



Universidade Federal do Rio de Janeiro  
Centro de Ciências Matemáticas e da Natureza  
Instituto de Química



Projeto Final de Curso  
IQWX01

Nathalia Lins da Silva

# **ESTUDO DA ESTABILIDADE E ESPECIFICIDADE DE INTERAÇÃO DO DOMÍNIO PWWP DA HISTONA METIL-TRANSFERASE NSD3 COM PEPTÍDEOS DERIVADOS DE HISTONAS**

Rio de Janeiro

2019

## **Agradecimentos**

Agradeço à Deus por ser a luz que ilumina o meu caminho.

À minha mãe, Lêda Lins, por ser meu maior exemplo de força e determinação, me ensinando sempre a acreditar e ser melhor todos os dias.

À minha irmã, Karoline Lins, pela parceria e companheirismo sempre, segurando as pontas na ChocoLins nas horas em que eu estava estudando e/ou fazendo experimento.

Aos amigos que fiz durante a graduação, por ter sofrido comigo durante toda essa trajetória dando sempre forças para seguir em frente e chegar até aqui.

Ao meu namorado, Leonardo Muniz, por ser paciente e compreensivo em todas as horas que eu não pude estar com ele por estar trabalhando ou estudando.

Aos amigos do LaBMol, por estarem sempre dispostos a ajudar no que era preciso e por compartilhar comigo, tornando mais leves, as longas horas na bancada. Em especial, à Germana Rona, por me passar todos os conhecimentos das técnicas de laboratório com amor.

Ao professor Anderson Pinheiro, por ter acreditado em mim naquela aula de Bioquímica e por todo o suporte durante essa jornada no laboratório.

**RESUMO****PROJETO DE CURSO**

**TÍTULO:** ESTUDO DA ESTABILIDADE E ESPECIFICIDADE DE INTERAÇÃO DO DOMÍNIO PWWP DA HISTONA METIL-TRANSFERASE NSD3 COM PEPTÍDEOS DERIVADOS DE HISTONAS

**ALUNA:** Nathalia Lins Da Silva

**ORIENTADOR:** Anderson De Sá Pinheiro, DBq – Instituto de Química – UFRJ

**DATA DA DEFESA:** 27/06/2019

**PALAVRAS-CHAVE:** NSD3; PWWP; histona; fluorescência; estrutura.

NSD3 (Nuclear Receptor Binding SET-domain containing protein 3) é um membro da família NSD de histona metil-transferases humanas. NSD3 possui uma arquitetura modular formada por um domínio catalítico SET, dois domínios PWWP (Pro-Trp Trp-Pro) e cinco domínios dedos de zinco PHD. A superexpressão de NSD3S é suficiente para a aquisição de fenótipos tumorais por células epiteliais de mama, sugerindo que o domínio PWWP desempenha papel crucial na patogênese do câncer de mama. O presente trabalho teve como objetivo a determinação da especificidade de interação de NSD3-PWWP com peptídeos derivados de histonas, bem como caracterização da estabilidade deste domínio frente a diferentes condições experimentais, a fim de facilitar a sua determinação estrutural. A especificidade de interação de NSD3-PWWP com 9 peptídeos (13-mer) derivados de caudas amino-terminais de histonas foi investigada através de espectroscopia de fluorescência. O espectro de fluorescência intrínseca de NSD3-PWWP apresentou máximo de emissão em 344 nm, sugerindo que os resíduos de triptofano são parcialmente expostos ao solvente. A adição de concentrações crescentes de peptídeos promoveu a supressão da fluorescência de modo dose-dependente. As isotermas de ligação derivadas dos dados de fluorescência exibiram duas inflexões distintas, sugerindo dois sítios de ligação com afinidades diferentes. A interação de NSD3-PWWP com os diferentes peptídeos se deu com afinidades semelhantes (~3-10 micromolar), sugerindo que NSD3-PWWP é capaz de reconhecer metil-lisinas com alta afinidade, porém com pouca seletividade. A fim de se determinar a estabilidade termodinâmica, NSD3-PWWP e NSD3-PWWP:H3K36me<sub>3</sub>, foram submetidos ao ensaio de fluorimetria diferencial de varredura. Curvas de desnaturação térmica foram construídas em diferentes condições de pH e força iônica, a fim de determinar a condição ótima de estabilidade e solubilidade da proteína. NSD3-PWWP livre se mostrou mais estável (T<sub>m</sub> ~63°C) em condições de pH básico (>8,0) e alta força iônica (NaCl 250 mM). Por sua vez, o complexo NSD3-PWWP:H3K36me<sub>3</sub> apresentou T<sub>m</sub> ~73°C, sugerindo que a interação com o peptídeo estabiliza a proteína.

Autorizamos a divulgação no Repositório Institucional da UFRJ-Pantheon.

Aluno: \_\_\_\_\_ Orientador: \_\_\_\_\_  
Assinatura Assinatura

## Sumário

Agradecimentos.....	i
Resumo.....	ii
Índice de Abreviaturas e siglas.....	v
Índice de Tabelas.....	vii
Índice de Figuras.....	viii
1. Introdução.....	1
1.1. Cromatina e histonas.....	1
1.2. Metilação de lisinas em histonas.....	4
1.3. Família NSD.....	5
1.4. NSD3.....	6
1.5. Domínio PWWP.....	8
2. Objetivos .....	10
2.1. Objetivo geral.....	10
2.2. Objetivos específicos.....	10
3. Material e métodos.....	11
3.1. Material.....	11
3.1.1 Meio LB .....	11
3.1.2 Meio mínimo .....	11
3.1.3 Cepas bacterianas .....	11
3.1.4 Peptídeos derivados de histonas .....	11
3.2. Métodos.....	12
3.2.1. Clonagem do domínio NSD3-PWWP.....	12
3.2.2. Transformação bacteriana.....	13

3.2.3. Expressão de NSD3-PWWP.....	14
3.2.4. Eletroforese em gel de poliacrilamida sob condições desnaturantes (SDS PAGE) .....	14
3.2.5. Dosagem de proteínas .....	15
3.2.6. Purificação de NSD3-PWWP .....	15
3.2.6.1. Lise bacteriana .....	15
3.2.6.2. Primeira cromatografia de afinidade a níquel .....	15
3.2.6.3. Diálise e clivagem com TEV .....	16
3.2.6.4. Segunda cromatografia de afinidade a níquel .....	16
3.2.6.5. Cromatografia de exclusão molecular .....	17
3.2.7. Aquisição de dados de RMN .....	17
3.2.8. Modelagem molecular .....	17
3.2.9. Espectroscopia de fluorescência .....	18
3.2.10. Fluorimetria diferencial de varredura .....	19
4. Resultados e discussão.....	20
4.1. Purificação de NSD3-PWWP.....	20
4.2. Caracterização estrutural preliminar de NSD3-PWWP.....	24
4.2.1. $^1\text{H}$ 1D RMN .....	24
4.2.2. [ $^1\text{H}$ , $^{15}\text{N}$ ] HSQC .....	25
4.3. Afinidade e especificidade de interação NSD3-PWWP com peptídeos de histonas.....	26
4.4. Estabilidade térmica de NSD3-PWWP.....	33
4.5. Estabilidade térmica do complexo NSD3-PWWP:H3K36me3 .....	37
5. Conclusões .....	43
6. Referências .....	44

## Índice de abreviaturas e siglas

<b>1D</b>	Unidimensional
<b>2D</b>	Bidimensional
<b>ADA</b>	<i>N-(2-Acetamido)iminodiacetic acid</i>
<b>ADP</b>	Adenosina difosfato
<b>Bis-Tris</b>	<i>2,2-Bis(hydroxymethyl)-2,2',2''-nitrilotriethanol</i>
<b>Bis-Tris propano</b>	<i>1,3-Bis[tris(hydroxymethyl)methylamino]propane</i>
<b>BSA</b>	Albumina bovina sérica
<b>CAPS</b>	<i>3-(Cyclohexylamino)-1-propanesulfonic acid</i>
<b>CHES</b>	<i>N-Cyclohexyl-2-aminoethanesulfonic acid</i>
<b>C-terminal</b>	Carboxi terminal
<b>D.O.600</b>	Densidade ótica a 600 nm
<b>DNA</b>	Ácido desoxirribonucléico
<b>DTT</b>	Ditiotreitol
<b>EPSPS</b>	<i>4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid</i>
<b>GST</b>	Glutathione-S-transferase
<b>H1</b>	Histona um
<b>H2A</b>	Histona dois A
<b>H2B</b>	Histona dois B
<b>H3</b>	Histona três
<b>H4</b>	Histona quatro
<b>HEPES</b>	<i>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</i>
<b>His6</b>	Cauda de seis histidinas
<b>HSQC</b>	Heteronuclear Single Quantum Coherence
<b>IPTG</b>	Isopropil- $\beta$ -D-tiogalactopiranosídeo
<b>I-TASSER</b>	<i>Iterative Threading ASSEMBly Refinement</i>
<b>KDMs</b>	Demetilases de lisina, do inglês <i>Lysine Histone Demethylases</i>
<b>KMTs</b>	Lisinas metil-transferases, do inglês <i>Lysine Histone Methyl-Transferases</i>
<b>LB</b>	Lúria-Bertani, meio rico para cultivo de micro-organismos
<b>MES</b>	<i>2-(N-Morpholino)ethanesulfonic acid</i>
<b>MOPS</b>	<i>3-(N-Morpholino) Propanesulfonic Acid</i>

<b>NSD</b>	<i>Nuclear receptor SET Domain-containing protein</i>
<b>N-terminal</b>	Amino-terminal
<b>PHD</b>	<i>Plant HomeoDomain</i>
<b>PI</b>	Ponto isoelétrico
<b>PIPPS</b>	
<b>PMSF</b>	Fluoreto de fenilmetanosulfonila
<b>PSA</b>	Persulfato de amônio
<b>PTM</b>	Modificações pós-traducional, do inglês <i>Post-Translational Modifications</i>
<b>PWWP</b>	prolina-triptofano-triptofano-prolina
<b>RMN</b>	Ressonância Magnética Nuclear
<b>RT-PCR</b>	Reação em cadeia da polimerase em tempo real
<b>SDS</b>	Dodecil sulfato de sódio
<b>SET</b>	<i>Suppressor of variegation 3-9, Enhancer of zeste and Thritorax</i>
<b>shRNA</b>	<i>Short Hairpin RNA</i>
<b>TEMED</b>	N,N,N',N'- tetrametiletilenodiamina
<b>TEV</b>	<i>Tobacco Etch Virus</i>
<b>Tm</b>	Temperatura de fusão, do inglês <i>melting temperature</i>
<b>Tris</b>	Tris (hidroximetil) aminometano
<b>WHSC1L1</b>	Wolf-Hirschhorn Syndrome Candidate 1-Like 1

## Índice de tabelas

<b>Tabela 1:</b> Peptídeos derivados de histonas utilizados neste trabalho.....	12
<b>Tabela 2:</b> Tampões utilizados nos experimentos de DSF.....	19
<b>Tabela 3:</b> Constantes de dissociação aparente obtidas para a interação de NSD3-PWWP com peptídeos de histonas.....	31
<b>Tabela 4:</b> Afinidades de ligação de domínios PWWP humanos com peptídeos metilados das histonas H3 e H4.....	32
<b>Tabela 5:</b> Temperatura de fusão obtida para NSD3-PWWP em 72 condições experimentais contendo diferentes pHs e concentração de sal.....	37
<b>Tabela 6:</b> Temperatura de fusão obtida para NSD3-PWWP:H3K36me3 na presença de 50 mM NaCl em diferentes pHs.....	42



## Índice de figuras

<b>Figura 1:</b> Estrutura cristalográfica do nucleossomo.....	1
<b>Figura 2:</b> Modificações pós-traducionais em histonas.....	2
<b>Figura 3:</b> Proteínas escritoras, apagadoras ou leitoras de PTMs em histonas.....	3
<b>Figura 4:</b> O processo de metilação e desmetilação de lisinas em proteínas.....	4
<b>Figura 5:</b> Arquitetura modular das KMTs da família NSD.....	6
<b>Figura 6:</b> Isoformas de <i>splicing</i> de NSD3.....	7
<b>Figura 7:</b> Estrutura tridimensional de LEDGF-PWWP.....	9
<b>Figura 8:</b> Representação esquemática da construção His <sub>6</sub> GST-NSD3-PWWP.....	13
<b>Figura 9:</b> Purificação de His <sub>6</sub> GST-NSD3-PWWP por cromatografia de afinidade a níquel.....	21
<b>Figura 10:</b> Clivagem de His <sub>6</sub> GST-NSD3-PWWP com a protease TEV.....	22
<b>Figura 11:</b> Purificação de NSD3-PWWP através de uma segunda etapa de cromatografia de afinidade a níquel.....	23
<b>Figura 12:</b> Purificação de NSD3-PWWP por cromatografia de exclusão molecular..	24
<b>Figura 13:</b> Espectro 1D <sup>1</sup> H RMN de NSD3-PWWP.....	25
<b>Figura 14:</b> Espectro 2D [ <sup>1</sup> H, <sup>15</sup> N] HSQC de NSD3-PWWP.....	26
<b>Figura 15:</b> Titulação de NSD3-PWWP com H3K36me3 investigada por espectroscopia de fluorescência.....	28
<b>Figura 16:</b> Isotermas de ligação de NSD3-PWWP com peptídeos de histonas.....	30
<b>Figura 17:</b> Sobreposição do modelo estrutural de NSD3-PWWP com estruturas tridimensionais de complexos PWWP:peptídeos de histonas previamente descritos.	33
<b>Figura 18:</b> Curva de desnaturação térmica de NSD3-PWWP nos tampões A1 (glicina, pH 3,0), A2 (Ácido cítrico, pH 3,2), A3 (PIPPS, pH 3,7), A4 (Ácido cítrico, pH 4,0) e A5 (Acetato de sódio, pH 4,5).....	35

<b>Figura 19:</b> Análise da estabilidade térmica de NSD3-PWWP por fluorimetria diferencial de varredura.....	36
<b>Figura 20:</b> Purificação e desnaturação térmica do complexo NSD3-PWWP:H3K36me3.....	38
<b>Figura 21:</b> Análise da estabilidade térmica do complexo NSD3-PWWP:H3K36me3 na presença de 50 mM de NaCl por fluorimetria diferencial de varredura.....	40

## 1 Introdução

### 1.1 Cromatina e histonas

O material genético de células eucarióticas encontra-se organizado em um polímero dinâmico denominado cromatina. O grau de compactação da fibra de cromatina varia de acordo com o desenvolvimento, ciclo celular e em resposta a mudanças ambientais. A regulação da estrutura da cromatina, ou seja, o remodelamento da cromatina desempenha um papel chave no controle de diversos processos nucleares, tais como: transcrição, replicação, reparo de DNA e apoptose (DE LA CRUZ et al., 2015).

O nucleossomo, a unidade fundamental da cromatina, é constituído por 147 pares de bases de DNA envolvido em um octâmero de histonas. As histonas (H3, H4, H2A, H2B e H1) são proteínas básicas, pequenas, constituídas por um domínio globular, ao qual o filamento de DNA encontra-se associado, e uma cauda N-terminal flexível e carregada. O octâmero central do nucleossomo é formado por um tetrâmero de H3-H4 e dois dímeros de H2A e H2B (BANNISTER & KOUZARIDES, 2011; IIZUKA & SMITH, 2003) (Figura 1). A histona H1, por sua vez, atua como estabilizadora da organização estrutural do nucleossomo, interagindo com o DNA remanescente, não ligado ao complexo (DE LA CRUZ et al., 2015; TSAO et al., 2009)

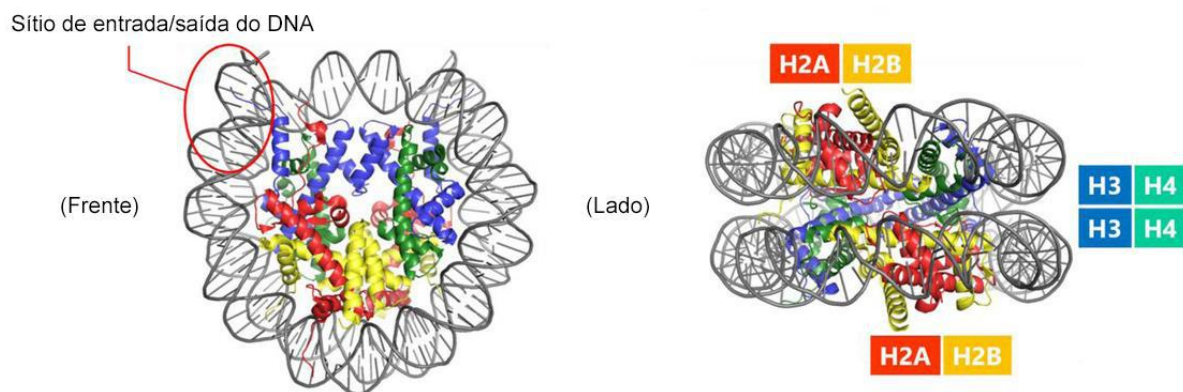


Figura 1: Estrutura cristalográfica do nucleossomo. (PDB ID: 3AFA) (Adaptado de SUEOKA et al., 2018)

As histonas H2A, H2B, H3 e H4 possuem um domínio globular formado por três  $\alpha$ -hélices e duas alças responsáveis pela organização da partícula central do nucleossomo. Os domínios globulares das histonas apresentam uma estrutura rígida devido a formação de motivos '*hand-shake*', referentes à dimerização de dois monômeros envolvendo o entrelaçamento das hélices das histonas com associação

de maneira cabeça-cauda (PLACEK & GLOSS, 2002), e contatos eletrostáticos com o DNA. No entanto, a região N-terminal das histonas é intrinsecamente desordenada e flexível e se projeta para fora do nucleossomo (SUEOKA et al., 2018).

As histonas participam diretamente da manutenção dinâmica da estrutura da cromatina (MARGUERON et al., 2005). A cauda N-terminal das histonas está sujeita a uma série de modificações químicas, tais como: fosforilação (Ph), acetilação (Ac), metilação (Me), propionilação (Pr), butirilação (Bu), crotonilação (Cr), 2 hidroxilisobutirilação (Hib), malonilação (Ma), succinilação (Succ), glutarilação (Gl), formilação (Fo), hidroxilação (OH), ubiquitinação (Ub), SUMOilação (Sumo), O-GlcNAcilação (Og), ADP-ribosilação (Ar) e citrulinação (Cit). (HUANG et al., 2015) (Figura 2). Até hoje, mais de 20 modificações pós-traducionais (PTMs—*post-translational modifications*) diferentes foram descobertas em diversos sítios em histonas (SUEOKA et al., 2018).

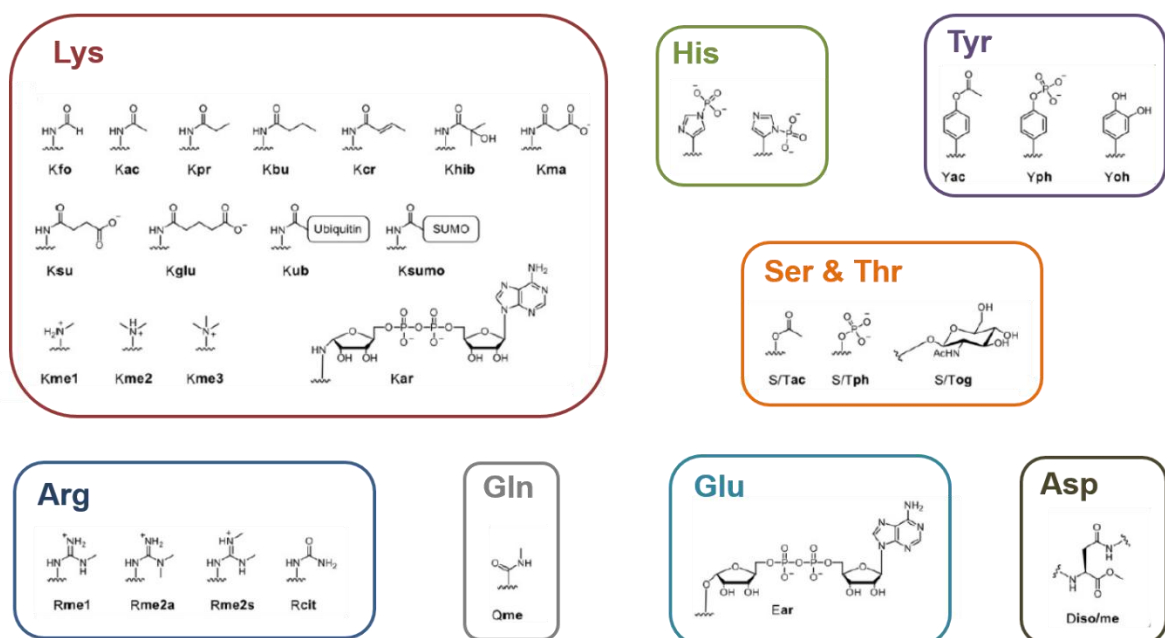


Figura 2: Modificações pós-traducionais em histonas (Adaptado de SUEOKA et al., 2018).

