

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
ÁREA DE CONCENTRAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

JOSÉ RENATO PATTARO JÚNIOR

**EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOFÍSICA DA PROTEÍNA
KIN DE *HOMO SAPIENS***

Maringá
2017

JOSÉ RENATO PATTARO JÚNIOR

**EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOFÍSICA DA PROTEÍNA
KIN DE *HOMO SAPIENS***

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas (área de concentração - Biologia Celular e Molecular), da Universidade Estadual de Maringá para a obtenção do grau de Mestre em Ciências Biológicas

Orientador: Prof. Dr. Flávio A. V. Seixas
Coorientador: Prof^a. Dr^a. Maria A. Fernandez
Coorientador: Dr. Quirino A. de Lima Neto

Maringá
2017

Dados Internacionais de Catalogação na Publicação (CIP)
(Biblioteca Central - UEM, Maringá, PR, Brasil)

P315e Pattaro Júnior, José Renato
Expressão heteróloga e caracterização biofísica da Proteína KIN de *Homo sapiens* / José Renato Pattaro Júnior. -- Maringá, 2017.
50 f. : il., color., figs., tabs.

Orientador(a): Prof. Dr. Flávio Augusto Vicente Seixas.
Orientador(a): Prof^a. Dr^a. Maria Aparecida Fernandez.
Orientador(a): Dr. Quirino Alves de Lima Neto.
Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Biológicas, Programa de Pós-Graduação Ciências Biológicas, 2017.

1. Proteína KIN. 2. Dicroísmo circular. 3. Expressão heteróloga. 4. Biomarcador tumoral. I. Seixas, Flávio Augusto Vicente, orient. II. Fernandez, Maria Aparecida, coorient. III. Lima Neto, Quirino Alves de, coorient. IV. Universidade Estadual de Maringá. Centro de Ciências Biológicas. Programa de Pós-Graduação Em Ciências Biológicas. V. Título.

CDD 21.ed. 572.84

AHS-CRB-9/1065

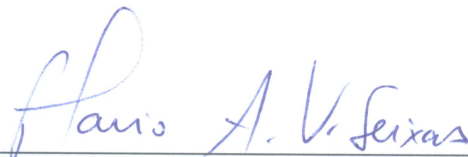
JOSÉ RENATO PATTARO JÚNIOR

**EXPRESSÃO HETERÓLOGA E CARACTERIZAÇÃO BIOFÍSICA DA
PROTEÍNA KIN DE *HOMO SAPIENS***

Dissertação apresentada à Universidade
Estadual de Maringá, como requisito
parcial para a obtenção do título de mestre.

Aprovado em: 24/02/2017

BANCA EXAMINADORA



Prof. Dr. Flavio Augusto Vicente Seixas - Orientador
Universidade Estadual de Maringá - UEM



Prof.^a Dr.^a Fátima Pereira de Souza
Universidade Estadual Paulista - UNESP



Prof.^a Dr.^a Érika Seki Kioshima Cotica
Universidade Estadual de Maringá - UEM

Biografia

José Renato Pattaro Júnior, filho de Leonice Aparecida Carmona Pattaro e José Renato Pattaro, nasceu no dia 11 de Março de 1989 na cidade de Maringá, Paraná. Graduado em Bioquímica (Bacharelado) em 2014 pela Universidade Estadual de Maringá. Tem experiência na área de Biologia Celular e Bioquímica Molecular, atuando principalmente nos seguintes temas: Tecnologias do DNA Recombinante e Expressão Heteróloga de Proteínas.

*Dedico este trabalho aos meus pais,
meus guerreiros e maiores exemplos.*

AGRADECIMENTO

Gostaria muito de agradecer a minha família e minha namorada por todo o apoio e compreensão. Ao meu orientador Prof. Dr. Flávio Seixas por me mostrar a direção certa nos tortuosos caminhos da ciência. A professora Maria Aparecida Fernandez por ter me aceito em seu laboratório, por todos esses anos de parceria e aprendizado, que foram essenciais para a minha formação. Ao Quirino A. de Lima Neto por todo o companheirismo e os puxões de orelha na bancada. A todos os colegas do LORF, especialmente ao Francisco Duarte Junior, a Fabiana dos Santos Rando e a Anelise Ramos pela amizade. Paulo Sérgio Alves Bueno por todos esses anos de amizade e as quebradas de galho. Professor Dr. Marcelo Fossey, Professora Dr.^a Fátima Pereira de Souza e ao Dr. Ícaro Caruso do CMIB do Ibilce-Unesp de São José do Rio Preto por todo o apoio nas análises de Dicroísmo Circular. Agradeço ao CNPq, CAPES e Fundação Araucária pelo apoio financeiro.

Apresentação

Esta dissertação de mestrado é composta por um artigo completo de pesquisa original, que será submetido a revista *Protein Science*, **ISSN**: 1469-896X e Fator de Impacto de 3.039.

Expressão heteróloga e caracterização biofísica da proteína KIN de *Homo sapiens*

José R. Pattaro Júnior¹, Quirino A. Lima Neto², Ícaro P. Caruso³, Marcelo A. Fossey³, Maria A. Fernandez², Flávio A. V. Seixas¹

¹Departamento de Bioquímica. Universidade Estadual de Maringá, Umuarama, PR, Brasil

²Departamento de Biotecnologia, Genética e Biologia Celular, Universidade Estadual de Maringá, Maringá, PR, Brasil

³Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas – Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, SP, Brasil.

Resumo Geral

A proteína KIN foi identificada em 1989 em células de rato a partir da reação cruzada de anticorpos policlonais monoespecíficos, dirigidos contra a proteína RecA de *Escherichia coli*. O gene *KIN* codifica uma proteína nuclear de 45 kDa que possui a capacidade de se ligar a ácidos nucleicos. Entre os eucariotos, o gene *KIN* é conservado filogeneticamente, o que indica atividade funcional em processos biológicos indispensáveis. Em murinos, a proteína KIN está associada à cromatina formando focos de replicação por todo o nucleoplasma, similar àqueles observados com outras proteínas envolvidas no reparo do DNA, na replicação e no *splicing* de RNA.

KIN é ubiquamente expressa em mamíferos e mostra alta expressão em células humanas de músculo esquelético, coração, testículo e células com altas taxas de proliferação, como tumores. Estes dados apontando que KIN tenha um papel na replicação. Adicionalmente, vem sendo associada as funções de fator de transcrição em bactérias e leveduras; reprogramação da expressão gênica global na adaptação de plantas à disponibilidade limitada de cobre, pela interação com SPL7; e em murinos mostrou-se que KIN é importante para um eficiente reparo e resolução das quebras de fita dupla, ocasionadas por radiação ionizante e pela troca de isotipo de imunoglobulinas pelos linfócitos B. Em neurônios existe a evidência de um papel alternativo para a KIN ao invés da replicação, já que esses não se dividem mais depois de maduros.

A proteína KIN é composta estruturalmente por motivos particulares: um dedo de zinco, um domínio de hélice alada (*Winged Helix Domain*), um domínio homólogo a proteína recA de *E. coli* e um motivo KOW. No entanto, a completa estrutura tridimensional da proteína KIN humana ainda não foi resolvida por métodos experimentais e nem por modelagem *in silico*, devido à falta de moldes estruturais para a maioria da sequência, indicando uma estrutura ímpar para essa proteína. Diante destas características, nosso grupo se propôs a realizar a análise filogenética e a caracterização biofísica da proteína KIN, utilizando técnicas de bioinformática e espectroscopia de Dicroísmo Circular (CD). A determinação dos parâmetros biofísicos da ${}_{\text{HSA}}\text{KIN}$ como Entalpia de desnaturação (ΔH_m) e Temperatura de *melting* (T_m) ainda não foi efetuada, assim, o objetivo deste trabalho foi o de realizar

ensaios de expressão, análise filogenética e análise biofísica desta proteína de modo a determinar parâmetros termodinâmicos que possam contribuir para a elucidação de suas funções celulares.

A análise computacional de parâmetros biofísicos e estruturais foi realizada a partir da sequência traduzida da proteína HSAKIN(His)_{6x} , utilizando programas e servidores on-line, os quais, também foram utilizados para os alinhamentos de múltiplas sequências. Foi prevista uma massa de 48.032,65 Da, ponto isoelétrico teórico de 9,07 e um NLS Bipartido (NLS+) para a proteína HSAKIN(His)_{6x} . Com base nos alinhamentos pôde-se verificar que HSAKIN é bastante conservada entre os animais utilizados, passando de 90% de identidade. Quando se analisou a identidade entre os domínios das proteínas homólogas à HSAKIN separadamente, notou-se um maior grau de identidade para os domínios dedo de zinco e hélice alada, sugerindo que podem ser responsáveis pelo processo mais fundamental à função desta proteína, a ligação à cromatina, ao qual já foi visto que esses dois domínios podem se associar, bem como a interação proteína-proteína. O motivo NLS+ encontrado apresentou maior identidade para a maioria dos eucariotos superiores em relação ao anteriormente descrito na literatura.

Bactérias *Escherichia coli* BL21 (DE3) transformadas com o plasmídeo pET23a-d(+) HSAKIN(His)_{6x} , contendo a sequência codificante da proteína HSAKIN(His)_{6x} , foram induzidas com IPTG para promover a superprodução da proteína de interesse. Posteriormente, as proteínas foram purificadas com cromatografia de afinidade e tiveram seus tampões trocados em coluna de exclusão molecular. Depois de purificadas e concentradas, as amostras foram submetidas a análise por dicroísmo circular.

A determinação da porcentagem dos elementos de estrutura secundária da proteína purificada foi realizada para três pHs (6,4; 7,4 e 8,4), por meio da deconvolução do espectro de CD a 20 °C. Em pH fisiológico a proteína HSAKIN(His)_{6x} possui mais da metade de sua composição de elementos não estruturados, voltas- β e *random coils*, seguido de folhas- β (30%) e em menor porcentagem de α -hélices (17%). Esses resultados sugerem, com base que regiões não enoveladas podem se inter-relacionar com uma coleção de parceiros na organização de complexas redes de interação proteína-proteína, assim, essas regiões não estruturadas de KIN

poderiam auxiliar na formação de complexas redes de interação proteína-proteína, promovendo a formação de complexos proteicos de alto peso molecular que agem nos processos aos quais HSAKIN foi relacionada. Esses dados podem ser extrapolados para as homólogas analisadas que apresentam altos valores de identidade de sequência.

Os resultados de desnaturação térmica para os pHs 6,4 e 7,4 no comprimento de onda 222 nm, mostraram os parâmetros biofísicos ΔH_m e T_m relativamente baixos em comparação aos encontrados na literatura. Esses valores substancialmente baixos, provavelmente se devem a esta proteína não possuir nenhuma ponte dissulfeto intra-cadeia, mesmo contando com 7 cisteínas em sua sequência. Possivelmente, a não ocorrência de ligações pode ser devido a diversos impedimentos estéricos impostos pela estrutura enovelada da proteína.

General Abstract

KIN protein was identified in 1989 in mouse cells by means cross-reaction of monospecific polyclonal antibodies directed against the RecA protein from *Escherichia coli*. The *KIN* gene encodes a 45 kDa nuclear protein which has nucleic acid binding capacity. Among eukaryotes, *KIN* gene is conserved phylogenetically, which indicates functional activity in indispensable biological processes. In murine, KIN protein is associated with chromatin forming replication foci throughout the nucleoplasm, similar to those observed with other proteins involved in replication, in DNA repair and RNA splicing.

KIN is ubiquitously expressed in mammals and shows high expression in human cells of skeletal muscle, heart, testis and cells with high proliferation rates, such as tumors. These data point out that KIN has a role in replication. In addition, it has been associated to other functions like transcription factor in bacteria and yeasts; reprogramming of global gene expression in plant adaptation to limited copper availability by interaction with SPL7; and in murines it has been shown that KIN is important for efficient repair and resolution of double-stranded breaks caused by ionizing radiation and for the immunoglobulins isotype exchange by B-lymphocytes. In neurons, there is evidence of an alternative role for KIN rather than replication, since they do not divide more after mature.

KIN protein is structurally composed for particular motifs: a zinc finger, a winged helix domain, a domain homologous to the recA protein from *E. coli* and a KOW motif. However, the complete three-dimensional structure of human KIN protein has not yet been resolved by experimental methods or by *in silico* modeling, due to the lack of structural templates for most of the sequence, indicating a unique structure for this protein. In view of these characteristics, our group proposed to carry out the phylogenetic analysis and biophysical characterization of human KIN protein, using bioinformatics techniques and Circular Dichroism spectroscopy (CD). The determination of the biophysical parameters of hSAKIN as denaturation enthalpy (ΔH_m) and melting temperature (T_m) has not been performed, so the aim of this work was to perform expression, phylogenetic and biophysical analysis of this protein in order to determine thermodynamic parameters that may contribute to the elucidation of their cellular functions.

The computational analysis of biophysical and structural parameters was performed from the HSAKIN(His)_{6x} protein sequence, using on-line programs and servers, which were also used for multiple sequence alignments. It was predicted a molecular weight of 48,032.65 Da, theoretical isoelectric point of 9.07 and a bipartite NLS (NLS+) for the HSAKIN(His)_{6x} protein. Based on the alignments it could be verified that HSAKIN is highly conserved among the tested animals, passing from 90% of identity. When the identity among HSAKIN homologous protein domains was analyzed separately, a higher degree of identity was noted for the zinc finger and winged helix domain, suggesting that they may be responsible for the most fundamental process to the function of this protein, the binding to chromatin, to which it has already been seen that these two domains can be associated, as well as the interaction protein-protein. The NLS+ motif found showed higher identity for most eukaryotes compared to that previously described in the literature.

Bacteria *Escherichia coli* BL21 (DE3) transformed with the plasmid pET23a-d (+) HSAKIN(His)_{6x} , containing the HSAKIN(His)_{6x} protein coding sequence, were induced with IPTG to promote overproduction of the protein of interest. Subsequently, the proteins were purified with affinity chromatography and had their buffers exchanged into a molecular exclusion column. After being purified and concentrated, the samples were analyzed by circular dichroism.

