

Universidade Estadual Paulista “Júlio de Mesquita Filho” - UNESP
Instituto de Biociências de Botucatu
Departamento de Física e Biofísica

**Efeitos de deleções pontuais em resíduos da região
linker das sequências de localização nuclear
clássicas nas interações com a Importina- α**

Monografia apresentada no curso de Ciências Biomédicas do Instituto de Biociências de Botucatu da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) como trabalho de conclusão do curso para obtenção do bacharelado.

Aluna: Thábatta Karollynne Estevam Nakamura

Orientador: Prof. Dr. Ney Lemke

Botucatu – SP
Dezembro/2016

**Efeitos de deleções pontuais em resíduos da região *linker*
das sequências de localização nuclear clássicas nas
interações com a Importina- α**

Thábatta Karollynne Estevam Nakamura

Monografia apresentada no curso de Ciências Biomédicas do Instituto de Biociências de Botucatu da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) como trabalho de conclusão do curso para obtenção do bacharelado.

Orientador: Prof. Dr. Ney Lemke

Ciências Biomédicas/ IBB – UNESP
Botucatu – SP
Dezembro/2016

Dedicatória

Dedico esse trabalho àqueles que por toda
a vida me deram asas para sonhar e alcançar
meus sonhos: meus pais

Agradecimentos

Agradeço aos meus pais, que me permitiram embarcar nesse universo que é a vida universitária – nova cidade com novos amigos a se conhecer. Por me permitirem explorá-lo sozinha, ora errando ora acertando, mas sempre crescendo, o que me permitiu tornar-me uma adulta responsável, que sonha alto e sempre segue em frente para torna-los realidade.

Agradeço ao Prof Dr. Ney Lemke pela oportunidade e confiança de realizar meu estágio de conclusão de curso junto ao grupo LBBC. Cheguei a sua porta sem conhecimento algum na área de bioinformática, levando comigo apenas a curiosidade de aprender sobre – já que é uma área de pesquisa ainda negligenciada no curso de Ciências Biomédicas – e mesmo assim fui muito bem recebida e orientada ao longo de todo ano.

Agradeço também a todo grupo LBBC pelo apoio e suporte durante todo ano. Agradeço em especial ao Marcos T. Geraldo pela paciência que demonstrou ao me ensinar sobre programação. Por detalhar passo a passo como utilizar um terminal, o porque do uso de cada comando, por me auxiliar no desenvolvimento do projeto e durante a escrita do trabalho de conclusão de curso e, principalmente, pela ajuda nos momentos que erre e/ou tive dúvidas, que foram muitos.

Agradeço a todos aqueles que adentraram minha vida e rotina durante esses anos de graduação. Aos amigos que Botucatu me trouxe e que levarei por toda a vida, independente da distância em que estivermos, amigos de festa, PET, da XLVII, e da República Curva de Rio. Por todo suporte, conselhos, conversas, desabafos e cervejas compartilhadas principalmente nos momentos em que a ausência da família se fez mais forte. Agradeço a todos os professores presentes no meu dia-a-dia durante a graduação, responsáveis pelo ensino e orientação. Agradeço principalmente às professoras: Dra. Doralina Guimarães Brum de Souza que me auxiliou no início da minha carreira acadêmica através da primeira iniciação científica na pesquisa clínica área de Neurologia; Profa. Dra. Cláudia Pellizzon e Profa. Dra. Luciene Mascarini Serra por sempre me ajudarem no que precisei, desde

o desenvolvimento de um pensamento crítico e social durante minha participação no Programa de Educação Tutorial PET-CBB até a finalização da graduação.

Agradeço ao ARMI (Australian Regenerative Medicine Institute) e à Monash University, instituto e universidade a qual tive oportunidade de fazer parte durante meu ano de intercâmbio em Melbourne, AUS. A convivência com cientistas e professores do ARMI permitiu que eu me encontrasse na carreira de Biomedicina. Em especial agradeço aos professores Jan Kaslin e Minni Anko por me aceitarem em seus grupos de pesquisa para desenvolvimento de projeto, mesmo sem experiência prévia na área de pesquisa experimental, e ainda sob fase de aprendizado da língua inglesa. Por me ensinarem diferentes técnicas e pela paciência que tiveram em me auxiliar/corrigir no inglês, no desenvolvimento do projeto de pesquisa e nas minhas primeiras experiências de apresentação de seminário, escrita e defesa de tese inteiramente em inglês.

Agradeço também a todos os programas de auxílio financeiro (Programa de Educação Tutorial (PET-CBB 2012) com o auxílio do CNPq; à bolsa de Iniciação Científica PIBIC (2013/2014) e FAPESP (2016); e ao Ciências sem Fronteiras (2014/2015) a qual obtive suporte financeiro durante meus anos de graduação.

Resumo

A via de importação nuclear é importante objeto de estudos porque está relacionada a diversos processos da funcionalidade celular. Na via clássica de importação nuclear as proteínas são importadas pelo complexo constituído pelas proteínas Importina- α e Importina- β . A Importina- α é proteína de formato solenoide que apresenta dez repetições do tipo Armadillo (ARM), sendo cada repetição formada por três α -hélices. A Importina- α é capaz de adotar diferentes conformações dependendo da proteína a qual se ligou principalmente por conta de seu formato solenoide. O processo de importação nuclear clássico se inicia a partir do reconhecimento das sequências de localização nuclear (NLS), presentes nas proteínas a serem importadas, que se ligam diretamente à Importina- α e são classificadas em monopartidas ou bipartidas. NLSs clássicas (cNLSs) bipartidas possuem dois sítios de ligação (principal e secundário) separadas por uma região intermediária chamada de *linker*, que é constituída por dez a doze resíduos. Dados da literatura mostram que mutações pontuais na região *linker* não apresentam efeitos na importação nuclear, contudo *linker* menor que dez resíduos impede este processo. Neste projeto, objetivamos analisar como o complexo Importina- α :cNLS bipartida é afetado, quando realizamos uma deleção pontual na região *linker*. Três modelos de estudo foram utilizados: a estrutura canônica da Importina- α ligada a NLS da Nucleoplasmina e duas estruturas com deleção pontual no *linker* nos resíduos ^{Npl}A159 (contato instável) e ^{Npl}K162 (contato estável). Interações entre resíduos da cNLS e da Importina- α foram analisadas a partir dos dados gerados por modos normais, no qual calculamos a frequência de contatos, valores de flutuação, correlação entre resíduos e a flexibilidade da Importina. Resultados mostram que a retirada de um resíduo do *linker* pode aumentar a interação entre resíduos do sítio principal, mas diminuindo a interação dos resíduos no sítio secundário, e vice-versa. Resíduos da região *linker* possuem maior variabilidade na flutuação dos resíduos indicando tentativa da Importina- α e da cNLS de se acomodarem. Assim, a ausência de cNLSs com região *linker* menor que dez resíduos poderia ser explicada pela dificuldade de acomodação entre os resíduos gerando uma instabilidade no complexo Importina- α :cNLS.

Abstract

The nuclear import pathway is an important subject of studies because is related to several cell functionality processes. In the classical nuclear import pathway, proteins are translocated via the complex comprised of Importin- α and Importin- β proteins. Importin- α is a solenoid protein composed of ten Armadillo (ARM) repeats, each one formed by three α -helices. Importin- α is capable of adopting different conformations, depending on the protein bound to it because of its solenoid architecture. The classical process of nuclear import starts with the recognition of nuclear localization sequences (NLS) within the cargo proteins. The classical NLSs (cNLSs) bind directly to Importin- α and are classified as monopartite or bipartite. Bipartite cNLSs have two binding sites (major and minor) separated by an intermediary region named as linker, which is comprised of ten to twelve residues. Literature data shows that point mutations at the linker region do not affect nuclear import, however, deletions in this region prevent the nuclear import process. In this project, we analyzed the effect of linker point mutations on the formation of the Importin- α :cNLS complex. Three models were used: the canonical structure of Importin- α bound to the Nucleoplasmin NLS and two structures with point deletions in the linker at residues ^{Npl}A159 (instable contact) and ^{Npl}K162 (stable contact). All models were submitted to normal modes calculations, and the interaction analyses between the bipartite cNLS and Importin- α residues were carried out accordingly to their interaction frequencies, fluctuation values, residues correlations and Importin- α flexibility. The results showed that interactions in the major site enhanced by reducing the contacts in the minor site, and vice-versa, as a consequence of a shorter linker region. Moreover, the residues fluctuations were increased, primarily in the linker, and could be related to an attempt of Importin- α to accommodate the cNLS. Thus, the absence of NLSs with linker region smaller than ten residues could be explained by the accommodation restraint between residues, generating instability to the complex Importin- α :cNLS complex.

Lista de Figuras

Figura 1. Estrutura do complexo poro nuclear (NPC).....	2
Figura 2. Via clássica de importação nuclear.....	3
Figura 3. Estrutura da NLS da Nucleoplasmina ligada à Importina- α	5
Figura 4. Estrutura de contatos na Importina- α	6
Figura 5. Sequências de localização nuclear.....	7
Figura 6. Modelo de Importina- α complexada com cNLS de Nucleoplasmina (Marfori <i>et al.</i> , 2012).....	10
Figura 7. Fluxograma do planejamento de atividades.....	11
Figura 8. Energias de restrição da estrutura canônica para os modos normais 7 a 50.....	16
Figura 9. Interações entre resíduos da região <i>linker</i> do NLS da Nucleoplasmina com Importina- α com frequência maiores que 50%.....	17
Figura 10. Energias de restrição dos modelos de deleção da ^{Npl} A159 e ^{Npl} K162 para os modos normais 7 a 50.....	19
Figura 11. Valores de flutuação entre resíduos da Importina- α e NLS da Nucleoplasmina.....	20
Figura 12. Correlação entre resíduos da Importina- α e NLS da Nucleoplasmina.....	21

Figura 13. Correlação gerado por modos normais e dinâmica molecular nas estruturas de deleção da ^{Npl} A159 e ^{Npl} K162.....	23
Figura 14. Representação no Pymol da movimentação da Importina- α	26
Figura 15. Ângulo entre α -hélices H3 de repetições ARM vizinhas da Importina- α ao longo do modo.....	27
Figura 16. Raio de curvatura da Importina- α	28

Tabelas e Figuras Apêndice

Figura A1. Alinhamento dos resíduos para modelagem.....	36, 37
Figura A2. Interações entre NLS da Nucleoplasmina e Importina- α com frequência acima de 50% no modelo canônico.....	38
Figura A3. Interações entre a região <i>linker</i> da NLS da Nucleoplasmina e Importina- α com frequência abaixo de 50% no modelo canônico.....	39, 40
Figura A4. Interações entre NLS da Nucleoplasmina e Importina- α com frequência acima de 50% no modelo de deleção no cNLS do resíduo ^{Npl} A159.....	41, 42
Figura A5. Interações entre NLS da Nucleoplasmina e Importina- α com frequência acima de 50% no modelo de deleção no cNLS do resíduo ^{Npl} K162.....	43, 44
Figura A6. Ângulo entre α -hélices H3 de repetições ARM vizinhas da Importina- α ao longo do modo.....	45, 46

Lista de Abreviações

Å – Ångstrom

ARM – Repetições *Armadillo*

cNLS – Sequência de Localização Nuclear clássica

Repetições FG – Repetições Fenilalanina-Glicina

IBB – Domínio de ligação à Importina- β

Imp- α – Importina- α

kDa – Kilo-Dalton

MMFP – *Miscellaneous Mean Field Potential*

MRMS – *Mass-weight Root Mean Square*

NLS – Sequência de Localização Nuclear

Nm – Nanômetros

NPC – Complexo Poro Nuclear

Npl – Nucleoplasmina

Ns – Nanossegundos

Ps - Picossegundos

RMSF – *Root Mean Square Fluctuation*

SV40-TAg – *Simian Virus 40 large T-antigen*

VMD – *Visual Molecular Dynamics*

Lista de Amino Ácidos

ALA – Alanina (A)

ARG – Arginina (R)

ASN – Asparagina (N)

ASP – Ácido Aspártico (D)

CYS – Cisteína (C)

GLN – Glutamina (Q)

GLU – Ácidos Glutâmico (E)

GLY – Glicina (G)

HIS – Histidina (H)

ILE – Isoleucina (I)

LEU – Leucina (L)

LYS – Lisina (K)

MET – Metionina (M)

PHE – Fenilalanina (F)

PRO – Prolina (P)

SER – Serina (S)

THR – Treonina (T)

TRP – Triptofano (W)

TYR – Tirosina (Y)

VAL – Valina (V)

Sumário

Dedicatória	III
Agradecimentos	IV
Resumo	VI
Abstract	VII
Lista de Figuras	VIII
Tabelas e Figuras Apêndice	X
Lista de Abreviações	XI
Lista de Amino Ácidos	XII
Sumário	XIII
1. Introdução	1
1.1. Sistemas de importação nuclear	1
1.2. Via clássica da importação nuclear	2
1.3. Importina- α e sequências de localização nuclear clássicas	4
1.4. Motivação	8
2. Objetivos	9
3. Metodologia	10
3.1. Modelo de estudo	10
3.2. Análises de Modos Normais e estruturas ao longo do modo	11
3.3. Frequência de interações entre os resíduos da Importina- α e a cNLS da Nucleoplasmina e valores de flutuação	12
3.4. Modelagem das estruturas com deleção pontual na região linker da NLS da Nucleoplasmina	13
3.5. Ângulos entre α -hélices H3 de ARMs vizinhos e raio de curvatura da Importina- α	14
3.6. Dinâmica Molecular	14
4. Resultados	15
4.1. ^{Npl} K162 é contato de característica estável na região linker e ^{Npl} A159 é contato instável	15
4.2. Interações entre resíduos da Importina- α e NLS da Nucleoplasmina nos modelos de deleção da ^{Npl} K162 e ^{Npl} A159	18

4.3. Valores de flutuação entre resíduos da Importina- α e NLS da Nucleoplasmina	19
4.4. Diferenças entre Modos Normais e Dinâmica Molecular segundo a análise de correlação dos resíduos da Importina- α e da NLS.....	22
4.5. Análises dos movimentos da Importina- α	23
5. Discussão e Conclusão	29
6. Referências	31
7. Apêndice	36

1. Introdução

1.1. Sistemas de importação nuclear

O desenvolvimento de organelas citoplasmáticas e do envelope nuclear difere os seres vivos eucariontes dos procariontes. O envelope nuclear é formado por membrana dupla que engloba o material genético e funciona como barreira para processos celulares como a transcrição e tradução. Apesar da separação núcleo-citoplasmática pelo envelope nuclear, o constante transporte de moléculas ocorre controladamente entre os compartimentos através de poros presentes conectando as membranas nucleares ao longo do envelope (Dingwall e Laskey, 1986; Stewart, 2007). Esses poros são conhecidos como complexo poro nuclear (NPC). NPCs possuem função e estruturas conservadas em diferentes organismos. Formados por aproximadamente 30 formas diferentes de nucleoporinas dispostas em simetria octogonal conferindo às NPCs formato de cesta com região central de aproximadamente 25 a 30 nanômetros (nm) de diâmetro e fibras se alongando em direção ao citoplasma formando os filamentos citoplasmáticos e em direção ao núcleo formando a cesta nuclear (Figura 1) (Rout e Wente, 1994; Cronshaw *et al.*, 2002).

O transporte núcleo-citoplasmático ocorre por difusão simples ou transporte ativo (Dingwall e Laskey, 1986). Comumente, moléculas pequenas, com peso de até 40 kilo-Daltons (kDa) são transportadas por difusão simples pelos poros (Dingwall e Laskey, 1986). Em casos de difusão simples, a concentração da molécula dentro e fora do núcleo determina a direção do transporte. Em contrapartida, macromoléculas que possuem peso maior que 40 kDa são transportadas ativamente com a ajuda de moléculas transportadoras ou fatores de transporte solúveis (Stewart, 2007). No transporte ativo, importação e exportação ocorrem com a ajuda de proteínas transportadoras denominadas importinas (Importina- α e Importina- β) e exportinas, respectivamente.