Desde 1964, quando Vincent Allfrey descreveu pela primeira vez os processos de metilação e acetilação em histonas, diversos estudos foram conduzidos na tentativa de correlacionar essas PTMs com a regulação da transcrição (ANDREOLI & DEL RIO, 2014). Avanços significativos nesta área decorreram da descoberta de proteínas capazes de incorporar, reconhecer e remover PTMs em histonas. Podemos entender essas proteínas como “escritoras” (*writers*), “apagadoras” (*erasers*) ou

“leitoras” (*readers*) da informação contida no padrão de PTMs em histonas (BROWNELL et. al. 1996) (Figura 3). A descoberta concomitante de uma histona acetil-transferase (p55/GCN5), capaz de regular o evento da transcrição, e uma histona deacetilase (HDAC1/Rpd3), capaz de remover a marca de acetilação, forneceu a primeira evidência da correlação entre PTMs em histonas e regulação transcricional (BROWNELL et. al, 1996; TAUNTON et al., 1996). Um marco importante nesta área foi a proposição do código de histonas por Strahl e Allis em 2000 (STRAHL & ALLIS, 2000), onde o padrão de PTMs sofridas pelas histonas constituiria um segundo código genético, capaz de determinar características hereditárias.

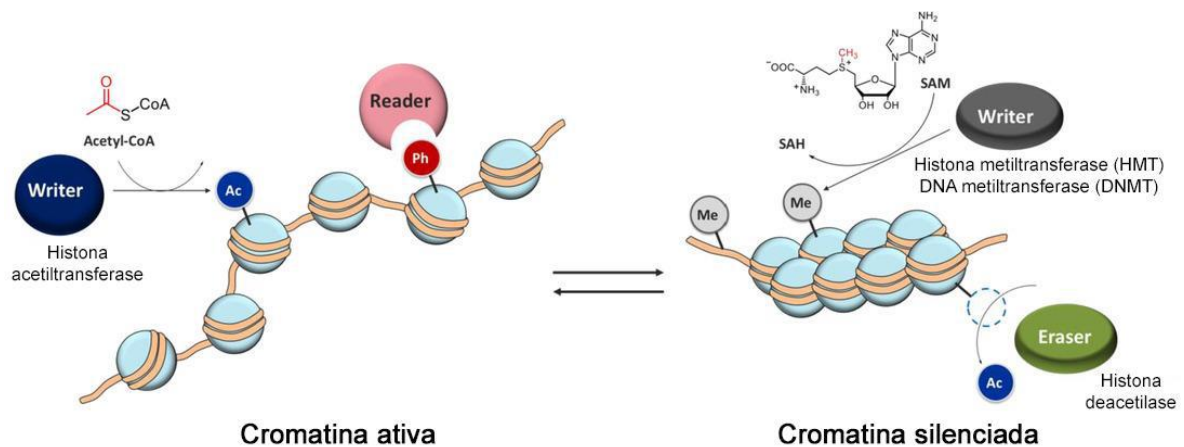


Figura 3: Proteínas escritoras, apagadoras ou leitoras de PTMs em histonas (SUEOKA et al., 2018).

As PTMs de histonas podem atuar individualmente ou em combinação umas com as outras, a fim de gerar um determinado resultado biológico (STRAHL & ALLIS, 2000). Estas podem afetar a função dos cromossomos através da alteração da carga eletrostática das histonas, o que acarreta em mudanças estruturais nas próprias histonas e na interação com o DNA genômico (IIZUKA & SMITH, 2003). Além disso, as PTMs podem criar, estabilizar, romper ou ocluir domínios da cromatina responsáveis pela interação com proteínas regulatórias, como fatores de transcrição, proteínas envolvidas na condensação da cromatina e reparo do DNA. Sendo assim, as PTMs de histonas desempenham um papel central na regulação da expressão gênica e constituem a principal forma de controle transcricional epigenético (BANNISTER & KOUZARIDES, 2011; IIZUKA & SMITH, 2003; SANTOS-ROSA & CALDAS, 2005).

## 1.2 Metilação de lisinas em histonas

As PTMs mais estudadas em histonas são metilação e acetilação. As histonas podem ser metiladas em resíduos de lisina e arginina através da ação de histona metiltransferases (BANNISTER & KOUZARIDES, 2011).

Algumas PTMs estão diretamente associadas ao estado de ativação da transcrição gênica. A acetilação, por exemplo, está associada a genes transcricionalmente ativos, enquanto a metilação está associada tanto a genes ativos quanto a genes silenciosos (IIZUKA & SMITH, 2003; MELLOR, 2006). A posição das PTMs também está relacionada a alterações da transcrição gênica. A metilação das lisinas K4, K36 e K79 na histona H3 está frequentemente associada a ativação da transcrição. Por outro lado, a metilação das lisinas K9 e K27 na histona H3 e da lisina K20 na histona H4 estão associadas a supressão da transcrição (MELLOR, 2006).

O grau de metilação de lisinas depende tanto da atividade das enzimas que adicionam o grupo metila, as lisinas metiltransferases (KMTs - *lysine methyltransferases*), quanto daquelas que removem o grupo metila, as demetilases específicas para lisinas (KDMs - *lysine demethylases*). Mais de 50 KMTs e 20 KDMs humanas já foram relatadas na literatura científica. O grupo  $\epsilon$ -amino da lisina pode aceitar até três grupos metila, resultando em mono-, di- ou trimetil-lisina, com os vários estados de metilação exercendo funções distintas (Figura 4) (HAN et al., 2019).

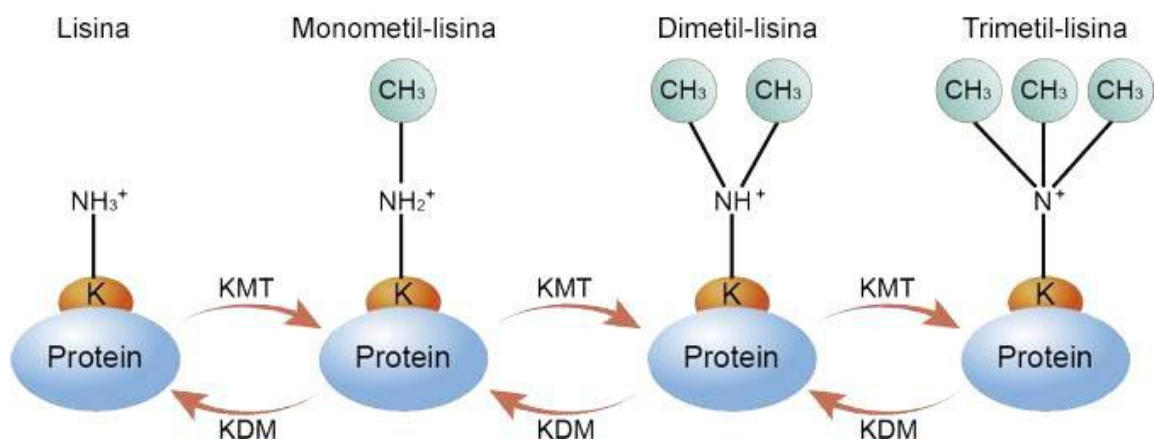


Figura 4: O processo de metilação e desmetilação de lisinas em proteínas (HAN et al., 2019).

Modificações epigenéticas estão envolvidas na patogênese de diversas doenças humanas, incluindo o câncer. Tais modificações levam a aparição de funções gênicas anômalas, como alterações nos níveis de expressão de genes e estabilidade do

genoma, características marcantes do processo de tumorização. Ao contrário de alterações genéticas, modificações epigenéticas são reversíveis e, portanto, apresentam um elevado potencial terapêutico (EGGER et al 2004; LOPEZ et al., 2009). Estudos prévios mostram que mudanças nos níveis de metilação de histonas estão relacionadas ao aumento da recorrência de tumores e a baixa sobrevivência de pacientes (GREER & SHI, 2012). Diversas KMTs encontram-se superexpressas ou mutadas em vários tipos de câncer, sugerindo um papel central na patogênese dessa doença. Neste contexto, inibidores de KMTs têm emergido como estratégia efetiva na terapia contra o câncer (ALBERT & HELIN, 2010).

### 1.3 Família NSD

Dentre as KMTs, destaca-se a família das proteínas NSD (*Nuclear receptor SET Domain-containing protein*), composta pelos seguintes membros: NSD1/KMT3B, NSD2/WHSC1/MMSET e NSD3/WHSC1L1. Estas proteínas possuem um papel fundamental na manutenção da estrutura da cromatina através da metilação da lisina K36 na histona H3 (H3K36me), um conhecido marcador de ativação transcricional (WAGNER & CARPENTER, 2012). NSD1 e NSD2 foram primeiramente identificadas como genes associados a doenças do desenvolvimento, como as síndromes de Soto e Wolf-Hirschhorn, respectivamente. Frequentemente, estas proteínas encontram-se alteradas ou superexpressas em diversos tecidos tumorais. Além disso, a redução dos níveis de NSDs leva a supressão do crescimento e progressão de tumores (MORISHITA & DI LUCCIO, 2011; VOUGIOUKLAKIS et al., 2015). Apesar da importância desta família de proteínas para uma gama de patologias humanas, os mecanismos moleculares associados às NSDs ainda permanecem pouco entendidos.

Estruturalmente, as NSDs apresentam uma arquitetura modular composta por domínios funcionalmente distintos e comuns às três proteínas, são eles: um domínio SET (*Suppressor of variegation 3-9, Enhancer of zeste and Thritorax*), dois domínios PWWP (prolina-triptofano-triptofano-prolina) e de quatro a cinco domínios dedos de zinco PHD (*Plant HomeoDomain*) (Figura 5). O domínio SET é responsável pela atividade catalítica das NSDs, enquanto os domínios PWWP e dedos de zinco PHD estão envolvidos no reconhecimento da cromatina e interações proteína-proteína (VOUGIOUKLAKIS et al., 2015).

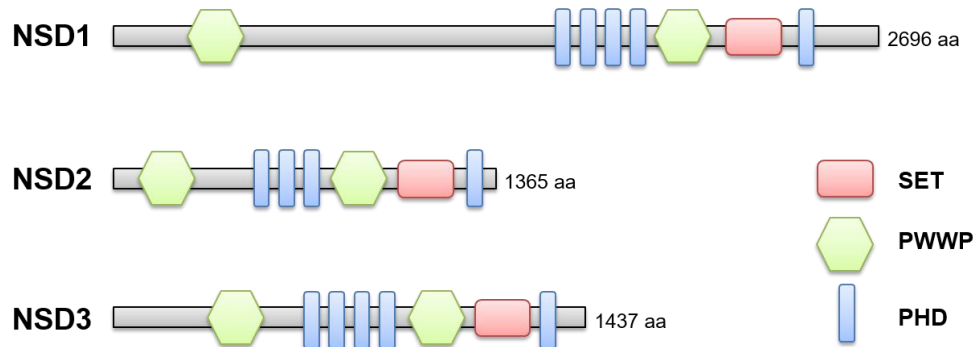


Figura 5: Arquitetura modular das KMTs da família NSD. PWWP: prolina-triptofano-triptofano-prolina; PHD: *Plant HomeoDomain*; SET: *Supressor of variegation 3-9, Enhancer of zeste and Thritorax*. (Adaptado de VOUGIOUKLAKIS et al., 2015).

#### 1.4 NSD3

O gene *NSD3*, o terceiro membro da família NSD, também conhecido como *WHSC1L1* (*Wolf-Hirschhorn Syndrome Candidate 1-Like 1*) está localizado na região cromossômica 8p11-12 (ANGRAND et al, 2001). Uma característica marcante das três proteínas NSD é que são altamente semelhantes dentro de uma região de cerca de 700 aminoácidos, que contém o domínio catalítico SET, um domínio PWWP, cinco domínios PHD e um domínio rico em Cys-His específico de NSD (C5HCH). No entanto, o módulo PHD5-C5HCH, quinto homeodomínio (PHD5) e o domínio rico em Cys-His adjacente (C5HCH), de NSD3 apresenta preferência por peptídeos da histona H3, tais como K4 não-metilado e K9 tri-metilado, o que é diferente dos outros dois membros (HAN et al., 2018). O módulo PHD5-C5HCH de NSD2 apresenta forte preferência por H3K9 não-metilada (H3K9me0), enquanto o PHD5-C5HCH de NSD1 parece não interagir com peptídeos derivados da histona H3 (HE et. al., 2013).

NSD3 foi primeiramente descrita como o produto de um oncogene frequentemente amplificado em diversas linhagens celulares e tumores primários de mama (ANGRAND et al, 2001). A região cromossômica onde está localizado o gene NSD3 encontra-se amplificada em 15-18% dos casos de câncer de mama e tem sido associada ao baixo prognóstico e a dificuldade de tratamento de tumores mamários (GELSI-BOYER et al., 2005; HAVERTY et al., 2008). Estudos de genômica e transcriptômica apontam para uma correlação entre o aumento do número de cópias gênicas e a superexpressão de NSD3 em diversos tipos de câncer, sobretudo o de mama. Liu et al. (2015) identificaram a NSD3 como a quarta KMT mais amplificada em 958 amostras de tumores mamários. Chen et al. (2014) descreveram a NSD3



como um oncogene líder putativo da região 8p11-12, capaz de dirigir a expressão de uma série de outros oncogenes presentes nesta região. A superexpressão de NSD3 promove a imortalização de células epiteliais de mama e induz a aquisição de fenótipos tumorais. Além disso, o *knockdown* de NSD3 leva a diminuição do crescimento e da sobrevivência de linhagens tumorais, sugerindo o potencial oncogênico desta enzima no câncer de mama (YANG et al., 2010; KANG et al., 2013). Por fim, NSD3 também tem sido implicada na translocação cromossomal t(8;11)(p11.2;p15) que resulta na proteína de fusão NSD3-NUP98 relacionada à leucemia mielóide aguda e à síndrome mielodisplásica (ROSATI et al., 2002; TAKETANI et al., 2009).

NSD3 possui três isoformas resultantes de *splicing* alternativo do éxon 10: NSD3L, NSD3S e WHISTLE (Figura 6). A isoforma longa (NSD3L) codifica a proteína íntegra formada por 1.437 aminoácidos, contendo dois domínios PWWP, cinco PHDs, um domínio rico em cisteína associado ao SET (SAC) e um domínio SET (HAN et al., 2018). Comparada com NSD3L, a isoforma curta de NSD3 codifica uma proteína de 645 aminoácidos que não possui o domínio catalítico SET, mas reserva o domínio PWWP N-terminal (HAN et al., 2018). Já a isoforma WHISTLE é formada por 506 aminoácidos C-terminais, contendo o domínio catalítico SET e apenas o segundo domínio PWWP. (VOUGIOUKLAKIS et al., 2015)

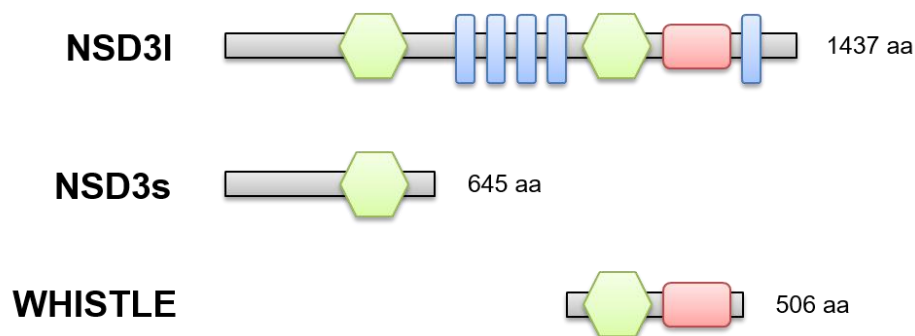


Figura 6: Isoformas de *splicing* de NSD3 (Adaptado de VOUGIOUKLAKIS et al., 2015).

Em determinadas linhagens de câncer de mama, NSD3S apresenta níveis de expressão maiores do que NSD3L. A superexpressão de NSD3S é suficiente para a transformação de células epiteliais de mama MCF10A em células com características tumorais, ressaltando a importância do domínio PWWP N-terminal para a atividade transformante de NSD3 (YANG et al., 2010).

## 1.5 Domínio PWWP

O domínio PWWP é encontrado exclusivamente em eucariotos, variando de eucariotos inferiores, como protozoários e leveduras até humanos. O genoma humano codifica mais de 20 proteínas contendo PWWP, que estão localizadas no núcleo e desempenham um papel importante na divisão, crescimento e diferenciação celular. Estes domínios estão implicados em várias funções da cromatina, incluindo modificação e reparo do DNA, além de regulação da transcrição (RONA et al., 2016). O domínio PWWP possui cerca de 100 a 150 resíduos de aminoácidos e contém o motivo conservado Pro-Trp-Trp-Pro. A estrutura deste domínio consiste em um barril- $\beta$  antiparalelo composto por cinco fitas- $\beta$  e um feixe helicoidal que possui de uma a seis  $\alpha$ -hélices (WEAVER et al., 2018). O domínio PWWP se liga tanto a sequências de DNA quanto a resíduos de lisinas metiladas em histonas. Sendo assim, este domínio funciona como um motivo de reconhecimento de metil-lisinas no nucleossomo, desempenhando um papel central na regulação epigenética (STEC et al., 2000; WU et al., 2011).

Diversas estruturas tridimensionais de domínios PWWP e seus complexos já foram determinadas e reportadas (RONA et al., 2016). Os dados estruturais sugerem que o modo de interação destes domínios com o DNA é amplamente conservado. Da mesma maneira, com exceção do domínio PWWP da proteína HDGF, todos os domínios PWWP estudados até o momento mostraram ligação às caudas de histonas metiladas (WEAVER et al., 2018). A ligação do domínio PWWP com o DNA se dá através de uma superfície básica, contendo resíduos de Arg e Lys, adjacente à região de ligação a metil-lisinas, incluindo a região do arco  $\beta$ 1- $\beta$ 2 (a qual inclui o motivo PWWP) e porções da região helicoidal imediatamente após a fita- $\beta$  final (Figura 7A) (WEAVER et al., 2018). Para a maioria dos domínios PWWP, a ligação ao DNA ocorre de maneira inespecífica através de interações eletrostáticas com os grupos fosfato do esqueleto principal do DNA (RONA et al., 2016). Quanto a interação com metil-lisinas, PWWP interage através de uma cavidade hidrofóbica composta por três resíduos aromáticos localizados no arco  $\beta$ 1- $\beta$ 2, motivo Pro-Trp-Trp-Pro e na fita  $\beta$ 3. As cadeias laterais destes aminoácidos estão perpendicularmente orientadas umas às outras, formando uma gaiola aromática que acomoda o grupo trimetil (Figura 7B) (VEZZOLI et al. 2010; WU et al., 2011; QIU et al., 2012; EIDAHN et al., 2013; VAN NULAND et al., 2013; WEN et al., 2014; WANG et al., 2014).