The percentage determination of the secondary structure elements of the purified protein was performed for three pHs (6,4; 7,4 e 8,4), by deconvolution of the CD spectrum at 20 °C. In physiologic pH HSAKIN(His)_{6x} has more than half of its composition of non-structured elements, β -turns and random coils, followed by β -sheets (30%) and a lower percentage of α -helices (17%). These results suggest, based on that non-envelope regions can interrelate with a collection of partners in the organization of complex protein-protein interaction networks, thus, these non-structured KIN regions could aid in the formation of complex protein-protein, promoting the formation of high molecular weight protein complexes that act on the processes to which HSAKIN has been related. These data can be extrapolated to the analyzed counterparts having high sequence identity values.

The thermal denaturation results for pHs 6.4 and 7.4 at the wavelength 222 nm showed relatively low biophysical parameters ΔHm and Tm compared to the literature. These substantially low values are possibly due to this protein having no

intra-chain disulfide bridge, even with 7 cysteines in its sequence. Possibly, the non-occurrence of bonds may be due to several steric hindrances imposed by the fold structure.

Expressão heteróloga e caracterização biofísica da proteína KIN de *Homo sapiens*

José R. Pattaro Júnior¹, Quirino A. Lima Neto², Ícaro P. Caruso³, Marcelo A. Fossey³, Maria A. Fernandez², Flávio A. V. Seixas¹

¹Departamento de Bioquímica. Universidade Estadual de Maringá, Umuarama, PR, Brasil

²Departamento de Biotecnologia, Genética e Biologia Celular, Universidade Estadual de Maringá, Maringá, PR, Brasil

³Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas – Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, SP, Brasil.

RESUMO

A proteína KIN foi identificada em 1989 em células de rato a partir da reação cruzada de anticorpos policlonais mono-específicos, dirigidos contra a proteína RecA de *Escherichia coli*. É ubiquamente expressa em mamíferos e se mostra superexpressa em células com altas taxas proliferativas. Diversos trabalhos na literatura apontam que KIN tenha um papel na replicação, reparo e no *splicing* do RNA. É estruturalmente composta por 5 motivos particulares: um dedo de zinco, um domínio de hélice alada (*Winged Helix Domain*), um domínio homólogo a proteína recA de *E. coli* e um motivo KOW. No entanto, a função e a completa estrutura tridimensional da proteína KIN humana ainda não foi resolvida por métodos experimentais e nem por modelagem *in silico*. Diante disso, o objetivo deste trabalho foi o de realizar ensaios de expressão, análise filogenética e análise biofísica (ΔH_m e T_m) desta proteína de modo a determinar parâmetros termodinâmicos que possam contribuir para a elucidação de suas funções celulares. A partir de análises computacionais da sequência da proteína HSAKIN(His)_{6x} obtivemos a massa de 48.032,65 Da e um PI de 9,07. Adicionalmente, com as análises de identidade por alinhamentos múltiplos, verificou-se que KIN é bastante conservada nos animais analisados, principalmente, seus domínios dedo de zinco e hélice alada. Identificou-se um NLS com maior identidade para os eucariotos superiores do que o proposto anteriormente na literatura. A proteína recombinante HSAKIN(His)_{6x} foi expressa heterologicamente em *E. coli* BL21 e purificada por cromatografia de afinidade. Com a deconvolução do espectro de CD pôde-se verificar que HSAKIN(His)_{6x} possui maior porcentagem em composição de elementos não estruturados, que podem auxiliar a promover a formação de complexos proteicos de alto peso molecular que agem nos processos aos quais HSAKIN foi relacionada. Os resultados de desnaturação térmica para os pHs 6,4 e 7,4 no comprimento de onda 222 nm, mostraram os parâmetros biofísicos ΔH_m e T_m relativamente baixos em comparação aos comparados com a literatura, possivelmente, ocasionado por a proteína não possuir nenhuma ponte dissulfeto intra-cadeia para a estabilização da estrutura.

Palavras-chave: Proteína KIN. Dicroísmo Circular. Expressão Heteróloga. Biomarcador Tumoral.

LISTA DE ILUSTRAÇÕES

Figura 1 -	Domínios funcionais da proteína KIN	39
Figura 2 -	Análise da purificação da proteína HSAKIN(His)_{6x} por meio <i>SDS-Page</i> e <i>Western Blotting</i>	40
Figura 3 -	Espectro de dicroísmo circular para HSAKIN(His)_{6x} em diferentes pHs	41
Figura 4 -	Curva de desnaturação térmica a 222 nm para a proteína recombinante HSAKIN(His)_{6x} em três diferentes pHs	42
Figura 5 -	Determinação dos parâmetros físico-químicos ΔHm e Tm para a proteína quimérica HSAKIN(His)_{6x}	43
Figura 6 -	Verificação por PCR da transformação de <i>E. coli</i> BL21 (DE3) utilizando o plasmídeo pET23a-d(+) HSAKIN(His)_{6x}	44
Figura S1 -	Sequência da proteína quimérica HSAKIN(His)_{6x}	45
Figura S2 -	Alinhamento das sequências homólogas à proteína KIN humana	46

LISTA DE TABELAS

Tabela I -	Análise da identidade das proteínas homólogas a HSAKIN	34
Tabela II -	Análise da conservação dos domínios das proteínas homólogas a KIN humana	35
Tabela III -	Porcentagens de estrutura secundária de HSAKIN(His) _{6x} em diferentes condições de pH	36
Tabela IV -	Determinação dos parâmetros ΔH_m e T_m para dois distintos pHs	37

LISTA DE ABREVIATURAS E SIGLAS

BTCD	<i>Binding To Curved DNA</i>
CD	<i>Circular Dichroism</i>
CTAB	<i>Cetyl trimethylammonium bromide</i>
EDBCP	<i>Ensemble-based Disulfide Bonding Connectivity Pattern Prediction</i>
FPLC	<i>Fast Protein Liquid Chromatography</i>
IPTG	Isopropil- β -D-1-tiogalactopiranosídeo
KOW	Kyrpides Ouzouns Woese
LB	Luria Bertani
NLS	<i>Nuclear Localization Signal</i>
ORF	<i>Open Reading Frame</i>
PARP	<i>Poly ADP-ribose Polymerase</i>
PCR	<i>Polymerase Chain Reaction</i>
PI	Ponto Isoelétrico
PMSF	<i>Phenylmethylsulfonyl Fluoride</i>
RMN	Ressonância Magnética Nuclear
SDS	<i>Sodium Dodecyl Sulfate</i>
SV40	<i>Simian Virus 40</i>
T_m	Temperatura de <i>Melting</i>
ZF	<i>Zinc Finger</i>

SUMÁRIO

Introdução	13
Resultados	16
<i>Análise computacional da sequência de $HSAKIN.(His)_6$</i>	16
<i>Superexpressão e purificação de $HSAKIN(His)_{6x}$</i>	18
<i>Predição dos componentes de estrutura secundária em diferentes condições de pH</i>	18
<i>Determinação dos parâmetros ΔH_m e T_m</i>	19
Discussão	20
Material e Métodos	24
<i>Material biológico</i>	24
<i>Análise computacional</i>	24
<i>Transformação das bactérias</i>	25
<i>Expressão e purificação da proteína $HSAKIN(His)_{6x}$</i>	26
<i>Avaliação dos parâmetros biofísicos por dicroísmo circular</i>	27
Referências Bibliográficas	28
Tabelas	34
Legendas das Figuras	38
Figuras (arte)	39
Material Suplementar	45
Trabalhos Publicados no Período	50

Introdução

Células são as unidades básicas, funcionais e biológicas da vida, e são regidas por uma intrincada rede de processos físico-bioquímicos. Numerosos processos celulares dependem da interação entre proteínas e ácidos nucleicos. Tais proteínas têm papéis cruciais na replicação, transcrição, *splicing* e reparo do DNA. Na proliferação celular estes processos devem ser estritamente coordenados, para assegurar a propagação do genoma íntegro durante as divisões celulares e, desta forma, prevenir mutações e rearranjos no DNA.^{1,2}

A proteína KIN é uma das proteínas que possuem interação com ácidos nucleicos, mas que ainda não tem função elucidada. Foi inicialmente identificada em 1989 em células de rato (FR 3T3) a partir da reação cruzada de anticorpos policlonais monoespecíficos, dirigidos contra a proteína RecA de *Escherichia coli*.³ Foi denominada inicialmente como KIN17, mas, posteriormente, recebeu o nome de BTCD (*Binding To Curved DNA*), ou seja, proteína que se liga à DNAs curvos (*bent*), encontrados em *hotspots* de recombinação ilegítima em cromossomos eucarióticos.^{4,5} Mas em publicações recentes tem sido referida como KIN.

O gene *KIN* Humano codifica uma proteína nuclear de 393 resíduos de aminoácidos (*HSAKIN*), com massa molecular de 45 kDa, e está localizado no braço curto do cromossomo 10.⁶ Entre os eucariotos, o gene *KIN* é conservado filogeneticamente, de leveduras a humanos, o que indica atividade funcional em processos biológicos indispensáveis (<https://tel.archives-ouvertes.fr/tel-00432998>). Em murinos, a proteína KIN está associada à cromatina formando focos de replicação por todo o nucleoplasma, similar àqueles observados com outras proteínas envolvidas no reparo do DNA,⁷⁻⁹ na replicação¹⁰ e no *splicing* de RNA.^{11,12} No entanto, foi verificado em células HEK 293 quando KIN é metilada em sua lisina

135 pela metiltransferase METTL22, desloca-se para o citoplasma, assim, perdendo sua capacidade de se associar à cromatina.^{13,14}

KIN é ubiquamente expressa em mamíferos e mostra alta expressão em células humanas de músculo esquelético, coração, testículo e células com altas taxas de proliferação, incluindo diversas células tumorais. Estes dados apontam que KIN tenha um possível papel na replicação.^{12,15-17} Adicionalmente, outras evidências mostram que KIN está envolvida em outros processos importantes. Em bactérias, ela atua como um substituto para a proteína H-NS, que age como fator de transcrição.⁵ Em leveduras, onde é denominada Rts2, regula a transcrição direta ou indiretamente, pela interação com fatores de transcrição em cabeça de garfo.¹⁸ Em plantas, ela interage com SPL7, uma proteína que é parte de um complexo que reprograma a expressão gênica global, que atua na adaptação da planta à disponibilidade limitada de cobre.^{19,20} Em neurônios, existe a evidência de um papel alternativo para a KIN ao invés da replicação. Estas células possuem altos níveis de expressão de KIN comparadas as outras células do sistema nervoso central, mesmo estas não mais se dividindo depois de diferenciadas.²¹ Em murinos, mostrou-se que KIN é importante para um eficiente reparo e resolução das quebras de fita dupla, ocasionadas por radiação ionizante e pelo processo de troca de isotipo de imunoglobulinas pelos linfócitos B.²²

A proteína KIN é composta estruturalmente por motivos particulares [Fig. 1]: um dedo de zinco (resíduos de aminoácidos 28-50), permitindo a interação da proteína KIN com DNA e RNA;^{4,23} um domínio de hélice alada (*Winged Helix Domain*) (resíduos 51-160), geralmente suportam interações com o DNA, mas, neste caso, devido ao posicionamento e a carga superficial, a sugestão é de que este domínio poderia mediar interações proteína-proteína;^{24,25} um domínio homólogo a

proteína recA de *E. coli* (resíduos 162-201), apresenta uma homologia de 49% com a extremidade C-terminal da proteína RecA bacteriana;^{26,27} um sinal de localização nuclear, NLS (*Nuclear Localization Signal*) (resíduos 240-257), assegura a localização nuclear da proteína, a qual possui uma distribuição intranuclear em compartimentos específicos onde ocorre síntese de DNA;¹² e um motivo KOW (*Kyrpides Ouzouns Woese*) (resíduos 335-373), relacionados com interações proteína-proteína e proteína-RNA.²⁸⁻³²

A completa estrutura tridimensional da proteína KIN humana ainda não foi resolvida por métodos experimentais e nem por modelagem *in silico*, devido à falta de moldes estruturais para a maior parte da sequência, indicando uma estrutura ímpar para essa proteína. O estudo da estrutura proteica, bem como de suas propriedades biofísicas podem ser ferramentas importantes para prever sua função. Diante disso, o objetivo trabalho foi de realizar a análise filogenética dos diferentes domínios e a caracterização biofísica da proteína KIN humana, por meio de técnicas de bioinformática e espectroscopia de Dicroísmo Circular (CD, para a determinação da composição dos elementos de estrutura secundária e dos parâmetros biofísicos de HSAKIN, Entalpia de desnaturação (ΔH_m) e Temperatura de *melting* (T_m), ainda não realizados. Tais dados poderão ser ferramentas importantes para a elucidação da função de HSAKIN nos inúmeros processos onde foi proposta sua participação e potencial aplicação clínica na determinação de protocolos onde proteínas associadas à replicação possam servir de biomarcadores facilitando o diagnóstico, prognóstico e direcionando o tratamento para uma série de tipos de tumores.

Resultados

Análise computacional da sequência de HSAKIN.(His)₆

O gene clonado de HSAKIN, possuindo 1257 pares de bases (pb), foi inserido no plasmídeo pET23a-d(+). Esse gene foi traduzido em uma proteína quimérica (HSAKIN(His)_{6x}) que possui 418 resíduos de aminoácidos, com uma massa molecular de 48.032,65 Da e ponto isoelétrico teórico de 9,07 [Fig. S1]. Quando se analisou a existência de um peptídeo sinal, só foi verificada a presença de uma sequência de trânsito quando foi utilizado o software NLS Mapper,³³ um programa específico para a previsão de sinais de localização nuclear dependentes de importinas α . Foi verificado um NLS bipartido com a sequência sinal KRKESSQSSTQSKEKKKKKS, que foi denominado NLS+, com localização condizente com o apresentado por Kannouche *et al* (2000), mas que possui uma sequência com dois resíduos a mais no lado carboxi-terminal que o proposto anteriormente.¹² Não foi observado a presença de nenhuma ligação dissulfeto intra-cadeia para a sequência analisada.

As identidades entre as sequências das proteínas homólogas à HSAKIN para os 22 organismos analisados, em conjunto com os respectivos comprimentos das sequências, são mostradas na Tabela I. Com base nesta análise, pode-se observar um alto grau de identidade com a HSAKIN entre os animais, chegando a mais de 90% entre os mamíferos analisados. A menor identidade com a HSAKIN foi observada para a levedura *Saccharomyces cerevisiae*, que foi de 33,63%, o que, em modelagem por homologia *in silico*, já seria possível produzir um modelo estrutural com moderado grau de confiabilidade.³⁴⁻³⁶ Quando se comparou os comprimentos entre as sequências, obteve-se uma maior flutuação entre os valores dentro do grupo dos eucariotos inferiores e protozoários, contudo, a maioria das proteínas homólogas teve números perto de 390 resíduos de aminoácidos. Curiosamente, dentre os

cordados, pode-se observar a KIN homóloga de *Poeciliopsis prolifica* como um *outlier*, possuindo apenas 220 resíduos.

