já depositadas em banco de dados. Estatísticas de coleta de dados e processamento encontram-se na **Tabela 2**.

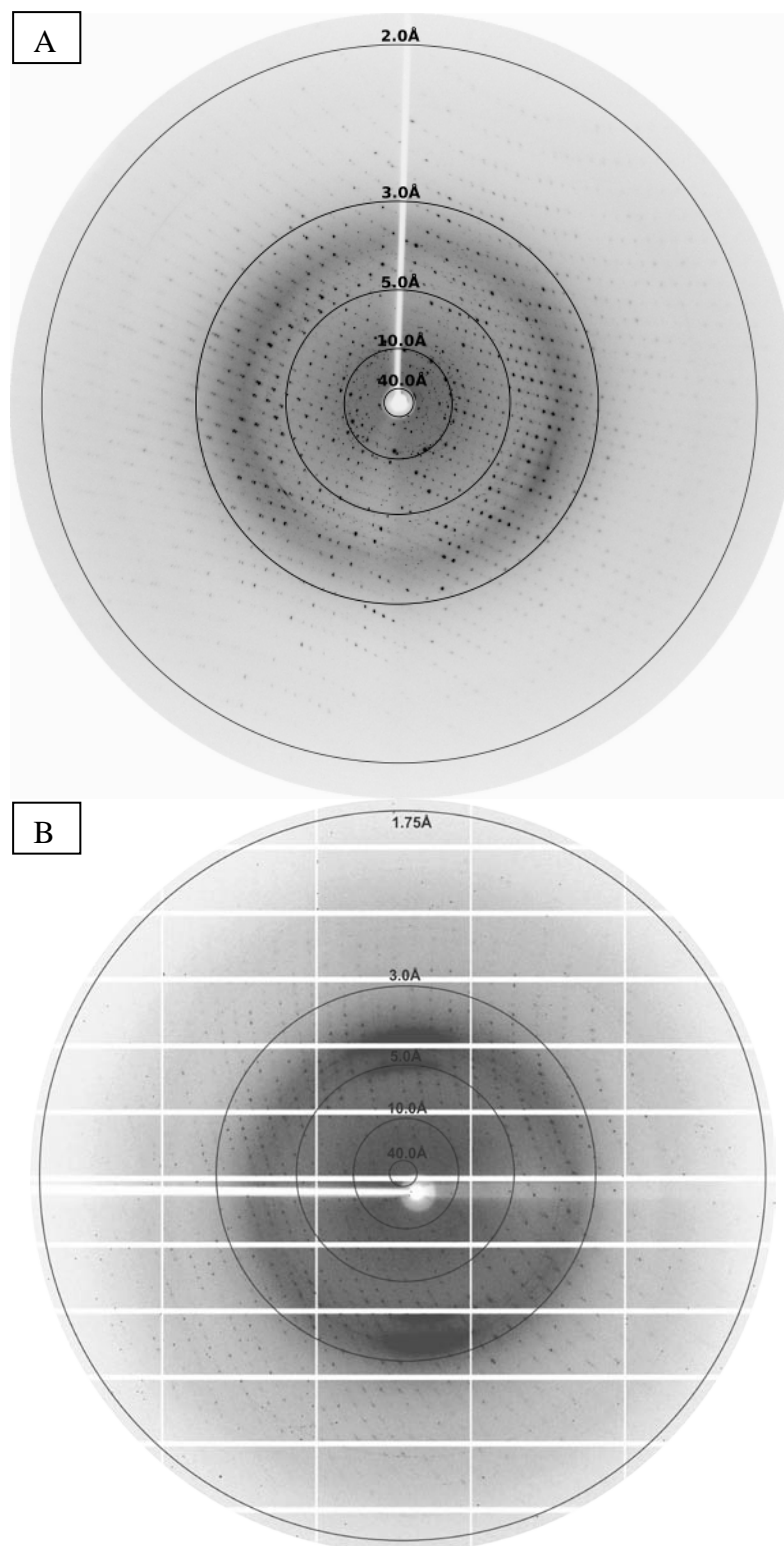


Figura 11. Imagem de difração de raios X dos cristais Imp α Nc-1 e Imp α Nc-2 do complexo Imp α Nc/ SV40 NLS. (A) Imagem coletada na Linha D03B-MX1 do LNL, em um comprimento de onda de 1,58 Å a 100 K, utilizando detector MAR CCD (MAR Research). A distância entre o cristal do detector foi de 70 mm e as imagens foram adquiridas a cada 1,5° de oscilação. **(B)** Imagem coletada na Linha X25 do NSLS, em um comprimento de onda de 1,0 Å a 100 K, utilizando detector Pilatus. A distância entre o cristal do detector foi de 190 mm e as imagens foram adquiridas a cada 0,5° de oscilação.

Tabela 2. Estatísticas da coleta de dados de difração de raios X e refinamento inicial dos modelos Imp α Nc-1 e Imp α Nc-2 do complexo Imp α Nc /SV40.

	Imp α Nc-1	Imp α Nc-2
Cela unitária (Å)	a=44,6; b=63,6; c=180,4	a= 45,1; b=64,4; c=185,6
Grupo espacial	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Resolução (Å)	40,0-2,05 (2,12-2,05) ^a	37,67 - 1,75 (1,81 – 1,75) ^a
Número de reflexões com F>0	932951	2450827
Reflexões únicas	30277	54988
Dados completos (%)	91,4 (89,8) ^a	99,80 (98,3) ^a
Mosaicidade	0,73	0,40
Redundância	10,5 (9,6) ^a	12,4 (9,3) ^a
R_{merge}^a (%)	9,7(66,4) ^a	6,8 (84,3) ^a
Média I/σ(I)	20,5(2,0) ^a	37,5 (2,0) ^a
R_{cryst}^c (%)	37,07	27,50
R_{free}^d (%)	42,45	46,39

^a

Números entre parênteses representam a faixa de mais alta resolução.

^b $R_{\text{merge}} = \frac{\sum_{hkl} (\sum_i (I_{hkl,i} - \langle I_{hkl} \rangle))}{\sum_{hkl} \langle I_{hkl} \rangle}$, onde $I_{hkl,i}$ é a intensidade de uma medida individual da reflexão com os índices de Miller h, k and l, e $\langle I_{hkl} \rangle$ é a intensidade média de cada reflexão. Calculado para $I > -3\sigma(I)$.

^c $R_{\text{cryst}} = \frac{\sum_{hkl} (|F_{\text{obs},hkl} - F_{\text{calc},hkl}|)}{\sum_{hkl} |F_{\text{obs},hkl}|}$, onde $|F_{\text{obs},hkl}|$ e $|F_{\text{calc},hkl}|$ são as amplitudes dos fatores de estrutura observados e calculados.

^d **R_{free} é equivalente ao R_{cryst}, mas calculado com base em 5% do total de reflexões obtidas.**

5.2.3. Elucidação da estrutura, modelagem e refinamento.

A primeira estrutura do complexo Imp α Nc/SV40NLS (Imp α Nc-1) foi elucidada pelo método de Substituição Molecular. Como modelo para a substituição foi utilizado o arquivo de coordenadas 3UL1_B da Imp α Mm, truncada em sua porção N-terminal e co-cristalizada com o peptídeo NLS da Nucleoplasmina (Marfori *et al.*, 2012). Nessa etapa foi possível confirmar a presença de um monômero na unidade assimétrica.

Posteriormente, o modelo do complexo foi submetido a um refinamento de posição dos átomos e do fator de temperatura das estruturas. O programa “Coot” foi utilizado para inspeção e modelagem do complexo (Emsley and Cowtan, 2004). Em virtude da identidade de 49 % entre as sequências, as estatísticas de refinamento iniciaram-se altas ($R_{\text{cryst}} = 37,07$ e

$R_{\text{free}} = 42,45$), sendo necessário efetuar a mutação de 203 resíduos como uma das primeiras etapas de etapa do processo de modelagem.

Uma primeira avaliação do mapa de densidade eletrônica indicou a presença de densidade eletrônica nos sítios de reconhecimento de NLS. A modelagem do peptídeo SV40 NLS em ambos os sítios, com posterior ciclo de refinamento confirmou a presença do peptídeo nos sítios principal e secundário da proteína. Durante os ciclos de modelagem, cadeias laterais de resíduos que não puderam ser modeladas, devido a ausência de densidade eletrônica, foram removidas. Em etapas finais, houve a adição de moléculas de água a estrutura. Estatísticas finais de refinamento encontram-se na **Tabela 3**.

O segundo modelo do complexo Imp α Nc/SV40NLS (Imp α Nc-2), também foi elucidado pelo método de Substituição Molecular. Dessa vez, utilizou-se como modelo para a substituição o primeiro modelo já elucidado (Imp α Nc-1).

Assim como no modelo anterior, a Imp α Nc-2 foi submetida a ciclos de modelagem e refinamento, onde resíduos e/ou cadeias laterais que não apresentaram densidade eletrônica foram removidas. Águas estruturais também foram adicionadas ao modelo. Estatísticas finais de refinamento encontram-se na **Tabela 3**.

Tabela 3. Estatísticas finais de refinamento dos modelos Imp α Nc-1 e Imp α Nc-2 do complexo ImpANc /SV40.

	ImpαNc-1	ImpαNc-2
Resolução (Å)	40,0-2,05 (2,12-2,05) ^a	36.98 - 1.75 (1.81 - 1.75)
Número de reflexões	932951	2450827
Reflexões únicas	32551 (3130) ^a	54993 (5195) ^a
Dados Completos (%)	91,4 (89,8) ^a	99.58 (95.81) ^a
R_{cryst}^b (%)	19,35	18,35
R_{free}^c (%)	23,41	20,97
Número de átomos	3510	3719
proteína	3235	3229
peptídeo	114	114
água	159	376
Resíduos nas regiões mais favoráveis	97	97
(não permitidas) (%)	1,2	1,4
Fator de temperatura médio (Å²)^e	30,2	37,7

^a Números em parênteses correspondem aos dados de mais alta resolução.

^b $R_{cryst} = \frac{\sum hkl (||F_{obs}hkl| - |F_{calc}hkl||)}{\sum hkl |F_{obs}hkl|}$, onde $|F_{obs}hkl|$ e $|F_{calc}hkl|$ são as amplitudes dos fatores de estrutura observados e calculados. ^c Rfree é equivalente ao Rcryst ,mas calculado com base em 5% do total de reflexões obtidas. ^d Calculado com o programa PROCHECK (Laskowski, et al., 1993).

5.2.4. Estruturas da proteína Imp α Nc complexada ao peptídeo SV40 NLS.

As estruturas da Imp α Nc foram modeladas de acordo com a presença de densidade eletrônica para os resíduos. Sendo assim, a estrutura Imp α Nc-1 resultou na modelagem dos resíduos no intervalo de 81 a 507 e a estrutura Imp α Nc-2, resultou na modelagem dos resíduos 79 a 507. A região entre os aminoácidos 461 e 465 não apresentou densidade eletrônica em nenhuma das estruturas, portanto, os resíduos nessa região não foram modelados. Em ambas as estruturas obtidas, a proteína Imp α Nc apresenta um domínio único e alongado, constituído por dez repetições em ARMs, cada um contendo três hélices- α (H3) conectadas por *loops* (**Figura 12**), assim como observado em estruturas da Imp α de outros organismos (Conti *et. al.*, 1998; Kobe, 1999; Terendau *et. al.*, 2007; Chang *et. al.*, 2012). Tanto em Imp α Nc-1, quanto Imp α Nc-2, o peptídeo SV40 NLS (¹²⁵PPKKR¹³²KV) se ligou ao sítio principal e secundário da proteína (**Figuras 13 e 14**), em uma conformação estendida e antiparalela a direção das hélices dos motivos ARMs (**Figura 12**). Durante a modelagem do peptídeo no sítio principal das estruturas, foi possível observar densidade eletrônica para todos os aminoácidos do

peptídeo (**Figura 13**) enquanto no sítio secundário, não foi possível visualizar densidade eletrônica para três resíduos da porção C-terminal do peptídeo (P125, P126, K127; **Figura 14**). Uma comparação entre as estruturas Imp α Nc-1 e Imp α Nc-2, permitiu verificar que a maior resolução da estrutura Imp α Nc-2, resultou no aparecimento de uma densidade mais definida para os resíduos dos peptídeos SV40NLS ligados aos sítios principal e secundário da proteína (**Figuras 13 e 14**). Devido a maior qualidade dos dados de Imp α Nc-2, esse modelo foi escolhido para as análises posteriores da estrutura do complexo Imp α Nc/SV40 NLS.

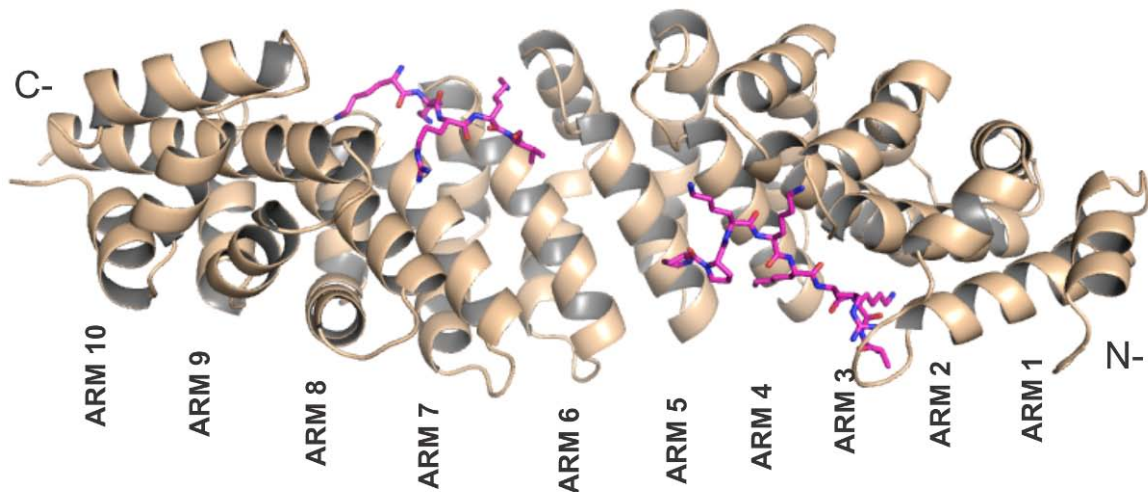


Figura 12. Estrutura da proteína Imp α Nc complexada à peptídeos SV40 NLS. Representação em diagrama de fitas (bege) da estrutura alongada e côncava da Imp α Nc (obtida a partir dos dados Imp α Nc-2), truncada em sua porção N-terminal, formada por dez motivos ARMs. Os peptídeos SV40 NLS, representados por bastões na cor rosa, estão ligados tanto ao sítio principal (próximo a região N-terminal da proteína) quanto secundário (próximo a região C-terminal da proteína).

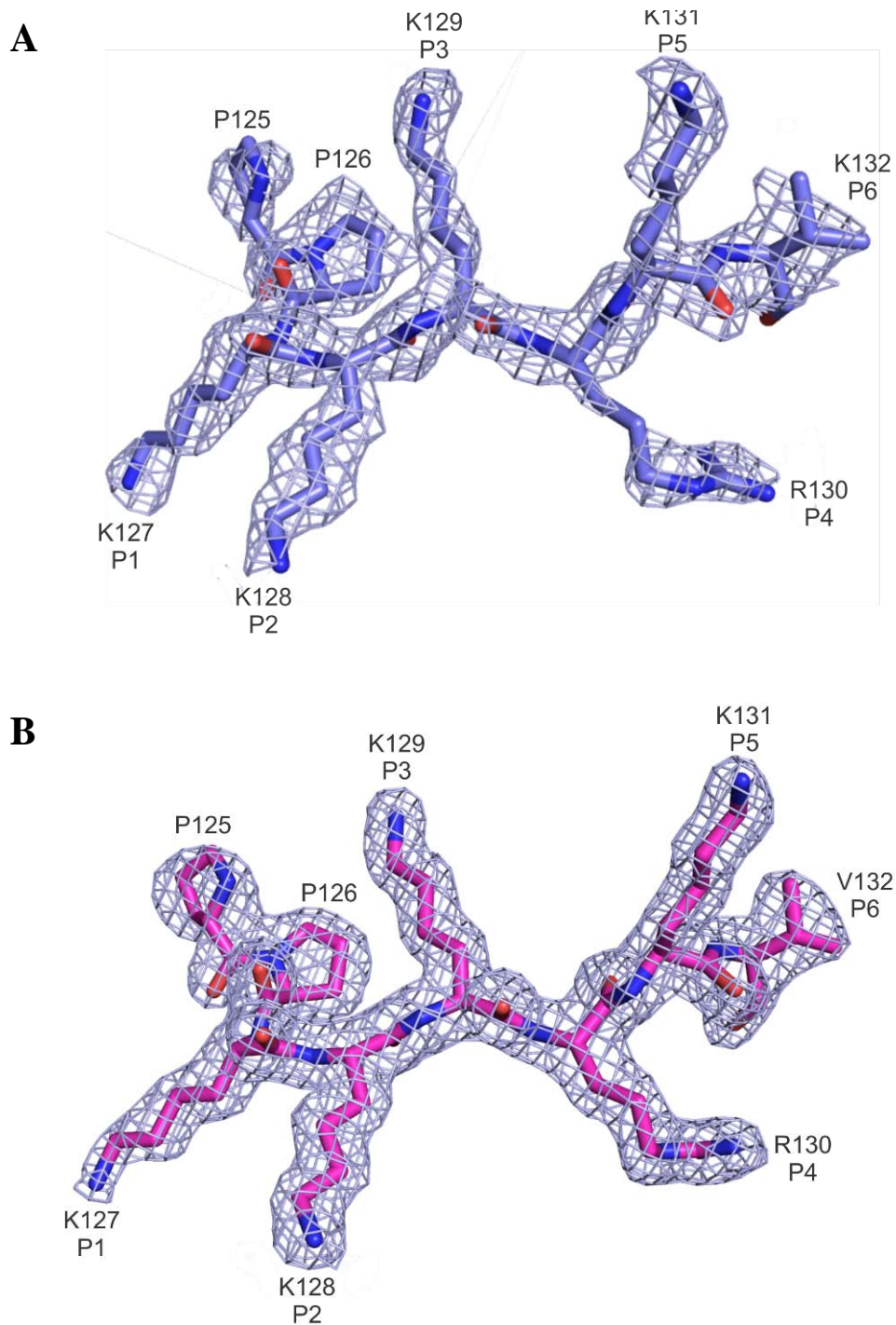


Figura 13. Mapa de densidade eletrônica do peptídeo SV40 NLS ligado aos sítios principal das estruturas Imp α Nc-1 e Imp α Nc-2. Mapa $2|F_{obs}| - |F_{calc}|$, gerado com desvio padrão de $\sigma=1,0$, correspondendo ao peptídeo SV40 NLS no sítio principal da proteína Imp α Nc, nas estruturas Imp α Nc-1 (A; azul) e Imp α Nc-2 (B; rosa). Essa figura foi gerada através do programa PyMOL (Delano, 2002)

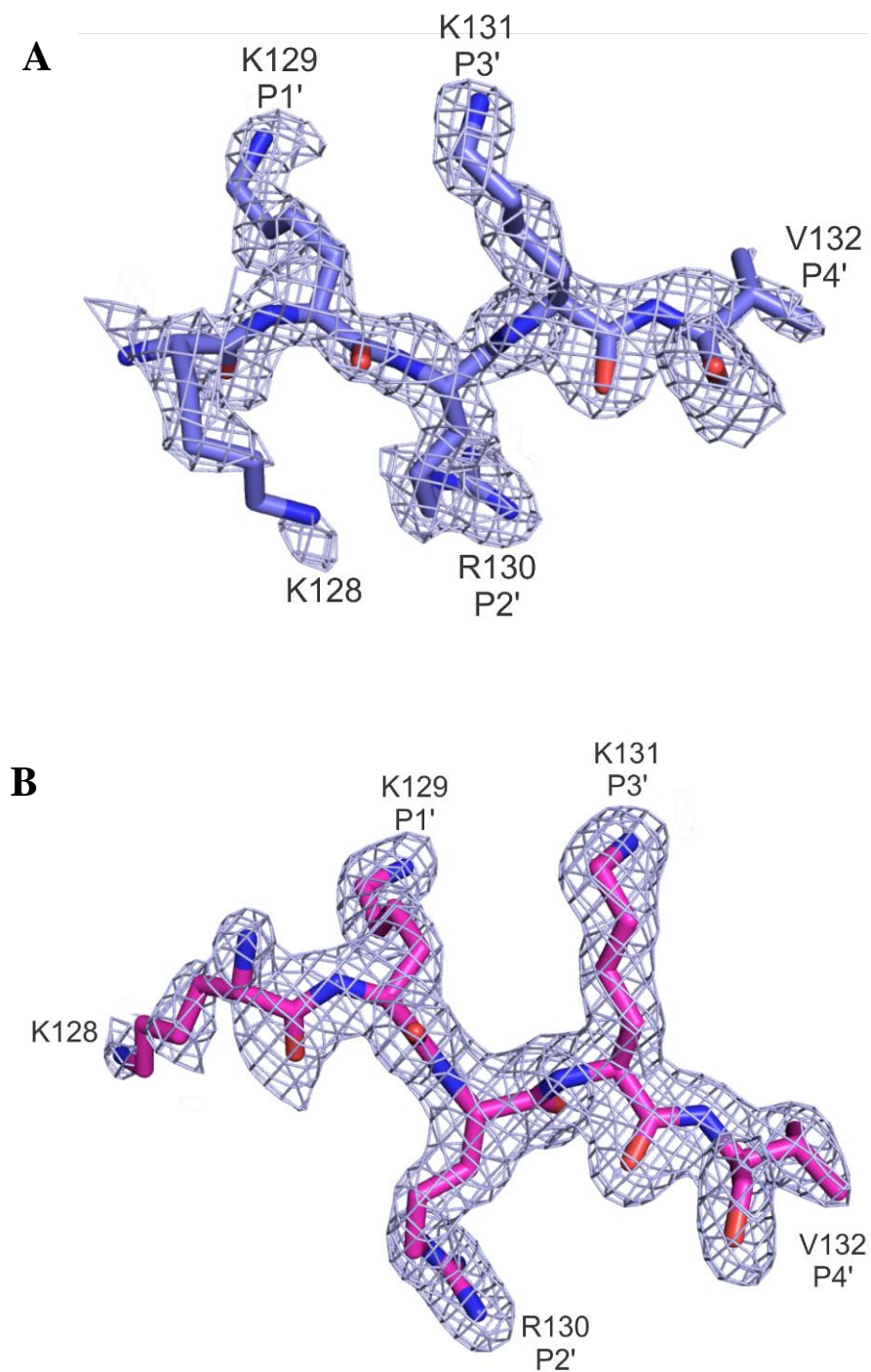


Figura 14. Mapa de densidade eletrônica do peptídeo SV40 NLS ligado ao sítio secundário das estruturas Imp α Nc-1 e Imp α Nc-2. Mapa $2|F_{obs}| - |F_{calc}|$, gerado com desvio padrão de $\sigma=1,0$, correspondendo ao peptídeo SV40 NLS no sítio principal da proteína Imp α Nc, nas estruturas Imp α Nc-1 (A; azul) e Imp α Nc-2 (B; rosa). Essa figura foi gerada através do programa PyMOL (Delano, 2002)

5.3. Interação entre o peptídeo SV40 NLS e resíduos presentes no sítio principal da Imp α Nc.

Os resíduos de asparaginas N150, N192, N235, no sítio principal, definem a direção e conformação da ligação do peptídeo ao sítio, através da interação, via ligações de hidrogênio, entre os grupos carbonil dos resíduos na proteína e a cadeia principal do peptídeo. As cadeias laterais de asparaginas, por serem menos flexíveis, minimizam a perda de entropia do complexo Imp α Nc/ SV40 NLS, após a ligação do peptídeo.

No sítio principal, a cadeia- lateral alifática da K131 do peptídeo se acomoda entre pequenos sulcos apolares formados pelos resíduos hidrofóbicos dos triptofanos W146 e W188, enquanto a cadeia-lateral da K129 se encaixa entre os resíduos W231 e W188 da proteína, portanto, as lisinas K129 e K131 ocupam as posições P3 e P5, respectivamente (**Figura 16**). Enquanto isso, as cadeias-laterais dos aminoácidos básicos K127, K128 e R130 do peptídeo se encaixam em alguns sulcos com carga negativa formados entre o final das hélices H3, ocupando as posições P1, P2 e P4, respectivamente (**Figura 16**).

