

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM PATOLOGIA MOLECULAR

VINÍCIUS ALVES FERNANDES

**ESTUDOS *IN SILICO* DE MOLÉCULAS LIGANTES DE
NUCLEOSSOMO SOBRE A CROMATINA**

BRASÍLIA - DF
2020

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM PATOLOGIA MOLECULAR

**ESTUDOS *IN SILICO* DE MOLÉCULAS LIGANTES DE
NUCLEOSSOMO SOBRE A CROMATINA**

VINÍCIUS ALVES FERNANDES

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Patologia Molecular da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Doutor em Patologia Molecular

Orientador (a) : Prof. Guilherme Martins Santos
Co-orientador (a): Prof. Dr. Werner Leopoldo Treptow

BRASÍLIA - DF
2020

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Ae Alves Fernandes, Vinícius
 Estudos in silico de moléculas ligantes de nucleossomo
 sobre a cromatina / Vinícius Alves Fernandes; orientador
 Guilherme Martins Santos; co-orientador Werner Leopoldo
 Treptow. -- Brasília, 2020.
 121 p.

 Tese (Doutorado - Doutorado em Patologia Molecular) --
 Universidade de Brasília, 2020.

 1. Cromatina. 2. Dinâmica Molecular. 3. Lipídios. 4.
 Peptídeos. I. Martins Santos, Guilherme, orient. II.
 Leopoldo Treptow, Werner, co-orient. III. Título.

Vinícius Alves Fernandes

Estudos *in silico* de moléculas ligantes de nucleossomo sobre a cromatina

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Patologia Molecular da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Doutor em Patologia Molecular .

Aprovada em de

Banca Examinadora

Prof. Dr. Guilherme Martins Santos (presidente)
Universidade de Brasília

Prof. Dr. Carlos André Ornelas Ricart
Universidade de Brasília

Prof. Dr. Leandro Martínez
Unicamp - SP

Prof. Dr. Paulo Sérgio Lopes de Oliveira
CNPEM/LNBio - SP

Prof.^a Dr.^a Angélica Amorim Amato (suplente)
Universidade de Brasília

AGRADECIMENTOS

Inicialmente, gostaria de agradecer à minha família, pelo suporte e ajuda na minha formação pessoal. Em especial, a minha mãe Neyde, pelo suporte emocional e afetivo que me deram uma base sólida para seguir em busca dos meus objetivos.

Agradeço ao meu orientador, Guilherme Santos, por me guiar durante esses anos iniciais de caminhada científica. A sua motivação e entusiasmo demonstrados em cada resultado são extraordinariamente revigoradores do espírito científico. Agradeço também ao meu co-orientador prof. Werner Treptow, pelos conhecimentos transmitidos e pela paciência demonstrada em todos os questionamentos. Agradeço também o seu esforço para que eu tivesse uma formação complementar em biofísica. Não são todos que tem a sorte de ter dois orientadores excepcionais, agradeço a eles pelos ensinamentos e por guiarem meus passos nessa jornada.

Agradeço aos meus colegas do Laboratório de Biologia Teórica e Computacional (LBTC) Leonardo, Miguel, Caio, Fiorote, Camila, Letícia, Alessandra, Mônica, Natália e João pela prontidão em auxiliar em todos os momentos do desenvolvimento desse trabalho e pelos momentos de descontração do dia-dia.

Agradeço ao pessoal do Farmol, Kaian, Isabel, Natália também por momentos de descontração e apoio no desenvolvimento do projeto.

Agradeço à minha noiva, Clariane Ramos, que sempre me incentivou a trilhar o caminho do crescimento pessoal e me deu suporte emocional nos momentos de dúvida e anseio da vida acadêmica. Agradeço também pela compreensão nos momentos que tive que me ausentar e pelo imenso carinho demonstrado.

Agradeço à banca examinadora, que disponibilizou seu tempo para tornar este trabalho ainda melhor.

Por fim, agradeço também ao CNPq e a FAP-DF pelo apoio financeiro.

RESUMO

Mudanças estruturais na cromatina regulam a expressão gênica e definem os resultados fenotípicos. Tais mudanças são ocasionadas por uma variedade de pequenas moléculas e proteínas que, em sua maioria, atuam em nível nucleossomal. Embora toda a superfície do nucleossomo seja alvo de diferentes proteínas ligantes de nucleossomo (NBPs), uma região conhecida como *patch* ácido é o alvo preferencial dessas moléculas. Dessa forma, experimentos computacionais foram realizados a fim de compreender o impacto ao nível atômico que vários peptídeos baseados em NBPs exercem sob o nucleossomo. Os resultados indicam que tais peptídeos podem interagir com o nucleossomo e desencadear mudanças estruturais. Foi observado efeitos sobre o perfil de hidratação do nucleossomo, assim como alterações nas flutuações dos resíduos de aminoácidos e nucleotídeos do nucleossomo. Além do estudo com os peptídeos, dados biofísicos indicaram que o colesterol interage com o nucleossomo através de múltiplos sítios de ligação e afeta a estrutura da cromatina *in vitro*. Atualmente há pouca consciência de que os lipídios possam ser armazenados na cromatina e influenciar seu estado. Aqui, com foco em lipídios que se ligam a receptores nucleares, fatores de transcrição clinicamente relevantes, foram exploradas as interações do nucleossomo com hormônios esteróides, ácidos biliares e ácidos graxos. Os resultados preliminares do encaixe molecular sugerem que outras moléculas lipídicas podem também ter impacto na estrutura da cromatina através da ligação a locais de colesterol comuns no nucleossomo.

Palavras-chave: cromatina, nucleossomo, lipídios, NBP, peptídeos, dinâmica molecular

ABSTRACT

Structural changes in chromatin regulate gene expression and define phenotype outcomes. Such changes are caused by a variety of small molecules and proteins that act, mostly, at the nucleosomal level. Although the entire surface of the nucleosome is target for different nucleosome binding proteins (NBPs), an interface, known as the acidic patch, is the NBP's preferred target. Herein, computational experiments were performed in order to understand the impact of various peptides based on NBPs on the nucleosome, at the atomistic level. Results indicate those peptides interact with the nucleosome and trigger structural changes. Effects on the nucleosome hydration profile were observed, as well as fluctuations changes in the amino acid residues and nucleotide. In addition to the study with peptides, biophysical data indicates that cholesterol interacts with the nucleosome through multiple binding sites and affects the structure of chromatin in vitro. Currently, there is little awareness that lipids can be stored in chromatin and influence its state. Here, focusing on lipids that bind to nuclear receptors, clinically relevant transcription factors, nucleosome interactions with steroid hormones, bile acids and fatty acids were explored. Preliminary docking results suggest that other lipid molecules may also impact the chromatin structure by binding to common cholesterol sites on the nucleosome surface.

Keywords: chromatin, nucleosome, lipids, NBP, peptides, molecular dynamics

LISTA DE FIGURAS

Figura 1: Ilustração da compactação do DNA.....	16
Figura 2: Representação do NCP.....	17
Figura 3: Regiões particulares do nucleossomo.....	19
Figura 4: <i>Patch</i> ácido.....	21
Figura 5: Ilustração dos potenciais entre átomos ligados empregados na DM..	27
Figura 6: Esquema que ilustra as condições periódicas de contorno (PBC).....	28
Figura 7: NBPs que tem como alvo a superfície proteica do nucleossomo.....	34
Figura 8: Sítio de ancoragem de argininas no <i>patch</i> ácido.....	35
Figura 9: Modelo de remodelamento da cromatina mediante a presença de NBPs.....	36
Figura 10: Etapas do estudo da interação de NBPePs com o nucleossomo (TELES et al., 2020).....	38
Figura 11: Desenho do peptídeo GMIP1.....	39
Figura 12: Aminoácidos retirados das caudas das histonas.....	43
Figura 13: RMSD GMIP1 e CtrlNeg.....	49
Figura 14: Clusters conformacionais.....	50
Figura 15: <i>Docking</i> da cauda da H4 na conformação alfa-hélice.....	51
Figura 16: <i>Docking</i> GMIP1 e CtrlNeg.....	52
Figura 17: RMSD de todas as trajetórias produzidas.....	53
Figura 18: Distância dependente do tempo entre os centros de massa de cada NBPeP e do <i>patch</i> ácido.....	57

Figura 19: Distância dependente do tempo entre regiões específicas dos NBPeps e o <i>patch</i> ácido.....	58
Figura 20: Alteração na hidratação do <i>patch</i> ácido.....	59
Figura 21: Perfil de RMSF do domínio globular do nucleossomo.....	60
Figura 22: RMSF do carbono alfa dos aminoácidos que formam o <i>patch</i> ácido.	62
Figura 23: Mudanças proporcionais média no RMSF do nucleossomo induzidas pelos NBPeps.....	63
Figura 24: Mudanças proporcionais no RMSF do nucleossomo induzidas pelos NBPeps.....	64
Figura 25: RMSD dos átomos da cadeia principal das histonas completas (preto) e sem as caudas (vermelho).....	66
Figura 26: RMSF do <i>patch</i> ácido do nucleossomo simulado com (preto) e sem as caudas (vermelho).....	66
Figura 27: Perfil do RMSD de diversas seleções de átomos do nucleossomo....	68
Figura 28: Perfil de distância entre o GMIP1 e o nucleossomo.....	70
Figura 29: Estrutura química dos lipídios testados nesse projeto.....	78
Figura 30: Possíveis locais de interação de algumas moléculas lipídicas com o nucleossomo.....	79
Figura 31: Número de moléculas de colesterol ligados aos nucleossomos.....	85

LISTA DE TABELAS

Tabela 1. Detalhes dos sistemas estudados.....	44
Tabela 2. Aminoácidos do nucleossomo formadores das cavidades de cada NB Pep.....	55
Tabela 3. Aminoácidos dos NB Peps responsáveis pela interação com o nucleossomo.....	55
Tabela 4. Receptores nucleares e seus ligantes lipídicos.....	73
Tabela 5. Comparação dos sítios de ligação do colesterol com os demais lipídios.....	80

SIGLAS

CENP-C: *Centrome Protein C*

CENP-N: *Centrome Protein N*

CHARMM: *Chemistry at Harvard Macromolecular Mechanics.*

DM: *Dinâmica Molecular*

GMIP1: *Genetic Modified Inducible Peptide 1*

HMGN2: *High Mobility Group Nucleosomal 2*

HO: *Histone Octamer*

KSHV: *Kaposi's Sarcoma Herpesvirus*

LANA: *Latency Associated Nuclear Antigen*

NAMD: *Nanoscale Molecular Dynamics program.*

NBP: *Nucleosome Binding Proteins.*

NBPeps: *Nucleosome Binding Peptides.*

NCP: *Nucleosome Core Particle*

NMR: *Nuclear magnetic resonance*

NPT: *ensemble com número de partículas, pressão e temperatura constantes.*

NVE: *ensemble com número de partículas, volume e energia constantes.*

NVT: *ensemble com número de partículas, volume e temperatura constantes.*

PDB: *Protein Data Bank*

PTM: *Post-translational Modifications*

RCC1: *Regulator of Chromosome Condensation 1*

RMSD: *Root Mean Squared Deviation.*

RMSF: *Root Mean Squared Fluctuation.*

SIR3: *Silent Information Regulator 3*

VdW: *van der Waals.*

SUMÁRIO

AGRADECIMENTOS	5
RESUMO	6
ABSTRACT	7
LISTA DE FIGURAS	8
LISTA DE TABELAS	10
SIGLAS	11
SUMÁRIO	12
I . CAPÍTULO I. INTRODUÇÃO E MÉTODOS	14
I.1 INTRODUÇÃO.....	14
I.1.1 Cromatina.....	14
I.1.2 Nucleossomo.....	16
I.1.3 Moduladores da arquitetura da cromatina.....	21
I.2 OBJETIVOS.....	24
I.2.1 Objetivo geral.....	24
I.2.1 Objetivos específicos.....	24
I.3 TEORIAS E MÉTODOS.....	25
I.3.1 Simulações de Dinâmica Molecular.....	25
I.3.1.1 Campo de força.....	26
I.3.1.2 Condições Periódicas de Contorno.....	28
I.3.1.3 Equações de movimento.....	29
I.3.1.4 Ensemble NPT e <i>annealing</i>	30
I.3.2 RMSD e RMSF.....	31
I.3.3 <i>Docking</i> molecular.....	32
II . CAPÍTULO II. PEPTÍDEOS LIGANTES DE NUCLEOSSOMO (NBPeps)	34
II.1 INTRODUÇÃO.....	34
II.1.1 Proteínas que se ligam ao nucleossomo (NBPs).....	34
II.1.2 Nucleossomo como alvo terapêutico.....	36
II.2 OBJETIVOS.....	40
II.2.1 Objetivo geral.....	40
II.2.1 Objetivos específicos.....	40
II.3 PROTOCOLOS.....	41
II.3.1 Protocolo de simulações de DM.....	41
II.3.1.1 Controle (nucleossomo livre em solução).....	42

II.3.1.2 NBPeps.....	43
II.3.2 Protocolo de Clusterização.....	45
II.3.3 Protocolo de <i>Docking</i> Molecular.....	45
II.3.3.1 H4-hx.....	46
II.3.3.2 GMIP1 e CtrlNeg.....	47
II.4 RESULTADOS.....	48
II.4.1 Busca do modo de interação com o nucleossomo dos NBPeps CtrlNeg, GMIP1 e H4-hx.....	48
II.4.1.1 <i>Clusters</i> conformacionais.....	49
II.4.1.2 <i>Docking</i> dos NBPeps ao nucleossomo.....	50
II.4.2 Estabilidade das simulações.....	52
II.4.3 Mapa de contatos.....	53
II.4.4 Dinâmica dos NBPeps no sítio.....	56
II.4.5 Hidratação do <i>patch</i> ácido.....	58
II.4.6 Impacto estrutural.....	59
II.4.7 Considerações gerais e discussão.....	65
II.4.8 Limitações e perspectivas.....	70
III . CAPÍTULO III. LIPÍDIOS.....	72
III.1 INTRODUÇÃO.....	72
III.1.1 Lipídios no núcleo.....	72
III.2 OBJETIVOS.....	75
III.2.1 Objetivo geral.....	75
III.2.1 Objetivos específicos.....	75
III.3 PROTOCOLOS.....	76
III.3.1 Protocolo de <i>docking</i> molecular.....	76
III.4 RESULTADOS.....	78
III.4.1 Interface de interação dos lipídios no nucleossomo.....	78
III.4.2 Comparação de locais de interação entre lipídios.....	80
III.4.3 Considerações gerais e discussão.....	81
III.4.4 Limitações e Perspectivas.....	84
IV . CONCLUSÕES.....	87
V . REFERÊNCIAS.....	88
ANEXOS.....	99

I . CAPÍTULO I. INTRODUÇÃO E MÉTODOS

O presente documento está subdividido em quatro capítulos, de forma que o capítulo I fornece uma introdução geral dos temas abordados e das teorias empregadas. O capítulo II está voltado para os estudos dos peptídeos ligantes de nucleossomo (NBPeps). No capítulo III é investigada a ligação de lipídios com o nucleossomo e discutido seu papel na cromatina. O capítulo IV apresenta as conclusões oriundas deste trabalho.

I.1 INTRODUÇÃO

I.1.1 Cromatina

O genoma de todos os seres eucariotos está organizado em conjunto com proteínas formando um complexo polimérico chamado cromatina. A cromatina desempenha um papel estrutural fundamental ao promover a compactação do genoma no núcleo eucariótico. No *Homo sapiens*, por exemplo, cada célula armazena em seu núcleo por volta de 6 bilhões de pares de bases, que se esticado, seriam aproximadamente 2 metros de DNA. Dentre as proteínas presentes na cromatina, as histonas são essenciais para a compactação do genoma, pois apresentam alta afinidade com o DNA. A interação das histonas com pequenos segmentos de DNA permite a formação da unidade repetitiva básica da cromatina, o nucleossomo (KORNBERG, 1977; OLINS; OLINS, 1974).

A forte interação entre DNA e histonas presente no nucleossomo restringe o acesso de várias proteínas ao DNA, regulando negativamente diversas funções celulares dependentes de tais interações, como transcrição gênica, replicação e reparo do DNA. É evidente que tais funções celulares deve ocorrer, para tanto, diversos processos que alteram a organização de nucleossomos podem ser ativados permitindo a modificação da cromatina local e o acesso de proteínas a elementos regulatórios (Revisado em: LAI; PUGH, 2017). Vários fatores podem desencadear

alterações na compactação da cromatina, como a presença de cátions mono e divalentes (KOROLEV et al., 2010), modificações pós-traducionais das histonas como acetilação e desacetilação (Revisado em: (TOLSMA; HANSEN, 2019), incorporação de histonas não canônicas ao nucleossomo (Revisado em: TALBERT; HENIKOFF, 2017) e interação com proteínas remodeladoras da cromatina (Revisado em: ZHOU; GAULLIER; LUGER, 2019).

A cromatina é uma estrutura altamente dinâmica e complexa que está em processo constante de compactação e descompactação. Entretanto, podemos dividi-la, didaticamente, em dois estados estruturais extremos: a heterocromatina e a eucromatina (HSU, 1962). A heterocromatina representa um estado compactado geralmente associado a baixa atividade transcricional e ao silenciamento gênico. No estado denominado eucromatina, a cromatina apresenta estrutura mais relaxada e conseqüentemente o DNA está mais acessível à maquinaria transcricional basal (Revisado em: SOLOVEI; THANISCH; FEODOROVA, 2016).