Como os sítios de ligação a DNA e histona não se sobrepõem extensivamente (Figura 7), sugere-se que a ligação simultânea a ambos é possível e que os domínios PWWP possam associar-se ao nucleossomo de forma multivalente. De fato, o domínio PWWP de Pdp1 se liga simultaneamente a H4K20me3 e ao DNA genômico (WEAVER et al., 2018), sem que a afinidade de ligação ao DNA seja alterada pela interação com H4K20me3, sugerindo eventos de ligação independentes (WEAVER et al., 2018).

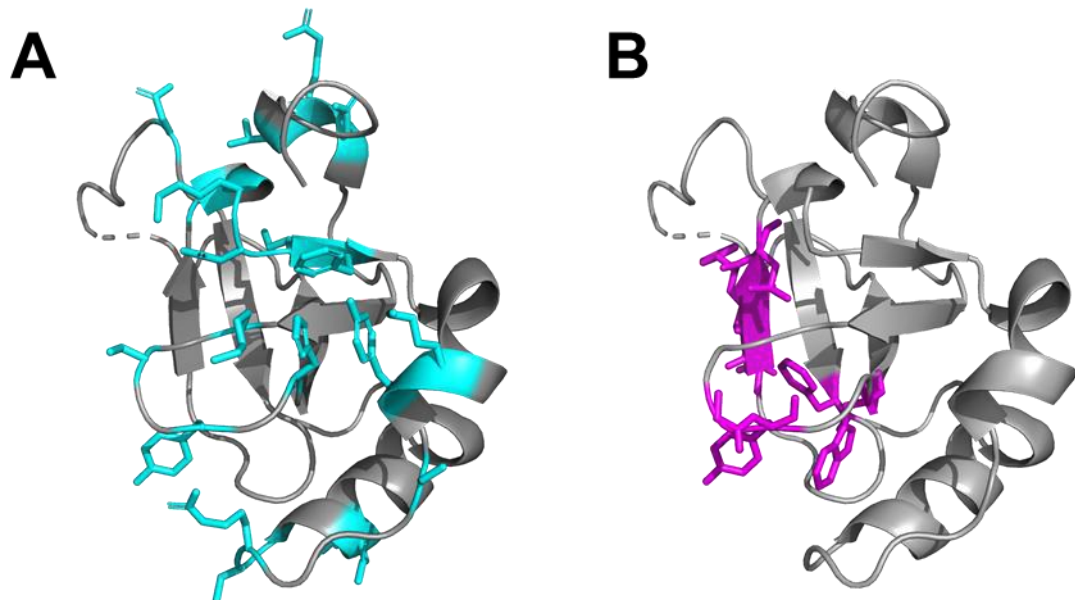


Figura 7: Estrutura tridimensional de LEDGF-PWWP (PDBID 4FU6). (A) Os resíduos que participam da ligação ao ácido nucleico estão mostrados em azul. (B) Os resíduos que participam da ligação com histonas estão mostrados em violeta. (Adaptado de WEAVER et al., 2018).

Apesar da gama de informações bioquímicas sobre domínios PWWP, pouco se conhece sobre o mecanismo pelo qual o domínio PWWP da histona metil-transferase NSD3 reconhece o nucleossomo.

Neste contexto, o presente trabalho teve como objetivo a determinação da especificidade de interação de NSD3-PWWP com metil-lisinas em histonas. Para isso, foi utilizado um grupo de 9 peptídeos (13-mer) derivados de caudas de histonas (H3 e H4) contendo diferentes lisinas metiladas e diferentes graus de metilação. Ainda, foi objetivo deste trabalho investigar a estabilidade térmica de NSD3-PWWP, tanto livre quanto complexado ao peptídeo H3K36me3, frente a diferentes condições experimentais, como pH e força iônica, a fim de facilitar a sua determinação estrutural futura.

## **2 Objetivos**

### **2.1 Objetivo geral**

O presente trabalho teve como objetivo a determinação da afinidade de ligação do domínio PWWP da histona metil-transferase NSD3 a peptídeos derivados de histonas, a fim de compreender a especificidade desta interação, bem como a determinação da estabilidade deste domínio em diferentes condições experimentais, a fim de facilitar a resolução da sua estrutura tridimensional.

### **2.2 Objetivos específicos**

- Expressão do domínio NSD3-PWWP;
- Purificação de NSD3-PWWP;
- Caracterização estrutural preliminar de NSD3-PWWP;
- Determinação da especificidade de interação entre NSD3-PWWP e peptídeos derivados de histonas;
- Determinação da estabilidade térmica de NSD3-PWWP em diferentes condições de pH e força iônica;
- Determinação da estabilidade térmica do complexo NSD3-PWWP:H3K36me3 em diferentes condições de pH e força iônica;

### 3 Material e métodos

#### 3.1 Material

##### 3.1.1 Meio LB

As cepas bacterianas foram cultivadas em meio Luria-Bertani (LB) composto de triptona 10 g/L, extrato de levedura 5 g/L e NaCl 5 g/L (Sigma). O meio LB sólido foi acrescido de ágar 15 g/L. Os meios seletivos foram adicionados de canamicina 100 µg/mL. O estoque do antibiótico foi preparado em água milliQ e esterilizado por filtração em membrana de 0,22 µm em uma concentração de 100 mg/mL.

##### 3.1.2 Meio mínimo

O meio mínimo de cultura (M9) foi constituído de Na<sub>2</sub>HPO<sub>4</sub> 48 mM, KH<sub>2</sub>PO<sub>4</sub> 22 mM, NaCl 8 mM, MgSO<sub>4</sub> 2 mM, CaCl<sub>2</sub> 0,1 mM, Tiamina (vitamina B1) 0,04 mM, glicose 3g e <sup>15</sup>NH<sub>4</sub>Cl 1g.

##### 3.1.3 Cepas bacterianas

Duas cepas bacterianas foram utilizadas, uma para armazenamento e propagação dos plasmídeos e outra para a expressão da proteína. Ambas as cepas foram cedidas pela professora Bianca Cruz Neves do Departamento de Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro. A cepa *Escherichia coli* DH5-α [genótipo: F-φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 end A1 hsdR17 (rk-, mk+) phoA supE44 λ- thi-1 gyrA96 relA1] foi utilizada para estoque plasmidial, enquanto a cepa *E. coli* BL21(DE3) [genótipo: F-ompT hsdSB(rB-, mB-) gal dcm (DE3)] foi utilizada para expressão das proteínas recombinantes.

##### 3.1.4 Peptídeos derivados de histonas

Nove peptídeos derivados das histonas H3 e H4 (Tabela 1), contendo as formas mais estudadas de metilação em lisinas, foram obtidos comercialmente da empresa GenScript (EUA) e diluídos em água MiliQ autoclavada. Uma alíquota de cada peptídeo foi tirada e diluída para a concentração final de 2 mM em tampão [20 mM fosfato de sódio (pH7,4), 100 mM NaCl].

Peptídeo	Sequência
<b>H3K4me3</b>	MART <b>Kme3</b> QTARKSTG-Amida
<b>H3K9me3</b>	Ac-TKQTARK <b>Kme3</b> STGGKA-Amida
<b>H3K27me3</b>	Ac-ATKAARK <b>Kme3</b> SAPATG-Amida
<b>H3K36</b>	Ac-PATGGV <b>KK</b> PHRYR-Amida
<b>H3K36me</b>	Ac-PATGGV <b>Kme</b> KPHRYR-Amida
<b>H3K36me2</b>	Ac-PATGGV <b>Kme2</b> KPHRYR-Amida
<b>H3K36me3</b>	Ac-PATGGV <b>Kme3</b> KPHRYR-Amida
<b>H3K79me3</b>	Ac-EIAQDF <b>Kme3</b> TDLRFQ-Amida
<b>H4K20me3</b>	Ac-GAKRHR <b>Kme3</b> VLRDNI-Amida

Tabela 1: Peptídeos derivados de histonas utilizados neste trabalho.

## 3.2 Métodos

### 3.2.1 Clonagem do domínio NSD3-PWWP

A sequência de DNA que codifica o domínio NSD3-PWWP (aminoácidos 259-402) foi clonada no vetor de expressão bacteriano pETM30-GST pela empresa Genscript (EUA). O vetor fusiona a proteína de interesse a uma cauda de seis histidinas (His<sub>6</sub>) seguida da proteína glutationa-S-transferase (GST) e de um sítio de clivagem para a protease TEV (*Tobacco Etch Virus*) (ENLYFQG).

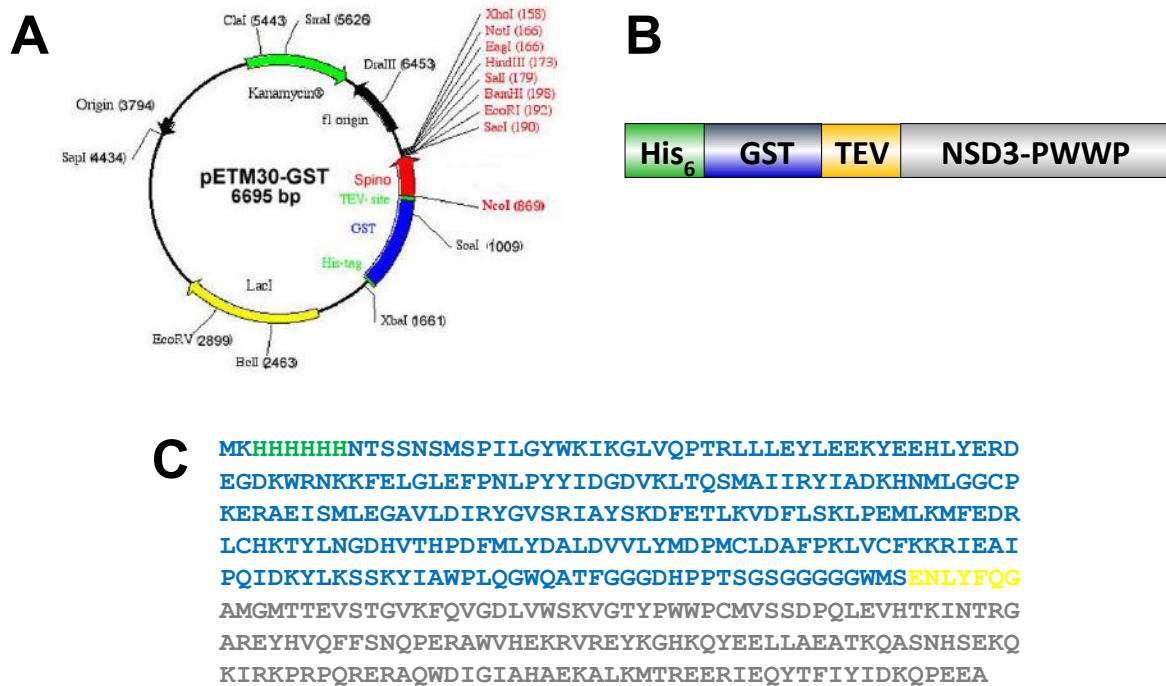


Figura 8: Representação esquemática da construção His<sub>6</sub>GST-NSD3-PWWP. (A) Mapa de restrição do vetor pETM30-GST utilizado na clonagem do domínio PWWP de NSD3. (B) Representação esquemática da proteína de fusão expressa em *E. coli*. O domínio PWWP foi fusionado a uma cauda N-terminal contendo seis histidinas, seguida de uma cauda GST e um sítio de clivagem para a protease TEV. (C) Sequência primária da proteína de fusão expressa em *E. coli* (~45 kDa). Em verde, a cauda de histidinas N- terminal. Em azul, a cauda de GST. Em amarelo, o sítio de clivagem para TEV. Após clivagem da cauda N-terminal, a construção NSD3-PWWP<sub>259-402</sub> passou a apresentar a massa molecular de 17,5 kDa.

### 3.2.2 Transformação bacteriana

Células de *Escherichia coli* BL21(DE3) competentes foram preparadas de acordo com o método descrito por HANAHAN *et al* (1983) e mantidas em ultrafreezer -80°C. Para a transformação, as alíquotas de células competentes foram descongeladas em banho de gelo. Um volume de 2 µL de DNA plasmidial foi adicionado a cada alíquota de 200 µL de célula. A mistura foi incubada em gelo por 30 min, seguido de choque térmico a 42°C por 45 s e novamente mantida em banho de gelo por 2 min. Em seguida, foram adicionados 800 µL de meio LB líquido às amostras, que foram então incubadas a 37°C por 1 h sob agitação a 200 rpm. Após esse período, as células foram centrifugadas a 8000 Xg por 10 min a 4°C (Sigma 2-16KL). Um volume de 970 µL do sobrenadante foi descartado e o precipitado formado

ressuspendido nos 30  $\mu$ L de meio restantes. As bactérias transformadas foram plaqueadas em meio LB sólido seletivo contendo 100  $\mu$ g/mL de canamicina e incubadas a 37°C por 16 h.

### **3.2.3 Expressão de NSD3-PWWP**

Uma única colônia de bactéria, resultante da transformação de *E. coli* BL21(DE3) com o plasmídeo recombinante que codifica a construção His<sub>6</sub>GST-NSD3-PWWP, foi inoculada em 5 mL de meio LB adicionado de 100  $\mu$ g/mL de canamicina e incubada a 37°C sob agitação a 200 rpm por 6 h. Após esse período, 250  $\mu$ L deste inóculo foram transferidos para 25 mL de meio LB adicionado de 100  $\mu$ g/mL de canamicina, e a cultura foi crescida a 37°C sob agitação a 200 rpm por 16 h. Após esse período, a cultura bacteriana foi transferida para 1 L de meio LB adicionado de 100  $\mu$ g/mL de canamicina e crescida a 37°C sob agitação a 200 rpm até que a densidade ótica a 600 nm (D.O.<sub>600</sub>) atingisse aproximadamente 0,6. Nesta D.O.<sub>600</sub>, a expressão de His<sub>6</sub>GST-NSD3-PWWP foi induzida com a adição de isopropil- $\beta$ -D-tiogalactopiranosídeo (IPTG) 1 mM às culturas (Sigma). Após a adição de IPTG, a cultura bacteriana cresceu por 20 h adicionais a 18°C. Ao final da expressão, as células foram coletadas por centrifugação a 8000 Xg por 30 min a 4°C (Sigma 2-16KL) e o precipitado armazenado em freezer -20 °C.

### **3.2.4 Eletroforese em gel de poli(acrilamida) sob condições desnaturantes (SDS PAGE)**

A técnica de SDS PAGE foi realizada conforme descrito por LAEMMLI (1970) utilizando o Mini-PROTEAN Tetra System 20 (BioRad). O gel de corrida foi constituído de 14,5% de acrilamida, 0,5% de bis-acrilamida, Tris-Glicina 390 mM (pH 8,8), dodecil sulfato de sódio (SDS) 0,1%, persulfato de amônio (PSA) 0,1% e TEMED 0,01%. O gel de concentração foi composto de 5% de acrilamida, Tris-Glicina 130 mM (pH 6,8), SDS 0,1%, PSA 0,1% e TEMED 0,01%. Foram aplicados 15  $\mu$ L de cada amostra adicionadas de 5  $\mu$ L de tampão de amostra [Tris-HCl 12,5 mM,  $\beta$ -mercaptoetanol 25 mM, SDS 0,5%, azul de bromofenol 0,25%, glicerol 2,5% (pH 6,8)] nos géis. Os géis correram a 30 mA por 1 h em tampão de corrida [Tris-HCl 25 mM, glicina 250 mM e SDS 0,1% (pH 8,3)]. Após a separação eletroforética, as proteínas foram visualizadas por coloração com Azul brilhante Comassie R (Vetec).



### **3.2.5 Dosagem de proteínas**

A concentração das amostras proteicas foi estimada através dos métodos colorimétricos realizados em placa de 96 poços utilizando os reagentes *Pierce 660nm Protein Assay* (Thermo Scientific) e *Bradford* (Bio-Rad). Para dosagem com Pierce, foram adicionados 150 µL do reagente a um volume de 10 µL de proteína em triplicata. Após 5 min de reação na ausência de luz, a absorvância das amostras foi medida a 660 nm utilizando um leitor de placas (SpectraMax M2). Para dosagem com Bradford, foram adicionados 200 µL do reagente, previamente diluído na proporção de 1:3 em água, a um volume de 10 µL de proteína em triplicata. Após 5 min de reação na ausência de luz, a absorvância das amostras foi medida a 590 nm utilizando um leitor de placas (Polaris-Celer). Os valores de concentração das amostras de proteína foram calculados a partir de uma curva padrão feita com albumina bovina sérica (BSA) na faixa de 0,025 a 2 mg/mL.

### **3.2.6 Purificação de NSD3-PWWP**

#### **3.2.6.1 Lise bacteriana**

O precipitado bacteriano proveniente da expressão de His<sub>6</sub>GST-NSD3-PWWP foi descongelado em banho de gelo e ressuspendido em tampão de lise [Tris-HCl 50 mM, NaCl 500 mM, imidazol 5 mM, β-mercaptoetanol 10 mM, triton-X 100 0,1%, fluoreto de fenilmetanosulfonila (PMSF) 250 µM (pH 8,0)]. As células foram lisadas em banho de gelo utilizando um sonicador (Sonics Vibracell) em modo pulsado (30 s de sonicação com intervalos de 50 s) e 20% de amplitude durante o tempo total de 50 min. O lisado foi centrifugado a 8000 Xg por 40 min a 4°C (Sigma 2-16KL), por fim o sobrenadante contendo a proteína de interesse foi filtrado em membrana de 0,22 µm (Millipore).

#### **3.2.6.2 Primeira cromatografia de afinidade à níquel**

Para a primeira etapa de purificação, a amostra do sobrenadante da etapa de lise celular foi submetida a uma cromatografia de afinidade a níquel. O perfil de eluição das proteínas presentes na amostra foi acompanhado por absorvância a 280 nm. Utilizou-se uma coluna HisTrap HP de 5 mL (GE Healthcare) acoplada ao sistema de purificação ÄktaPrime Plus (GE Healthcare), equilibrada em tampão A [Tris-HCl 50

mM, NaCl 500 mM, imidazol 5 mM,  $\beta$ -mercaptoetanol 10 mM (pH 8,0)]. A amostra proteica foi aplicada na coluna utilizando um fluxo de 1 mL/min e, em seguida, a coluna foi lavada com tampão A, tampão 2M NaCl [Tris-HCl 50 mM, NaCl 2 M, imidazol 5 mM,  $\beta$ -mercaptotanol 10 mM (pH 8,0)] e tampão A novamente em fluxo de 3 mL/min. Após a etapa de lavagem, a coluna foi submetida a um gradiente linear de 0 a 100% de tampão B [Tris-HCl 50 mM, NaCl 500 mM, imidazol 500 mM,  $\beta$ -mercapta 10 mM (pH 8,0)], ou seja, um gradiente linear de imidazol de 5 a 500 mM, utilizando um fluxo de 3 mL/min. O pico observado no cromatograma foi coletado em tubos cônicos. Após o término do gradiente, a coluna foi lavada com solução de imidazol 2 M para que as moléculas que persistissem ligadas à coluna fossem removidas. Amostras recolhidas ao longo de todo o processo de purificação foram submetidas à SDS PAGE.

### **3.2.6.3 Diálise e clivagem com TEV**

A amostra contendo His<sub>6</sub>GST-NSD3-PWWP proveniente da primeira etapa de cromatografia, identificada por SDS PAGE, teve primeiramente sua concentração determinada. A cauda N-terminal His<sub>6</sub>GST, presente no domínio PWWP de NSD3 recombinante, foi removida através de incubação com a protease His<sub>6</sub>-TEV N1a (S219V), que apresenta uma atividade sítio-específica e é capaz de reconhecer a sequência de resíduos ENLYFQG. A incubação foi feita em razão molar de 5:1 (NSD3-PWWP:TEV) a 25°C durante 16h. Simultaneamente à clivagem, a amostra foi dialisada contra tampão [Tris-HCl 50 mM, NaCl 500 mM,  $\beta$ -mercaptoetanol 10 mM (pH 8,0)] em membrana SnakeSkin (3000 Da; Thermo Scientific) para remoção do imidazol oriundo da eluição cromatográfica. A eficiência da clivagem foi avaliada por SDS PAGE.