A avaliação dos dados estruturais do complexo Imp/ SV40 NLS através dos programas LigProt e PISA (Krissinel & Henrick, 2007), mostrou interações de outros resíduos da proteína com as cadeias laterais dos resíduos do peptídeo SV40 NLS, que garantem a estabilidade da interação do peptídeo com a proteína (**Figuras 15 e 16**). No sítio principal, o resíduo K127, na posição P1 participa de interações hidrofóbicas com G195. Já na posição P2, a K128 forma uma ponte salina com o aminoácido D196 e ligações de hidrogênio com G154, A152 e T159. A cadeia- lateral da lisina K129, na posição P3, não apresenta interações bem definidas, além das interações com os triptofanos W231 e W188, no entanto, segundo Conti *et. al.*, o resíduo pode interagir com os aminoácidos D270 e E266 através de moléculas de água (Conti *et. al.*, 1998). A cadeia-lateral da arginina R130, em P4, forma ligações de hidrogênio com os resíduos de L109 e K111 e interações hidrofóbicas com P115 e S153, enquanto a cadeia-lateral da lisina K131, na posição P5, apresenta ligações de hidrogênio com a glutamina Q185 da proteína e apresenta interações hidrofóbicas com a F142. Por fim, V132, em P6, é estabilizado através de ligações de hidrogênio com S110.

Um resíduo prolina na porção N terminal do NLS (P125) também interage com os resíduos R238 e D270 da proteína, via ligações de hidrogênio e interações hidrofóbicas. Próximo ao sítio principal também foi possível observar a presença conservada dos aminoácidos tirosina (Y277) e arginina (R315), que formam a região de ligação entre os dois sítios da proteína durante a ligação de NLSs bipartidos (Fontes *et. al.*, 2000).

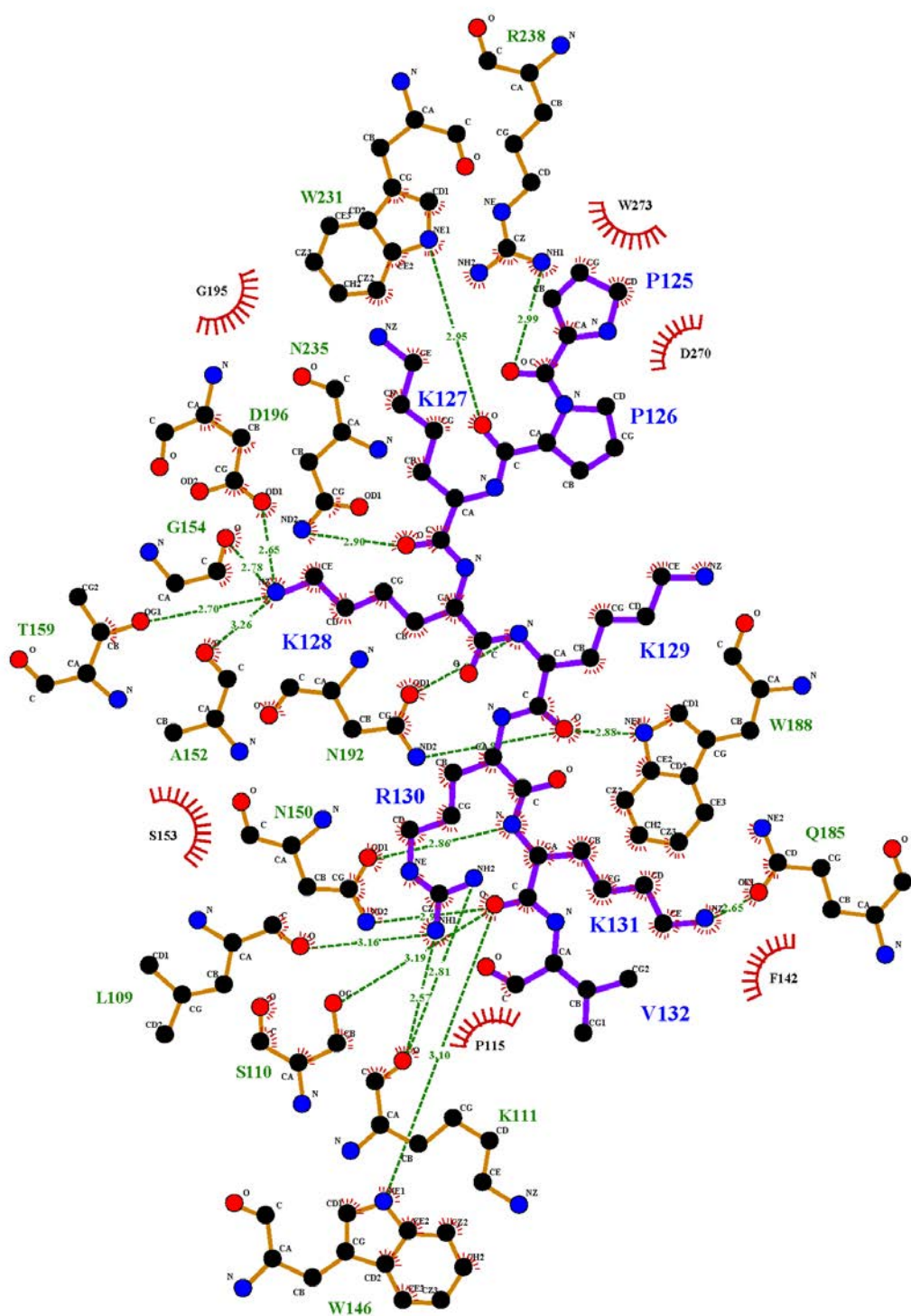


Figura 15. Diagrama de interações dos peptídeos SV40 NLS com resíduos do sítio principal da Imp α Nc. O peptídeo SV40 NLS está representado em bolas-e-bastões, na cor violeta enquanto resíduos da proteína estão em amarelo. Átomos de carbono, nitrogênio e oxigênio são mostradas em preto, azul e vermelho, respectivamente. As interações do tipo ligações de hidrogênio são representadas por linhas tracejadas na cor verde, onde também podem ser observadas as distâncias de ligação. Os resíduos da proteína que promovem interações hidrofóbicas estão destacadas na forma de “arcos” vermelhos. Essa figura foi preparada com o programa LIGPLOT (Wallace *et al.*, 1995).

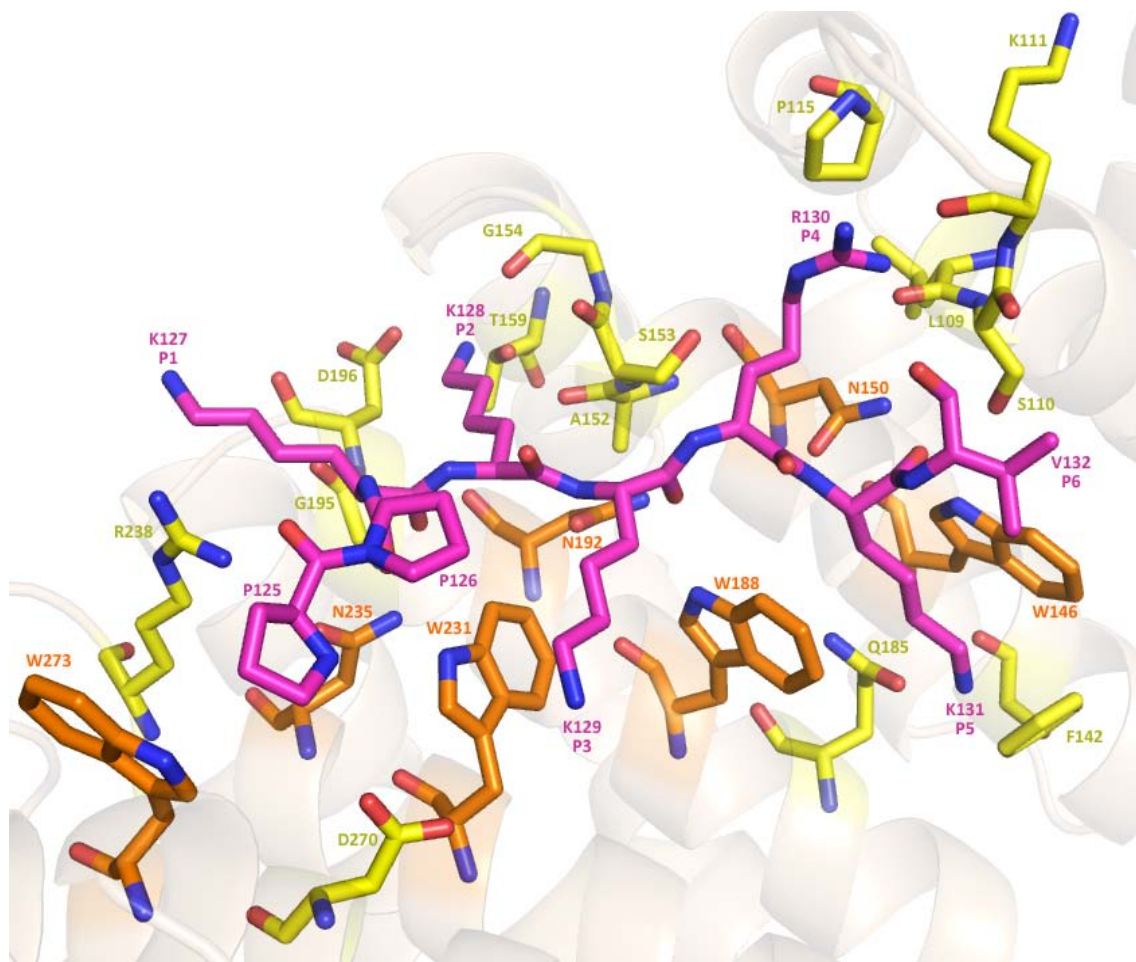


Figura 16. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio principal. Os resíduos de triptofano e asparagina que orientam e estabilizam a interação do peptídeo (rosa) ao sítio estão representados em laranja. Resíduos que interagem com as cadeias laterais do peptídeo são representados em amarelo. Átomos de oxigênio e nitrogênio são representados em vermelho e azul, respectivamente. Essa figura foi gerada através do programa PyMOL (Delano, 2002).

A distribuição dos fatores de temperatura (β -factor) ao longo da cadeia do peptídeo SV40 NLS reflete diferenças na estabilidade de cada resíduo durante a interação com a Imp α (Tabela 4). No sítio principal, a média do fator de temperatura do peptídeo (33,21 Å) é menor do que a média do fator de temperatura da proteína (37,71 Å). Os resíduos que estão posicionados nas extremidades apresentam um fator de temperatura mais alto, enquanto os resíduos K128 e K129, nas posições P2 e P3, apresentam valor de fator de temperatura mais baixo. Esses dados são condizentes com as interações que esses resíduos realizam com a proteína.

Tabela 4: Distribuição do fator de temperatura (β -factor), para cada resíduo do peptídeo SV40 NLS que interage com a proteína Imp α Nc no sítio principal.

Posição no sítio principal	Resíduos do peptídeo SV40 NLS	β -factor (Å)
	P125	39,98
	P126	33,79
P1	K127	34,11
P2	K128	28,25
P3	K129	28,15
P4	R130	32,04
P5	K131	29,38
P6	V132	40,02
Total		33,21

5.3.1. Interação entre o peptídeo SV40 NLS e resíduos presentes no sítio secundário da Imp α Nc.

No sítio secundário da Imp α Nc, de maneira similar ao principal, são as asparaginas N319 e N361 que orientam a cadeia principal do peptídeo na ligação com a proteína, via ligações de hidrogênio. Sendo que, nesse sítio, os resíduos K129 e K131 do peptídeo se encaixam entre os sulcos com carga negativa formados pelas hélices- α da proteína, enquanto os anéis aromáticos dos triptofanos W357 e W399 acomodam a cadeia lateral da R130 (**Figura 18**).

A ligação do peptídeo SV40 NLS ao sítio secundário da Imp α Nc apresentou algumas interações que podem ser significantes para o reconhecimento de peptídeos nesse sítio (**Figuras 17 e 18**), como: ligações de hidrogênio da K129, em P1', com os resíduos G323, V321, T328 e do resíduo K131, em P3', com N283, G281, T322, além de interações hidrofóbicas com I286. O resíduo R130, em P2', também participa de interações importantes, formando uma ponte salina com E396 e ligação de hidrogênio com S360. Por fim, a V132 em P4' realiza interações hidrofóbicas com D280 e R315 e N319. O resíduo de K128, apesar da alta flexibilidade que apresentou na interação, realizou interações de hidrogênio com N403 e interações hidrofóbicas com A364 e S406.

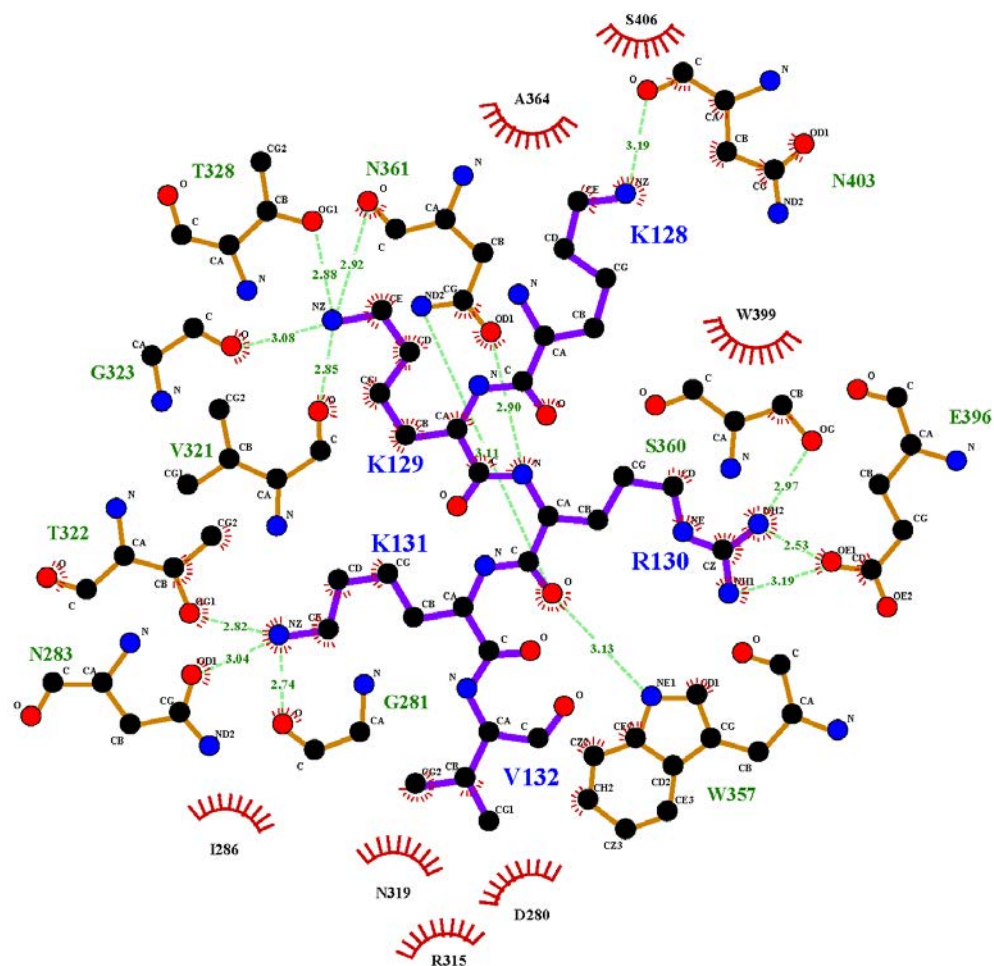


Figura 17. Diagrama de interações dos peptídeos SV40 NLS com resíduos do sítio secundário da Imp α Nc. O peptídeo SV40 NLS está representado em bolas-e-bastões, na cor violeta enquanto resíduos da proteína estão em amarelo. Átomos de carbono, nitrogênio e oxigênio são mostrados em preto, azul e vermelho, respectivamente. As interações do tipo ligações de hidrogênio são representadas por linhas tracejadas na cor verde, onde também podem ser observadas as distâncias de ligação. Os resíduos da proteína que promovem interações hidrofóbicas estão destacados na forma de “arcos” vermelhos. Essa figura foi preparada com o programa LIGPLOT (Wallace *et al.*, 1995).

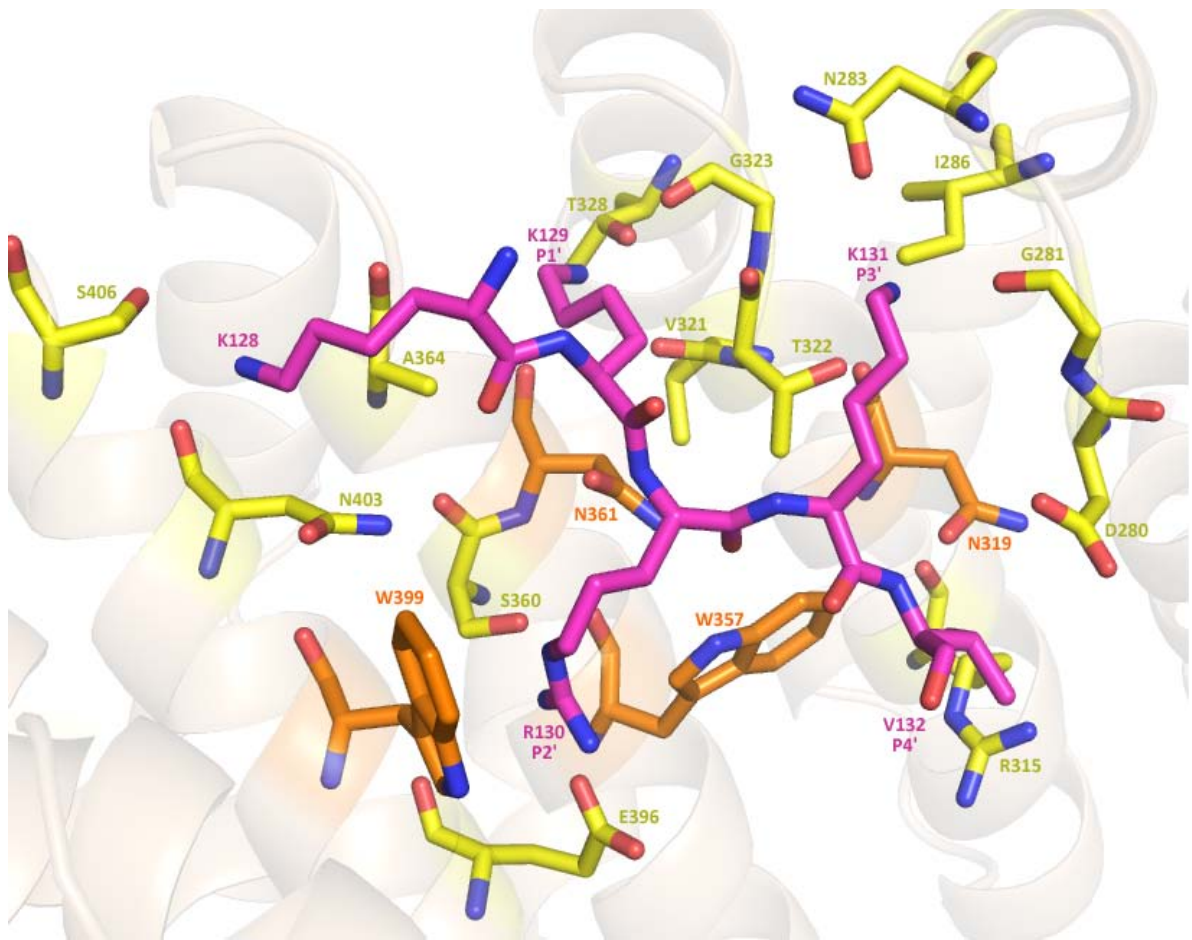


Figura 18. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio secundário. Os resíduos de triptofano e asparagina que orientam e estabilizam a interação do peptídeo (rosa) ao sítio estão representados em laranja. Resíduos que interagem com as cadeias laterais do peptídeo são representados em amarelo. Átomos de oxigênio e nitrogênio são representados em vermelho e azul, respectivamente. Essa figura foi gerada através do programa PyMOL (Delano, 2002).

A distribuição dos fatores de temperatura (β -factor) ao longo da cadeia do peptídeo SV40 NLS que interage com o sítio secundário da proteína, mostra características específicas dessa interação (**Tabela 5**). Nesse sítio, a média do fator de temperatura do peptídeo (39,28 Å) é maior do que a média do fator de temperatura da proteína (37,71 Å). O valor elevado do fator de temperatura nesse sítio é devido, principalmente, ao elevado fator de temperatura da K128, que apresentou uma alta flexibilidade. Dentre os resíduos do peptídeo que interagem com a proteína no sítio secundário, o que apresenta menor valor de fator de temperatura é o peptídeo R130, na posição P2', seguido da K129, na posição P1'.

Tabela 5. Distribuição do fator de temperatura (β -factor), para cada resíduo do peptídeo SV40 NLS que interage com a proteína Imp α Nc no sítio secundário.

Posição no sítio secundário	Resíduos do SV40 NLS	β -factor (Å)
	K128	50,87
P1'	K129	37,76
P2'	R130	36,22
P3'	K131	39,17
P4'	V132	38,49
Total		39,28

5.4. Discussão

5.4.1. Interações significativas entre resíduos do peptídeo SV40 NLS e resíduos presentes nos sítios principal e secundário da Imp α Nc.

Os resíduos de aminoácidos presentes na porção côncava da Imp α Nc e conservados na Imp α de outros organismos, formam uma superfície ideal para a ligação de peptídeos NLS. Na estrutura da Imp α Nc, os peptídeos SV40 NLS são observados tanto no sítio principal quanto no secundário, no entanto, é possível observar diferenças na ocupação dos dois sítios.

As posições dos resíduos do peptídeo interagindo com a proteína no sítio principal segue o mesmo padrão observado em outras estruturas do complexo Imp α /SV40NLS (**Figura 19**). As ligações de hidrogênio que ocorrem na posição P2, através da K128, garantem uma maior estabilidade do resíduo nessa região, conforme observado pelo menor fator de temperatura da K128 em comparação com resíduos em outras posições, e destacam a importância do resíduo para o reconhecimento de NLSs no sítio principal, como já foi demonstrado em estudos onde peptídeos SV40 NLS mutados nessa região apresentaram alterações no processo de transporte (Colledge *et. al.*,1986). As interações hidrofóbicas da K129, em P3, com resíduos de triptofano da proteína, também conferem uma maior

estabilidade na interação deste resíduo, que se torna significativo para a estabilização do peptídeo no sítio principal. As interações do resíduo de prolina na porção N terminal do NLS (P125) mostra a contribuição de outros resíduos do peptídeo na estabilidade da interação com a proteína.

Figura 19. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio principal, nas posições P1 a P6. Esquema mostrando as principais interações do peptídeo no sítio principal da proteína. O peptídeo é representado em rosa, resíduos de triptofano e asparagina em laranja, e outros resíduos que interagem a cadeia lateral de resíduos específicos do peptídeo são mostrados em amarelo. As linhas pontilhadas representam ligações de hidrogênio, e as interações hidrofóbicas estão representadas por arcos tracejados.

A ligação do peptídeo SV40 NLS no sítio secundário é similar a ligação do mesmo peptídeo no sítio secundário da Imp de arroz (Chang et. al., 2012 ; **Tabela 6; Figura 20**), com os aminoácidos K129 e R130, nas posições P1' e P2', respectivamente, enquanto nas Imp α de camundongo e levedura são os aminoácidos K128 e K129 que se ligam na posição P1' e P2' (**Tabela 6**), respectivamente (Conti et. al., 1998; Fontes et. al., 2000).

Tabela 6. Ligação do peptídeo SV40 NLS a Imp α de diferentes organismos. Em vermelho, estão destacadas as regiões onde o peptídeo SV40 NLS se liga de forma diferencial, entre as proteínas de diferentes organismos. Os aminoácidos que apresentaram densidade eletrônica na estrutura obtida dos complexos estão sublinhados.