Pode-se classificar os estados organizacionais da cromatina também a partir de níveis sucessivos de condensação do DNA e interações internucleossomais (Figura 1). Iniciando por arranjos de nucleossomos de comprimento variável que se assemelham a um “colar de contas” também conhecido como fibra de 10 nm, devido ao diâmetro do nucleossomo. Nesse estado os nucleossomos são conectados por segmentos de DNA de comprimentos múltiplos denominados *linker* DNA. Variações no comprimento do *linker* DNA interferem na estrutura e estabilidade das fibras de cromatina (COLLEPARDO-GUEVARA; SCHLICK, 2014; HUANG et al., 2018). Mediante a presença da *linker* histona, H1 em humanos ou H5 em eritrócitos de aves, e interações internucleossomais de curto alcance, forma-se uma fibra mais compactada com 30 nm de diâmetro (ROBINSON et al., 2006). Para formar esse nível mais compactado da cromatina a *linker* histona interage com o *linker* DNA e com o nucleossomo. Mais de um modo de ligação entre a *linker* histona e o nucleossomo foi observado, o que resulta em alterações na estrutura e na eficiência de compactação da fibra de cromatina (ZHOU et al., 2015). Entretanto, a dobragem específica da fibra de 30 nm em si é bastante controversa, pois apesar de a sua ocorrência ser suportada por vários autores *in vitro*, ainda não foi possível visualizá-

las no contexto celular (FUSSNER; CHING; BAZETT-JONES, 2011; LI; ZHU, 2015; ZHOU; BAI, 2019). As interações entre os nucleossomos e entre fibras segue aumentando a densidade de DNA empacotado (Figura 1).

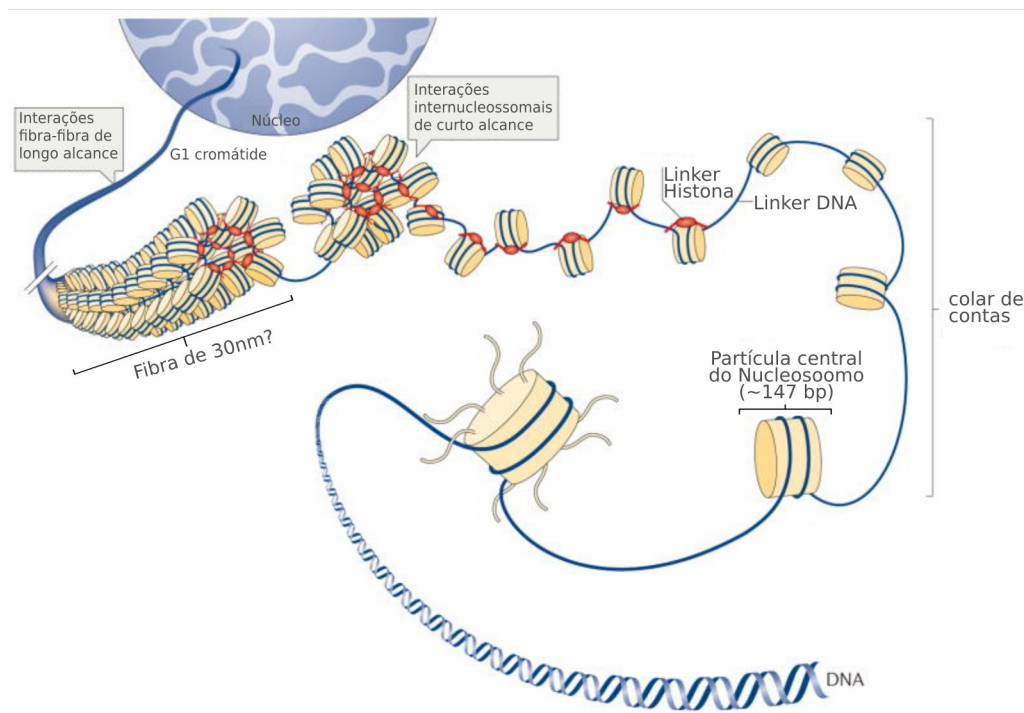


Figura 1: Ilustração da compactação do DNA.

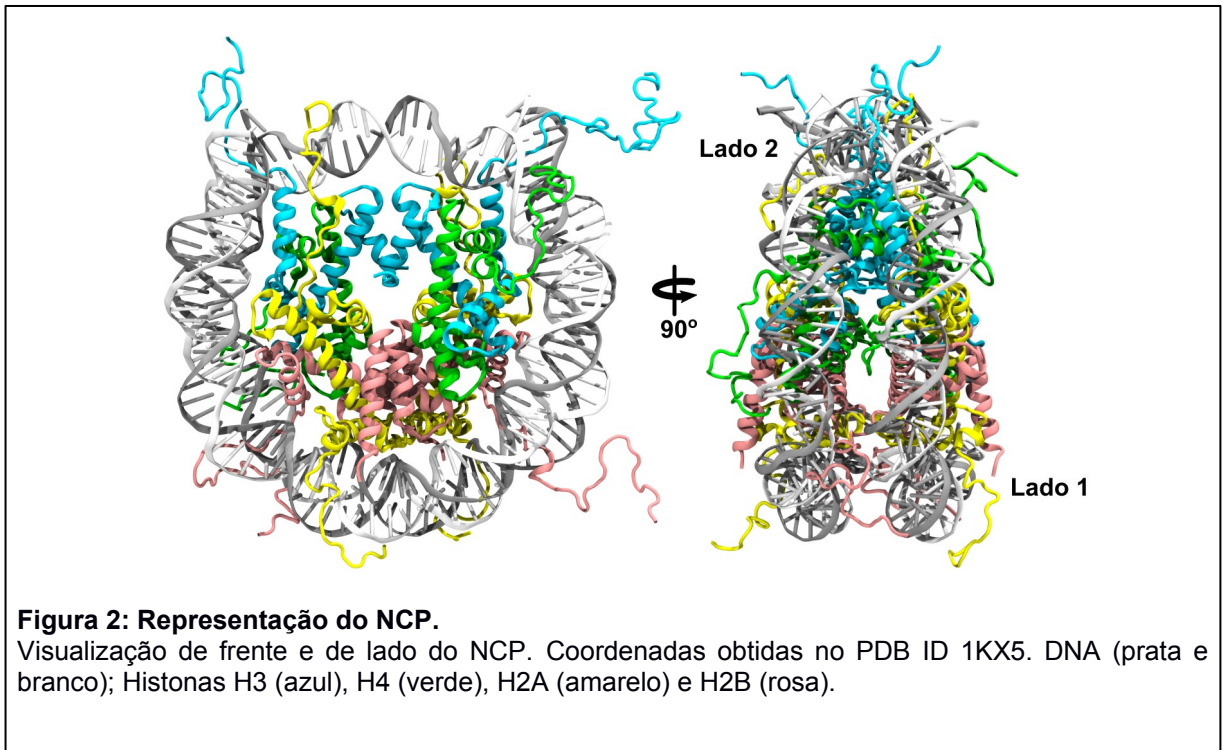
Representação dos níveis hierárquicos de compactação do DNA. Adaptado de (FYODOROV et al., 2018).

I.1.2 Nucleossomo

Os nucleossomos são os blocos primários de formação da cromatina, organizando aproximadamente 147 pares de bases de DNA em torno de um complexo octamérico de histonas (HO). O nucleossomo canônico é formado por quatro tipos diferentes de histonas (H2A, H2B, H3 e H4), com duas cópias de cada (FINCH et al., 1977).

Após aproximadamente 20 anos desde os primeiros trabalhos estruturais sobre o nucleossomo, em 1997 a primeira estrutura cristalográfica de alta resolução (2,8 Å) foi publicada (LUGER et al., 1997). A estrutura resolvida da partícula central

do nucleossomo (NCP do inglês *Nucleosome Core Particle*), nome dado ao nucleossomo sem a presença de *linker* DNA, revela a simetria bilateral do nucleossomo com eixo sobre a *dyad*, região de entrada e saída do DNA. Dessa forma, se o nucleossomo for rotacionado em 180° sobre esse eixo, o resultado será uma visualização praticamente idêntica à anterior (Figura 2).



Uma característica marcante das histonas é sua composição majoritária de aminoácidos básicos, o que permite uma associação entre o octâmero de histonas e o DNA forte o suficiente para protegê-lo de digestões enzimáticas por endonucleases, como a nuclease microcócica (VAN HOLDE, 1989). Na ausência de DNA são formados oligômeros de histonas estáveis, como os dímeros das histonas H2A e H2B e o tetrâmero, formado entre as histonas H3 e H4. Durante o processo de formação do nucleossomo, o tetrâmero $(H3)_2-(H4)_2$ e ~ 120 pares de base DNA organizam-se em um complexo estável. Posteriormente, os dímeros (H2A-H2B) se ligam a cada um dos lados do tetrâmero resultando em ~ 147 pares de base de DNA enrolado no HO (KARANTZA et al., 1995; KARANTZA; FREIRE; MOUDRIANAKIS,

1996; SAUER et al., 2018). A neutralização das cargas das histonas proporcionada pelo DNA é essencial para que o octâmero se mantenha estável em solução fisiológica. Além disso, foi observado que moléculas de água contribuem para a estabilidade da ligação entre DNA e histonas, desempenhando um papel importante para a estabilidade do nucleossomo (DAVEY et al., 2002).

O octâmero de histonas (HO) pode ser dividido estruturalmente em dois domínios distintos. Abrangendo ~ 75% da massa total das histonas, o domínio globular central é responsável pelas interações entre histonas no HO. Este domínio forma um arcabouço octamérico, constituído majoritariamente por alfa-hélices, no qual o DNA se enrola (LUGER et al., 1997). Os 25% de massa das histonas restantes formam as extensões N-terminais e C-terminais, também conhecido como caudas de histonas (Figura 3a) (ARENTS et al., 1991).

O domínio globular central é marcado estruturalmente pela presença do motivo de dobra da histona. Cada motivo de dobra, constituído por três alfa-hélices (α_1 , α_2 , α_3) conectadas por duas alças (L1 e L2), é unido no sentido antiparalelo ao motivo complementar de outra histona. De forma que heterodímeros são formados com a interação da H3 com a H4 e a H2A com a H2B. A interação entre os motivos de dobra das histonas é mantida por várias ligações de hidrogênio e interações hidrofóbicas. Como mencionado anteriormente, os heterodímeros das histonas H3 e H4 formam um tetrâmero que tem como elo principal uma interface conhecida como feixe de quatro hélices, “4-helix bundle” (Figura 3c-d). Essa interface constituída pelas hélices α_2 e α_3 de cada motivo de dobra, é estabilizada por interações hidrofóbicas e é fundamental para a estabilização estrutural do nucleossomo. A associação dos dímeros H2A:H2B em cada um dos lados do tetrâmero também é mantida pelo “4-helix bundle” (LUGER et al., 1997).

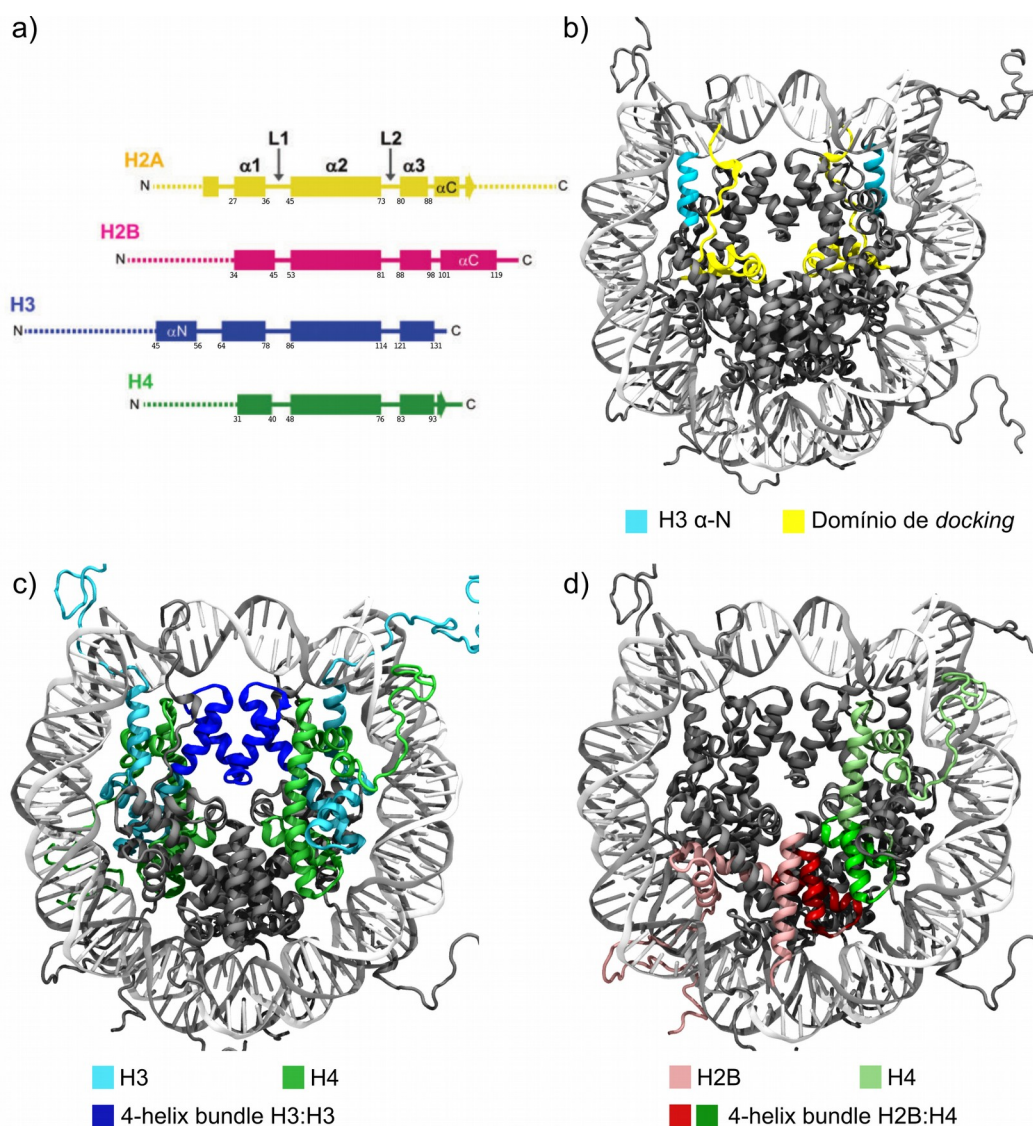


Figura 3: Regiões particulares do nucleossomo.

a) Esquema mostrando os domínios estruturais de cada histona. As hélices $\alpha 1$ a $\alpha 3$ formam o domínio global central. Esse esquema foi adaptado de (KOYAMA; KURUMIZAKA, 2018). b) Extensões do domínio global central da H3, hélice N-terminal (α -N), e da H2A, domínio de *docking*. c-d) NCP destacando os 4-helix bundles (H3:H3 em azul escuro, H2B vermelho escuro e H4 verde escuro).

As histonas H3, H2B e H2A apresentam regiões com estruturas definidas adicionais que contribuem para a estabilidade do nucleossomo. A H2B apresenta hélice C-terminal que define o limite longitudinal do octâmero de histonas. Na H3, uma hélice N-terminal (α -N) adicional auxilia o empacotamento do DNA na região da *dyad* (Figura 3b). Essa hélice é guiada para a *dyad* pela estrutura adicional C-terminal da H2A, conhecida como domínio de *docking*, aminoácidos 80 a 119 da

H2A (Figura 3b) (LUGER et al., 1997).

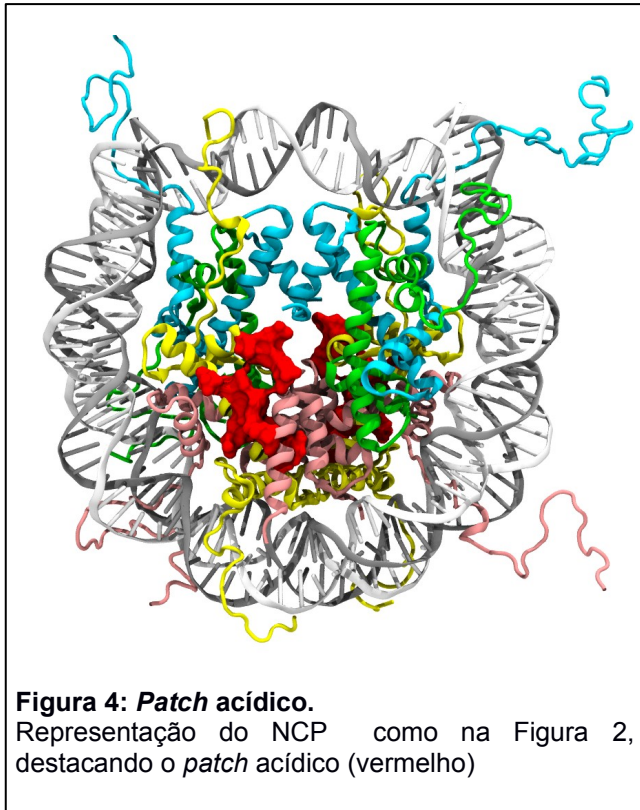
Alguns trabalhos sugerem que o enfraquecimento das interações no domínio do *docking* pode reduzir a estabilidade do nucleossomo e a formação de estados mais compactados da cromatina. A deleção do domínio de *docking* altera o ângulo da região de entrada e saída do DNA (*dyad*) e promove o desenrolamento do DNA (SHUKLA et al., 2010). Mesmo a presença do domínio de *docking* parcial está associado com o desempacotamento do DNA no nucleossomo (PENG et al., 2020). Também foi observado que a ausência do *docking domain* ou sua deleção parcial, atenua a ligação da *linker* histona (H1) ao nucleossomo (SHUKLA et al., 2010; VOGLER et al., 2010) o que provavelmente terá impacto na formação das fibras de cromatina.