### **3.2.6.4 Segunda cromatografia de afinidade à níquel**

Após a etapa de clivagem, foi realizada uma segunda etapa de cromatografia de afinidade a níquel, com o objetivo de separar a protease His<sub>6</sub>-TEV da proteína de interesse. Esta segunda cromatografia de afinidade foi realizada nas mesmas condições experimentais da primeira, descrita no item 3.2.5.2. Nesta etapa, a protease His<sub>6</sub>-TEV permaneceu ligada à resina de níquel enquanto a proteína NSD3-PWWP, desprovida da cauda de afinidade, foi diretamente eluída no lavado. Posteriormente, His<sub>6</sub>-TEV foi eluída através de lavagem da coluna com tampão B [Tris-HCl 50 mM,

NaCl 500 mM, imidazol 500 mM,  $\beta$ -mercaptoetanol 10 mM (pH 8,0)]. Amostras recolhidas ao longo de todo o processo de purificação foram submetidas à SDS PAGE.

### **3.2.6.5 Cromatografia de exclusão molecular**

Identificadas as frações referentes à NSD3-PWWP, estas foram reunidas e concentradas por ultrafiltração em sistema CENTRICON (3000 Da; Millipore) até o volume de 10 mL. A amostra concentrada foi submetida a uma terceira etapa de purificação por cromatografia de exclusão molecular utilizando uma coluna Superdex 75 (HiLoad 16/600; GE Healthcare) acoplada ao sistema ÄKTAPrime Plus (GE Healthcare). A coluna foi equilibrada em tampão de interesse utilizando um fluxo de 1 mL/min. O perfil de eluição de NSD3-PWWP foi avaliado através do traçado de absorvância a 280 nm. As frações referentes aos picos de eluição foram avaliadas por SDS PAGE.

### **3.2.7 Aquisição dos dados de RMN**

Os espectros de RMN foram adquiridos a 25°C em espectrômetros Bruker Avance III 900 e 800 MHz. Os experimentos foram adquiridos com uma amostra purificada de  $^{15}\text{N}$ -NSD3-PWWP a concentração de 300  $\mu\text{M}$  em tampão fosfato de sódio 20 mM (pH 6,5), NaCl 50 mM, DTT 5 mM,  $\text{D}_2\text{O}$  10%. Os espectrômetros estão localizados no Centro Nacional de Ressonância Magnética Nuclear da Universidade Federal do Rio de Janeiro (CNRMN-RJ). A aquisição e o processamento dos espectros de RMN foram feitos com os programas Top Spin 3.1 (Bruker Biospin) e NMRPipe.

### **3.2.8 Modelagem molecular**

O modelo de NSD3-PWWP foi gerado utilizando a plataforma I-TASSER (Iterative Threading ASSEmbly Refinement) e analisado com o programa Pymol. As sobreposições foram feitas utilizando os modelos do banco de dados PDB 2xy4, 5ciy e 6iis sobrepostos ao modelo de NSD3 utilizando o Pymol.

### 3.2.9 Espectroscopia de fluorescência

Os experimentos de espectroscopia de fluorescência foram realizados a 25°C no espectrofluorímetro Agilent Technologies (Cary Eclipse Fluorescence Spectrophotometer). Os espectros foram obtidos utilizando 2 µM de NSD3-PWWP em tampão 20 mM fosfato de sódio (pH 7,4), 100 mM NaCl. Foi utilizada um cubeta de quartzo de 500 µL de volume interno. Alíquotas de cada um dos nove peptídeos (Tabela 1), diluídos no mesmo tampão a uma concentração estoque de 2 mM, foram adicionadas à solução de proteína, de forma que a diluição da amostra não ultrapassasse 10% do volume total, resultando em concentrações finais de 0 – 276 µM na cubeta. Em cada concentração de peptídeo, foi coletado um espectro de fluorescência intrínseca da proteína. A excitação foi feita em 295 nm, a fim de se excitar somente os resíduos de triptofano, enquanto a emissão de fluorescência foi varrida de 310 a 450 nm. A largura das fendas de excitação e emissão foi de 5,0 nm e a tensão da fotomultiplicadora foi ajustada para 880 V. Os espectros foram coletados com 5 varreduras utilizando o programa Scan, fornecido pelo fabricante. Os valores de fluorescência foram obtidos a partir da área sob a curva dos espectros e corrigidos pela diluição em cada ponto da titulação. A supressão de fluorescência foi quantificada de acordo com a equação:

$$y = \frac{F_0 - F_i}{F_0 - F_s}$$

Onde  $F_0$  é a fluorescência inicial (NSD3-PWWP livre),  $F_i$  é a fluorescência em cada ponto e  $F_s$  a fluorescência saturada (no último ponto da titulação). As isotermas de ligação foram ajustadas utilizando o programa Sigmaplot 10.0 de acordo com a equação:

$$f = \left( a * \frac{(Kd1 + (x^n) + 2) - (((Kd1 + (x^n) + 2)^2) - 4 * (x^n) * 2)^{0,5}}{4} \right) + b * \left( \frac{x^n}{Kd2 + (x^n)} \right)$$

Onde  $Kd1$  e  $Kd2$  são as constantes de dissociação para dois sítios de ligação, e  $a$  e  $b$  são os máximos de  $y$  resultantes da saturação dos sítios I e II com o ligante (LEITGEB et al., 2009). Os experimentos foram realizados em triplicata e os resultados foram expressos como a média  $\pm$  erro padrão.

### 3.2.10 Fluorimetria diferencial de varredura (DSF)

A amostra purificada de NSD3-PWWP em tampão 10 mM fosfato de sódio (pH 7,0), 50 mM NaCl foi diluída a 5  $\mu$ M. 4  $\mu$ L do fluoróforo SYPRO Orange 5000X foram adicionados a 5 mL da amostra diluída. Foram transferidos 40  $\mu$ L de amostra para cada poço da placa de 96 poços e adicionados 10  $\mu$ L de cada solução tampão (Tabela 2). 0,5  $\mu$ L, 2,5  $\mu$ L e 5  $\mu$ L de solução aquosa de NaCl 5M foram adicionados de forma a se ter concentrações finais de 50 mM, 125 mM e 250 mM de NaCl em determinados poços. A placa de 96 poços foi levada ao aparelho de RT-PCR (Bio-Rad CFX96) e submetida a aquecimento de 25 a 95°C em passos de 1°C, fazendo leituras de fluorescência a cada 1 min, utilizando o canal FRET. Os dados foram coletados e processados utilizando o programa Bio-Rad CFX Manager 3.1.

Referência	Tampão (100 mM)	pH
A1	Glicina	3,0
A2	Ácido Cítrico	3,2
A3	PIPPS	3,7
A4	Ácido Cítrico	4,0
A5	Acetato de Sódio	4,5
A6	Fosfato de Sódio/Potássio	5,0
B1	Citrato de Sódio	5,5
B2	Fosfato de Sódio/Potássio	6,0
B3	Bis-Tris	6,0
B4	MES	6,2
B5	ADA	6,5
B6	Bis-Tris Propano	6,5
C1	Acetato de Amônio	7,0
C2	MOPS	7,0
C3	Fosfato de Sódio/Potássio	7,0
C4	HEPES	7,5
C5	Tris	7,5
C6	EPPS	8,0
D1	Imidazol	8,0
D2	Bicina	8,5
D3	Tris	8,5
D4	CHES	9,0
D5	CHES	9,5
D6	CAPS	10,0

Tabela 2: Tampões utilizados nos experimentos de DSF. Os tampões são pertencentes ao kit de solubilidade JBS (Jena Bioscience).

## 4 Resultados e discussão

### 4.1 Purificação de NSD3-PWWP

A proteína recombinante foi purificada por uma combinação de técnicas cromatográficas. A primeira etapa consistiu em uma cromatografia de afinidade a níquel, uma vez que a construção His<sub>6</sub>GST-NSD3-PWWP apresenta uma cauda N-terminal de His<sub>6</sub>. As histidinas possuem um anel imidazólico na cadeia lateral capaz de fazer ligações de coordenação com o metal imobilizado na resina, possibilitando a separação da proteína de interesse das demais proteínas presentes no extrato. Uma vez que a proteína é separada do extrato bruto, esta pode ser eluída ao se submeter o sistema a um gradiente de imidazol, que irá competir com as histidinas da proteína pela interação com o níquel da resina.

A construção His<sub>6</sub>GST-NSD3-PWWP foi eluída da coluna cromatográfica utilizando um gradiente contínuo de imidazol de 5 a 500 mM. A Figura 9A mostra o cromatograma de eluição, onde é observado um pico de eluição com máximo em aproximadamente 120 mM de imidazol. As frações referentes as etapas da primeira cromatografia de afinidade foram avaliadas por SDS PAGE. A eletroforese mostrou a presença de uma banda majoritária em aproximadamente 45 kDa, correspondente à proteína de fusão His<sub>6</sub>GST-NSD3-PWWP (Figura 9B).

O próximo passo no processo de purificação consistiu em submeter a amostra de proteína à clivagem com a protease TEV. Esta protease é utilizada para a remoção de caudas de purificação de proteínas recombinantes por ser extremamente específica, reconhecendo a sequência heptapeptídica (EXXYXQ\*S/G) (X pode ser qualquer resíduo), além de ser ativa em baixas temperaturas e na presença de vários inibidores normalmente adicionados durante o processo de purificação.

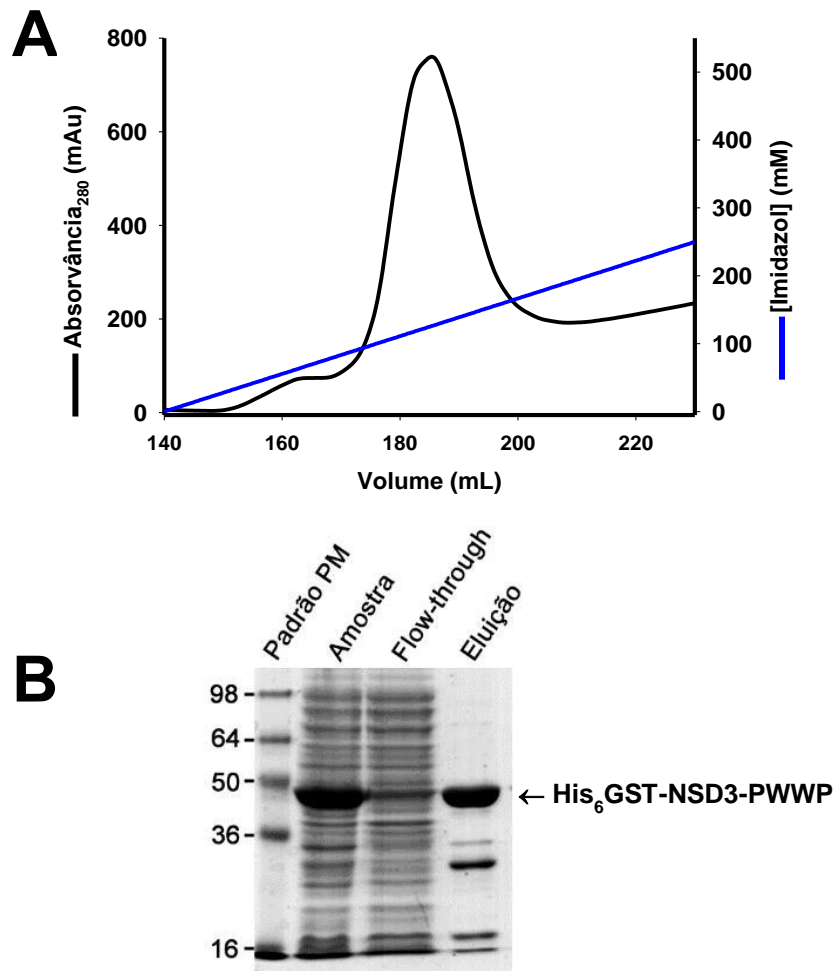


Figura 9: Purificação de His<sub>6</sub>GST-NSD3-PWWP por cromatografia de afinidade a níquel. (A) Cromatograma de eluição de His<sub>6</sub>GST-NSD3-PWWP da coluna de afinidade His-Trap HP (GE Healthcare), evidenciando o volume máximo de eluição em ~120 mM de imidazol. Em preto, o traçado de absorvância a 280 nm. Em azul, o gradiente de imidazol. (B) Análise por SDS PAGE das frações coletadas em (A). PM: padrão de massa molecular; Amostra: His<sub>6</sub>GST-NSD3-PWWP antes da cromatografia; *Flow-through*; Eluição: constituição proteica do pico da primeira cromatografia de afinidade. A proteína de fusão foi eluída ao longo do pico da primeira cromatografia de afinidade e corre com banda de aproximadamente 45 kDa no gel de SDS PAGE.

A eficiência da clivagem de His<sub>6</sub>GST-NSD3-PWWP pela protease TEV foi verificada por SDS PAGE (Figura 10). Após o tratamento com TEV, por 18 h, a banda de 45 kDa, referente ao pico de eluição da primeira cromatografia de afinidade a níquel, separou-se em duas bandas, uma com ~28 kDa referente à cauda His<sub>6</sub>GST, mas também à própria TEV que apresenta a mesma massa molecular (28 kDa), e outra com ~17 kDa referente à NSD3-PWWP. Esses resultados sugerem que

His<sub>6</sub>GST-NSD3-PWWP sofreu clivagem completa dentro das condições experimentais utilizadas.

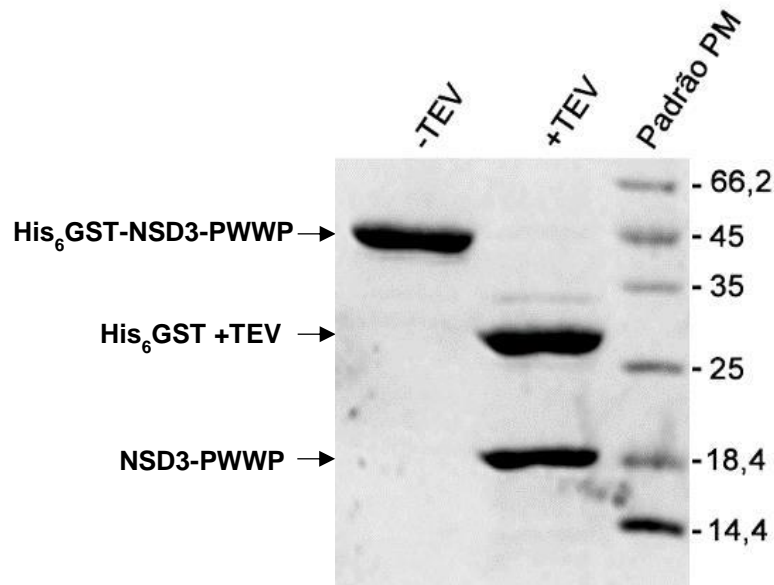


Figura 10: Clivagem de His<sub>6</sub>GST-NSD3-PWWP com a protease TEV. A clivagem da cauda N-terminal His<sub>6</sub>GST foi realizada por incubação com a protease TEV em razão molar 5:1 (His<sub>6</sub>GST-NSD3-PWWP :TEV) em temperatura ambiente durante 18 h. A eficiência da clivagem foi analisada através de SDS PAGE. PM: padrão de massa molecular; -TEV: amostra de proteína na ausência de TEV; +TEV: amostra após 18 h de incubação com TEV. A migração da banda proteica para menor massa molecular (~17 kDa) indica clivagem completa após 18 h de incubação. A banda em ~28 kDa no poço (+TEV) se refere tanto à protease TEV quanto à cauda His<sub>6</sub>GST.

A amostra contendo a construção NSD3-PWWP clivada foi submetida a outra cromatografia de afinidade a níquel tendo como objetivo separar a proteína de interesse da cauda His<sub>6</sub>GST livre e da protease His<sub>6</sub>-TEV. A Figura 11A mostra o cromatograma de eluição, onde a proteína NSD3-PWWP livre da cauda foi eluída diretamente no *flow-through*. O pico de eluição, em concentração de 180 mM de imidazol, refere-se à TEV e à cauda His<sub>6</sub>GST. A constituição proteica das frações obtidas durante a segunda cromatografia de afinidade a níquel foi avaliada por SDS PAGE, revelando a presença de NSD3-PWWP livre (~17 kDa) no *flow-through* e da cauda His<sub>6</sub>GST juntamente com a protease TEV (~28 kDa) na eluição com imidazol (Figura 11B).



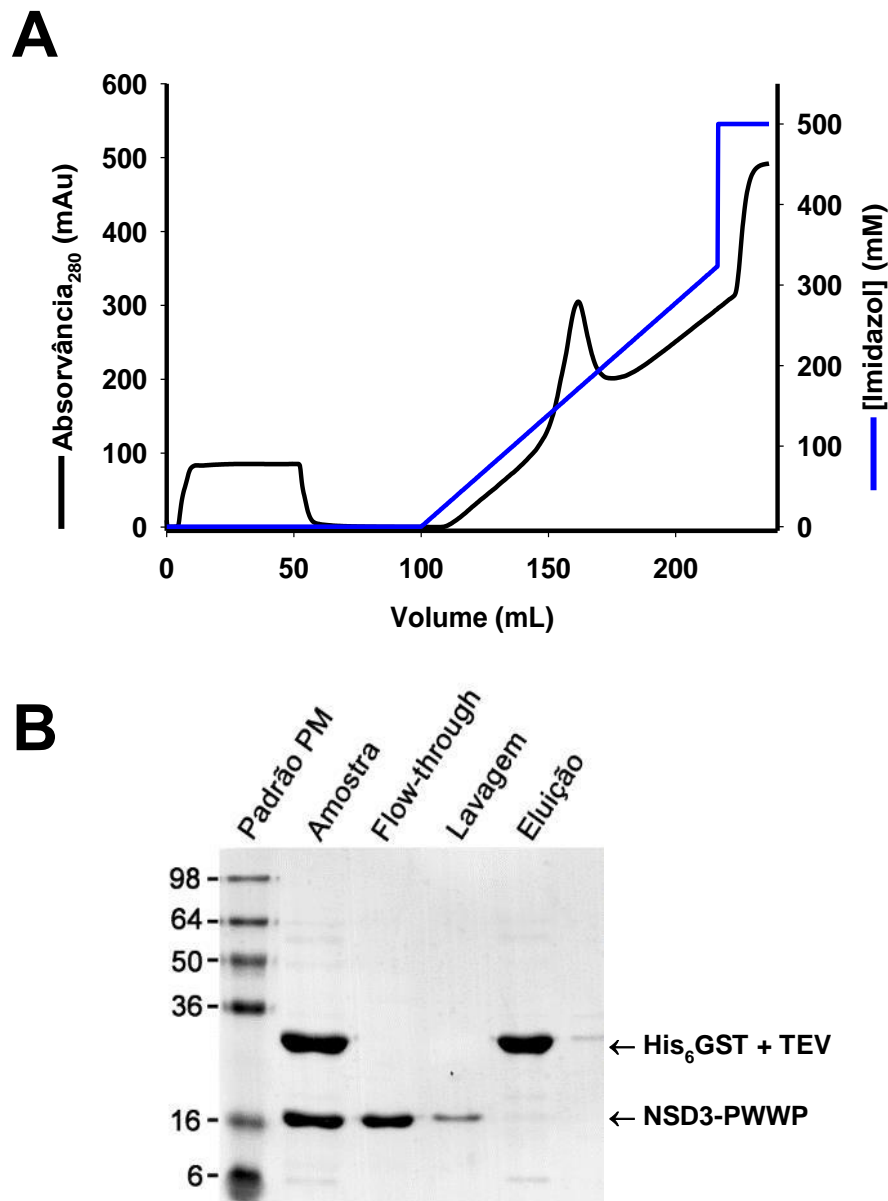


Figura 11: Purificação de NSD3-PWWP através de uma segunda etapa de cromatografia de afinidade a níquel. (A) Cromatograma de eluição de NSD3-PWWP da coluna de afinidade His-Trap HP (GE Healthcare). Em preto, o traçado de absorvância a 280 nm. Em azul, o gradiente de imidazol. (B) Análise por SDS PAGE das frações coletadas em (A). PM: padrão de massa molecular; Amostra: His<sub>6</sub>GST-NSD3-PWWP após clivagem com TEV; *Flow-through*; Lavagem: lavagem com tampão A [Tris-HCl 50 mM (pH 8,0), NaCl 500 mM, imidazol 5 mM, β-mercaptoetanol 10 mM]; Eluição: constituição proteica do pico da segunda cromatografia de afinidade. A proteína foi obtida no *flow-through* e corre com banda de aproximadamente 17 kDa no gel de SDS PAGE.