A análise dos domínios também foi realizada tomando a sequência de HSAKIN como referência (Tabela II). Dentre os domínios, o que possui maior porcentagem de identidade é o motivo dedo de zinco, responsável por permitir interação da proteína KIN com DNA e RNA. Foi observado em KIN de murinos que este domínio é similar ao da PARP (*poly ADP-ribose polymerase*), uma proteína que está envolvida no reparo do DNA.^{23,25} Entretanto, não foi observado este domínio na KIN de *Phytophthora nicotianae*. Similarmente ao motivo dedo de zinco, mas em menor grau, o motivo hélice alada se mostrou bastante conservado entre os animais e relativamente conservado entre os organismos do Reino *Plantae*, com mais de 50% de identidade para estes últimos. Para o motivo homólogo à RecA, se observou uma queda na identidade para todas as proteínas homólogas, menos para a da alga *Ostreococcus lucimarinus*. Tal queda de similaridade se manteve para o motivo NLS para todas as homólogas, menos para a KIN de *Brugia malayi*, que se manteve praticamente constante.

Também foram realizados testes para a averiguação da identidade com o NLS que encontramos (NLS+). Para a grande maioria das proteínas homólogas, foi verificado um aumento de identidade, em cerca de 18 das 22 HSAKIN homólogas, principalmente, entre os eucariotos mais desenvolvidos. Diferentemente, não foi observada nenhuma tendência de aumento ou diminuição da identidade para o motivo KOW, mas verificou-se entre os cordados que este se mantém conservado, com um mínimo de 51,28% de identidade para o urocordado *Ciona intestinalis*. Não foi observado os motivos NLS e KOW para as KIN homólogas de *Poeciliopsis prolifica* e *Saccharomyces cerevisiae* possivelmente devido as curtas cadeias de

suas proteínas.

Superexpressão e purificação de $HSAKIN(His)_{6x}$

A proteína recombinante $HSAKIN$ fusionada a uma cauda de histidina (*His-tag*) foi superexpressa *Escherichia coli* BL21 (DE3) e purificada em coluna de afinidade, resultando em um rendimento de 100-500 $\mu\text{g/mL}$. A proteína foi eluída com imidazol em dois passos, no primeiro foi feito um gradiente de concentração de imidazol de 10-40 mM para eluir as proteínas ligadas inespecificamente e, em seguida, foi feito outro gradiente de 40-300 mM, para a eluição das proteínas mais especificamente ligadas. A proteína depois de purificada apresentou-se insolúvel em altas concentrações, possivelmente, devido à presença de 7 resíduos de cisteína em sua cadeia, com potencial para a formação de pontes dissulfeto intercadeias. A proteína foi mantida em uma concentração de aproximadamente 1 mg/mL. A análise da purificação foi realizada por meio *SDS-Page* e *Western Blotting* [Fig. 2].³⁷

Predição dos componentes de estrutura secundária em diferentes condições de pH

A determinação da porcentagem dos elementos de estrutura secundária da amostra purificada foi realizada por meio da deconvolução do espectro de CD a 20 °C (Tabela III), por meio do programa CDPro.^{38,39} A análise dos espectros [Fig. 3] revelou que a proteína recombinante $HSAKIN(His)_{6x}$ possui a maior parte de sua composição de elementos não estruturados, como voltas β e *random coil*, perfazendo uma porcentagem de mais de 50% da estrutura, verificado em todos os três pHs analisados. Considerando os elementos estruturados α -hélices e folhas- β , a $HSAKIN(His)_{6x}$ possui uma porcentagem de folhas- β significativamente maior, cerca

de 30% para os pHs 6,4 e 7,4, e de 40% para pH 8,4. Já para α -hélice foram encontrados valores de 18, 17 e 9% para os respectivos pHs 6,4; 7,4 e 8,4. Estes dados corroboram com as composições estruturais das partes de HSAKIN já solucionadas por RMN e cristalografia de raios-X.^{24,25,31,32}

Determinação dos parâmetros ΔH_m e T_m

A variação espectral em 222 nm foi utilizada para acompanhar a variação do espectro de absorção em função da temperatura, de modo a determinar a temperatura de *melting* (T_m) e a entalpia de desnaturação (ΔH_m) [Fig. 4]. A determinação de ΔH_m foi realizada no ponto de transição ($\Delta G=0$). No pH 8,4, devido à proximidade do PI, a proteína se mostrou quase que completamente desnaturada e formou grânulos de precipitação, havendo uma grande perda na intensidade do sinal comparado aos outros pHs. Não foi observado um comportamento sigmoidal para a curva em pH 8,4, indicando que não houve transição entre o estado enovelado para o desnaturado. Para os pHs 6,4 foi observado uma perda de sinal em menor escala, se comparado ao pH 8,4, mas que ainda possibilitou a observação de uma faixa de transição. Para o pH 7,4 foi verificado um comportamento sigmoidal característico, com uma ampla faixa de transição entre os estados enovelado e desnaturado. Assim, foi realizada a determinação dos parâmetros físico-químicos apenas para as amostras nos pHs 6,4 e 7,4.

As curvas de desnaturação a 222 nm para os pHs 6,4 e 7,4 podem ser vistas na Fig. 5. Para determinação de ΔH_m e T_m utilizou-se a equação (1).⁴⁰ As curvas indicam uma transição endotérmica de desnaturação e os valores obtidos (Tabela IV) refletem a estabilidade térmica da proteína. Foram encontrados os valores de 241 e 279 kJ.mol⁻¹ para a ΔH_m e 45,6 e 50,5 °C de T_m para os pHs 6,4 e 7,4,

respectivamente. Esses valores demonstram uma perda de estabilidade térmica pela proteína em pH 6,4 em comparação ao pH 7,4. Tais dados corroboram com o esperado, onde HSAKIN(His)_{6x} demonstrou maior estabilidade em pH 7,4, devido a este ser o pH celular.

Discussão

Os resultados mostrados neste trabalho para os alinhamentos múltiplos tiveram o intuito de identificar a conservação evolutiva das posições dos aminoácidos em moléculas de proteínas com base nas relações filogenéticas entre as sequências homólogas. O grau em que a posição do aminoácido é evolutivamente conservada é fortemente dependente da sua importância estrutural e funcional, posições em rápida evolução são variáveis, enquanto as posições de evolução lenta são conservadas. Assim, a análise de conservação da posição entre os membros da mesma família pode frequentemente revelar a importância de cada posição para a estrutura ou função da proteína.⁴¹⁻⁴⁴ Desta forma, nossos dados obtidos a partir dos alinhamentos com a sequência completa das proteínas homólogas à KIN humana, mostraram que KIN é bastante conservada nos animais analisados, com mais de 50% de identidade, indicando que esta proteína está envolvida em processos biológicos fundamentais para a célula. Estes dados corroboram com a possível participação de KIN em diversos processos indispensáveis à manutenção e sobrevivência celular listados na literatura.^{15-17,22,27,45}

As análises de identidade dos domínios homólogos de HSAKIN , mostraram uma maior conservação dos domínios dedo de zinco e hélice alada. Assim, tais domínios podem ser os principais elementos responsáveis pela manutenção da estrutura e função da proteína KIN. Possivelmente, o alto grau de identidade desses

domínios pode sugerir que são responsáveis pelo processo mais fundamental à função desta proteína, a ligação à cromatina, ao qual já foi visto que esses dois domínios podem se associar, bem como a interação proteína-proteína.²³⁻²⁵ As proteínas homólogas de $_{\text{HSA}}\text{KIN}$ de *Saccharomyces cerevisiae* e *Poeciliopsis prolifica* reforçam essa hipótese, por possuírem apenas estes dois domínios, assim, possivelmente participando dos processos mais elementares ao qual KIN é requisitada.

Com base na comparação entre os motivos NLS+ e o NLS, foi verificado uma maior identidade para a grande maioria dos eucariotos superiores, principalmente entre os animais, quando se analisou NLS+. Esses dados indicam que a sequência sinal de localização nuclear seja maior em eucariotos superiores do que o proposto anteriormente.¹² Há uma relação de proporcionalidade direta, com o aumento da sequência, há um aumento da efetividade desta região como sinal para importinas α .⁴⁶ Embora seja difícil se estabelecer uma sequência consenso para os sinais de localização nuclear, os *clusters* do NLS de $_{\text{HSA}}\text{KIN}$ apresentaram uma composição bastante similar aos dos NLS bipartidos das proteínas nucleoplasmina de *Xenopus laevis* e nucleolina de *Gallus gallus*, e ao do NLS monopartido do antígeno-T do *simian virus 40* (SV40).⁴⁶⁻⁴⁹

Os dados oriundos das deconvoluções dos espectros de CD para os diferentes pHs, demonstram que a proteína $_{\text{HSA}}\text{KIN}(\text{His})_{6x}$ é composta em maior parte de elementos não estruturados. Tais elementos são comuns em proteínas que possuem multidomínios conectados por regiões flexíveis, o que sugere uma maior dinâmica estrutural para estas proteínas.^{50,51} Essas regiões não enoveladas desempenham funções biológicas fundamentais, participando em várias vias de sinalização e regulação, através de interações específicas proteína-proteína,

proteína-ácido nucleico, proteína-ligante e locais de modificações pós-translacionais, como metilações.⁵²⁻⁵⁵ A capacidade de regiões não enoveladas de se inter-relacionar com uma coleção de parceiros é utilizada na organização de complexas redes de interação proteína-proteína, servindo como uma âncora ou agindo como um andaime globular estável que interage com regiões intrinsecamente desordenadas dos seus alvos.⁵⁶⁻⁵⁹ Pode ser que estes elementos não estruturados de HSAKIN, com ação conjunta com os domínios hélice alada e KOW, ajudem-na a formar complexas redes de interação proteína-proteína, como uma proteína acessória, promovendo a formação de complexos proteicos de alto peso molecular que agem nos processos aos quais HSAKIN foi relacionada.^{45,60-62} Esses dados referentes à composição da estrutura secundária de HSAKIN(His)_{6x} podem ser extrapolados para as homólogas analisadas que apresentam altos valores de identidade de sequência, devido a proteínas de mesma família se enovelarem de maneira similar.⁶³

Analisando os dados da desnaturação térmica no comprimento de onda 222 nm, foi possível observar uma discrepância significativa entre os valores dos parâmetros físico-químicos para os dois pHs analisados. Uma diferença de 38 kJ.mol⁻¹ para ΔH_m e 4,9 °C para T_m foi observada, indicando uma maior fragilidade térmica para a proteína HSAKIN(His)_{6x} quando se encontra em pH 6,4. Valores positivos de ΔH_m são compatíveis com quebra de pontes de hidrogênio, bem como a exposição de grupos hidrofóbicos do interior do estado nativo (enovelado) da proteína durante o processo de desnaturação.⁶⁴ Desta maneira, mudanças na ionização dos resíduos de aminoácidos decorrente do baixo pH, assim, com suas cargas alteradas, diversas interações antes formadas podem estar previamente quebradas em pH 6,4, como por exemplo, ligações de hidrogênio, pontes salinas, contatos de van der Waals e outras interações hidrofóbicas.

Se compararmos os parâmetros físico-químicos encontrados para a proteína HSAKIN(His)_{6x} com o de outras proteínas, veremos valores de T_m relativamente baixos.⁶⁵ Quando relacionou-se os parâmetros biofísicos encontrados na literatura para ribonuclease H de *Escherichia coli*, proteína que também interage com ácidos nucleicos, observou-se valores bastante distintos, 502,5 KJ.mol⁻¹ e 66 °C para ΔH_m e T_m , respectivamente.⁶⁶ Comparou-se também estes parâmetro para a proteína porcina de *Sus scrofa*, uma proteína secretada no muco nasal de vertebrados, pertencente a supermfamília das lipocalinas, que interage especificamente com pequenos ligantes hidrofóbicos, assumindo o papel de carregadora durante a percepção olfatória.⁶⁷⁻⁶⁹ Foram encontrados os valores de 374,2 KJ.mol⁻¹ para ΔH_m e 69,15 °C para T_m para a proteína porcina.⁷⁰ O servidor EDBCP previu,⁷¹ a partir da sequência de aminoácidos, que não há pontes dissulfeto na estrutura de HSAKIN . Portanto, estes valores substancialmente baixos encontrados para os parâmetros físico-químicos ΔH_m e T_m de HSAKIN(His)_{6x} , muito se devem a esta proteína não possuir nenhuma ponte dissulfeto intra-cadeia, mesmo contando com 7 cisteínas em sua sequência. Possivelmente, a não ocorrência dessas ligações em HSAKIN(His)_{6x} pode ser devido a diversos impedimentos estéricos impostos pela estrutura enovelada da proteína.

Este trabalho fornece resultados que permitem concluir que a proteína HSAKIN é bastante conservada entre os animais, possuindo um maior grau de identidade nos domínios dedo de zinco e hélice alada, indicando que estes domínios são os responsáveis por funções mais elementares de KIN, como a ligação ao DNA e interação proteína-proteína no complexo de replicação. Quanto ao motivo NLS+ mostrou-se mais conservado entre a maioria das proteínas homólogas analisadas em relação ao NLS anteriormente descrito. A proteína recombinante HSAKIN(His)_{6x}

mostrou possuir mais do que 50% de sua estrutura secundária de elementos não estruturados, como voltas- β e *random coil*. Tais elementos podem contribuir para a formação de complexos de alto peso molecular envolvidos nos processos aos quais $HSAKIN$ foi correlacionada. Estes resultados podem ser extrapolados para as proteínas homólogas analisadas que apresentaram alta identidade com $HSAKIN$. As análises de desnaturação térmicas em 222 nm mostraram que em pH 7,4 $HSAKIN(His)_{6x}$ foram encontrados os mais altos valores de ΔH_m e T_m , indicando uma maior estabilidade térmica neste pH.