NLS: Imp α	Sítio secundário								Sítio principal											
	P1'	P2'	P3'	P4'					P1	P2	P3	P4	P5	P6						
SV40:Imp α Nc	P	P	K	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>		<u>P</u>	<u>P</u>	<u>K</u>	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>			
SV40:Imp α Os	P	P	K	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>	G	<u>G</u>	<u>S</u>	<u>P</u>	<u>P</u>	<u>K</u>	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>	<u>G</u>
SV40:Imp α Mm			<u>P</u>	<u>K</u>	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>			<u>P</u>	<u>K</u>	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>		
SV40:Imp α Sc	S	P	<u>K</u>	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>			S	P	<u>K</u>	<u>K</u>	<u>K</u>	<u>R</u>	<u>K</u>	<u>V</u>			

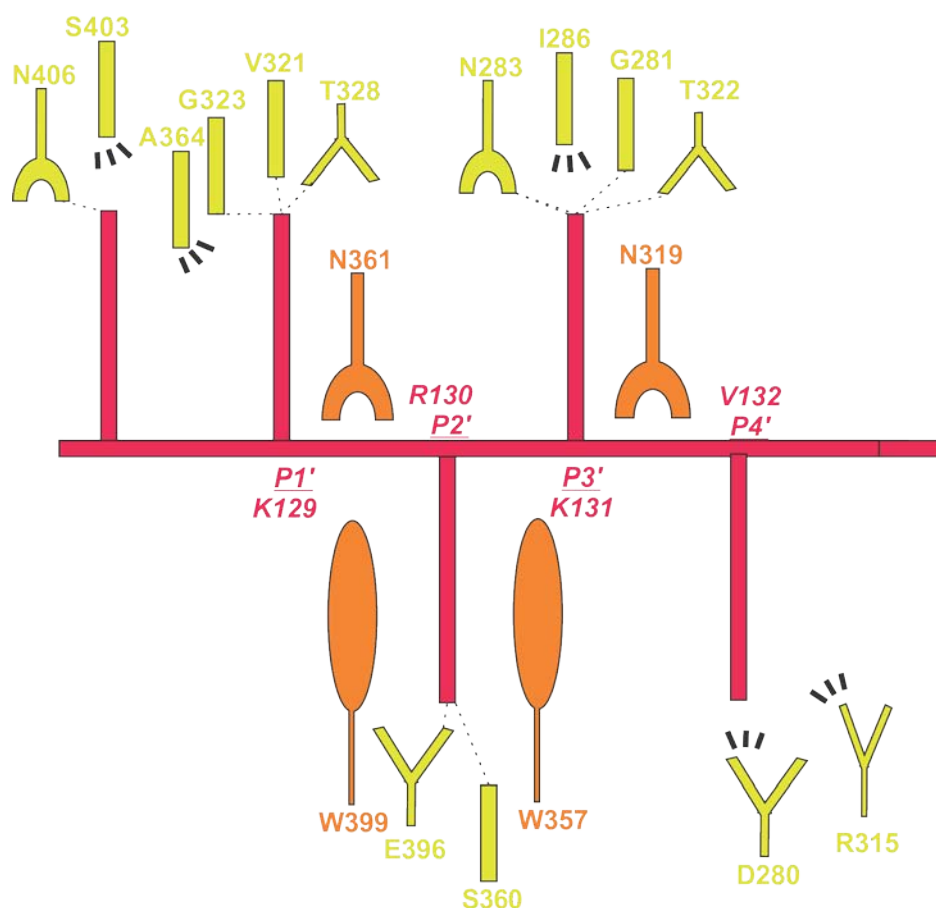


Figura 20. Representação dos resíduos da Imp α Nc que interagem com o peptídeo SV40 NLS no sítio secundário nas posições P1' a P4'. Esquema mostrando as principais interações do peptídeo no sítio principal da proteína. O peptídeo é representado em rosa, resíduos de triptofano e asparagina em laranja, e outros resíduos que interagem a cadeia lateral de resíduos específicos do peptídeo são mostrados em amarelo. As linhas pontilhadas representam ligações de hidrogênio, e as interações hidrofóbicas estão representadas por arcos tracejados.

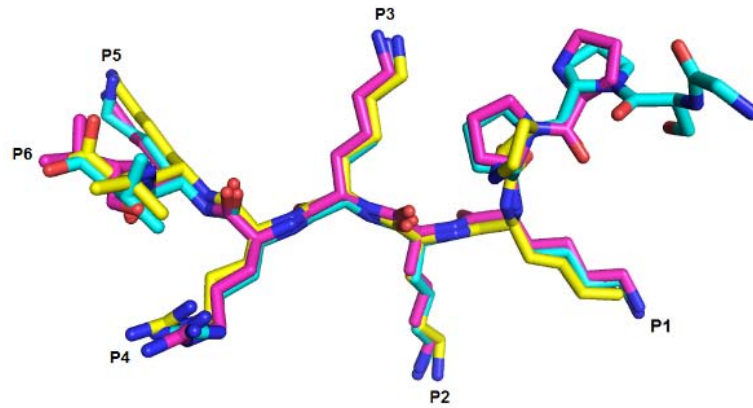
As interações eletrostáticas no sítio secundário não são tão bem definidas quanto no sítio principal (Conti *et. al.*, 1998). No entanto, a ligação do peptídeo SV40 NLS ao sítio secundário da Imp α Nc apresentou algumas interações que podem ser significantes para o reconhecimento de peptídeos nesse sítio. Assim como ocorre com a K128, na posição P3 do sítio principal, as interações hidrofóbicas com os triptofanos da cadeia da proteína podem garantir maior estabilidade ao resíduo R130 na posição P2', o que resultou em um menor fator de temperatura para esse resíduo. O resíduo na posição P1' também apresentou fator de temperatura relativamente baixo, indicando que as ligações hidrogênicas da K129, em P1' são significativas para o reconhecimento e estabilidade no reconhecimento de NLSs através do sítio secundário da proteína.

5.4.2. Interação preferencial do peptídeo SV40 NLS ao sítio principal da proteína Imp α Nc.

A sobreposição dos peptídeos de SV40 NLS observados nas estruturas de Imp α Mm e Imp α Os, mostra que a conformação do peptídeo é conservada no sítio principal, na interação com diferentes organismos, diferente do que ocorre no sítio secundário (**Figuras 21A e 21B**).

Na Imp α Nc o fator de temperatura do peptídeo no sítio principal é 33,21 Å e no sítio secundário é de 39,28 Å, enquanto o fator de temperatura da proteína é de 37,71 Å. O menor fator de temperatura no sítio principal indica a maior estabilidade do peptídeo SV40 NLS interagindo com esse sítio da proteína. Esse fator, aliado a falta de densidade para alguns resíduos do peptídeo no sítio secundário, a variação da conformação do peptídeo na interação ao sítio secundário de proteínas Imp α de diferentes organismos e o maior número de interações bem definidas entre peptídeo e proteína no sítio principal, sugerem a ligação preferencial do peptídeo ao sítio principal, assim como já sugerido nas análises estruturais das proteínas Imp α Mm, Imp α Sc e Imp α Os complexadas ao peptídeo Imp α Nc/ SV40 NLS (Fontes *et. al.*, 2000; Conti *et. al.*, 1998; Chang *et. al.*, 2012) . Apesar da maior flexibilidade na interação do peptídeo SV40 NLS com o sítio secundário, sabe-se que esse se torna significativo no reconhecimento de NLS's bipartidos.

A



B

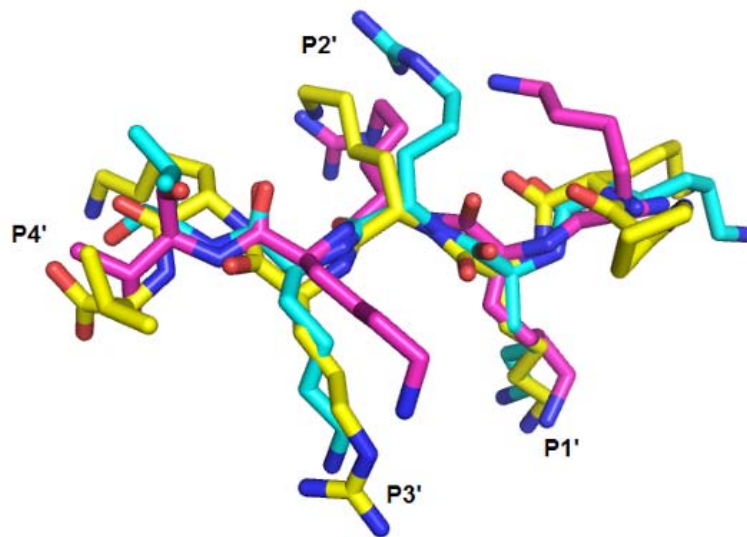


Figura 21. Sobreposição do peptídeo SV40 NLS interagindo com os sítios principal e secundário das proteínas ImpαNc, ImpαMm e ImpαOs. A) Sobreposição dos peptídeos SV40 NLS ligados ao sítio principal das proteínas ImpαNc (rosa), ImpαMm (amarelo) e ImpαOs (azul). B) Sobreposição dos peptídeos SV40 NLS ligados ao sítio secundário das proteínas ImpαNc (rosa), ImpαMm (amarelo) e ImpαOs (azul).

***6. Comparação da estrutura da
ImpαNc com Impα de outros
organismos.***

6.1. Material e Métodos

6.1.1. Dicroísmo circular.

Com o intuito de verificar o enovelamento da proteína Imp α Nc e comparar com o perfil de enovelamento da proteína Imp α Mm, foram realizados experimentos de Dicroísmo circular no aparelho Jasco J-815 (JASCO Inc., Tokyo, Japan) com o processamento dos dados pelo programa Spectra Manager II. Os experimentos foram realizados com as proteínas concentradas a 0,37 mg/mL, em cubeta de quartzo com comprimento óptico de 0,5 mm. As medidas foram realizadas em um espectro de 190 – 260 nm a 20 °C, velocidade de varredura de 100 nm/min; comprimento de bandas de 2 nm e tempo de resposta de 1 s. O espectro final obtido corresponde a média de 20 medidas acumuladas e corrigidas em relação a linha de base referente ao tampão (10 mM de Tris HCl pH 8 e 20 mM de NaCl) e normalizadas em razão da elipicidade molar residual $[\theta]$. A estimativa de conteúdo de estrutura secundária foi realizada com a utilização do programa CONTINN do pacote CDPPro (Sreerama & Woody, 2000; Sreerama & Woody, 2004a).

No mesmo equipamento foi realizada análise de desnaturação térmica para avaliar a estabilidade térmica da proteína Imp α Nc, pelo monitoramento das mudanças na elipicidade em comprimento de onda fixo de 222 nm, enquanto a amostra foi submetida a uma variação térmica de 10-90 °C. Os dados foram coletados em uma rampa de temperatura de 1 °C/min, com intervalo de equilíbrio de 5s entre cada temperatura, utilizando uma cubeta de quartzo com comprimento óptico de 2,0 mm.

A curva de desnaturação foi normalizada através da equação $[\theta]=[(\epsilon_F- \epsilon_U)]$, onde θ corresponde a elipicidade em qualquer temperatura, ϵ_F é o valor da elipicidade quando a proteína está completamente dobrada, e ϵ_U corresponde a elipicidade quando a proteína está totalmente desnaturada (Greenfield, 2004). A temperatura de desnaturação (*T melting* - *Tm*) foi calculada com o programa *Denatured Protein Analysis* do pacote de programas *Spectra ManagerTM II*.

6.1.2. Alinhamento das sequências da Imp α de diferentes organismos.

A sequência de aminoácidos das Imp de diferentes organismos que já tiveram suas estruturas cristalográficas elucidadas (Imp α Mm, Imp α Sc, Imp α Os, Imp α Hs) foram alinhadas

com a ferramenta *online Clustal W2*, (disponível em: <http://www.ebi.ac.uk/Tools/msa/clustalw2/>). As sequências de aminoácidos das proteínas foram obtidas a partir das sequências depositadas no banco de dados *Protein Data Bank* (PDB), referente a estrutura cristalográfica de cada Imp α utilizada no trabalho.

6.2. Resultados

6.2.1. Estrutura secundária da Imp α Nc.

O espectro de Dicroísmo Circular da Imp α Nc, evidencia a predominância de hélices- α na estrutura, indicada pelos picos negativos característicos de alfa-hélices a 208 nm (-5644,7 [θ]) e 222 nm (-15,438,8 [θ]) bem como um pico positivo a 195 nm (**Figura 22**). A Imp α Mm também apresenta os picos característicos de hélices- α (**Figura 22**). Adicionalmente, a predição da estrutura secundária da proteína Imp α Nc através do processamento dos dados pelo algoritmo CONTINN (Sreerama & Woody, 2000) resultou no seguinte perfil estrutural: 52,7 % de hélices- α , 16,4 % de loops e 24,1 % de elementos desordenados, valores muito próximos aos obtidos através do processamento dos dados da proteína Imp α Mm: 51,1 % de α -hélice, 16,5 % de *loops* e 25,8 % de elementos desordenados, corroborando o alto grau de similaridade estrutural entre as duas proteínas e a viabilidade da amostra para os experimentos.

A temperatura de desenovelamento da proteína Imp α Nc foi obtida submetendo a amostra a uma rampa de temperatura de 10 a 90 °C, a 222 nm. Os dados mostraram uma temperatura de desenovelamento de (T_m) de 39,8 °C. A partir dessa temperatura a proteína desnatura e perde o enovelamento. O experimento inverso, com a amostra submetida a uma variação de temperatura de 90 a 10 °C, a 222 nm, mostrou que a proteína não voltou a seu estado original, permanecendo desenovelada (dados não mostrados).

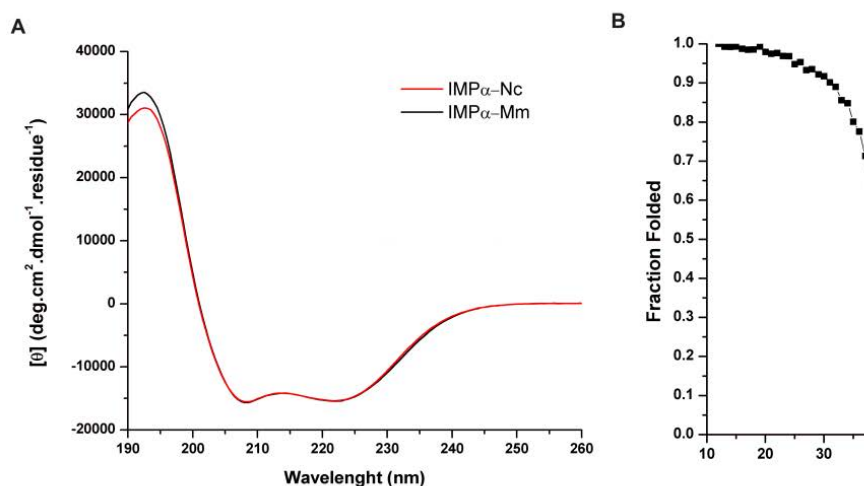


Figura 22. Verificação da estrutura secundária e temperatura de desnaturação térmica da Imp α Nc. (A) Espectro de dicroísmo circular da proteína Importina- α *Neurospora crassa* (vermelho) e *Mus musculus* (preto), obtido a 190 – 260 nm, 20 °C. (B) Gráfico de desnaturação térmica da proteína Importina- α de *Neurospora crassa* quando exposta a uma variação de temperatura de 10 a 90°C, a 222 nm, com temperatura de *desnaturaç*ão (T_m) = 39,2 °C.

6.2.2. Alinhamento das sequências de aminoácidos e sobreposição das estruturas das Imp α de diferentes organismos.

A sobreposição da estrutura da proteína Imp α Nc com estruturas já elucidadas da Imp α de arroz (*Oryza sativa*; Imp α Os; Chang *et. al.*, 2012; PDB: 4B80), levedura (*Saccharomyces cerevisiae*, Imp α Sc; Conti *et. al.*, 1998; PDB ID: IBK6), camundongo (*Mus musculus*; Imp α Mm; Fontes *et. al.*, 2000; PDB ID: IEJL) e de homem (*Homo sapiens*, Imp α Hs; PDB ID: 4B18) mostra a conservação da estrutura alongada e cônica da proteína Imp α (**Figura 24A, 24B e 24C**). O desvio dos C α das estruturas sobrepostas (r.m.s.d./ *root mean square deviation*), em relação a Imp α Nc, calculado através do programa Coot, é de: 2,66 Å para Imp α Sc, 3,49 Å para Imp α Os, 4,81 Å para a Imp α Hs e 6,37 Å para Imp α Mm.

Já o alinhamento das sequências das proteínas mostrou a conservação dos aminoácidos nas regiões correspondentes aos sítios principal e secundário da Imp α presença de aminoácidos não conservados ao redor dessas regiões (**Figura 23**).

ImpαNc	75	-----TESQLSEDLPKMVEGVFSSEIDKQIQATTKFRKLLSKERN	114	
ImpαSc	92	-----LPQMTQQLNSDDMQEQLSATVKFRQILSREHR	123	
ImpαOs	66	-----AMADIGSSLPAMIGGVYSDDNQLQLEATQFRKLLSIEERS	105	ARM1
ImpαHs	76	-----EMAPGGVITSDMIEMIFSKSPEQQLSATQKFRKLLSKEPN	115	
ImpαMm	73	-----NQGTVNWSVEDIVKGINSNNLESQLQATQAARKLLSREKQ	112	
ImpαNc	115	----PPIEEVIKT-GVVGRFVEFLR-SPHTLVQFEAAWALTNIASG	154	
ImpαSc	124	----PPIIDVVIQA-GVVPRLVEFMRENQPEMLQLEAAWALTNIASG	164	
ImpαOs	106	----PPIEEV IQS-GVVPRFVQFLTREDFPQLQFEAAWALTNIASG	146	ARM2
ImpαHs	116	----PPIIDEVISTPGVVARFVEFLKRKENCTLQFESAWVLTNIASG	157	
ImpαMm	113	----PPIIDNIIRA-GLIPKFVSVFLGKTDCSPIQFESAWALTNIASG	153	
ImpαNc	155	--SATQTQVVIIEAGAVPIFVLELLGSPEPDVREQAQVWALGNIAGDSP	198	
ImpαSc	165	--TSAQTKVVVDADAVPLFIQLLYTGSVEVKQAIWALGNVAGDST	208	
ImpαOs	147	--TSENTKVVIDHGAVPFIQVLLGSSSSDDVREQAQVWALGNVAGDST	190	ARM3
ImpαHs	158	--NSLQTRIVIQAGAVPIFIELLSSEFEDVQEQAVWALGNIAGDST	201	
ImpαMm	154	--TSEQTKAVVDGGAIPAFISLLASPHAHISEQAQVWALGNIAGDGS	197	
ImpαNc	199	QC RDYVLS CGALR P L L T L L G D S R K L S M --- L R N A T W T L S N F C R G K	240	
ImpαSc	209	DY RDYVLC QNAMEP I L G L F N - S N K P S L --- I R T A T W T L S N L C R G K	249	
ImpαOs	191	KC RD L V L A N G A L L P L L A Q L N E H T K L S M --- L R N A T W T L S N F C R G K	232	ARM4
ImpαHs	202	MC RDYVLD C N I L P L L Q L F S K Q N R L T M --- T R N A V W A L S N L C R G K	243	
ImpαMm	198	AF RD L V I K H G A I D P L L A L L A V P D L S T L A C G Y L R N L T W T L S N L C R N K	243	
ImpαNc	241	-TPQPDWNTIAPALPVLAQLVYSLDDEVLIDACWAIISYSDGSDNK	285	
ImpαSc	250	-KPQPDWSVVSQALPTLAKLIYSMDTE TLVDACWAIISYSDGQPQA	294	
ImpαOs	233	--PQPSPQTRPALPALARLIHSNDEEVLTDACWALSYSDDGTNDK	276	ARM5
ImpαHs	244	--SPPPEFAKVSPCLNVLSWLLFVSDTDVLADACWALSYSDDGPNDK	288	
ImpαMm	244	-NPAPPLDAVEQILPTLVRLHNDPEVLADSCWAIISYLDGPNER	288	
ImpαNc	286	-----IQAVIEAGIPRRLEVELMHASTSVQTPALRSVGNIVTGDD	325	
ImpαSc	295	-----IQAVIDVRIPKRLEVELSHESTLVQTPALRAVGNIVTGND	334	
ImpαOs	277	-----IQAVIEAGVCPRLVELLLHPSPSVLI PALRTVGNIVTGDD	316	ARM6
ImpαHs	289	-----IQAVIDAGVCRRLVELLMHNDYKVVSPALRAVGNIVTGDD	328	
ImpαMm	289	-----IEMVVKGVVPQLVKLLGATELPIVTPALRAI GNIVTGTD	328	
ImpαNc	326	---VQTQVIINCGALPCLLSLLSS-NKDGIRKEACWTI SNITAGNS	367	
ImpαSc	335	---LQTQVVINAGVLPALRLLSS-PKENIKKEACWTI SNITAGNT	376	
ImpαOs	317	---AQTQCIIDHQA LPCLLSLLTQNLKKS IKKEACWTI SNITAGNK	359	ARM7
ImpαHs	329	---IQTQVILNCSALQSLHLLSS-PKESIKKEACWTI SNITAGNK	370	
ImpαMm	329	---EQTQKVIDAGALAVFP SLLTN-PKTNIQKEATWMSNITAGRQ	370	
ImpαNc	368	AQIQSVIDANI IPP LIHL LSHA DLKTRKEACWAI SNAT SGG LQKPD	413	
ImpαSc	377	EQIQAVIDANL IPP LVKL LEVA EYKTKKEACWAI SNAS SGG LQRPD	422	
ImpαOs	360	DQIQAVINAGI IGP LVNLLQTA EFDIKKEAAWAI SNAT SGG SH--D	403	ARM8
ImpαHs	371	AQIQTVIDANI F P A L I S I L Q T A E F R T R K E A A W A I T N A T S G G S A --E	414	
ImpαMm	371	DQIQQVVNHGLVFP LVGV LSKA DFKTQKEAAWAI TNYT SGG TV--E	414	
ImpαNc	414	QIRYLVAQGC IKPLCDLLACPDNKIIQVALDGLENI LKVG ELDKNA	459	
ImpαSc	423	IIRYLVSQGC IKPLCDLLEIADNRIIEVTLDALENI IKMGEADKEA	468	
ImpαOs	404	QIKYLVSEGC IKPLCDLLICPDIRIVTVCLEGLENI LKVG ETDKTL	449	ARM9
ImpαHs	415	QIKYLVELGC IKPLCDLLTVMDSKIVQVALNGLENI LRLGEQAEKR	460	
ImpαMm	415	QIVYLVHCGI IEPLMNL SAKDTKIIQVILDAISNIFQAEEK----	456	
ImpαNc	460	AGDGPDS INRYALF IECCGMEKIHDCQTNANEEIYMKAYNII EKY	505	
ImpαSc	469	RG---LNINENADF IEKAGGMEKIFNCQQNENDKIYEKAYKII ETY	511	
ImpαOs	450	AAG---DVNVFSQMIDEAEGLEKIENLQSHDNNEIYEKAVKILEAY	492	ARM10
ImpαHs	461	NGT---GINPYCALIEEAYGLDKIEFLQSHENQEIYQKAFDLIEHY	503	
ImpαMm	457	LG----ETEKLSIMIECCGLDKIEALQRHENESVYKASLNLI EKY	498	
ImpαNc	506	FS----- 507		
ImpαSc	512	FG----- 513		
ImpαOs	493	WM----- 494		
ImpαHs	504	FG----- 503		
ImpαMm	499	FSVEEEE 505		

Figura 23. Comparação das sequências de aminoácidos de Impde diferentes organismos. Alinhamento das sequências de aminoácidos das proteínas ImpαNc, ImpαSc, ImpαOs, ImpαHs e ImpαMm, destacando os motivos ARMs e os principais resíduos que formam os sítios de ligação principal (ARMs 2 a 4) e secundário (ARMs 7 a 8). Em vermelho estão destacados os aminoácidos conservados e em preto as regiões não conservadas.

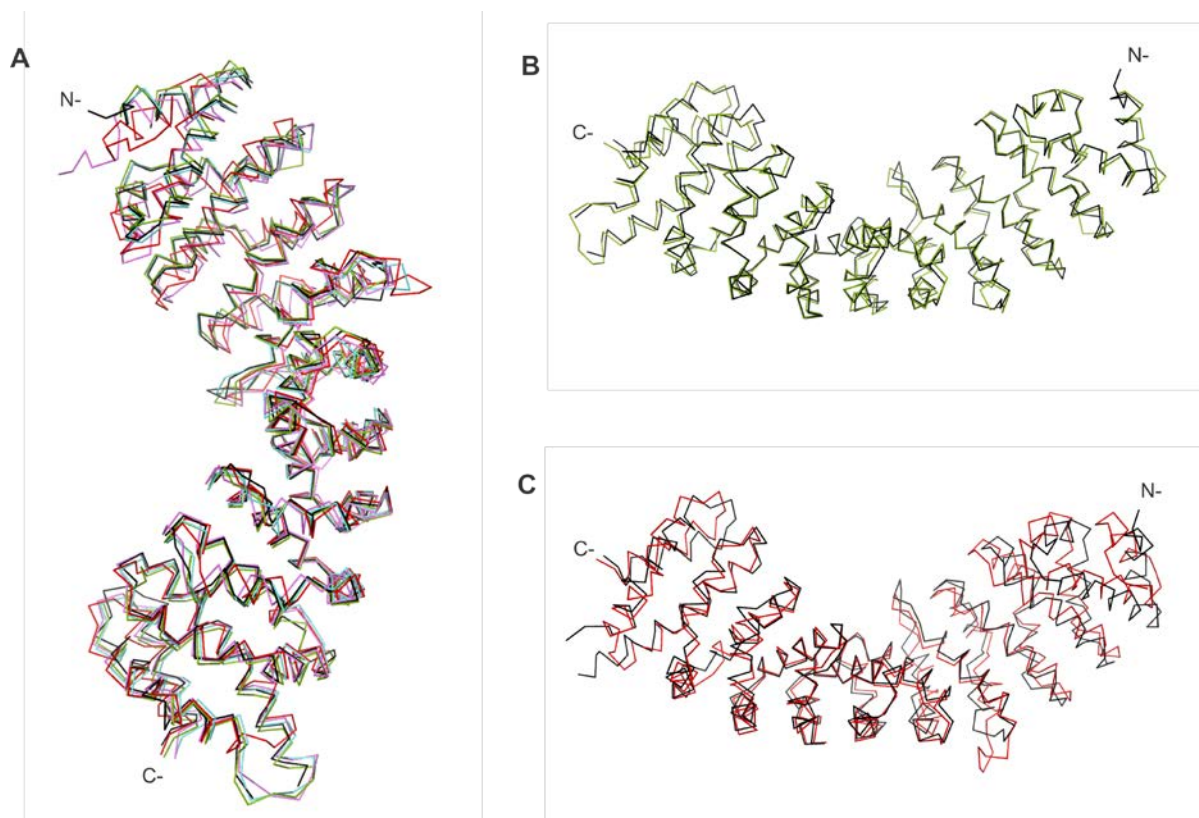


Figura 24. Comparação das estruturas de Imp de diferentes organismos. A) Estruturas das proteínas Imp α Nc (preto), Imp α Sc (verde), Imp α Mm (vermelho), Imp α Hs (rosa) e Imp α Os (azul), em formato de *ribbons*. **B)** Sobreposição da Imp α Nc com a Imp α Sc, mostrando a maior semelhança entre as proteínas. Em **C** é possível observar o maior desvio da Imp α Mm em relação a Imp α Nc.