A amostra contendo NSD3-PWWP foi concentrada por ultrafiltração e submetida a última etapa de purificação por cromatografia de exclusão molecular. A Figura 12

mostra o cromatograma de eluição da etapa de gel-filtração. Observa-se a presença de um único pico de eluição com máximo em volume de 77 mL, o que é compatível com o esperado para uma proteína de ~17 kDa no seu estado monomérico. A análise por eletroforese revelou a presença de uma banda proteica de ~17 kDa correspondente à proteína (Figura 12).

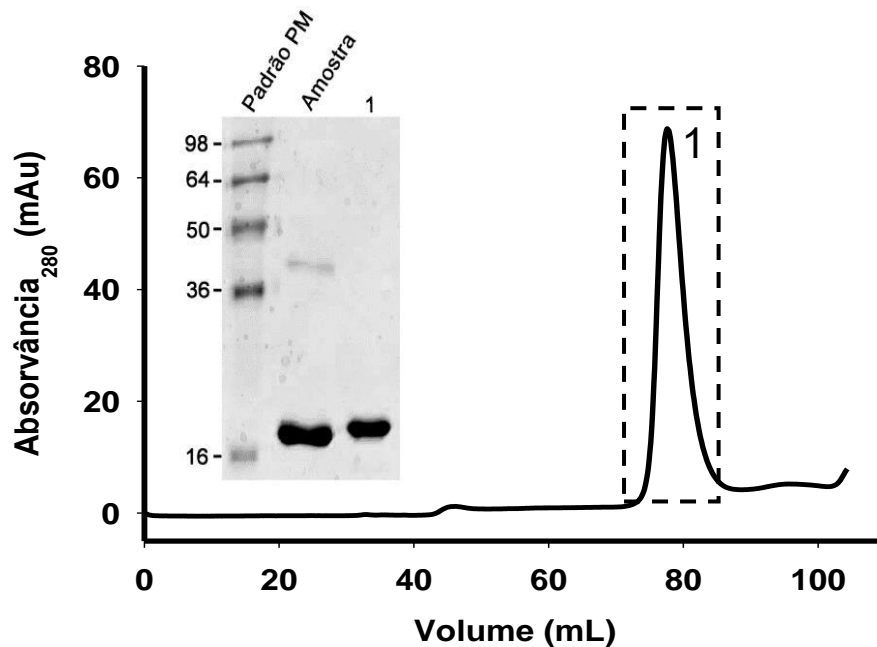


Figura 12: Purificação de NSD3-PWWP por cromatografia de exclusão molecular. Cromatograma de eluição de NSD3-PWWP da coluna de gel-filtração Superdex 75 16/600 (GE Healthcare), evidenciando um único pico em 77 mL. A purificação foi realizada em tampão fosfato de sódio 20 mM (pH 7,4), 100 mM NaCl. Em preto, o traçado de absorvância a 280 nm. Inset: Análise por SDS PAGE das frações coletadas durante a etapa de purificação. PM: padrão de massa molecular; Amostra: NSD3-PWWP após a segunda cromatografia de afinidade a níquel; 1: constituição proteica do pico 1 (77 mL), evidenciando a amostra de NSD3-PWWP purificada.

## 4.2 Caracterização estrutural preliminar de NSD3-PWWP

### 4.2.1 $^1\text{H}$ 1D RMN

A Figura 13 mostra o espectro unidimensional de RMN de hidrogênio obtido para a amostra purificada de NSD3-PWWP. Este espectro é compatível com o de uma proteína enovelada, mostrando sinais de RMN com linhas finas e uma grande dispersão de deslocamento químico na região dos hidrogênios amídicos (10-6 ppm). A presença de sinais abaixo de 0 ppm, referentes a hidrogênios de metilas próximos

a anéis aromáticos (*ring current shifts*) confirma o estado enovelado deste domínio. As ressonâncias em 5-6 ppm confirmam a presença de estruturas em folha- $\beta$ . Além disso, as ressonâncias acima de 10 ppm correspondem aos hidrogênios indólicos dos triptofanos presentes no domínio.

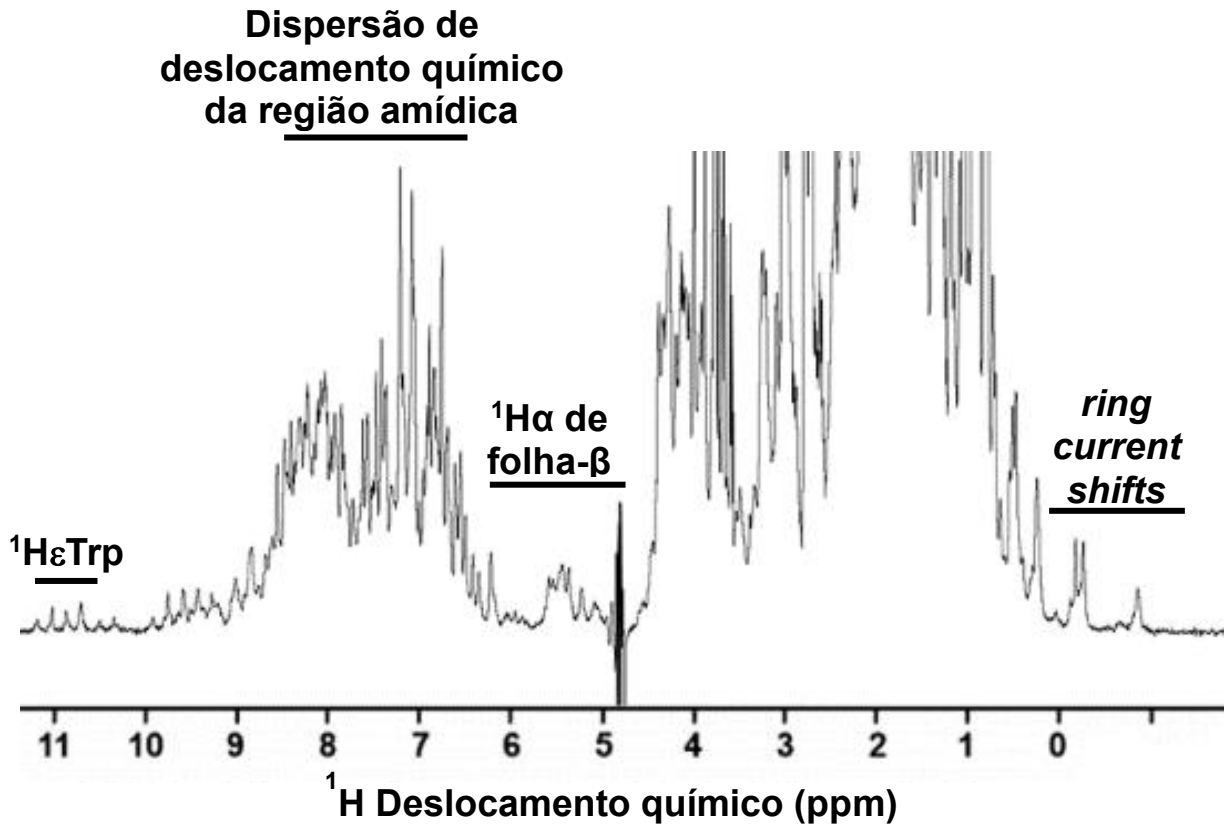


Figura 13: Espectro 1D  $^1\text{H}$  RMN de NSD3-PWWP. O espectro foi adquirido com  $^{15}\text{N}$ -NSD3-PWWP purificado (0,3 mM) em tampão 20 mM fosfato de sódio (pH 6,5), 50 mM NaCl, 5 mM DTT, 10%  $\text{D}_2\text{O}$  a temperatura de 25°C.

#### 4.2.2 [ $^1\text{H}$ , $^{15}\text{N}$ ] HSQC

A Figura 14 mostra o espectro bidimensional de correlação heteronuclear [ $^1\text{H}$ , $^{15}\text{N}$ ] HSQC (*Heteronuclear Single Quantum Coherence*) adquirido para a amostra purificada de  $^{15}\text{N}$ -NSD3-PWWP. Esse espectro é referido como a impressão digital de uma proteína e fornece informações acerca das correlações entre hidrogênios covalentemente ligados a nitrogênios na sequência primária. Ou seja, esse espectro mostra as correlações entre os hidrogênios e nitrogênios amídicos da proteína, de forma que é gerado um pico por aminoácido da cadeia polipeptídica, com exceção das prolinas, que não possuem hidrogênio amídico. Além disso, esse espectro mostra picos de correlação para os grupos  $\text{NH}_2$  das cadeias laterais de asparagina e

glutamina, cujos valores de deslocamento químico de  $^{15}\text{N}$  caem em torno de  $\sim 113$  ppm. De um modo geral, o espectro mostra sinais de ressonância bem definidos com intensidade relativamente homogênea e boa dispersão de deslocamento químico, o que é condizente com uma proteína enovelada.

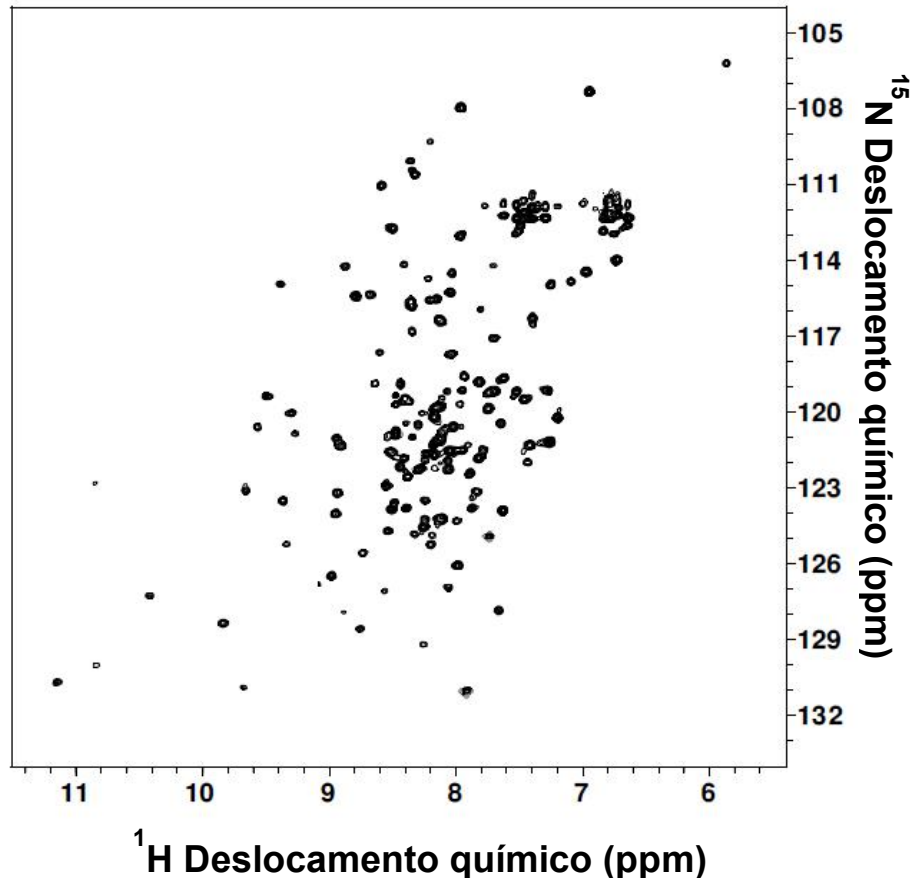


Figura 14: Espectro 2D [ $^1\text{H}$ ,  $^{15}\text{N}$ ] HSQC de NSD3-PWWP. O espectro foi adquirido com  $^{15}\text{N}$ -NSD3-PWWP (0,3 mM) em tampão 20 mM fosfato de sódio (pH 6,5), 50 mM NaCl, 5 mM DTT, 10%  $\text{D}_2\text{O}$  a temperatura de  $25^\circ\text{C}$ .

### 4.3 Afinidade e especificidade de interação de NSD3-PWWP com peptídeos de histonas

Estudos prévios revelaram os mecanismos de reconhecimento e especificidade de ligação de domínios leitores com marcas epigenéticas em histonas (BANNISTER & KOUZARIDES, 2011; MUSSELMAN et al., 2012; PATEL & WANG, 2013; ROTHBART & STRAHL, 2014; ANDREWS et. al., 2016) . No entanto, este tipo de estudo empregando o nucleossomo íntegro tem se mostrado muito difícil, pela dificuldade de se controlar os sítios específicos de PTMs em histonas utilizando preparações de nucleossomos purificados de células, mas também pela dificuldade

de se remontar o nucleossomo *in vitro* a partir de seus componentes recombinantes (WEAVER et al., 2018). Assim, utilizamos a abordagem na qual a interação do domínio leitor com a marca epigenética é investigada no contexto de fragmentos peptídicos derivados de histonas.

Para compreender as preferências de ligação do domínio PWWP de NSD3, investigamos a afinidade de interação deste domínio com nove peptídeos derivados das caudas N-terminais das histonas H3 e H4 contendo diferentes lisinas metiladas e graus de metilação (Tabela 1) através de espectroscopia de fluorescência. Esta técnica foi utilizada uma vez que domínios PWWP conhecidamente utilizam seus resíduos de triptofano presentes no motivo de sequência conservado Pro-Trp-Trp-Pro para interagir com metil-lisinas. O espectro de fluorescência intrínseca de NSD3-PWWP exibiu emissão máxima em 344 nm (Figura 15A). A emissão de fluorescência do triptofano é altamente dependente da polaridade do meio, de modo que triptofanos expostos a solventes polares apresentam um desvio para o vermelho no seu espectro de fluorescência (EFTINK, 1991). Portanto, o comprimento de onda máximo de emissão observado para NSD3-PWWP (344 nm) sugere a presença de triptofanos parcialmente expostos ao solvente. Este resultado é suportado pelo modelo tridimensional de NSD3-PWWP que apresenta ao menos três resíduos de triptofano cujas cadeias laterais estão voltadas para fora da superfície da proteína (Figura 15C).

A titulação dos peptídeos de histonas promoveu a supressão parcial e dose-dependente da fluorescência intrínseca de NSD3-PWWP (Figura 15A). Este resultado sugere que ao menos um dos cinco resíduos de triptofano presentes em NSD3-PWWP participa diretamente e compõe o sítio de interação com os peptídeos de histonas. A partir dos valores de intensidade de fluorescência total, obtidos a partir da medida da área sob os espectros, em função da concentração de ligante, foi possível obter as isotermas de ligação para cada um dos nove peptídeos (Figura 15B e Figura 16). Todas as curvas de ligação apresentaram um caráter bifásico, com duas inclinações distintas, sugerindo a presença de dois sítios de ligação com afinidades diferentes, um de alta e outro de baixa afinidade. De fato, os dados experimentais só puderam ser corretamente ajustados empregando uma equação que leva em consideração duas constantes de dissociação aparentes distintas. Os valores das constantes de dissociação estimadas para cada peptídeo estão resumidos na Tabela 3.

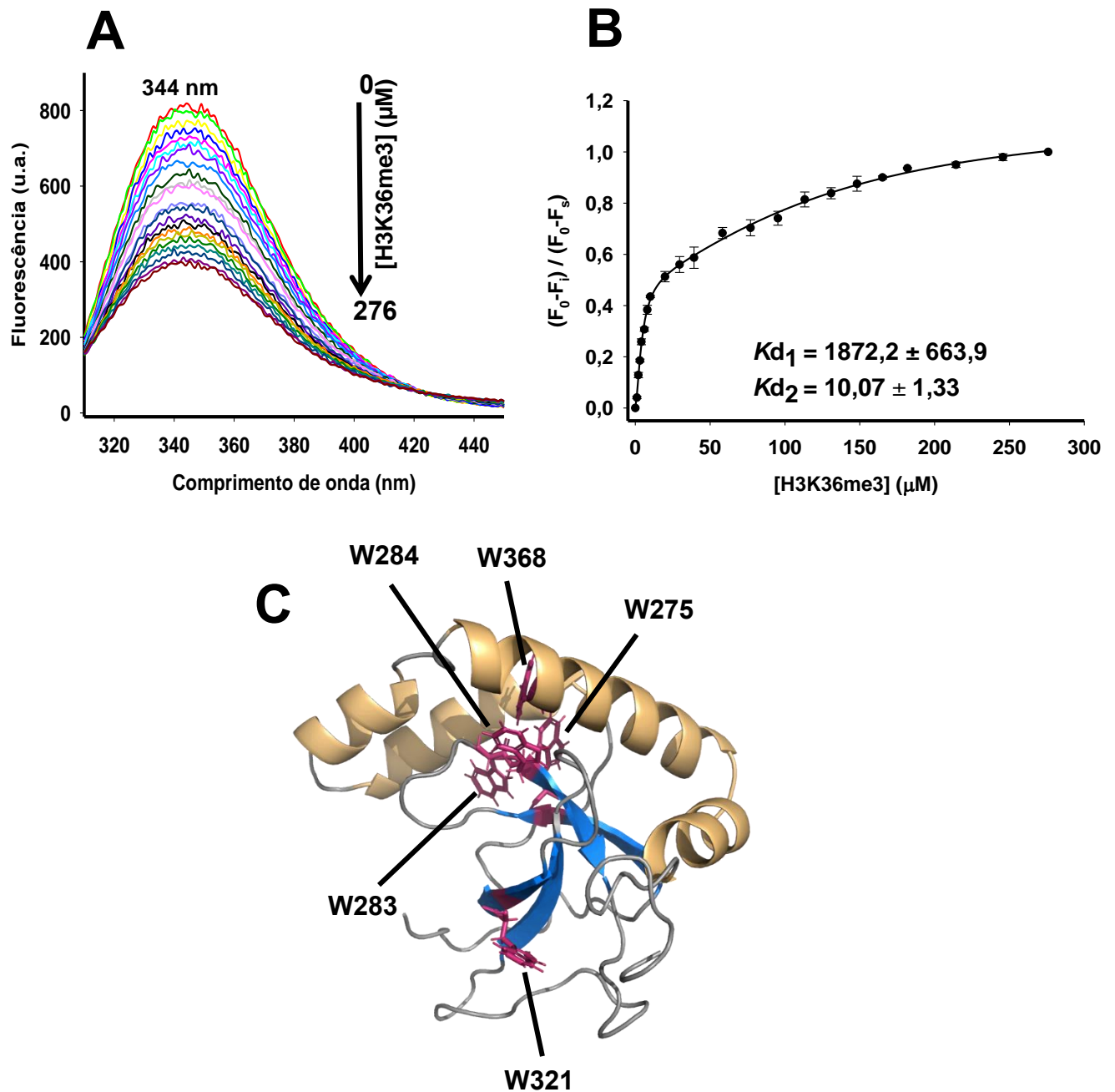


Figura 15: Titulação de NSD3-PWWP com H3K36me3 investigada por espectroscopia de fluorescência. Concentrações crescentes de peptídeo (0 a 276  $\mu\text{M}$ ) foram adicionadas a 2  $\mu\text{M}$  de NSD3-PWWP em tampão 20 mM fosfato de sódio (pH 7,4), 100 mM NaCl. (A) Espectros de fluorescência de NSD3-PWWP evidenciando o máximo de emissão em 344 nm e a supressão de fluorescência dose-dependente em função da concentração de peptídeo. (B) Isotherma de ligação derivada dos dados de fluorescência. O ajuste dos dados experimentais originou duas constantes de afinidade ( $K_{d1}$  e  $K_{d2}$ ). (C) Modelo estrutural de NSD3-PWWP. As fitas- $\beta$  estão representadas em azul, as  $\alpha$ -hélices em laranja e as regiões de alça em cinza. As cadeias laterais dos cinco triptofanos presentes na proteína estão representadas em rosa e marcadas.