Importinas são proteínas flexíveis de formato solenoide com capacidade de translocação núcleo-citoplasma através dos NPCs. Constituídas principalmente pela subfamília das Importinas- β , reconhecem as proteínas a serem transportadas

por específicos sítios de ligação. O transporte é direcionado pelas proteínas Ran, RanGTP e RanGDP, encontradas altamente concentradas no núcleo e citoplasma, respectivamente, fornecendo energia para o transporte e auxiliando na reciclagem das importinas (Chook e Blobel, 2001; Forwood *et al.*, 2008). Nos seres humanos a Importina- β 1 é a principal responsável pela importação. As proteínas a serem importadas podem se ligar diretamente à Importina- β 1 ou através de proteínas adaptadoras como Importina- α , Snuportina-1 ou Importina- β 7 (Jakel *et al.*, 1999; Stewart, 2007). A maioria das proteínas é importada com o auxílio da Importina- α , que atua como proteína adaptadora da Importina- β , classificando esta via como via clássica da importação nuclear (Figura 2) (Stewart, 2007).

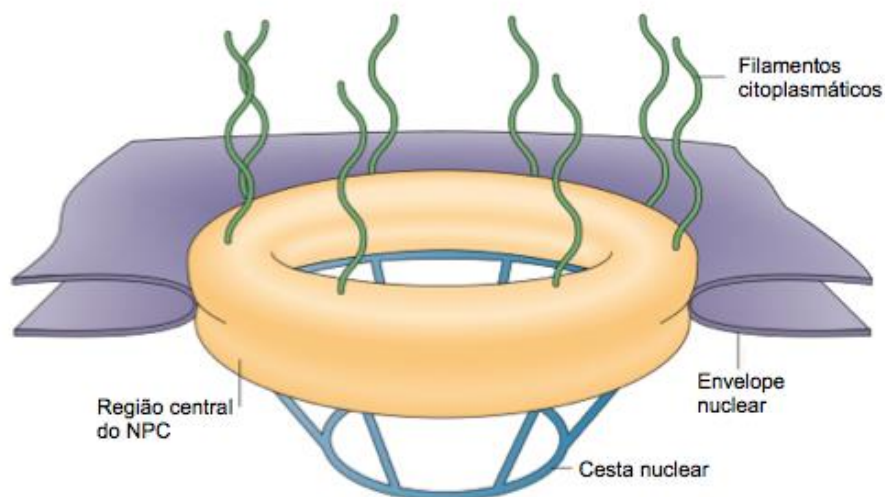


Figura 1. Estrutura do complexo poro nuclear (NPC) (adaptado de Stewart (2007)). Formado por aproximadamente 30 diferentes nucleoporinas dispostas em simetria octogonal conferindo formato de cesta com região central de aproximadamente 25 a 30 nm de diâmetro e fibras se alongando em direção ao citoplasma (filamentos citoplasmáticos) e em direção ao núcleo (cesta nuclear).

1.2. Via clássica da importação nuclear

A via clássica da importação nuclear ocorre com o auxílio da Importina- α , proteína adaptadora que forma o complexo de importação através da ligação à Importina- β e à proteína a ser importada. A literatura descreve o processo de

importação nuclear em quatro fases: a ligação da proteína com o transportador; translocação através das NPCs; liberação da proteína no núcleo e reciclagem das importinas (Figura 2).

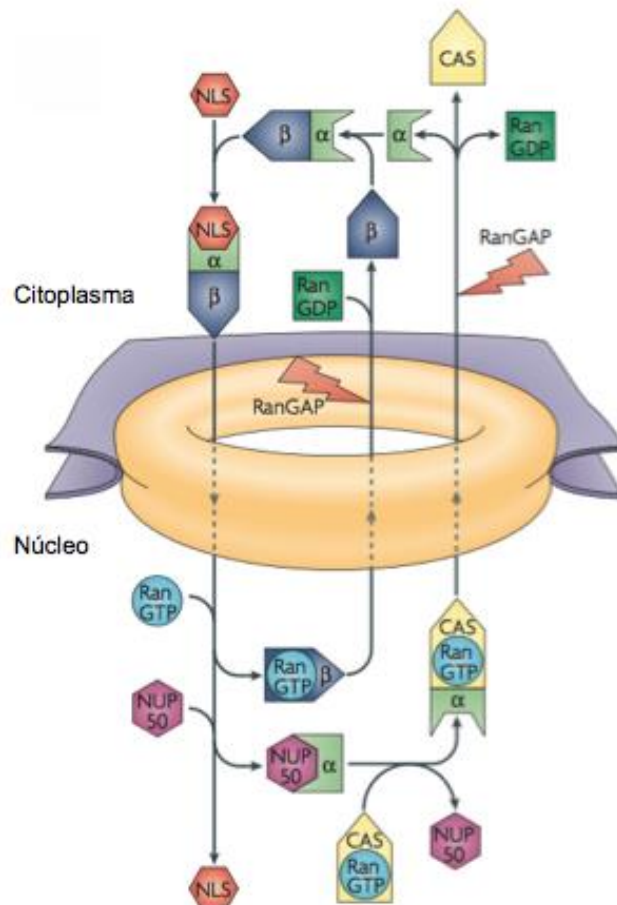


Figura 2. Via clássica de importação nuclear (adaptado de Stewart (2007)). O complexo de importação é formado no citoplasma entre proteínas a serem importadas que apresentam NLSs, Importina-α e Importina-β. Depois de passar através das NPCs, a ligação de RanGTP na Importina-β dissocia esta da Importina-α. A proteína contendo NLS é então deslocada da Importina-α e esta é transportada para o citoplasma por seu fator de exportação nuclear (i.e. CAS) complexado com RanGTP. No citoplasma, RanGAP estimula a hidrólise de GTP, liberando as importinas para um outro ciclo de importação.

A Importina-α é normalmente encontrada inativada no citoplasma através de processo de auto-inibição, mas ativa-se através de ligação com a Importina-β. Quando ativada, liga-se à proteína a ser carregada através de dois possíveis sítios de

ligação, denominados sítios principal e secundário. Moléculas importadas são reconhecidas através de sequências de localização nuclear (NLS), ricas em resíduos básicos, principalmente lisina e arginina. Proteínas importadas pela via clássica da importação nuclear apresentam as chamadas NLS clássicas (cNLS) que se ligam à Importina- α concomitantemente pelos dois sítios de reconhecimento (sítio principal e secundário), ou apenas pelo sítio principal, o que caracteriza a cNLS como bipartida ou monopartida, respectivamente (Stewart, 2007).

Com o complexo formado, o transporte para o núcleo ocorre através de interações realizadas por resíduos hidrofóbicos da Importina- β com nucleoporinas ricas em repetições fenilalanina-glicina (repetições FG) das NPCs (Fahrenkrog e Aebi, 2003). No núcleo, a Importina- β se liga à RanGTP, a qual possui afinidade maior comparada à Importina- α . A ligação entre Importina- β e RanGTP muda a conformação da Importina- β liberando a Importina- α que então se auto-inibe, inativando-se e facilitando a liberação da carga no núcleo. Ambas importinas são então recicladas com o auxílio da RanGTP para realizarem novos transportes. A RanGTP exporta a Importina- β diretamente e a Importina- α com auxílio da proteína nuclear CAS. No citoplasma, RanGTP se liga à RanBP1 formando uma estrutura intermediária que causa mudança conformacional na estrutura da RanGTP liberando as Importinas. Essa mudança estrutural da RanGTP permite o acesso à RanGAP que realiza a hidrolização da RanGTP em RanGDP.

1.3. Importina- α e sequências de localização nuclear clássicas

A Importina- α é uma proteína de formato solenoide com estrutura conservada entre espécies. Contudo, a quantidade de genes decodificantes da Importina- α ampliou-se com a evolução da complexidade dos organismos. Bactérias *Saccharomyces cerevisiae*, por exemplo, possuem apenas um gene decodificante de Importina- α , enquanto que *Drosophila melanogaster* e *Caenorabditis elegans* possuem três Importinas- α , e seres humanos, sete (Kohler *et al.*, 1999; Arjomand *et al.*, 2014; Chen e Madura, 2014). Estruturalmente, a Importina- α é formada por uma região N-terminal e uma região C-terminal (Figura 3). A região N-terminal da Importina- α apresenta o domínio de ligação à Importina- β (IBB). Quando a Importina- α está inativa no citoplasma, o sítio IBB está realizando interação com a própria importina- α na região do sítio principal. No entanto, a Importina- α é ativada através

da interação dessa região com a Importina- β , que libera o sítio principal para realizar interações com as cNLSs das proteínas a serem importadas. (Pang e Zhou, 2014; Christie *et al.*, 2015; Miyatake *et al.*, 2015).

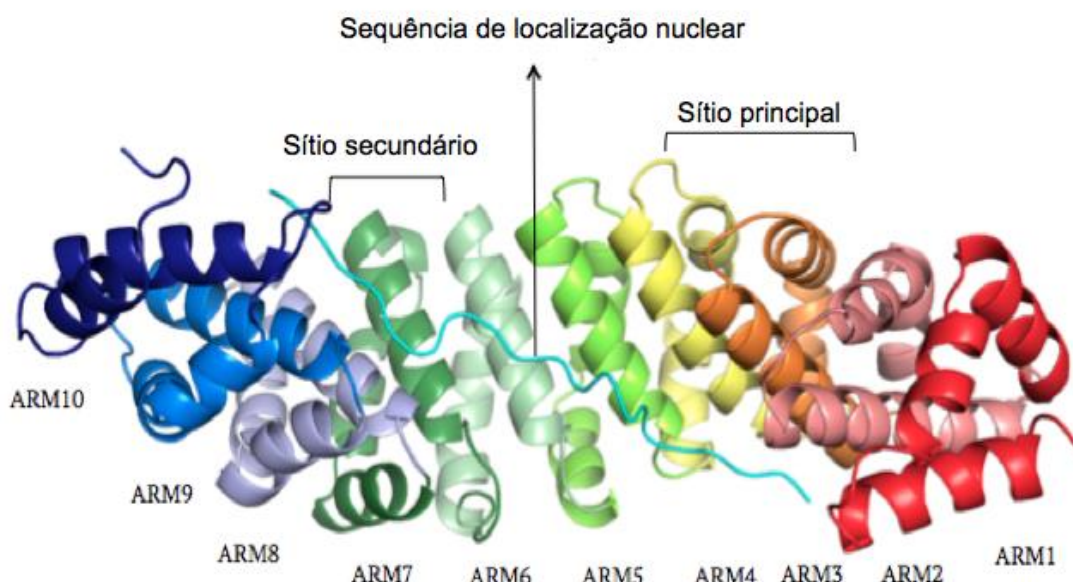


Figura 3. Estrutura da NLS da Nucleoplasmina (ciano) ligada à Importina- α . As 10 repetições ARM da Importina- α são exibidas com cores diferentes: ARM1 = “tv_red”, ARM2 = “salmon”, ARM3 = “tv_orange”, ARM4 = “tv_yellow”, ARM5 = “chartreuse”, ARM6 = “pale green”, ARM7 = “forest”, ARM8 = “lightblue”, ARM9 = “marine”, ARM10 = “density”. Os locais de ligação da NLS estão localizados na superfície cônica da Importina- α : o sítio principal se estende das repetições ARM 2-4, ao passo que o sítio secundário se estende das repetições ARM 6-8.

A região C-terminal é constituída por 10 repetições conhecidas como repetições *Armadillo* (ARM). Cada repetição ARM é composta aproximadamente 40 aminoácidos que constituem 3 α -hélices (H1, H2 e H3). As α -hélices H1 e H2 são responsáveis pela curvatura convexa da proteína enquanto que a α -hélice H3 é responsável pela curvatura cônica da Importina- α . Ademais, as α -hélices H3 compreendem os sítios de reconhecimento das proteínas a serem importadas (Conti *et al.*, 1998; Pang e Zhou, 2014; Christie *et al.*, 2015; Pumroy e Cingolani, 2015). Como dito anteriormente, existem dois sítios de reconhecimento da cNLS na Importina- α , o sítio principal que está localizado entre as repetições ARM2 e ARM4,

e o sítio secundário entre as repetições ARM6 e ARM8 (Marfori *et al.*, 2012; Pang e Zhou, 2014). As posições em que a cNLSs e a Importina- α interagem no sítio principal são denominadas P2 a P5, e no sítio secundário, P1' e P2' (Figuras 4 e 5B). Em cNLS bipartidas a região intermediária aos dois sítios de ligação é chamada de região *linker* e é constituída normalmente por 10 a 12 resíduos.

Dentre as formas de ligação entre os resíduos, quatro são classificadas como principais: os contatos hidrofóbicos entre cadeias laterais de asparagina da Importina- α e cNLS que determina a direção da cadeia; contatos hidrofóbicos da cNLS com cadeias laterais dos triptofanos da Importina- α ; pontes salinas e ligações de hidrogênio entre os resíduos das cadeias (Conti *et al.*, 1998; Fontes *et al.*, 2000).

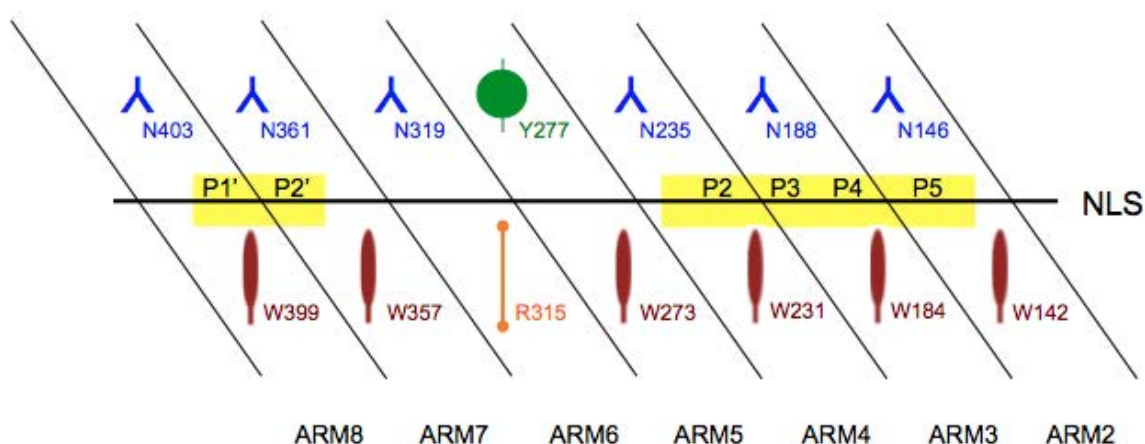


Figura 4. Estrutura de contatos na Importina- α . Entre as repetições ARM2 e ARM8 a Importina- α possui estrutura conservada composta por resíduos de triptofano e asparagina nas α -hélices H3, com exceção do ARM5 em que a asparagina é substituída por uma tirosina e no ARM6 em que o triptofano é substituído por arginina. O sítio principal da Importina- α está localizado entre os ARMs 2 e 4 contendo os pontos de contato P2 a P5. Já o sítio secundário está localizado entre os ARMs 6 e 8 contendo os pontos de contato P1' e P2'.

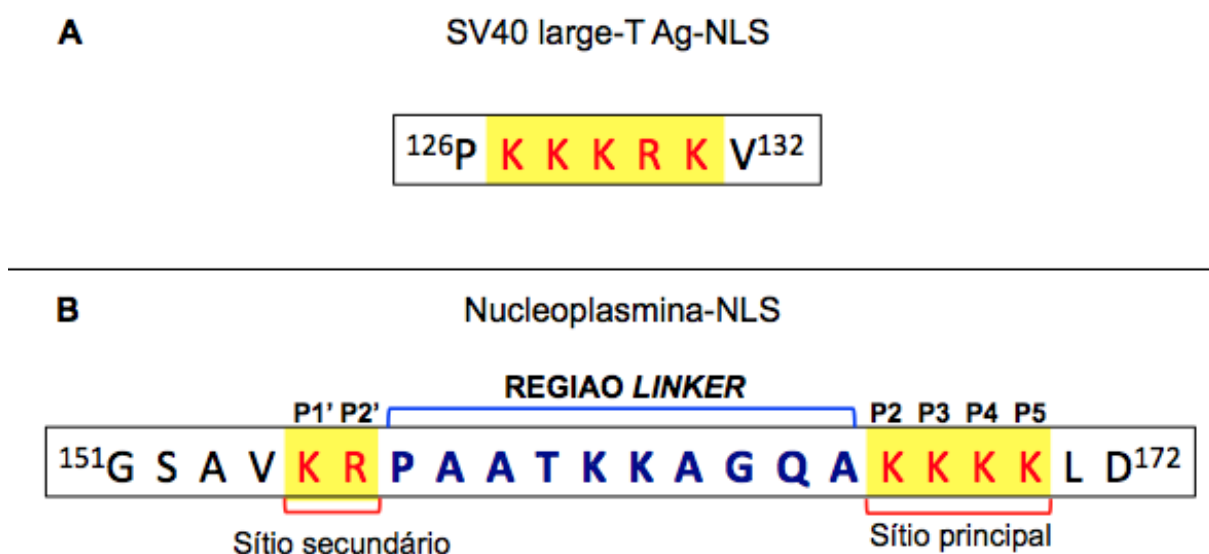


Figura 5. Sequências de localização nuclear. **A.** cNLS da SV40 large-T Ag, primeira cNLS monopartida a ser descrita. Região em vermelho e amarelo representa a região de ligação da cNLS com o sítio principal da Importina- α . **B.** cNLS da Nucleoplasmina, primeira cNLS bipartida a ser descrita. O sítio principal realiza ligações com a Importina- α nas posições P2 a P5 e o sítio secundário nas posições P1' e P2' (representados em amarelo e vermelho) A região *linker* compreende os resíduos presentes entre os dois sítios na NLS bipartida, representados em azul.