6.3. Discussão

6.3.1. Comparação das estruturas das Imp de diferentes organismos.

Uma comparação do espectro de dicroísmo circular das proteínas Imp α Nc e Imp α Mm permitiu destacar semelhanças estruturais entre as proteínas. Apesar de pertencerem a famílias diferentes (Imp α Nc pertence a família α 1 das Importinas e a Imp α Mm, pertence a família α 2), as proteínas apresentaram uma alta conservação da estrutura secundária. Entretanto, a análise das sequências de aminoácidos e sobreposição das proteínas destacaram algumas diferenças significativas entre as proteínas Imp α de diferentes famílias.

Na sobreposição das estruturas das proteínas Imp α Os, Imp α Sc, Imp α Mm e Imp α Hs, foi possível observar o maior valor de desvio r.m.s.d. da Imp α Mm, decorrente de sua conformação mais côncava em relação ao eixo central (**Figura 24C**). Tal diferença

conformacional também é refletida na mudança dos parâmetros de rede do cristal da Imp α Nc em relação ao da Imp α Mm (Fontes *et. al.*, 2000).

A proteína Imp α Nc apresenta maior identidade com a Imp α de levedura (64,61 %), de arroz (58,98 %) e humana (60,22 %) do que com a Imp α de camundongo (46,99 %). Verificando a filogenia das proteínas Imp α de células eucarióticas, observa-se que os genes das proteínas de fungo e planta, pertencem ao mesmo grupo de família mais similares com as proteínas Imp α de animais pertencentes ao grupo α 1 do que as Imp α de animais do grupo α 2 ou α 3 (Goldfarb *et. al.*, 2004). As proteínas humana (Imp α HS), de levedura (Imp α Sc), arroz (Imp α Os) e fungo (Imp α Nc) já elucidadas e utilizadas nesse estudo, pertencem a família α 1, enquanto a de camundongo (Imp α Mm) pertence a família α 2. Sendo assim, a proximidade filogenética da proteína Imp α Nc com Imp α HS, Imp α Os e Imp α Sc justifica a maior similaridade observada entre essas proteínas do que com a Imp α Mm.

A análise das estruturas sobrepostas junto ao alinhamento das sequências das proteínas permite observar a presença de resíduos de aminoácidos conservados, principalmente nas regiões que compreendem os sítios de ligação principal e secundário das proteínas (**Figura 23**), o que indica que são os resíduos não-conservados, presentes no entorno dessas regiões, que podem ser responsáveis pelas especificidades de cada proteína no reconhecimento de NLSs.

6.3.2. Resíduos conservados apenas entre Imp α da família α 1 relacionados a especificidades na ligação a NLSs.

A elucidação do complexo da Imp α de arroz, Imp α Os, e um peptídeo NLS monopartido específico de proteínas de plantas (B54NLS: SVLGKRKRHPKV), mostrou que esse peptídeo se ligou preferencialmente ao sítio secundário da proteína, diferente do esperado para uma NLS monopartida (Chang *et. al.*, 2012). O mesmo peptídeo, quando co-cristalizado com a Imp α Mm, ligou-se, preferencialmente, ao sítio principal. Comparando as estruturas e as sequências de aminoácidos das proteínas Imp α Os e Imp α Mm, foi possível observar a presença de alguns aminoácidos não-conservados entre as duas proteínas (S394, R427, E434, E480, K484, em Imp α Os; T404, K435, D442, S483, A487, em Imp α Mm), localizados na região C-terminal, próxima ao sítio secundário, provavelmente responsáveis pelas especificidades das proteínas na ligação de determinadas NLSs monopartidas (Chang, *et. al.*, 2012). Alguns destes resíduos, encontrados na região C-terminal da Imp α Os, também estão presentes na Imp α Nc (S402, E493, K497) (**Figura 25**) e outras proteínas da família α 1:

Imp α Sc (S408, R443, K500) e Imp α Hs (E491, K495). A presença de um resíduo de treonina (T404) na Imp α Mm no lugar das serinas encontradas na Imp α Os e Imp α Nc (S394 e S402, respectivamente) provoca um impedimento estérico para determinados resíduos presentes na porção N-terminal do peptídeo de NLS específico de plantas (Chang, et. al., 2012; **Figura 25**). Enquanto que, os resíduos de lisina e ácido glutâmico na Imp α Os, Imp α Hs, Imp α Nc e Imp α Sc formam uma região favorável para ligação de resíduos de valina e leucina do peptídeo da NLS específico de plantas (**Figura 25**).

Considerando a maior proximidade filogenética da proteína Imp α Nc com a proteína Imp α Os do que com a Imp α Mm (Goldfarb *et. al.*, 2004) é possível que esses mesmos resíduos estejam relacionados com especificidades na interação de peptídeos NLSs de fungos com a proteína Imp α Nc. Através da cristalização da Imp α Nc com NLSs monopartidas específicas, essa possível interação diferencial poderá ser verificada.

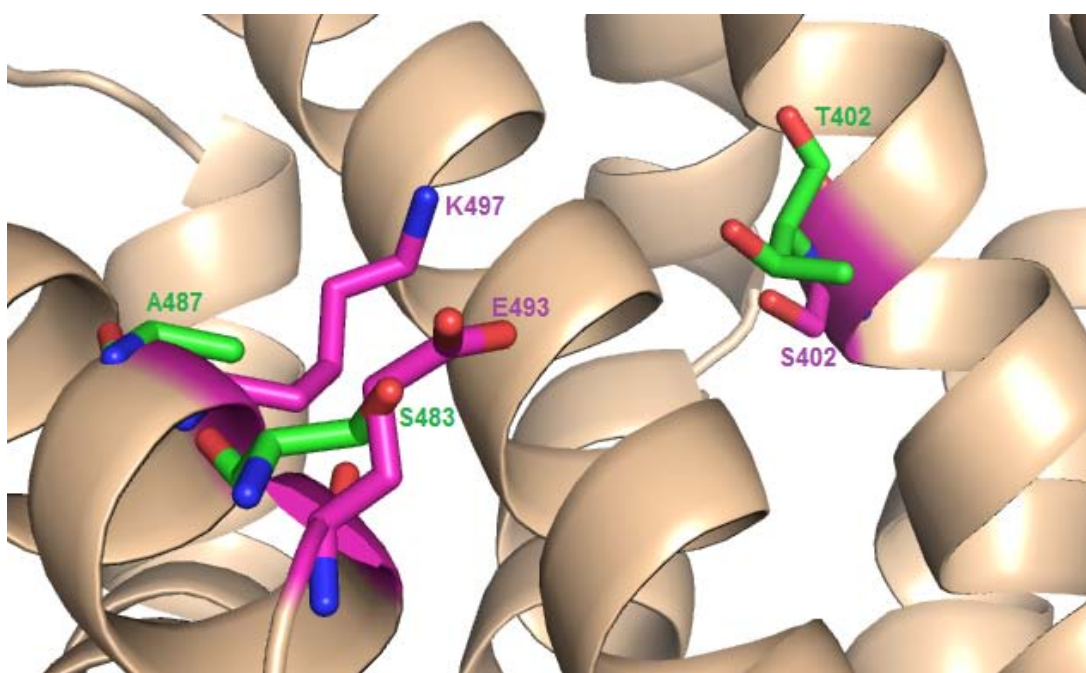


Figura 25. Resíduos não-conservados entre as estruturas da Imp α Mm e Imp α Nc que podem estar relacionados a especificidades na interação com NLSs. Estrutura da Imp α Nc (bege, representada em diagrama de fitas) com os resíduos do sítio secundário que foram identificados na Imp α Os e podem estar relacionados com especificidades na ligação a NLSs (rosa, representados em bastões). A sobreposição da estrutura da Imp α Mm (verde, representada em bastões) com a Imp α Nc mostra resíduos não conservados na mesma posição. O resíduo T402 na Imp α Mm provoca um impedimento estérico para determinados resíduos de NLSs, enquanto os resíduos S402, K497 e E493, formam uma região favorável para a interação com NLSs específicos.

***7. Afinidade do peptídeo SV40 NLS a
ImpαNc.***

7.1. Material e Métodos

7.1.1. Calorimetria de titulação isotérmica.

Nos experimentos de calorimetria de titulação isotérmica com a proteína Imp α Nc foi utilizado o peptídeo SV40 NLS, como padrão para experimentos posteriores de caracterização de potenciais NLSs.

As medidas foram realizadas utilizando o microcalorímetro MicroCal iTC200 (GE Healthcare). As amostras de Imp α Nc a 50 μ M foram tituladas com os peptídeos em concentração 20 vezes maior em relação a proteína. As titulações foram realizadas em 20 injeções de 2,0 μ L com intervalo de 240 s entre cada titulação, a temperatura de 10 °C, e 1000 rpm de velocidade de homogeneização. Foi realizada a titulação do peptídeo no mesmo tampão da amostra da proteína (20 mM Tris HCl pH 8,0 e 100 mM de NaCl) para descontar o calor de diluição das titulações do peptídeo SV40 NLS na Imp α Nc. Os dados obtidos foram processados nos programas NITPIC (Keller *et. al.*, 2012) e SEDPHAT (Houtman *et. al.*, 2007) para a obtenção das constantes de dissociação, entalpia, entropia e energia livre de Gibbs.

7.2. Resultados

7.2.1. Curva de titulação e constantes termodinâmicas da interação entre Imp α Nc e SV40NLS.

Através da técnica de calorimetria de titulação isotérmica, foram realizados ensaios de afinidade entre a proteína Imp α Nc e o peptídeo SV40 NLS, em uma proporção molar de 1:20 entre proteína e peptídeo, respectivamente. A interação entre proteína e peptídeo gerou uma curva de titulação adequada para o processamento dos dados (**Figura 26**), que permitiu o cálculo das constantes termodinâmicas envolvidas nessa interação (**Tabela 7**).

A Imp α apresenta dois sítios de ligação (principal e secundário) e o peptídeo SV40 NLS é uma NLS monopartida, ou seja, pode se ligar tanto ao sítio principal quanto secundário da proteína. Sendo assim, os dados foram processados considerando um modelo de interação

que indica dois sítios de ocupação não-simétricos. Devido ao bom ajuste da curva quando considerado esse modelo de interação, será dada preferência a esse modelo nas próximas análises.

Tabela 7. Valores das constantes termodinâmicas que definem a afinidade da interação da Imp α Nc com o peptídeo SV40 NLS. Os valores foram obtidos a partir de experimentos de Calorimetria de Titulação Isotérmica (ITC), através da titulação da Imp α Nc com SV40 NLS. Os dados apresentados foram gerados através do programa Sedphat.

Constantes termodinâmicas	Sítio 1	Sítio 2
Kd (μ M)	$0,15 \pm 0,1$	$1,4 \pm 0,04$
ΔH (Kcal/Mol)	$-2,18 \pm 0,12$	$-7,14 \pm 0,18$
ΔG (Kcal/Mol)	$-8,82$	$-7,56$
ΔS (cal/Mol*K)	$23,44$	$23,44$

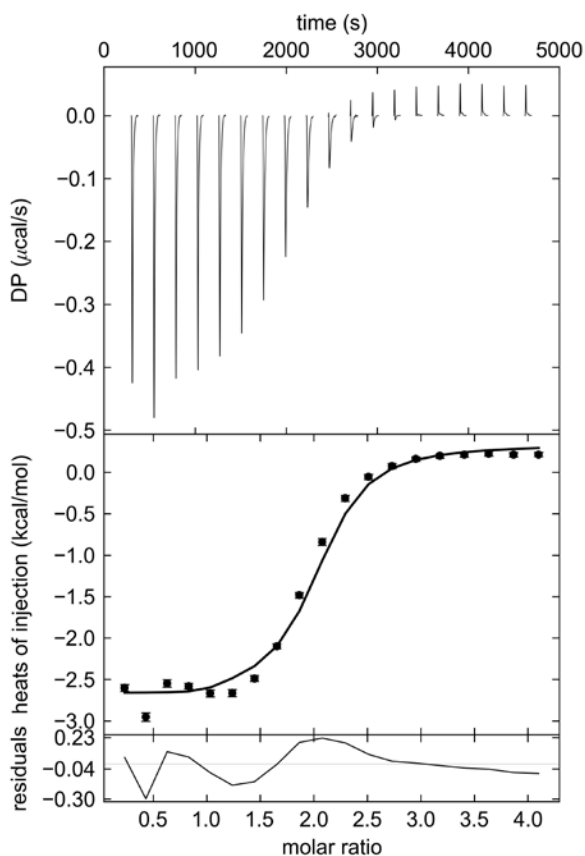


Figura 26. Termograma de titulação isotérmica do complexo Imp/SV40 NLS. O termograma é referente a titulação da proteína Imp α Nc a 50 mM com o peptídeo SV40 NLS a 1 mM, 10 °C.

7.3. Discussão

7.3.1. Afinidade entre peptídeos SV40 NLS e a proteína Imp α Nc.

Os dados de calorimetria de titulação isotérmica (ITC) mostraram que a variação da entropia (ΔS) nos dois sítios, contribui mais do que a variação da entalpia (ΔH) na variação total da energia de interação (ΔG), o que indica a espontaneidade da interação da proteína com o peptídeo (**Tabela 7**). No sítio 1, a contribuição da variação da entropia é maior do que no sítio 2, assim como a variação da energia total da interação. Além disso, foi possível observar que o valor da constante de dissociação (K_d) no sítio 1 é menor do que no sítio 2. Como o aumento da entropia indica um maior equilíbrio da interação e o menor valor de K_d indica uma maior afinidade entre as moléculas, sugere-se que os sítios 1 e sítio 2 correspondem aos sítios principal e secundário da proteína, respectivamente, pois nas análises estruturais foi possível observar a ligação preferencial do peptídeo SV40 NLS ao sítio principal.

As variações de energia e entropia são decorrentes de diferentes tipos de interação entre a proteína e o ligante: interações polares do peptídeo com as cadeias laterais de asparaginas da Imp α , interações hidrofóbicas com as cadeias laterais de triptofanos da Imp α , interações de carga das cadeias laterais básicas da NLS com resíduos negativos da Imp α , interações dos resíduos básicos da NLS com os dipolos das hélices da Imp α .

Os valores obtidos no experimento com a Imp α Nc são próximos aos da literatura referente à interação da Imp α de mamífero com o mesmo peptídeo (Cutres *et. al.*, 2008), no entanto, percebe-se algumas variações na afinidade, que podem ser decorrentes tanto da variação do equipamento e *software* utilizado nas análises ou de especificidades na interação da Imp α com NLSs.

Também foi possível observar que ao mesmo tempo em que ocorria a interação entre as amostras do peptídeo e a proteína um segundo processo estava ocorrendo, provavelmente uma oligomerização de alguns dos complexos formados, levando ao aparecimento de picos indicando uma reação endotérmica nas injeções finais. Mesmo com a diminuição da concentração da proteína em relação a do peptídeo, o perfil da curva permaneceu o mesmo e um perfil semelhante pode ser observado em experimentos com a isoforma de mamífero junto ao peptídeo FEN1 NLS, indicando que esse processo é característico da interação de peptídeos NLSs com a Imp α .

As variações termodinâmicas observadas na interação entre o SV40 NLS e a proteína Imp α Nc são utilizados como referência para a verificação da interação da proteína com potenciais NLSs específicos de fungos.

8. Conclusões

A proteína Imp α Nc mostrou uma maior estabilidade na presença de peptídeos NLS, que garantem a conservação do perfil monomodal da proteína em solução. Sendo assim, a cristalização da Imp α Nc só foi possível com a presença do peptídeo SV40 NLS.

A estrutura da Imp α Nc apresenta motivos ARMs conservados que levam a uma conformação alongada e côncava, conforme já visto na estrutura de Imp α de outros organismos, sendo que, a Imp α Nc apresenta uma conformação mais estendida do que a Imp α Mm. O peptídeo SV40 NLS se ligou aos sítios principal e secundário da Imp α Nc, sendo que, no sítio secundário, o peptídeo se ligou de forma similar a Imp α Os, com K129 e R130 nas posições P1' e P2', respectivamente.

A Imp α Nc apresentou mais similaridades com outras proteínas da família α 1 (Imp α Os, Imp α Sc e Imp α Hs) do que com a Imp α Mm, da família α 2. Inclusive foi possível observar a presença de aminoácidos conservados no sítio secundário, apenas entre as proteínas da família α 1, que podem estar relacionados com a ligação preferencial de peptídeo NLSs específicos, a esse sítio.

Por fim, ensaios de afinidade com a Imp α Nc junto ao peptídeo SV40 NLS permitiram quantificar a interação entre proteína e peptídeo. Tais valores poderão ser utilizados como padrão na análise de potenciais peptídeos NLSs específicos de fungos.

9. Perspectivas

Os próximos passos para caracterizar especificidades da Imp α Nc incluem a identificação de NLSs específicos de fungos, através de buscas em banco de dados e ensaios de afinidade, com destaque para NLSs de proteínas envolvidas com o metabolismo de glicogênio.

Serão conduzidos experimentos de cristalização com a Imp α Nc junto aos NLSs identificados, seguidos de análises estruturais e funcionais que permitirão a caracterização da interação desses peptídeos com a proteína, auxiliando na compreensão do mecanismo de transporte intracelular de macromoléculas em fungos e, por consequência, do metabolismo de fungos.

10. Referências bibliográficas

ADAM, S. A., GERACE, L. Cytosolic proteins that specifically bind nuclear localization signals are receptors for nuclear import. **Cell**, v.66, 837-847.19991.

ADAM, S. A.; STERNE-MARR, R. E.; GERACE, L. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. **J. Cell Biol**, v.777, p.807-816. 1990.

ADAMS, P. D.; AFONINE, P. V.; BUNKÓCZI, G; CHEN, V. B.; DAVIS, I. W.; ECHOLS, N.; HEADD, J. J.; HUNG, L.-W.; KAPRAL, G. J.; GROSSE-KUNSTLEVE, R. W.; MCCOY, A. J.; MORIARTY, N. W.; OEFFNER, R.; READ, R. J.; RICHARDSON, D. C.; RICHARDSON, J. S.; TERWILLIGER, T. C.; ZWART, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Cryst**, v.D66, p.213-221. 2010.

AFONINE, P. V.; GROSSE-KUNSTLEVE, R. W.; ADAMS, P. D. The Phenix refinement framework. **CCP4 Newslett**. 42, contribution 8, 2005.

AFONINE, P. V.; GROSSE-KUNSTLEVE, R.W.; ECHOLS, N.; HEADD, J. J.; MORIARTY, N. W.; MUSTYAKIMOV, M.; TERWILLIGER, T. C.; URZHUMTSEV, A.; ZWART P. H., ADAMS, P. D. Towards automated crystallographic structure refinement with *phenix.refine*. **Acta Cryst**, v.68, p.352-367.2012.

AKEY, C. W.; AND GOLDFARB, D. S.. Protein import through the nuclear pore complex is a multistep process. **J. Cell Biol**, v.709, p.971-982. 1989.

ALTAN, B.; YOKOBORI, T.; MOCHIKI, E.; OHNO, T.; OGATA, K.; OGAWA, A.; YANAI, M.; KOBAYASHI, T.; LUVSANDAGVA, B.; ASAO, T.; KUWANO, H. Nuclear karyopherin- $\alpha 2$ expression in primary lesions and metastatic lymph nodes was associated with poor prognosis and progression in gastric cancer. **Carcinogenesis**, vol.34, p.2314–2321.2013.

ANDRADE, M. A.; PETOSA, C.; O'DONOGHUE, S. I.; MULLER, C. W.; BORK, P. Comparison of ARM and HEAT protein repeats. **J Mol Biol**, v.309, p.1-18. 2001.

ARAUJO-BAZAN, L.; FERNANDEZ-MARTINEZ, J.; DE LOS RÍOS, V. M.; ETXEESTE, O.; ALBAR, J. P.; PENALVA, M. A.; ESPESO, E. A. NapA and NapB are the *Aspergillus nidulans* Nap/SET family members and NapB is a nuclear protein specifically interacting with importin a. **Fung. Gen. and Biol**, v.45, p.278–291. 2008.

ARMS, K. Cytonucleoproteins in cleaving eggs of *Xenopus laevis*. **J. Embryol. Exp. Morphol**, v.20, p.367-374.1968.

BAYLISS, R.; LITTLEWOOD, T.; STEWART, M. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. **Cell**, v.102, p.99-108. 2000.

BARROS, A. C.; TAKEDA, A. A. S.; CHANG, C. -W.; KOBE, B.; FONTES, M. R. M. Structural basis of nuclear import of flap endonuclease 1 (FEN1). **Acta Cryst**. v.D68, p.743–750. 2012.

- BEADLE, G.W.; TATUM, E. L. Genetic control of biochemical reactions in *Neurospora*. **Proc. Natl Acad. Sci. USA** 27, 499–506. 1941.
- BONNER, W. M. Protein migration into nuclei . II . Frog oocyte nuclei accumulate a class of microinjected oocyte nuclear proteins and exclude a class of microinjected oocyte cytoplasmic proteins . **J . Cell Biol**, v.64, p.431- 437. 1975.
- BRISTOWE, J. S. Transactions of the Pathological Society of London. v.5, p.38-41.1854.
- CALLAN, H. G.; TOMLIN, S. G. Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. **Proc. R. Soc.**, v.B137, pp. 367–378. 1950.
- CHANG, C.-W.; COUÑAGO, R. L. M.; WILLIAMS, S. J.; BODÉN, M.; KOBE, B. Crystal Structure of Rice Importin- α and Structural Basis of Its Interaction with Plant-Specific Nuclear Localization Signals. **The Plant Cell**, v. 24, p.5074–5088. 2012.
- CHEN, M-H.; BEN-EFRAIM, I.; MITROUSIS, G.; WALKER-KOPP, N.; SIMS, P. J.; CINGOLANI, G. Phospholipid Scramblase 1 Contains a Nonclassical Nuclear Localization Signal with Unique Binding Site in Importin α . **J. Biol. Chem**, v.280, p..10599-10606. 2005.
- CHOOK, Y. M.; BLOBEL, G. Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. **Nature**, v.399, p.230-7. 1999.
- CINGOLANI, G.; BEDNENKO, J.; GILLESPIE, M. T.; GERACE, L. Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta. **Mol Cell**, v.10, p.1345-53. 2002.
- CINGOLANI, G.; PETOSA, C.; WEIS, K.; MULLER, C. W. Structure of importin-beta bound to the IBB domain of importin-alpha. **Nature**, v.399, p.221-9. 1999.
- COLLEDGE, W. H.; RICHARDSON, W. D.; EDGE, M. D.; SMITH, A. E. Extensive Mutagenesis of the Nuclear Location Signal of Simian Virus 40 Large-T Antigen. **Mol. Cell. Biol**, v.6, p.4136 - 4139. 1986.
- CONTI, E.; IZAURRALDE, E. Nucleocytoplasmic transport enters the atomic age. **Curr Opin Cell Biol**, v.13, p.310-9. 2001.
- CONTI, E.; KURIYAN, J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. **Structure**, v.8, p.329-38. 2000.
- CONTI, E.; UY, M.; LEIGHTON, L.; BLOBEL, G.; KURIYAN, J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. **Cell**, v.94, p.193-204. 1998.
- COOK, A.; BONO, F.; JINEK, M.; CONTI, E. Structural biology of nucleocytoplasmic transport. **Annu Rev Biochem**, v.76, p.647-71. 2007.
- CUTRESS, M. L.; WHITAKER, H. C.; MILLS, I. G.; STEWART, M.; NEAL, D. E. Structural basis for the nuclear import of the human androgen receptor. **J Cell Sci**, v.121, p.957-68. 2008.