Todas as histonas do HO (H2A, H2B, H3 e H4) apresentam regiões N-terminais que se estendem do domínio central e a histona H2A na região C-terminal também. Assim cada nucleossomo dispõe de dez regiões extensoras conhecidas como caudas das histonas. As caudas das histonas apresentam alta flexibilidade e geralmente não formam estruturas estáveis, exceto em casos especiais, como em presença do DNA ou interações com outras macromoléculas (CUTTER; HAYES, 2015).

As caudas das histonas são formadas, em sua maioria, por resíduos de arginina e lisina que lhes conferem carga positiva, que por conseguinte promovem atração pelo grupo fosfato da cadeia principal do DNA. As caudas são regiões importantes para a arquitetura da cromatina, visto que já foram observadas sua associação com a organização local do DNA (DAVEY et al., 2002; LUGER et al., 1997) e são alvo de alterações pós traducionais que promovem mudanças na arquitetura da cromatina (Revisado em: LAWRENCE; DAUJAT; SCHNEIDER, 2016). As caudas também estão relacionadas com a acessibilidade de fatores de transcrição, pois sua presença evita o desenrolamento do DNA (LEE et al., 1993; VETTESE-DADEY et al., 1994; YANG et al., 2005). Além disso, as caudas das histonas contribuem para a estabilidade do nucleossomo. A remoção das caudas da H3 e H2A impactaram estruturalmente a última hélice da H2A (BISWAS et al., 2011).

Entretanto, quando todas as caudas foram retiradas ao mesmo tempo, não foram identificadas alterações estruturais no núcleo central do nucleossomo (SHAYTAN et al., 2016).

Apesar do nucleossomo ser majoritariamente constituído por aminoácidos básicos, em sua superfície destaca-se uma região de oito aminoácidos com características opostas (E56, E61, E64, D90, E91, E92 da H2A e E102, E110 da



H2B). Esses aminoácidos formam uma região bem delimitada e carregada negativamente conhecida como *patch* ácido (LUGER et al., 1997). O *patch* ácido forma um interface na superfície do nucleossomo que é alvo de diversas proteínas. Já foi observada a interação dessa interface com a cauda da histona H4 do nucleossomo adjacente (DAVEY et al., 2002; LUGER et al., 1997) e com diversas proteínas ligantes de nucleossomo (NBP *Nucleosome Binding Proteins*) (Revisado em: (ZHOU; GAULLIER; LUGER, 2019).

I.1.3 Moduladores da arquitetura da cromatina

No ambiente nuclear, a cromatina está exposta a diversas moléculas e proteínas que podem modificar sua organização e compactação. Tais alterações estão intimamente relacionadas a processos celulares cruciais, como o correto funcionamento do complexo maquinário de transcrição, replicação e reparo do DNA (BARTMAN; BLOBEL, 2015; LUGER; DECHASSA; TREMETHICK, 2012). O magnésio (Mg^{2+}) é um exemplo, mesmo em uma concentração relativamente pequena (1 mM) é o suficiente para promover a condensação máxima das fibras de

cromatina (KOROLEV et al., 2010).

Além das histonas que compõem o HO, há também uma quinta histona que desempenha papel-chave na formação de estados mais compactados da cromatina, a *linker* histona. Em 2008, Robinson e colaboradores demonstraram que a completa descompactação da cromatina (formação da eucromatina) só é atingida quando ocorre também a evicção da *linker* histona (ROBINSON et al., 2008). O modelo mais aceito atualmente é o de que a ligação dessa histona ocorre na região de entrada e saída do DNA (região dyad) (SONG et al., 2014; ZHOU et al., 2015), porém, algumas pequenas diferenças foram observadas em relação à orientação da *linker* histona nessa região (BEDNAR et al., 2017).

A cauda da H4 é outro fator de extrema importância para a formação dos estados mais compactados da cromatina. Nos primeiros trabalhos estruturais de alta resolução do nucleossomo (DAVEY et al., 2002; LUGER et al., 1997), foi observado a interação de alguns resíduos (16 a 25) da cauda da H4 com o *patch* ácido do nucleossomo adjacente. Dorigo e colaboradores analisaram a compactação da cromatina em diversos arranjos de nucleossomos com deleção individual das caudas das histonas e deleções combinadas. Eles verificaram que a cauda da H4, principalmente a região entre os resíduos 14-19, além de ser necessária, é a cauda mais relevante para a formação da fibra de 30 nm *in vitro* (DORIGO et al., 2003).

Trabalhos posteriores investigaram a importância da lisina 16 da cauda da H4 para a interação entre nucleossomos. Foi observado que, após a neutralização de cargas da lisina 16 (acetilação), a formação *in vitro* da fibra de 30 nm foi incompleta (ALLAHVERDI et al., 2011; ROBINSON et al., 2008). Em 2017, foi investigada a interação da cauda da H4, com e sem a lisina 16 acetilada, por meio de simulações computacionais e análise do perfil energético. Além de observarem a ancoragem da K16 no *patch* ácido, este trabalho sugere que a acetilação é suficiente para romper a ligação da cauda da H4 com o *patch* ácido, reduzindo a interação internucleossomal (ZHANG; ERLER; LANGOWSKI, 2017). Yang e Arya, após extensivas simulações de DM, *docking* e cálculos de afinidade chegaram a conclusão que o peptídeo da cauda da H4 (resíduos 16 a 23) isolado tem forte

tendência a formar alfa-hélice e, nessa conformação, exibe alta afinidade com o *patch* ácido (YANG; ARYA, 2011). Ainda mais, foi mostrado que a cauda da H4 é importante para o *stacking* de nucleossomos tanto mediante interação com o patch ácido (CHEN et al., 2017) como pela interação com o DNA (SAURABH et al., 2016).

I.2 OBJETIVOS

I.2.1 Objetivo geral

Investigar as características da interação de moléculas ligantes de nucleossomo (lipídios e peptídeos), avaliando os efeitos dessa interação em nível atômico.

I.2.1 Objetivos específicos

a) NBPePs– Capítulo 2

- Construção dos sistemas atomísticos de cada peptídeo com o nucleossomo;
- Simular por DM os complexos;
- Observar como a ligação dos NBPePs interfere na estrutura do nucleossomo;
- Correlacionar os achados computacionais com os dados biofísicos e biológicos

b) Lipídios e cromatina – Capítulo 3

- Buscar os sítios de ligação de vários lipídios no nucleossomo eucariótico;
- Comparar os resultados obtidos com a análise realizada com colesterol;
- Analisar os possíveis efeitos da presença de lipídios na cromatina.

I.3 TEORIAS E MÉTODOS

Diante dos objetivos propostos e o nível de detalhamento almejado para esse estudo, foram utilizadas metodologias capazes de tratar interações moleculares com resolução atômica. Assim, métodos teóricos e computacionais foram empregados, como a dinâmica molecular (DM) e *docking* molecular. No que diz respeito à DM, os métodos e teorias utilizados estão bem estabelecidos e registrados pela comunidade científica, portanto detalhes técnicos podem ser encontrados em várias referências de excelência como (FRENKEL; SMIT, 2001; SCHLICK, 2010). Abaixo segue um resumo das técnicas utilizadas.

I.3.1 Simulações de Dinâmica Molecular

A dinâmica molecular é uma metodologia computacional capaz de produzir trajetórias para um sistema composto por um número finito de partículas (átomos). Através das simulações de DM é possível obter a evolução temporal de cada átomo, representado pelo centro do seu núcleo, com carga elétrica, massa e interações bem definidas. A cada avanço temporal, as forças resultante das interações interatômicas do sistema estudado são calculadas e utilizadas para descrever o movimento de cada partícula, segundo as leis de Newton (mecânica clássica).

Em posse das trajetórias e assumindo a ergodicidade dos sistemas estudados em DM, é possível calcular propriedades dinâmicas médias de um dado sistema microscópico. A ergodicidade permite estabelecer correspondência entre as médias temporais de uma trajetória e as médias por todo o espaço de fase de um sistema. Assim, é possível calcular médias termodinâmicas \bar{A}_t de propriedades de interesse

A do sistema (e.g. temperatura, pressão, energia livre, etc) na forma de médias temporais \bar{A}_t segundo:

$$\lim_{t \rightarrow \infty} \bar{A}_t = \langle A \rangle \quad (1)$$

A descrição do sistema a ser estudado é condicionada as configurações iniciais das partículas presentes no sistema e ao campo de força. Entre as diversas

formas de se obter as coordenadas iniciais de sistemas biológicos complexos, como as proteínas e ácidos nucleicos, destacam-se a cristalografia de raios X, a ressonância magnética nuclear (NMR, do inglês *Nuclear magnetic resonance*) e a Crio-microscopia eletrônica. As estruturas dos sistemas resolvidos por tais técnicas podem ser obtidas no banco de dados *Protein Data Bank* (PDB), um repositório de coordenadas atômicas de estruturas moleculares.

1.3.1.1 Campo de força

O campo de força é um conjunto de descrições das interações intra e intermoleculares do sistema. Esta descrição é realizada por meio do computo da função de energia potencial, da qual derivam as forças agindo em todas as partículas do sistema. A função de energia potencial é associada a parâmetros moleculares para compôr o campo de força.

Assim, para efetivamente movimentar um átomo i , primeiro é necessário realizar o cálculo das forças exercidas pelos demais átomos j sobre o átomo i , através da função de energia potencial. Esta função, deve aliar precisão para gerar resultados fidedignos e eficiência, evitando alta demanda computacional para resolvê-la. Além disso, a fim de reproduzir as propriedades do sistema em estudo, são empregados os parâmetros moleculares que descrevem as constantes de equilíbrio das interações entre os átomos e as cargas parciais atômicas. Existem diversos campos de força clássicos disponíveis para simulações de dinâmica molecular, como o AMBER (WANG et al., 2004), CHARMM (FOLOPPE; MACKERELL, 2000; MACKERELL et al., 1998; MACKERELL; FEIG; BROOKS, 2004) e GROMOS (SCHULER; DAURA; VAN GUNSTEREN, 2001).

O campo de força CHARMM 36 (HUANG; MACKERELL, 2013), incluindo atualizações para os parâmetros do colesterol (LIM; ROGASKI; KLAUDA, 2012), foi empregado nas presentes simulações. Nesse campo de força, a energia potencial de uma partícula ($U(r)$) dependente da sua posição r é resultado da soma de diferentes termos ($U = E_{\text{ligação}} + E_{\text{ângulo}} + E_{\text{diedro}} + E_{\text{Coulomb}} + E_{\text{vdw}} + E_{\text{impróprio}} + E_{\text{Urey-Bradley}}$). Estes termos podem ser divididos entre aqueles que descrevem as energias de interação

entre os átomos ligados covalentemente ($E_{\text{ligação}} + E_{\text{ângulo}} + E_{\text{diedro}}$) e átomos não ligados ($E_{\text{Coulomb}} + E_{\text{vdw}}$) (Figura 5). Assim a energia potencial de uma partícula ($U(r)$) é calculada de acordo com:

$$U(r) = \sum_{\text{ligação}} k_r (r - r_0)^2 + \sum_{\text{ângulo}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{diedro}} k_d [1 + \cos(n\omega - \omega_0)] + \sum_{\substack{\text{coulomb} \\ i < j}} \frac{q_i q_j}{4\pi \epsilon r_{ij}} + \sum_{\substack{\text{vdW} \\ i < j}} \epsilon_{ij} \left[\left(\frac{R_{ij}^0}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^0}{r_{ij}} \right)^6 \right] \quad (2)$$

onde, respectivamente, k_r e k_θ são constantes de força de ligação e angular; r_0 e θ_0 são distâncias e ângulos de equilíbrio; k_d , n e ω_0 são barreira torsional, sua periodicidade e fase; ϵ é a constante dielétrica; q_i é a carga parcial do átomo i ; R_{ij}^0 e ϵ_{ij} correspondem aos parâmetros e à força de van der Waals para o par ij .

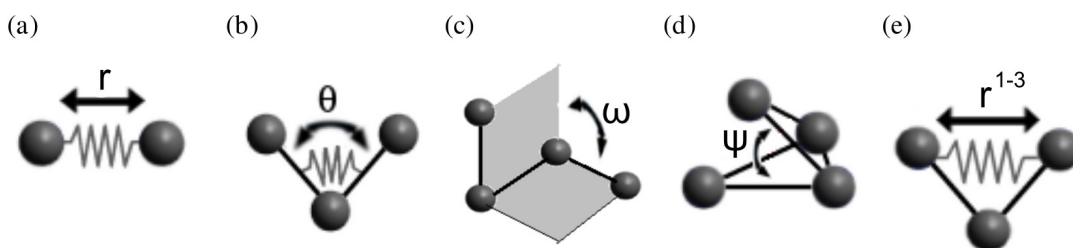


Figura 5: Ilustração dos potenciais entre átomos ligados empregados na DM.

a) Energia de ligação; b) Energia do ângulo c) Energia do diedro d) Energia do diedro impróprio e) Energia Urey-Bradley.

O CHARMM ainda fornece dois termos adicionais relacionados com restrições geométricas nas moléculas. A $E_{\text{impróprio}}$ é energia potencial da torção dos diedros impróprios entre quatro átomos. Esse termo é empregado quando se necessita manter a quiralidade em tetraedros ou planaridade em anéis, por exemplo, anéis aromáticos. Já $E_{\text{Urey-Bradley}}$ é energia potencial similar a energia potencial de ligação, porém descreve o potencial harmônico da interação entre dois átomos separados por um átomo e duas ligações covalentes, ou seja, os átomos 1 e 3 (Figura 5e).

Vale ressaltar que não existem termos referentes a formação ou quebra de

ligações covalentes, pois na dinâmica molecular esses eventos não são modelados.

I.3.1.2 Condições Periódicas de Contorno

Geralmente, nos sistemas simulados por DM não há variação no número de partículas e o tamanho do sistema é limitado. Portanto, uma parcela significativa de átomos encontra-se nos limites do sistema e podem sofrer efeitos de borda. Frequentemente, o tratamento de tais efeitos consiste na utilização de condições periódicas de contorno (PBC, do inglês *Periodic Boundary Conditions*). Nesse esquema, o sistema original é inserido em uma célula unitária mínima e réplicas desse compartimento mínimo circundam a célula original em todas as direções tridimensionais cartesianas ($\pm x$; $\pm y$; $\pm z$). É criado uma espécie de espaço infinito para o sistema, contudo, na prática apenas uma caixa é simulada, já que todas são idênticas. Portanto, se uma partícula atravessar os limites da caixa de simulação, entrando em uma das imagens do sistema, ela será representada retornando a mesma caixa no limite oposto (Figura 6). As interações interatômicas são calculadas de acordo com a distância mínima entre os átomos, descontando as dimensões da caixa.

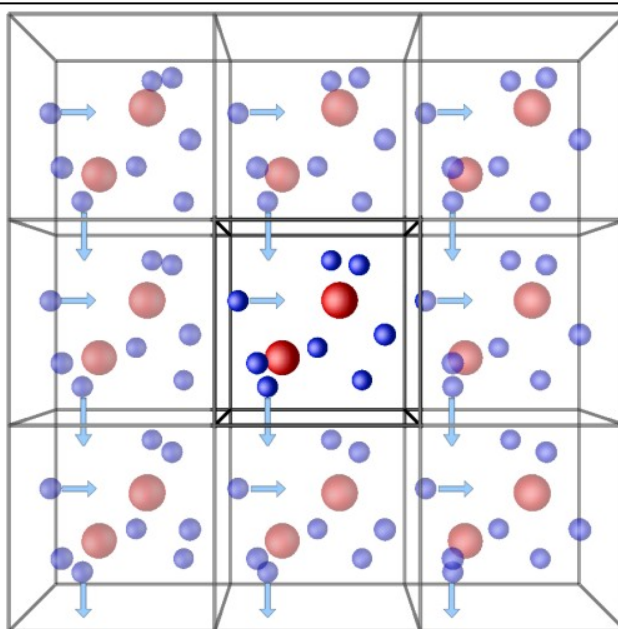


Figura 6: Esquema que ilustra as condições periódicas de contorno (PBC).