Dentre as formas de interação entre cNLSs e Importina- α já descritas na literatura, a posição P2 no sítio principal é estritamente conservada entre espécies, contendo um resíduo de lisina. Por sua vez, o sítio secundário, apesar de mais diversificado entre as proteínas, possui pontos altamente conservados como a posição P1' e P2' que usualmente apresentam os resíduos lisina e arginina (Stewart, 2007; Chang *et al.*, 2013; Pang e Zhou, 2014). Sequências consenso foram designadas para proteínas que apresentam cNLSs. A cNLS monopartida foi descrita a partir da NLS do *Simian Virus 40 large T-antigen* [^{SV40-TAg}PKKKRRV¹³²], primeira cNLS monopartida identificada (Figura 5A) (Smith *et al.*, 1985; Chang *et al.*, 2013). Assim sendo, cNLSs monopartidas normalmente possuem sequência K[K/R]X[K/R], no qual o X representa qualquer resíduo. A cNLS bipartida, por sua vez, foi descrita pela primeira vez com base na estrutura da NLS da proteína Nucleoplasmina (Npl) [^{Npl}VKRPAATKKAGQAKKKKLD¹⁷²], no qual os resíduos sublinhados KR e KKKK

ligam-se à Importina- α no sítio secundário e principal, respectivamente (Figura 5B) (Fontes *et al.*, 2000; Marfori *et al.*, 2012; Chang *et al.*, 2013). A partir da NLS da Nucleoplasmina, a sequência consenso gerada para cNLSs bipartidas corresponde a $[K/R][K/R]X_{10-12}[K/R]_{3/5}$, no qual $[K/R]_{3/5}$ representa ao menos a presença de lisinas ou argininas de três a cinco resíduos consecutivos e X_{10-12} representa a quantidade usual de resíduos presentes na região *linker* (Bernardes *et al.*, 2015).

1.4. Motivação

A literatura tem mostrado por análises experimentais que simples mutações pontuais ou substituição de resíduos nos sítios principal e secundário afeta substancialmente a importação de cNLSs bipartidas podendo acarretar na diminuição ou interrupção da importação de uma determinada proteína por conta da diminuição das interações da NLS com a Importina- α (Dingwall e Laskey, 1986; Robbins *et al.*, 1991; Conti *et al.*, 1998; Leung *et al.*, 2003; Roman *et al.*, 2013; Chen e Madura, 2014). Kalderon *et al.* (1984) por exemplo, apresentou um dos primeiros estudos com o SV40 TAg [^{SV40-TAg}PKKKRRV¹³²], no qual através de mutações pontuais na ^{SV40-TAg}K128 (em negrito), observava a interrupção da importação desta proteína para o núcleo.

Além disso, dados estruturais mostram a ocorrência de interações a partir da região *linker* de algumas cNLSs bipartidas com resíduos da Importina- α , destacando a preferência por *linker* de resíduos mais ácidos e com um mínimo de 10 resíduos, no entanto mutações nos resíduos não parecem ter efeito na importação (Conti *et al.*, 1998; Conti e Kuriyan, 2000; Fontes *et al.*, 2000; Fontes *et al.*, 2003; Kuroda *et al.*, 2006; Christie *et al.*, 2015). Robbins *et al.* (1991), por exemplo, realizando mutações pontuais em diferentes resíduos da região *linker* da Nucleoplasmina [^{Npl}VKRPAATKKAGQAKKKKLD¹⁷²], especificamente as mutações ^{Npl}P→T157, ^{Npl}A→P163, ^{Npl}K→N161 e ^{Npl}K→N162, observou que tais alterações não tiveram efeitos na importação nuclear. No entanto, deleções nessa região impediram a importação, o que indica a possível importância desta região neste processo e a necessidade de uma região *linker* de no mínimo dez resíduos.

Mais recentemente, Geraldo *et al.* (2016) com base em Análise de Modos Normais e Dinâmica Molecular mostrou a existência de contatos estáveis entre resíduos da região *linker* com a Importina- α . Modos normais são oscilações

harmônicas simples a partir de um mínimo local de energia, característico de um sistema (Konrad, 2005). Por meio da análise de modos normais de menor frequência e, conseqüentemente maior amplitude, se procura prever o tipo de mudança conformacional de uma proteína a fim de cumprir a sua função. Exemplos da aplicação dos modos normais demonstram mudanças conformacionais seguindo as trajetórias de um ou mais modos normais. Normalmente, aproximadamente 50% dos movimentos observados em uma proteína podem ser descritos por um ou dois modos de baixa frequência (Suhre e Sanejouand, 2004). Há duas vantagens principais que torna a análise de modos normais uma técnica frequentemente utilizada em estudos de dinâmica de macromoléculas. Primeiro, os resultados são precisos e sem erros estatísticos. Segundo, a técnica de modos normais permite a observação de movimentos de maior amplitude (Konrad, 2005).

2. Objetivos

Apesar da existência de estudos analisando a importância da região *linker*, pouco ainda foi explorado em relação aos contatos realizados por seus resíduos, especialmente sobre a implicação de deleções nesta região. Assim, o presente trabalho teve como objetivo analisar os efeitos das deleções pontuais da região *linker* da NLS para a importação nuclear.

Uma maneira de abordar tal questão é através do uso de técnicas de simulação de biomoléculas, sendo que a Análise de Modos Normais um dos métodos que tem adquirido destaque nas últimas décadas. Desta maneira, os objetivos específicos foram: (i) determinar através da análise de modos normais a frequência de interações hidrofóbicas, ligações de hidrogênio e pontes salinas entre resíduos da região *linker* da NLS da Nucleoplasmina e Importina- α ; (ii) a partir dessas frequências, selecionar dois resíduos no *linker*, um de contato estável e um de contato instável; e (iii) através de deleção pontual destes resíduos, gerar um *linker* menor que dez resíduos para determinar se as interações com a Importina- α são mantidas e quais os efeitos para a importação nuclear.

3. Metodologia

3.1. Modelo de estudo

Nosso modelo de estudo foi a estrutura cristalográfica da Importina- α truncada no sítio IBB (*Protein Data Bank* ID: 3UL1; (Marfori *et al.*, 2012) complexada com a cNLS bipartida da proteína Nucleoplasmina (^{Npl}GSAVKRPAATKKAGQAKKKKLD¹⁷²; os resíduos sublinhados representam as duas regiões de ligação com a Importina- α : os resíduos KR se ligam ao sítio secundário da Importina- α nas posições P1' e P2' e os resíduos KKKK se ligam ao sítio principal nas posições P2 a P5) (Figura 6).

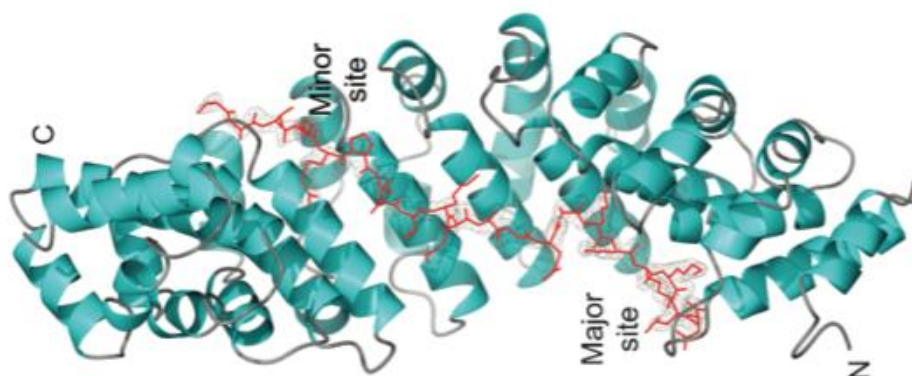


Figura 6. Modelo de Importina- α complexada com cNLS de Nucleoplasmina (Marfori *et al.*, 2012). Importina- α (em verde) com truncamento no sítio IBB realizando ligações com resíduos da cNLS bipartida (vermelho) da proteína Nucleoplasmina.

Resíduos da Importina- α e da NLS da Nucleoplasmina serão apresentados com nome da estrutura seguidos por nome e posição do resíduo. Por exemplo, ^{Npl}K162 que representa a lisina na posição 162 da NLS da Nucleoplasmina.

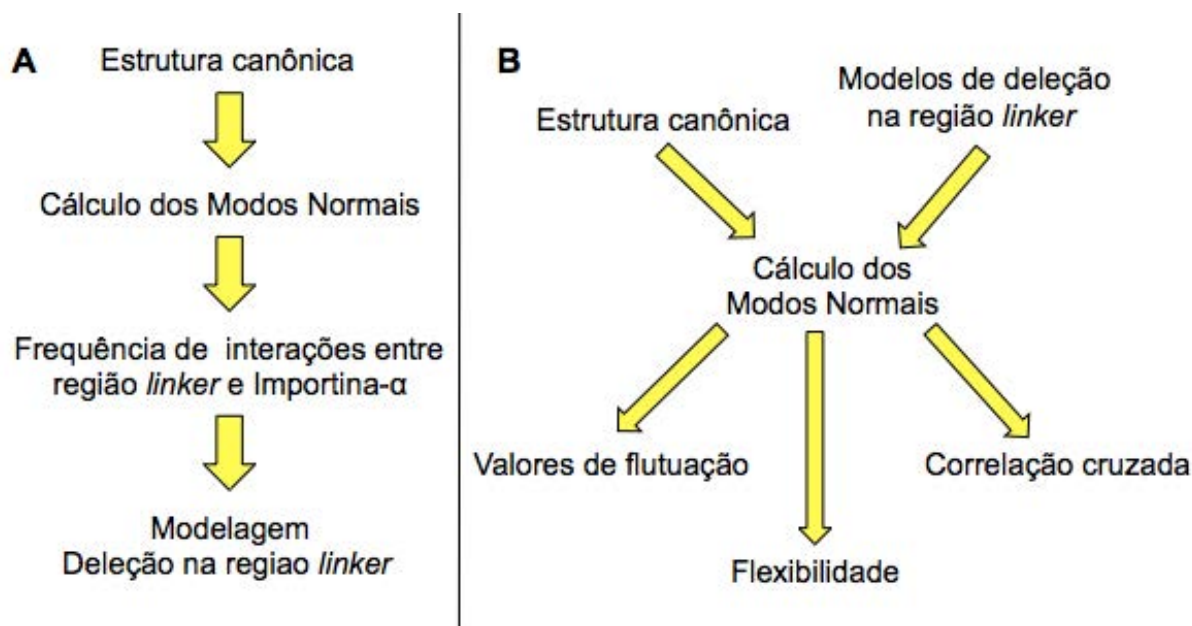


Figura 7. Fluxograma do planejamento de atividades. **A.** Análise da frequência de interações hidrofóbicas, ligações de hidrogênio e pontes salinas para seleção dos resíduos que serão deletados na região *linker*. **B.** Análise dos efeitos da deleção pontual na região *linker* nas interações com Importina- α e para importação nuclear através dos cálculos dos valores de flutuação dos resíduos, correlação cruzada e flexibilidade da Importina- α .

3.2. Análises de Modos Normais e estruturas ao longo do modo

Os modos normais foram gerados utilizando o programa CHARMM v.37b1 (Brooks *et al.*, 2009) seguindo protocolo apresentado por Huang e Mackerell (2013). Na primeira etapa do protocolo, realizamos a minimização de energia no vácuo com campo de força charmm36, utilizando restrições harmônicas e o método de gradientes conjugados. Após a minimização de energia, foi efetuada a etapa de diagonalização e cálculo dos primeiros 50 modos normais de baixa frequência utilizando o módulo VIBRAN do programa CHARMM. Os primeiros seis modos foram descartados, pois se tratavam de movimentos de rotação e translação da molécula, que não representam movimentos de interesse biológico. Após esse procedimento, utilizamos os 44 modos restantes para gerar as estruturas ao longo de cada um deles utilizando o módulo VMOD do programa CHARMM, seguindo metodologia semelhante a que foi descrita por (Philot *et al.*, 2013). A amplitude do deslocamento foi de -3 a +3Å de MRMS (*mass-weight root mean square*) ao longo do modo

normal, utilizando um passo de 0,1Å. A cada passo, realizamos uma dinâmica molecular curta de 10 picossegundos (ps), seguida de uma minimização de energia, que ao final, gerou uma estrutura. Ao todo geramos 61 estruturas por modo normal.

3.3. Frequência de interações entre os resíduos da Importina- α e a cNLS da Nucleoplasmina e valores de flutuação

A energia de restrição foi calculada, segundo o módulo MMFP (*Miscellaneous Mean Field Potential*) do programa CHARMM utilizando os dados das estruturas ao longo de cada modo normal. O módulo MMFP nos permitiu gerar um gráfico na forma de mapa de calor, com a variação de energia de restrição das estruturas ao longo dos modos, no programa MatLab. O mapa de calor mostra em tons de laranja e vermelho estruturas com energia de restrição elevada (0.8-1 kcal/mol). Seleccionamos como estruturas favoráveis para análise as que apresentavam energia menores que 0.8 kcal/mol, em tons de azul no mapa de calor e com estas estruturas calculamos a frequência das interações hidrofóbicas, pontes salinas e ligações de hidrogênio.

A frequência das pontes salinas e ligações de hidrogênio foram determinadas utilizando o programa VMD (*Visual Molecular Dynamics*) (Humphrey *et al.*, 1996). Para pontes salinas utilizamos como critério para ocorrência de ligações, distância menor ou igual a 3,5 Å entre os resíduos. Nas ligações de hidrogênio, também utilizamos um critério de 3,5 Å de distância entre os resíduos e adicionamos o critério de desvio de angulação entre o doador e acceptor de hidrogênio como sendo menor ou igual a 60 graus. A frequência das interações hidrofóbicas, por sua vez, foi realizada no programa CHARMM com adoção de distância menor ou igual a 4 Å entre os átomos mais próximos das cadeias laterais hidrofóbicas. A frequência de ocupação de cada tipo de interação foi calculada a fim de determinarmos os contatos mais estáveis e instáveis na interface do complexo.

Igualmente utilizando os modos seleccionados pelo módulo MMFP, fizemos o cálculo dos valores de RMSF (*root mean square fluctuation*) para os carbonos- α da NLS da Nucleoplasmina e da proteína Importina- α no programa Wordom (Seeber *et al.*, 2007). Em seguida, calculamos os valores de correlação de movimento dos resíduos na interface do complexo proteína-peptídeo. Igualmente ao

módulo MMFP, geramos um mapa de calor no programa MatLab com os valores de correlação.

3.4. Modelagem das estruturas com deleção pontual na região linker da NLS da Nucleoplasmina

Os modelos com deleção no *linker* foram gerados com base no método de modelagem comparativa do programa Modeller v9.11 (Eswar *et al.*, 2006). A modelagem comparativa permite utilizar um modelo estrutural gerado a partir de métodos experimentais (e.g. cristalografia de raios-X) para a predição de proteínas de interesse que tenham alta similaridade com esta estrutura molde. Utilizando como molde (*template*) a estrutura da sequência canônica da NLS da Nucleoplasmina ligada à Importina- α (PDB ID 3UL1), realizamos a modelagem das sequências da cNLS com as deleções nos resíduos ^{Npl}A159 e ^{Npl}K162. O alinhamento das sequências que foram utilizadas para modelagem se encontram no Apêndice 1.