- DAVIS, R. H. & PERKINS, D. D. Neurospora: a model of model microbes. **Nature Rev. Genet.** v.3, 397–403. 2002.
- DAVIS, R. H. Neurospora: Contributions of a Model Organism. **Oxford Univ. Press.** New York. 2000.
- DEGRASSE, J. A.; DUBOIS, K. N.; DEVOS, D.; SIEGEL, T. N.; SALI, A.; FIELD, M. C.; ROUT, M. P.; CHAIT, B. T. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. **Mol Cell Proteomics**, v.8, p.2119-2130. 2009.
- DELANO, W.L. The PyMOL User's Manual. **DeLano Scientific LLC**, San Carlos, CA .2002.
- DINGWALL, C.; LASKEY, R. A. Nuclear targeting sequences--a consensus? **Trends Biochem Sci**, v.16, p.478-81. 1991.
- DINGWALL, C.; SHARNICK, S. V.; LASKEY, R. A. A polypeptide domain that specifies migration of nucleoplasmin into the nucleus. **Cell**, v.30, p.449-58. 1982.
- DRENTH, J. Principles of protein X-ray crystallography. New York: Springer-Verlag. v.311. 1994.
- EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Crystallogr. D.**, v.60, p.2126-2132. 2004.
- FANARA, P.; HODEL, M. R.; CORBETT, A. H.; HODEL, A. E. Quantitative analysis of nuclear localization signal (NLS)-importin alpha interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding. **J Biol Chem**, v.275, p.21218-23. 2000.
- FANG, X.; CHEN, T.; TRAN, K.; PARKER, C.S. Developmental regulation of the heat shock response by nuclear transport factor karyopherin-alpha3. **Development**, v.128, p.3349–3358. 2001.
- FERNANDEZ-MARTINEZ, J.; ROUT, M. P. A jumbo problem: mapping the structure and functions of the nuclear pore complex. **Curr Opin in Cell Biol**, v.24, p.92–99. 2012.
- FONTES, M. R.; TEH, T.; JANS, D.; BRINKWORTH, R. I.; KOBE, B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. **J Biol Chem**, v.278, p.27981-7. 2003a.
- FONTES, M. R.; TEH, T.; KOBE, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. **J Mol Biol**, v.297, p.1183-94. 2000.
- FRANKE, W. W.; SCHEER, U. Pathways of nucleocytoplasmic translocation of ribonucleoproteins. **Symp Soc Exp Biol**, p.249-82. 1974.

- FREIRE, E.; MAYORGA, O. L.; STRAUME, M. Isothermal titration calorimetry. **Anal. Chem.**, v.62 (18), p. 950A–959A. 1990.
- GALAGAN, J.E.; CALVO, S.E.; BORKOVICH, K.A.; SELKER, E.U.; READ, N.D.; JAFFE, D.; FITZHUGH, W.; MA, L.J.; SMIRNOV, S.; PURCELL, S.; ET AL. The genome sequence of the filamentous fungus *Neurospora crassa*. **Nature**, v.422, p.859–868. 2003.
- GALL, J. G. Octagonal Nuclear Pores. **J. Cell Biol**, v.32, p.391-399. 1967.
- GOLDFARB, D. S.; CORBETT, A. H.; MASON, D. A.; HARREMAN, M. T.; ADAM, S. A. Importin alpha: a multipurpose nuclear-transport receptor. **Trends Cell Biol**, v.14, p.505-14. 2004.
- GONÇALVES R.D.; CUPERTINO F.B.; FREITAS, F.Z.; LUCHESSI, A.D.; BERTOLINI M.C. A genome-wide screen for *Neurospora crassa* transcription factors regulating glycogen metabolism. *Mol Cell Proteomics* 10: DOI 10.1074/mcp.M111.007963-2. 2011.
- GORLICH, D. Transport into and out of the cell nucleus. **EMBO J**, v.17, p.2721-7. 1998.
- GORLICH, D.; KUTAY, U. Transport between the cell nucleus and the cytoplasm. **Annu Rev Cell Dev Biol**, v.15, p.607-60. 1999.
- GORLICH, D.; MATTAJ, I. W. Nucleocytoplasmic transport. **Science**, v.271, p.1513-8. 1996.
- GORLICH, D.; PREHN, S.; LASKEY, R. A.; HARTMANN, E. Isolation of a protein that is essential for the first step of nuclear protein import. **Cell**, v.79, p.767-78. 1994.
- GORLICH, D.; KOSTKA, S.; KRAFT, R.; DINGWALL, C.; LASKEY, R.A.; HARTMANN, E.; PREHN, S. Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. **Curr. Biol.**, v. 5, p.383–392. 1995a.
- GORLICH, D.; VOGEL, F.; MILLS, A. D.; HARTMANN, E.; LASKEY, R. A. Distinct functions for the two importin subunits in nuclear protein import. **Nature**, v.377, p.246-8. 1995b.
- GREENFIELD, N.J. Analysis of circular dichroism data. **Methods Enzymol**, 383, 282-317. 2004
- GURDON, J. B. Nuclear transplantation and the control of gene activity in animal development . **Proc. Royal Soc . Lond**, v.B176, p.303-314.1970.
- HANNEMANN, H.; SUNG, P.; CHIU, H.; YOUSUF, A.; BIRD, J.; LIM, S. P.; DAVIDSON, A. D. Serotype-specific Differences in Dengue Virus Non-structural Protein 5 Nuclear Localization. **J. Biol. Chem**, v.288, p.22621-22635.2013.
- HODEL, M. R.; CORBETT, A. H.; HODEL, A. E. Dissection of a nuclear localization signal. **J. Biol. Chem**, v.276, p.1317-1325. 2000.
- HOUTMAN, J. C. D.; BROWN, P. H.; BOWDEN, B.; YAMAGUSHI, H.; APPELLA, E.;

SAMELSON, L.E.; SCHUCK, P. Studying multi-site binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling. **Protein Science**, v.16, 30-42. 2007.

HUBNER, S; SMITH, H. M. S.; HU, W.; CHAN, C. K.; RIHS, H.; PASCHALI, B. M.; RAIKHEL, N. V.; JANS, D. A. Plant Importin α Binds Nuclear Localization Sequences with High Affinity and Can Mediate Nuclear Import Independent of Importin- β . **THE JOURNAL OF BIOLOGICAL CHEMISTRY**, v.274, pp. 22610–22617.1999.

ISGRO, T. A.; SCHULTEN, K. Binding dynamics of isolated nucleoporin repeat regions to importin- α . **Structure**, v.13, p.1869–1879. 2005.

ISGRO, T. A.; SCHULTEN, K. Cse1p-binding dynamics reveal a binding pattern for FG-repeat nucleoporins on transport receptors. **Structure**, v.15, p.977–991. 2007b.

ISGRO, T. A.; SCHULTEN, K. Association of nuclear pore FG-repeat domains to NTF2 import and export complexes. **J Mol Biol**, v.366, p.330–345. 2007a.

JANCARIK, J.; KIM, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. **J. Appl. Cryst**, v.24, p.409-411.1991.

KALDERON, D.; ROBERTS, B.L.; RICHARDSON, W.D.; SMITH, A.E. A short amino acid sequence able to specify nuclear location. **Cell**, v.39, p.499–509. 1984.

KELLER, S.; VARGAS, C.; ZHAO, H.; PISZCZEK, G.; BRAUTIGAM, C. A.; SCHUCK, P. High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis. **Analytical Chemistry**, v.84, 5066–5073. 2012.

KEMINER, O.; SIEBRASSE, J. P.; ZERF, K.; PETERS, R. Optical recording of signal-mediated protein transport through single nuclear pore complexes. **Proc Natl Acad Sci U S A**, v.96, p.11842-7. 1999.

KESSEL, R. G. Fine structure of the pore-annulus complex in the nuclear envelope and annulate lamellae of germ cells. **Z. ellforsch**, v.94, pp. 441–453. 1969.

KNUDSEN, N. O.; ANDERSEN, S. D.; LUTZEN, A.; NIELSEN, F. C.; RASMUSSEN, L. J. Nuclear translocation contributes to regulation of DNA excision repair activities. **DNA Repair (Amst)**, v.8, p.682-9. 2009.

KOBE, B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin α . **Nat Struct Biol**, v.6, p.388-97. 1999.

KÖHLER, M.; ANSIEAU, S.; PREHN, S.; LEUTZ, A.; HALLER, H.; *ET AL.* Cloning of two novel human importin- α subunits and analysis of the expression pattern of the importin- α protein family. **FEBS Lett**, v.417, p.104–108. 1997.

KOHLER, M.; HALLER, H.; HARTMANN, E. Nuclear protein transport pathways. **Exp Nephrol**, v.7, p.290-4. 1999.

KOSUGI, S.; HASEBE, M.; MATSUMURA, N.; TAKASHIMA, H.; MIYAMOTO-SATO, E.; TOMITA, M.; YANAGAWA, H. Six classes of nuclear localization signals specific to different binding grooves of importin alpha. **J. Biol. Chem**, v.284, p.478–485. 2009.

KRISSINEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. **J. Mol. Biol.** v.372, p.774--797. 2007.

KRULL, S.; THYBERG, J.; BJORKROTH, B.; RACKWITZ, H. R.; CORDES, V. C. Nucleoporins as components of the nuclear pore complex core structure and Tpr as the architectural element of the nuclear basket. **Mol Biol Cell**, v.15, p.4261-4277. 2004.

KUHN, J. Die Krankheiten der Culturgewuchsp, ihre Uhrsachen und ihre. 1858.

LANFORD, R. E.; BUTEL, J. S. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. **Cell**, v.37, p.801-13. 1984.

LANFORD, R. E.; KANDA, P.; KENNEDY, R. C. Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. **Cell**, v.46, p.575-82. 1986.

LANGE, A.; MILLS, R. E.; LANGE, C. J.; STEWART, M.; DEVINE, S. E.; CORBETT, A. H. Classical nuclear localization signals: definition, function, and interaction with importin alpha. **J Biol Chem**, v.282, p.5101-5. 2007.

LEE, S. J.; SEKIMOTO, T.; YAMASHITA, E.; NAGOSHI, E.; NAKAGAWA, A.; IMAMOTO, N.; YOSHIMURA, M.; SAKAI, H.; CHONG, K. T.; TSUKIHARA, T.; YONEDA, Y. The structure of importin-beta bound to SREBP-2: nuclear import of a transcription factor. **Science**, v.302, p.1571-5. 2003.

LEWIS, E. A.; MURPHY, K. P. Isothermal Titration Calorimetry. **Met. Mol. Biol.**, v.305, p.1-15. 2005.

LIM, R. Y. H.; AEBI, U.; FAHRENKROG, B. Towards reconciling structure and function in the nuclear pore complex. **Histochem Cell Biol**, v.129, p.105–116. 2008.

MARFORI, M.; MYNOTT, A.; ELLIS, J. J.; MEHDI, A. M.; SAUNDERS, N. F. W.; CURMI, P.; FORWOOD, J.; BODEN, M.; KOBE, B. Molecular basis for specificity of nuclear import and prediction of nuclear localization. **Biochimica and Biophysica Acta**, v. 1813, 1562-1577. 2011.

MASON, D. A.; FLEMING, R. J.; GOLDFARB, D. S. Drosophila melanogaster Importin $\alpha 1$ and $\alpha 3$ can replace Importin $\alpha 2$ during spermatogenesis but not oogenesis. **Genetics**, v.161, p.157–170. 2002.

MCCOY, A. J.; GROSSE-KUNSTLEVE, R. W.; STORONI, L. C.; READ, R. J. Phaser crystallographic software. **Acta Crystallogr**, D61, 458–464. 2007.

MCREE, D. E. Practical protein crystallography: San Diego: **Academic Press**. 1993.

MELCHIOR, F.; PASCHAL, B.; EVANS, E.; GERACE, L. Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. **J. Cell Biol**, v.723, 1649-1659.1993.

MERRIAM, R. W. Movement of cytoplasmic proteins into nuclei induced to enlarge and initiate DNA or RNA synthesis. **J. Cell Sci**, v.5, p.333-349.1969.

MOORE, M. S., BLOBEL, G. A G protein involved in nucleocytoplasmic transport: the role of Ran. **Trends Biochem. Sci**, v.79, p.211-216. 1994.

MOORE, M. S.; BLOBEL, G. The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. **Nature**, v.365, p.661-3. 1993.

NEWMAYER, D. D.; AND FORBES, D. J. Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation. **Cell**, v.52, p.641-653. 1988.

OTWINOWSKI, Z.; MINOR, W. Processing of X-ray diffraction data collected in oscillation mode. **Methods Enzymol**, v.276, p.307-326. 1997.

PFEIFER, M.; BERG, S.;REYNOLDS, A.B. A repeating amino acid motif shared by proteins with diverse cellular roles. **Cell**, v.76, p.789-791. 1994

PUMROY, R.A.; NARDOZZI, J.D.; HART, D.J.; ROOT, M.J.; CINGOLANI, G. Nucleoporin Nup50 stabilizes closed conformation of armadillo repeat 10 in importin alpha5. **J. Biol. Chem.**, v.287, pp. 2022–2031. 2012.

RICHARDSON, W. D.; MILLS, A. D.; DILWORTH, S. M.; LASKEY, R. A.; DINGWALL, C. Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. **Cell**, v.52, p.655-664. 1988.

ROUT, M. P.; AITCHISON, J. D.; SUPRAPTO, A.; HJERTAAS, K.; ZHAO, Y.; CHAIT, B. T. The yeast nuclear pore complex: composition, architecture, and transport mechanism. **J Cell Biol**, v.148, p.635-51. 2000.

SEKIMOTO,T.; IMAMOTO, N.; NAKAJIMA, K.; HIRANO, T.; YONEDA, Y. Extracellular signal-dependent nuclear import of Stat1 is mediated by nuclear pore-targeting complex formation with NPI-1, but not Rch1. **EMBO J**. v.16, p.7067–7077. 1997.

SHABMAN, R. S.; GULCICEK, E. E.; STONE, K. L.; BASLER, C. F. The Ebola Virus VP24 Protein Prevents hnRNP C1/C2 Binding to Karyopherin $\alpha 1$ and Partially Alters its Nuclear Import. **J. Inf. Dis.**, v.204, p.S904–S910.2011.

SOROKIN, A. V.; KIM, E. R.; OVCHINNIKOV, L. P. Nucleocytoplasmic transport of proteins. **Biochemistry (Mosc)**, v.72, p.1439-57. 2007.

SREERAMA, N.; WOODY, R.W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. **Anal. Biochem.**, 287(2), p.252-260. 2000.

STEWART, M. Molecular mechanism of the nuclear protein import cycle. **Nat Rev Mol Cell Biol**, v.8, p.195-208. 2007.

STEWART, M. Structural basis for the nuclear protein import cycle. **Biochem Soc Trans**, v.34, p.701-4. 2006.

STINNETT, S. M.; ESPESO, E. A.; COBEÑO, L.; ARAÚJO-BAZÁN, L. E CALVO, A. M. *Aspergillus nidulans* VeA subcellular localization is dependent on the importin a carrier and on light. **Mol. Microb.** v.63, p.242–255.2007.

SWEET, B.H.; HILLEMAM, M. R. The Vacuolating Virus, S.V.⁴⁰. **Exp Biol Med.** v. 105, p.2 420-427. 1960.

TAKEDA, A. A. S.; DE BARROS, A. C.; CHANG, C-W; KOBE, B. ; FONTES, M. R. M. Structural Basis of Importin- α -Mediated Nuclear Transport for Ku70 and Ku80. **J. Mol. Biol.** v.412, p.226–234. 2011.

TAKEDA, A. A. S.; FREITAS, F. Z.; MAGRO, A. J.; BERNARDES, N. E.; FERNANDES, C. A. H.; GONCALVES, R. D.; BERTOLINI, M. C.; FONTES, M. R. M. Biophysical characterization of the recombinant importin- α from *Neurospora crassa*. **Prot. Pep. Letters.** 20(1), 8-16.2013.

TARENDEAU, F., BOUDET, J., GUILLIGAY, D., MAS, P. J., BOUGAULT, C. M., BOULO, S., BAUDIN, F., RUIGROK, R. W., DAIGLE, N., ELLENBERG, J., CUSACK, S., SIMORRE, J. P.; HART, D. J. Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. **Nat. Struct. Mol. Biol.**, v.14, p.229–233. 2007.

TEH, T.; TIGANIS, T.; KOBE, B. Crystallization of importin alpha, the nuclear-import receptor. **Acta Crystallogr D Biol Crystallogr**, v.55, p.561-3. 1999.

TERWILLIGER, T.C.; GROSSE-KUNSTLEVE, R.W.; AFONINE, P.V.; MORIARTY, N.W.; ZWART, P.H.; HUNG, L.-W.; READ, R.J.; ADAMS, P.D. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. **Acta Cryst.** v.64, p.61-69.2008.

WALLACE, A. C.; LASKOWSKI, R. A.; THORNTON, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. **Protein Eng**, v.8, p.127-34. 1995.

ANDRADE, M. A.; BORK, P. HEAT repeats in the Huntington's disease protein. **Nature Genet.** v.11, p.115–116. 1995.

Apêndices

Apêndice A

Resumo das técnicas biofísicas utilizadas no presente trabalho

1. ESPALHAMENTO DE LUZ DINÂMICO.

O Espalhamento de Luz Dinâmico é uma técnica espectroscópica que permite analisar o tamanho, formato, comportamento, flexibilidade e interações de macromoléculas em solução. A técnica consiste na incidência de um feixe monocromático de laser sobre a amostra e os feixes de luz espalhados são detectados por uma placa posicionada em um ângulo respeitando o feixe transmitido. O detector faz medições das flutuações dinâmicas de luz causadas pelo movimento Browniano de partículas em solução. Através da equação de Stokes-Einstein (1) e das medidas acima mencionadas, é calculado o raio hidrodinâmico da partícula, ou seja, é feita uma medição global da molécula envolta a uma camada de solvatação. A partir da medida do raio hidrodinâmico é possível calcular o perfil oligomérico da partícula e sua polidispersidade (Berne e Pecora, 1976).

$$D = \frac{kT}{C\pi\eta r} \quad (1)$$

Na equação 1, k é a constante de Boltzman, C é uma constante relacionada à fricção da partícula com o líquido, r é o raio da partícula difusora, T , a temperatura absoluta e n é a densidade em número de moléculas do fluido

2. DICROÍSMO CIRCULAR.

A espectroscopia de Dicroísmo Circular (*circular dichroism* -CD) é um método muito eficiente para analisar a conformação de proteínas e peptídeos em solução (Greenfield, 1996). A técnica consiste na incidência de uma luz polarizada circular sob uma cubeta contendo a amostra. A luz polarizada interage com grupos cromóforos da proteína ou peptídeo, que não apresentam centro de simetria. O espectro resultante é uma somatória vetorial da quantidade de luz polarizada circular absorvida nos dois sentidos, direita e esquerda (Greenfield, 2000). Os principais grupos cromóforos encontrados em proteínas são as ligações amida do esqueleto peptídico e as cadeias laterais aromáticas, sendo que, tais grupos são encontrados de forma ordenada em algumas regiões, como as folhas- β e α - hélices. Assim, o perfil de absorção de luz nessas conformações é característico (**Figura 27**) e pode ser observado no espectro de dicroísmo circular da proteína (Greenfield, 1996).

Para que os experimentos de dicroísmo circular sejam bem sucedidos, as amostras devem estar livres de contaminantes como outras proteínas, moléculas cromóforas, como nucleotídeos e tampões opticamente ativos, para que não ocorra interferência nas medidas. Além disso, antes das medidas, deve-se injetar um fluxo de nitrogênio no aparelho para diminuir a concentração de oxigênio, o qual pode vir a formar moléculas de ozônio durante as medições, interferindo nas medidas.

O comprimento de onda utilizado para gerar o espectro de CD pode ser dividido em três faixas: abaixo de 250 nm onde as contribuições do esqueleto peptídico são dominantes; entre 250 e 300 nm, onde a maior contribuição para o espectro provém das cadeias laterais aromáticas e entre 300 e 700 nm onde os cromóforos extrínsecos contribuem (Sreerama e Woody, 2004).

A definição de CD é dada pela equação 2, onde ϵ_L e ϵ_R são os coeficientes de extinção da luz polarizada a direita e a esquerda, respectivamente e l o comprimento de onda. O cálculo do espectro de CD, para um dado comprimento de onda λ é feito através de algoritmos que assumem a faixa Gaussiana para todas as transições da luz e utilizam a relação entre CD molar $\Delta\epsilon_k$, e força de rotação, R_k para uma dada transição k utilizam a seguinte equação: $\Delta\epsilon_k = 2.278 R_k \lambda^2 / \Delta k$ (Sreerama e Woody, 2004).

$$\Delta\epsilon(\lambda) = \epsilon_L(\lambda) - \epsilon_R(\lambda) \quad (2)$$

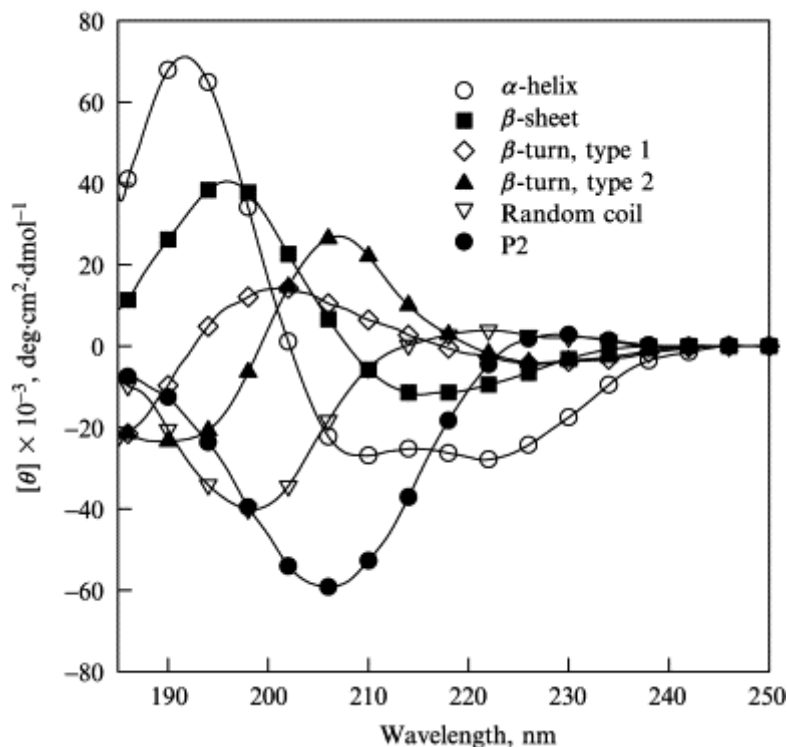


Figura 27. Representação das curvas de Dicroísmo circular correspondentes a diferentes elementos de estrutura secundária (Greenfield, 2000).

3. CRISTALIZAÇÃO DE PROTEÍNAS

A cristalização de proteínas consiste, essencialmente, em criar um microambiente favorável para o empacotamento e conseqüente formação e crescimento de cristais de uma determinada proteína. Trata-se de uma área da ciência em constante desenvolvimento, em que, através de um processo de tentativa e erro, tenta-se precipitar a proteína em sua solução (Drenth, 1994).

O princípio da cristalização é a organização das moléculas de forma tridimensional, unidas por interações não-covalentes. Tal arranjo de moléculas pode ser dividido em partes idênticas denominadas células unitárias. A célula unitária é a menor e mais simples unidade do cristal, representante de todo o restante do arranjo cristalino. Pode-se imaginar o cristal como um eficiente conjunto de várias células unitárias, onde estas podem ser vistas como dentro de “caixas”, unidas por suas linhas e vértices. Portanto, conhecendo a organização das moléculas dentro de uma célula unitária, pode-se prever o arranjo de todo o restante do cristal. Submetendo esse cristal a difração de raio-x, obtém-se uma imagem da nuvem eletrônica que

circunda as moléculas no meio da célula unitária, permitindo localizar todos os átomos da célula.

O requisito básico para a cristalização de uma proteína é a pureza da amostra, pois nenhum outro componente ou proteína estranha deve estar presente, para que não ocorra interação ou formação de complexo com a proteína de interesse, o que pode prejudicar sua cristalização. Se esta não está extremamente pura, outras etapas de purificação devem ser executadas.

Além disso, para que ocorra a formação de cristais, deve-se ter uma solução supersaturada de proteína, a qual tende ao equilíbrio. Nessa situação, a energia livre da proteína diminui através do aumento de interações favoráveis, o que leva a formação do arranjo cristalino. No entanto, existe uma linha tênue entre a formação de cristais e precipitação de amorfos, em solução supersaturada. Para que se encontre um equilíbrio entre essas fases deve-se encontrar a proporção ideal entre a concentração da proteína e a concentração do agente precipitante (**Figura 28**).