A presença de dois sítios de ligação entre NSD3-PWWP e os peptídeos de histonas é improvável, devido ao pequeno tamanho do domínio em relação ao peptídeo. Estruturas tridimensionais de complexos entre domínios PWWP e peptídeos derivados de histonas (HRP3-PWWP:H3K36me3, DNMT3B-PWWP: H3K36me3, BRPF1-PWWP:H3K36me3) revelaram que a ligação se dá com estequiometria de 1 (1 domínio PWWP para 1 peptídeo) (WANG & TIAN, 2019; VEZZOLI et al., 2010; RONDELET et al., 2016). Portanto, o comportamento bifásico das isotermas de ligação obtidas a partir dos dados de fluorescência não se deve a ligação de 2 moléculas de peptídeo por molécula de NSD3-PWWP.

A fim de melhor compreender a natureza da interação responsável pelas duas inclinações presentes nas isotermas de ligação, o modelo estrutural de NSD3-PWWP foi sobreposto a diversas estruturas tridimensionais de complexos entre domínios PWWP e peptídeos derivados de histonas (HRP3-PWWP:H3K36me3, DNMT3B-PWWP: H3K36me3, BRPF1-PWWP:H3K36me3) (Figura 17). A sobreposição sugere que a cadeia lateral da lisina trimetilada do peptídeo encontra-se voltada diretamente para o resíduo Trp284 presente na gaiola aromática de NSD3-PWWP, evidenciando a interação. Além disso, observa-se no modelo estrutural do complexo, a proximidade entre a valina da cadeia lateral do peptídeo e o resíduo Trp321, sugerindo que o peptídeo também faz contatos com este triptofano. Portanto, o modelo estrutural do complexo NSD3-PWWP:H3K36me3 sugere que o comportamento bifásico da isoterma de ligação pode estar relacionado ao fato do peptídeo interagir de maneira diferente com ao menos dois resíduos de triptofano da proteína. A interação da metil-lisina com o Trp284 da gaiola aromática constituiria o sítio de maior afinidade, enquanto a interação de resíduos do peptídeo com o Trp321 constituiria um sítio de menor afinidade.

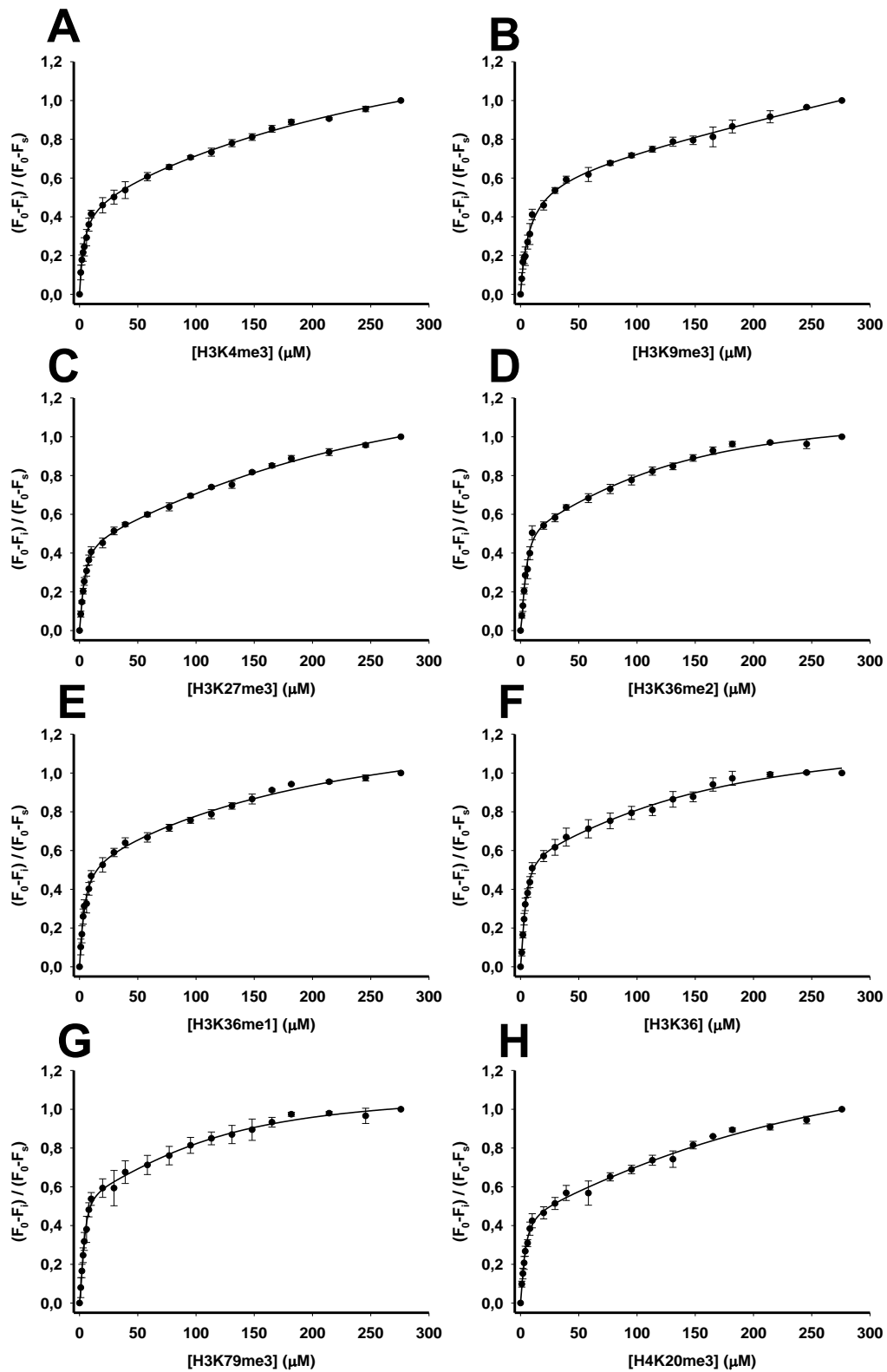


Figura 16: Isothermas de ligaç o de NSD3-PWWP com pept deo de histonas. As curvas mostram a supress o de fluoresc ncia em funç o da concentraç o de ligante para os pept deos (A) H3K4me3, (B) H3K9me3, (C) H3K27me3, (D) H3K36me2, (E) H3K36me1, (F) H3K36, (G) H3K79me3, (H) H4K20me3. Os dados experimentais se referem a m dia  $\pm$  erro padr o de tr s experimentos independentes.



Peptídeo	$K_d1$ ( $\mu\text{M}$ )	$K_d2$ ( $\mu\text{M}$ )
H3K36me3	1872,2 $\pm$ 663,9	10,1 $\pm$ 1,3
H3K36me2	1901,7 $\pm$ 1117,3	9,7 $\pm$ 2,1
H3K79me3	2358,9 $\pm$ 1174,6	8,5 $\pm$ 1,4
H3K36	1512,9 $\pm$ 615,8	7,2 $\pm$ 1,0
H3K9me3	11203,1 $\pm$ 144422,1	6,9 $\pm$ 0,9
H3K27me3	1265,6 $\pm$ 329,5	5,6 $\pm$ 0,6
H4K20me3	1597,6 $\pm$ 725,3	5,6 $\pm$ 1,0
H3K36me1	577,6 $\pm$ 243,9	4,9 $\pm$ 0,7
H3K4me3	473,8 $\pm$ 149,9	3,8 $\pm$ 0,5

Tabela 3: Constantes de dissociação aparente obtidas para a interação de NSD3-PWWP com peptídeos de histonas.

Os valores de  $K_d1$  são da ordem de mM, sugerindo um sítio de ligação de baixa afinidade. Entretanto, estes valores devem ser interpretados com cautela, uma vez que o erro experimental associado a eles é alto, indicando que esta afinidade não pôde ser bem determinada dentro da faixa de concentração de ligante utilizada. Os valores de  $K_d2$  são da ordem de baixo  $\mu\text{M}$ , sugerindo um sítio de alta afinidade (levando em consideração o tamanho do domínio isolado e do peptídeo, bem como o número de contatos no complexo). É interessante notar que todos os peptídeos, independente da natureza da sequência ou do grau de metilação, interagiram com NSD3-PWWP com afinidades similares (3,8 – 10,1  $\mu\text{M}$ ), sugerindo que o domínio PWWP de NSD3 interage com peptídeos derivados de histonas com alta afinidade, porém com baixa seletividade. Entretanto, NSD3-PWWP apresentou afinidades maiores para H3K4me3 e para a forma monometilada de H3K36 (afinidade duas vezes maior quando comparada as formas di- e trimetiladas de H3K36), sugerindo ligeira especificidade frente a estas sequências.

Quando comparado com outros domínios PWWP, como exemplo os citados na Tabela 4 (WU et al., 2011), NSD3-PWWP interage com maior afinidade com os peptídeos (baixo  $\mu\text{M}$ ) do que os demais domínios, sugerindo que NSD3-PWWP é um módulo de reconhecimento de metil-lisinas de alta afinidade.

Peptídeo	Sequência	BRPF1 Kd (mM)	BRPF2 Kd (mM)	HDGF2 Kd (mM)	MUM1 Kd (mM)	WHSC1 Kd (mM)
<b>H3K36me2</b>	PATGGVK(me)2KPHRY	0,8 ± 0,2	Muito fraco	1,0 ± 0,2	0,4 ± 0,1	0,5 ± 0,1
<b>H3K36me3</b>	PATGGVK(me)3KPHRY	0,8 ± 0,1	0,9 ± 0,1	Muito fraco	0,1 ± 0,07	0,6 ± 0,1
<b>H3K79me2</b>	EIAQDFK(me)2TDLRY	0,2 ± 0,01	0,2 ± 0,02	Muito fraco	ND	ND
<b>H3K79me3</b>	EIAQDFK(me)3TDLRY	0,1 ± 0,01	0,4 ± 0,1	Muito fraco	ND	ND
<b>H4K20me2</b>	AKRHRK(me)2VLRDN	Muito fraco	ND	Muito fraco	ND	ND
<b>H4K20me3</b>	AKRHRK(me)3VLRDN	Muito fraco	ND	Muito fraco	ND	ND

Tabela 4: Afinidades de ligação de domínios PWWP humanos com peptídeos metilados das histonas H3 e H4. WHSC1 contém dois domínios PWWP os dados listados são apenas o domínio PWWP do N-terminal. ND: não determinado. Muito fraco: a ligação é muito fraca para ser ajustada de forma confiável. (Adaptado WU et. al., 2011)

Sendžikaitė et al. (2019) sugere que, em geral, os domínios PWWP interagem com caudas de histonas metiladas, incluindo, mas não se limitando a H3K36me3 (SENDŽIKAITĖ et al., 2019). Wang et al. (2019) destaca que análises de deslocamento químico permitiram determinar constantes de dissociação de baixo mM para a ligação do domínio TbTFIIS2-2-PWWP ao peptídeo H4K17 trimetilado e a H3K32, que são comparáveis com os  $K_d$ s obtidos para a ligação de Brpf1-PWWP a H3K36me3 e para a interação entre Pdp1-PWWP a H4K20me3 (WANG et al., 2019). Desta forma, podemos observar que outros domínios PWWP também mostraram pouca especificidade de interação frente a peptídeos de histonas, sugerindo que o domínio PWWP não é exatamente um leitor específico de H3K36me3 como sugerido previamente (VERMEULEN et al., 2010) ou a especificidade de interação pode ser dada pela interação simultânea de outros domínios da proteína com o nucleossomo como um todo.

Apesar da pouca especificidade, NSD3-PWWP mostrou afinidade ligeiramente maior para H3K4me3 e H3K36me1. O fato da afinidade ser duas vezes maior para a forma monometilada do que as formas di- e trimetilada de H3K36 sugere que o domínio PWWP N-terminal reconhece H3K36me1 para que o domínio catalítico SET promova a trimetilação desta lisina.

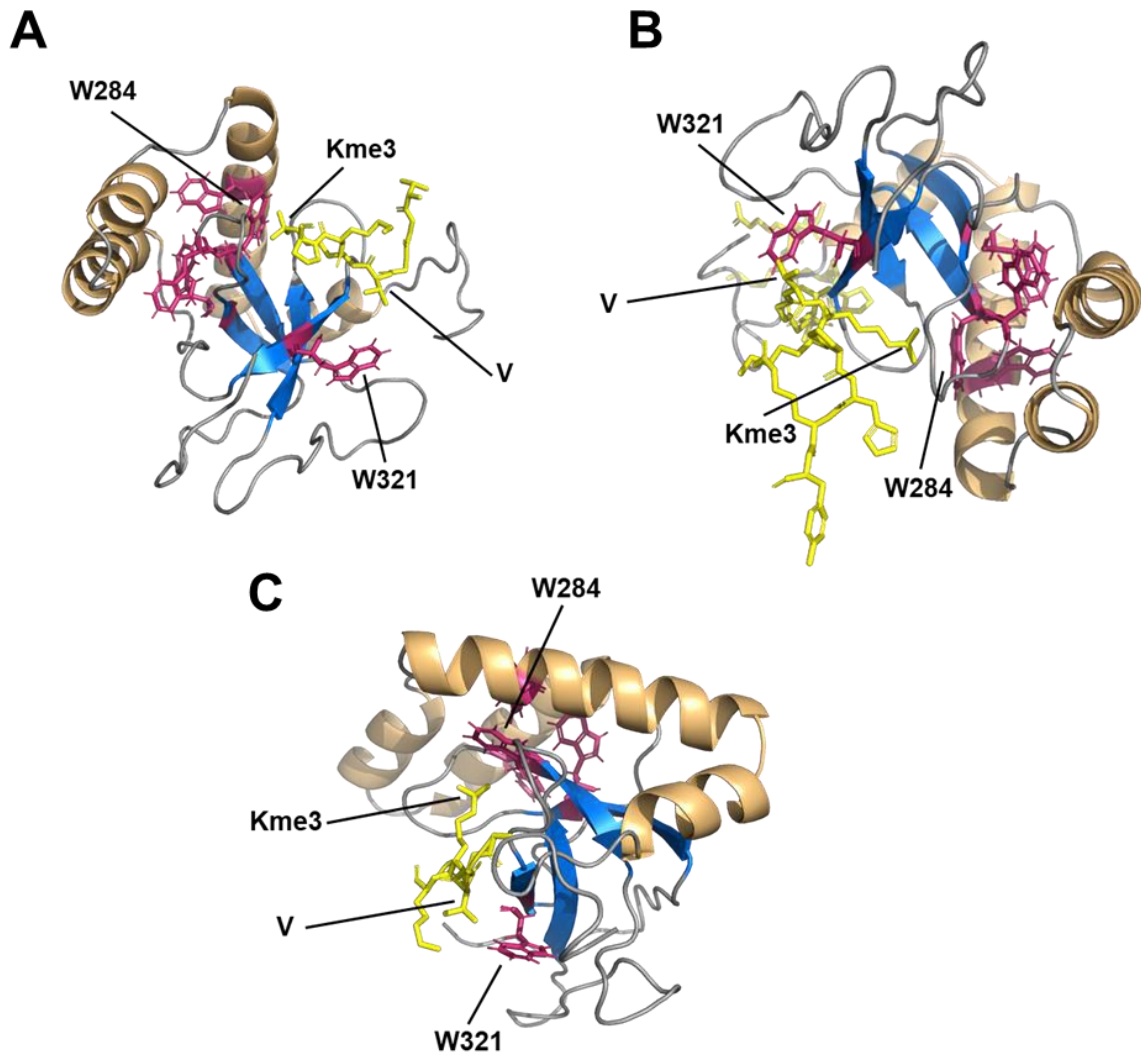


Figura 17: Sobreposição do modelo estrutural de NSD3-PWWP com estruturas tridimensionais de complexos PWWP:peptídeos de histonas previamente descritos. As fitas- $\beta$  estão representadas em azul, as  $\alpha$ -hélices em laranja e as regiões de alça em cinza. As cadeias laterais dos resíduos de triptofano estão representadas em rosa. Os peptídeos estão representados em bastões amarelos. (A) Sobreposição do modelo de NSD3-PWWP com o complexo DNMT3B-PWWP: H3K36me3 (PDB 5ciu). (B) Sobreposição do modelo de NSD3-PWWP com o complexo BRPF1-PWWP:H3K36me3 (PDB 2x4y). (C) Sobreposição do modelo de NSD3-PWWP com o complexo HRP3-PWWP:H3K36me3 (PDB 6iis). Para facilitar a visualização, após sobreposição, a estrutura tridimensional dos domínios PWWP de DNMT3B, BRPF1 e HRP3 não estão representadas.

#### 4.4 Estabilidade térmica de NSD3-PWWP

Apesar dos resultados estruturais obtidos para NSD3-PWWP nas condições experimentais descritas no item 3.2.7, a determinação da estrutura tridimensional deste domínio através da espectroscopia de Ressonância Magnética Nuclear foi

impossibilitada pela baixa solubilidade e estabilidade da proteína purificada. NSD3-PWWP, nas condições experimentais testadas, mostra um limite de solubilidade de aproximadamente 100  $\mu\text{M}$ , entretanto é muito pouco estável e sofre agregação e precipitação em função do tempo, o que impediu a aquisição dos dados de RMN. Portanto, faz-se necessária a otimização das condições experimentais, como pH e força iônica, a fim de estabilizar a proteína recombinante purificada e facilitar a sua determinação estrutural.

A expressão e purificação de proteínas recombinantes pode ser significativamente melhorada com a otimização das condições do meio, como pH e força iônica, bem como a adição de ligantes que podem reduzir a agregação proteica. A identificação de tampões e aditivos estabilizadores de proteínas é tipicamente feita através da medida da temperatura de fusão ( $T_m$ ) em experimentos de desnaturação térmica (HUYNH; PARTCH, 2015). Para isso, foi utilizado o corante SYPRO Orange que sofre aumento significativo no seu rendimento quântico quando ligado a regiões hidrofóbicas expostas em proteínas, permitindo o monitoramento do processo de desnaturação proteica. Além disso, os comprimentos de onda de excitação e emissão de fluorescência do SYPRO Orange são compatíveis com conjuntos de filtros comumente encontrados em instrumentos de PCR em tempo real, permitindo a adaptação destes equipamentos para ensaios de desnaturação térmica de proteínas. Isto permite que o ensaio seja realizado em placas de 96 poços, possibilitando o acompanhamento simultâneo da desnaturação térmica em várias condições experimentais (NIESEN; 2007).

O perfil das curvas de desnaturação térmica obtidas nos tampões A1, A2, A3, A4 e A5 (Figura 18), todos em pH ácido (3,0 a 4,5), sugere que, a temperatura de 25  $^{\circ}\text{C}$ , NSD3-PWWP já se encontra desenovelada, pois nestas condições a amostra já apresenta valor de fluorescência elevado, de modo que as regiões hidrofóbicas estão voltadas para fora da proteína. Portanto, condições de pH ácido, abaixo de 4,5, promovem desnaturação de NSD3-PWWP, provavelmente por repulsão eletrostática.

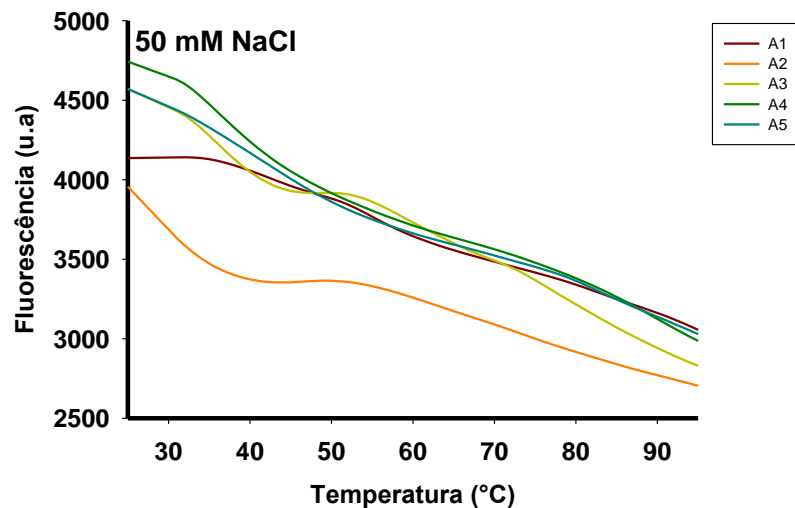


Figura 18: Curva de desnaturação térmica de NSD3-PWWP nos tampões A1 (glicina, pH 3,0), A2 (Ácido cítrico, pH 3,2), A3 (PIPPS, pH 3,7), A4 (Ácido cítrico, pH 4,0) e A5 (Acetato de sódio, pH 4,5).