Fonte: <http://isaacs.sourceforge.net/phys/pbc.html>

I.3.1.3 Equações de movimento

A trajetória de cada átomo é obtida a partir da integração das equações newtonianas de movimento. Para resolver tais equações, a integração é dividida em pequenos estágios, separados por um intervalo de tempo fixo dt ou passo de integração. Um dos algoritmos mais utilizados por apresentar ótima estabilidade (conservação da energia total do sistema) em simulações longas é o *Verlet* (VERLET, 1967). Nesse algoritmo, após o cálculo das forças $F_i(t)$ que atuam sobre uma partícula i no instante atual t (equação 4), as coordenadas dessa partícula no momento $(t+dt)$ são obtidas a partir da sua posição atual r_i e a sua posição no momento anterior $(t-dt)$ segundo a equação:

$$r_i(t+dt) = 2r_i(t) - r_i(t-dt) + \frac{F_i(t)}{m_i} dt^2 + O(dt^4) \quad (3)$$

$$F_i(t) = - \frac{\partial U(r_i)}{\partial r_i(t)} \quad (4)$$

onde $U(r_i)$ é a função da energia potencial da partícula i , m_i sua massa e $O(dt^4)$ é o erro associado ao cálculo da posição. As velocidades das partículas, indispensável para o cálculo da energia total do sistema são obtidas pela equação:

$$v_i(t) = \frac{r_i(t+dt) - r_i(t-dt)}{2dt} + O(dt^2) \quad (5)$$

com erro associado de $O(dt^2)$.

A escolha do passo de integração dt é fundamental para a DM, pois durante esse intervalo as interações entre os átomos são consideradas constantes e caso este seja elevado, levará ao aumento do erro associado. Portanto o dt deve ser infinitesimal o bastante para assegurar a conservação da energia do sistema e amostrar o movimento mais rápido no deslocamento dos átomos. Por sua vez, um dt muito pequeno resulta no aumento significativo do custo computacional. Comumente, em sistemas biológicos, utiliza-se dt de 2 fs, (10^{-15} segundos), valor próximo da frequência de vibração dos átomos de hidrogênio (FRENKEL; SMIT, 2001).

I.3.1.4 Ensemble NPT e *annealing*

Simulações de DM, por padrão, geram trajetórias em uma amostra microcanônica, isto é, com número de partículas, volume e energia total constantes (NVE). Assim, as demais propriedades do sistema, por exemplo temperatura e pressão, poderão oscilar para valores extremos. Um dos objetivos da aplicação de DM no estudo de sistemas biológicos é reproduzir um sistema experimental real, portanto é interessante que as simulações se aproximem das condições experimentais, como isotérmicas (NVT, amostra canônica) ou isotérmicas e isobáricas (NPT). Para tanto, são adicionados barostatos e ou termostatos ao algoritmo utilizado na dinâmica.

Vários modelos de termostato para DM já foram desenvolvidos, possibilitando gerar trajetórias com temperatura constante. Entre os mais utilizados estão os termostatos de Berendsen (BERENDSEN et al., 1984), Nosé-Hoover (HOOVER, 1985; NOSÉ, 1984) e Langevin (QUIGLEY; PROBERT, 2004).

O termostato de Langevin (QUIGLEY; PROBERT, 2004) pode ser utilizado para simular uma amostra em temperatura constante. Em linhas gerais, forças aleatórias são adicionadas ao sistema de forma a simular seu acoplamento a um banho térmico. As equações de movimento de cada partícula do sistema são modificadas assim como a energia cinética de tais partículas. Desta maneira o termostato de Langevin, utilizado nesse trabalho, possibilita a geração de *ensembles* isotérmicos. O controle da pressão é possível utilizando um esquema similar ao termostato de Langevin, porém é adicionado às equações de movimentos um termo que simula o acoplamento de um pistão ao sistema. Assim, a manutenção da pressão é atingida pela alteração no volume da caixa de simulação (MARTYNA; TOBIAS; KLEIN, 1994). A combinação dessas duas abordagens permite simular sistemas NPT.

Nos momentos iniciais de uma simulação de dinâmica molecular forças fora do equilíbrio podem promover deformações estruturais das proteínas. Para evitar tal processo é utilizado um protocolo de *annealing*. O *annealing* é realizado pelo acréscimo de forças harmônicas que restringem o movimento de um grupo de

átomo, comumente os átomos da cadeia principal das proteínas. A intensidade dessas forças harmônicas é reduzida gradativamente ao longo de um intervalo de tempo e esquema definido. A partir deste ponto a simulação é realizada sem restrições.

1.3.2 RMSD e RMSF

O cálculo do RMSD (do inglês *Root Mean Square Deviation*) no tempo é uma análise comumente utilizada para avaliar a estabilidade de estruturas moleculares obtidas das trajetórias de DM. É computado o desvio médio de um conjunto de átomos das estruturas instantâneas da simulação $t > 0$ relativo ao mesmo conjunto de átomos da estrutura de referência, habitualmente a estrutura inicial $t = 0$. Um fator crucial para a determinação do RMSD é a sobreposição tridimensional das estruturas, que deve interferir o mínimo possível nas flutuações regionais da proteína. Esta etapa, permite a avaliação apenas dos desvios internos da molécula estudada, eliminando a translação e rotação do cálculo. Após a sobreposição das estruturas a serem comparadas o RMSD é calculado pela equação:

$$RMSD(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N \|x_i(t) - x_i^{ref}\|^2} \quad (6)$$

onde N é número total de átomos da seleção, $x_i(t)$ e x_i^{ref} são as coordenadas do átomo i , no tempo t e na referência.

Quando o RMSD de uma trajetória atinge um platô e flutua em torno de uma média conformacional estável, o sistema é tido em equilíbrio. Após a constatação do equilíbrio estrutural do sistema pode-se avaliar as flutuações atômicas locais do sistema, o RMSF (do inglês, *Root Mean Square Fluctuation*). Para tanto é calculado o desvio de um átomo em relação a sua posição atômica média em um intervalo de tempo definido. Com o RMSF é possível obter uma representação quantitativa de oscilações no sistema e identificar regiões mais ou menos flexíveis. O RMSF é calculado pela seguinte equação:

$$RMSF = \sqrt{\frac{1}{T} \sum_{t_j=1}^T (x_i(t_j) - \bar{x}_i)^2} \quad (7)$$

onde $x_i(t_j)$ corresponde a posição do átomo i no instante t_j e \bar{x}_i a posição média no tempo T do átomo i .

1.3.3 Docking molecular

O *docking* ou ancoragem molecular é uma técnica de modelagem molecular utilizada para encontrar regiões de ligação não covalente de ligantes a receptores macromoleculares, assim como estimar a afinidade de tais ligações. Duas etapas podem ser caracterizadas nos estudos de *docking*: (1) busca exaustiva das orientações e conformações possíveis do ligante (também denominadas poses) no receptor, conforme um algoritmo de busca. A capacidade de cobrir todo o espaço conformacional do ligante e do receptor é desejável no algoritmo de busca, entretanto o aumento na quantidade de graus de liberdade, principalmente relacionada ao receptor, na busca está correlacionado com a sua demanda computacional. (2) Classificação das soluções encontradas em mais ou menos favoráveis, de acordo com a *scoring function* do *software*.

Algumas *scoring functions* são baseadas em campos de força, com conceitos físicos, e outras utilizam parâmetros experimentais em bancos de dados para pesar termos da equação utilizada, conhecidas como *knowledge-based* e *empirical-based*. Entretanto, em ambos os casos, a precisão do ranqueamento das poses ainda é um dos maiores desafios das *scoring functions* (DHANIK; KAVRAKI, 2012). O algoritmo de busca e a *scoring functions* variam de acordo com o *software* de *docking* utilizado. Nesse trabalho foi utilizado o *AutoDock Vina* (TROTT; OLSON, 2010).

O *AutoDock Vina* foi o programa utilizado nesse trabalho, este é um dos programas mais citados de *docking*. Vários fatores contribuíram para seu sucesso, como uma evolução na precisão dos modos de ligação, facilidade de uso, disponibilização gratuita, alta performance e a capacidade de paralelização, o que o significativamente mais rápido do que o *AutoDock 4* (SOUSA et al., 2013).

Para que o resultado da ancoragem molecular seja fiel à realidade biológica a flexibilidade do receptor e do ligante deve ser avaliada em sua totalidade. Entretanto o tempo necessário para que o algoritmo percorra todos os graus de liberdade,

principalmente do receptor, tornaria inviável a flexibilização de ambos, receptor e ligante. No Auto Dock Vina é possível realizar o *docking* com a flexibilização das moléculas de entrada, porém devido ao limite máximo de rotações permitidas pelo algoritmo ser de 32 torções, a flexibilização apenas é aplicada a moléculas muito pequenas (como lipídios). Nesse sentido o *ensemble docking* surge como alternativa para lidar com a flexibilidade do receptor e de ligantes com rotações superiores a 32. No *ensemble docking* a busca das poses do ligante não é realizada em uma estrutura estática do receptor, mas sim em um conjunto de conformações derivadas normalmente de trajetória de DM (GUEDES; DE MAGALHÃES; DARDENNE, 2014).

II . CAPÍTULO II. PEPTÍDEOS LIGANTES DE NUCLEOSSOMO (NBPeps)

II.1 INTRODUÇÃO

II.1.1 Proteínas que se ligam ao nucleossomo (NBPs)

Uma das formas encontradas pela natureza para promover o remodelamento da cromatina é mediante a interação com proteínas capazes de se ligarem diretamente ao nucleossomo, NBPs (do inglês *Nucleosome Binding Proteins*). Diversas proteínas e/ou complexos proteicos já tiveram os detalhes da interação com o nucleossomo revelados. Recentemente, Zhou, Gaullier e Luger (2019), publicaram um trabalho de revisão comparando o modo de ligação de complexos proteínas:nucleossomo (ZHOU; GAULLIER; LUGER, 2019). São listados vinte complexos estruturais e metade destes interagem diretamente com a superfície proteica do nucleossomo (Figura 7). Interessantemente, apenas a proteína CENP-N não depende do *patch* ácido para manter a ligação com o nucleossomo (PENTAKOTA et al., 2017). As demais NBPs reconhecem e compartilham o modo

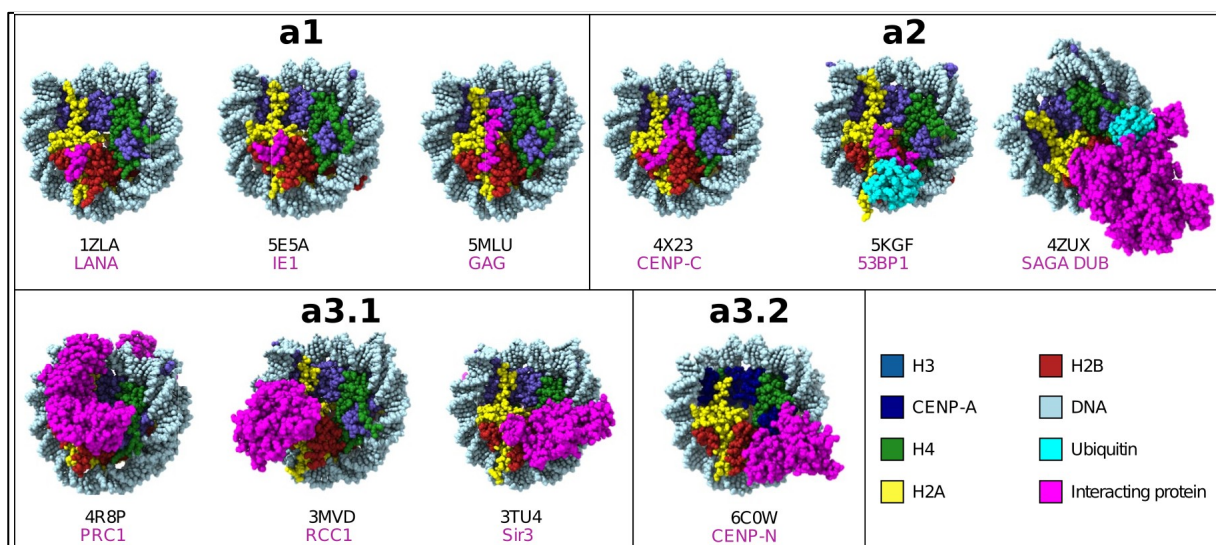


Figura 7: NBPs que tem como alvo a superfície proteica do nucleossomo.

O código PDB de cada complexo está indicado. **a1**, fragmentos de proteínas que reconhecem o *patch* ácido; **a2**, proteínas que reconhecem o *patch* ácido e outras regiões das histonas; **a3**, proteínas que interagem com histonas e o DNA nucleossomal (a interação com o *patch* ácido está presente em **a3.1** mas não em **a3.2**). Adaptado de (ZHOU; GAULLIER; LUGER, 2019).

de contato com o *patch* ácido. Argininas ou lisinas das NBPs se projetam na cavidade ácida do nucleossomo em um reconhecimento molecular semelhante a uma “âncora” (Figura 8) (LESBATS et al., 2017; MCGINTY; TAN, 2016; MORGAN et al., 2016; WILSON et al., 2016).

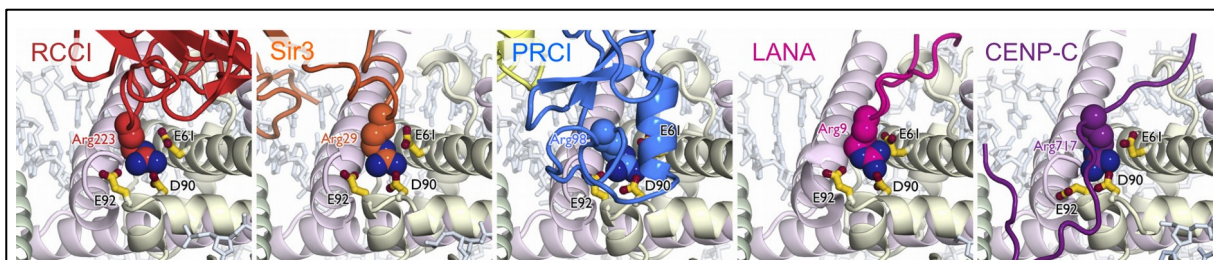


Figura 8: Sítio de ancoragem de argininas no *patch* ácido.

Cavidade formada pelos resíduos E61, D90 e E92 da H2A, responsável pela ligação da âncora de arginina das NBPs RCC1, Sir3, PRC1, LANA, CENP-C. Adaptado de (MCGINTY; TAN, 2016).

Dentre as NBPs já identificadas, três são particularmente importantes para este trabalho.

(1) O antígeno nuclear associado a latência, ou LANA (*Latency-Associated Nuclear Antigen*), do KSHV (*Kaposi's Sarcoma-Associated Herpesvirus*). No cristal é revelado um peptídeo formado por apenas 23 aminoácidos que se encaixa na cavidade do *patch* ácido formando um *hairpin* nessa região (BARBERA et al., 2006).

(2) O RCC1 (*Regulator of Chromosome Condensation 1*), importante proteína para o processo de mitose e transporte de macromoléculas entre o citoplasma e o núcleo. Foi a primeira proteína intacta cristalizada em complexo com o nucleossomo interagindo com o *patch* ácido nos dois lados do nucleossomo e também realizando contato com o DNA (MAKDE et al., 2010).

(3) A HMGN2 (*High Mobility Group Nucleosome Binding Protein 2*), as proteínas do grupo HMGN interferem na compactação da cromatina mediada pela *linker* histona, aumentando a atividade transcricional. Por meio de ressonância magnética nuclear (NMR) os pesquisadores chegaram a um modelo de interação do

nucleossomo com a HMGN2. Foi verificado que um peptídeo de 24 aminoácidos da HMGN2 formava um “grampo” entre o DNA nucleossomal (ligando-se na *dyad*) e o núcleo proteico do nucleossomo (ligando-se no *patch* ácido) (KATO et al., 2011).

II.1.2 Nucleossomo como alvo terapêutico

Diante das evidências que revelaram a existência de proteínas e outras moléculas que interagem com o nucleossomo, como as NBPs e a cauda da H4, é possível perceber que a modulação da estrutura da cromatina pode ser fortemente influenciada pelo equilíbrio entre a interação de tais moléculas ligantes de nucleossomos e a superfície nucleossomal. Em 2013, Hansen e colaboradores propuseram um modelo de regulação da cromatina no contexto das NBPs e *patch* ácido. Na ausência de NBPs, o estado da cromatina é definido pela ligação da cauda da histona H4 ao *patch* ácido, denominado estado padrão. NBPs alterariam o estado padrão da cromatina para uma conformação específica condizente com

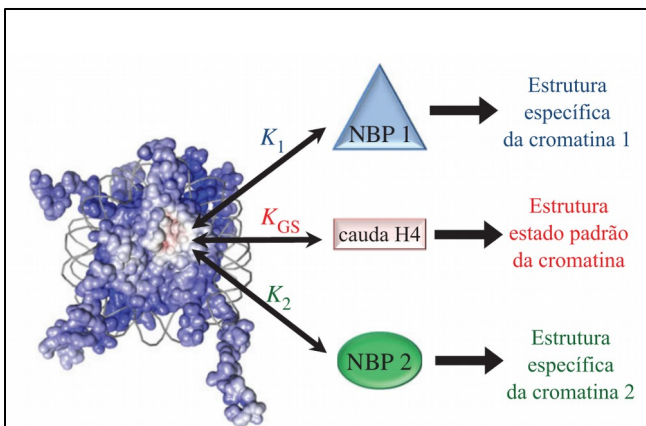


Figura 9: Modelo de remodelamento da cromatina mediante a presença de NBPs.