Foram gerados dez diferentes modelos que foram avaliados segundo sua energia livre. O programa Modeller realiza durante a modelagem algumas etapas de dinâmica molecular curta seguida de minimização de energia, como tentativa de otimizar conformacionalmente o modelo gerado. Dentre os modelos gerados, realizamos a análise da qualidade estereoquímica através do website Rampage Ramachandran (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) a fim de selecionarmos o melhor modelo para o cálculo de modos normais. A análise de qualidade normalmente é determinada por um conjunto de fatores como: comprimento de ligações, planaridade das ligações peptídicas, planaridade dos anéis, ângulos de torção nas cadeias principal e laterais, quiralidade, impedimento estérico, e energia funcional. Com a estrutura do modelo gerado validada, escolhemos aquele de menor energia e realizamos alinhamento estrutural no programa Pymol com a estrutura molde a fim de checar a integridade da estrutura final.

3.5. Ângulos entre α -hélices H3 de ARMs vizinhos e raio de curvatura da Importina- α

Para a nossa molécula em estudo, os três primeiros modos normais apresentam movimentos básicos como capacidade de representação da movimentação da molécula como todo. Realizamos o cálculo dos ângulos entre α -hélices H3 de ARMs vizinhos, utilizando um script disponível para Pymol (<https://pymolwiki.org/index.php/AngleBetweenHelices>), sendo que para cada par de α -hélices foram analisados os vetores ao longo dos átomos de carbono- α . Já o raio de curvatura da Importina- α foi calculado segundo a metodologia descrita por Geraldo *et al.* (2016), no qual três vértices são selecionados na estrutura da proteína a fim de calcularmos o raio da circunferência que intersecta esses três pontos.

3.6. Dinâmica Molecular

A dinâmica molecular foi realizada no programa GROMACS v.4.5 utilizando o campo de força charmm36. Inicialmente, as estruturas foram colocadas individualmente em caixas de simulação de formato cúbico em que foram adicionadas moléculas de água (modelo TIP3P) e as cargas do sistema foram neutralizadas com íons Na^+ e Cl^- . A minimização de energia foi gradativa e seguiu quatro etapas: (i) limitação do movimento de toda proteína e peptídeo para possibilitar a minimização de energia das moléculas de solvente do sistema pelo método *steepest descent*; (ii) limitação do movimento das cadeias principais da proteína e do peptídeo para a minimização de energia das cadeias laterais pelo método *steepest descent*; (iii) minimização de energia do sistema pelo método *steepest descent*, sem restrição de movimento; (iv) minimização do sistema, sem restrição de movimento, pelo método de gradientes conjugados. Esta metodologia de minimização é uma tentativa de garantir uma estrutura inicial razoável em termos de geometria e orientação do solvente.

Com as estruturas minimizadas, realizamos a equilibração do sistema. Esta etapa é muito importante, pois nela adicionamos condições de temperatura e pressão ao sistema. Consequentemente, precisamos garantir que após adição destas condições, o solvente e os íons ao redor da proteína estejam equilibrados, antes de prosseguirmos com a simulação de dinâmica molecular. A equilibração

seguir duas etapas. A primeira etapa envolveu a adição de uma condição NVT (número constante de partículas, volume e temperatura), em que a temperatura adicionada ao sistema foi de 300 K (Kelvin) e o sistema simulado nesta condição por 100ps. A segunda etapa envolveu a adoção de uma condição NPT (número constante de partículas, pressão e temperatura), em que a pressão foi adicionada ao sistema e se manteve relativamente constante oscilando próximo ao valor de 1 bar. Esta etapa também foi simulada por 100 ps.

Após a equilibração, o sistema foi simulado por 100 ns. O tempo de simulação foi escolhido com base na capacidade de processamento da nossa arquitetura computacional. Ao final, as simulações de dinâmica molecular geradas foram visualizadas no software VMD (*Visual Molecular Dynamics*) (Humphrey *et al.*, 1996).

4. Resultados

4.1. *NpI*K162 é contato de característica estável na região linker e *NpI*A159 é contato instável.

A estrutura de Marfori *et al.* (2012) utilizada como modelo de estudo, apresenta a Importina- α com truncamento no sítio IBB (resíduos 70 a 529) complexada com a cNLS da Nucleoplasmina. O truncamento no sítio IBB permite maior afinidade entre proteína e Importina- α através da diminuição do conflito no sítio principal entre o domínio IBB e cNLS. A Nucleoplasmina foi escolhida como modelo de estudo para o nosso projeto porque é uma proteína chaperona altamente caracterizada na literatura, possuindo uma NLS bipartida [*NpI*GSAVKRPAATKKAGQAKKKKLD¹⁷²] com região *linker* constituída por 10 resíduos. Para a análise das interações entre Importina- α e cNLS com região *linker* menor que dez resíduos, estabelecemos a geração de duas estruturas com deleção na região *linker*, uma com deleção de resíduo com contato forte e estável entre cNLS e Importina- α , e outra estrutura, com deleção de resíduo de contato fraco e, mais instável, gerando estruturas com regiões *linker* de apenas 9 resíduos. Assim, caso as interações do complexo fossem desestabilizadas, especialmente nos sítios

de ligação principal e secundário, poderíamos avaliar se tal efeito foi devido à ausência de um contato forte ou pela redução no comprimento do *linker*.

Para a seleção dos resíduos que seriam deletados, primeiramente, foram calculados os modos normais dos 50 primeiros modos de baixa frequência no modelo canônico, seguidos da geração das estruturas ao longo de cada modo. A escolha de estruturas para análise das interações foi feita com base em energia de restrição, segundo o módulo MMFP do programa CHARMM. Estruturas energeticamente desfavoráveis – valores iguais ou maiores que 0.8 kcal/mol, representadas nas cores laranja e vermelho – foram descartadas da análise (Figura 8). A partir do gráfico, selecionamos dos modos 7 a 27 como estruturas energeticamente favoráveis no modelo canônico e analisamos a frequência da formação de pontes salinas, ligações de hidrogênio e interações hidrofóbicas que são os principais tipos de interações presentes na interface dos complexos de importação nuclear Importina- α :NLS (Fontes *et al.*, 2000).

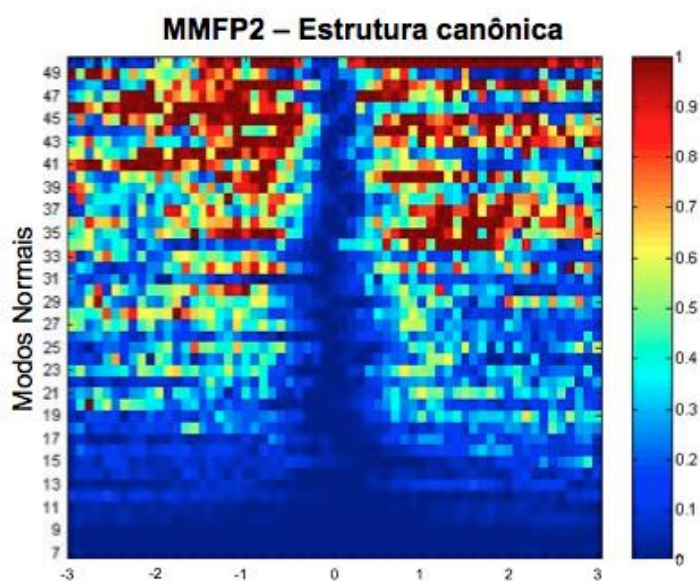


Figura 8. Energias de restrição da estrutura canônica para os modos normais 7 a 50. Energias de restrição calculadas segundo o módulo MMFP para as estruturas ao longo dos modos dos modos normais 7 a 50. Cores mais leves representam energia de restrição mais baixa enquanto que cores fortes representam energias de restrição mais elevadas. Modos normais foram selecionados a partir da presença de energias de restrição maiores ou iguais a 0.8 kcal/mol.

Para a escolha de um resíduo com contato estável na região *linker* para deleção, selecionamos aqueles que apresentaram frequência de interação maior que 50% (Apêndice Figura A2). Dos contatos realizados pela região *linker* com a Importina- α com frequência de interações maiores que 50%, escolhemos por utilizar a ^{Npl}K162 como contato estável pois realizava maior número contatos hidrofóbicos (^{Imp- α} W273, ^{Imp- α} P311 e ^{Imp- α} R315 e ainda apresentou ligação de hidrogênio com o ^{Imp- α} N361 (Figura 9, Apêndice Figura A2). Em contrapartida, para a escolha do resíduo com contato instável na região *linker* para a deleção, selecionamos dentre os resíduos que apresentaram frequência de interação menores que 50%. No caso, escolhemos a ^{Npl}A159 que, apesar de realizar contatos fracos, interagiu por ligações de hidrogênio com a Importina- α , nos contatos ^{Imp- α} E396, ^{Imp- α} K240, ^{Imp- α} K392 e ^{Imp- α} W357 como observado na Tabela A3 do Apêndice.

Contatos estáveis

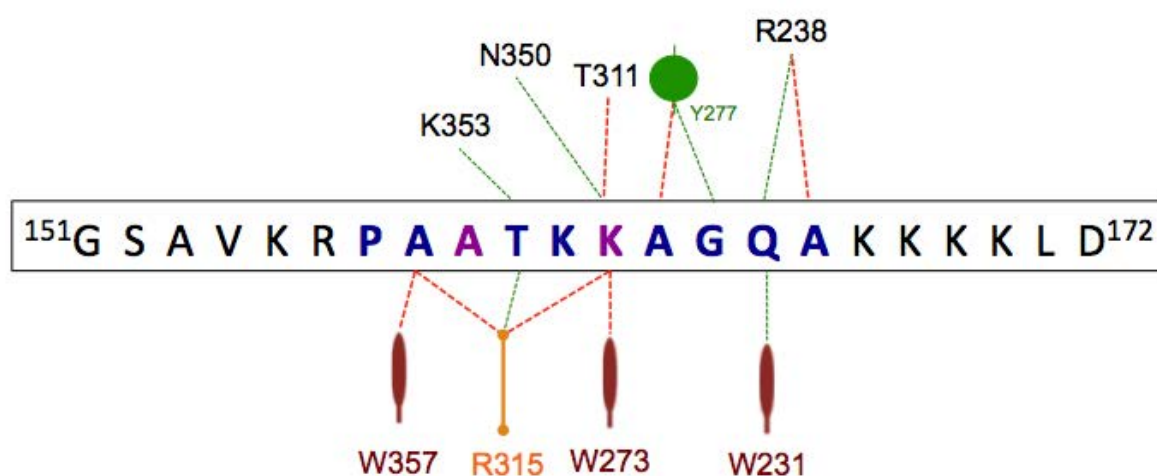


Figura 9. Interações entre resíduos da região *linker* da NLS da Nucleoplasmina com Importina- α com frequência maiores que 50%. Região *linker* destacada em azul e negrito. No linker estão destacados em roxo os resíduos deletados (^{Npl}A159 e ^{Npl}K162). Ligações de hidrogênio estão representadas por retas pontilhadas em verde. Retas pontilhadas em vermelho representam os contatos hidrofóbicos entre os resíduos.

4.2. Interações entre resíduos da Importina- α e NLS da Nucleoplasmina nos modelos de deleção da ^{Npl}K162 e ^{Npl}A159

Após selecionarmos os resíduos ^{Npl}A159 e ^{Npl}K162 para deleção realizamos a modelagem das duas estruturas. Com o modelo de menor energia de cada estrutura, resolvemos analisar as interações presentes entre as cNLSs com deleção e a Importina- α para determinar se a deleção de um resíduo no *linker* alterava as interações com a Importina- α . Seguindo o *script* realizado com a estrutura canônica, primeiramente calculamos os modos normais e geramos as estruturas ao longo de cada modo. Em seguida, analisamos as interações nas estruturas que foram selecionadas com base nos valores de energias de restrição. Para o modelo de deleção da ^{Npl}A159, encontramos os modos normais de energias mais favoráveis do modos normais 7 a 21 e para o modelo de deleção da ^{Npl}K162, os modos normais 7 a 25 (Figura 10).

Com a escolha dos modos com estruturas que apresentaram energia favorável, calculamos a frequência das interações hidrofóbicas, pontes salinas e ligações de hidrogênio para cada modelo (Apêndice Figura A4 e A5). Selecionamos os contatos com frequência maiores que 50% e comparamos aos obtidos no modelo canônico. Ambas as deleções apresentaram em sua maioria, interações semelhantes aos observados no modelo canônico. Entretanto, dentre as diferenças nos contatos realizados observamos que no modelo de deleção da ^{Npl}A159 houve diminuição da presença de contatos hidrofóbicos, por exemplo, os contatos ^{Npl}K155 com ^{Imp- α} A364 no sítio secundário; ^{Npl}K167 com ^{Imp- α} D192 e ^{Npl}K170 com ^{Imp- α} W184 no sítio principal e ^{Npl}A158 com ^{Imp- α} R315, ^{Npl}K162 com ^{Imp- α} W357 e ^{Imp- α} T311 na região *linker* estiveram ausentes no modelo de deleção da ^{Npl}A159. Em contrapartida, houve aumento no número ocorrências de pontes salinas, como formação das pontes ^{Npl}K155 com ^{Imp- α} D160 no sítio secundário, ^{Npl}K162 com ^{Imp- α} E354 na região *linker* e ^{Npl}K168 com ^{Imp- α} D270 no sítio principal.

Em relação ao modelo de deleção da ^{Npl}K162 observamos que houve uma mudança nos contatos do tipo ligação de hidrogênio com relação ao modelo canônico. Por exemplo, deixamos de obter interações com frequência maior que 50% para os contatos ^{Npl}K155 e ^{Imp- α} G323, ^{Npl}R156 e ^{Imp- α} S360 no sítio secundário; mas obtivemos contato da ^{Npl}R156 com ^{Imp- α} E396. No *linker* não observamos os contatos ^{Npl}T160 e ^{Imp- α} R315; ^{Npl}G164 e ^{Imp- α} Y277; ^{Npl}Q165 e ^{Imp- α} W231, mas

obtivemos contatos entre ^{Npl}K161 e ^{Imp-α}R315, ^{Npl}Q165 e ^{Imp-α}R238 e ^{Npl}A166 com ^{Imp-α}W231. Por fim, no sítio principal não houve alta frequência de contatos entre ^{Npl}K167 e ^{Imp-α}T155, ^{Npl}K169 e ^{Imp-α}R106, mas formaram-se os contatos entre ^{Npl}K168 e ^{Imp-α}S149, ^{Npl}K169 com ^{Imp-α}E107 e ^{Imp-α}N146.

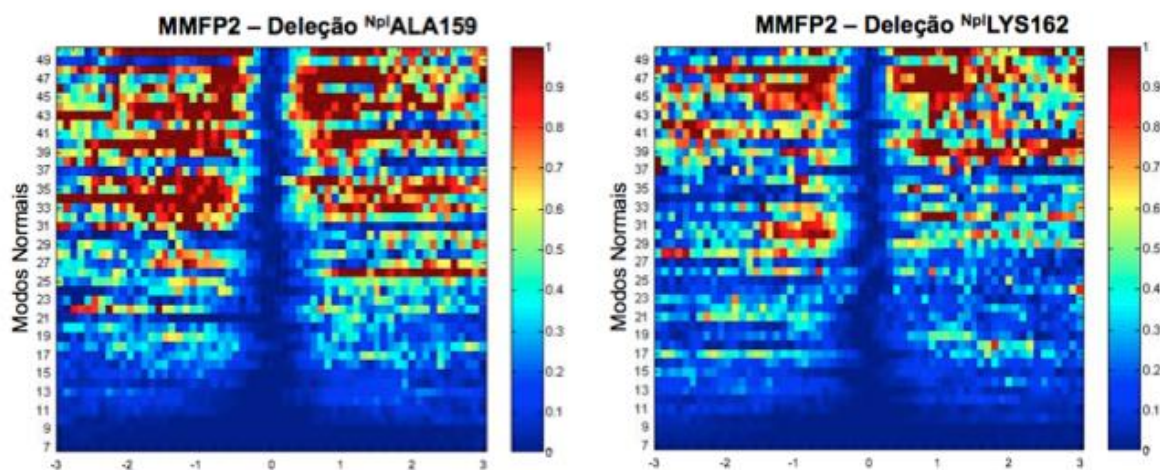


Figura 10. Energias de restrição dos modelos de deleção da ^{Npl}A159 e ^{Npl}K162 para os modos normais 7 a 50. Energias de restrição calculadas segundo o módulo MMFP para as estruturas ao longo dos modos dos modos normais 7 a 50 dos modelos de deleção da ^{Npl}A159 e ^{Npl}K162 respectivamente. Cores mais leves representam energia de restrição mais baixa enquanto que cores fortes representam energias de restrição mais elevadas. Modos normais foram selecionados a partir da presença de várias energias de restrição maiores ou iguais a 0.8. O modelo de deleção da ^{Npl}A159 possui modos normais de energias mais favoráveis do modos normais 7 a 21 e o modelo de deleção da ^{Npl}K162, os modos normais 7 a 25.