Para as demais condições (Figura 19), as curvas apresentaram um perfil sigmoide típico de desnaturação térmica de proteínas na presença do fluoróforo SYPRO Orange. À temperatura inicial de 25°C, o baixo valor de fluorescência indica que a proteína encontra-se enovelada. O decaimento do valor de fluorescência a temperaturas muito elevadas é provocado pela agregação proteica (HUYNH; PARTCH, 2015). A temperatura de fusão ( $T_m$ ) da proteína em cada condição experimental pode ser identificada a partir da primeira derivada do valor de fluorescência em função da temperatura, sendo o  $T_m$  representado pelo mínimo desta curva. Os valores de  $T_m$  obtidos para NSD3-PWWP em todas as condições analisadas estão listados na Tabela 5.

Dentre as condições analisadas, NSD3-PWWP apresentou temperaturas de fusão maiores quanto maior a concentração de NaCl no meio, visto que os valores de  $T_m$  para as condições contendo 250 mM de sal foram até 7°C maiores do que aqueles para as condições com 50 mM de sal (Tabela 5). Este resultado pode ser explicado pelo efeito de “*salting in*” desempenhado pelo cloreto de sódio, o qual interage com a camada mais externa da proteína aumentando sua interação com o meio aquoso e diminuindo a interação proteína-proteína não-específica, o que leva a diminuição de agregação e aumento da solubilidade proteica (Lehninger, 2014).

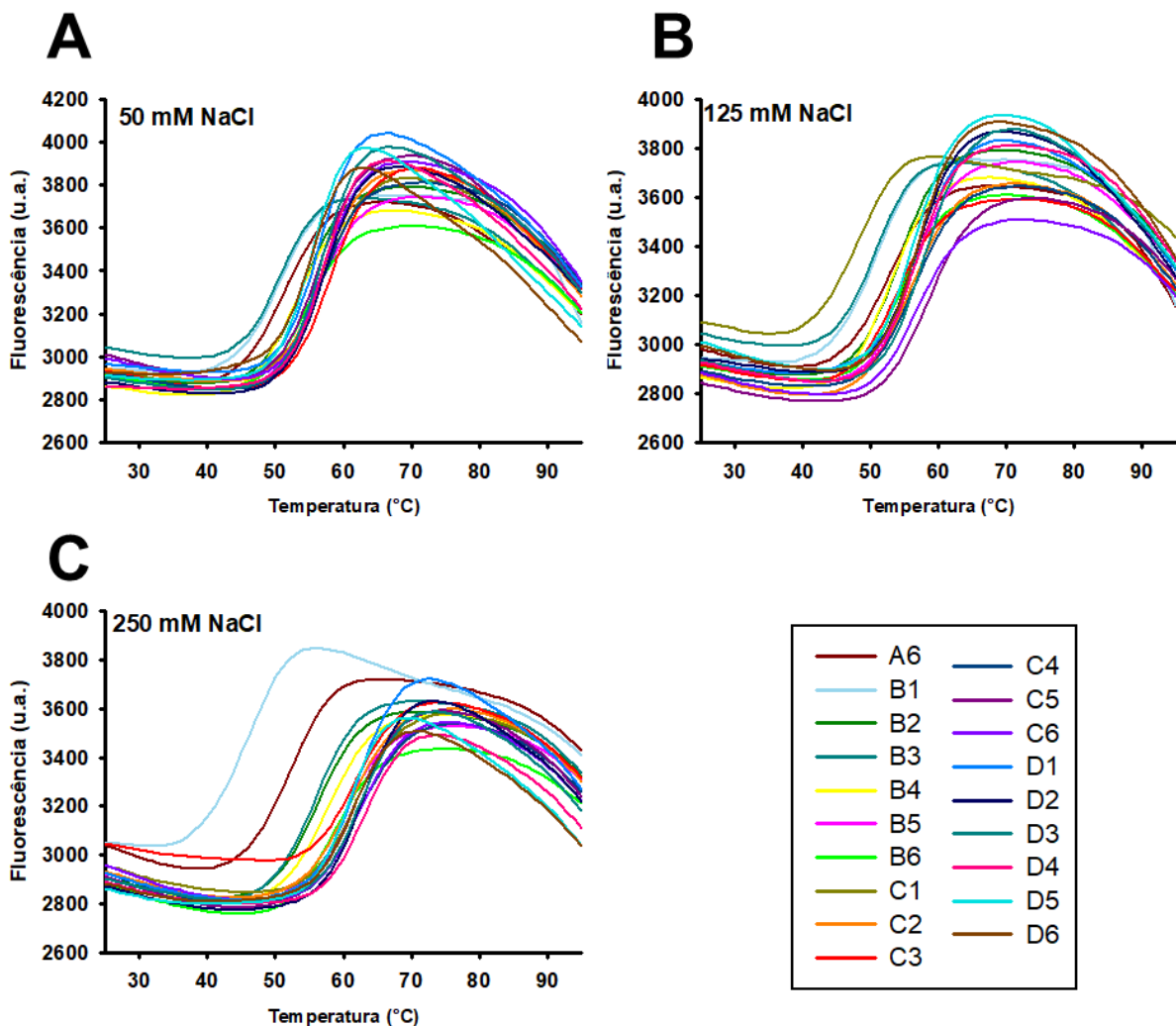


Figura 19: Análise da estabilidade térmica de NSD3-PWWP por fluorimetria diferencial de varredura. (A) Curvas de desnaturação de NSD3-PWWP na presença de 50 mM NaCl. (B) Curvas de desnaturação de NSD3-PWWP na presença de 125 mM NaCl. (C) Curvas de desnaturação de NSD3-PWWP na presença de 250 mM NaCl. A legenda indica os tampões utilizados em cada condição.

Com relação ao pH, NSD3-PWWP se mostrou mais estável (maior  $T_m$ ) nas condições de pH mais básico, especificamente em pH 8,5 e 9 (Tabela 5). Este resultado é diferente do esperado, visto que a proteína apresenta PI (ponto isoelétrico) teórico de 8,95. Sendo o PI de uma proteína o pH no qual a carga líquida é zero, neste ponto a proteína é menos solúvel e sofre precipitação isoelétrica, o que não se observa nos resultados encontrados.

Ao se associar os resultados de pH e força iônica, verifica-se que quanto maior o pH e a concentração de sal, maior o  $T_m$  de NSD3-PWWP, logo mais estável é a proteína. As melhores condições, dentro das analisadas, foram os tampões D2 [20

mM Bicina (pH 8,5), 250 mM NaCl], D3 [20 mM Tris (pH 9,0), 250 mM NaCl] e D4 [20 mM CHES (pH 9,0), 250 mM NaCl].

Tampão	pH	[NaCl] (mM)	T <sub>m</sub> (°C)	[NaCl] (mM)	T <sub>m</sub> (°C)	[NaCl] (mM)	T <sub>m</sub> (°C)
Fosfato de Sódio/Potássio	5,0	50	51	125	53	250	52
Citrato de Sódio	5,5	50	51	125	48	250	46
Fosfato de Sódio/Potássio	6,0	50	54	125	56	250	56
Bis-Tris	6,0	50	51	125	54	250	56
MES	6,2	50	53	125	57	250	57
ADA	6,5	50	56	125	59	250	60
Bis-Tris Propano	6,5	50	54	125	57	250	59
Acetato de Amônio	7,0	50	57	125	60	250	62
MOPS	7,0	50	57	125	58	250	61
Fosfato de Sódio/Potássio	7,0	50	58	125	60	250	62
HEPES	7,5	50	57	125	60	250	62
Tris	7,5	50	56	125	60	250	62
EPPS	8,0	50	57	125	61	250	62
Imidazol	8,0	50	56	125	60	250	62
Bicina	8,5	50	57	125	61	250	63
Tris	8,5	50	57	125	61	250	63
CHES	9,0	50	57	125	61	250	63
CHES	9,5	50	55	125	59	250	61
CAPS	10,0	50	55	125	59	250	62

Tabela 5: Temperatura de fusão obtida para NSD3-PWWP em 72 condições experimentais contendo diferentes pHs e concentração de sal.

#### 4.5 Estabilidade térmica do complexo NSD3-PWWP:H3K36me3

A fim de investigar se a interação de NSD3-PWWP com peptídeos derivados de histonas aumenta a estabilidade e, conseqüentemente, a solubilidade da proteína, NSD3-PWWP foi complexada com o peptídeo H3K36me3, derivado da histona H3 contendo o resíduo de lisina K36 trimetilado. Primeiramente, o complexo foi formado *in vitro* na proporção molar 1:1 e purificado por cromatografia de exclusão molecular, a fim de garantir a ausência de peptídeo livre na amostra. O perfil de eluição do complexo apresentou um único pico com máximo em ~77,2 mL, levemente menor que o máximo de eluição da proteína livre em ~77,6mL (Figura 20A), sugerindo que o complexo é ligeiramente maior que a proteína livre. A intensidade do pico de eluição

correspondente ao complexo NSD3-PWWP:H3K36me3 foi maior do que o da proteína livre, em injeções de mesma concentração de proteína, supostamente devido a presença de tirosina na sequência do peptídeo, aminoácido capaz de absorver luz no comprimento de onda de 280 nm.

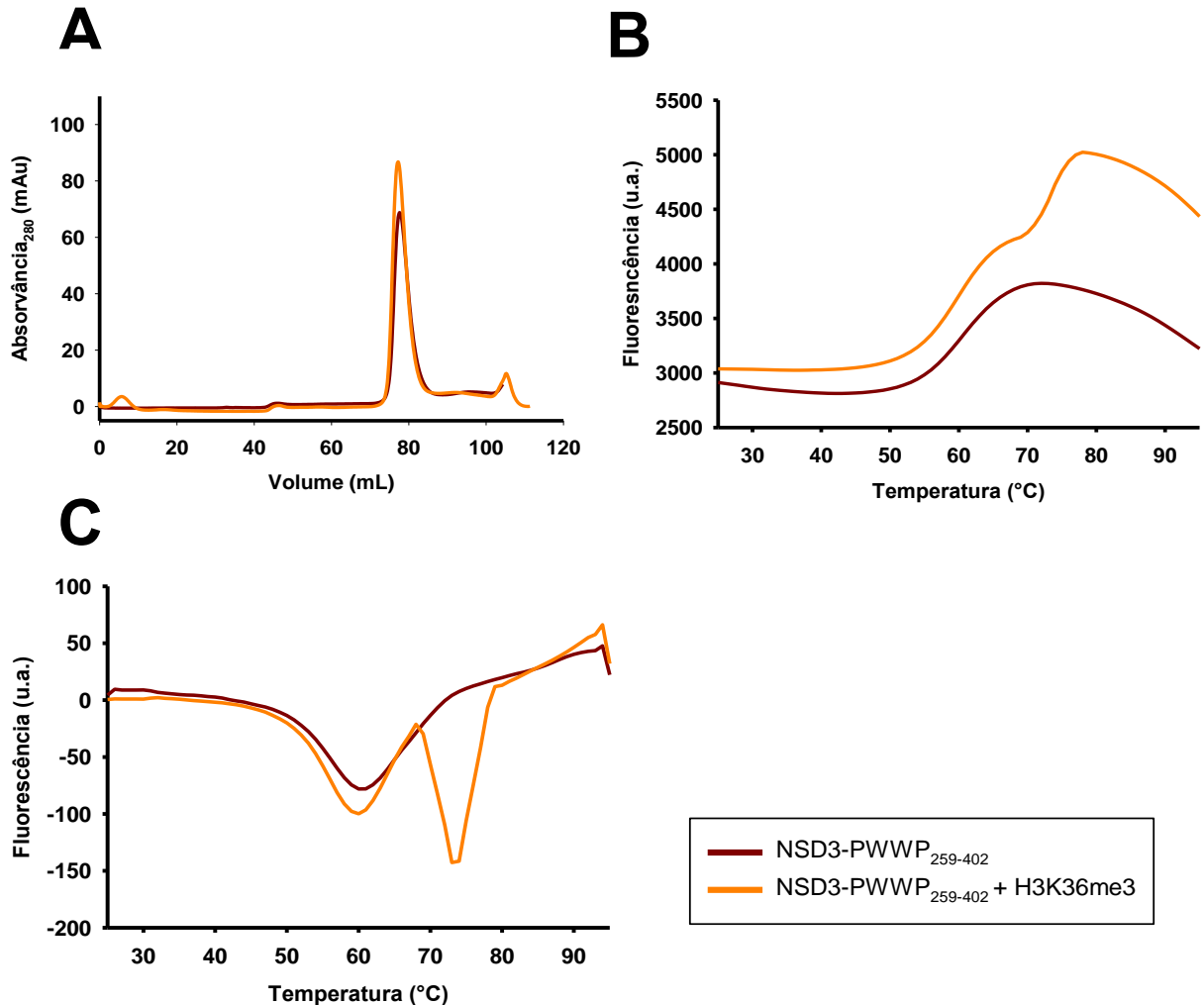


Figura 20: Purificação e desnaturação térmica do complexo NSD3-PWWP:H3K36me3. (A) Perfil de eluição da cromatografia de exclusão molecular em coluna Superdex 75 16/600 (GE Healthcare). Cromatograma de eluição de NSD3-PWWP evidenciando um único pico em 77,2 mL (marrom). Cromatograma de eluição do complexo NSD3-PWWP:H3K36me3, evidenciando um único pico em 77,6mL. O perfil de eluição da amostra livre e complexada são típicos de proteínas globulares e a presença de um único pico com volume de eluição menor sugere formação do complexo. (B) Curva de desnaturação de NSD3-PWWP livre (marrom) e complexada a H3K36me3 (laranja) em tampão [10 mM fosfato de sódio (pH 7,0), 175 mM NaCl]. (C) Primeira derivada da curva de desnaturação obtida em (B). O perfil da curva sugere que o primeiro pico na amostra complexada se refere à proteína livre e o segundo pico em ~73°C se refere ao complexo.



NSD3-PWWP livre e complexada a H3K36me3 foram submetidas ao experimento de fluorimetria diferencial de varredura em tampão [10 mM fosfato de sódio (pH 7,0), 175 mM NaCl], resultando em duas curvas de desnaturação distintas (Figura 20B). O perfil de desnaturação da proteína complexada apresentou duas transições sugerindo que, nesta condição, há a presença de duas espécies proteicas. Comparando a curva de NSD3-PWWP:H3K36me3 com a de NSD3-PWWP livre, sugere-se que a primeira transição é referente à proteína livre, uma vez que apresenta o mesmo  $T_m$  (60 °C) na condição testada, e a segunda transição, por apresentar  $T_m$  mais elevado (73 °C), é referente ao complexo. É possível que tenha havido dissociação parcial do complexo na condição analisada devido ao aumento da força iônica em relação a condição na qual o complexo foi isolado. Sendo o segundo pico referente à proteína complexada ao peptídeo H3K36me3 e o valor de  $T_m2$  à temperatura de fusão do complexo, a ligação entre NSD3-PWWP e H3K36me3 provoca um aumento de  $T_m$  em comparação à proteína livre de 13°C, indicando um aumento na estabilidade da proteína quando complexada.

Em mãos do resultado de que NSD3-PWWP complexada é mais estável do que livre, um novo experimento de fluorimetria diferencial de varredura foi realizado utilizando o complexo NSD3-PWWP:H3K36me3 em 24 condições diferentes, a fim de se determinar o tampão no qual a amostra complexada é mais estável. A Figura 21 mostra o perfil das curvas de desnaturação do complexo NSD3-PWWP:H3K36me3 nos diferentes tampões em baixa concentração de NaCl (50 mM). A amostra apresentou perfil de proteína desnaturada para os tampões com pHs entre 3 e 5,5, em temperatura ambiente, confirmando a alta instabilidade em pH ácido, como a proteína livre (figura 21A). Dentre as demais condições analisadas, as curvas, de uma forma geral, apresentaram duas transições (Figura 21C), sugerindo a presença de duas espécies proteicas em solução.

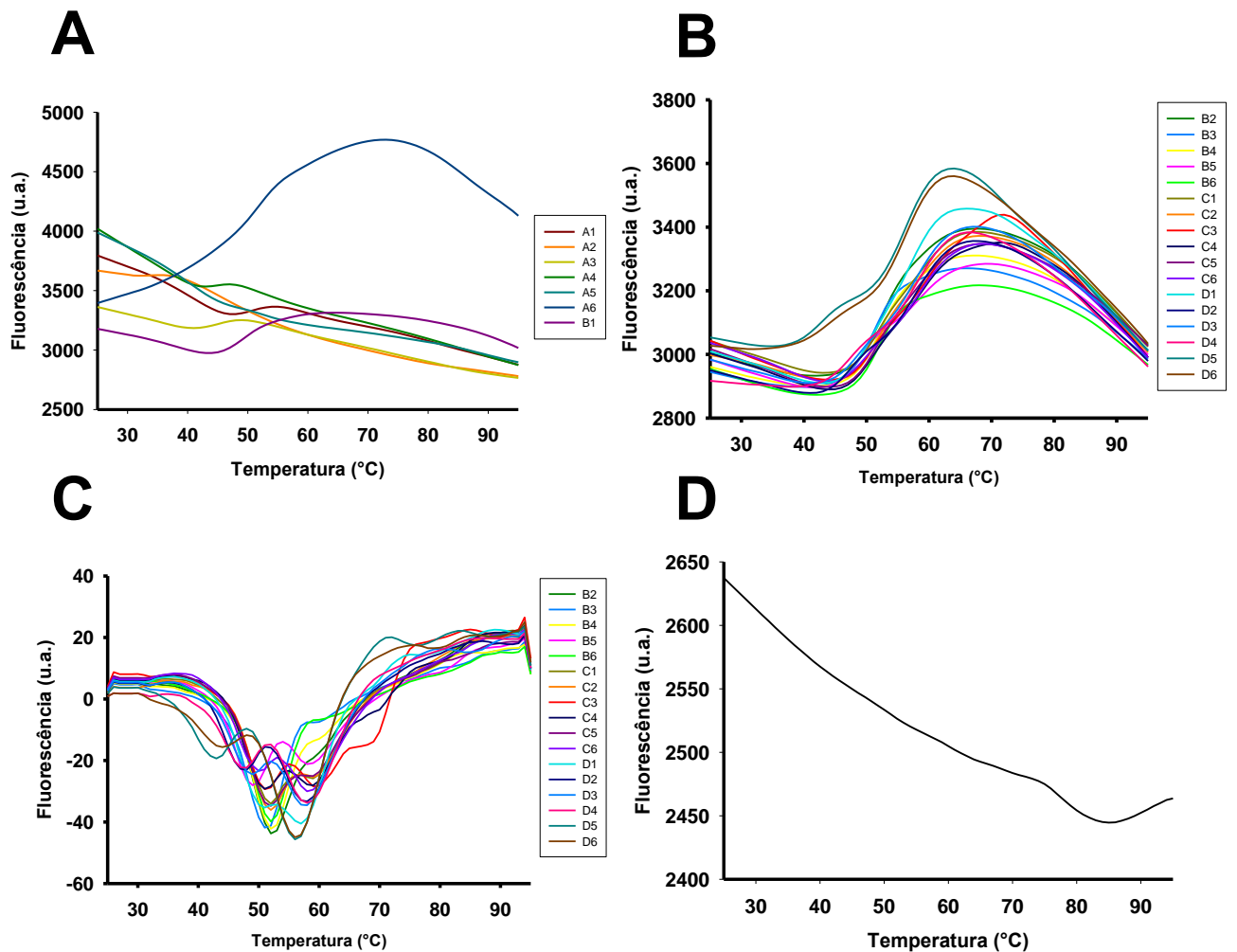


Figura 21: Análise da estabilidade térmica do complexo NSD3-PWWP:H3K36me3 na presença de 50 mM de NaCl por fluorimetria diferencial de varredura. A legenda indica os tampões utilizados em cada condição. (A) Curva de desnaturação térmica nos tampões com pH de 3 à 5,5. (B) Curvas de desnaturação térmica nos tampões com pH 6 à 10. (C) Primeira derivada da curva de desnaturação obtida em (B). (D) Curva de desnaturação térmica de H3K36me3 livre.