Material e Métodos

Material biológico

O Vetor de expressão pET23a-d(+) (Novagen, USA) contendo o cDNA de $HSAKIN$ [pET23a-d(+) $HSAKIN(His)_{6x}$], foi cedido gentilmente pelo Professor PhD Benoit Coulombe do Institut de Recherches Cliniques de Montréal, Québec, Canadá. O cDNA da $HSAKIN$ foi clonado no vetor pET23a-d(+) utilizando as endonucleases de restrição NheI e NotI, gerando um plasmídeo que está sob o controle do promotor T7 e codifica uma cauda com 6 resíduos de histidina na extremidade C-terminal de $HSAKIN$, gerando a proteína quimérica $HSAKIN.(His)_6$. A cepa *Escherichia coli* BL21 (DE3) que possui o gene T7 RNA Polimerase sob o controle do promotor *lacUV5* induzível por IPTG (*Isopropyl β -D-1-thiogalactopyranoside*) foi utilizada para a superexpressão da proteína $(HSA)KIN.(His)_6$.

Análise computacional

O gene da proteína $HSAKIN(His)_{6x}$ foi traduzida, para o conhecimento de sua sequência e quantidade exata de resíduos em sua cadeia principal, pela ferramenta

EMBOSS Transeq.⁷² Para a determinação da massa de HSAKIN(His)_{6x} e ponto isoelétrico teórico, utilizou-se a ferramenta ProtParam.⁷³ A predição de peptídeos de trânsito nuclear foi realizada pelo servidor SignalP 4.1⁷⁴ e o software NLS Mapper.³³ Adicionalmente a aferição de sequências sinais, utilizou-se a ferramenta InterProScan sequence search,⁷⁵ para verificar os domínios conservados da proteína KIN. Utilizou-se para a predição de ligações dissulfeto o servidor EDBCP (*Ensemble-based Disulfide Bonding Connectivity Pattern Prediction*).⁷¹

A análise de proteínas homólogas à HSAKIN foi realizada usando o programa Clusatal Omega,⁷⁵ para o alinhamento e verificação da identidade entre as sequências dos domínios dos diferentes organismos retirados da base de dados *National Center for Biotechnology Information*, NCBI (<https://www.ncbi.nlm.nih.gov/>).

Transformação das bactérias

As bactérias *Escherichia coli* BL21 (DE3) foram transformadas com o vetor pET23a-d(+)_{HSAKIN(His)_{6x}} por choque térmico.³⁷ A verificação da transformação foi obtida por PCR convencional [Fig. 6], devido ao vetor pET23a-d(+)_{HSAKIN(His)_{6x}} não possuir sistema de seleção azul/branco, após cultura por 12-16 horas a 37°C em meio LB (Luria Bertani) com 50 µg/mL de ampicilina e por seguinte extração pelo método do CTAB⁷⁶. Utilizou-se os *primers* universal 5'-AGCGCTAGCATGGGGAAGTCGGATTTTCT-3' e reverso 5'- AGTGCGGCCGCGGCAAGTTTAGAAATGTCTT-3', *Taq* DNA Polimerase (Invitrogen, USA) e condições padrão: desnaturação inicial 94 °C por 5 min; seguida de 30 ciclos de 94 °C por 1 min, 60 °C por 1 min e 72 °C por 1 min; e uma extensão final de 72 °C por 10 min. A análise visual da PCR foi realizada em gel de agarose 0.7% com tampão TBE 0.5X (Tris/Boreto/EDTA) e brometo de etídio 1 µg/mL.

Expressão e purificação da proteína HSAKIN(His)_{6x}

Culturas de *Escherichia coli* BL21 (DE3) transformadas com pET23a-d(+) HSAKIN(His)_{6x} foram crescidas aerobicamente em fresco Erlenmeyer contendo meio LB com 100 $\mu\text{g/mL}$ de ampicilina. A quantidade de meio utilizada não ultrapassou 20% do volume total do frasco para uma melhor oxigenação do meio, e incubou-se a 37°C em shaker a 75 rpm até atingir densidade óptica em 600 nm (OD_{600}) de 0,8. Então, as culturas foram induzidas aerobicamente com 0,4 mM de IPTG e mantidas por 8h a 30 °C e 75 rpm. Depois da indução a cultura foi centrifugada a 10.000 x g por 10 min a 4 °C e o pellet lavado com Tris-HCl 20 mM, pH 8,0 e armazenado a -20 °C por 12-16 horas antes da extração.

O pellet da cultura foi ressuscitado em tampão de lise (50 mM NaH_2PO_4 , 300 mM NaCl e 10 mM Imidazol, pH 7.4 com 0,5 mM de PMSF e inibidor de protease 1X) e sonicado no gelo por 4 minutos com pulsos de 20 seg On e 20 seg Off com amplitude de 60%, para o rompimento das células em sonicador Fisher Scientific. As células rompidas foram centrifugadas a 10.000 x g por 10 minutos a 4°C e tiveram o sobrenadante filtrado em filtro de seringa Millex® de 0,22 μm (Merck-Millipore), para a separação da fração de proteínas solúveis.

A purificação da proteína HSAKIN(His)_{6x} ocorreu a partir da fração solúvel por meio de cromatografia de afinidade em FPLC, *Fast Protein Liquid Chromatography* (ÄKTA Purifier, GE), e coluna de níquel-sefarose de 1 mL (HisTrap, GE). Anteriormente a purificação, a coluna foi lavada com 10 volumes de coluna de água e equilibrada com 10 volumes de tampão de ligação (50 mM NaH_2PO_4 , 300 mM NaCl e 10 mM Imidazol, pH 7.4). A amostra foi injetada no equipamento utilizando um Superloop de 10 mL. A proteína (HSA)KIN.(His)_6 foi eluída com gradiente de tampão de eluição (50 mM NaH_2PO_4 , 300 mM NaCl e 300 mM Imidazol, pH 7.4). A

amostra eluída foi injetada em coluna de exclusão molecular Sefadex G25 (HiTrap *Desalting*, GE) para a troca de tampão (20 mM NaHPO₄/NaH₂PO₄, 10 mM NaF, pH 7,4). As amostras foram concentradas a 1 mg/mL usando filtros de centrifuga com *Cut Off* de 10 kDa (Amicon®, Merck-Millipore) em centrifuga a 1.000 x g à 4°C e quantificadas usando o kit Pierce™ BCA Protein Assay (Thermo Fisher Scientific). A análise da purificação foi realizada por meio da técnica SDS-Page (Gel de empilhamento 5% e gel de resolução 12%) e *Western Blotting*.³⁷ Para a imunodeteção, foram utilizados o anticorpo primário K58 (Santa Cruz Biotech) e o secundário *polyclonal goat anti-mouse* com HRP (Dako).

Avaliação dos parâmetros biofísicos por dicroísmo circular

A análise da estrutura secundária e termodinâmica de desnaturação em 222 nm foi realizada em espectro-polarímetro de dicroísmo circular (CD), modelo Jasco J815 equipado com um controlador de temperatura do tipo Peltier e um suporte de célula termostática, interligado a um banho termostático. Para todas as análises utilizou-se cubeta de quartzo com caminho óptico de 5,0 mm e com uma concentração de proteína de 7,75 µM. Foram coletados espectros de CD e Absorção de luz entre os comprimentos de onda 190 a 260 nm. Cada espectro foi coletado em uma temperatura distinta entre 20 °C e 90 °C em intervalos de 2 °C. A estabilidade da proteína foi determinada em tampão fosfato (20 mM NaHPO₄/NaH₂PO₄, 10 mM NaF), para três diferentes pHs (6,4; 7,4 e 8,4). As curvas de desnaturação térmica foram ajustadas à Equação (1),⁴⁰ assumindo uma reação de dois estados, nativo e desnaturado.

$$S_{obs} = \frac{S_f + m_f T + (S_u + m_u T)K_u}{1 + K_u} \quad (1)$$

Onde

$$K_u = \exp\left(\frac{\Delta H_m \left(1 - \frac{T}{T_m}\right)}{RT}\right) \quad (2)$$

Aqui, S_{obs} é o sinal espectroscópico dependente da temperatura, m_f , m_u , S_f e S_u são as inclinações e intercepções das linhas de base nativa e desnaturada, respectivamente, T é a temperatura absoluta, T_m é o ponto de *melting*, K_u é a constante de equilíbrio de desnaturação e ΔH_m é a entalpia de van't Hoff de desnaturação em T_m .

Referências Bibliográficas

1. Aguilera A, Gomez-Gonzalez B. Genome instability: a mechanistic view of its causes and consequences. *Nature reviews. Genetics* 2008;9(3):204-17.
2. Segurado M, Tercero JA. The S-phase checkpoint: targeting the replication fork. *Biology of the cell* 2009;101(11):617-27.
3. Angulo JF, Moreau PL, Maunoury R, Laporte J, Hill AM, Bertolotti R, Devoret R. KIN, a mammalian nuclear protein immunologically related to E. coli RecA protein. *Mutation research* 1989;217(2):123-34.
4. Mazin A, Milot E, Devoret R, Chartrand P. KIN17, a mouse nuclear protein, binds to bent DNA fragments that are found at illegitimate recombination junctions in mammalian cells. *Molecular & general genetics : MGG* 1994;244(4):435-8.
5. Timchenko T, Bailone A, Devoret R. Btcd, a mouse protein that binds to curved DNA, can substitute in Escherichia coli for H-NS, a bacterial nucleoid protein. *The EMBO journal* 1996;15(15):3986-92.
6. Angulo JF, Rouer E, Mazin A, Mattei MG, Tissier A, Horellou P, Benarous R, Devoret R. Identification and expression of the cDNA of KIN17, a zinc-finger gene located on mouse chromosome 2, encoding a new DNA-binding protein. *Nucleic acids research* 1991;19(19):5117-23.
7. Misteli T, Spector DL. The cellular organization of gene expression. *Current opinion in cell biology* 1998;10(3):323-31.
8. Haaf T, Golub EI, Reddy G, Radding CM, Ward DC. Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage

- and its localization in synaptonemal complexes. *Proceedings of the National Academy of Sciences of the United States of America* 1995;92(6):2298-302.
9. Nelms BE, Maser RS, MacKay JF, Lagally MG, Petrini JH. In situ visualization of DNA double-strand break repair in human fibroblasts. *Science* 1998;280(5363):590-2.
 10. Newport J, Yan H. Organization of DNA into foci during replication. *Current opinion in cell biology* 1996;8(3):365-8.
 11. Zeng C, Kim E, Warren SL, Berget SM. Dynamic relocation of transcription and splicing factors dependent upon transcriptional activity. *The EMBO journal* 1997;16(6):1401-12.
 12. Kannouche P, Mauffrey P, Pinon-Lataillade G, Mattei MG, Sarasin A, Daya-Grosjean L, Angulo JF. Molecular cloning and characterization of the human KIN17 cDNA encoding a component of the UVC response that is conserved among metazoans. *Carcinogenesis* 2000;21(9):1701-10.
 13. Cloutier P, Coulombe B. Regulation of molecular chaperones through post-translational modifications: decrypting the chaperone code. *Biochimica et biophysica acta* 2013;1829(5):443-54.
 14. Cloutier P, Lavalleye-Adam M, Faubert D, Blanchette M, Coulombe B. Methylation of the DNA/RNA-binding protein Kin17 by METTL22 affects its association with chromatin. *Journal of proteomics* 2014;100:115-24.
 15. Biard DS, Saintigny Y, Maratrat M, Paris F, Martin M, Angulo JF. Enhanced expression of the Kin17 protein immediately after low doses of ionizing radiation. *Radiation research* 1997;147(4):442-50.
 16. Biard DS, Miccoli L, Despras E, Harper F, Pichard E, Creminon C, Angulo JF. Participation of kin17 protein in replication factories and in other DNA transactions mediated by high molecular weight nuclear complexes. *Molecular cancer research : MCR* 2003;1(7):519-31.
 17. Despras E, Miccoli L, Creminon C, Rouillard D, Angulo JF, Biard DS. Depletion of KIN17, a human DNA replication protein, increases the radiosensitivity of RKO cells. *Radiation research* 2003;159(6):748-58.
 18. Yang YL, Suen J, Brynildsen MP, Galbraith SJ, Liao JC. Inferring yeast cell cycle regulators and interactions using transcription factor activities. *BMC genomics* 2005;6:90.
 19. Garcia-Molina A, Xing S, Huijser P. The Arabidopsis KIN17 and its homolog KLP mediate different aspects of plant growth and development. *Plant signaling & behavior* 2014;9(3):e28634.
 20. Garcia-Molina A, Xing S, Huijser P. A conserved KIN17 curved DNA-binding domain protein assembles with SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE7 to adapt Arabidopsis growth and development to limiting copper availability. *Plant physiology* 2014;164(2):828-40.
 21. Araneda S, Angulo J, Touret M, Sallanon-Moulin M, Souchier C, Jouvét M. Preferential expression of kin, a nuclear protein binding to curved DNA, in the neurons of the adult rat. *Brain research* 1997;762(1-2):103-13.
 22. Le MX, Haddad D, Ling AK, Li C, So CC, Chopra A, Hu R, Angulo JF, Moffat J, Martin A. Kin17 facilitates multiple double-strand break repair pathways that govern B cell class switching. *Scientific reports* 2016;6:37215.
 23. Mazin A, Timchenko T, Menissier-de Murcia J, Schreiber V, Angulo JF, de Murcia G, Devoret R. Kin17, a mouse nuclear zinc finger protein that binds preferentially to curved DNA. *Nucleic acids research* 1994;22(20):4335-41.