K_{GS} representa a constante de equilíbrio para o estado padrão, K_1 e K_2 são as constantes de equilíbrio para a interação da NBP1 e NBP2 com o *patch* ácido. A estrutura específica da cromatina induzida pelas NBPs pode ser desde um estado completamente compactado ou descompactado. Adaptado de (KALASHNIKOVA et al., 2013).

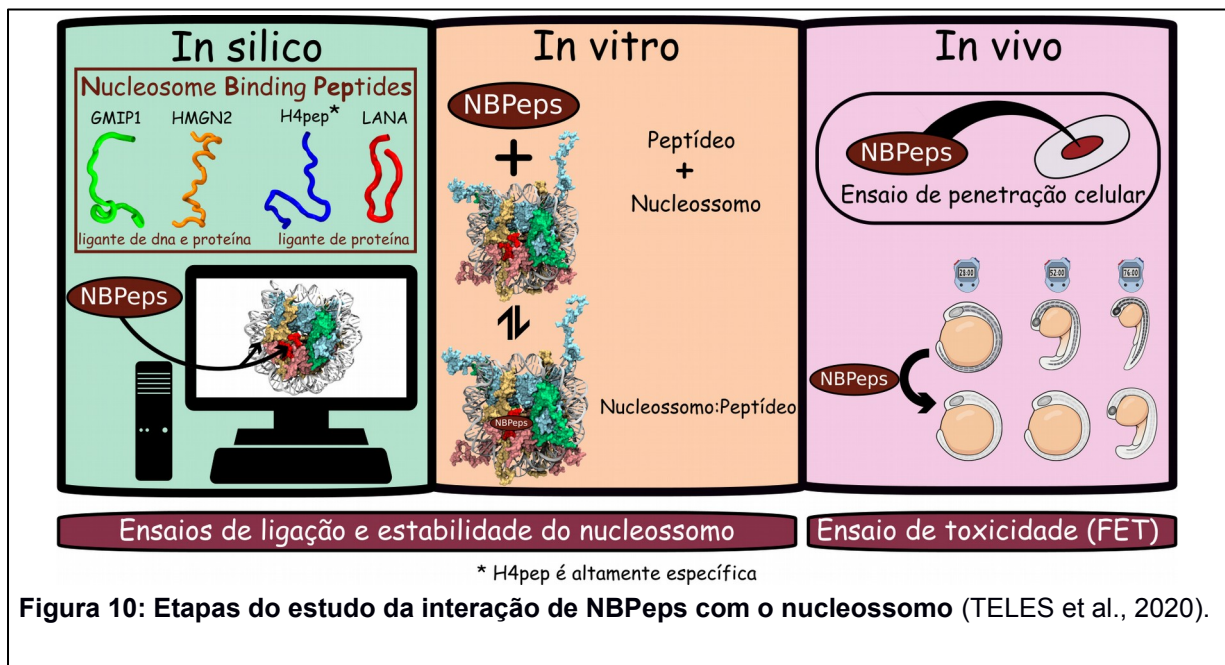
sua funcionalidade biológica. O modelo também sugere que NBPs competiriam com a cauda da H4 e com outras NBPs para a ligação ao *patch* ácido, o que se traduz em constantes de ligação específicas para cada caso. Desta forma, a configuração predominante da cromatina local seria dependente da NBP que apresentasse maior constante de ligação e/ou concentração local. Outra possibilidade seria o estado padrão (Figura 9) (KALASHNIKOVA et al., 2013).

Em 2014, um estudo investigou a possibilidade de inibição da ligação do LANA ao nucleossomo com intuito de tratar infecções latentes por KSVH. Foram

testadas mais de 350.000 pequenas moléculas, entretanto nenhuma passou pela contraprova nos ensaios de competição com LANA. Os autores sugerem que compostos mais complexos sejam mais eficientes em deslocar o LANA do nucleossomo (BEAUCHEMIN et al., 2014).

Baseado no modelo proposto por Hansen e colaboradores, nos achados de Beachemin e colaboradores e na natureza regulatória do *patch* ácido, o grupo do Professor Guilherme Santos (2015) sugeriu que a superfície do nucleossomo poderia ser vista como um alvo farmacológico (SILVA; OLIVEIRA; SANTOS, 2015). Nesse trabalho é proposto que a modulação da cromatina por moléculas exógenas ligantes de nucleossomo (eNBMs) seja um mecanismo valioso do ponto de vista farmacológico, podendo visar a regulação de várias desordens celulares. Em um trabalho posterior foi ampliada a discussão para o desenho racional de eNBMs. É apresentado que eNBMs poderiam ser peptídeos ou peptidomiméticos, provavelmente com aminoácidos modificados, desenhados a partir da estrutura dos motivos ligadores ao nucleossomo presentes nas NBPs (CABRAL; MACHADO; SANTOS, 2016). Neste documento será adotada a nomenclatura NBPePs (peptídeos ligantes de nucleossomo) para designar peptídeos, com aminoácido não modificados, que se ligam ao nucleossomo. Recentemente, foi demonstrado que compostos binucleares de rutênio que se ligam ao *patch* ácido induziram compactação anormal da cromatina (DAVEY et al., 2017). A anormalidade resultou em morte celular e os autores propõem que tais compostos apresentam potencial aplicação para o desenvolvimento de drogas.

Diante das informações aqui apresentadas, fica evidente que existe um campo a ser abordado acerca do desenvolvimento e caracterização de eNBMs. Portanto, foi desenvolvido um estudo da interação de quatro NBPePs com o nucleossomo. Esse trabalho apresentou etapas *in silico*, *in vitro* e *in vivo* (Figura 10). Entretanto, serão apresentados neste capítulo somente os resultados *in silico*, de caracterização atômica da interação entre NBPePs e nucleossomo. Os ensaios *in vitro* e *in vivo* foram realizados por outro membro do laboratório (Kaian Teles).



Os quatro NBPeps investigados foram: (i) o GMIP1 (do inglês, *Genetic Modified Inducible Peptide*), um peptídeo formado pelo motivo de ligação ao nucleossomo da HMGN2, aqui denominado apenas (ii) HMGN2pep, (iii) o H4pep (peptídeo formado pelo motivo de ligação ao nucleossomo presente na cauda N-terminal da H4) e (iv) o LANA. H4pep e LANA se ligam somente ao *patch* ácido e HMGN2 e GMIP1 interagem com o *patch* ácido e com o DNA.

O peptídeo GMIP1 foi desenvolvido pelo grupo do Prof. Guilherme Santos. O desenho *in silico* foi realizado com o programa KVFinder (OLIVEIRA et al., 2014) que dentre as estruturas cristalográficas das NBPs disponíveis no PDB em 2012, apontou que a interação do RCC1 com o nucleossomo apresentava uma melhor acomodação na superfície do nucleossomo. O RCC1, apresenta duas porções distintas de ligação ao nucleossomo e ambas porções foram utilizadas para o desenho do GMIP1. Para unir estas duas porções, uma ligante do *patch* ácido e outra que interage com o DNA, foi adicionado um espaçador de três alaninas (Figura 11). O pedido de patente do GMIP1 e outras 111.105 variantes do mesmo, foi depositada em março de 2015 no INPI (Instituto Nacional de Propriedade Intelectual).

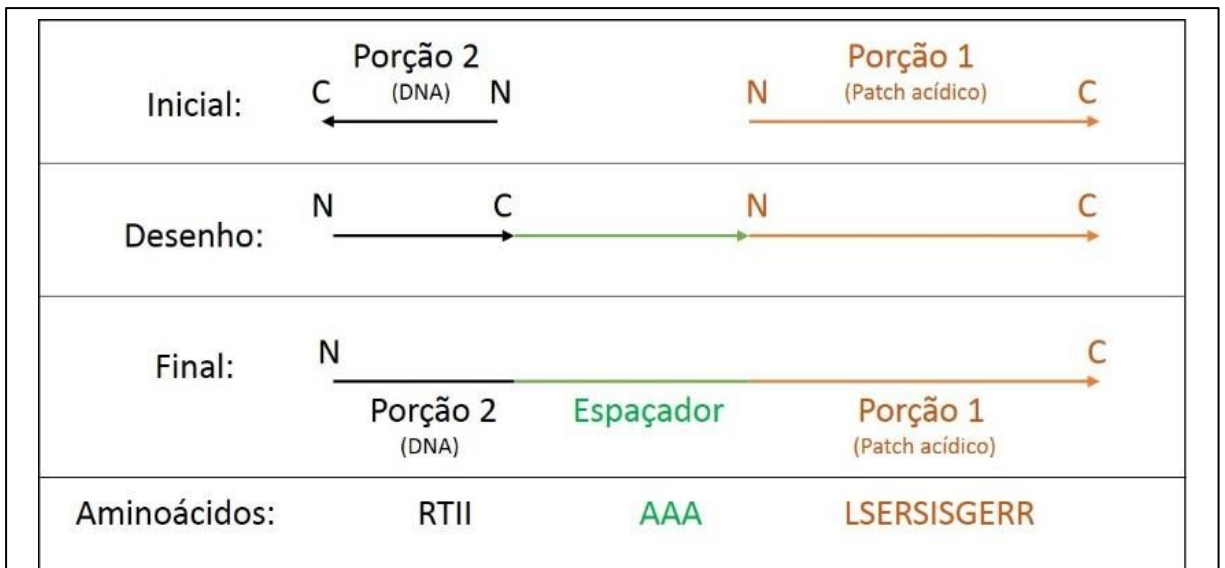


Figura 11: Desenho do peptídeo GMIP1.

Quadro I) Orientação das porções 1 e 2 da proteína RCC1 escolhidas para o desenho do peptídeo GMIP1, sendo a porção 1 aquela que faz contato com o DNA, e a porção 2 a que faz ligação com o *patch* ácido. Quadro II) Desenho do peptídeo com a orientação da porção 1 invertida e adição do espaçador para a construção de GMIP1. A porção 1 foi invertida para mimetizar a orientação do contato do RCC1 com a superfície do nucleossomo. Quadro III) Construção final de GMIP1 mostrando a sequência de aminoácidos das porções 1 e 2 conectados pelo espaçador de três alaninas.

II.2 OBJETIVOS

II.2.1 Objetivo geral

Investigar as características da interação de moléculas ligantes de nucleossomo (lipídios e peptídeos), avaliando os efeitos dessa interação em nível atômico.

II.2.1 Objetivos específicos

a) NBPeps

- Construção dos sistemas atomísticos de cada peptídeo com o nucleossomo;
- Simular por DM os complexos;
- Observar como a ligação dos NBPeps interfere na estrutura do nucleossomo;
- Correlacionar os achados computacionais com os dados biofísicos e biológicos

II.3 PROTOCOLOS

No total foram criados modelos atomísticos de interação com o nucleossomo de cinco peptídeos diferentes (NBPePs) e um controle negativo. O peptídeo da cauda H4 (H4pep) foi analisado nas conformações: aberta (H4-coil) e alfa-hélice (H4-hx). Os demais NBPePs foram LANA, HMGN2pep e GMIP1. O peptídeo controle negativo (CtrlNeg) foi desenvolvido com o intuito de não se ligar ao nucleossomo utilizando como modelo o GMIP1. Os aminoácidos do GMIP1 foram modificados por outros com características físico-químicas contrárias, assim os pontos de contatos essenciais para a interação do GMIP1 com o nucleossomo foram modificados.

GMIP1, H4-hx e CtrlNeg passaram pela etapa de *docking* molecular em busca da configuração inicial de contato com o nucleossomo. GMIP1 e CtrlNeg ainda passaram por um tratamento adicional de clusterização de estruturas antes da etapa de *docking*. Cada NBPeP e o CtrlNeg foram amostradas juntamente com o nucleossomo durante 115 ns em três condições diferentes. NBPeP apenas em um lado do nucleossomo (sistema lado 1), NBPeP no outro lado do nucleossomo (sistema lado 2) e NBPeP em ambos os lados do nucleossomo (sistema ambos). Abaixo segue os detalhes das simulações DM, do processo de clusterização e da etapa de *docking*.

II.3.1 Protocolo de simulações de DM

As simulações de DM foram realizadas com o programa NAMD (versão 2.10 e 2.13), um código de dinâmica molecular de alto desempenho desenvolvido para simulação de grandes sistemas biomoleculares (PHILLIPS et al., 2005). O campo de força utilizado foi o CHARMM 36, o qual contém parâmetros para proteínas e ácidos nucleicos (HUANG; MACKERELL, 2013).

Todos os sistemas foram simulados empregando *ensembles* NPT, utilizando o algoritmo de Nosé-Hoover para manter pressão e o termostato de Langevin para manter a temperatura constante (respectivamente em 1 atm e 300 K). Foi utilizado um passo de integração de 2 fs e as estruturas (*frames*) foram armazenadas a cada 5000 passos, 10 picossegundos. As interações de van der Waals foram calculadas a

cada passo. Já as interações eletrostáticas foram obtidas a cada 2 passos empregando o método PME (DARDEN; YORK; PEDERSEN, 1993). As ligações covalentes com átomos de hidrogênio foram mantidas rígidas pelo algoritmo SHAKE. Os sistemas foram simulados em presença explícita de água, descrita pelo modelo TIP3 (JORGENSEN et al., 1983). Os sistemas foram inseridos em uma caixa de simulação que determinou os limites do sistema e foi utilizada para as definições das condições periódicas de contorno (PBC) nos eixos x, y e z. As dimensões da caixa de simulação divergiu entre os sistemas, mas foi utilizada uma distância mínima de 20 Å entre as extremidades da proteína e da caixa de simulação.

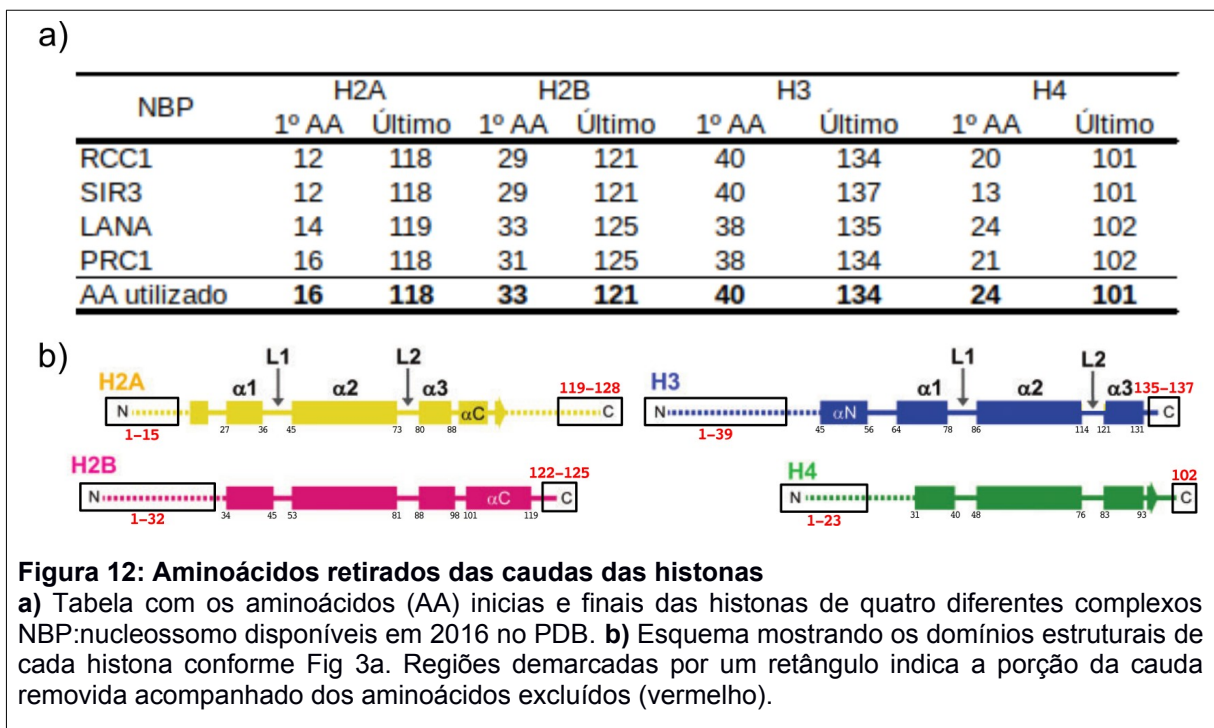
As simulações pertinentes foram conduzidas em supercomputadores de médio porte, disponibilizados pelo Laboratório de Biologia Teórica e Computacional (LBTC) da Universidade de Brasília sob supervisão do prof. Dr. Werner Treptow e auxílio do grupo de trabalho do laboratório. Em média, para um sistema típico de 300 mil átomos, podem ser gerados 10 ns (nanossegundos) de trajetórias a cada 24 horas.

Para analisar e visualizar os modelos computacionais e os resultados das simulações de DM foram criados vários códigos nas linguagens de programação TCL e *Python*. Também foram utilizados códigos desenvolvidos no LBTC. Todos esses códigos foram utilizados em conjunto com o VMD (*Visual Molecular Dynamics*) (HUMPHREY; DALKE; SCHULTEN, 1996), um software de análise gráfica 3D.