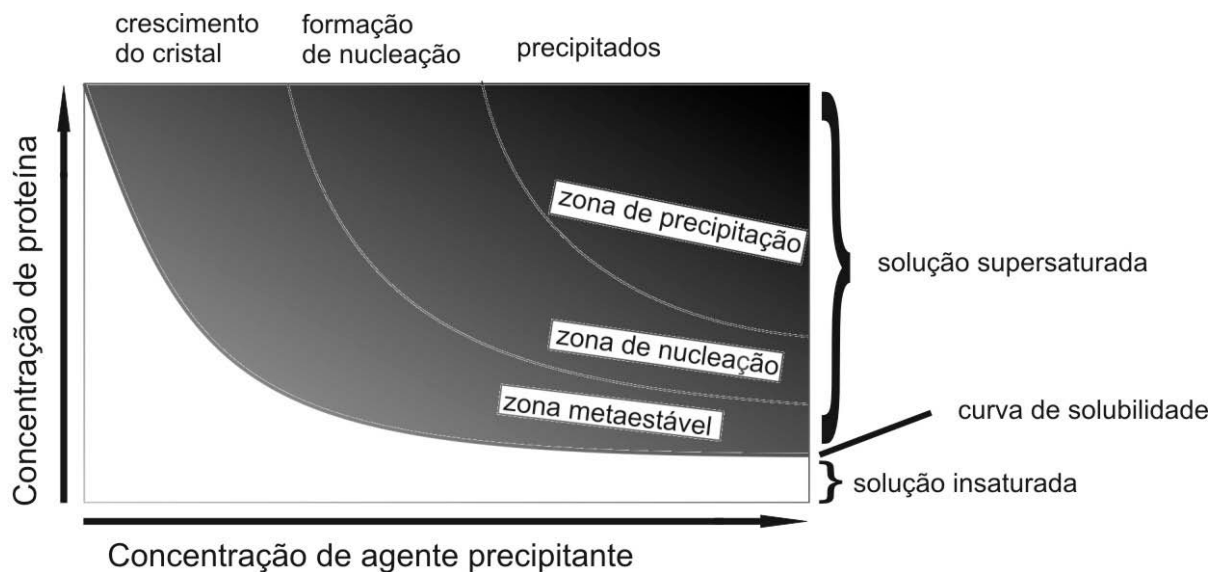


Figura 28. Diagrama de fases do processo de cristalização de proteínas. Através da variação da concentração de proteína e de agente precipitante é possível encontrar o ponto de equilíbrio que permite a formação de cristais.

1.1. PARÂMETROS PARA A CRISTALIZAÇÃO

- Amostra protéica: pureza, concentração, ligantes
- Solução: tipo de precipitante, concentração do precipitante, tipo e concentração do sal, PH, aditivos e outros.

- Ambiente: temperatura, pressão, gravidade.
- Método de cristalização: razão do equilíbrio, volumes.

1.2. TÉCNICAS DE CRISTALIZAÇÃO.

Nos experimentos de cristalização com a Importina- α de *Neurospora crassa*, as técnicas empregadas foram as seguintes:

- Difusão de Vapor (*Vapor Diffusion*)

O método de difusão de vapor é baseado no processo de *salting-out*, onde se prepara uma solução contendo a proteína e o agente precipitante e, gradativamente, aumenta-se a concentração do agente precipitante até a solução atingir o grau de supersaturação, ocorrendo, então, precipitação ou cristalização.

Na difusão de vapor, prepara-se uma gota contendo a proteína e a solução precipitante que fica sobre um poço contendo apenas a solução precipitante. Assim, essas soluções, em um sistema fechado, só podem realizar trocas gasosas, logo, a gota vai perdendo líquido para a solução do poço, saturando-a. Dentro dessa metodologia, as seguintes técnicas podem ser empregadas:

- *Hanging Drop*.

Nessa técnica de cristalização, um sistema fechado e composto por uma gota contendo a amostra de proteína e a solução precipitante, a qual fica suspensa, ancorada por tensão superficial a uma lâminula sobre um poço contendo a solução precipitante. A lâminula é selada sobre o poço, utilizando graxa, para que não ocorram trocas com o meio (**Figura 29A**).

- *Sitting Drop*

Se a solução da proteína tem uma baixa tensão superficial, recomenda-se a técnica de *sitting drop*, onde a gota fica “assentada” sobre um apoio, no centro da solução precipitante (**Figura 29B**).

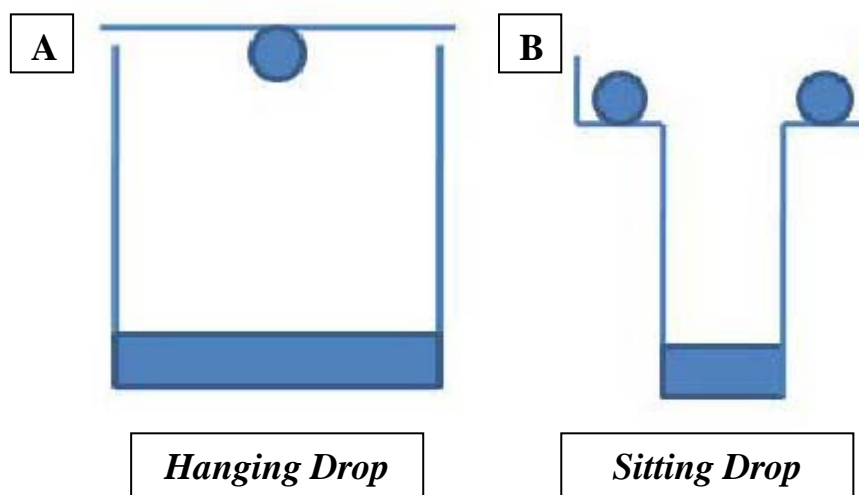


Figura 29. Ilustração das técnicas de Hanging Drop (A) e Sitting Drop (B) (Cosmobrand).

- *Microseeding*

A técnica de *microseeding* consiste na aplicação de fragmentos de cristais, obtidos em experimentos anteriores, na gota da proteína que se deseja cristalizar, para que esses fragmentos funcionem como “sementes”. Essas “sementes” aceleram a etapa de nucleação dos cristais e impulsionam o crescimento dos mesmos.

- *Microbath*

A técnica de *microbath* consiste na imersão da gota de cristalização, contendo a proteína e solução precipitante, em óleo.

4. CALORIMETRIA DE TITULAÇÃO ISOTÉRMICA

A compreensão das interações que ocorrem entre as macromoléculas biológicas, sejam associações entre si ou de uma macromolécula com ligantes, é parte essencial dos estudos de macromoléculas a nível molecular. As abordagens existentes para o estudo dessas interações podem ser agrupadas em dois principais grupos: medição direta da concentração de moléculas livres ou ligadas ou a utilização de técnicas que consideram as mudanças físicas do ambiente

onde ocorre a interação, as quais são proporcionais a extensão da saturação da ligação (Freire *et. al.*, 1990).

A Calorimetria de Titulação Isotérmica faz parte da segunda abordagem, pois mede a variação de energia, na forma de calor, de reações bioquímicas ou interações moleculares. Devido a característica do ITC de medir o calor de interação, é possível obter uma medida direta das constantes termodinâmicas envolvidas na interação, através da equação 3 (Freire *et. al.*, 1990; Lewis e Murphy, 2005).

$$\Delta G = \Delta H - T\Delta S \quad (3)$$

A técnica consiste em titular, em intervalos regulares previamente estabelecidos, amostras do ligante em uma célula contendo a amostra da macromolécula em solução, com concentração conhecida. A variação de temperatura dessa célula, a cada injeção do ligante, é detectada pela célula de referência que promove a compensação térmica da célula de amostra, através do aumento ou diminuição da potência fornecida, para que ambas as células fiquem em equilíbrio ($\Delta T=0$). Os sinais de compensação térmica são traduzidos em um termograma no decorrer da titulação. A integração desses dados de fluxo de calor em função do tempo resultam na variação do calor total por injeção. Considerando a razão molar entre ligante e macromolécula, calcula-se as constantes termodinâmicas envolvidas na interação em estudo.

Apêndice B

Artigo publicado na revista *Acta Crystallographica Section F*

Crystallization and preliminary X-ray crystallographic analysis of importin- α from *Neurospora crassa*

Natalia E. Bernardes,^a
Agnes A. S. Takeda,^a Fernanda Z.
Freitas,^b Maria Célia Bertolini^b
and Marcos R. M. Fontes^{a*}

^aDepartamento de Física e Biofísica, Instituto de Biociências, UNESP – Universidade Estadual Paulista, Botucatu, SP, Brazil, and

^bDepartamento de Bioquímica e Tecnologia Química, Instituto de Química, UNESP – Universidade Estadual Paulista, Araraquara, SP, Brazil

Correspondence e-mail: fontes@ibb.unesp.br

Received 30 September 2013

Accepted 5 March 2014

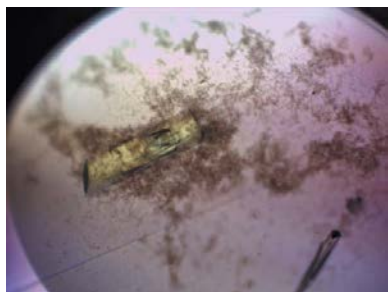
Importin- α recognizes cargo proteins that contain classical nuclear localization sequences (NLS) and, in complex with importin- β , is able to translocate nuclear proteins through the nuclear pore complex. The filamentous fungus *Neurospora crassa* is a well studied organism that has been widely used as a model organism for fundamental aspects of eukaryotic biology, and is important for understanding the specific mechanisms of protein transport to the cell nucleus. In this work, the crystallization and preliminary X-ray diffraction analysis of importin- α from *N. crassa* (IMP α -Nc) complexed with a classical NLS peptide (SV40 NLS) are reported. IMP α -Nc–SV40 NLS crystals diffracted X-rays to 2.0 Å resolution and the structure was solved by molecular-replacement techniques, leading to a monomeric structure. The observation of the electron-density map indicated the presence of SV40 NLSs interacting at both the minor and major NLS-binding sites of the protein.

1. Introduction

The nuclear envelope present in eukaryotic cells allows the separation of the genetic material and the transcriptional machinery from the translational and metabolic processes in the cytoplasm. Nuclear proteins are synthesized in the cytoplasm and need to be transported into the nucleus through the nuclear pore complexes (NPCs). This translocation process usually occurs by active transport through factors termed karyopherins (importins and exportins; Görlich & Kutay, 1999). A pathway involved in this process is the 'classical' nuclear transport pathway which depends on an importin- α –importin- β heterodimer and requires a positively charged sequence within the cargo protein, known as a nuclear localization sequence (NLS; Lange *et al.*, 2007). Importin- α (IMP α) recognizes the cargo proteins that contain an NLS and importin- β (IMP β) is responsible for transport of the importin–cargo protein complex through the NPC.

Classical NLSs contain one (monopartite) or two (bipartite) clusters of positively charged amino acids, which are recognized by the IMP α . This protein receptor has two NLS-binding sites formed by conserved residues, known as the major and minor NLS-binding sites. The N- and C-terminal clusters of a bipartite NLS interact with the minor and major NLS-binding sites, while a monopartite NLS interacts only with the major binding site. The crystal structure of full-length IMP α showed that the major NLS-binding site is occupied by residues from its N-terminal region (IMP β -binding domain) that resembles an NLS (Kobe, 1999); thus, IMP α is auto-inhibited in the absence of IMP β . Therefore, N-terminally truncated IMP α has been used as a model for an IMP α –IMP β heterodimer in several crystal structures, since the truncated protein may simulate an IMP α –IMP β heterodimer (Rexach & Blobel, 1995; Conti & Kuriyan, 2000; Fontes *et al.*, 2000).

Crystal structures of IMP α from *Saccharomyces cerevisiae* (Conti *et al.*, 1998), *Oryza sativa* (Chang *et al.*, 2012), *Mus musculus* (Kobe, 1999) and *Homo sapiens* (Dias *et al.*, 2009) have been solved, showing they are built from repetitive sequences (armadillo repeats) and belong structurally to the class of proteins termed solenoid proteins (Kobe & Kajava, 2011), having sequence identities in the range 41–64%. The structural differences among IMP α from different organisms, although small, have fundamental importance for the specificity



in the transport of cargo proteins. The presence of different IMP α isoforms that are involved in the transport of different proteins highlights the importance of specificity in nuclear transport.

Neurospora crassa has been well studied in terms of classical and molecular genetics, and together with the yeasts *S. cerevisiae* and *Schizosaccharomyces pombe* has been much used as a model organism for fundamental aspects of eukaryotic biology. Nevertheless, compared with yeasts, *N. crassa* has developed greater morphological and developmental complexity as a consequence of its larger number of genes. Comprehension of the nuclear import mechanism in fungal cells could be an important tool for understanding their metabolism and growth (Stinnett *et al.*, 2007; Cupertino *et al.*, 2012), enabling the future applications of fungi in the development of specific drugs and treatments for infections and for applications in the biotechnological industry.

In order to obtain structural and functional information on the nuclear transport process in *N. crassa*, we recently cloned the cDNA encoding the N-terminally truncated importin- α from *N. crassa* (IMP α -Nc), produced the recombinant protein in *Escherichia coli* and performed biophysical experiments and bioinformatics studies (Takeda *et al.*, 2013). On the basis of the theoretical model proposed in this previous study and a sequence alignment among the IMP α s from different organisms, it was suggested that the presence of nonconserved amino acids relatively close to the NLS-binding region may be related to the binding specificity of the IMP α -Nc to cargo proteins (Takeda *et al.*, 2013).

In the present work, we report the crystallization experiments, X-ray diffraction data collection at 2 Å resolution and molecular-replacement solution of the N-terminally truncated importin- α from *N. crassa* bound to a classic NLS peptide (simian virus 40 large tumour antigen NLS peptide; SV40 NLS), aiming to obtain structural insights into the specific mechanisms of protein transport to the cell nucleus in *N. crassa*.

2. Materials and methods

2.1. Protein production and purification

The recombinant N-terminally truncated IMP α -Nc (75–549) was produced in LB medium as a hexa-His-fusion protein using the *E. coli* host strain Rosetta (DE3) pLysS (Novagen) as described previously (Takeda *et al.*, 2013). The protein was purified on a nickel-affinity column using an ÄKTAprime purification system (GE Healthcare) and eluted in a 0.15–3.0 M imidazole linear gradient, followed by

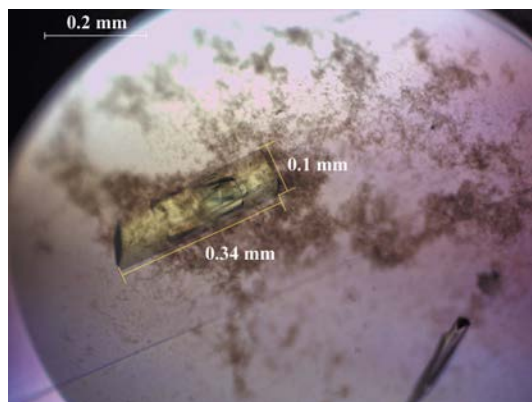


Figure 1
Crystal of the IMP α -Nc-SV40 NLS peptide complex.

dialysis to remove imidazole and to change the buffer. Samples were concentrated to 12 mg ml⁻¹ and stored in a buffer composed of 20 mM Tris-HCl pH 8.0, 100 mM NaCl at -20°C. The purity was estimated to be 98% by SDS-PAGE.

2.2. Synthesis of the NLS peptide

The peptide corresponding to the simian virus 40 large tumour antigen NLS peptide (SV40 NLS; ¹²⁶PKKKRKV¹³²) was synthesized by Proteimax (Brazil) with purity higher than 95%. The peptide contains N- and C-terminal residues additional to the minimal identified NLSs to avoid artifactual binding at the termini (Fontes *et al.*, 2000).

2.3. Crystallization experiments

Crystallization experiments were performed using a protein concentration of 12 mg ml⁻¹ and a peptide:protein molar ratio of 8 for SV40 NLS peptide. Initial attempts to crystallize native and NLS-complexed IMP α -Nc were conducted with similar crystallization conditions used for IMP α from different organisms (Fontes *et al.*, 2000), but no crystal hits were obtained. The sparse-matrix method (Jancarik & Kim, 1991) was then used to screen crystallization conditions using several commercial kits, including Crystal Screen, Crystal Screen 2 and PEG/Ion from Hampton Research, The JSCG Suite from Qiagen and the PACT-HTS crystal screen from Jena Bioscience, using both native and NLS-complexed IMP α -Nc. The experiments were performed using sitting-drop vapour diffusion (McPherson, 1982), with MRC 2 Well Crystallization Plates (Swissci) and the Oryx4 automated crystallization system from Douglas Instruments.

No crystals were obtained for native IMP α -Nc, but protein crystals for IMP α -Nc with NLS SV40 peptide were found in two different conditions: (i) 0.2 mM sodium phosphate dibasic dihydrate, 20% (w/v) polyethylene glycol 3350 and (ii) 0.2 mM Bicine pH 8.5, 20% (w/v)

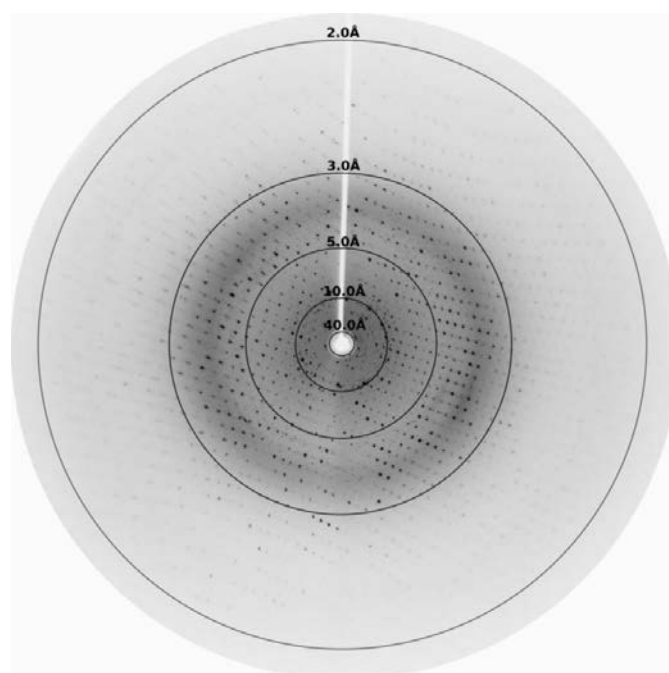


Figure 2
Diffraction pattern image of the IMP α -Nc-SV40 NLS peptide complex.

Table 1

X-ray diffraction data-collection and processing statistics.

Values in parentheses are for the highest resolution shell. Data were processed using the *HKL-2000* suite (Otwinowski & Minor, 1997).

Unit-cell parameters (Å)	$a = 44.7, b = 63.6, c = 180.4$
Space group	$P2_12_12_1$
Resolution (Å)	40.0–2.00 (2.07–2.00)
Unique reflections	32551
Completeness (%)	91.4 (89.8)
$R_{\text{merge}}^{\dagger}$ (%)	9.8 (9.23)
Radiation source	D03B-MX1 station, LNLS
Data-collection temperature (°C)	–173
Average $I/\sigma(I)$	14.7 (4.7)
Multiplicity	10.4 (9.9)
Mosaicity	0.71
Matthews coefficient V_M (Å ³ Da ^{–1})	2.42
Molecules in the asymmetric unit	1
Solvent content (%)	49.1

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual measurement of the reflection with Miller indices hkl and $\langle I(hkl) \rangle$ is the mean intensity of that reflection. Calculated using reflections with $I > -3\sigma(I)$.

polyethylene glycol 6000 [conditions from the PEG/Ion crystallization kit (Hampton Research) and The JSCG Suite (Qiagen), respectively]. Despite the crystals from these two conditions showing a diffraction pattern when submitted to X-ray diffraction experiments, crystals from condition (ii) only diffracted to low resolution. After optimization of condition (i), the best crystals were obtained using reservoir solution (50 μ l) consisting of 0.15 mM sodium phosphate dibasic dihydrate, 20% (w/v) polyethylene glycol 3350 at 277 K after approximately 1 week. The crystallization drops were composed of a mixture of 1 μ l protein solution, 0.5 μ l peptide solution and 0.5 μ l reservoir solution. Fig. 1 shows a typical crystal obtained from this condition.

2.4. X-ray data collection and processing

X-ray diffraction data were collected from a single IMP α -Nc–SV40 NLS crystal at a wavelength of 1.58 Å using a synchrotron-radiation source (D03B-MX1 beamline, Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brazil) and a MAR CCD image-plate detector (MAR Research). A crystal was mounted in a nylon loop and flash-cooled in a stream of nitrogen at –173°C using 10% (v/v) polyethylene glycol 400 as a cryoprotectant. The crystal-to-detector distance was 70 mm and an oscillation range of 1.5° was used, resulting in the collection of a total of 241 images (Fig. 2). The data were processed using the *HKL-2000* suite (Otwinowski & Minor, 1997).

3. Results and discussion

The data-collection statistics are shown in Table 1. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 44.7, b = 63.6, c = 180.4$ Å. They are not isomorphous to the *M. musculus* importin- α –SV40 NLS complex (Fontes *et al.*, 2000) or to other *M. musculus* importin- α –monopartite NLS complexes (Takeda *et al.*, 2011). The diffraction images showed a pattern probably related to a major crystal surrounded by satellite crystals (Fig. 2); therefore, a wide-angle data set (360°) was collected in order to increase the completeness. The data set is 91.4% complete at 2.0 Å resolution, with an R_{merge} of 9.8%. Calculations based on the protein molecular weight indicate the presence of one molecule in the asymmetric unit. This corresponds to a Matthews coefficient (Matthews, 1968) V_M of 2.42 Å³ Da^{–1} with a calculated solvent content of 49.1%. These values are within the range for typical

protein crystals, assuming a value of 0.74 cm³ g^{–1} for the protein partial specific volume.

The crystal structure of IMP α -Nc–SV40 NLS was determined using molecular-replacement techniques implemented in *Phaser* (McCoy *et al.*, 2007) using as a model the coordinates of importin- α from *M. musculus* complexed to nucleoplasmin NLS (PDB entry 3ull, chain B; Marfori *et al.*, 2011) as the search model, which confirmed the presence of a monomer in the asymmetric unit. Inspection of the electron-density map indicated the presence of the SV40 NLS peptide at both minor and major NLS sites of the protein. The monomeric conformation obtained in the IMP α -Nc–SV40 NLS complex (this study) is in agreement with previous results for IMP α -Nc bound to NLS peptides (Takeda *et al.*, 2013) using analytical size-exclusion chromatography experiments, electrophoresis and dynamic light-scattering experiments. The monomeric conformation is a common characteristic of several N-terminally truncated importins- α from several organisms (Kobe, 1999; Dias *et al.*, 2009; Chang *et al.*, 2012). N-terminally truncated yeast importin- α is an exception, in which stabilization of the dimeric structure is achieved by the formation of an extensive dimeric interface which includes the linker-binding region of bipartite NLSs (Conti *et al.*, 1998).

In the present study, we also observed that it was impossible to crystallize N-terminally truncated IMP α -Nc without a ligand. This result is in agreement with our previous study (Takeda *et al.*, 2013), in which we demonstrated that N-terminally truncated unbound IMP α -Nc presented lower stability and has the tendency to form aggregates. This protein behaviour was also observed for N-terminally truncated importin- α from *Xenopus laevis* (Falces *et al.*, 2010). Indeed, attempts to crystallize importin- α from *M. musculus* without any ligand (Takeda, 2009; Mynott, 2009) were also unsuccessful.

Finally, despite the conservation of key residues among importins- α from different organisms (Takeda *et al.*, 2013), the presence of nonconserved amino acids close to the NLS-binding site may influence the binding of IMP α -Nc to specific cargo proteins. The crystal structure of IMP α -Nc may add important elements to the role of importin- α in the protein transport to the nucleus.

This work was financially supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) providing research grants and fellowships.

References

- Chang, C.-W., Couñago, R. L. M., Williams, S. J., Bodén, M. & Kobe, B. (2012). *Plant Cell*, **24**, 5074–5088.
- Conti, E. & Kuriyan, J. (2000). *Fold. Des.* **8**, 329–338.
- Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. (1998). *Cell*, **94**, 193–204.
- Cupertino, F. B., Freitas, F. Z., de Paula, R. M. & Bertolini, M. C. (2012). *PLoS One*, **7**, e44258.
- Dias, S. M., Wilson, K. F., Rojas, K. S., Ambrosio, A. L. & Cerione, R. A. (2009). *Nature Struct. Mol. Biol.* **16**, 930–937.
- Falces, J., Arregi, I., Konarev, P. V., Urbaneja, M. A., Svergun, D. I., Taneva, S. G. & Bañuelos, S. (2010). *Biochemistry*, **49**, 9756–9769.
- Fontes, M. R., Teh, T. & Kobe, B. (2000). *J. Mol. Biol.* **297**, 1183–1194.
- Görlich, D. & Kutay, U. (1999). *Annu. Rev. Cell Dev. Biol.* **15**, 607–660.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Kobe, B. (1999). *Nature Struct. Biol.* **6**, 388–397.
- Kobe, B. & Kajava, A. V. (2011). *Curr. Opin. Struct. Biol.* **11**, 725–732.
- Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E. & Corbett, A. H. (2007). *J. Biol. Chem.* **282**, 5101–5105.
- Marfori, M., Mynott, A., Ellis, J. J., Mehdi, A. M., Saunders, N. F., Curmi, P. M., Forwood, J. K., Bodén, M. & Kobe, B. (2011). *Biochim. Biophys. Acta*, **1813**, 1562–1577.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*. New York: Wiley.
- Mynott, A. V. (2009). PhD thesis, University of New South Wales, Australia.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rexach, M. & Blobel, G. (1995). *Cell*, **83**, 683–692.
- Stinnett, S. M., Espeso, E. A., Cobeño, I., Araújo-Bazán, L. & Calvo, A. M. (2007). *Mol. Microbiol.* **63**, 242–255.
- Takeda, A. A. S. (2009). PhD thesis, Universidade Estadual Paulista Júlio De Mesquita Filho – UNESP, Brazil.
- Takeda, A. A., de Barros, A. C., Chang, C.-W., Kobe, B. & Fontes, M. R. (2011). *J. Mol. Biol.* **412**, 226–234.
- Takeda, A. A. S., Freitas, F. Z., Magro, A. J., Bernardes, N. E., Fernandes, C. A. H., Gonçalves, R. D., Bertolini, M. C. & Fontes, M. R. M. (2013). *Protein Pept. Lett.* **20**, 8–16.