4.3. Valores de flutuação entre resíduos da Importina-α e NLS da Nucleoplasmina

Também objetivando a comparação das mudanças entre os contatos realizados pela região *linker* e pelos sítios principal e secundário das estruturas canônica e com as deleções – ^{Npl}A159 e ^{Npl}K162, realizamos as análises da flutuação dos resíduos da NLS utilizando os valores de RMSF dos carbonos-α da

NLS calculados para os modos normais (Figura 11). Também calculamos os valores de correlação de movimento entre os resíduos da interface do complexo, sendo este resultado exibido na forma de mapa de calor com base nos valores de correlação (Figura 12).

Os resultados da análise das flutuações dos resíduos demonstram pequenas variações entre a interação dos resíduos da cNLS e Importina- α dos modelo canônico e de deleções. É possível observar diminuição na flutuação dos resíduos do sítio secundário e região *linker* em ambas estruturas com deleção. Contudo, tal diminuição na flutuação dos resíduos é mais expressiva no modelo com deleção da ^{Npl}K162 que na deleção da ^{Npl}A159. Estas diminuições nos valores de flutuação dos resíduos poderia indicar aumento na adesão entre NLS e Importina- α nessas regiões.

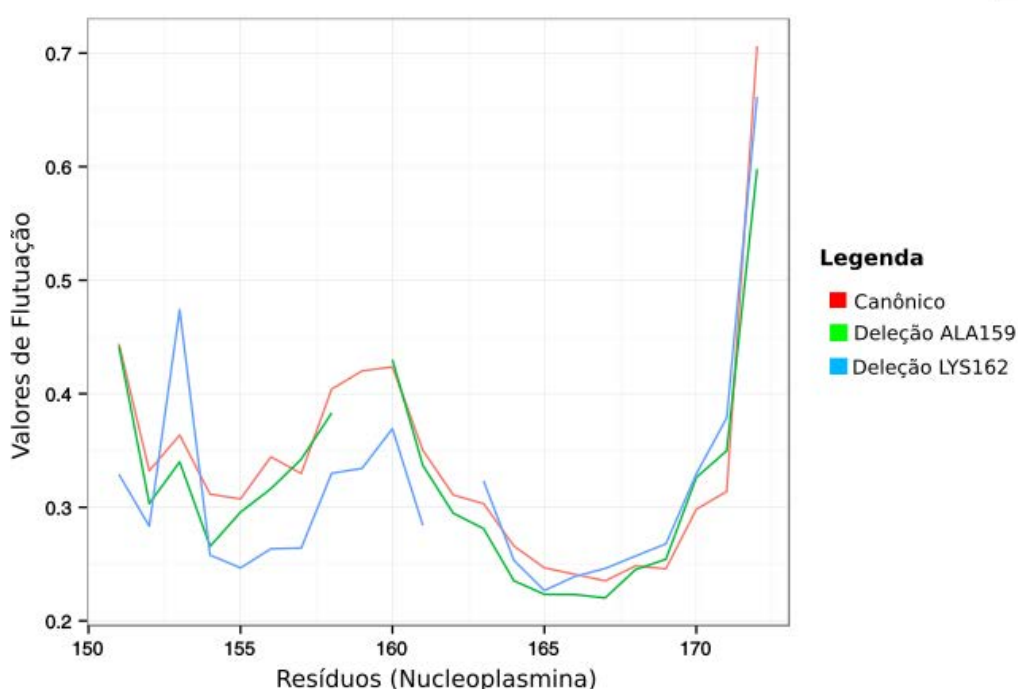


Figura 11. Valores de flutuação entre resíduos da Importina- α e NLS da Nucleoplasmina. Valores de flutuação da estrutura canônica, modelo com deleção da ^{Npl}A159 e com deleção da ^{Npl}K162 está representado no gráfico em vermelho, verde e azul respectivamente. Com relação à estrutura canônica, não houve variações expressivas entre a interação dos resíduos da cNLS e Importina- α no modelo com deleção da ^{Npl}A159, mas no modelo com deleção da ^{Npl}K162 houve diminuição na flutuação dos

resíduos na região do sítio secundário e *linker*, e aumento na flutuação dos resíduos NplA153, região anterior ao sítio secundário.

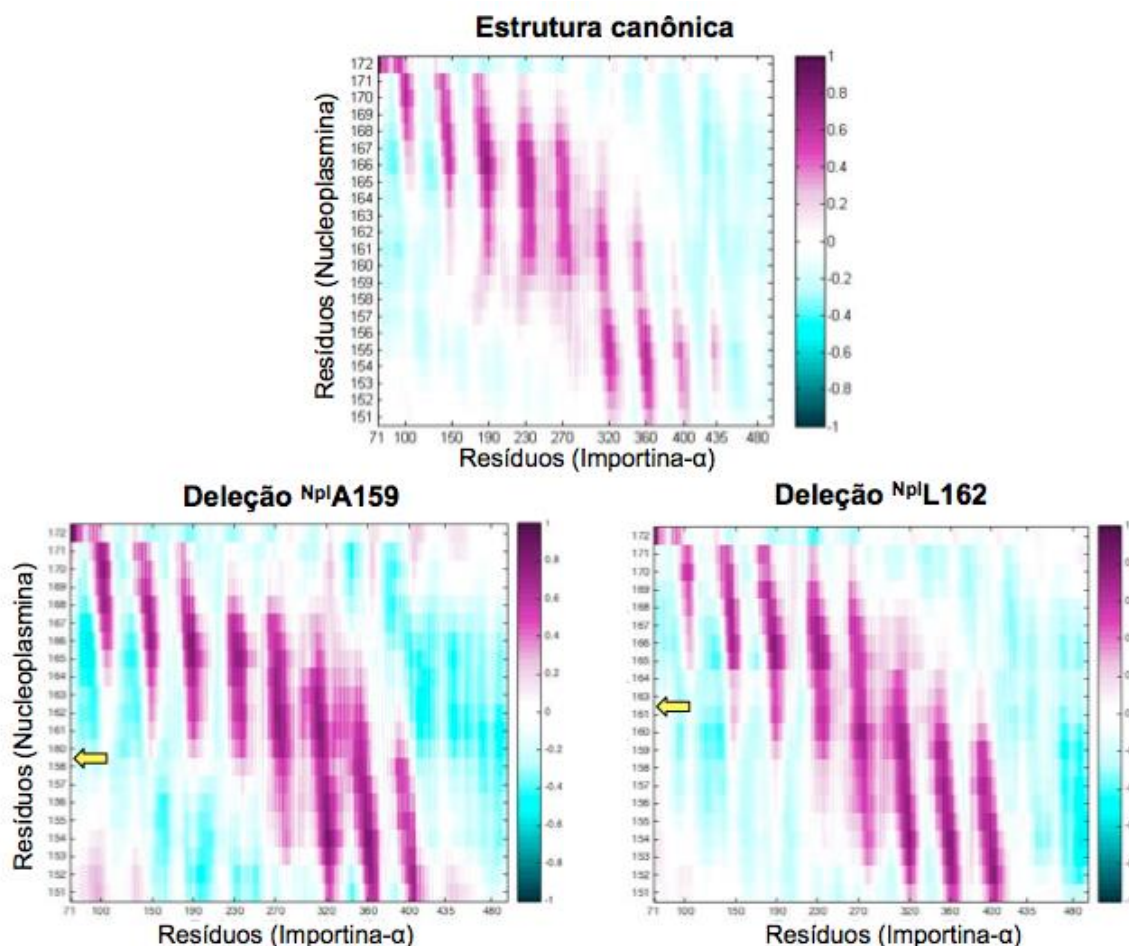


Figura 12. Correlação entre resíduos da Importina- α e NLS da Nucleoplasmina. Os gráficos de correlação mostram a relação entre os resíduos da cNLS com os resíduos da Importina- α . Os resultados apresentados em tons de roxo e em azul demonstram as correlações positivas e negativas, respectivamente, sendo que quanto mais intensa a cor, maior o valor de correlação. Setas amarelas indicam o local do resíduo deletado da estrutura.

Para os resultados de correlação, observamos que, congruente aos valores de flutuação, nossos dados mostram correlações positivas – representadas em tons de roxo nos mapas de calor – que parecem indicar maior interação entre as regiões *linker* e sítio secundário em ambos os casos de deleção em relação ao modelo canônico.

4.4. Diferenças entre Modos Normais e Dinâmica Molecular segundo a análise de correlação dos resíduos da Importina- α e da NLS

Semelhante às análises apresentadas por Geraldo *et al.* (2016), realizamos também a simulação de dinâmica molecular para as nossas estruturas com deleção e calculamos as correlações a fim de comparação destes resultados com os nossos dados de modos normais (Figura 13).

A comparação dos resultados de correlação da estrutura com deleção da ^{Npl}A159 mostram baixa correlação entre os resíduos do sítio principal na dinâmica molecular comparado aos resultados obtidos nos modos normais, mas a correlação entre os resíduos do sítio secundário e da região *linker* permaneceram semelhantes. Agora, em relação à estrutura com deleção da ^{Npl}K162, pudemos observar que a correlação da dinâmica molecular mostrou uma diminuição na relação entre resíduos ao longo de toda a NLS, mas com destaque na região *linker*. Tal diferença de resultados da análise de correlação por Modos Normais e Dinâmica Molecular pode estar relacionada a capacidade dos modos normais de representar uma tendência movimento de alta amplitude da estrutura e, assim, as possíveis interações ocorrendo nesse movimento, ao caso que a dinâmica molecular permite a visualização da conformação da estrutura ao longo de um determinado período de tempo. Entretanto, por conta de seu alto custo computacional, o período de tempo possível de ser calculado e analisado é curto, o que limita a dinâmica molecular à observação de mudanças de baixa amplitude, assim não possuindo capacidade da representação de outras possíveis interações como visualizado pelo modos normais.

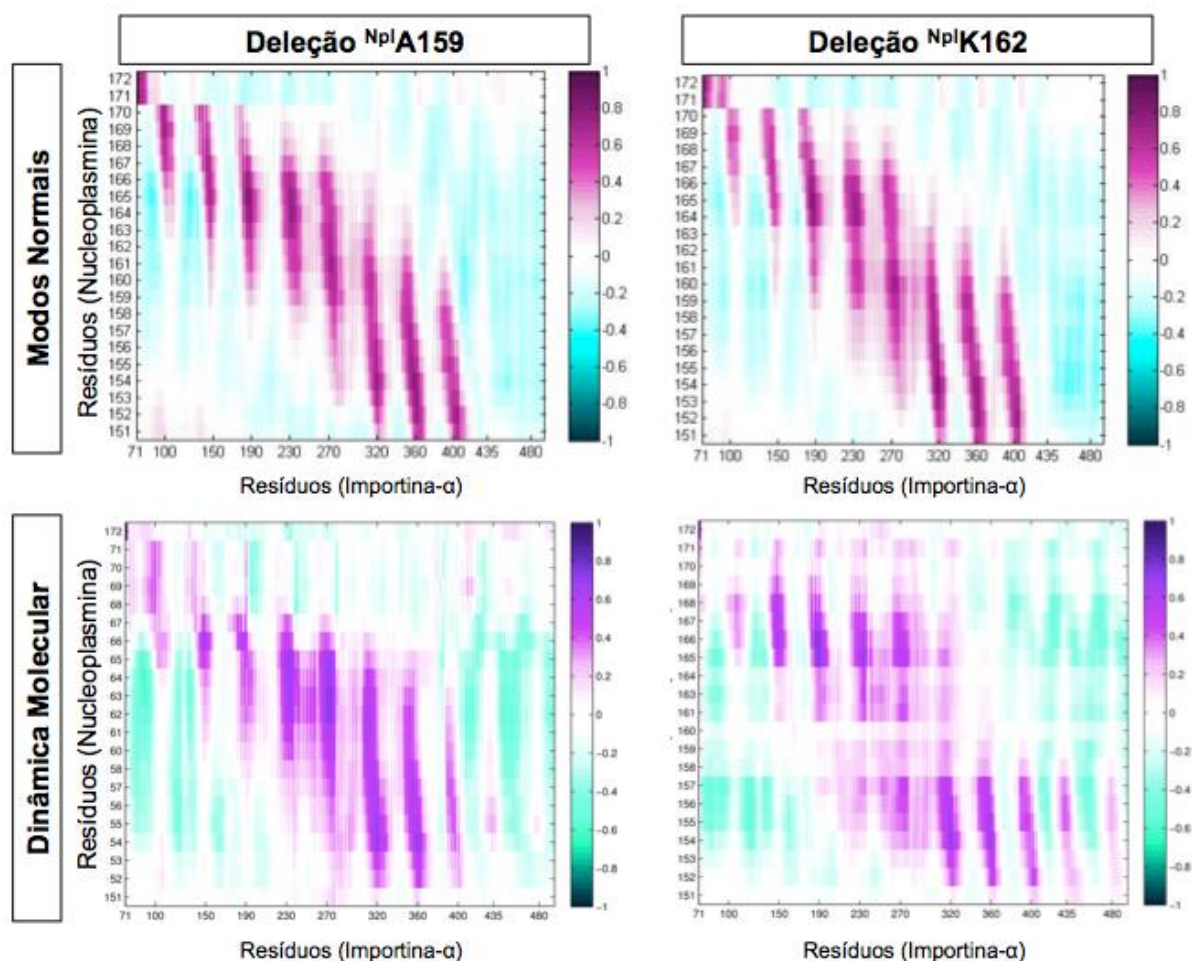


Figura 13. Correlação dos resíduos, calculada para os modos normais e a dinâmica molecular nas estruturas com deleção da ^{Npl}A159 e ^{Npl}K162. Gráfico com os mapas de calor de correlação gerados para a análise de modos normais e dinâmica molecular entre os resíduos da NLS da Nucleoplasmina e os resíduos da Importina- α . Os resultados de correlação gerados para a dinâmica molecular mostram menor interação entre os resíduos da NLS com a Importina- α no sítio principal do modelo com deleção da ^{Npl}A159 e ao longo de toda a NLS no modelo com deleção da ^{Npl}K162, principalmente na região *linker*.

4.5. Análises dos movimentos da Importina- α

Conti *et al.* (1998) apresentou cálculos que mostram que existe um movimento de rotação de aproximadamente 30 graus entre unidades estruturais adjacentes causando a torção solenoide. Desta forma, com o objetivo de analisar se a cNLS bipartida com *linker* de nove resíduos alterava a flexibilidade da Importina- α , resolvemos analisar geometricamente a movimentação da Importina- α . Utilizando os

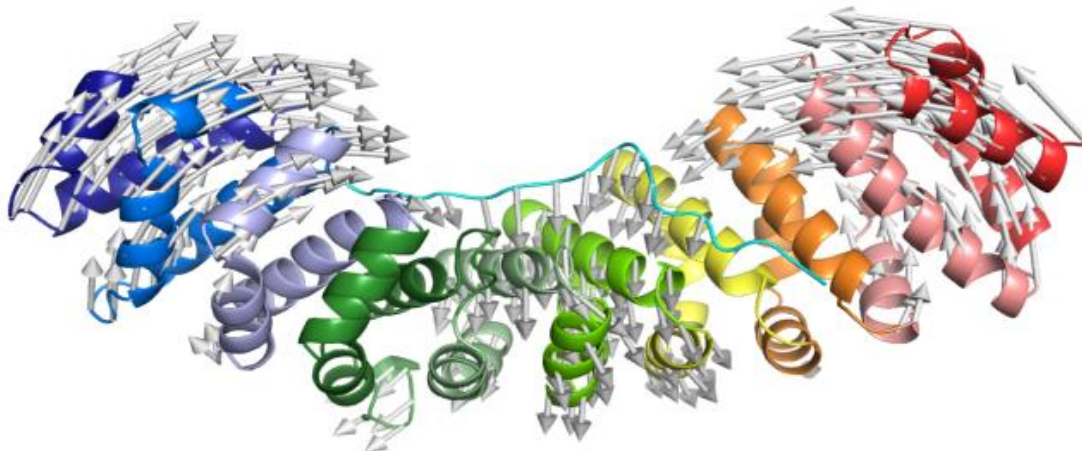
modos normais de menor frequência (modos 7, 8 e 9) que descrevem os movimentos de torção, dobramento e dobramento lateral com torção da Importina- α (Figura 14), realizamos os cálculos dos ângulos entre α -hélices H3 de repetições ARMs vizinhas e calculamos o raio de curvatura da Importina- α .