24. Carlier L, le Maire A, Braud S, Masson C, Gondry M, Zinn-Justin S, Guilhaudis L, Milazzo I, Davoust D, Gilquin B and others. NMR assignment of region 51-160 of human KIN17, a DNA and RNA-binding protein. *Journal of biomolecular NMR* 2006;36 Suppl 1:29.
25. Carlier L, Couprie J, le Maire A, Guilhaudis L, Milazzo-Segalas I, Courcon M, Moutiez M, Gondry M, Davoust D, Gilquin B and others. Solution structure of the region 51-160 of human KIN17 reveals an atypical winged helix domain. *Protein science : a publication of the Protein Society* 2007;16(12):2750-5.
26. Angulo JF, Rouer E, Benarous R, Devoret R. Identification of a mouse cDNA fragment whose expressed polypeptide reacts with anti-recA antibodies. *Biochimie* 1991;73(2-3):251-6.
27. Kannouche P, Pinon-Lataillade G, Mauffrey P, Faucher C, Biard DS, Angulo JF. Overexpression of kin17 protein forms intranuclear foci in mammalian cells. *Biochimie* 1997;79(9-10):599-606.
28. Steiner T, Kaiser JT, Marinkovic S, Huber R, Wahl MC. Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *The EMBO journal* 2002;21(17):4641-53.
29. Kyrpides NC, Woese CR, Ouzounis CA. KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends in biochemical sciences* 1996;21(11):425-6.
30. Ponting CP. Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic acids research* 2002;30(17):3643-52.
31. le Maire A, Schiltz M, Braud S, Gondry M, Charbonnier JB, Zinn-Justin S, Stura E. Crystallization and halide phasing of the C-terminal domain of human KIN17. *Acta crystallographica. Section F, Structural biology and crystallization communications* 2006;62(Pt 3):245-8.
32. le Maire A, Schiltz M, Stura EA, Pinon-Lataillade G, Couprie J, Moutiez M, Gondry M, Angulo JF, Zinn-Justin S. A tandem of SH3-like domains participates in RNA binding in KIN17, a human protein activated in response to genotoxics. *Journal of molecular biology* 2006;364(4):764-76.
33. Kosugi S, Hasebe M, Tomita M, Yanagawa H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106(25):10171-6.
34. D'Alfonso G, Tramontano A, Lahm A. Structural conservation in single-domain proteins: implications for homology modeling. *Journal of structural biology* 2001;134(2-3):246-56.
35. Sali A. 100,000 protein structures for the biologist. *Nature structural biology* 1998;5(12):1029-32.
36. Vitkup D, Melamud E, Moulton J, Sander C. Completeness in structural genomics. *Nature structural biology* 2001;8(6):559-66.
37. Sambrook J, Russell DW. *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2001.
38. Johnson WC. Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* 1999;35(3):307-12.
39. Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical biochemistry* 2000;287(2):252-60.

40. Olsson U, Wolf-Watz M. Overlap between folding and functional energy landscapes for adenylate kinase conformational change. *Nature communications* 2010;1:111.
41. Paiardini A, Bossa F, Pascarella S. Evolutionarily conserved regions and hydrophobic contacts at the superfamily level: The case of the fold-type I, pyridoxal-5'-phosphate-dependent enzymes. *Protein science : a publication of the Protein Society* 2004;13(11):2992-3005.
42. Jones S, Thornton JM. Searching for functional sites in protein structures. *Current opinion in chemical biology* 2004;8(1):3-7.
43. Pils B, Copley RR, Schultz J. Variation in structural location and amino acid conservation of functional sites in protein domain families. *BMC bioinformatics* 2005;6:210.
44. Vogel C, Bashton M, Kerrison ND, Chothia C, Teichmann SA. Structure, function and evolution of multidomain proteins. *Current opinion in structural biology* 2004;14(2):208-16.
45. Biard DS, Miccoli L, Despras E, Frobert Y, Creminon C, Angulo JF. Ionizing radiation triggers chromatin-bound kin17 complex formation in human cells. *The Journal of biological chemistry* 2002;277(21):19156-65.
46. Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *The Journal of cell biology* 1988;107(3):841-9.
47. Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* 1991;64(3):615-23.
48. Schmidt-Zachmann MS, Nigg EA. Protein localization to the nucleolus: a search for targeting domains in nucleolin. *Journal of cell science* 1993;105 (Pt 3):799-806.
49. Kalderon D, Richardson WD, Markham AF, Smith AE. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 1984;311(5981):33-8.
50. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nature reviews. Molecular cell biology* 2005;6(3):197-208.
51. Dunker AK, Oldfield CJ. Back to the Future: Nuclear Magnetic Resonance and Bioinformatics Studies on Intrinsically Disordered Proteins. *Advances in experimental medicine and biology* 2015;870:1-34.
52. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, Obradovic Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *Journal of proteome research* 2007;6(5):1882-98.
53. Vucetic S, Xie H, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *Journal of proteome research* 2007;6(5):1899-916.
54. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *Journal of proteome research* 2007;6(5):1917-32.
55. Tompa P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS letters* 2005;579(15):3346-54.

56. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *The FEBS journal* 2005;272(20):5129-48.
57. Oldfield CJ, Meng J, Yang JY, Yang MQ, Uversky VN, Dunker AK. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC genomics* 2008;9 Suppl 1:S1.
58. Patil A, Nakamura H. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS letters* 2006;580(8):2041-5.
59. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS computational biology* 2006;2(8):e100.
60. Masson C, Mena F, Pinon-Lataillade G, Frobert Y, Chevillard S, Radicella JP, Sarasin A, Angulo JF. Global genome repair is required to activate KIN17, a UVC-responsive gene involved in DNA replication. *Proceedings of the National Academy of Sciences of the United States of America* 2003;100(2):616-21.
61. Miccoli L, Frouin I, Novac O, Di Paola D, Harper F, Zannis-Hadjopoulos M, Maga G, Biard DS, Angulo JF. The human stress-activated protein kin17 belongs to the multiprotein DNA replication complex and associates in vivo with mammalian replication origins. *Molecular and cellular biology* 2005;25(9):3814-30.
62. Miccoli L, Biard DS, Creminon C, Angulo JF. Human kin17 protein directly interacts with the simian virus 40 large T antigen and inhibits DNA replication. *Cancer research* 2002;62(19):5425-35.
63. Brändén C-I, Tooze J. *Introduction to protein structure*. New York: Garland Pub.; 1999. xiv, 410 p. p.
64. Cantor CR, Schimmel PR. *Biophysical chemistry*. San Francisco: W. H. Freeman; 1980. v. <1 > p.
65. Ku T, Lu P, Chan C, Wang T, Lai S, Lyu P, Hsiao N. Predicting melting temperature directly from protein sequences. *Computational biology and chemistry* 2009;33(6):445-50.
66. Hollien J, Marqusee S. A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. *Biochemistry* 1999;38(12):3831-6.
67. Breer H. Odor recognition and second messenger signaling in olfactory receptor neurons. *Seminars in cell biology* 1994;5(1):25-32.
68. Pelosi P. Odorant-binding proteins. *Critical reviews in biochemistry and molecular biology* 1994;29(3):199-228.
69. Flower DR. The lipocalin protein family: structure and function. *The Biochemical journal* 1996;318 (Pt 1):1-14.
70. Burova TV, Choiset Y, Jankowski CK, Haertle T. Conformational stability and binding properties of porcine odorant binding protein. *Biochemistry* 1999;38(45):15043-51.
71. Wedemeyer WJ, Welker E, Narayan M, Scheraga HA. Disulfide bonds and protein folding. *Biochemistry* 2000;39(15):4207-16.
72. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research* 2013;41(Web Server issue):W597-600.

73. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology* 1999;112:531-52.
74. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* 2011;8(10):785-6.
75. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G and others. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 2014;30(9):1236-40.
76. Del Sal G, Manfioletti G, Schneider C. The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. *BioTechniques* 1989;7(5):514-20.

Tabelas

Tabela I – Análise da identidade das proteínas homólogas a HSAKIN

Organismo	Comprimento da Cadeia	Identidade	Identificação do Gene (GI)
1. <i>Homo sapiens</i>*	393	100%	3850704
2. <i>Macaca mulatta</i>	393	98,73%	387763229
3. <i>Mus musculus</i>	391	92,84%	33859684
4. <i>Gallus gallus</i>	390	88,97%	363727398
5. <i>Aptenodytes forsteri</i>	362	85,36%	686588472
6. <i>Xenopus tropicalis</i>	389	81,49%	52345552
7. <i>Xenopus laevis</i>	387	78,81%	148235927
8. <i>Danio rerio</i>	385	77,08%	82524382
9. <i>Poeciliopsis prolifica</i>	220	82,65%	958279811
10. <i>Alligator mississippiensis</i>	391	90,28%	951038934
11. <i>Malus domestica</i>	394	37,17%	657944252
12. <i>Cucumis melo</i>	397	47,30%	659098345
13. <i>Arabidopsis thaliana</i>	411	50,26%	18405389
14. <i>Drosophila melanogaster</i>	390	51,96%	7296288
15. <i>Culex quinquefasciatus</i>	386	55,50%	167874937
16. <i>Ciona intestinalis</i>	390	62,92%	699240770
17. <i>Brugia malayi</i>	432	47,68%	170585660
18. <i>Plasmodium berghei</i>	443	35,57%	1063418344
19. <i>Toxoplasma gondii</i>	516	41,43%	820697156
20. <i>Phytophthora nicotianae</i>	329	45,31%	970648244
21. <i>Saccharomyces cerevisiae</i>	232	33,63%	1420235
22. <i>Schizosaccharomyces pombe</i>	304	40,86%	19112830
23. <i>Ostreococcus lucimarinus</i>	327	38,17%	144576639

* As sequências foram alinhadas tendo como organismo referência o *Homo sapiens*.

Tabela II – Análise da conservação dos domínios das proteínas homólogas a KIN humana

Organismo	Dedo de Zinco (28-50)	Hélice Alada (51-160)	RecA (162-201)	NLS (240-258)	NLS+ (240-260)	KOW (335-373)
1. <i>Homo sapiens</i> *	100	100	100	100	100	100
2. <i>Macaca mulatta</i>	100	100	97,50	88,89	90,00	100
3. <i>Mus musculus</i>	100	100	92,50	81,25	83,33	94,87
4. <i>Gallus gallus</i>	100	95,45	82,50	82,35	84,21	79,49
5. <i>Aptenodytes forsteri</i>	54,55	96,36	82,50	82,35	84,21	79,49
6. <i>Xenopus tropicalis</i>	100	91,82	75,00	53,33	58,82	69,23
7. <i>Xenopus laevis</i>	100	89,09	75,00	53,33	58,82	66,67
8. <i>Danio rerio</i>	100	88,18	67,50	62,50	66,67	64,10
9. <i>Poeciliopsis prolifica</i>	100	88,18	67,50	naln	naln	naln
10. <i>Alligator mississippiensis</i>	100	99,09	87,50	82,35	84,21	79,49
11. <i>Ciona intestinalis</i>	100	71,82	46,15	27,78	35,00	51,28
12. <i>Malus domestica</i>	100	50,91	30,00	0,00	8,33	7,69
13. <i>Cucumis melo</i>	100	50,91	30,00	27,27	15,38	38,46
14. <i>Arabidopsis thaliana</i>	100	50,91	30,00	16,67	35,71	33,33
15. <i>Drosophila melanogaster</i>	91,30	58,18	48,72	7,14	12,50	41,03
16. <i>Culex quinquefasciatus</i>	91,30	59,09	56,76	28,57	31,25	53,85
17. <i>Brugia malayi</i>	78,26	53,70	32,50	33,33	30,00	41,03
18. <i>Plasmodium berghei</i>	82,61	44,44	25,00	10,00	9,09	33,33
19. <i>Toxoplasma gondii</i>	86,96	47,22	35,14	22,22	15,00	33,33
20. <i>Phytophthora nicotianae</i>	naln	66,67	35,14	25,00	11,11	38,46
21. <i>Saccharomyces cerevisiae</i>	60,87	32,38	21,05	naln	naln	naln
22. <i>Schizosaccharomyces pombe</i>	73,91	43,64	20,00	14,29	21,43	naln
23. <i>Ostreococcus lucimarinus</i>	78,26	36,70	38,46	0,00	0,00	33,33

* As sequências foram alinhadas tendo como organismo referência o *Homo sapiens*.

naln = não alinhado

Tabela III – Porcentagens de elementos de estrutura secundária de HSAKIN(His)_{6x} em diferentes condições de pH.

Elemento de Estrutura (%)	pH 6.4	pH 7.4	pH 8.4
	20 °C	20 °C	20 °C
α -helice	18	17	9
Folha- β	30	31	40
Volta- β	30	30	29
<i>Random coil</i>	22	22	22

Tabela IV – Determinação dos parâmetros ΔH_m e T_m para dois distintos pHs

pH	ΔH_m (kJ.mol⁻¹)	T_m (°C)
6,4	241	45,6
7,4	279	50,5

Legendas das Figuras

Figura 1 - Domínios funcionais da proteína KIN. ZF (dedo de zinco); WH (domínio hélice alada); RH (domínio de homologia a proteína RecA de *E. coli*); NLS (sinal de localização nuclear); KOW (motivo KOW).

Figura 2 - Análise da purificação da proteína HSAKIN(His)_{6x} por meio de SDS-Page e *Western Blotting*. **A** - Purificação de HSAKIN(His)_{6x} em extrato de *E. coli* BL21 (DE3) verificada em dois géis SDS-Page 12%, gel A e B, corados com Coomassie Brilliant Blue g250. MM e FT correspondem, respectivamente, a marcador de massa molecular e *Flow Through*. Os números indicam a ordem das frações coletadas na fase de eluição da purificação. Os géis foram revelados no equipamento *ImageQuant LAS 500* (GE). O marcador de massa molecular utilizado foi o *Full Range Rainbow* (GE). **B** – Imunodeteção de HSAKIN(His)_{6x} utilizando o anticorpo K58 (Santa Cruz Biotech).

Figura 3 – Espectro de dicroísmo circular para HSAKIN(His)_{6x} em diferentes pHs. Todas as amostras estavam em tampão fosfato 20 mM e 10 mM de NaF.

Figura 4 – Curva de desnaturação térmica a 222 nm para a proteína recombinante HSAKIN(His)_{6x} em três diferentes pHs.

Figura 5 - Determinação dos parâmetros físico-químicos ΔH_m e T_m para a proteína quimérica HSAKIN(His)_{6x} . **A**- Mostra a curva de desnaturação para o pH 6,4 e **B** – representa a curva de desnaturação para o pH 7,4.

Figura 6 - Verificação por PCR da transformação de *E. coli* BL21 (DE3) utilizando o plasmídeo pET23a-d(+) HSAKIN(His)_{6x} . O marcador molecular (MM) utilizado foi o 2-log (NEB).