Seguem os detalhes dos sistemas estudados:

II.3.1.1 Controle (nucleossomo livre em solução)

As coordenadas iniciais do nucleossomo foram obtidas no PDB, código 3TU4 (ARMACHE et al., 2011) com a exclusão de parte das caudas das histonas. A escolha do extensão de cada cauda foi baseado nas estruturas atômicas de NBPs em complexo com o nucleossomo disponíveis no PDB em 2016 (Tabela 2). De forma que cada cauda de histona apresentasse a menor extensão presente no PDB. Todas as histonas tiveram suas caudas reduzidas a no máximo 5 aminoácidos.



O nucleossomo, com a redução das caudas das histonas, foi solvatado em solução aquosa com 150 mM de NaCl em uma caixa de 141 x 147 x 110 Å³, totalizando 214.224 átomos. Desses, 182 íons cloro, 414 íons sódio e 64.156 moléculas de água. Foi utilizado um protocolo de *annealing*, no qual um potencial harmônico de 1 kcal/mol foi aplicado aos átomos da cadeia principal do nucleossomo. A intensidade do potencial foi reduzida (de 1 a 0,1 a 0,01 kcal/mol) sucessivamente ao longo de 4 ns. Ao final, foi produzida uma trajetória total de 115 ns.

II.3.1.2 NBPePs

No total foram simulados cinco NBPePs em complexo com o nucleossomo, são elas: cauda da H4 (nas conformações aberta e alfa-hélice), LANA, HMG2, GMIP1 e controle negativo (CtrlNeg). Cada peptídeo foi simulado em quatro condições diferentes: (i) livre em solução, (ii) nas duas faces do nucleossomo, (iii) apenas em uma face e (iv) apenas na outra face. Tanto as estruturas atômicas de cada NBPeP quanto o encaixe com o nucleossomo foram obtidos por métodos diferentes, entretanto, para todos os sistemas as coordenadas do nucleossomo

foram as mesmas, adquirida da trajetória de equilíbrio do sistema controle (10 ns).

Na tabela abaixo segue os detalhes da formação de cada sistema:

Tabela 1. Detalhes dos sistemas estudados

NBM	Sequência	Coordenadas	Complexo	Referências
H4 (aberta)	RGKGGKGLGKGGAKR XRK VLR	Cristalografia de raio X (1KX5)	Baseado na estrutura da cristalografia de raio X	(DAVEY et al., 2002)
H4 (hélice)	RGKGGKGLGKGGAKRHRK VLR	Servidor on line (I-tasser)	Baseado em Arya et. Al (<i>Docking</i>)	(YANG; ARYA, 2011)
HMG2	DEPQRRSARLSAKPAPPKP EPKPKK	RMN	Baseado na estrutura da RMN	(KATO et al., 2011)
LANA	PGMRLRSGRSTGAP	Cristalografia de raio X (1ZLA)	Baseado na estrutura da cristalografia de raio X	(BARBERA et al., 2006)
GMIP1	RTIIAAALSERSISGEGR	Desenho racional (Kvfinder)	<i>Ensemble Docking</i>	(MAKDE et al., 2010)
CtrlNeg	LIFFGKAXAGXAVAAPIVK GV	Servidor on line (I-tasser)	<i>Ensemble Docking</i>	

O protocolo de *annealing* empregado foi: primeiramente, para que o banho de eletrólitos se equilibrasse, foi imposto potencial harmônico de 1 kcal/mol contrário ao movimentos dos átomos do complexo (nucleossomo e NBPeps). Posteriormente, foram aplicadas restrições harmônicas, por aproximadamente 7 ns aos átomos dos NBPeps e 5 ns aos átomos do nucleossomo. Os potências harmônicos foram removidos sucessiva e gradualmente (de 1 a 0,1 a 0,01 kcal/mol) nos átomos da cadeia lateral e da cadeia principal dos aminoácidos. Para o nucleossomo, o padrão de átomos com potenciais harmônicos foi: Cadeia principal do nucleossomo → aminoácidos a 8 Å de qualquer átomo do NBPep (cadeia lateral e depois a cadeia principal) → aminoácidos a 3 Å de qualquer átomo do NBPep (cadeia lateral e depois a cadeia principal). Já para os átomos dos NBPeps, o potencial harmônico começou a ser reduzido após 5,5 ns de simulação (cadeia lateral e depois a cadeia principal). Dessa forma, as cadeias laterais do nucleossomo tiveram mais de 1,5 ns para acomodar-se aos NBPeps com a movimentação restringida. Para cada NBPep em qualquer condição de simulação foram produzidos 115 ns de trajetória de DM, incluindo o período do *annealing*. Ao todo, 2760 ns de simulações foram amostrados.

II.3.2 Protocolo de Clusterização

A *clusterização* é uma ferramenta que permite a separação de dados heterogêneos em grupos mais similares entre si. Portanto, pode ser considerada como uma mineração de dados, que aplicada a um conjunto de elementos os divide em grupos que compartilham características similares (ABRAMYAN et al., 2016). Tais características são avaliadas por alguma função que permite a diferenciação dos elementos do conjunto de dados. A *clusterização* pode ser utilizada nos mais variados campos da ciência que produzem uma elevada quantidade de dados, sendo utilizada em trajetórias de DM desde os primórdios dessa técnica. Diversos algoritmos de clusterização já foram desenvolvidos empregando modelos matemáticos distintos para separar os grupos de elementos (SHAO et al., 2007).

Trajetórias de DM de equilíbrio em solução do GMIP1 e do peptídeo controle negativo foram agrupadas utilizando o programa VMD. Nesse programa é utilizado o algoritmo de Daura (DAURA et al., 1999) para dividir as conformações em grupos similares. Esse algoritmo é um versão simplificada e mais rápida do algoritmo *quality threshold* (QT). O RMSD entre as estruturas foi utilizado para separar ou agrupar as estruturas em grupos distintos. O valor de corte (*cutoff*) do programa define os valores máximos de RMSD entre duas estruturas para que estas sejam consideradas similares. Os valores de corte utilizados foram 4.5 Å para o GMIP1 e 3.5 Å para o controle negativo. Ambos valores são diferentes do padrão do programa (1.0 Å), mas foram empregados devido a presença de regiões desestruturada e flexíveis nos peptídeos NBPePs, principalmente para o GMIP1, as quais contribuem para o aumento do RMSD. O programa dividiu as estruturas em quatro grupos ou *clusters*, além de um grupo adicional de estruturas não agrupadas. Totalizando cinco *clusters* para cada peptídeo.

II.3.3 Protocolo de *Docking* Molecular

Foram realizados experimentos de *docking* ou ancoragem molecular dos peptídeos com o nucleossomo em busca da conformação inicial do complexo para posteriores simulações de DM. Os peptídeos investigados por essa técnica foram: o peptídeo da cauda da H4 na conformação em alfa-hélice, o GMIP1 e o peptídeo

controle negativo (CtrlNeg).

Os experimentos de *docking* foram realizados com o programa Auto Dock Vina (TROTT; OLSON, 2010). Nesse programa o espaço de busca de soluções é chamado de *grid*. Como foi investigada a interação dos NBPeps com o *patch* ácido, essa região determinou a escolha do *grid*. Este compreendeu uma caixa de 45 x 45 x 45 Å³ cujo centro correspondeu ao centro de massa do *patch* ácido. Em geral, nesse programa é possível considerar as torções de pequenos ligantes, porém proteínas (tanto receptor quanto ligante) são tratadas como estruturas fixas. Assim para considerar a flexibilidade inerente de moléculas biológicas em solução, foram fornecidas ao programa estruturas em diferentes conformações dessas moléculas advindas de trajetórias de equilíbrio de DM (*Ensemble Docking*). Como o *patch* ácido apresenta pequenas flutuações ao longo da trajetória de DM (RMSD), não se fez necessário a utilização de um conjunto de conformações do receptor. Entretanto, devido a ausência de estruturas secundárias estáveis nos NBPeps aqui testados, a utilização de diversas estruturas para os ligantes foi imprescindível. Os parâmetros foram deixados no padrão do programa, com exceção apenas da exaustividade, que foi aumentada para 200. Esse parâmetro, quando aumentado, permite que o programa aumente o tempo de procura de encaixe e a quantidade de conformações do ligante (TROTT; OLSON, 2010).

II.3.3.1 H4-hx

O *docking* da cauda da H4 na conformação alfa-hélice (H4-hx) foi realizado de maneira a reproduzir o resultado obtido por (YANG; ARYA, 2011). Em seu trabalho, Yang e Arya realizaram exaustivos cálculos de DM e propôs a conformação em alfa-hélice para o peptídeo da cauda da H4. Ainda mais, eles verificaram que a energética de interação da cauda da H4 com o *patch* ácido é consideravelmente mais favorável para o peptídeo na conformação alfa-hélice.

Assim, foi utilizada apenas uma conformação desse peptídeo e todas as 400 soluções geradas pelo programa foram selecionadas. Foi utilizado um *cutoff* de 1Å no agrupamento das nuvens de soluções com intuito de gerar o maior número de modelos de interação possível. Posteriormente, a nuvem de solução que mais se

assemelhou com o resultado do artigo de Yang & Arya foi escolhida.

II.3.3.2 GMIP1 e CtrINeg

GMIP1 e o peptídeo controle negativo apresentaram alta flexibilidade conformacional na simulação de DM em solução. Assim, foram fornecidas ao programa 20 configurações independentes (*frames*) dos ligantes obtidas aleatoriamente das trajetórias de DM em solução. A análise de clusterização das conformações da trajetória de DM determinou os intervalos de tempo 51ns – 60ns para o GMIP1 e 47ns – 56ns para peptídeo controle negativo (ver tópico II.4.1.1).

O Auto Dock Vina investigou as formas de encaixe de cada uma das 20 estruturas de ambos peptídeos no *patch ácido*. As 20 melhores soluções classificadas pelo software foram selecionadas, totalizando 400 modelos de encaixe para cada peptídeo no nucleossomo. Esses modelos foram agrupados em nuvens de soluções utilizando um critério de proximidade com *cutoff* de 5 Å de distância entre os resultados.

II.4 RESULTADOS

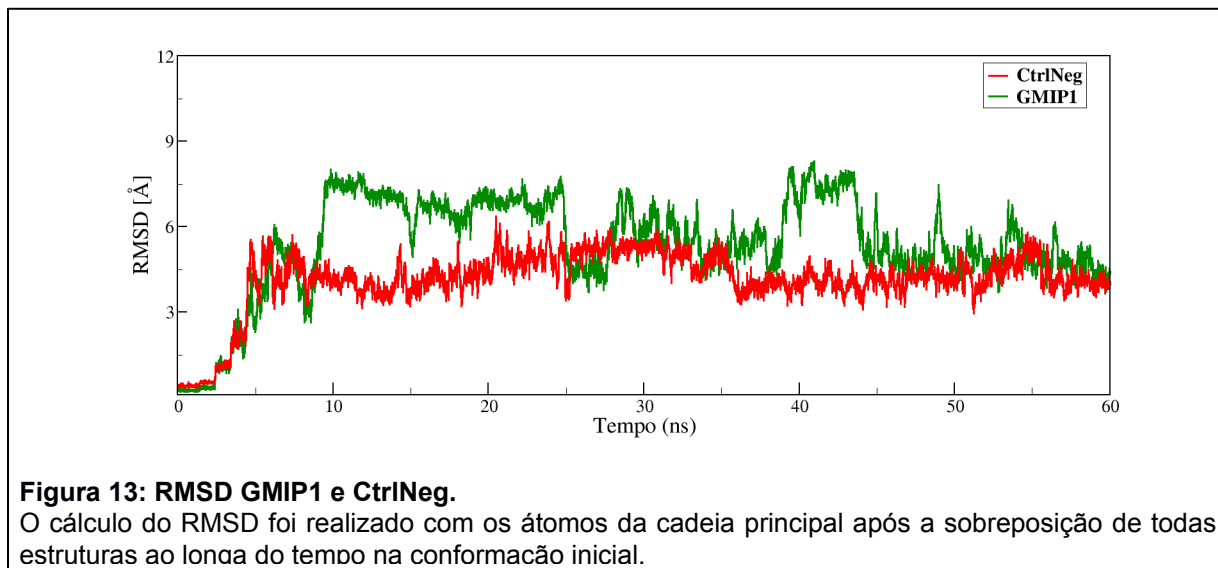
Inicialmente foi buscado o modo de interação dos NBPePs que não dispunham de um modelo de interação com o nucleossomo validado, isto é, estrutura tridimensional disponível na literatura. Posteriormente, foram analisadas as trajetórias do nucleossomo na presença dos NBPePs. Relembrando que cada peptídeo foi amostrado em três condições diferentes, individualmente no *patch* ácido em cada um dos lados do nucleossomo e conjuntamente em ambos os lados. Após o período de equilíbrio estrutural do complexo foram realizadas diversas análises com o intuito de caracterizar a interação dos NBPePs com o nucleossomo.

II.4.1 Busca do modo de interação com o nucleossomo dos NBPePs CtrlNeg, GMIP1 e H4-hx

Alguns peptídeos, como o CtrlNeg, H4-hx e GMIP1 não dispunham de modelos estruturais de contato com o nucleossomo. Para esses ligantes foram realizados estudos prévios de *docking* em busca da forma de interação com o nucleossomo. Para os outros NBPePs esse passo não foi necessário, visto que já existiam as coordenadas iniciais do complexo.

O *docking* do CtrlNeg e GMIP1 foi realizado empregando várias estruturas desses ligantes (*ensemble docking*), para que as flutuações inerentes destes fossem levadas em consideração. Assim, ambos peptídeos foram simulados em solução aquosa com 150 mM de NaCl por 60 ns. A figura 13 apresenta o RMSD dessas trajetórias após sobreposição das estruturas da trajetória. Os átomos das cadeias principais da estrutura inicial de cada peptídeo foram a referência para a sobreposição. É evidente as flutuações do RMSD desses peptídeos, principalmente para o GMIP1 (Figura 13). Tais flutuações sugerem que ao longo da trajetória vários estados conformacionais distintos e provavelmente distantes entre si são alcançados. Isso era esperado, visto que ambos peptídeos possuem estruturas tridimensionais de *random coil*, a qual está relacionada a um espaço conformacional extenso. Entretanto, isto se torna um obstáculo para a escolha das estruturas a

serem utilizadas nas análises de *docking*, objetivo da produção dessas trajetórias. A fim de reduzir o espaço conformacional para a busca da interação com o nucleossomo as trajetórias, após o período de estabilização, foram divididas em grupos ou *clusters* conformacionais (descrito no tópico II.3.2).



II.4.1.1 *Clusters* conformacionais

O GMIP1 se divide basicamente em três grupos (Figura 14). Nos *clusters* 1 e 3 o peptídeo encontra-se em uma conformação mais aberta, enquanto que no *cluster* 2 o GMIP1 visita estados conformacionais mais retraídos. Já o *cluster* 4 se assemelha a um estado intermediário entre os grupos anteriores. Como mencionado anteriormente, o GMIP1 foi desenhado baseado no motivo de ligação do RCC1, o qual se liga ao nucleossomo mediante contatos com o DNA e o *patch* ácido. Dentre os grupos encontrados, somente as conformações mais abertas podem satisfazer tal modo de interação, *clusters* 1 e 3. Como o *cluster* 1 é o predominante e satisfaz o requisito estrutural para a ligação ao nucleossomo, o intervalo 50-60 ns (*cluster* 1) foi o escolhido para a amostragem do *ensemble docking*. Já o CtrlNeg (Figura 14), apresenta grupos conformacionais mais similares do que o GMIP1. Pequenas variações são responsáveis pelas diferenças estruturais e consequente classificação em grupos distintos. No geral, as conformações assemelham-se a um laço, e na região interna desse “laço” que as pequenas diferenças são percebidas. É evidente que o primeiro *cluster* é o mais frequente durante o intervalo de tempo

analisado. Logo, o período final (47-56 ns) desse grupo (*cluster 1*) foi o escolhido para a amostragem do *ensemble docking*.