Assim, para os movimentos de torção, analisamos a flexibilidade da Importina- α segundo os ângulos formados pelas α -hélices H3 das repetições ARM vizinhas. Como dito anteriormente, as α -hélices H3 são responsáveis pelos contatos da Importina- α com a cNLS e pela curvatura côncava da Importina- α . A Figura 15 apresenta os gráficos com os ângulos formados pelas α -hélices das repetições ARMs vizinhas em relação as estruturas ao longo do modo dos ARMs que apresentaram mudança na flexibilidade da Importina- α . Os gráficos com os ângulos de todos os ARMs pode ser encontrado na Figura A6 do Apêndice. Os gráficos da figura 15A, B e C apresentam os ângulos da entre as α -hélices H3 durante o movimento de dobramento lateral com torção representado pelo modo 7 na estrutura canônica e modos 8 nos modelos de deleção. Dos resultados podemos observar diminuição da flexibilidade da Importina- α na região *linker*. Obtivemos diminuição do ângulo entre as α -hélices H3 dos ARMs 4 e 5 nos modelos de deleção quando comparados ao canônico. Houve também diminuição entre os ângulos formados pelas α -hélices H3 dos ARMs 5 e 6 nos modelos com deleção, principalmente no modelo com deleção da ^{Npl}A159. Além disso, ocorreu um aumento na flexibilidade entre os ângulos dos ARMs 6 e 7 no modelo com deleção da ^{Npl}K162 em relação a estrutura canônica, alterações ausentes no modelo de deleção da ^{Npl}A159. Igualmente, foi obtido maior flexibilidade no ângulo entre os ARMs 8 e 9 da região posterior ao sitio secundário, no modelo com deleção da ^{Npl}A159 em relação ao modelo canônico.

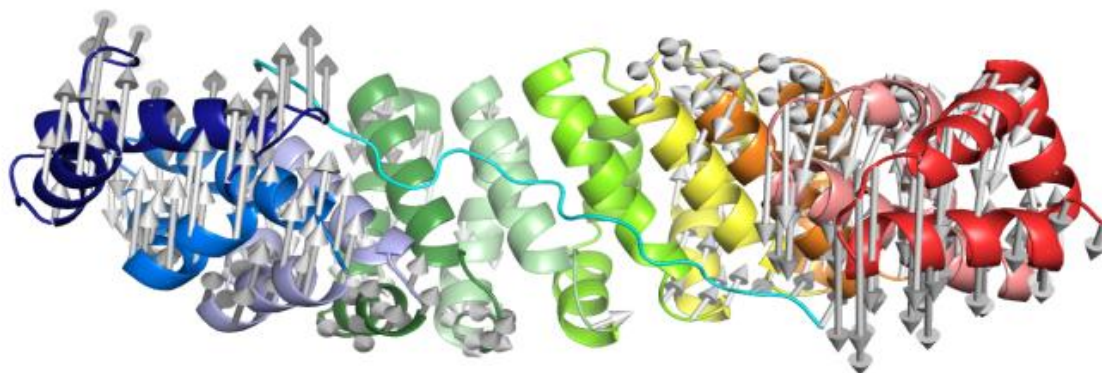
Os gráficos da figura 15D, E e F apresentam os ângulos entre as α -hélices H3 durante o movimento de torção representados pelos modos 9 de cada estrutura. Destes resultados observamos que nos modelos com deleção houve diminuição da flexibilidade entre os ângulos dos ARMs 2 e 3, demonstrando diminuição da flexibilidade no sitio principal. Também é observável maior flexibilidade entre os ângulos dos ARMs 5 e 6 e ARMs 6 e 7 – região *linker* – nos modelos com deleção em relação ao modelo canônico.

Agora em relação ao movimento de dobramento, analisamos o raio de curvatura da Importina- α (Figura 16) e observamos maior variação no raio de curvatura no modelo com deleção da ^{Npl}A159 em relação aos demais modelos.

Dobramento



Torção



Dobramento lateral com torção

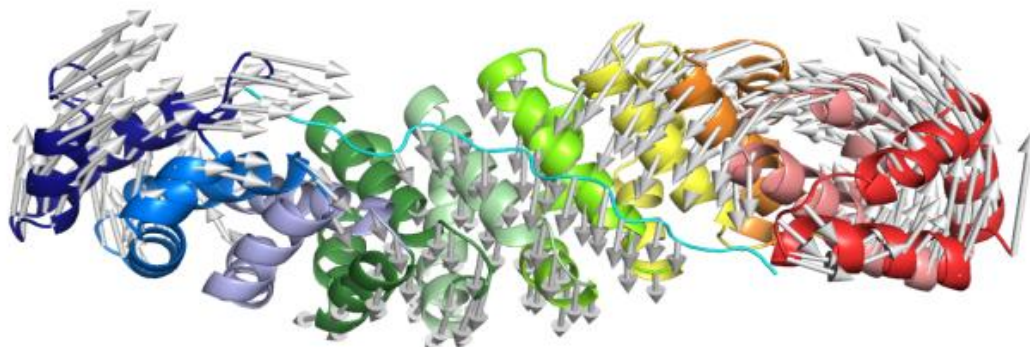


Figura 14. Representação no Pymol da movimentação da Importina- α . A figura apresenta as movimentações da Importina- α nos modos de baixa frequência 7, 8 e 9: movimento de dobramento (“abre-e-fecha”); movimento de torção e movimento de dobramento lateral com torção. Peptídeo em azul ciano representa a NLS da Nucleoplasmina ligada à Importina- α .

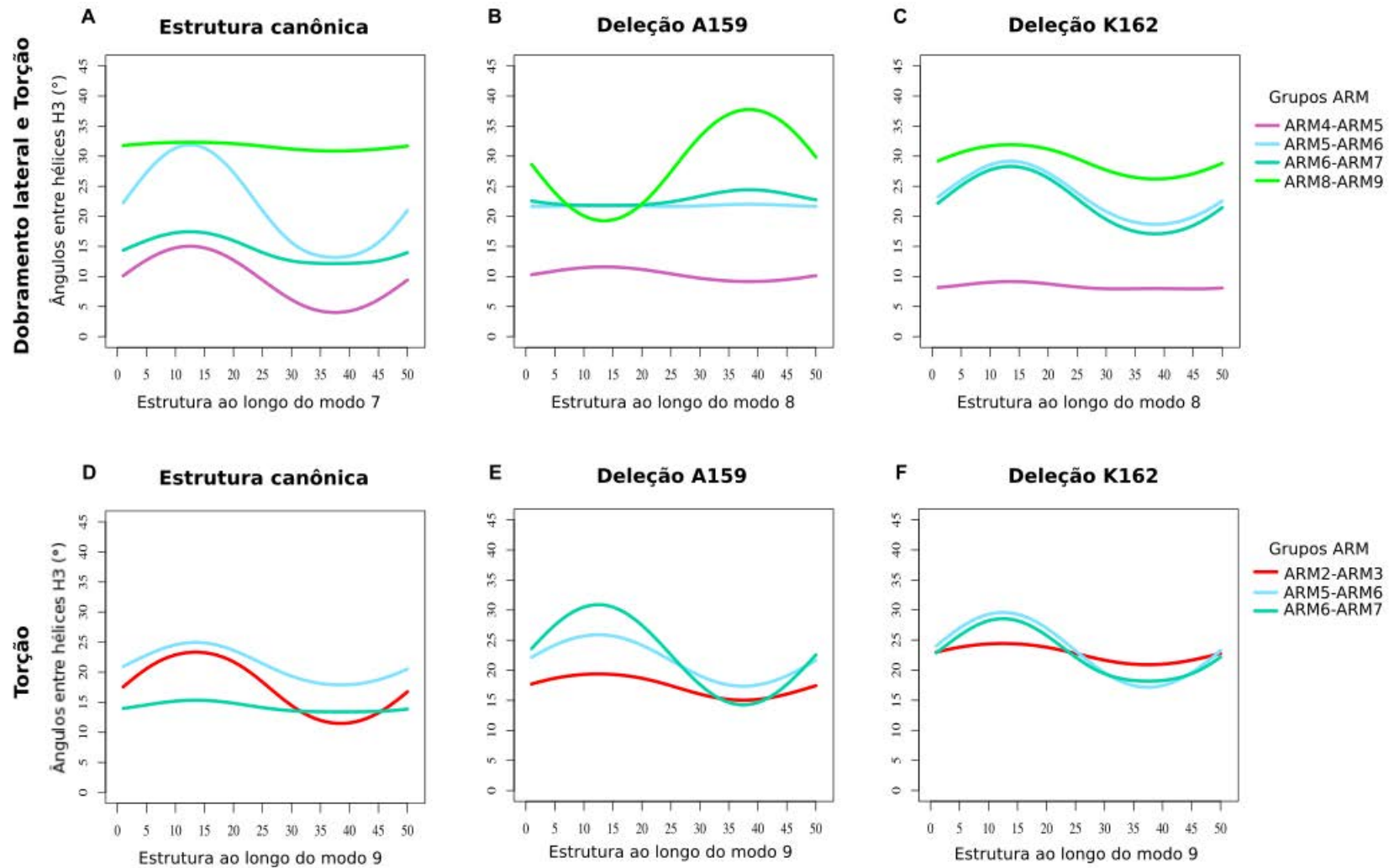


Figura 15. Ângulo entre α -hélices H3 de repetições ARM vizinhas da Importina- α ao longo dos modos normais com movimento de torção. Os gráficos apresentam a flexibilidade da Importina- α calculados através dos ângulos formados pelas α -hélices H3 de repetições ARM vizinhas que sofreram variação entre as estruturas com deleção pontual e o modelo a variação angular durante o

movimento de dobramento lateral com torção representado pelo modo 7 na estrutura canônica e modos 8 nos modelos com deleção. Gráficos 15D, E e F apresentam a variação angular durante o movimento de torção representados pelos modos 9 de cada estrutura.

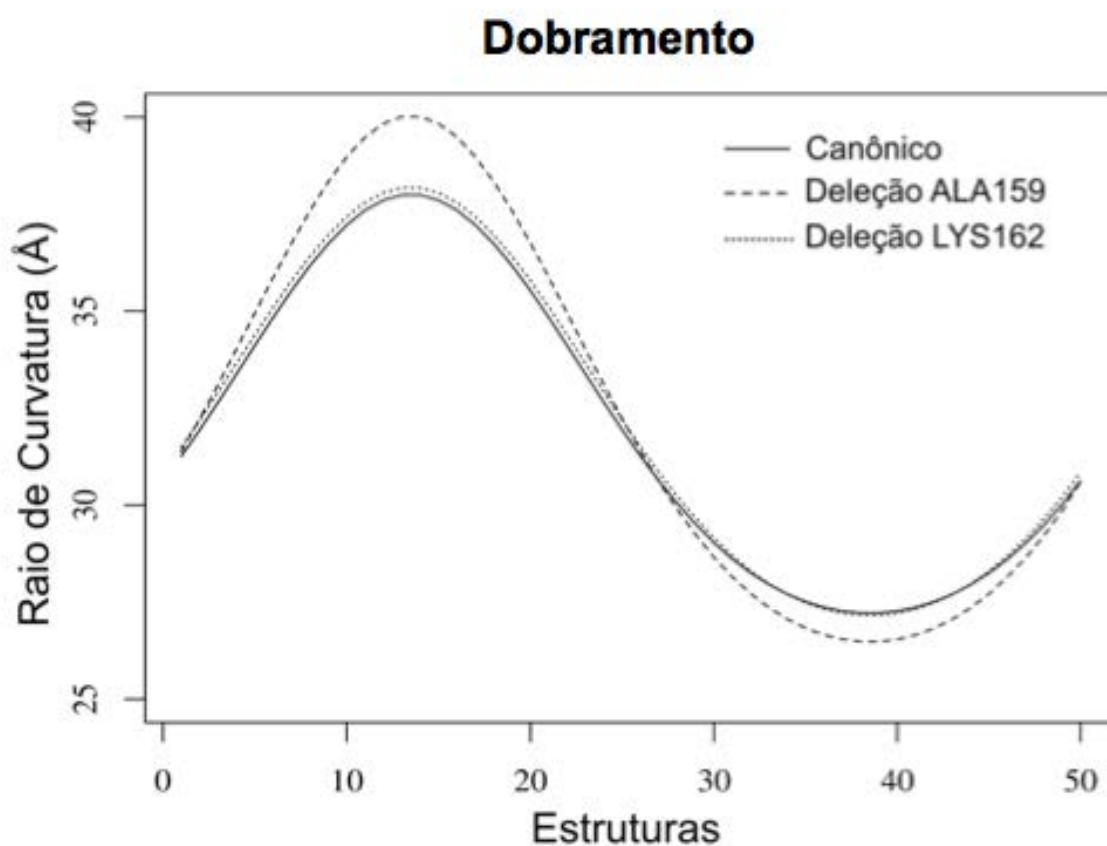


Figura 16. Raio de curvatura da Importina-α. Calculamos o raio de curvatura da Importina-α ao longo do movimento de dobramento. O gráfico demonstra que o raio de curvatura na Importina-α foi maior no modelo com deleção da ^{Npl}A159. O modelo com deleção da ^{Npl}K162 não apresentou alterações em relação ao raio de curvatura da estrutura canônica.

5. Discussão e Conclusão

Neste projeto, analisamos a influência de um *linker* de apenas nove resíduos para as interações e o comportamento conformacional da Importina- α . A literatura apresenta desde suas primeiras análises cristalográficas de Importinas- α ligadas a peptídeos NLS, que as Importinas são capazes de adotar diferentes conformações dependendo da proteína a qual se ligou principalmente por conta de seu formato solenoide/helicoidal (Pemberton e Paschal, 2005). Assim, analisamos as interações entre os resíduos da cNLS da Nucleoplasmina e da Importina- α segundo frequências de contatos, valores de flutuação, correlação positiva e flexibilidade da Importina- α para determinar o que ocorre entre resíduos da Importina- α e cNLS bipartida quando esta possui região *linker* menor que dez resíduos.

Semelhante a descrições encontradas na literatura (Fontes *et al.*, 2000; Fontes *et al.*, 2003; De Barros *et al.*, 2012; Marfori *et al.*, 2012), analisando as interações entre as estruturas, observamos a realização de ligações de hidrogênio entre os contatos ^{Npl}G164 e ^{Imp- α} R315, ^{Npl}Q165 com ^{Imp- α} W231 e ^{Npl}K162 com ^{Imp- α} R315 na região *linker* (Fontes *et al.*, 2000); ligações de hidrogênio e contatos hidrofóbicos entre resíduos do *linker* com ^{Imp- α} R315, ^{Imp- α} R238 e ^{Imp- α} Y277; ligações de hidrogênio entre resíduos do sítio secundário com os resíduos ^{Imp- α} W357, ^{Imp- α} S360 e ^{Imp- α} V321; e no sítio principal, interações entre ^{Npl}K167 (posição P2) com ^{Imp- α} T155 e ^{Imp- α} N192 (Fontes *et al.*, 2003). Ainda observamos a formação de diferentes contatos nos modelos de deleção, principalmente na deleção da ^{Npl}K162, como por exemplo a formação das ligações de hidrogênio ^{Npl}R156 com ^{Imp- α} E396 no sítio secundário, ^{Npl}K161 e ^{Imp- α} R315, ^{Npl}Q165 e ^{Imp- α} R238 e ^{Npl}A166 com ^{Imp- α} W231 na região *linker* e ^{Npl}K168 e ^{Imp- α} S149, ^{Npl}K169 com ^{Imp- α} E107 e ^{Imp- α} N146 no sítio principal. Entretanto, também observamos uma diminuição nas interações hidrofóbicas entre resíduos da Importina- α com da NLS do modelo com deleção da ^{Npl}A159, mas aumento na quantidade de pontes salinas. Estas diferenças nos contatos realizados pela Importina- α com os modelos canônicos e com deleções na região *linker* pode ser consequência da alteração na flexibilidade da Importina- α para se ajustar aos diferentes modelos cNLS.