Os valores de  $T_m$  para cada uma das condições mostradas na Figura 21C estão listados na Tabela 6. Podemos verificar que a diferença entre  $T_{m1}$  e  $T_{m2}$  variam de 6 a 12 °C, indicando que as duas espécies em solução possuem estabilidades bem distintas. Partindo do fato de que a amostra inicial possui apenas NSD3-PWWP, H3K36me3 e o complexo, os valores de  $T_m$  observados no experimento devem ser referentes a essas espécies. Podemos verificar que o peptídeo livre apresentou alta intensidade de fluorescência desde o início da curva de temperatura (Figura 21D), resultado sustentado pelo fato de o peptídeo livre possui uma sequência linear, sem estrutura, com todos os resíduos presentes na cadeia expostos ao solvente, sendo

assim a ligação com o fluoróforo é imediata, apresentando curva de temperatura característica de proteína desenovelada. Diante disso, sugere-se que as transições observadas na Figura 21C sejam referentes à proteína livre e complexada, podendo-se comparar ao resultado do experimento das amostras livre e complexada em tampão [10 mM fosfato de sódio (pH 7,0), 175 mM NaCl] (Figura 20).

Por comparação entre os resultados, os valores relativos ao  $T_{m1}$  na Tabela 6 referem-se à temperatura de fusão de NSD3-PWWP livre e os valores de  $T_{m2}$  à temperatura de fusão de NSD3-PWWP complexada. A diferença entre os  $T_{ms}$  encontrados no experimento de DSF para a proteína livre (Figura 19A) e para a primeira transição da curva do complexo (Figura 21) pode ser atribuída a diferenças nas condições experimentais do aparelho, como intensidade de luz e etc., uma vez que os experimentos foram realizados em dias diferentes. Portanto, embora os valores absolutos de  $T_m$  sejam diferentes, os  $T_{ms}$  presentes na Tabela 6 foram somente comparados entre si e empregados na otimização da condição experimental do complexo. Assim, os valores de temperatura de fusão presentes na Tabela 6 indicam que o complexo NSD3-PWWP:H3K36me3 se encontra mais estável nas condições com pH entre 7 e 7,5, são elas: 20 mM acetato de amônio (pH 7,0), 20 mM MOPS (pH 7,0), 20 mM fosfato de sódio/potássio (pH 7,0), 20 mM HEPES (pH 7,5) e 20 mM Tris (pH 7,5), todos com 50 mM de NaCl.

A escolha de um tampão no qual a proteína é mais estável auxilia na determinação estrutural de NSD3-PWWP, uma vez que a obtenção dos espectros de RMN necessita que a amostra seja estável durante dias (~7 dias) dentro do espectrômetro a temperatura ambiente, sem que haja degradação ou agregação. A condição ideal para a obtenção de dados de RMN é tampão fosfato de sódio/potássio em baixa concentração de sal, uma vez que o fosfato não apresenta hidrogênio, portanto o tampão não interfere na coleção dos sinais da amostra, e altas concentrações de sal promovem diminuição da relação sinal/ruído. Sendo assim, a condição de 20 mM fosfato sódio, 50 mM de NaCl encontra-se dentro dessas especificações, podendo ser utilizada para a coleção dos dados de RMN do complexo NSD3-PWWP:H3K36me3.

<b>Tampão (20 mM)</b>	<b>pH</b>	<b>Tm1(°C)</b>	<b>Tm2 (°C)</b>
Fosfato de Sódio/Potássio	6,0	52	-
Bis-Tris	6,0	51	-
MES	6,2	52	-
ADA	6,5	49	58
Bis-Tris Propano	6,5	52	-
Acetato de Amônio	7,0	52	59
MOPS	7,0	52	59
Fosfato de Sódio/Potássio	7,0	51	59
HEPES	7,5	51	59
Tris	7,5	52	59
EPPS	8,0	50	58
Imidazol	8,0	51	57
Bicina	8,5	48	58
Tris	8,5	48	58
CHES	9,0	48	58
CHES	9,5	43	56
CAPS	10,0	44	56

Tabela 6: Temperatura de fusão obtida para NSD3-PWWP:H3K36me3 na presença de 50 mM NaCl em diferentes pHs.

## 5 Conclusões

- ✓ Através de técnicas cromatográficas foi possível obter o domínio PWWP da histonas metil-transferase NSD3 recombinante purificado.
- ✓ Os espectros de RMN apresentaram sinais característicos de proteína enovelada, como grande dispersão dos hidrogênios amídicos, sugerindo que a proteína recombinante encontra-se enovelada.
- ✓ O espectro de fluorescência intrínseca de NSD3-PWWP apresentou máximo em 344 nm quando excitado em 295 nm (excitação seletiva dos resíduos de triptofano), sugerindo a presença de triptofanos expostos ao solvente na estrutura da proteína.
- ✓ As isotermas de ligação derivadas dos dados de fluorescência exibiram duas inflexões distintas, sugerindo dois sítios de interação com afinidades diferentes.
- ✓ Os valores das constantes de dissociação aparente encontrados para a interação de NSD3-PWWP com nove peptídeos de histonas (derivados das histonas H3 e H4 contendo diferentes lisinas metiladas e graus de metilação) foram similares entre si e variaram na faixa de baixo  $\mu\text{M}$  (3,8 a 10,1  $\mu\text{M}$ ), sugerindo que NSD3-PWWP se liga a peptídeos de histona com alta afinidade, porém com pouca especificidade.
- ✓ Entretanto, a afinidade para H3K36me1 foi duas vezes maior do que as encontradas para H3K36me2 e H3K36me3, sugerindo que NSD3-PWWP consegue discernir entre as diferentes formas metiladas de H3K36, apresentando maior especificidade para a forma monometilada.
- ✓ O domínio PWWP de NSD3 se mostrou instável em condições de pH ácido (< 5,0).
- ✓ Os resultados de fluorimetria diferencial de varredura nas 72 diferentes condições revelaram que NSD3-PWWP apresenta maior estabilidade em pHs básicos (> 8,0) e alta concentração de sal (250 mM).
- ✓ NSD3-PWWP apresentou temperatura de fusão mais alta quando complexado ao peptídeo H3K36me3 (73 °C) em relação à proteína livre (60 °C) em 10 mM fosfato de sódio (pH 7,0), 175 mM NaCl, sugerindo que a interação com o peptídeo aumenta a estabilidade do domínio.
- ✓ O complexo NSD3-PWWP:H3K36me3 mostrou-se mais estável em tampões com pHs entre 7 e 7,5, em baixa concentração de sal, sendo esta uma condição propícia para aquisição de dados de RMN.

## 6 Referências

- ALBERT M.; HELIN K. Histone methyltransferases in cancer. **Semin Cell Dev Biol**, v. 21, n. 2, p. 209-220, 2010.
- ANDREOLI, F.; DEL RIO, A. Physicochemical modifications of histones and their impact on epigenomics. **Drug Discovery Today**, v. 19, n. 9 p. 1372-1379, 2014.
- ANGRAND P.O.; APIOU F.; STEWART A.F.; DUTRILLAUX B.; LOSSON R.; CHAMBON P. NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. **Genomics**, v. 74, n. 1, p. 79-88, 2001.
- BANNISTER, A.J.; KOUZARIDES, T. Regulation of chromatin by histone modifications. **Cell Res**, v. 21, n. 3, p. 381-395, 2011.
- BROWNELL, J. E.; ZHOU, J.; RANALLI, T.; KOBAYASHI, R.; EDMONDSON, D. G.; ROTH, S. Y.; ALLIS, C. D. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. **Cell**, v. 84, n. 6, p. 843–851, 1996.
- CHEN Y.; MCGEE J.; CHEN X.; DOMAN T.N.; GONG X.; ZHANG Y.; HAMM N.; MA X.; HIGGS R.E.; BHAGWAT S.V.; BUCHANAN S.; PENG S.B.; STASCHKE K.A.; YADAV V.; YUE Y.; KOUROS-MEHR H. Identification of druggable cancer driver genes amplified across TCGA datasets. **PLoS One**, v. 9, n.5, e98293, 2014.
- DE LA CRUZ, X.; LOIS, S.; SÁNCHEZ-MOLINA, S.; MARTÍNEZ-BALBÁS, M.A. Do protein motifs read the histone code? **BioEssays**, v. 27, n. 2, p. 164-175, 2015.
- DENNIS E. EPPS, THOMAS J. RAUB, VALERIA CAIOLFA, AUGUSTO CHIARI, MORENO ZAMAI. Determination of the Affinity of Drugs toward Serum Albumin by Measurement of the Quenching of the Intrinsic Tryptophan Fluorescence of the Protein. **J. Pharm**, v. 51, p. 41-48, 1999.
- DONG HAN; MENGXI HUANG; TING WANG; ZHIPING LI; YANYAN CHEN; CHAO LIU; ZENGJIE LEI; XIAOYUAN CHU. Lysine methylation of transcription factors in cancer. **Cell Death Dis**, v. 10, n. 4, 290, 2019.

EFTINK M.R. Fluorescence techniques for studying protein structure. **Methods Biochem Anal.**, v. 35, p. 127-205, 1991.

EGGER G., LIANG G., APARICIO A., JONES P.A. Epigenetics in human disease and prospects for epigenetic therapy. **Nature**, v. 429, n. 6990, p. 457-463, 2004.

EIDAH, J. O.; CROWE, B. L.; NORTH, J. A.; MCKEE, C. J.; SHKRIABAI, N.; FENG, L.; PLUMB, M.; GRAHAM, L.R.; GORELICK, J.R.; HESS, S.; POIRIER, G.M.; FOSTER, P.M.; KVARATSKHELIA, M. Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. **Nucleic Acids Research**, v. 41, n. 6, p. 3924–3936, 2013.

GELSI-BOYER V.; ORSETTI B.; CERVERA N.; FINETTI P.; SIRCOULOMB F.; ROUGÉ C.; LASORSA L.; LETESSIER A.; GINESTIER C.; MONVILLE F.; ESTEYRIÈS S.; ADÉLAÏDE J.; ESTERNI B.; HENRY C.; ETHIER S.P.; BIBEAU F.; MOZZICONACCI M.J.; CHARAFE-JAUFFRET E.; JACQUEMIER J.; BERTUCCI F.; BIRNBAUM D.; THEILLET C.; CHAFFANET M. Comprehensive profiling of 8p11-12 amplification in breast cancer. **Mol Cancer Res**, v. 3, n. 12, p. 655-667, 2005

GREER E.L.; SHI Y. Histone methylation: a dynamic mark in health, disease and inheritance. **Nat Rev Genet**, v. 13, n. 5, p. 343-357, 2012.

HAN X.; PIAO L.; ZHUANG Q.; YUAN X.; LIU Z.; HE X. The role of histone lysine methyltransferase NSD3 in cancer. **Oncotargets Ther**, v. 11, p. 3847-3852, 2018.

HAVERTY P.M.; FRIDLAND J.; LI L.; GETZ G.; BEROUKHIM R.; LOHR S.; WU T.D.; CAVET G.; ZHANG Z.; CHANT J. High-resolution genomic and expression analyses of copy number alterations in breast tumors. **Genes Chromosomes Cancer**, v. 47, n. 6, p. 530-542, 2008.

HE, C.; LI, F.; ZHANG, J.; WU, J.; SHI, Y. The Methyltransferase NSD3 Has Chromatin-binding Motifs, PHD5-C5HCH, That Are Distinct from Other NSD (Nuclear Receptor SET Domain) Family Members in Their Histone H3 Recognition. **Journal of Biological Chemistry**, v. 288, n. 7, p. 4692–4703, 2012.

HUANG, H.; LIN, S.; GARCIA, B. A.; ZHAO, Y. Quantitative Proteomic Analysis of Histone Modifications. **Chemical Reviews**, v. 115, n. 6, p. 2376–2418, 2015.

IZUKA, M.; SMITH, M.M. Functional consequences of histone modifications. **Current Opinion in Genetics & Development**, v.13, n. 2, p.154-160, 2003.

KANG D.; CHO HS.; TOYOKAWA G.; KOGURE M.; YAMANE Y.; IWAI Y.; HAYAMI S.; TSUNODA T.; FIELD H.I.; MATSUDA K.; NEAL D.E.; PONDER B.A.; MAEHARA Y.; NAKAMURA Y.; HAMAMOTO R. The histone methyltransferase Wolf-Hirschhorn syndrome candidate 1-like 1 (WHSC1L1) is involved in human carcinogenesis. **Genes Chromosomes Cancer**, v. 52, n. 2, p. 126-139, 2013.

LEITGEB S.; STRAGANZ G.D.; NIDETZKY B. Biochemical characterization and mutational analysis of the mononuclear non-haem Fe<sup>2+</sup> site in Dke1, a cupin-type dioxygenase from *Acinetobacter johnsonii*. **Biochem J**, v. 418, n. 2, p. 403-411, 2009.

LIU L.; KIMBALL S.; LIU H.; HOLOWATYJ A.; YANG Z-Q. Genetic alterations of histone lysine methyltransferases and their significance in breast cancer. **Oncotarget**, v. 6, n. 4, p. 2466-2482, 2015.

LOPEZ J., PERCHARDE M., COLEY H.M., WEBB A., CROOK T. The context and potential of epigenetics in oncology. **Br J Cancer**, v. 100, n. 4, p. 571-577, 2009.

MARGUERON, R.; TROJER, P.; REINBERG, D. The key to development: Interpreting the histone code? **Curr Opin Genet Dev**, v.15, n. 2, p.163-176, 2005.

MARTA PABIS, GRZEGORZ M. POPOWICZ, RALF STEHLE, DAVID FERNANDEZ-RAMOS, SAM ASAMI, LISA WARNER, SOFÁA M. GARCÍA-MAURINO, ANDREAS SCHLUNDT, MAROAA L. MARTÍNEZ-CHANTAR, IRENE D'IAZ-MORENO, MICHAEL SATTLER. HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRM3s. **Nucleic Acids Research**, v. 47, n. 2, p. 1011–1029, 2019.

MARTIN, B. J. E.; MCBURNEY, K. L.; MALTBY, V. E.; JENSEN, K. N.; BRIND'AMOUR, J.; HOWE, L. J. Histone H3K4 and H3K36 Methylation Independently Recruit the NuA3 Histone Acetyltransferase in *Saccharomyces cerevisiae*. **Genetics**, v. 205, n. 3, p. 1113-1123, 2017.



MELLOR, J. Dynamic nucleosomes and gene transcription. **Trends in Genetics**, v. 22, n. 6, p. 320-329, 2006.

MORISHITA M, DI LUCCIO E. Cancers and the NSD family of histone lysine methyltransferases. **Biochim Biophys Acta**, v. 1816, n. 2, p. 158-163, 2001.

NELSON, D. L.; COX, M.M. Princípios de bioquímica de Lehninger. 6. ed. Porto Alegre: Artmed, 2014.

PLACEK, B. J.; GLOSS, L. M. The N-Terminal Tails of the H2A–H2B Histones Affect Dimer Structure and Stability. **Biochemistry**, v. 41, n. 50, p. 14960–14968, 2002.

RONA, G. B.; ELEUTHERIO, E. C. A.; PINHEIRO, A. S. PWWP domains and their modes of sensing DNA and histone methylated lysines. **Biophys Rev**, v. 8, p. 63-74, 2016.

RONDELET, G.; DAL MASO, T.; WILLEMS, L.; WOUTERS, J. Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. **Journal of Structural Biology**, v. 194, n. 3, p. 357–367, 2016.

ROSATI R.; LA STARZA R.; VERONESE A.; AVENTIN A.; SCHWIENBACHER C.; VALLESPI T.; NEGRINI M.; MARTELLI M.F.; MECUCCI C. NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). **Blood**, v. 99, n. 10, p. 3857-3860, 2002.

SANTOS-ROSA, H.; CALDAS, C. Chromatin modifier enzymes, the histone code and cancer. **European Journal of Cancer**, v. 41, n.16, p. 2381-2402, 2005.

SENDŽIKAITĖ, G.; HANNA, C. W.; STEWART-MORGAN, K. R.; IVANOVA, E.; KELSEY, G. A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice. **Nature Communications**, v. 10, n. 1, 2019.

STEC I.; NAGL S.B.; VAN OMMEN G.J.; Dunnen Den J.T. The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? **FEBS Lett**, v. 473, n. 1, p. 1-5, 2000.

STRAHL, B. D.; ALLIS, C. D. The language of covalent histone modifications. **Nature**, v. 403, p. 41-45, 2000.

SUEOKA, T.; KOYAMA, K.; HAYASHI, G.; & OKAMOTO, A. Chemistry-driven Epigenetic Investigation of Histone and DNA Modifications. **The Chemical Record**, v.18, p. 1-19, 2018

TAKETANI T.; TAKI T.; NAKAMURA H.; TANIWAKI M.; MASUDA J.; HAYASHI Y. NUP98-NSD3 fusion gene in radiation-associated myelodysplastic syndrome with t(8;11)(p11;p15) and expression pattern of NSD family genes. **Cancer Genet Cytogenet**, v. 190, n. 2, p. 108-112, 2009.

TAUNTON, J.; HASSIG, C. A.; SCHREIBER, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. **Science**, v. 272, p. 408-411, 1996.

TSAO, H.S.; SPINELLA, S.A.; LEE, A.T.; ELMORE, D.E. Design of novel histone-derived antimicrobial peptides. **Peptides**, v. 30, n. 12, p. 2168-2173, 2009.

VERMEULEN, M.; EBERL, H. C.; MATARESE, F.; MARKS, H.; DENISSOV, S.; BUTTER, F.; LEE, K.K.; OLSEN, J.V.; HYMAN, A.A.; STUNNENBERG, H.G.; MANN, M. Quantitative Interaction Proteomics and Genome-wide Profiling of Epigenetic Histone Marks and Their Readers. **Cell**, v. 142, n. 6, p. 967–980, 2010.

VEZZOLI A.; BONADIES N.; ALLEN M.D.; FREUND S.M.; SANTIVERI C.M.; KVINLAUG B.T.; HUNTLY B.J.; GÖTTGENS B.; BYCROFT M. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. **Nat Struct Mol Biol**, v. 17, n. 5, p. 617-619, 2010.

VOUGIOUKLAKIS, T.; HAMAMOTO, R.; NAKAMURA, Y.; SALOURA, V. The NSD family of protein methyltransferases in human cancer. **Epigenomics**, v. 7, n. 5, p. 863-874, 2015.

WAGNER E.J.; CARPENTER P.B. Understanding the language of Lys36 methylation at histone H3. **Nat Rev Mol Cell Biol**, v. 13, n. 2, p. 115-126, 2012.

WANG, R.; GAO, J.; ZHANG, J.; ZHANG, X.; XU, C.; LIAO, S.; TU, X. Solution structure of TbTFIIS2-2 PWWP domain from *Trypanosoma brucei* and its binding to H4K17me3 and H3K32me3. **Biochemical Journal**, 2019.

WEAVER, T.; MORRISON, E.; MUSSELMAN, C. Reading More than Histones: The Prevalence of Nucleic Acid Binding among Reader Domains. **Molecules**, v. 23, n. 10, 2614, 2018.

WU H.; ZENG H.; LAM R.; TEMPEL W.; AMAYA M.F.; XU C.; DOMBROVSKI G.; QIU W.; WANG Y.; MIN J. Structural and histone binding ability characterizations of human PWWP domains. **PLoS One**, v. 6, n. 6, 6 e18919, 2011.

YANG Z.Q.; LIU G.; BOLLIG-FISCHER A.; GIROUX C.N.; ETHIER S.P. Transforming properties of 8p11-12 amplified genes in human breast cancer. **Cancer Res**, v. 70, n. 21, p. 8487-8497, 2010.