Figuras (Arte)

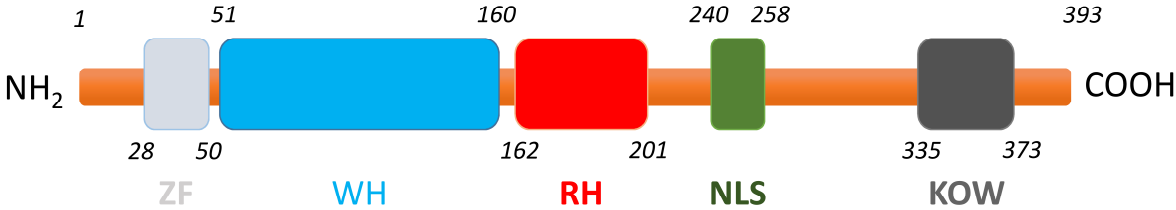


Fig. 1

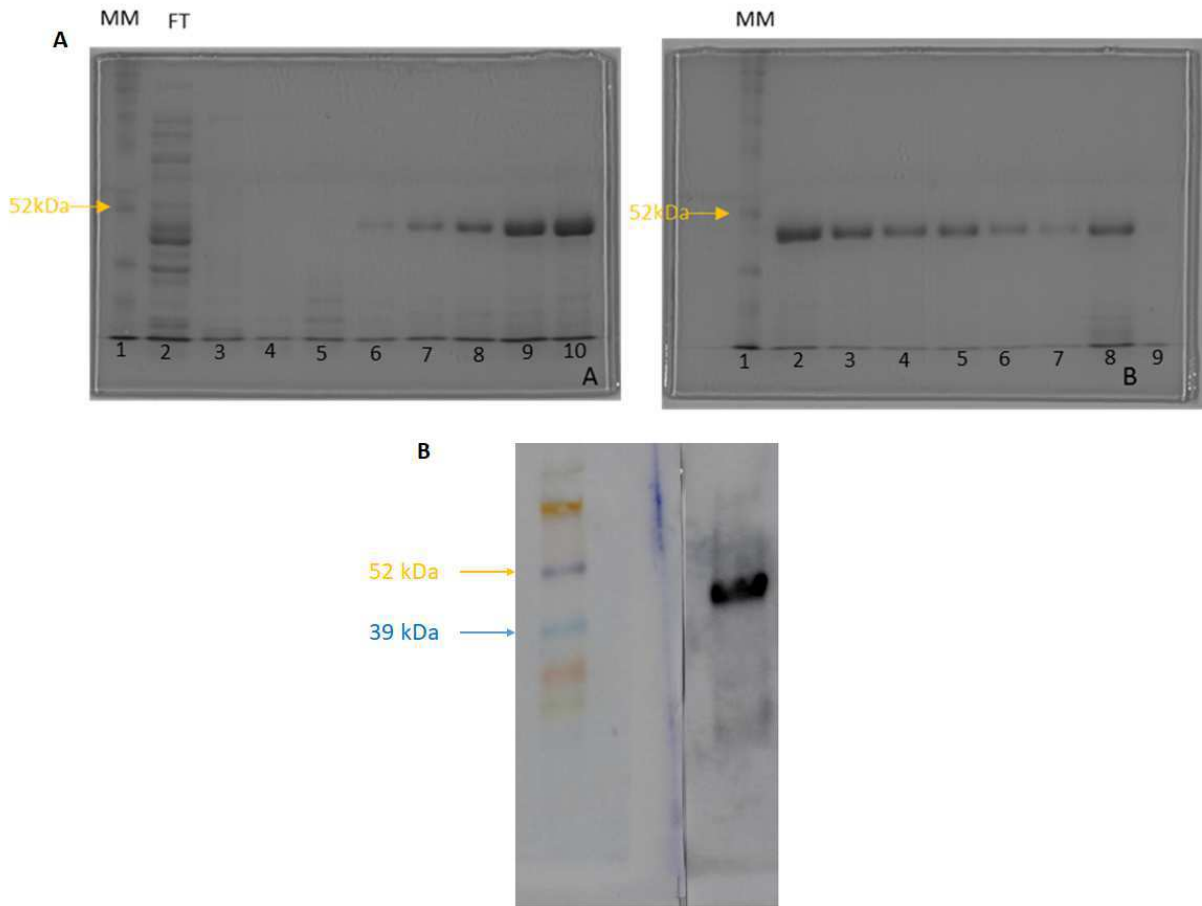


Fig. 2

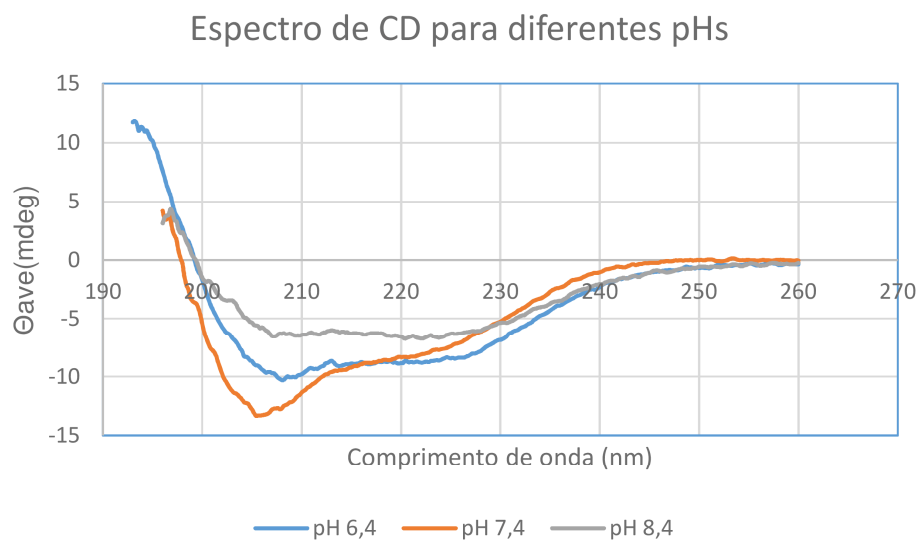


Fig. 3

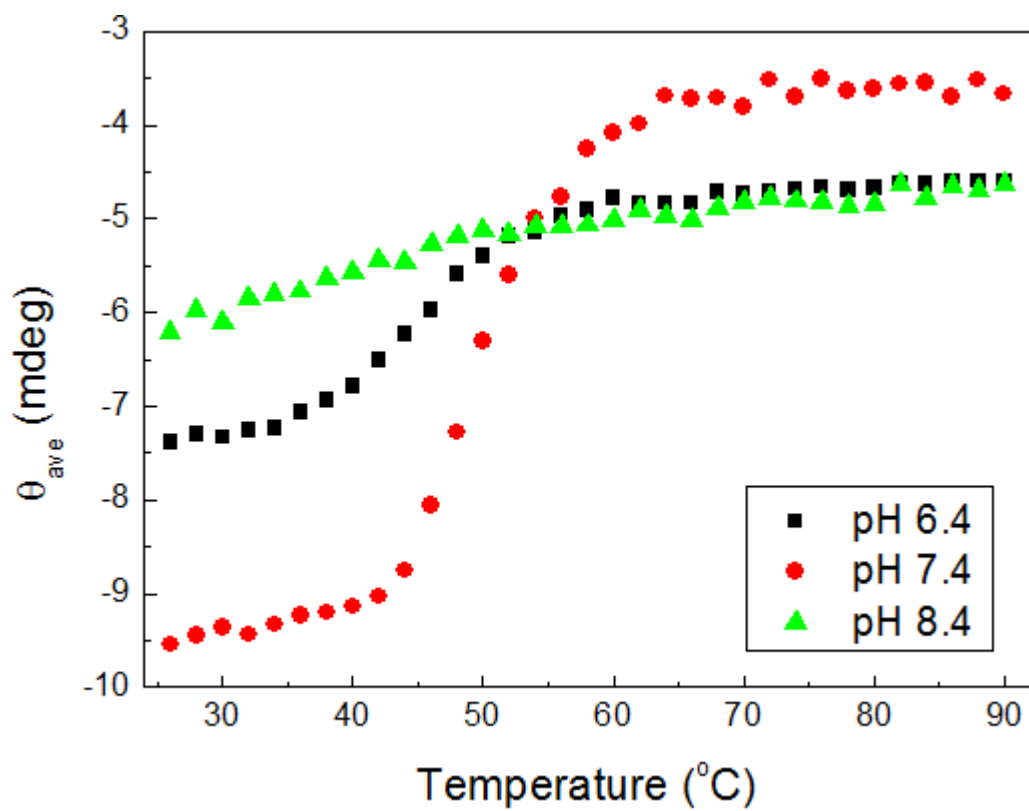


Fig. 4

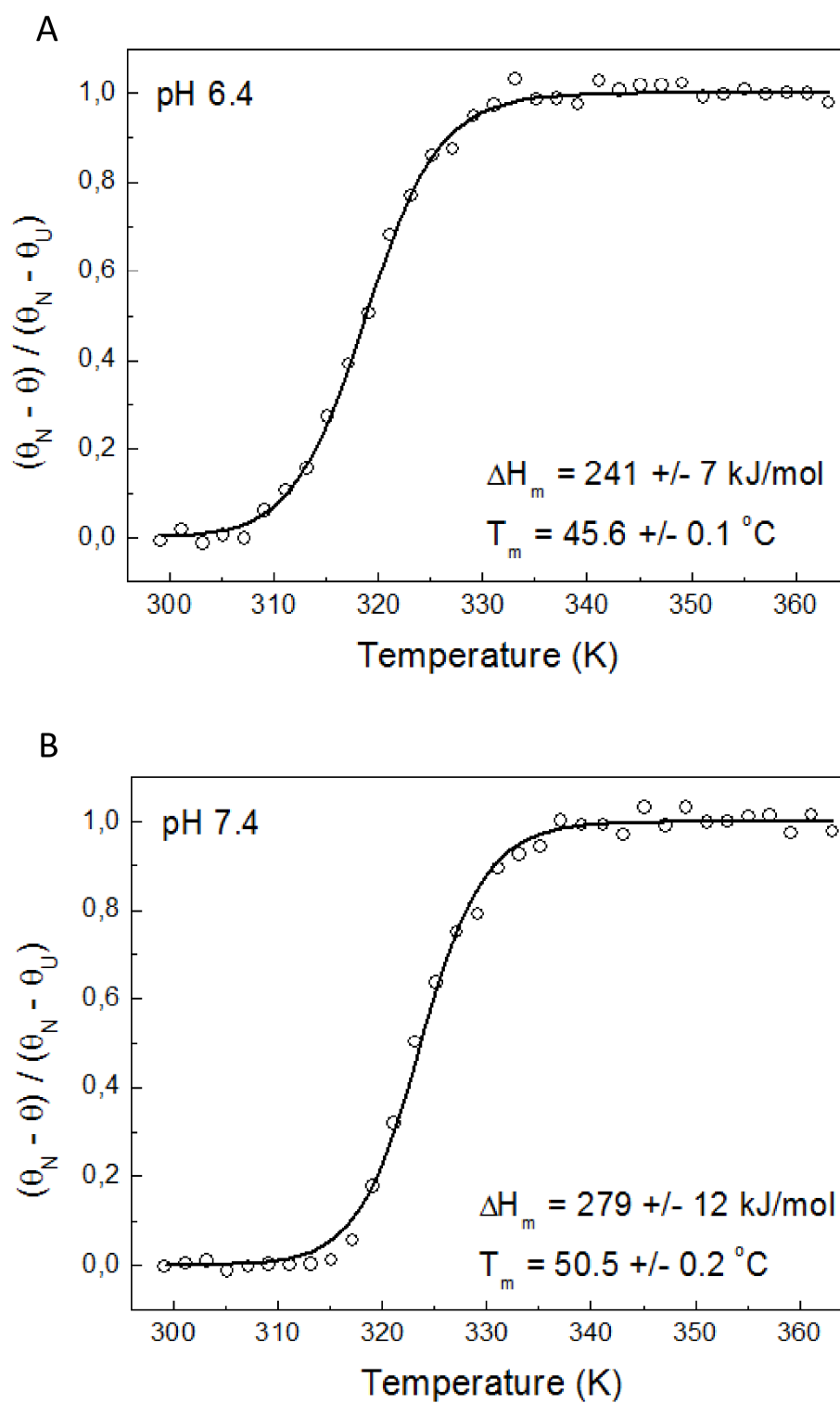


Fig. 5

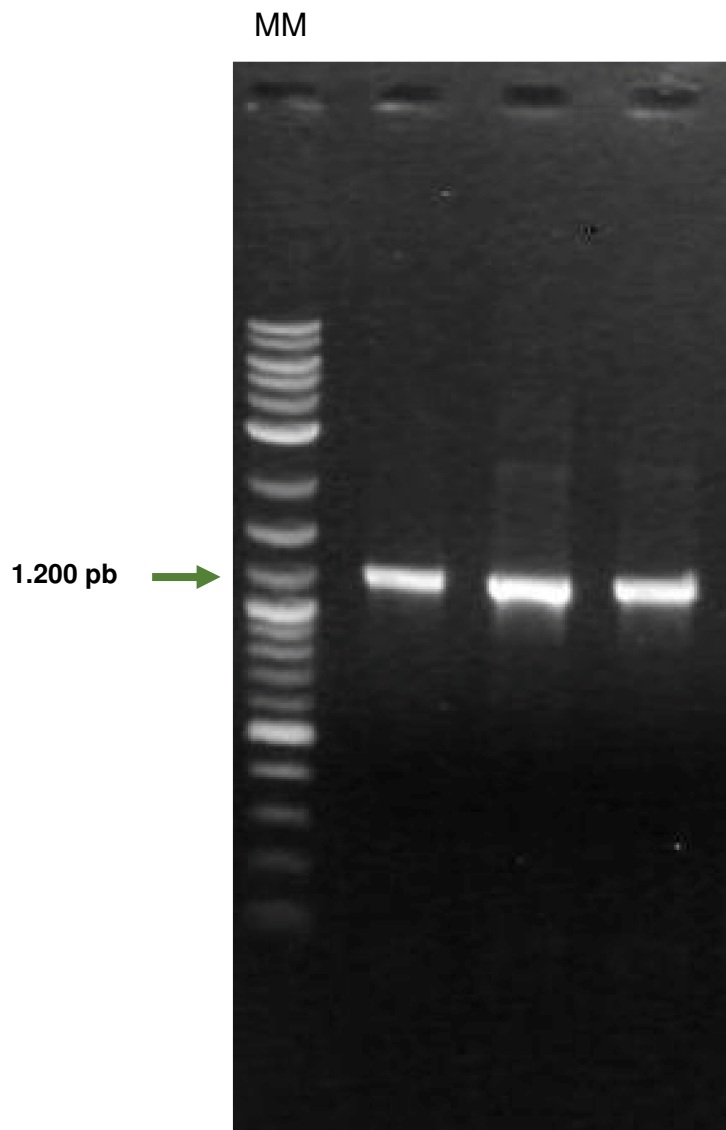


Fig. 6

Material Suplementar



Fig. S1 – Sequência da proteína quimérica HSAKIN(His)_{6x}. A região do *T7 gene 10 leader peptide* e da *His x6 Tail* são oriundos do vetor de expressão pET23a-d(+) (Novagen).