Ademais, analisando os valores de flutuação e correlação entre resíduos da NLS da Nucleoplasmina nos modelos de deleção e da Importina- α , comparados à

estrutura canônica, observamos interação mais intensa nos sítios secundários e região *linker*, congruente com nossos resultados obtidos na análise de flexibilidade da Importina- α no movimento de dobramento lateral. Desta análise observamos a diminuição da flexibilidade entre resíduos da região *linker* de ambas as deleções nos ARMs 4 e 5 e ARMs 5 e 6 no modelo com deleção da ^{Npl}A159.

Os ângulos representados pelo movimento de torção, por sua vez, apresentam diminuição da flexibilidade entre os ângulos dos ARMs 2 e 3, no sítio principal, em ambos modelos com deleção e aumento da flexibilidade entre ARMs 5 e 6, e ARMs 6 e 7, na região *linker*. Congruente com resultados de flutuação e correlação, no qual observamos o aumento da flexibilidade entre os ARMs 6 e 7 no modelo com deleção da ^{Npl}K162; ARMs 8 e 9 na deleção da ^{Npl}A159. Geraldo *et al.* (2016) observou que o aumento da torção, principalmente entre os ARMs 6 e 7, estaria relacionado a uma maior flexibilidade da Importina- α , sendo este um possível indicador de instabilidade da proteína, já que este comportamento estrutural também foi encontrado na Importina em sua forma truncada não ligada à NLS (Falces *et al.*, 2010; Takeda *et al.*, 2013).

Em conclusão, a ausência de cNLSs com região *linker* menor que dez resíduos poderia ser explicada pela dificuldade de acomodação entre os resíduos e a maior flexibilidade da Importina- α , gerando uma instabilidade no complexo. Assim, as cNLS bipartidas provavelmente requerem um mínimo de dez resíduos no *linker* para a estabilidade na ligação Importina- α :NLS.

Análises adicionais ainda precisam ser realizadas para confirmar estes dados. Uma abordagem útil seria a realização de experimentos de Calorimetria de Titulação Isotérmica (ITC). Esta técnica avalia as interações bioquímicas entre moléculas em virtude do calor absorvido ou gerado por essa interação. Desta maneira, é obtido um perfil termodinâmico da interação molecular em um único experimento. Assim, poderemos complementar nossos dados de simulação computacional referente à ligação entre NLS e Importina- α .

6. Referências

ARJOMAND, A. et al. The alpha-importome of mammalian germ cell maturation provides novel insights for importin biology. **Faseb j**, v. 28, n. 8, p. 3480-93, Aug 2014. ISSN 0892-6638. Disponível em: < <http://www.fasebj.org/content/28/8/3480.full.pdf> >.

BERNARDES, N. E. et al. Structure of Importin-alpha from a Filamentous Fungus in Complex with a Classical Nuclear Localization Signal. **PLoS One**, v. 10, n. 6, p. e0128687, 2015. ISSN 1932-6203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4474859/pdf/pone.0128687.pdf> >.

BROOKS, B. R. et al. CHARMM: the biomolecular simulation program. **J Comput Chem**, v. 30, n. 10, p. 1545-614, Jul 30 2009. ISSN 0192-8651. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2810661/pdf/nihms101618.pdf> >.

CHANG, C. W. et al. Distinctive conformation of minor site-specific nuclear localization signals bound to importin-alpha. **Traffic**, v. 14, n. 11, p. 1144-54, Nov 2013. ISSN 1398-9219. Disponível em: < <http://onlinelibrary.wiley.com/store/10.1111/tra.12098/asset/tra12098.pdf?v=1&t=ilsc313t&s=0d8ddb8e83de136f9e2978f62f97fbb3dd9d6510> >.

CHEN, L.; MADURA, K. Yeast importin-alpha (Srp1) performs distinct roles in the import of nuclear proteins and in targeting proteasomes to the nucleus. **J Biol Chem**, v. 289, n. 46, p. 32339-52, Nov 14 2014. ISSN 0021-9258. Disponível em: < <http://www.jbc.org/content/289/46/32339.full.pdf> >.

CHOOK, Y. M.; BLOBEL, G. Karyopherins and nuclear import. **Curr Opin Struct Biol**, v. 11, n. 6, p. 703-15, Dec 2001. ISSN 0959-440X (Print) 0959-440x. Disponível em: < http://ac.els-cdn.com/S0959440X01002640/1-s2.0-S0959440X01002640-main.pdf?_tid=9ca3d344-ef58-11e5-8609-00000aab0f6c&acdnat=1458560136_a417411bbe3efb98d0810924b11e1e4f >.

CHRISTIE, M. et al. Structural Biology and Regulation of Protein Import into the Nucleus. **J Mol Biol**, Oct 30 2015. ISSN 0022-2836. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0022283615006166> >.

CONTI, E.; KURIYAN, J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. **Structure**, v. 8, n. 3, p. 329-38, Mar 15 2000. ISSN 0969-2126 (Print) 0969-2126. Disponível em: < http://ac.els-cdn.com/S0969212600001076/1-s2.0-S0969212600001076-main.pdf?_tid=ca3c14ee-f80c-11e5-9500-00000aab0f01&acdnat=1459517132_46b17606919be14a11d2d23ea8a005f2 >.

CONTI, E. et al. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. **Cell**, v. 94, n. 2, p. 193-204, Jul 24 1998. ISSN 0092-8674 (Print) 0092-8674. Disponível em: < [http://www.cell.com/cell/pdf/S0092-8674\(00\)81419-1.pdf](http://www.cell.com/cell/pdf/S0092-8674(00)81419-1.pdf) >.

CRONSHAW, J. M. et al. Proteomic analysis of the mammalian nuclear pore complex. **J Cell Biol**, v. 158, n. 5, p. 915-27, Sep 2 2002. ISSN 0021-9525 (Print) 0021-9525. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2173148/pdf/200206106.pdf> >.

DE BARROS, A. C. et al. Structural basis of nuclear import of flap endonuclease 1 (FEN1). **Acta Crystallogr D Biol Crystallogr**, v. 68, n. Pt 7, p. 743-50, Jul 2012. ISSN 0907-4449. Disponível em: < <http://scripts.iucr.org/cgi-bin/paper?S0907444912010281> >.

DINGWALL, C.; LASKEY, R. A. Protein import into the cell nucleus. **Annu Rev Cell Biol**, v. 2, p. 367-90, 1986. ISSN 0743-4634 (Print) 0743-4634. Disponível em: < <http://www.annualreviews.org/doi/pdf/10.1146/annurev.cb.02.110186.002055> >.

ESWAR, N. et al. Comparative protein structure modeling using Modeller. **Curr Protoc Bioinformatics**, v. Chapter 5, p. Unit 5.6, Oct 2006. ISSN 1934-3396. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4186674/pdf/nihms-628459.pdf> >.

FAHRENKROG, B.; AEBI, U. The nuclear pore complex: nucleocytoplasmic transport and beyond. **Nat Rev Mol Cell Biol**, v. 4, n. 10, p. 757-66, Oct 2003. ISSN 1471-0072 (Print) 1471-0072. Disponível em: < <http://www.nature.com/nrm/journal/v4/n10/pdf/nrm1230.pdf> >.

FALCES, J. et al. Recognition of nucleoplasmin by its nuclear transport receptor importin alpha/beta: insights into a complete import complex. **Biochemistry**, v. 49, n. 45, p. 9756-69, Nov 16 2010. ISSN 0006-2960. Disponível em: < <http://pubs.acs.org/doi/abs/10.1021/bi101179g> >.

FONTES, M. R. et al. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. **J Biol Chem**, v. 278, n. 30, p. 27981-7, Jul 25 2003. ISSN 0021-9258 (Print) 0021-9258. Disponível em: < <http://www.jbc.org/content/278/30/27981.full.pdf> >.

FONTES, M. R.; TEH, T.; KOBÉ, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. **J Mol Biol**, v. 297, n. 5, p. 1183-94, Apr 14 2000. ISSN 0022-2836 (Print) 0022-2836. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0022283600936428> >.

FORWOOD, J. K. et al. Kap95p binding induces the switch loops of RanGDP to adopt the GTP-bound conformation: implications for nuclear import complex assembly dynamics. **J Mol Biol**, v. 383, n. 4, p. 772-82, Nov 21 2008. ISSN 0022-2836. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0022283608009704> >.

GERALDO, M. T. et al. Bending-Twisting Motions and Main Interactions in Nucleoplasmin Nuclear Import. **PLoS One**, v. 11, n. 6, p. e0157162, 2016. ISSN

1932-6203. Disponível em: <
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4892583/pdf/pone.0157162.pdf> >.

HUANG, J.; MACKERELL, A. D., JR. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. **J Comput Chem**, v. 34, n. 25, p. 2135-45, Sep 30 2013. ISSN 0192-8651. Disponível em: <
<http://onlinelibrary.wiley.com/store/10.1002/jcc.23354/asset/jcc23354.pdf?v=1&t=imm2b0c6&s=f3c7a5667e804f11c92cefd53b863ae16d9d502e> >.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: visual molecular dynamics. **J Mol Graph**, v. 14, n. 1, p. 33-8, 27-8, Feb 1996. ISSN 0263-7855 (Print) 0263-7855.

JAKEL, S. et al. The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. **Embo j**, v. 18, n. 9, p. 2411-23, May 4 1999. ISSN 0261-4189 (Print) 0261-4189. Disponível em: <
<http://emboj.embopress.org/content/embojnl/18/9/2411.full.pdf> >.

KALDERON, D. et al. Sequence requirements for nuclear location of simian virus 40 large-T antigen. **Nature**, v. 311, n. 5981, p. 33-8, Sep 6-11 1984. ISSN 0028-0836 (Print) 0028-0836.

KOHLER, M. et al. Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. **Mol Cell Biol**, v. 19, n. 11, p. 7782-91, Nov 1999. ISSN 0270-7306 (Print) 0270-7306. Disponível em: <
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC84838/pdf/mb007782.pdf> >.

KONRAD, H. Normal Mode Theory and Harmonic Potential Approximations. In: (Ed.). **Normal Mode Analysis**: Chapman and Hall/CRC, 2005. p.1-16. (Chapman & Hall/CRC Mathematical & Computational Biology). ISBN 978-1-58488-472-9.

KURODA, K. et al. The role of X/Y linker region and N-terminal EF-hand domain in nuclear translocation and Ca²⁺ oscillation-inducing activities of phospholipase C ζ , a mammalian egg-activating factor. **J Biol Chem**, v. 281, n. 38, p. 27794-805, Sep 22 2006. ISSN 0021-9258 (Print) 0021-9258. Disponível em: < <http://www.jbc.org/content/281/38/27794.full.pdf> >.

LEUNG, S. W. et al. Dissection of the karyopherin alpha nuclear localization signal (NLS)-binding groove: functional requirements for NLS binding. **J Biol Chem**, v. 278, n. 43, p. 41947-53, Oct 24 2003. ISSN 0021-9258 (Print) 0021-9258. Disponível em: < <http://www.jbc.org/content/278/43/41947.full.pdf> >.

MARFORI, M. et al. Structural basis of high-affinity nuclear localization signal interactions with importin-alpha. **Traffic**, v. 13, n. 4, p. 532-48, Apr 2012. ISSN 1398-9219. Disponível em: < <http://onlinelibrary.wiley.com/store/10.1111/j.1600-0854.2012.01329.x/asset/tra1329.pdf?v=1&t=ilsa6tlh&s=c0e9fa8d05c21bba34754aa279936094ffe0589e> >.

MIYATAKE, H. et al. Crystal structure of human importin- α 1 (Rch1), revealing a potential autoinhibition mode involving homodimerization. **PLoS One**, v. 10, n. 2, p. e0115995, 2015. ISSN 1932-6203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4319847/pdf/pone.0115995.pdf> >.

PANG, X.; ZHOU, H. X. Design rules for selective binding of nuclear localization signals to minor site of importin α . **PLoS One**, v. 9, n. 3, p. e91025, 2014. ISSN 1932-6203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3946659/pdf/pone.0091025.pdf> >.

PEMBERTON, L. F.; PASCHAL, B. M. Mechanisms of receptor-mediated nuclear import and nuclear export. **Traffic**, v. 6, n. 3, p. 187-98, Mar 2005. ISSN 1398-9219 (Print) 1398-9219. Disponível em: < <http://onlinelibrary.wiley.com/store/10.1111/j.1600-0854.2005.00270.x/asset/j.1600-0854.2005.00270.x.pdf?v=1&t=im1wra0s&s=8015a7f638e0c08a7325a02bc6672501a468e74f> >.

PHILOT, E. A. et al. Binding sites and hydrophobic pockets in Human Thioredoxin 1 determined by normal mode analysis. **J Struct Biol**, v. 184, n. 2, p. 293-300, Nov 2013. ISSN 1047-8477. Disponível em: < http://ac.els-cdn.com/S1047847713002323/1-s2.0-S1047847713002323-main.pdf?_tid=815a8b30-fa6d-11e5-917d-00000aab0f02&acdnat=1459778573_90d9cd33aab9be8cfd10520411b8786d >.

PUMROY, R. A.; CINGOLANI, G. Diversification of importin- α isoforms in cellular trafficking and disease states. **Biochem J**, v. 466, n. 1, p. 13-28, Feb 15 2015. ISSN 0264-6021. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4405237/pdf/nihms679310.pdf> >.

ROBBINS, J. et al. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. **Cell**, v. 64, n. 3, p. 615-23, Feb 8 1991. ISSN 0092-8674 (Print) 0092-8674.

ROMAN, N. et al. Structural characterisation of the nuclear import receptor importin α in complex with the bipartite NLS of Prp20. **PLoS One**, v. 8, n. 12, p. e82038, 2013. ISSN 1932-6203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3858281/pdf/pone.0082038.pdf> >.

ROUT, M. P.; WENTE, S. R. Pores for thought: nuclear pore complex proteins. **Trends Cell Biol**, v. 4, n. 10, p. 357-65, Oct 1994. ISSN 0962-8924 (Print) 0962-8924.

SEEBER, M. et al. Wordom: a program for efficient analysis of molecular dynamics simulations. **Bioinformatics**, v. 23, n. 19, p. 2625-7, Oct 1 2007. ISSN 1367-4803. Disponível em: < <http://bioinformatics.oxfordjournals.org/content/23/19/2625.full.pdf> >.

SMITH, A. E. et al. The nuclear location signal. **Proc R Soc Lond B Biol Sci**, v. 226, n. 1242, p. 43-58, Oct 22 1985. ISSN 0950-1193 (Print)

0950-1193.

Disponível

em:

<

<http://rspb.royalsocietypublishing.org/content/226/1242/43.long> >.

STEWART, M. Molecular mechanism of the nuclear protein import cycle. **Nat Rev Mol Cell Biol**, v. 8, n. 3, p. 195-208, Mar 2007. ISSN 1471-0072 (Print) 1471-0072.

SUHRE, K.; SANEJOUAND, Y. H. ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. **Nucleic Acids Res**, v. 32, n. Web Server issue, p. W610-4, Jul 1 2004. ISSN 0305-1048.

TAKEDA, A. A. et al. Biophysical characterization of the recombinant importin- α from *Neurospora crassa*. **Protein Pept Lett**, v. 20, n. 1, p. 8-16, Jan 2013. ISSN 0929-8665.