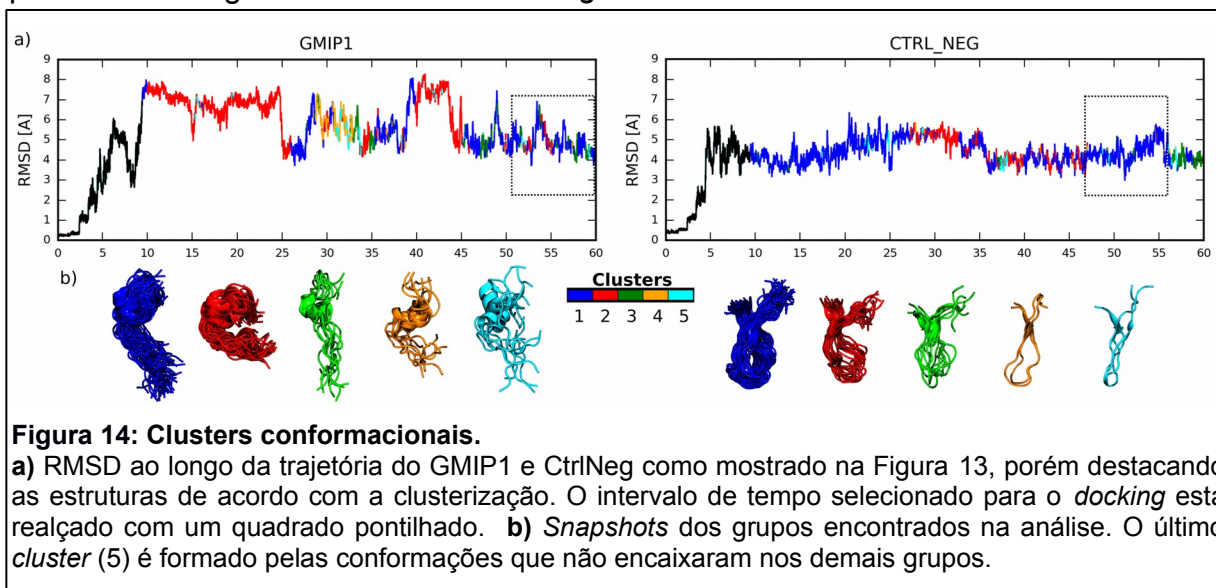


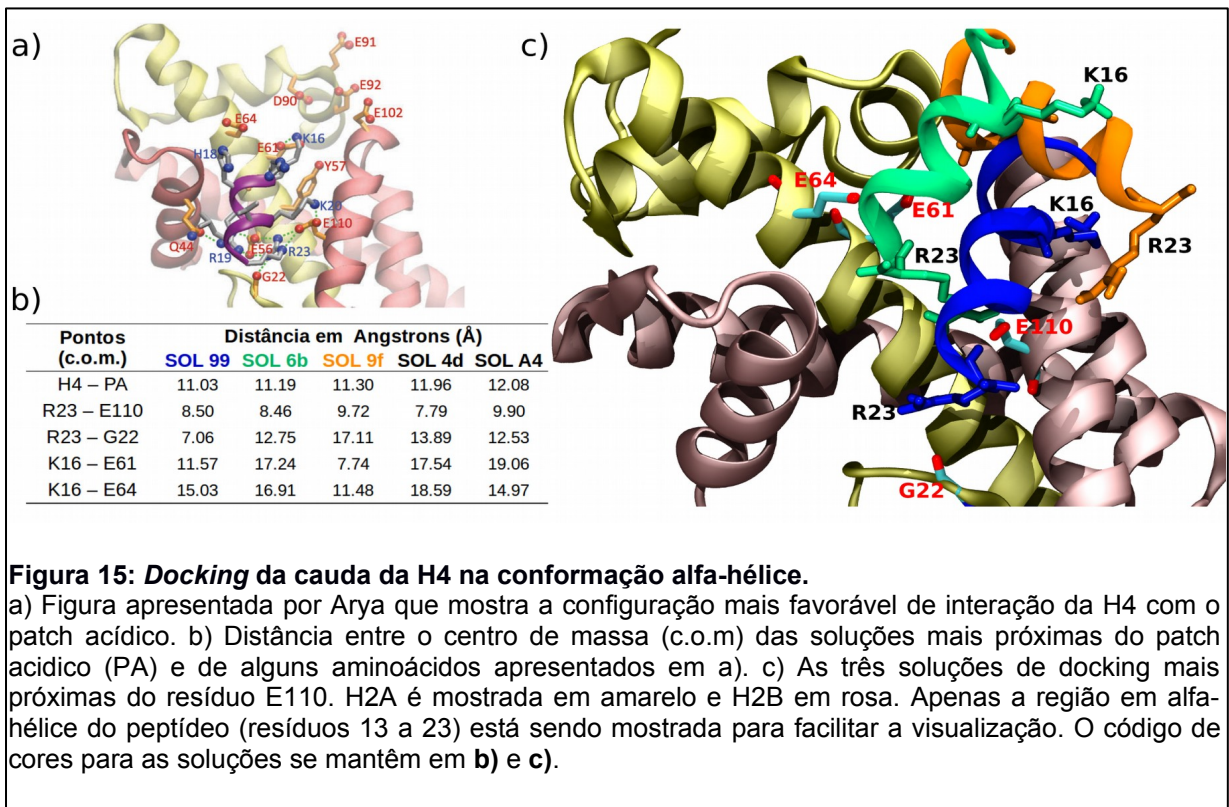
Figura 14: Clusters conformacionais.

a) RMSD ao longo da trajetória do GMIP1 e CtrlNeg como mostrado na Figura 13, porém destacando as estruturas de acordo com a clusterização. O intervalo de tempo selecionado para o *docking* está realçado com um quadrado pontilhado. **b)** *Snapshots* dos grupos encontrados na análise. O último *cluster* (5) é formado pelas conformações que não encaixaram nos demais grupos.

II.4.1.2 *Docking* dos NBPePs ao nucleossomo

Após a definição das estruturas a serem utilizadas no *ensemble docking* dos peptídeos, seguiram-se experimentos de *docking* molecular dos peptídeos GMIP1, H4-hx e CtrlNeg. O modo de ligação do peptídeo da cauda da H4 na conformação de alfa-hélice (H4-hx) com o nucleossomo foi proposto por Yang e Arya (YANG; ARYA, 2011) e foi utilizado como referência nesse estudo (Figura 15a).

Após o agrupamento dos resultados do *docking* em nuvens de soluções, 17 modos de interação da H4-hx com o nucleossomo foram obtidos. A Figura 15b exibe as soluções que mais se aproximaram do *patch* ácido (pelo menos 12.50 Å de distância entre os centros de massa). Tais soluções foram comparadas ao modelo de ligação referência mediante verificação da distância dos aminoácidos G22, E61 e E64 da H2A e E110 da H2B em relação a K16 e R23 da H4-hx. Nenhuma das soluções encaixou-se exatamente da mesma forma que a identificada por Arya. Porém a sol99 foi a que mais se aproximou da orientação utilizada como referência (Figura 15c).



Para o GMIP1, foram obtidos oito modelos de encaixe com o nucleossomo após as definições das nuvens de soluções. Os modelos foram comparados com o modo de ligação do RCC1, no qual o desenho do GMIP1 foi baseado. Apenas em duas soluções de *docking* a orientação estava similar à do RCC1. Entretanto, conforme mostrado pela Figura 16a, nenhuma manteve exatamente a forma de ligação do RCC1. A nuvem de soluções 141 foi a que mais se aproximou do *patch* ácido e por isso foi escolhida para as análises posteriores.

Já para o controle negativo, dentre as 15 nuvens de solução encontradas, nenhuma foi encontrada a 12.50Å, mesmo raio de corte do H4-hx, do *patch* ácido. Assim o raio de corte foi ampliado para 25Å de distância entre o centro de massa do peptídeo e do *patch* ácido. Assim, três nuvens de soluções foram observadas (Figura 16b). Porém a nuvem de soluções 21 foi claramente a que apresentou configuração com maior provavelmente de desenvolver contatos com o *patch* ácido. Aliás, nessa solução o resíduo H9 situa-se a 5Å do centro de massa do *patch* ácido (Figura 16b).

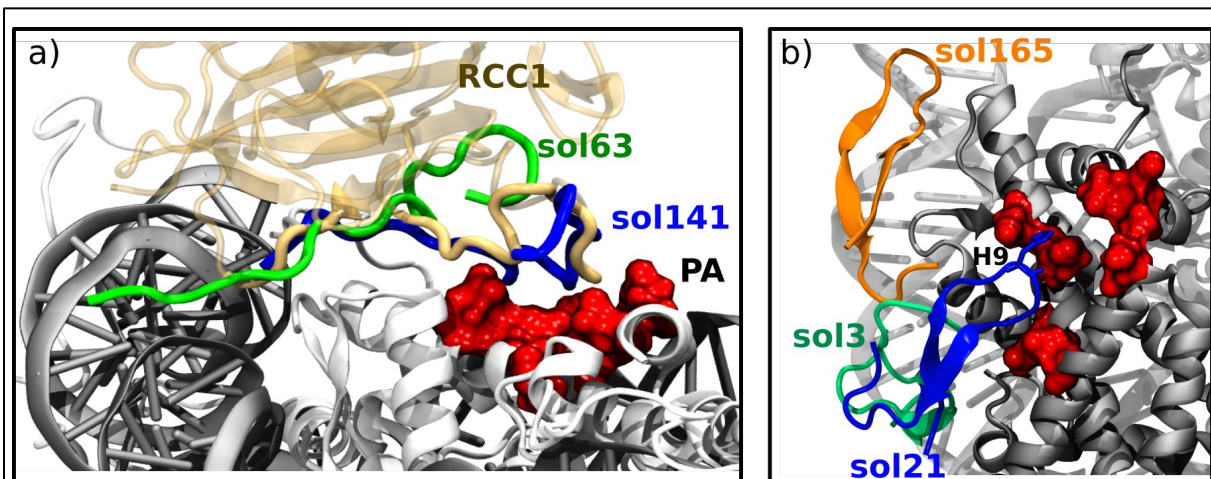
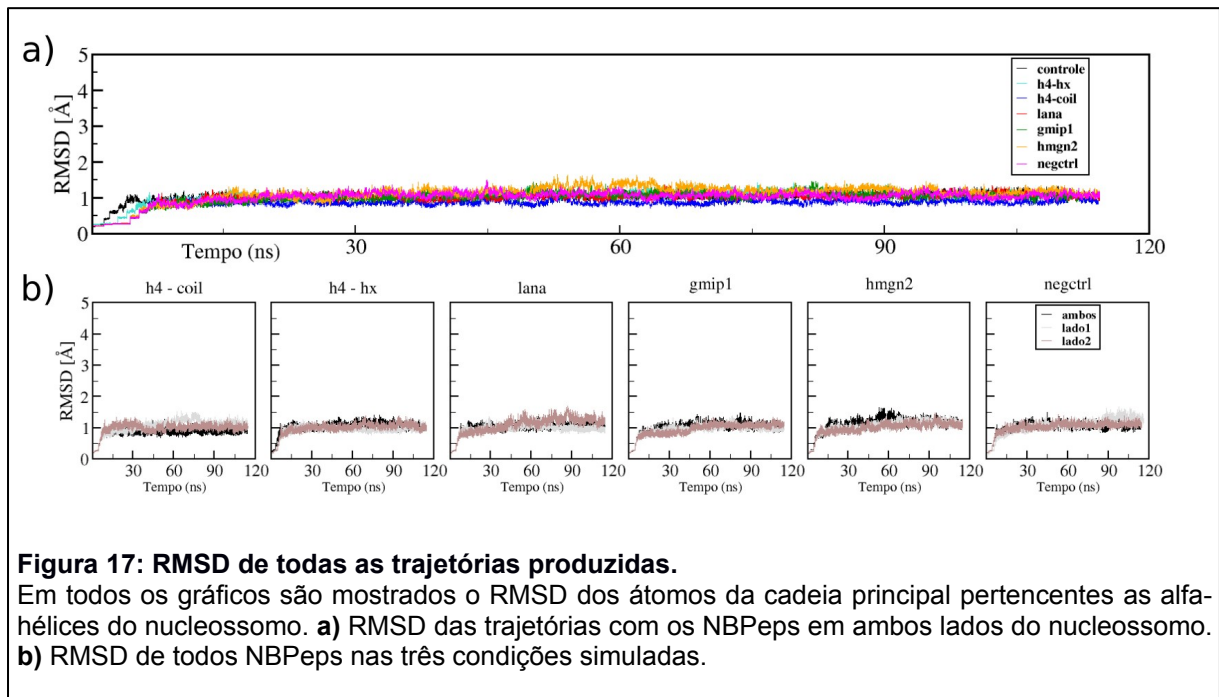


Figura 16: Docking GMIP1 e CtrlNeg

a) GMIP1: Forma de interação do RCC1 (beije) com o nucleossomo (histonas em branco e DNA em prata) observada por (MAKDE et al., 2010, p. 1). Os aminoácidos do RCC1 escolhidos para o desenho do GMIP1 são destacados região beje próxima ao PA). Também é mostrado a melhor solução de *docking* do GMIP1 (azul) no *patch* ácido (PA, vermelho). **b)** CtrlNeg: Representação das histonas H2A e H2B (cinza), do PA (vermelho) e das soluções de *docking* do CtrlNeg que mais se aproximaram do PA: sol21, com H9 destacado (azul), sol3 (verde) e sol165 (laranja).

II.4.2 Estabilidade das simulações

Após as trajetórias dos NBPEps, em todas as três condições, completarem 115 ns algumas análises estruturais foram realizadas. Inicialmente foi verificada a estabilidade conformacional das trajetórias. Para isto, foi avaliado o RMSD dos átomos pertencentes às alfa-hélices do nucleossomo. Como pode ser observado na Figura 17, o perfil do RMSD de todas as trajetórias converge para um platô no início das simulações. Após 15 ns de simulação, o nucleossomo em complexo com NBPEps em ambos os lados exibe um perfil de RMSD sem grandes variações e com valores baixos, indicativo de equilíbrio estrutural (Figura 17a). Esse mesmo perfil é observado em cada simulação com apenas um NBPEp (Figura 17b). Deste modo, para as análises posteriores, apenas os 100 ns finais das trajetórias foram avaliados.



II.4.3 Mapa de contatos

Após verificada a estabilidade estrutural do nucleossomo, a primeira análise realizada foi a definição do mapa de contatos de cada NBPeP com o nucleossomo. Para isto, foi criado um código de análise que definiu os aminoácidos do nucleossomo que formam os sítios de ligação para cada NBPeP, de acordo com os critérios listados a seguir. Um resíduo de aminoácido do nucleossomo é considerado em contato com um NBPeP caso estes estejam distantes em até 3,5 Å. Posteriormente, é calculada a probabilidade média de contato para cada resíduo levando em conta os 100 ns de trajetória de equilíbrio. Foram considerados como pertencentes aos sítios de ligação, apenas aqueles resíduos de aminoácidos com probabilidade maior que 20 % de contato ao longo da trajetória. Em termos de tempo de simulação, essa probabilidade se traduz em 20 ns ou mais dentro do raio de corte.

Como cada NBPeP foi simulado em três condições diferentes em complexo com o nucleossomo, seis mapas de contato para cada NBPeP foram produzidos. Os quatro primeiros são os sítios específicos de cada NBPeP em cada condição de

simulação. Exemplo, o LANA simulado do lado 1 forma um sítio, simulado do lado 2 forma outro sítio e simulado em ambos os lados forma-se mais dois sítios, totalizando quatro sítio “específicos”. O quinto sítio de ligação, com a alcunha de “combinado” é a soma de todos os aminoácidos para um determinado NBPeP que fizeram parte de qualquer um dos quatro sítios anteriores, os “específicos”. E os sítios “regulares” são formados pelos aminoácidos que interagiram com o mesmo NBPeP pelo menos em três dos sítios específicos.

Com o mapa de contato completo (Tabela 2), foram analisados os sítios de interação de cada NBPeP. Em todos os casos, pelo menos um contato com algum aminoácido do *patch* ácido (E64 da H2A) é mantido. É interessante o fato de que, com exceção do GMIP1, os resíduos E61, D90 e E92 da H2A também demonstraram alta frequência de interação entre os NBPePs. Também foram observados pontos de interação do HMGN2pep e H4-coil com a H3 e H4. Aliás, esses dois NBPePs e o GMIP1 também apresentaram contatos com o DNA. O LANA foi o peptídeo que apresentou contatos mais conservados com o nucleossomo. Praticamente em todas as simulações esse ligante interage com os mesmos aminoácidos. A cauda da H4 na conformação em alfa-hélice obteve sítios de interação similares ao LANA, porém com um número menor de contatos conservados. Já a cauda da H4 na conformação aberta, apesar de interagir com a H3 e o DNA, apresentou seis contatos regulares com o *patch* ácido. Mesmo número de contatos regulares com o *patch* ácido do LANA e da H4-hx. O peptídeo HMGN2 foi o que fez contatos com o maior número de aminoácidos do nucleossomo. O GMIP1 apresentou apenas um contato frequente com o *patch* ácido e foi o segundo com a maior distribuição de interações, evidenciado pelos sítios combinados.

A mesma análise efetuada com os aminoácidos do nucleossomo, foi feita com os aminoácidos pertencentes aos NBPePs. Assim, também foi criado um mapa com os aminoácidos dos NBPePs que mais se aproximaram do nucleossomo, utilizando os mesmo critérios da análise anterior (Tabela 3).

Tabela 2. Aminoácidos do nucleossomo formadores das cavidades de cada NB Pep.
Em vermelho estão os aminoácidos do *patch* ácido.