Apêndice C

Artigo publicado na revista *Protein and Peptides Letters*

Biophysical Characterization of the Recombinant Importin- α from *Neurospora crassa*

Agnes A. S. Takeda¹, Fernanda Z. Freitas², Angelo J. Magro¹, Natalia E. Bernardes¹, Carlos A. H. Fernandes¹, Rodrigo D. Gonçalves², Maria Célia Bertolini^{2,#} and Marcos R. M. Fontes^{1,*}

¹Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, SP, 18618-970, Brazil; ²Departamento de Bioquímica e Tecnologia Química, Instituto de Química, Universidade Estadual Paulista, Araraquara, SP, Brazil

Abstract: *Neurospora crassa* has been widely used as a model organism and contributed to the development of biochemistry and molecular biology by allowing the identification of many metabolic pathways and mechanisms responsible for gene regulation. Nuclear proteins are synthesized in the cytoplasm and need to be translocated to the nucleus to exert their functions which the importin- α receptor has a key role for the classical nuclear import pathway. In an attempt to get structural information of the nuclear transport process in *N. crassa*, we present herein the cloning, expression, purification and structural studies with N-terminally truncated IMP α from *N. crassa* (IMP α -Nc). Circular dichroism analysis revealed that the IMP α -Nc obtained is correctly folded and presents a high structural conservation compared to other importins- α . Dynamic light scattering, analytical size-exclusion chromatography experiments and molecular dynamics simulations indicated that the IMP α -Nc unbound to any ligand may present low stability in solution. The IMP α -Nc theoretical model displayed high similarity of its inner concave surface, which binds the cargo proteins containing the nuclear localization sequences, among IMP α from different species. However, the presence of non-conserved amino acids relatively close to the NLS binding region may influence the binding specificity of IMP α -Nc to cargo proteins.

Keywords: Biophysical characterization, classical nuclear import pathway, heterologous expression, homology modeling, importin- α , *Neurospora crassa*.

1. INTRODUCTION

The filamentous fungus *Neurospora crassa* is a well-studied organism, and has been widely used as a model organism for fundamental aspects of eukaryotic biology. These studies contributed for the development of biochemistry and molecular biology, allowing the identification of several metabolic pathways and mechanisms responsible for gene regulation. Proteins presenting activities in the nucleus such as transcription factors are synthesized in the cytoplasm and must to be translocated to the nuclear membrane through the nuclear pore complex (NPC) associated with protein transporters. One pathway involved in this process is the classical nuclear transport pathway, which depends on importin- α /importin- β heterodimer.

Importin- α (IMP α) recognizes the cargo proteins that contain a nuclear localization sequence (NLS) and, importin- β (IMP β) is responsible for translocation of the importin/cargo protein complex through the NPC. Classical NLSs (cNLSs) contain one or two clusters of positively charged amino acids, and are therefore often divided in monopartite

(containing a single cluster of basic amino acids), and bipartite cNLSs (containing two clusters of basic amino acids) [1]. Structural studies have shown that both classes are recognized by the receptor IMP α . This nuclear protein receptor has two NLS binding sites formed by conserved residues in its armadillo repeat-domain (ARM), the major and minor NLS-binding sites. The N- and C-terminal clusters of a bipartite NLS interact with the minor and major NLS-binding sites while a monopartite NLS interacts only with the major binding site [2,3,4]. The structure of full-length IMP α showed that the major NLS-binding site is occupied by residues 44-54 from its N-terminal region (IMP β binding domain) that resembles an NLS [5]; IMP α is therefore autoinhibited in the absence of IMP β . Then, N-terminally truncated IMP α has been used as a model for IMP α /IMP β heterodimer since the truncated may simulate an IMP α /IMP β heterodimer in several crystal structures [2,3,4]. Crystal structures of IMP α from *Saccharomyces cerevisiae* [6], *Mus musculus* [5] and *Homo sapiens* [7] have been reported in the literature. These structures present similar fold; include the 8 ARM repeat-domain, having sequence identities in the range of 41-61%. The structural differences among IMP α from these organisms, although small, should have fundamental importance for the specificity in the transport of cargo proteins to the cellular nuclei. It is noteworthy that there are different IMP α isoforms in the same organism, which are involved in the transport of different proteins, highlighting the importance of specificity in the nuclear transport.

[#]Address correspondence to this author at the Departamento de Bioquímica e Tecnologia Química, Instituto de Química, Universidade Estadual Paulista, Araraquara, SP, Brazil; Tel: 55-16-33019675; Fax: 55-16-33019692; E-mail: mcbertol@iq.unesp.br

*Address correspondence to this author at the Departamento de Física e Biofísica, Instituto de Biociências, UNESP, C. P. 510, CEP 18618-000, Botucatu-SP, Brazil; Tel: +55.14.38800271; Fax: +55.14.38153744; E-mail: fontes@ibb.unesp.br

In order to get structural and functional information of the nuclear transport process in *N. crassa*, we cloned the cDNA encoding the N-terminally truncated importin- α from *N. crassa* (IMP α -Nc), produced the recombinant protein in *Escherichia coli*, performed biophysical experiments (analytical size-exclusion chromatography, circular dichroism and dynamic light scattering) and bioinformatics studies. The results obtained in this work may add important elements to understand the nuclear transport process in the model microorganism *N. crassa*.

2. MATERIAL AND METHODS

2.1. Cloning of the cDNA Encoding the IMP α from *N. Crassa*

Conidia (2×10^8 cells/mL) from the *N. crassa* strain FGSC 9718 were cultivated in VM liquid medium [8] supplemented with sucrose 2%, at 30 °C, 250 rpm during 24 h. The mycelium was harvested and total RNA was prepared [9]. To synthesize the cDNA first-strand, 600 ng of total RNA were treated with RQ1 RNase-Free DNase (Promega) and reverse transcribed using the oligo (dT)₂₀ reverse primer (Invitrogen) and the commercial kit *Superscript*® III *First-strand Synthesis for RT-PCR* (Invitrogen), according to the manufacturer's protocols. The 1,425 bp cDNA fragment encoding the IMP α was amplified by PCR using 2.0 μ L of total cDNA, 1U PhusionTM DNA polymerase (Finnzymes), 10 mM dNTPs, 1x PhusionTM GC buffer, 1.5 μ L of DMSO, and 100 pmoles/ μ L of each specific oligonucleotides IMP α Δ 74-F (5'-CATATGACCGAGTCTCAGTTGAGC GAG-3', *NdeI* site underlined) and IMP α -R (5'-GAATC TTACATGTCCATCGACTCGGTG-3', *EcoRI* site underlined). The oligonucleotides were designed based on the NCU01249 ORF sequence in the genome database at the FGSC site (<http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html>). The oligonucleotide IMP α Δ 74-F was designed to amplify a cDNA encoding a truncated protein, in which the 74 N-terminal amino acid residues were missing. The amplified fragment was cloned into the *EcoRV* site of the pMOS-Blue vector (GE Healthcare) leading to the pMOS-Nc1249 plasmid construction. An *NdeI-EcoRI* DNA cassette was transferred from the pMOS-Nc1249 construction to pET28a vector (Novagen) to generate the pET28-Nc1249 construction. Both plasmid constructions were confirmed by DNA sequencing.

2.2. Production and Purification of Recombinant Importin- α from *N. Crassa* and *M. Musculus*

The recombinant IMP α -Nc protein (75-549) was produced in LB medium as a Hexa-His-fusion protein using the *Escherichia coli* host strain RosettaTM (DE3) pLysS (Novagen). Briefly, the pET28-Nc1249 transformants were cultured at 37 °C and 250 rpm until an OD_{600nm} of about 0.6. The recombinant protein was induced by addition of 0.1 mM IPTG final concentration, during 6 h at 28 °C and 200 rpm. After that, cells were harvested by centrifugation, suspended in lysis buffer (50 mM Hepes, pH 7.0, 150 mM imidazole, 500 mM NaCl, 5% glycerol, 0.5% Nonidet-P40, 10 mM benzamidine, 1 mM PMSF, 5 mM EDTA) and lysed by sonication (5 cycles of 30 sec sonication followed by 30 sec on ice). The cellular extract was clarified by centrifugation

(20,000 xg, 20 min at 4 °C), the supernatant was filtered in a 0.22 μ m pore membrane (Millipore) and applied onto a nickel-affinity column in an Akta Prime Purification System (GE Healthcare). The recombinant protein was eluted in a 0.15-1.5 M imidazole linear gradient, and fractions containing the purified protein were pooled and dialyzed against lyses buffer. The protein was quantified by the Hartree [10] method using BSA as standard.

Hexa-His-tagged truncated *Mus musculus* importin- α isoform 2 (IMP α -Mm), corresponding to amino acids 70-529 was expressed in *E. coli* and purified by nickel affinity chromatography, as described previously [11]. The protein was eluted using a 0.15-1.5 M linear gradient of imidazole followed by dialysis and the IMP α -Mm sample was stored in buffer 20 mM Tris-HCl pH 8.0, 100 mM NaCl and 10 mM DTT at -20°C.

2.3. Circular Dichroism Spectroscopy

Circular dichroism (CD) experiments were performed on a Jasco J-815 (JASCO Inc., Tokyo, Japan) Spectropolarimeter, equipped with a peltier temperature control PTC-423-S. The IMP α -Nc and IMP α -Mm samples were prepared at 370 μ g.mL⁻¹ in buffer 10 mM Tris-HCl, pH 8.0; 20 mM NaCl. All CD measurements were taken in a spectral range of 190–260 nm at 20 °C, using an optical path-length 0.05 mm with scanning speed 100 nm/min; a band width of 2 nm and response time of 1 s. The final spectra is resultant of 20 spectra that were accumulated, averaged and corrected from the baseline for buffer solution contribution and normalized to residual molar ellipticity [0].

The thermal denaturation analysis was obtained by monitoring the circular ellipticity changes at a fixed wavelength of 222 nm while the sample was heated from 10 to 90°C. Data points were acquired by a ramp rate of 1°C/min and an equilibration time of 5 seconds after each temperature adjustment using an optical path-length 2.0 mm. The denaturation curve was normalized to apparent fraction folded, according to the equation $[\theta] = [(\epsilon_F - \epsilon_U) \alpha] + \epsilon_U$, where $[\theta]$ is the ellipticity at any temperature, ϵ_F is the ellipticity when the protein is fully folded, and ϵ_U is the ellipticity when the protein is totally unfolded, as previously described [12]. The melting temperature (T_M) was calculated by Denatured Protein Analysis program from Spectra ManagerTM II software.

2.4. Dynamic Light Scattering

The dynamic light scattering (DLS) measurements were performed with native N-terminally truncated IMP α -Nc at concentration of 3.5 mg.mL⁻¹ using a DynaPro TITAN instrument (Wyatt Technology). The sample was filtered through a 0.22 μ m pore membrane (Millipore) prior to the measurement. The data were collected at a *fixed angle* of 90°, temperature of 4°C and measured one hundred times. All results were analyzed with the Dynamics v.6.10 software.

2.5. Analytical Size-exclusion Chromatography

Analytical size-exclusion chromatography was performed using a Superdex-200 10/30 column, equilibrated with a buffer containing 20 mM Tris HCl pH 8.0 and 100 mM

NaCl, attached to an ÄKTA purifier system (GE HealthcareTM). Two samples of the purified IMP α -Nc ($V = 0.1$ ml; $[5.7 \text{ mg}\cdot\text{ml}^{-1}]$) were injected at a flow rate of $0.5 \text{ ml}/\text{min}$: (i) A sample of native IMP α -Nc; (ii) A sample of IMP α -Nc mixed to FEN1 NLS peptide ($^{350}\text{SSAKRKEPEPKGS TKKKAKT}^{369}$) at a proportion of 1:8, incubated for 20 minutes before chromatography. The molecular weight standards were obtained from a high molecular weight gel filtration calibration kit (Sigma-AldrichTM) containing the following components: Blue dextran (2,000 kDa), Beta-amylase (200 kDa), Alcohol dehydrogenase (150 kDa), Albumin from bovine serum (66 kDa), Carbonic anhydrase (29 kDa) and Cytochrome C (12.4 kDa). The retention volumes for each standard and samples were measured and used to calculate the partition coefficients (K_{av}), which were defined as $K_{av} = (V_r - V_o)/(V_c - V_o)$, where V_r = retention volume, V_o = void volume (calculated based on the retention time of the blue dextran standard), and V_c = geometric bead volume for the column. The coefficient K_{av} obtained for each standard was plotted against the log of the molecular weight to generate a standard curve. The collected fractions were evaluated by 12% SDS-PAGE electrophoresis and western blotting assay using a Monoclonal Anti-polyHistidine antibody produced in mouse (1:1000) (Sigma-Aldrich).

2.6. Molecular Modeling and Dynamics

2.6.1. IMP α -Nc Modeling

The full-length sequence of IMP α -Nc was submitted to HHpred server [13] (<http://toolkit.tuebingen.mpg.de/hhpred>). According to the alignment data obtained in this server, the crystallographic model of complex formed by exportin (CSE1P), importin- α (Kap60p) from *S. cerevisiae* and RanGTP (PDB ID 1wa5_chain B) [14] was selected as the more adequate template for the initial modeling of the theoretical IMP α -Nc structural models. The program Modeller 8v2 [15] and the template Kap60p, truncated at N- and C-termini (residues Glu81 and Asp510, respectively) were used to generate two sets of ten IMP α -Nc structural models: (i) models with two segments corresponding to the region of the IBB domain (segment 1: $^{24}\text{ELRRRR}^{29}$; segment 2: $^{42}\text{EENLAKRRGI}^{51}$) bound to the IMP α -Nc major and minor binding sites (IMP α -Nc-IBB), and (ii) models devoid of these segments (IMP α -Nc- Δ IBB). Variable target function method (VTFM) with conjugate gradients (CG) [16] and molecular dynamics (MD) with simulated annealing (SA) [16] were used in order to refine the models. The best IMP α -Nc-IBB and IMP α -Nc- Δ IBB were selected according to stereochemical and energetic parameters calculated, respectively, with RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) [17] and ProSA-web servers (<https://prosa.services.came.sbg.ac.at/prosa.php>) [18].

2.6.2. Molecular Dynamics Simulations

The best IMP α -Nc-IBB and IMP α -Nc- Δ IBB theoretical models calculated with program Modeller 8v2 were submitted to MD simulations using the program GROMACS (Groningen Machine for Chemical Simulation) v.4.5.3 [19]. All simulations were executed in presence of explicit water molecules [20] using an Ubuntu 9.04 Linux operational system and eight threads of a dual processor and quad-core Intel Xeon E5520 CPU (2.27 GHz) with 24 Gb of RAM. GRO-

MOS 96 53a6 force field [21] was chosen to perform the MD simulations and the protonation states of the charged groups were set to pH 7.0. The minimum distance between any atom of the models and the box wall was 1.0 nm. An energy minimization (EM) using a steepest descent algorithm was performed to generate the starting configuration of the systems. After this step, 200 ps of MD simulation with position restraints applied to the protein (PRMD) were executed in order to relax the systems gently. Then, 20 ns of unrestrained MD simulation were calculated to evaluate the stability of the structures. All MD simulations were carried out in a periodic truncated cubic box under constant temperature (298 K) and pressure (1.0 bar), which were held by coupling to an isotropic pressure and external heat bath [22]. Overall stereochemical and fold quality of the theoretical IMP α structural models obtained after MD simulations were checked with the servers RAMPAGE [17] and ProSA-web [18].

2.6.3. Statistical Analysis

The Brown-Forsythe test was used to analyze the variance deviation of the backbone r.m.s.d. of IMP α -Nc-IBB and IMP α -Nc- Δ IBB models after 10 ns of MD simulation calculated by the program GROMACS v.4.5.3 [19].

3. RESULTS AND DISCUSSION

3.1. Production and Purification of the Recombinant IMP α -Nc

Recombinant IMP α -Nc was produced as a truncated protein missing the 74 N-terminal amino acid residues and fused to a His tag. The truncated protein lacks the amino acid sequence $^1\text{MADRYIPEHRRRTQFKAKSAFKPDELRRRREEQ QVEIRKAKREENLAKRRRGIGAGDSRPGASLGAAPDS-DDENPP}^{74}$, which corresponds to the auto-inhibitory domain. The protein was overexpressed in *E. coli* in a soluble form with a high yield. Single-step purification by immobilized metal affinity chromatography was enough to produce pure protein for all biophysical experiments (Fig. 1).

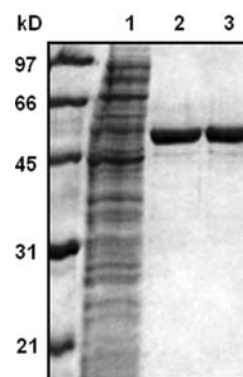


Figure 1. Production in *E. coli* and purification of the recombinant IMP α from *N. crassa*. The recombinant protein production in *E. coli* host strain RosettaTM was performed as described in Materials and Methods. Protein extract containing the recombinant His6-IMP α - Δ IBB was applied onto a HisTrapTM HP column and eluted in a 0.5 M linear imidazole gradient. Fractions containing the purified protein were pooled, dialyzed, and used in further experiments. Lanes 1, flow through (FT); 2 and 3, purified IMP α - Δ IBB after dialysis. The numbers on the left side indicate the protein molecular weight in kDa.

3.2. Circular Dichroism Indicates High Structural Similarity Between Importins- α

The circular dichroism spectrum of the recombinant IMP α -Nc exhibits negative peaks at 208 (-15644.7 [θ]) and 222 nm (-15,438.8 [θ]), which are typical of α -helix rich proteins, confirming the correct fold of the protein obtained (Fig. 2A). This spectrum was compared to that acquired from IMP α -Mm (-15374.1 and -15520.9 for 208 and 222 nm, respectively), for which the structure was elucidated by X-ray crystallography [5]. Both proteins displayed very similar CD spectra confirming that the IMP α -Nc and IMP α -Mm have similar folding. The secondary structure content prediction from the spectrum obtained of IMP α -Nc using CONTINN algorithm [23] resulted in 52.7% α -helix, 16.4% of loops and 24.1% of unordered elements. This result is similar to the values obtained for IMP α -Mm structure, 51.1% α -helix, 16.5% of loops and 25.8% of unordered elements.

To examine the thermal stability of IMP α -Nc, a thermal analysis was performed at 222 nm in the range of 10 – 90°C. Results indicate that the calculated melting point (T_m) for the truncated protein is 39.2°C, which is plausible considering the normal growth temperature for *N. crassa* (30°C) (Fig. 2B). Studies of heat-shock stress (45°C) with *N. crassa* have shown a reduction in glycogen syntase transcript levels (*gsn*) and glycogen levels upon raising the temperature [24]. Proteins involved in the regulation of this metabolic pathway may depend of IMP α -Nc to be transported to the nucleus to develop their function; however for temperatures higher than 40°C, IMP α -Nc is unfolded, impairing the nuclear translocation associated with it.

3.3. Molecular Modeling and Molecular Dynamics Simulations

The initial modeling of the IMP α -Nc-IBB and IMP α -Nc- Δ IBB was based on the template Importin- α (Kap60p) from *S. cerevisiae* (PDB code 1wa5_chain B). This structure pre-

sents a large part of IBB domain modeled in both major and minor NLS binding sites [14]. For the model selection was used the alignment generated by the HHPred server [13] (score = 674.3), the identity (60%) and similarity (98.9%) between the template and IMP α -Nc amino acid sequences [13]. The best IMP α -Nc-IBB and IMP α -Nc- Δ IBB models generated by the program Modeller 8v2 [15] showed a high overall stereochemical quality, with 100% of their amino acid residues distributed in the favorable and allowed regions of their respective Ramachandran plots [14]. Furthermore, the adequate folding of the models was confirmed by the Z-scores: -12.38 and -11.64 for IMP α -Nc-IBB and IMP α -Nc- Δ IBB models, respectively [15]. These models were submitted to molecular dynamics (MD) simulations in order to get potential insights into IMP α features, especially regarding the influence of ligands on the structural behavior of the protein in solution (Fig. 3A). After 20 ns-MD simulations, it was verified that the good stereochemical quality of the IMP α -Nc-IBB and IMP α -Nc- Δ IBB models is kept, since, respectively, 97.9% and 96.7% of their amino acid residues occupied the favorable and allowed Ramachandran plot regions. Moreover, the very similar Z-scores calculated after the MD simulations (-11.10 for IMP α -Nc-IBB and -11.91 for IMP α -Nc- Δ IBB) indicate the conservation of an appropriate folding in both models.

Despite the stabilization of both models after MD simulations, it can be observed by the (Fig. 3A) that IMP α -Nc- Δ IBB model present higher variation of the average backbone r.m.s.d. oscillation in comparison with IMP α -Nc-IBB model. The average backbone r.m.s.d. values calculated for IMP α -Nc-IBB and IMP α -Nc- Δ IBB (0.031 and 0.045 nm, respectively) during the last 10 ns of the MD simulations (Fig. 3A) were significantly different, as pointed out by the Brown-Forsythe test (p-value = $1,4 \cdot 10^{-20}$). Therefore, it is possible to confirm the higher degree of flexibility presented by the IMP α -Nc- Δ IBB model during the progress of the simulation.

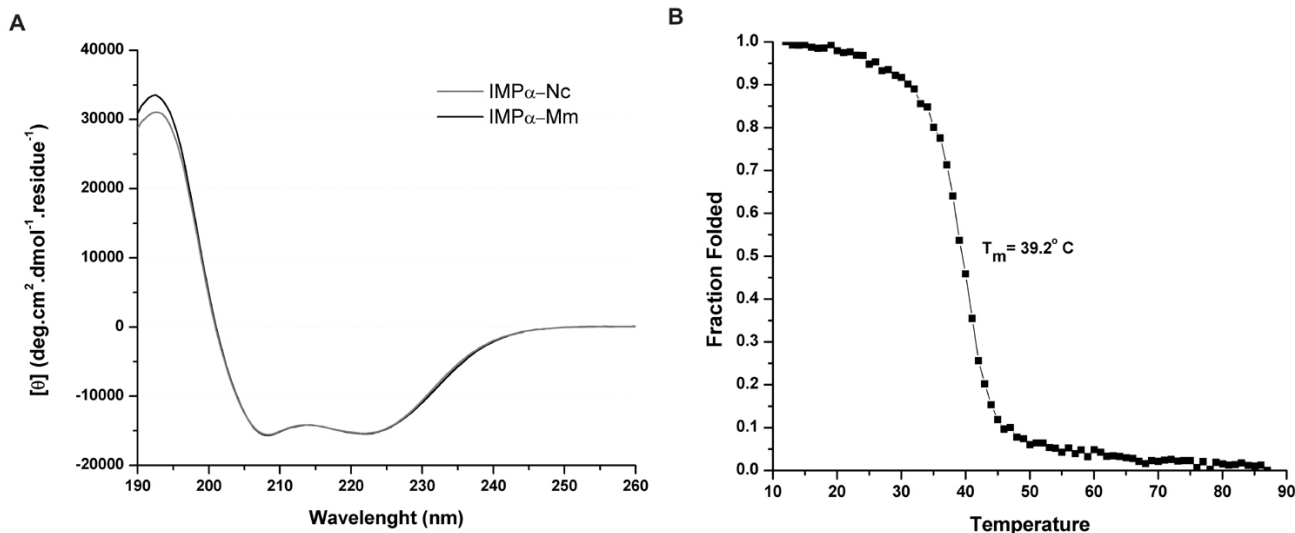


Figure 2. (A) Far-UV circular dichroism spectra of recombinants IMP- α from *N. crassa* (in grey) and *Mus musculus* (in black). (B) Thermal denaturation curve of IMP α from *N. crassa* obtained by circular dichroism spectra at a wavelength of 222 nm while heating from to 10° to 90°C.