7. Apêndice

A

#Deleção A₁₅₉

>P1;model

sequence: model: : A: : B :: : 0.00: 0.00

SAVKRPA-TKKAGQAKKKKLD

/

GTVNWSVEDIVKGINSNNLESQ LQATQAARKLLSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIAS
GTSEQTKAVVDGGAIPAFISLLASPHAHISEQAVWALGNIAGDGSFRDLVIKHGAIDPLLALLAVPDLSTLACGYLR
NLTWTLSNLCRNKNPAPPLDAVEQILPTLVRLHHNDPEVLADSCWASISYLTGPNRIEMVVKKGVVPQLVKLLG
ATELPIVTPALRAIGNIVTGTDEQTQKVIDAGALAVFPSLLTNPKTNIQKEATWTMSNITAGRQDQIQVVNHGLVPF
LVGVLSKADFQKQKEAAWAITNYTSGGTVEQIVYLVHCGIIEPLMNLLSAKDTKIIQVILDAISNIFQAAEKLGETEKL
IMIEECGGLDKIEALQRHENESVYKASLNLIKYS*

B

#Deleção K₁₆₂

>P1;model

sequence: model: : A: : B :: : 0.00: 0.00

SAVKRPAATK-AGQAKKKKLD

/

GTVNWSVEDIVKGINSNNLESQ LQATQAARKLLSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIAS
GTSEQTKAVVDGGAIPAFISLLASPHAHISEQAVWALGNIAGDGSFRDLVIKHGAIDPLLALLAVPDLSTLACGYLR
NLTWTLSNLCRNKNPAPPLDAVEQILPTLVRLHHNDPEVLADSCWASISYLTGPNRIEMVVKKGVVPQLVKLLG
ATELPIVTPALRAIGNIVTGTDEQTQKVIDAGALAVFPSLLTNPKTNIQKEATWTMSNITAGRQDQIQVVNHGLVPF
LVGVLSKADFQKQKEAAWAITNYTSGGTVEQIVYLVHCGIIEPLMNLLSAKDTKIIQVILDAISNIFQAAEKLGETEKL
IMIEECGGLDKIEALQRHENESVYKASLNLIKYS*

C

#modelo canônico

>P1;tudo

structureX: tudo: : A: : B :: : 0:0

SAVKRPAATKKAGQAKKKKLD

/

GTVNWSVEDIVKGINSNNLESQ LQATQAARKLLSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIAS
GTSEQTKAVVDGGAIPAFISLLASPHAHISEQAVWALGNIAGDGSFRDLVIKHGAIDPLLALLAVPDLSTLACGYLR
NLTWTLSNLCRNKNPAPPLDAVEQILPTLVRLHHNDPEVLADSCWASISYLTGPNRIEMVVKKGVVPQLVKLLG
ATELPIVTPALRAIGNIVTGTDEQTQKVIDAGALAVFPSLLTNPKTNIQKEATWTMSNITAGRQDQIQVVNHGLVPF
LVGVLSKADFQKQKEAAWAITNYTSGGTVEQIVYLVHCGIIEPLMNLLSAKDTKIIQVILDAISNIFQAAEKLGETEKL
IMIEECGGLDKIEALQRHENESVYKASLNLIKYS*

D

#Template

>P1;tudo

/

GTVNWSVEDIVKGINSNLLESQLQATQAARKLLSREKQPPIDNIIRAGLIPKFVSFLGKTDCSPIQFESAWALTNIAS
GTSEQTKAVVDGGAIPAFISLLASPHAHISEQAVWALGNIAGDGSAFRDLVIKHGAIDPLLALLAVPDLSTLACGYLR
NLTWTLNLCRNKNPAPPLDAVEQILPTLVRLHHNDPEVLADSCWAI SYLTDGPNERIEMVVKGVVPQLVKLLG
ATELPIVTPALRAIGNIVTGTDEQTQKVIDAGALAVFPSLLTNPKNTIQKEATWTMSNITAGRQDQIQQVNVNHLVPF
LVGVLSKADFKTQKEAAWAITNYTSGGTVEQIVYLVHCIIPLMNLLSAKDTKIIQVILDAISNIFQAAEKLGETEKL
IMIEECGGLDKIEALQRHENESVYKASLNLIK YFS

Figura A1. Alinhamento dos resíduos para modelagem. **A.** Alinhamento para modelagem da estrutura com deleção da alanina-159 (^{Npl}A159) na região *linker* da NLS bipartida da Nucleoplasmina. **B.** Alinhamento para modelagem da estrutura com deleção da lisina-162 (^{Npl}K162) na região *linker* da NLS bipartida da Nucleoplasmina. **C.** Alinhamento para modelagem do modelo canônico sem átomos de água. **D.** *Template* baseado na estrutura apresentada por Marfori *et al.* (2012).

INTERAÇÕES DA ESTRUTURA CANONICA (F>50%)				
Tipo de interação	Nucleoplasmina-NLS	Importina- α	Frequências > 50%	Região NLS
LIGAÇÕES DE HIDROGENIO	K155	V321	99.86%	S
	K155	G323	98.23%	S
	K155	N361	90.18%	S
	R156	W357	87.99%	S
	R156	S360	66.98%	S
	R156	N361	97.54%	S
	T160	R315	98.50%	L
	T160	K353	65.48%	L
	K162	N350	98.77%	L
	G164	R238	98.09%	L
	G164	Y277	97.00%	L
	Q165	W231	81.04%	L
	K167	A148	65.76%	P
	K167	G150	82.26%	P
	K167	T155	55.12%	P
	K167	N188	69.03%	P
	K167	D192	100.00%	P
	K168	W184	69.58%	P
	K168	N188	93.59%	P
	K169	L104	90.59%	P
	K169	R106	81.31%	P
	K170	W142	78.85%	P
CONTATOS HIDROFOBICOS	K155	A364	86.61%	S
	R156	W399	97.13%	S
	A158	R315	54.51%	L
	A158	W357	90.16%	L
	K162	W273	80.6%	L
	K162	T311	67.35%	L
	K162	R315	56.56%	L
	A163	Y277	86.48%	L
	A166	R238	89.89%	L
	K167	D192	78.83%	P
	K169	N146	69.67%	P
	K170	W142	97.54%	P
	K170	Q181	53.42%	P
	K170	W184	98.22%	P
PONTES SALINAS	R156	E396	100.00%	S
	K167	D192	100.00%	P
	K168	E266	52.93%	P

Figura A2. Interações entre NLS da Nucleoplasmina e Importina- α com frequência acima de 50% no modelo canônico. Tabela apresenta interações entre resíduos da NLS da Nucleoplasmina e Importina- α com frequência de ligações hidrofóbicas, ligações de hidrogênio e pontes salinas acima de 50%. Interações realizadas pelos resíduos da região *linker* estão destacados em azul. S = sítio secundário, P = sítio principal, L = região *linker*.

INTERAÇÕES DA ESTRUTURA CANONICA NA REGIAO <i>LINKER</i>			
Tipo de Interação	Nucleoplasmina-NLS	Importina- α	Frequências < 50%
LIGAÇÕES DE HIDROGENIO	P157	W357	20.87%
	P157	D280	0.14%
	P157	K395	0.14%
	A158	K240	9.69%
	A158	W357	13.64%
	A158	N319	8.87%
	A158	R315	4.37%
	A158	D280	1.09%
	A158	E396	0.27%
	A159	E396	4.50%
	A159	K240	0.95%
	A159	K392	0.14%
	A159	W357	0.14%
	T160	K392	0.68%
	T160	W357	0.69%
	T160	E396	0.68%
	T160	E354	0.14%
	K161	N239	3.96%
	K161	R315	0.27%
	K161	D280	0.27%
	K161	N241	0.14%
	K162	R315	7.63%
	K162	E354	0.41%
	K162	W273	0.41%
	K162	T311	0.27%
	K162	Y277	0.27%
	A163	Y277	11.87%
	A163	R238	13.23%
	A163	W273	0.41%
	G164	W273	0.82%
	Q165	R238	2.86%
	Q165	E107	0.27%
	A166	W231	48.43%
	A166	G191	12.55%
	A166	N235	9.96%
	A166	R238	9.27%
	A166	N188	1.23%
CONTATOS HIDROFOBICOS	K162	E354	33.33%
	A166	N235	3.01%
PONTES SALINAS	K161	D280	0.55%
	K162	E354	0.14%

Figura A3. Interações entre a região *linker* da NLS da Nucleoplasmina e Importina- α com frequência abaixo de 50% no modelo canônico. Tabela apresenta interações entre

resíduos da região *linker* da NLS da Nucleoplasmina e Importina- α com frequência de ligações hidrofóbicas, ligações de hidrogênio e pontes salinas menores que 50%.

INTERACOES ENTRE IMPORTINA- α E NLS COM DELEÇÃO NA ^{Npl} A159				
Tipo de Interação	Nucleoplasmina-NLS	Importina- α	Frequência > 50%	Região NLS
LIGAÇÃO DE HIDROGENIO	K155	N361	92.63%	S
	K155	G323	89.63%	S
	K155	V321	99.86%	S
	R156	W357	99.73%	S
	R156	S360	99.59%	S
	R156	N361	99.59%	S
	R156	E396	98.36%	S
	P157	W357	52.52%	L
	T160	R315	100.00%	L
	K162	E354	95.91%	L
	G164	R238	99.45%	L
	G164	Y277	97.54%	L
	Q165	W231	80.49%	L
	A166	W231	75.44%	L
	K167	D192	99.59%	P
	K167	N188	90.04%	P
	K167	G150	83.08%	P
	K167	A148	81.99%	P
	K167	T155	80.35%	P
	K168	N188	94.27%	P
	K168	S149	94.13%	P
	K168	W184	51.02%	P
	K168	E266	50.61%	P
	K169	L104	87.18%	P
	K169	R106	85.40%	P
	K169	W184	69.17%	P
	K170	Q181	70.53%	P
	K170	W142	95.63%	P
CONTATOS HIDROFOBICOS	R156	W399	97.54%	S
	A158	W357	99.73%	L
	K162	R315	62.57%	L
	A163	Y277	96.17%	L
	A166	R238	96.17%	L
	K169	N146	50.96%	P
	K170	W142	98.09%	P
	K170	W184	97.81%	P
PONTES SALINAS	K155	D160	100.00%	S
	R156	E396	98.36%	S
	K162	E354	99.18%	L
	K167	D192	99.59%	P
	K168	E266	59.62%	P
	K168	D270	50.48%	P

Figura A4. Interações entre NLS da Nucleoplasmina e Importina- α com frequência acima de 50% no modelo de deleção no cNLS do resíduo ^{Npl}A159. Tabela apresenta interações entre resíduos da NLS da Nucleoplasmina com deleção no resíduo ^{Npl}A159 e Importina- α com frequência de ligações hidrofóbicas, ligações de hidrogênio e pontes salinas acima de 50% nos sítios principal, secundário e região *linker*. S = sítio secundário, P = sítio principal, L = região *linker*.

INTERACOES ENTRE IMPORTINA- α E NLS COM DELEÇÃO NA ^{Npl} K162				
Tipo de interação	Nucleoplasmina-NLS	Importina- α	Frequência >50%	Região NLS
LIGACOES DE HIDROGENIO	K155	V321	99.70%	S
	K155	N361	95.54%	S
	R156	E396	100.00%	S
	R156	N361	97.62%	S
	R156	W357	90.03%	S
	R156	S360	73.36%	S
	T160	K353	54.02%	L
	K161	R315	58.63%	L
	G164	R238	91.07%	L
	Q165	R238	83.78%	L
	A166	W231	55.36%	L
	K167	D192	99.55%	P
	K167	A148	98.51%	P
	K167	G150	95.39%	P
	K167	N188	94.35%	P
	K168	N188	97.47%	P
	K168	W184	81.99%	P
	K168	S149	60.71%	P
	K169	L104	94.20%	P
	K169	E107	64.88%	P
	K169	N146	62.65%	P
	K170	W142	90.77%	P
CONTATOS HIDROFOBICOS	K155	A364	85.69%	S
	R156	W399	96.57%	S
	A158	W357	95.53%	L
	A163	Y277	79.58%	L
	A163	W273	68.70%	L
	A166	R238	82.71%	L
	K167	D192	88.52%	P
	K169	N146	88.67%	P
	K170	W142	99.85%	P
	K170	W184	98.96%	P
PONTES SALINAS	R156	E396	100.00%	S
	K167	D192	99.79%	P
	K169	E107	61.78%	P

Figura A5. Interações entre NLS da Nucleoplasmina e Importina- α com frequência acima de 50% no modelo de deleção no cNLS do resíduo ^{Npl}K162. Tabela apresenta interações entre resíduos da NLS da Nucleoplasmina com deleção no resíduo ^{Npl}K162 e

Importina- α com frequência de ligações hidrofóbicas, ligações de hidrogênio e pontes salinas acima de 50% nos sítios principal, secundário e região *linker*. S = sítio secundário, P = sítio principal, L = região *linker*.

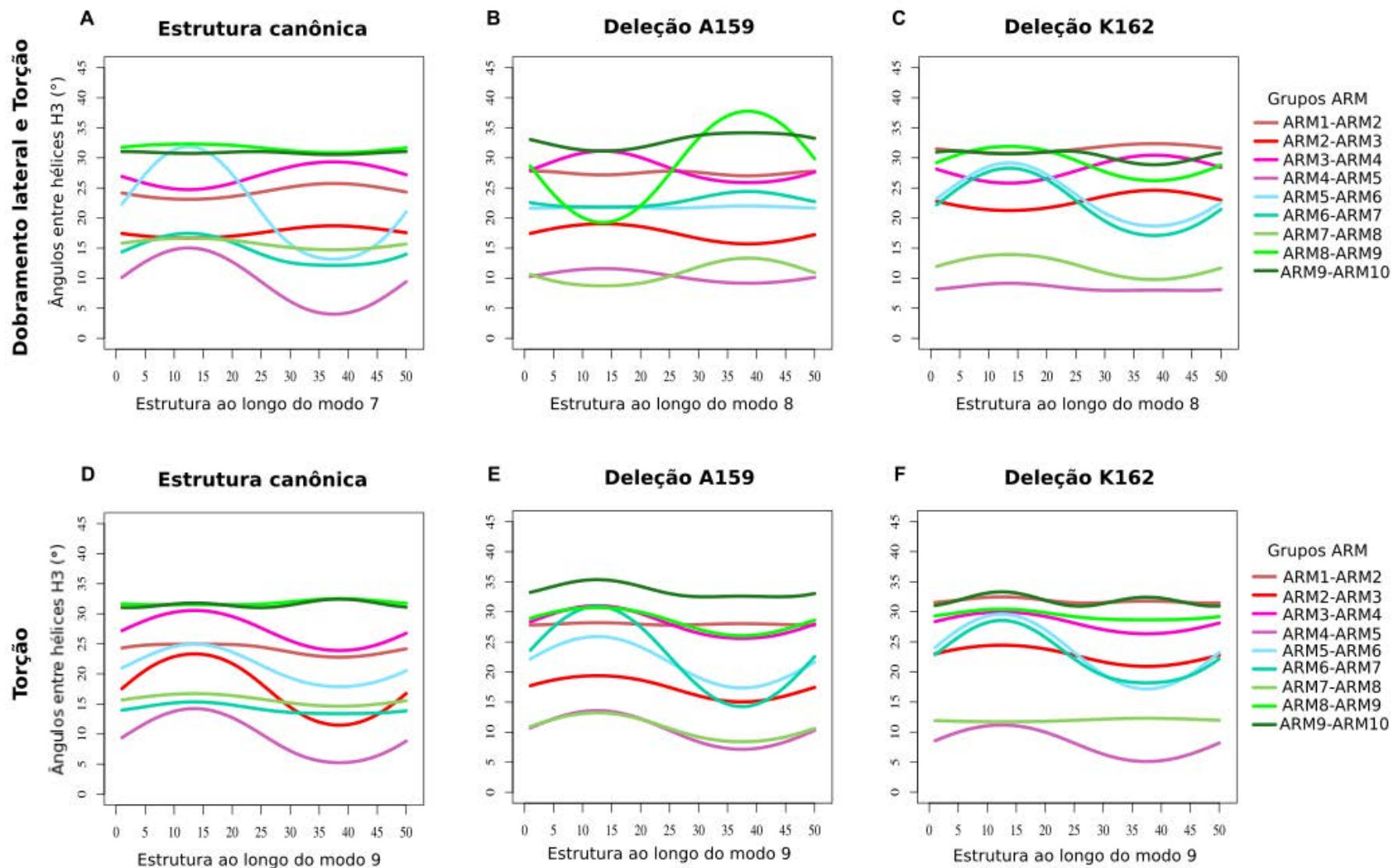


Figura A6. Ângulo entre α -hélices H3 de repetições ARM vizinhas da Importina- α ao longo do modo. Os gráficos apresentam como a

flexibilidade da Importina- α através do cálculo dos ângulos formados pelas α -hélices H3 de repetições ARM vizinhas. Os gráficos 5A, B e C apresentam a variação angular durante o movimento de dobramento lateral representado pelo modo 7 na estrutura canônica e modos 8 nos modelos de deleção. Gráficos 5D, E e F apresentam a variação angular durante o movimento de torção representados pelos modos 9 de cada estrutura