Saccharomyces_cerevisiae ---MADYDSAKYWSKQGARRGLQKTRYCQICQROCKDANGFQSHNKSPSHLRKISQVT
Ostreococcus_lucimarinus -----MSAKSLSNQLKSKGLQRLRWYCOLCEKQCRDENGFKCHCSPSAHVRLAVFA
Plasmodium_berghei --MPRAEPTGPKWLANMKAKGLQKLRWYQMCQKQCRDENGFKCHRLSETHORQMIFC
Phytophthora_nicotianae -----
Schizosaccharomyces_pombe --MGRAEAGTPKAI SNALKSKGLQRLRWYCSACQKQMRDENGFKCHTQSEGHIRQMNVIA
Arabidopsis_thaliana --MGKNDFLTPKAIANRIKAKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQOVVFG
Malus_domestica --MGKNDFLTPKAIANRIKAKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQOVVFG
Cucumis_melo --MGKNEFLTPKAIANRIKAKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQOVVFG
Toxoplasma_gondii --MPKAEFGTTKWSNKAAGLQKLRWYQMCQKQCRDENGFKCHRMSEGHORMQVFI
Brugia_malayi --MGKYEKGTPKAIANRCKSKGLQKLRWFQMCCKKQCRDQNGFKCHLSETHORQLLFA
Ciona_intestinalis MGKEKPGFLTAKAIGNRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLIAG
Xenopus_tropicalis --MGKADFLTPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Xenopus_laevis --MGKADFLSPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Danio_rerio --MGKADFLTPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Poeciliopsis_prolifica --MGKADFLSPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Homo_sapiens --MGKSDFLTPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Macaca_mulatta --MGKSDFLTPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Mus_musculus --MGKSDFLTPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Alligator_mississippiensis --MGKSDFLSPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Gallus_gallus --MGKSDFLSPKAIANRIKSKGLQKLRWYQMCQKQCRDENGFKCHCMSESHORQLLFA
Aptenodytes_forsteri -----MPKGMHVWSTNGFKCHCMSESHORQLLFA
Drosophila_melanogaster --MGRAEVTGPKYLANRCKSKGLQKLRWFQMCCKKQCRDENGFKCHCMSESHORQLLFA
Culex_quinquefasciatus MGKKGAEVGTGPKYLANRCKSKGLQKLRWFQMCCKKQCRDENGFKCHCMSESHORQLLFA

Saccharomyces_cerevisiae AE---DARRYNIQFEKGLQLLQKHGKWDANKVYNEYVQDRDHVHMNATMHRSLTQF
Ostreococcus_lucimarinus QNOKKVVDFGSSRFKTAFLSELNHRNSRVAANKVYNELIANKTHVHMNATRWTSLETF
Plasmodium_berghei QDANKFMDEYSMFKEKFMRLMKTCKYCRARILANTVYTNMISDKTHIHMNATVWVTLTDF
Phytophthora_nicotianae -----MFEEAFLENLRRRHGTRKMRATHVYNEYIQDKLHVHMNATQWTTLSGF
Schizosaccharomyces_pombe MNPGRKIQDFSNQFLRDFISLLRTAHGKIKIHFNFQYQYIIRDKNHVHMNATRWHTLSEF
Arabidopsis_thaliana QNPTRVVDGYSEEFQTFLLDMRRSHRFSRIAATVVYNEYINDRHHVHMNSTEWATLTFE
Malus_domestica ENSNRIVDGYSEEFQTFLLDMRRSHRFSRIAATVVYNEYINDRHHVHMNSTEWATLTFE
Cucumis_melo QNPHRIVEGYSEEFESFMEHIRKSHRFSRIAATVVYNEYINDRHHIHMNSTQWATLTFE
Toxoplasma_gondii QNPTKFMDDFSQFEREFMRLMKTCKYCRARILANTVYTNMISDKTHIHMNATVWVTLTDF
Brugia_malayi ENPNTYLKEYSEEFENFLKVLRSFTGTRKRVANEVYQYIIRDKMHTHMNSTKWHHTLTFE
Ciona_intestinalis ENPQGFHSFSSSFFKDFMQLLKTFRGTRKRVHNNIVYNEYISHKEHIMHNATKWVTLTAF
Xenopus_tropicalis ENPQQFMDYFSDERTEFLLELRRRFGTRKRVHNNIVYNEYISHREHVHMNATQWETLTFE
Xenopus_laevis ENPQQFMDYFSDERTEFLLELRRRFGTRKRVHNNIVYNEYISHREHVHMNATQWETLTFE
Danio_rerio ENPNTYLSYFSDERTEFLLELRRRFGTRKRVHNNIVYNEYISHREHVHMNATQWETLTFE
Poeciliopsis_prolifica ENPNTKFMDFSDERTEFLLELRRRFGTRKRVHNNIVYNEYISHREHVHMNSTQWETLTFE
Homo_sapiens ENPQQFMDYFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Macaca_mulatta ENPQQFMDYFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Mus_musculus ENPQQFMDYFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Alligator_mississippiensis ENPQQFMDYFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Gallus_gallus ENPQQFLNYFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Aptenodytes_forsteri ENPQQFMDYFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Drosophila_melanogaster DNPQKFLHSFSEEFNDFLELRRRFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE
Culex_quinquefasciatus GNAGRFIDFSSEFMTGYLQILRRQFGTRKRVHNNIVYNEYISHREHIMHNATQWETLTFE

* : : . . * : : : * * * : * : :

Saccharomyces_cerevisiae VRYLGRAGKVDVMDIDDTSENVEGPLLIRIHPSS--LSSPEDGMLRSQEEQEVIAAE
Ostreococcus_lucimarinus VHHLVABGVCEIDETDK-----GIFITYLHAESKIFQ-GAVARKRTHEDDGSARKQEK
Plasmodium_berghei VLYLGTGKCKIEQTER-----GWYLEYIDREKI--EREKEFNKKKIEYSYEMKEK
Phytophthora_nicotianae VQYLGRGKCVDETEK-----GWYIYIDRDPKAIARQKELEKKKKAELDHEERNKL
Schizosaccharomyces_pombe CKFLGRQGMCRVEENEK-----GFFISYIDKNPANILRNEANKRERQEKSDDEEQRLR
Arabidopsis_thaliana IKHLGKTGKCKVEETPK-----GWFITYIDRSETLFFKERLKNKRVKSDLAEEKQER
Malus_domestica VKHLGRTGKCKVDETPK-----GWFITYIDRSETLFFKERLKNKRLRADMADEEKQER
Cucumis_melo VKHLGRTGQCKVEETPK-----GWFITYIDRSETLFFKERLKNKRLRADMADEEKQER
Toxoplasma_gondii VQYLGSTNCKIELTPK-----GWYVEYIDHDER--RRLEQEKRRKVELSQEEAKAR
Brugia_malayi VIYLGSGKCHVDQTEK-----GWFIAWIDQEEEL-RKQE-ALHKVKAIHDDERLQQ
Ciona_intestinalis TKWLGREGHCKVDQTEK-----GWFIQYIDRDPVLARQKELDKQRERDDDERQAK
Xenopus_tropicalis TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRKQEQEKKKKQDLDEERTAK
Xenopus_laevis TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRKQEQEKKKKQDLDEERTSK
Danio_rerio TKWLGREGYCKVDEAPK-----GWYVQYIDRDPETIRRQEELEKKKKQDLDEERSAK
Poeciliopsis_prolifica TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Homo_sapiens TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Macaca_mulatta TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Mus_musculus TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Alligator_mississippiensis TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Gallus_gallus TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Aptenodytes_forsteri TKWLGREGYCKVDETPK-----GWYIYIDRDPETIRRQEELEKKKKQDLDEERTAK
Drosophila_melanogaster VKWLGRTGQVIADETEK-----GWFVYIDRSPAMERQAKADRKEKMEKDDDEERMAD
Culex_quinquefasciatus VKFLGRNGHCVADETEK-----GWYISYIDRDPETLAQEQEKMAKKQKMDKDAERLAE

* . : * : : .

Saccharomyces_cerevisiae
Ostreococcus_lucimarinus
Plasmodium_berghei
Phytophthora_nicotianae
Schizosaccharomyces_pombe
Arabidopsis_thaliana
Malus_domestica
Cucumis_melo
Toxoplasma_gondii
Brugia_malayi
Ciona_intestinalis
Xenopus_tropicalis
Xenopus_laevis
Danio_rerio
Poeciliopsis_prolifera
Homo_sapiens
Macaca_mulatta
Mus_musculus
Alligator_mississippiensis
Gallus_gallus
Aptenodytes_forsteri
Drosophila_melanogaster
Culex_quinquefasciatus

LLKRQLNRAKQTEKVYQPEM-----KSEISGDSTLKRQVTFHNGRNVNKKK----
LIERQVKRGAVPAKPEQHE-----PTGL-----KLQQ-----
KINQVIDEAKKKGFQI QSE-----YTGI EKKNDKII--ISSVKATSTNTSI---
FIKRQLQIASGKEGA--DTQ-----QOPTKLQRGENGKII--KMSLGLKRESAKKETV
LLDEQIKRAYESAONNEDNKDGSREQPVLHEIDLKSKGNPI--QLNLSSSSDSSHSAQNE
ETQRQIERAAEKLNGGGGEGETSG-----NDEVVDDGDDEK-----KDEDL--R
EIKKQIERVCQSMFVNGSDQG-----SEAE-----ANQEV--K
EIKRQIERAEQLAPLASASAE-----LEAE-----QTREL--K
RLEELVEEAKKRGY---RDA-E---YTPLLRRNGDEKI--VLKGLARPSEKSN--E
LLEEQAERAREKRQSSDMSKN-H---PTELRRAGDSDSKV--TFSLTT--KKLDE--E
AIEEMVKRGQ--DTA-KDSDSG-K---FTELVRND--EKI--AINLQM--NSEPKA--A
FIEEQVKKGL--EVT-EQN-TP-T---YTELSRQNEEEKV--AFNLNK--GASTSG--T
FIEDQVKKGL--EGV-EQN-TP-T---FTELSRQNEEEKV--AFNLNK--GAGTSG--T
FIEEQVRRGR--EGR-EPTTEP-V---FTELQRKSEDEKV--AFNLNLI--AAAGP---
FIEEQVRRGM--DGK-EAEAQP-V---YTELKRSEEEKV--AFNLST--PVASSS--K
FIEEQVRRGL--EGK-EQE-VP-T---FTELSRENDEEKV--TFNLNK--GACSSS--G
FIEEQVRRGL--EGK-EQE-AP-T---FTELSRENDEEKV--TFNLNK--GACSSS--G
FIEEQVRRGR--EGR-EQE-TP-V---FTELSRENEEEKV--TFNLNK--GAGGSA--G
FIEEQVRRGL--EGK-EQE-RP-V---YTELNRDNDDEEKV--AFNLNK--GASTSV--A
FIEEQVRRGL--EGK-ELE-EP-V---YTELNRENEEEKV--AFNLNK--GASTSV--A
FIEEQVRRGL--EGK-ELE-KP-V---YTELNRENEEEKV--TFNLNK--GASTSI--A
FIEEQIKNAKADGE-EDEGEK-K---FTELKREE-NEPL--KLDLRL--EKKFQ---
FIEEQVKRGHKEE---EPSTS-A---YSELKRENEEPI--KIALSL--GSSSK---

..

Saccharomyces_cerevisiae
Ostreococcus_lucimarinus
Plasmodium_berghei
Phytophthora_nicotianae
Schizosaccharomyces_pombe
Arabidopsis_thaliana
Malus_domestica
Cucumis_melo
Toxoplasma_gondii
Brugia_malayi
Ciona_intestinalis
Xenopus_tropicalis
Xenopus_laevis
Danio_rerio
Poeciliopsis_prolifera
Homo_sapiens
Macaca_mulatta
Mus_musculus
Alligator_mississippiensis
Gallus_gallus
Aptenodytes_forsteri
Drosophila_melanogaster
Culex_quinquefasciatus

-----KKVPPRKDG-----IKFR-----
-----TGFNMD-----LTRKEKLSQM-----
-----NHNNDVPK-----SNIFLDILKQNKIKSTN-----D--Q---
AFTAKASSVFGQEGE-----
FFQTRNTPTFSSSSSSQTSLKHKPKNVFAELNKSrk-----KN--NKDSL--
L-KSGVKVGFALGGGVK--QVATGKE-----RG-----
L-EPGVKIGFALGSSKL--GPKD-----KG-----
L-ESGVKLGfALGTTSK--LKEK-----KG-----
M-SDG-HSGERPANALKLEAELKRKRFAEMAKSSASSETGVDRSRRGSEDRPSLRQ
L-TRPCTSVFATS--YSIFHTKIKRETKL-----ELEFLDKSK-----KHQ
T-----SSQLAPR--V-----S-----
A-TTK-SSSLVSN--A-----LKSAAV-----
T-A---SSSHASN--A-----LKSAAV-----
---SKPSTTLGPS--A-----LKT-----
S-RY-----
A-TSSKSSTLGP--A-----LKTIGS-----
A-TSSKSSTLGP--A-----LKTIGS-----
A-TTSKSSSLGP--A-----LKLGS-----
A-SSK-ASVLGP--A-----LKMVG-----
A-SSKTSSILGP--A-----LKMV-----
A-SSKTSSVLPN--A-----LKMVD-----
-----PDTVLGKS--A-----LAKRPA-----
-----M--DIKPI--K-----LEKRPL-----

Saccharomyces_cerevisiae
Ostreococcus_lucimarinus
Plasmodium_berghei
Phytophthora_nicotianae
Schizosaccharomyces_pombe
Arabidopsis_thaliana
Malus_domestica
Cucumis_melo
Toxoplasma_gondii
Brugia_malayi
Ciona_intestinalis
Xenopus_tropicalis
Xenopus_laevis
Danio_rerio
Poeciliopsis_prolifera
Homo_sapiens
Macaca_mulatta
Mus_musculus
Alligator_mississippiensis
Gallus_gallus
Aptenodytes_forsteri
Drosophila_melanogaster
Culex_quinquefasciatus