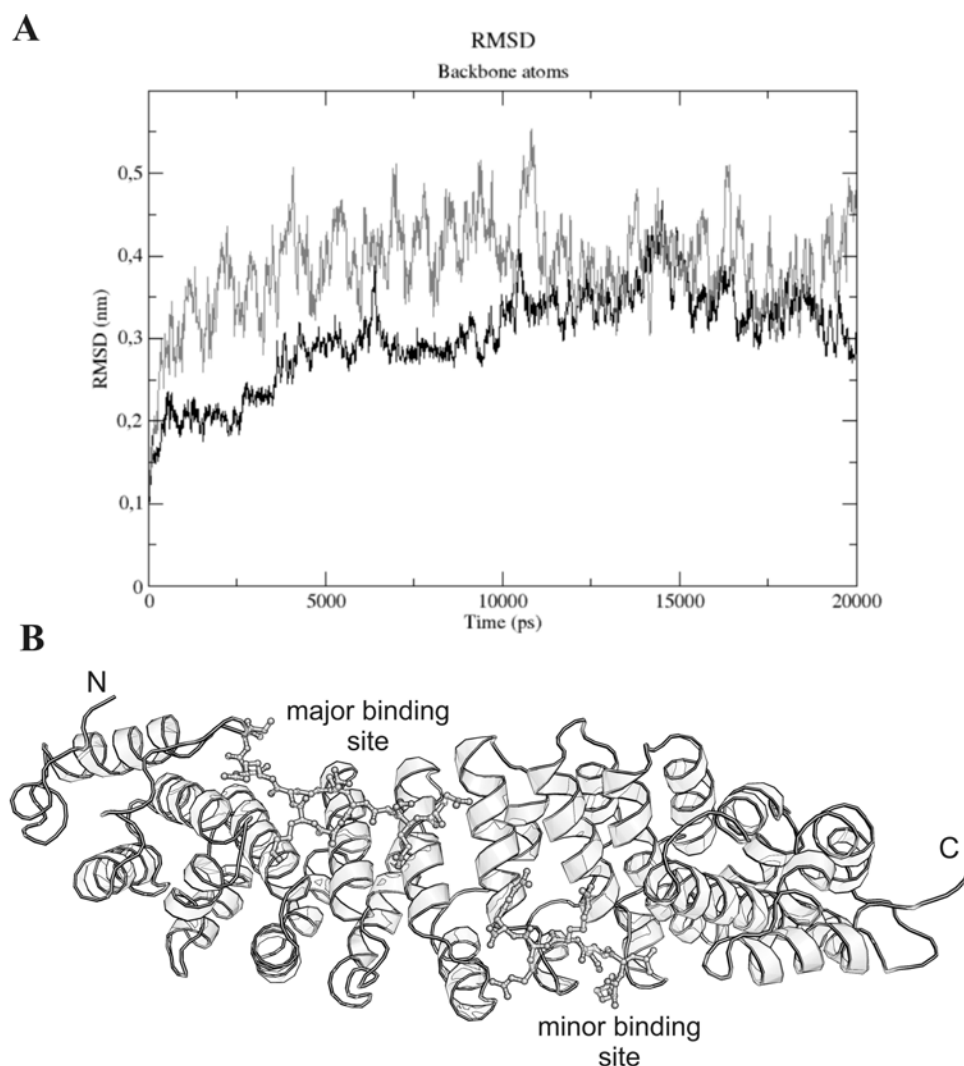


Figure 3. (A) Average backbone r.m.s.d. during the 20 ns-MD simulations of the theoretical models: N-terminally truncated importin- α from *N. crassa* (IMP α -Nc Δ IBB) in light grey and same protein bound to IBB segments (IMP α -Nc-IBB) in grey. Despite of the stabilization after 10 ns observed in both simulations, the model IMP α -Nc- Δ IBB presented a higher oscillation, as pointed out by the Brown-Forsythe test. (B) Ribbons diagram of importin- α theoretical model from *N. crassa* bound to IBB segments (IMP α -Nc-IBB). The IBB segments are shown in a ball-and-stick representation.

3.4. The Theoretical Structure of the IMP α -Nc and the Conservation of Key Residues

The final theoretical model of N-terminally truncated IMP α -Nc (Fig. 3B) is composed by the NLS binding domain and two segments of IBB domain. The structure conserves the elongated shape built from ten armadillo (ARM) repeat each containing three α -helices (H1, H2 and H3) connected by loops as occurs for all other IMP α structures solved to date [1,5,6,7]. In the IMP α -Nc model, the IBB domain is bound to both major and minor NLS-binding sites, which correspond to ARM repeats 2-4 (in red) and 6-8 (in blue), respectively. Additionally, the ARM repeat 5 (in green) is occupied by IBB residues of the region analogous to linker region of bipartite NLS sequences (Fig. 3B).

The (Fig. 4) shows a partial alignment of the sequences IMP α -Nc, IMP α -Mm and IMP α from *S. cerevisiae* (IMP α -Sc) corresponding to the NLS binding domain. This align-

ment shows the identical amino acids highlighted in magenta, whereas in cyan are represented the non-conserved amino acids. These residues are also highlighted in the IMP α -Nc surface model demonstrating that its inner concave surface (defined by the H3 α -helix of ARM repeat motif), is a highly conserved region among the IMP α from different organisms. Similar results were obtained in a recently review published by Marfori and colleagues [25]. These results confirm the conserved tryptophan, asparagines, aspartic and glutamic acids of IMP α -Nc, comprising cores to accommodate the positively charged classic NLS. By contrast, there is a high concentration of non-conserved amino acids (in cyan) relatively closed to the IBB segments (represented as grey sticks in (Fig. 4)). Some of these regions consist in loops connecting α -helices; however others contain organized secondary structures and may influence the binding of IMP α -Nc to specific cargo proteins.

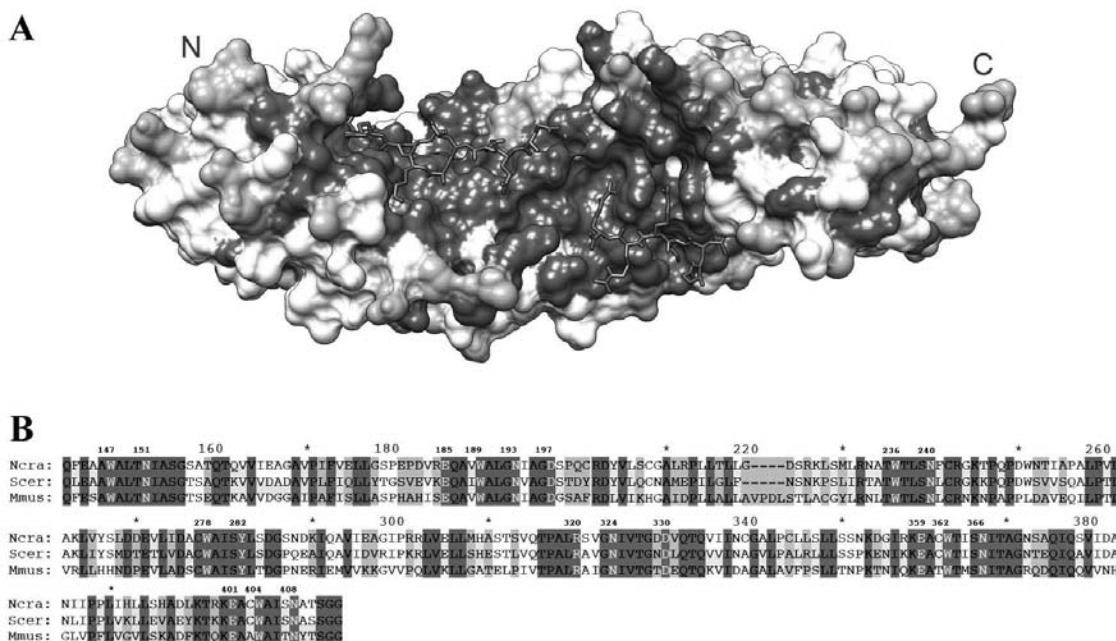


Figure 4. (A) Surface representation of Importin- α theoretical model from *N. crassa*. The IBB segments are shown in grey as stick representation. The conserved amino acids among importin- α from *N. crassa*, *S. cerevisiae* and *Mus musculus* are shown in dark grey and non-conserved ones are shown in light grey. (B) Partial alignment of importin- α from *N. crassa* (Nc), *S. cerevisiae* (Sc) and *M. musculus* (Mm). The conserved amino acids are highlighted in dark grey and the non-conserved in light grey. Some fundamental residues to NLS binding are written in white.

3.5. N-terminally Truncated IMP α -Nc Requires a Ligand at the NLS Binding Region to be Stable

Several crystal structures of IMP α have been solved by X-ray crystallography either in the full-length version, N-terminally truncated and N-terminally truncated bound to several peptides and ligands [2,3,5-7,26-29]. In all of these cases there was a ligand or its own auto-inhibitory sequence (in the case of full-length IMP α) bound at the major or minor binding sites of IMP α . The N-terminally truncated yeast IMP α (IMP α -Sc) is the unique exception, however in this case, the crystal structure presented as homodimer [5], forming an extensive dimeric interface which includes the linker binding region of bipartite NLSs. Co-crystallization experiments of IMP α -Sc and nucleoplasmin only were possible after of the mutation Tyr393Asp, since this is a critical residue to its dimerization process [5]. Attempts to crystallize IMP α -Mm without any ligand by us [30] or by other scientists [31] were unsuccessful, and the reason for that, may be because IMP α -Mm does not have tyrosine residue in the position equivalent to 393 as occur for IMP α -Sc, which may aid its dimerization. These experimental results indicate that it is impossible to crystallize N-terminally truncated IMP α without a ligand (or even other IMP α monomer) in the NLS binding site region. In order to get insights about this finding we performed analytical size-exclusion chromatography experiments with apo IMP α -Nc and IMP α -Nc complexed to FEN1 bipartite NLS peptide [32]. Size-exclusion chromatography indicated the arrangement of two populations of apo IMP α -Nc (peaks 1 and 2) in solution (Fig. 5). In attempt to estimate the molecular weight of these observed peaks, a calibration curve was calculated depending on the retention

volumes of each standard previously applied to the chromatography column. The peak 1 is too broad to infer its molecular weight, once it is comprised by aggregates of IMP α -Nc, whose molecular weights are higher than the recommended separation range of Superdex-200 10/30 column. The peak 2 also corresponds to apo IMP α -Nc as shown by SDS electrophoresis gel and western blot analysis (Fig. 5) and presented an estimated molecular weight of 83.7 kDa. This apparent molecular weight is higher than the expected for N-terminally truncated IMP α -Nc (52 kDa); but this can be explained by the elongated conformation of importins that may lead to reduction of its retention time at column during the chromatography. Dynamic light scattering (DLS) measurements corroborates these results, since that the DLS measures performed at 4°C indicated a mean hydrodynamic radius (R_H) of 4.1 nm with a polydispersity of 20.4 % and an estimated MW of 90 kDa.

By contrast, size-exclusion chromatography of IMP α -Nc complexed to FEN1 NLS peptide resulted in a single peak (peak 3) that corresponds to IMP α -Nc as shown by SDS electrophoresis gel and western blot analysis (Fig. 5). This peak presented an estimated molecular weight of 81.1 kDa, which is slightly smaller than the molecular weight presented by IMP α -Nc at peak 2. This result suggests that the interaction between the IMP α -Nc and the NLS peptide may lead to a protein conformation change and could explain this small difference of retention times between IMP α -Nc samples during the chromatography.

In addition to these data, the analytical size-exclusion chromatography experiments, combined to SDS electrophoresis were also capable to detect aggregation of IMP α -Nc in

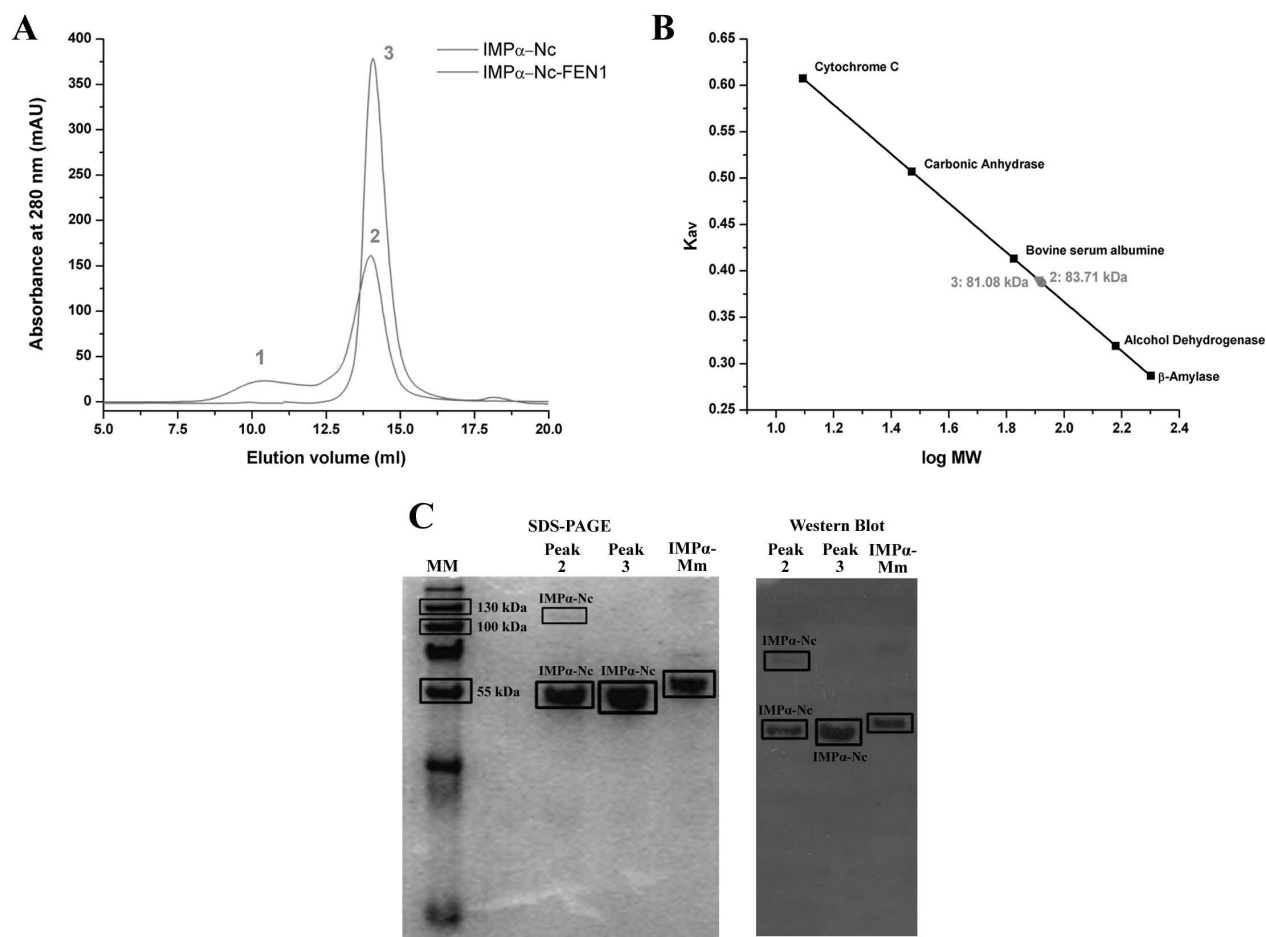


Figure 5. (A) Analytical size-exclusion chromatography of IMP α -Nc in absence and presence of FEN1 NLS peptide in a Superdex-200 10/30 column equilibrated with a buffer containing 20 mM Tris HCl pH 8.0 and 100 mM NaCl. (B) Calibration curve determined using standard proteins shows the estimated molecular weight of the peaks 2 and 3. The peak 1 is broad to infer the molecular weight, once it is comprised by aggregates of IMP α -Nc whose molecular weights are higher than the recommended separation range of Superdex-200 10/30 column. The protein standards were obtained from a high molecular weight gel filtration calibration kit (Sigma-Aldrich) containing: blue dextran (2,000 kDa), beta-amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin from bovine serum (66 kDa), carbonic anhydrase (29 kDa) and cytochrome C (12.4 kDa). (C) Coomassie blue stained 12% SDS-PAGE gel (MM: Page RulerTM Plus Prestained Protein Ladder from Thermo Scientific) and western blot analyses of the peaks 2 and 3 using IMP α -Mm as positive control. In all lanes were applied 15 μ g of pure protein.

absence of NLS peptide even under denaturing conditions. The fractions correspondent to peak 2 were pooled, concentrated and evaluated by SDS electrophoresis experiments. The apo IMP α -Nc sample (peak 2) presented a predominant band of approximately 55kDa and other band around 130 kDa (Fig. 5). In contrast, the sample from peak 3, which corresponds to IMP α -Nc complexed to FEN1 NLS peptide, showed a single band at the same level as the predominant apo IMP α -Nc band (55 kDa) (Fig. 5). The bands observed in the SDS electrophoresis gel were confirmed as IMP α -Nc by western blot analysis showing that even after the gel filtration the IMP α -Nc can form dimers (Fig. 5). Moreover, the presence of the single band from peak 3 indicates that the aggregation of truncated IMP α -Nc was abolished (or significantly reduced) by interaction between IMP α -Nc and FEN1 NLS peptide.

Then, taking into account all the results obtained by us in the MD simulations, size-exclusion chromatography and DLS experiments with the IMP α -Nc and the crystallography studies with several IMP α , we can hypothesize that the IMP α -Nc has lower stability when is unbound than when is bound to a ligand at its NLS binding region. Furthermore, due to the high reactive residues of the minor and major binding sites of IMP α -Nc [1,5,6,7], the N-terminally truncated IMP α -Nc may bind to any flexible segments containing clusters of positively charged amino acids which give stability to the IMP α -Nc. These results are in agreement with the high affinity constants of IMP α to proteins containing NLS basic clusters when IMP α is bound to IMP β [26], since N-terminally truncated IMP α may simulate an IMP α /IMP β heterodimer. These authors also demonstrated using several techniques that IMP α is monomeric in the auto-inhibited

state (when bound to the IBB domain) and stoichiometry of its association with importin- β is 1:1.

CONCLUSIONS

Importin- α from *N. crassa* was cloned, expressed, purified and, studied using biophysical and bioinformatics tools. For the first time, biophysical and bioinformatics studies were performed for importin- α for other organisms than mammalian or yeast. Dynamic light scattering, analytical size-exclusion chromatography experiments and molecular dynamics simulations showed that N-terminally truncated IMP α -Nc presents low stability when it is not bound to a ligand, suggesting thus an explanation to its partial aggregation (or dimerization) in solution. Additionally, the results showed high similarity of inner concave surface, which binds the cargo proteins containing the NLS sequences among the IMP α from different organisms. However, the presence of non-conserved amino acids relatively close to the NLS binding region may be related to the binding specificity of IMP α -Nc to cargo proteins. Additional experiments, including X-ray crystallography, site-directed mutagenesis, and cellular localization may add important elements to understand the fundamental role of IMP α in the protein transport.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

This work was financially supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) providing research grants and fellowships.

ABBREVIATIONS

IMP α	=	Importin- α
IMP β	=	Importin- β
IBB	=	Importin- β binding domain
IMP α -Mm	=	Importin- α from <i>Mus musculus</i>
IMP α -Nc	=	Importin- α from <i>Neurospora crassa</i>
IMP α -Nc-IBB	=	Importin- α from <i>N. crassa</i> bound to IBB segments
IMP α -Nc- Δ IBB	=	N-terminally truncated importin- α from <i>N. crassa</i>
IMP α -Sc	=	Importin- α from <i>Saccharomyces cerevisiae</i>

REFERENCES

[1] Dingwall, C.; Laskey, R.A. Nuclear Targeting Sequences - a Consensus, *Trends Biochem. Sci.*, **1991**, 16(12), 478-481.
 [2] Conti, E.; Kuriyan, J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Fold. Des.*, **2000**, 8(3), 329-338.

[3] Fontes, M.R.; Teh, T.; Kobe, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J. Mol. Biol.*, **2000**, 297(5), 1183-1194.
 [4] Rexach, M.; Blobel, G. Protein Import into Nuclei - Association and Dissociation Reactions Involving Transport Substrate, Transport Factors, and Nucleoporins. *Cell*, **1995**, 83(5), 683-692.
 [5] Conti, E.; Uy, M.; Leighton, L.; Blobel, G.; Kuriyan, J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*, **1998**, 94(2), 193-204.
 [6] Kobe, B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nat. Struct. Biol.*, **1999**, 6(4), 388-397.
 [7] Dias, S.M.; Wilson, K.F.; Rojas, K.S.; Ambrosio, A.L.; Cerione, R.A. The molecular basis for the regulation of the cap-binding complex by the importins. *Nat. Struct. Mol. Biol.*, **2009**, 16(9), 930-937.
 [8] Vogel, H. A convenient growth medium for *Neurospora* (medium N). *Microb. Genet. Bull.*, **1956**, 13, 2.
 [9] Sokolovsky, V.; Kaldenhoff, R.; Ricci, M.; Russo, V. E. A. Fast and reliable mini-prep RNA extraction from *Neurospora crassa*. *Fungal Genet. Newsl.*, **1995**, 37, 3.
 [10] Hartree, E.F. Determination of protein: a modification of the Lowry method that gives a linear photometric response. *Anal. Biochem.*, **1972**, 48(2), 422-427.
 [11] Fontes, M.R.; Teh, T.; Jans, D.; Brinkworth, R.I.; Kobe, B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *J. Biol. Chem.*, **2003**, 278(30), 27981-27987.
 [12] Greenfield, N.J. Analysis of circular dichroism data. *Meth. Enzymol.*, **2004**, 383, 282-317.
 [13] Soding, J.; Biegert, A.; Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, **2005**, 33, W244-248.
 [14] Matsuura, Y.; Stewart, M. Structural basis for the assembly of a nuclear export complex. *Nature*, **2004**, 432(7019), 872-877.
 [15] Marti-Renom, M.A.; Stuart, A.C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.*, **2000**, 29, 291-325.
 [16] Sali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **1993**, 234(3), 779-815.
 [17] Lovell, S.C.; Davis, I.W.; Arendall, W.B., 3rd; de Bakker, P.I.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins*, **2003**, 50(3), 437-450.
 [18] Wiederstein, M.; Sippl, M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.*, **2007**, 35, W407-410.
 [19] Berendsen, H.J.C.; Vanderspoel, D.; Vandrunen, R. Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput. Phys. Commun.*, **1995**, 91(1-3), 43-56.
 [20] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Hermans, J. *Interaction models for water in relation to protein hydration*, in: Pullman B. (Ed.), Intermolecular Forces, Publishing Company, Dordrecht, **1981**, pp. 331-342.
 [21] Oostenbrink, C.; Soares, T.A.; van der Vegt, N.F.; van Gunsteren, W.F. Validation of the 53A6 GROMOS force field. *Eur. Biophys. J.*, **2005**, 34(4), 273-284.
 [22] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **1984**, 81(8), 7.
 [23] Sreerama, N.; Woody, R.W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.*, **2000**, 287(2), 252-260.
 [24] Freitas, F.Z.; Bertolini, M.C. Genomic organization of the *Neurospora crassa* gsn gene: possible involvement of the STRE and HSE elements in the modulation of transcription during heat shock. *Mol. Genet. Genomics*, **2004**, 272(5), 550-561.
 [25] Marfori, M.; Mynott, A.; Ellis, J.J.; Mehdi, A.M.; Saunders, N.F.; Curmi, P.M.; Forwood, J.K.; Boden, M.; Kobe, B. Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim. Biophys. Acta.*, **2011**, 1813(9), 1562-1577.
 [26] Catimel, B.; Teh, T.; Fontes, M.R.; Jennings, I.G.; Jans, D.A.; Howlett, G.J.; Nice, E.C.; Kobe, B. Biophysical characterization of

- interactions involving importin-alpha during nuclear import. *J. Biol. Chem.*, **2001**, 276(36), 34189-34198.
- [27] Cutress, M.L.; Whitaker, H.C.; Mills, I.G.; Stewart, M.; Neal, D.E. Structural basis for the nuclear import of the human androgen receptor. *J. Cell. Sci.*, **2008**, 121(Pt7), 957-968.
- [28] Takeda, A.A.; de Barros, A.C.; Chang, C.W.; Kobe, B.; Fontes, M.R. Structural basis of importin-alpha-mediated nuclear transport for Ku70 and Ku80. *J. Mol. Biol.*, **2011**, 412(2), 226-234.
- [29] Lott, K.; Bhardwaj, A.; Sims, P.J.; Cingolani, G.. A minimal nuclear localization signal (NLS) in human phospholipid scramblase 4 that binds only the minor NLS-binding site of importin alpha1. *J. Biol. Chem.*, **2011**, 286(32), 28160-28169.
- [30] Takeda, A.A. *Structural studies with the importin*. Doctorate Thesis. Universidade Estadual Paulista Júlio De Mesquita Filho-UNESP, Botucatu-Brasil, **2009**.
- [31] Mynott, A.V. *Structural Investigations of CLIC Proteins and Importin- α Recognition of Nuclear Localisation Signals*. Doctorate Thesis. University of New South Wales, Sidney- Australia, **2009**.
- [32] Barros, A. C.; Takeda, A.A.S.; Chang, C.; Kobe, B. and Fontes, M.R.M. Structural basis of nuclear import of Flap endonuclease 1 (FEN1). *Acta Crystallogr. D Biol. Crystallogr.*, **2012**, D68(7), 743-750.

Received: May 2, 2012

Revised: May 31, 2012

Accepted: June 1, 2012

ANEXOS

Tabela de aminoácidos

	Aminoácido	Abreviação
Hidrofóbicos	Alanina	A
	Valina	V
	Leucina	L
	Isoleucina	I
	Prolina	P
	Fenilalanina	F
	Triptofano	W
	Metionina	M
Polares com carga nula	Glicina	G
	Serina	S
	Treonina	T
	Tirosina	Y
	Asparagina	N
	Glutamina	Q
	Cisteína	C
Polares ácidos	Ácido aspártico	D
	Ácido glutâmico	E
Polares básicos	Lisina	K
	Arginina	R
	Histidina	H