-----LKAESI-----KKEEPLSKDKPQNKRPMSLELLIMENEKKKLKN
-----SEEA-----
-----ESSKL-----LFGDEENDKVERG
-----GSAKL-----VDFEVEDETNERK
-----ESSRV-----VDFEDETETARK
DEdGVHSIKARSRLSAAERLRLIEIQEQARARGSPSTKRTNADELGRLLVGDREDKNESG
-----QWDLNLSAATLKREIKSY-----
-----LSS-----L-----
-----MGS-----V-----
-----MGS-----V-----
-----AS-----V-----
-----SAS-----V-----
-----SAS-----V-----
-----AAS-----G-----
-----TGS-----V-----
-----GA-----V-----
-----GA-----V-----
-----PEA-----EEK-----
-----EVA-----DE-----

Saccharomyces cerevisiae
Ostreococcus lucimarinus
Plasmodium berghei
Phytophthora nicotianae
Schizosaccharomyces pombe
Arabidopsis thaliana
Malus domestica
Cucumis melo
Toxoplasma gondii
Brugia malayi
Ciona intestinalis
Xenopus tropicalis
Xenopus laevis
Danio rerio
Poeciliopsis prolifica
Homo sapiens
Macaca mulatta
Mus musculus
Alligator mississippiensis
Gallus gallus
Aptenodytes forsteri
Drosophila melanogaster
Culex quinquefasciatus

```
-----  
-----EQSHLVGTIMRIY-----  
LL-----PNLNK-----NIN--ANNSKTEHHS--ITKHMNEN-----HEKRRK  
-----ESEDKRSNIRHVAVDIIEENERRKREKSRRLR  
-----DQGQNVKRIIRSAVEDIIAQET-----MR  
EK----RKRSG-----DSGRSEKEIRSADELIMKEE-----EK  
V----KKA-----KDGSGGDANISLAEELMREE-----EK  
KP----EGTS-----KNRIGTGGQLSSLEELMKEE-----EM  
HFRCSGGRQSDTHSVGLRGGEIEELTDRGSRGDCESHKTELHESDDDER-----KP  
-----RKEHDDSTVGSNKSRRREDL-----KSRNVLDEIMEME-----ER  
-----GKRKQPS-----SSGFKPS-----KRKKSALDEIIAAE-----EG  
-----RKDTSQV-----Q-----PKE-----KKKKSALDEIMEME-----EK  
-----RKDTSQG-----Q-----AKE-----KKKKSALDEIMEME-----EQ  
-----RKESLS-----G-E-FKD-----KKKKSALDEIIEME-----EQ  
-----RKESSQS-----STQSKEK-----KKKKSALDEIMEIE-----EE  
-----RKESAQS-----SAQSKEK-----KKKKSALDEIMEIE-----EE  
-----RKESSQS-----SAQP--A-----KKKKSALDEIMELE-----EE  
-----RKETGQS-----SGQSKEK-----KKKKSALDEIMEFE-----EE  
-----RKESAHS-----SGQSKEK-----KKKKSALDEIMELE-----EE  
-----RKESAHS-----SGQSKEK-----KKKKSALDEIMELE-----EE  
-----FKKPKS-----VAGD-----SQTRSVLDEIIOQE-----ES  
-----KKEKKH-----KPGD-----FKKPSALDEIIEKE-----ER
```

Saccharomyces cerevisiae
Ostreococcus lucimarinus
Plasmodium berghei
Phytophthora nicotianae
Schizosaccharomyces pombe
Arabidopsis thaliana
Malus domestica
Cucumis melo
Toxoplasma gondii
Brugia malayi
Ciona intestinalis
Xenopus tropicalis
Xenopus laevis
Danio rerio
Poeciliopsis prolifica
Homo sapiens
Macaca mulatta
Mus musculus
Alligator mississippiensis
Gallus gallus
Aptenodytes forsteri
Drosophila melanogaster
Culex quinquefasciatus

```
-----  
-----CAPGEQSKQKYIVTNVYK---GMATLEGIGPFNSAK  
KDKNSEDYDIWITKNIIVKIVDKN---HKYKSKGAIISISSTEKNKCEIKIKNTNKYTL  
EESKRQRENWVTGIVIKVMNKKVGDGAYYCKGKGVVDVEE---KFCATVELLDSGDVLR  
EKRR-NIK---L-----  
KKERMNRKDYWLFEGIVKVMKALAEKGYKQKGVVKKVID---NYVGEIKMLDSKHVLR  
KKERINRKDYWJCEGIVKVMKALADKGYKQKGVVVKVID---KYVGEIEMLESKHVLR  
KKEKLNKRDYWLCDGIVKIMSKDLAEKGYKQKGVVVKVID---KYVGEIEMLDGKHVLR  
PAHKYRNKDWLKEGCIIVKVMHAELGDGQYKQKGVVLRVEE---RYAAYVKMLDSGDILR  
RKEKRNRRDNWLHEGIVVKIITKKFGN-DFYKAKGVVQQLVD---DFTAILDVND--SLK  
RKKQLHTEHWLAKNIIVKIIITKRLGD-KYFKKAVVKEVQD---KFTAIVKMLDSGDQVK  
K-KKTERTDYWLQPEIIVKIVTKKLGE-KYYKKAVVKEVID---RYTAVVKLVDSDGDKLK  
K-KKTERTDYWLQPEIIVKIVTKKLGE-KYYKKAVVKEVID---RYTAVVKLVDSDGDKLK  
K-KKSVRSDNWIQENIIVKIVTKKLGE-KYYKKAVVKEVID---RYTAVVKLVDSDGDKLK  
-----  
K-KRTARTDYWLQPEIIVKIIITKKLGE-KYHKKAVVKEVID---KYTAVVKMIDSGDKLK  
K-KRTARTDYWLQPEIIVKIIITKKLGE-KYHKKAVVKEVID---KYTAVVKMIDSGDKLK  
K-KRTARTDAWLQPGIIVKIIITKKLGE-KYHKKAVVKEVID---RYTAVVKMIDSGDRLK  
K-KRTARTDYWLQPEIIVKIVTKKLGE-KYHKKAVVKEVID---KYTAVVKMIDSGDKLK  
K-KRTSRTDYWLQPEIIVKIVTKKLGE-KYHKKAVVKEVID---KYTAVVKMIDSGDKLK  
K-KRTSRTDYWLQPEIIVKIVTKKLGE-KYHKKAVVKEVID---RYTAVVKMIDSGDKLK  
KKEKRNKRDYWLHKGIVVKFISKSMGE-KFFKQKAVVLDVID---RYQKIKFLETGEKLK  
KKEKRNKRDYWLAEIIVKLVAKSLGE-KYYKQKAVVTEVID---KYRKRIRLLETGDKLK
```

Saccharomyces cerevisiae
Ostreococcus lucimarinus
Plasmodium berghei
Phytophthora nicotianae
Schizosaccharomyces pombe
Arabidopsis thaliana
Malus domestica
Cucumis melo
Toxoplasma gondii
Brugia malayi
Ciona intestinalis
Xenopus tropicalis
Xenopus laevis
Danio rerio
Poeciliopsis prolifica
Homo sapiens
Macaca mulatta
Mus musculus
Alligator mississippiensis
Gallus gallus
Aptenodytes forsteri
Drosophila melanogaster
Culex quinquefasciatus

```
-----  
VSTDHLRTVIPAIGASIKVIRGENAGCIGSVISVDASGGSLRVLLEGS-----GD  
AYQSQTQTVIPQIGRMVILKGNKGLKGIKIKISEEDYAVVSVLHKNSDE-----II  
LDQDDLETVIPKGRKVIIVNGLGRGCTAKLLDISVDDFCARIRIDSGSYR-----GE  
-----  
VDQKELETVLPQIGMVKIVNGAYRGSNARLLGVDTEKFCQVQIEKGVYD-----GR  
VDQAELETVIPNLGCLVKIVNGAYRGSNARLLAVDTDNFCQVQIEKGVYD-----GR  
VDQEELETVIPQIGLVRIVNGAYRGSNARLLSVDTDKFCQVQIEKGVYD-----GR  
LDQKMLETVIPNIGRVRIVLGEWRDHEGKMEGIDMDAFCVHVVDSDGPVFKVRSHLAP I  
VSQENVETVIPAVGREMLIVNGAYRGTAVLQAIIEKKFAVVLRLDES LAK-----GR  
FDQTHLETVIPSIGKTVLIVNGGYRDCVATLEKIDTSKFSATLIKITEGPLN-----GR  
LDQTHLETVIPAPGKRVFVNLGGYKGFHILDCINEKFSATIIIDS GPLK-----GR  
LDQSHLETIPAPGKRVLVNLGGYKGFHILDGIIIEKRF SATIIDS GPLK-----GR  
LDQSHVETVIPAPGKRVLILNGQYRGTAEILEGINEHKFSATLTLDSGRMK-----GK  
-----  
LDQTHLETVIPAPGKRIIVLNLGGYRGNEGTLESINEKTF SATIV IETGPLK-----GR  
LDQTHLETVIPAPGKRIIVLNLGGYRGNEGTLESINEKTF SATIV IETGPLK-----GR  
LDQTHLETVIPAPGKRVLVNLGGYRGNEGTLESINEKAFS SATIV IETGPLK-----GR  
LDQTHLETVIPAPGKRVLVNLGGYRGNEGTLESINEKFSATI IIDS GPLK-----GR  
LDQTHLETVIPAPGKVMVNLGGYRGNEGTLESINEKFSATI IIDS GPLK-----GR  
LDQTHLETVIPAPGKVMVNLGGYRGNEGTLESINEKFSATI IIDS GPLK-----GR  
VDQAHLETVIPALDKPVMVNGAYRGSALLRKLDERYSVSVELHGPLK-----GR  
VDQAHLETVIPSIGKQVVLNLGGYRGNATLKEIDTDNYCTTIEIASGPLK-----GR
```


<i>Saccharomyces_cerevisiae</i>	-----
<i>Ostreococcus_lucimarinus</i>	EMNTFSFEDVSKLDKKYSFNRDKM
<i>Plasmodium_berghei</i>	AEEHMIFFDDISKFQEK -----
<i>Phytophthora_nicotianae</i>	ILDRVEYEDICRLADE -----
<i>Schizosaccharomyces_pombe</i>	-----
<i>Arabidopsis_thaliana</i>	VIKSIEYEDICKLA -----
<i>Malus_domestica</i>	VLKAVEYEDICKLA -----
<i>Cucumis_melo</i>	VLKAVEYEDICKLAS -----
<i>Toxoplasma_gondii</i>	VVKNLPYEHVCKIA -----
<i>Brugia_malayi</i>	IL-RLPYEDACKIKQ -----
<i>Ciona_intestinalis</i>	IVEDVKYEDFSKKFDQS -----
<i>Xenopus_tropicalis</i>	KVEGIPYEDISKVA -----
<i>Xenopus_laevis</i>	KVEGIPYEDISKVA -----
<i>Danio_rerio</i>	TVEGIAVEDFSKLA -----
<i>Poeciliopsis_prolifica</i>	-----
<i>Homo_sapiens</i>	RVEGIQYEDISKLA -----
<i>Macaca_mulatta</i>	RVEGIQYEDFSKLA -----
<i>Mus_musculus</i>	RVEGIQYEDISKLA -----
<i>Alligator_mississippiensis</i>	RVDSIQYEDISKLA -----
<i>Gallus_gallus</i>	RVEGIQYEDISKLA -----
<i>Aptenodytes_forsteri</i>	RVEDIQYEDISKLA -----
<i>Drosophila_melanogaster</i>	IVDNVQYEDISKLHGA -----
<i>Culex_quinquefasciatus</i>	IVNNVQYEDISKLFG -----

Zinc Finger

Winget Helix

RecA Homologous

NLS

KOW

Fig. S2 – Alinhamento das sequências homólogas de _{HSA}KIN.



Ability of HMGB1 protein to bind to intrinsically bent and non-bent DNA sites in the *AMPD2* gene amplicon

K.J.R. Passos, A. Fiorini, F.R. Rosado, D.V.B. Freitas, Q.A. Lima Neto, J.R. Pattaro Junior, V.P. Gaspar and M.A. Fernandez

Departamento de Biologia Celular e Genética, Universidade Estadual de Maringá, Maringá, PR, Brasil

Corresponding author: M.A. Fernandez
E-mail: mafernandez@uem.br

Genet. Mol. Res. 15 (2): gmr.15027441
Received August 13, 2015
Accepted October 2, 2015
Published June 10, 2016
DOI <http://dx.doi.org/10.4238/gmr.15027441>

ABSTRACT. HMGB-like proteins are architectural chromatin factors, and their function is heavily dependent on their ability to interact with DNA (especially non-canonical DNA structures). HMGB1 is involved in many DNA processes, and dysregulation of HMGB protein expression has profound effects on cellular transcription, resulting in severe developmental defects as well as cancer. During DNA replication, elements that form the origin are still not well defined in metazoans. Sites with A (adenine) or T (thymine) repeats cause intrinsic curvatures in the DNA and are described to be involved in the replication machinery by providing binding sites to replication proteins. As a result, the DNA molecule shows intrinsically bent DNA sites, caused by periodic repeats of 2 or more As/Ts (dA/dT) as well as intrinsically non-bent DNA sites (INBDs), due to a succession of curvatures that cancel each other. In the present study, we mapped 11 INBDs present in the *AMPD2* gene that are related to each replication origin (*oriGNA13*, *oriC*, *oriB*, and *oriA*). Following characterization of INBDs, we tested the ability of HMGB1 to bind to the bent (*b1*, *b2*, *b4a*, *b4b*, *b5*, *b6*, *b7*, and *b8*) and non-bent DNA fragments (*nb7*, *nb11*, *nb1*, *nb2*, *nb4*, and *nb5*) via electrophoretic mobility shift assays. All fragments showed efficient binding to HMGB1. However, the non-bent DNA fragments *nb2*, *nb4*, and *nb5* showed slightly reduced binding efficiency.

Key words: HMGB1; Intrinsically bent DNA; *AMPD2* amplicon; Intrinsically non-bent DNA; DNA replication