NBM	L	H2B						H2A						H3	H4	DNA1				DNA2							
LANA	1	Q44	V45		L103	H106	E110		A53	E56	Y57	A60	E61	E64													
LANA	2	Q44	V45		L103	H106	E110		A53	E56	Y57	A60	E61	E64													
LANA.I1	1	Q44	V45		E102	L103	H106	E110		A53	E56	Y57	A60	E61	E64												
LANA.I2	2	Q44	V45		L103	H106	E110		A53	E56	Y57	A60	E61	E64													
combinado regular		Q44	V45		E102	L103	H106	E110		A53	E56	Y57	A60	E61	E64												
H4-hx	1	V41	Q44	V45		E102		E110	G22	Q24	E56		E61	E64			N89	D90	E91	E92							
H4-hx	2	Q44	V45	H46				E110	K113	E56		A60	E61	E64			N89	D90	E91	E92							
H4-hx.I1	1	Q44					H106					E61	E64				D90	E91	E92								
H4-hx.I2	2	Q44	V45			E102		E110		G22	E56		E61	E64				D90	E91	E92							
combinado regular		V41	Q44	V45	H46		E102	H106	E110	K113	G22	Q24	E56	A60	E61	E64		N89	D90	E91	E92						
		Q44	V45				E102		E110				E56		E61	E64			D90	E91	E92						
H4-coil	1		V45			E102	H106				A60	E61	E64	L65	A69	D72		N89	D90	E91	E92		K118				
H4-coil	2					E102		E110			E56		E61	E64		D72		N89	D90	E91	E92						
H4-coil.I1	1		V45			E102	H106	E110				E61	E64					N89	D90	E91	E92		Q125	R134	E53		
H4-coil.I2	2	Q44	V45			E102	H106	E110			E56	A60	E61	E64				N89	D90	E91	E92		R134				
combinado regular		Q44	V45			E102	H106	E110			E56	A60	E61	E64	L65	A69	D72		N89	D90	E91	E92		K118	Q125	R134	E53
		V45				E102	H106	E110					E61	E64				N89	D90	E91	E92						
HMGN2	1											E61	E64		D72		N89	D90	E91	E92	P109	R52			138T 139A 141A 142T		
HMGN2	2					E102	H106	S109	E110			E61	E64	L65	R71	D72		N89	D90	E91	E92					9T 10A 11T	
HMGN2.I1	1					E102						E61	E64		D72		N89	D90	E91	E92	P109	N110	I111	V114	140T 141T		
HMGN2.I2	2											E61	E64		D72		N89	D90	E91	E92	P109				10C 11C 12C		
combinado regular						E102	H106	S109	E110			E61	E64	L65	R71	D72		N89	D90	E91	E92	P109	N110	I111	V114	9T 10A 11T 140T 141T	
												E61	E64		D72		N89	D90	E91	E92	P109	R52	K56		10C 11C 12C 138T 139A 141A 142T		
															D72		N89	D90	E91	E92	P109						
GMIP1	1		V45	H46								R71		K74												12A 13T	
GMIP1	2				D48		H106		E110			E64	N68					N89									
GMIP1.I1	1	Q44	V45	H46	P47	D48						E61	E64	R71		K74	K75		N89	D90	E91	E92				12A 13T	
GMIP1.I2	2						H106					E61	E64	N68		D72		N89	D90	E91	E92				10C 11C		
combinado regular		Q44	V45	H46	P47	D48		H106	E110			E61	E64	N68	R71	D72	K74	K75	N89	D90	E91	E92				10C 11C	
												E64														12A 13T 140T 141T	

Tabela 3. Aminoácidos dos NB Peps responsáveis pela interação com o nucleossomo.
Em azul estão as argininas (escuro) e lisinas (claro).

LANA comb.	G5	M6	R7	L8	R9	S10	R12	T14	G15	A16											
LANA reg.		M6	R7	L8	R9		R12	T14	G15												
H4-hx comb.	R3	G4	K5					K8		K12	G13	G14	A15	K16	R17	R19		V21	L22	R23	
H4-hx reg.	R3															R19				R23	
H4-coil comb.	R3	G4	K5	G6	G7	K8	G9	K12			G14	A15	K16	R17	R19	K20	V21	L22	R23		
H4-coil reg.	R3	G4	K5					K12					K16		R19					R23	
HMGN2 comb.	D18	E19	P20	R22	R23	R26	L27	A29	K30	P31	A32	P33	P34	K35	E37	K39	P40	K41	K42		
HMGN2 reg.				R22	R23	R26				P31	A32	P33			E37	K39	P40				
GMIP1 comb.	R1	T2	I3	A5	A6	A7	L8	S9	R11	S12	I13	R18	R19								
GMIP1 reg.	R1		I3		A6	A7		S9	R11			R18									

Notavelmente, em todas as simulações, os NBPEps interagiram com o nucleossomo por meio de três argininas. A HMGN2 foi o NBPEp que apresentou o maior número de pontos de contato com o nucleossomo (nove), provavelmente por ser o mais extenso dos peptídeos. Os demais NBPEps, apresentaram sete aminoácidos interagindo regularmente com o nucleossomo. Salvo a H4-hx, que nas quatro condições avaliadas conservou apenas três argininas como pontos de interação.

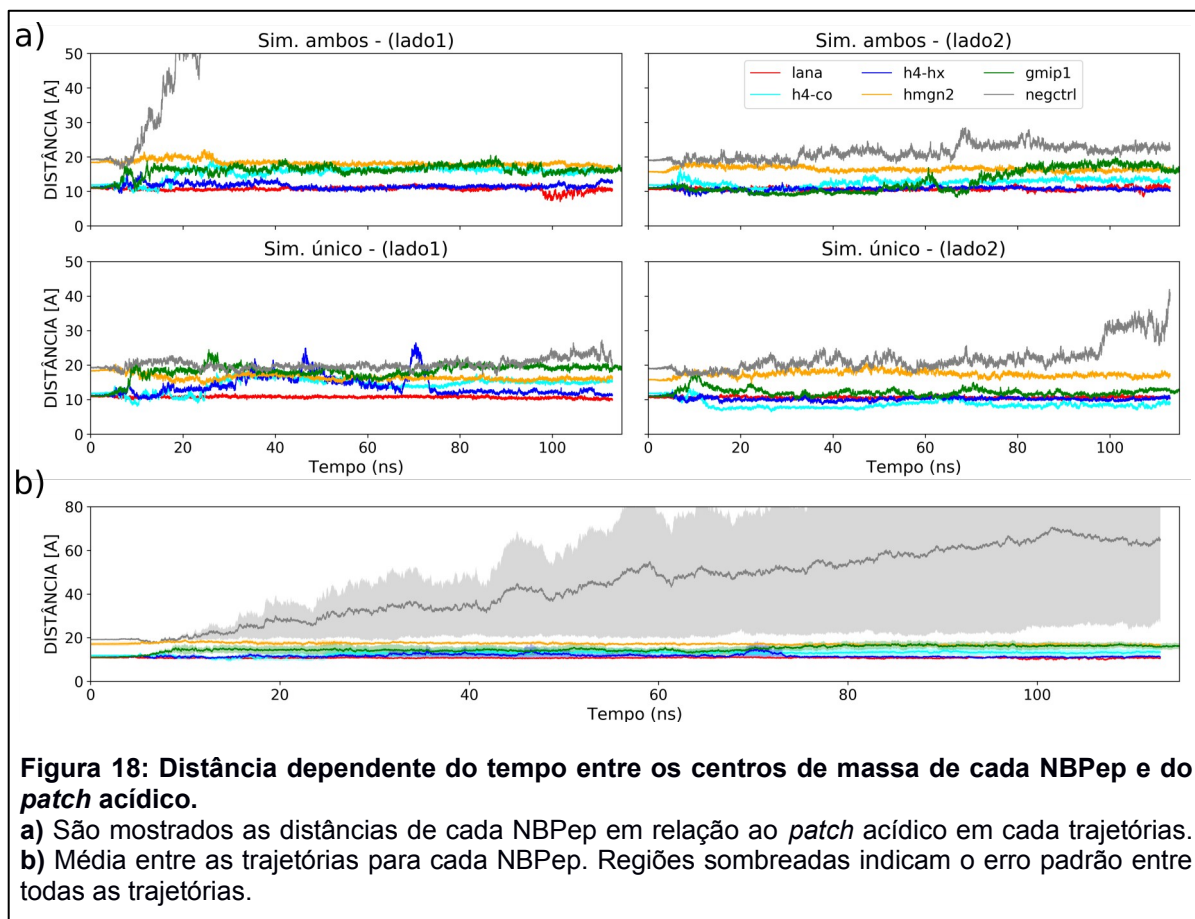
II.4.4 Dinâmica dos NBPEps no sítio

Para observar o comportamento dos NBPEps ao longo das trajetórias foi investigada a movimentação de cada NBPEp ao longo do tempo. Desta forma, foi calculada a distância dependente do tempo de cada ligante em relação ao centro de massa do *patch* ácido (Figura 18). A posição dos NBPEps também foi definida pelo centro de massa. Antes das distâncias serem calculadas, as estruturas foram sobrepostas, empregando o nucleossomo como referência, para evitar erros provenientes de rotação e translação.

Analisando as trajetórias individuais (Figura 18a), observa-se que o peptídeo CtrlNeg afasta-se do nucleossomo rapidamente em uma das simulações. Nas demais trajetórias, esse peptídeo também demonstra distanciamento do *patch* ácido. De forma que, após 80 ns de simulação no lado 2 do nucleossomo o CtrlNeg distancia-se em aproximadamente 30Å do nucleossomo. Os demais NBPEps permaneceram a pelo menos 20Å do *patch* ácido em todas as condições de simulação. O peptídeo LANA apresentou o perfil de distância mais próximo do nucleossomo em todas as trajetórias. Por sua vez, os peptídeos com estruturas conformacionais mais flexíveis (HMGN2, H4-coil e GMIP1) demonstraram maiores oscilações na distância em relação a região ácida do nucleossomo. Porém, a cauda da H4 na conformação aberta (H4-coil), na trajetória em que estava presente somente no lado 2 do nucleossomo, demonstrou perfil mais próximo do nucleossomo, mesmo comparado ao LANA.

Observando o perfil geral de todos NBPEps em todas as simulações é possível perceber que em algum ponto é atingido um platô de distância (Figura 18b).

Isto sugere que tais peptídeos estejam transitando para um estado de equilíbrio com o nucleossomo. Salvo o caso do CtrlNeg, no qual fica evidente a falta de afinidade com o nucleossomo.



Entretanto, dado que ambos peptídeo da cada da H4, GMIP1 e HMGN2 são constituídos por um número de aminoácidos maior do que o LANA e os dois últimos ainda realizam contatos com o DNA, o perfil de distância calculado levando em conta os peptídeos completos pode não ser suficiente para descrever a interação. Assim foi verificada a distância dos quatorze aminoácidos de cada NBPeP responsáveis pela interação com o *patch* ácido (Figura 19). O valor de quatorze resíduos foi escolhido por ser a totalidade de aminoácidos do LANA, e a escolha dos aminoácidos foi baseada na análise anterior do mapa de contatos. As principais alteração foram observadas para os NBPePs H4-coil e HMGN2pep. Ambos demonstraram maior proximidade do *patch* ácido. H4-coil, aliás, com perfil mais

próximo comparado ao LANA. No entanto, para GMIP1 não foram observadas alterações nessa nova análise, este manteve-se o mais distante do *patch* ácido de todos os NBPePs. A investigação da dinâmica dos NBPePs sugere que esses peptídeos estão de fato interagindo com os aminoácidos do nucleossomo e mantendo essa interação no período de tempo analisado. Em seguida, foi avaliado se os NBPePs modificaram a hidratação do *patch* ácido.

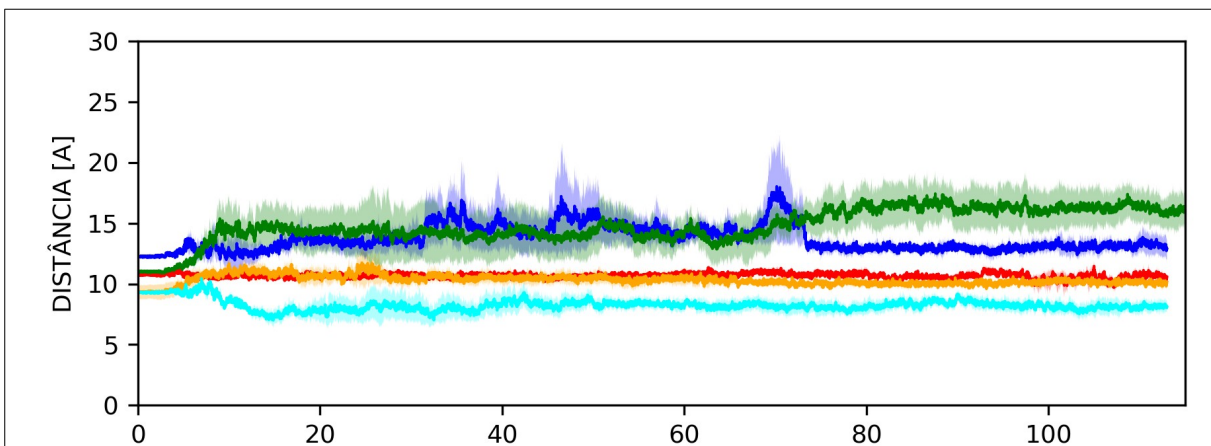


Figura 19: Distância dependente do tempo entre regiões específicas dos NBPePs e o *patch* ácido

Distância média entre as trajetórias para cada NBPeP. Os resíduos escolhidos para os NBPePs foram: 6 ao 19 GMIP1 (verde); 10 ao 23 H4-coil (ciano) e H4-hx (azul), 18 ao 31 HMGN2 (laranja), e todos do LANA (vermelho).

II.4.5 Hidratação do *patch* ácido

Para avaliar a influência dos NBPePs no perfil de hidratação do *patch* ácido foi computado o número médio de moléculas de água próximas ao *patch* ácido. Para tanto, foram criadas esferas com raios de 14 e 20Å a partir do centro do *patch* ácido e calculada a quantidade média de moléculas de água presente em cada esfera ao longo das trajetórias de equilíbrio, para cada NBPeP e sistema controle. Em seguida, foi computada a mudança proporcional no número de moléculas de água dos sistemas NBPePs em relação ao sistema controle (Figura 20).

Todos NBPePs, salvo GMIP1, reduziram a quantidade de água no *patch* ácido em pelo menos 25% na esfera de menor raio. A diferença entre os valores

observados para esses NBPePs é pequena e correlaciona parcialmente com o perfil de distância do *patch* ácido. Para GMIP1 é possível perceber que removeu uma menor quantidade de água condizente com seu perfil mais distante do *patch* ácido. HMGN2pep e LANA em ambas análises demonstraram resultados semelhantes e H4-coil, peptídeo mais próximo do *patch* ácido deslocou cerca de 30% das moléculas de água dessa região, maior redução para qualquer NBPeP. Apesar que na esfera com raio de 20Å foi observada maior redução para H4-hx, provavelmente devido a conformação em alfa-hélice ocupar um maior volume na esfera de 20Å. Assim, foi verificado que os NBPePs mantiveram contato próximo com *patch* ácido e deslocaram moléculas de água dessa interface. Em seguida foi avaliado o impacto estrutural promovido pelos NBPePs no nucleossomo.

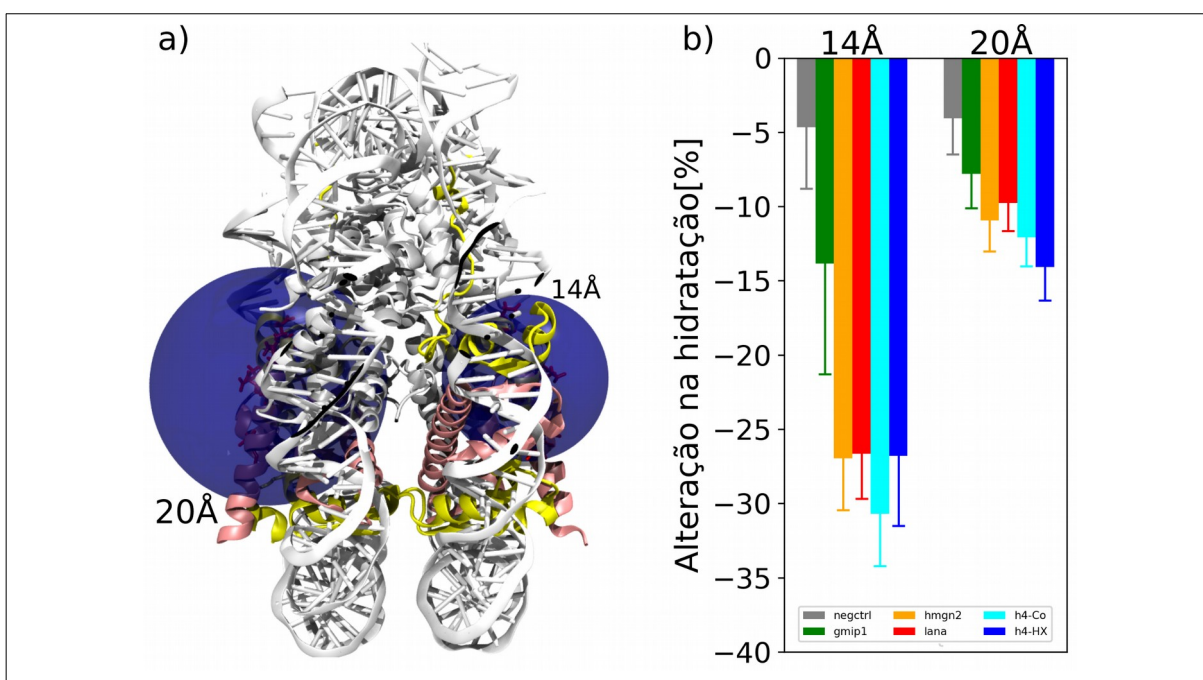


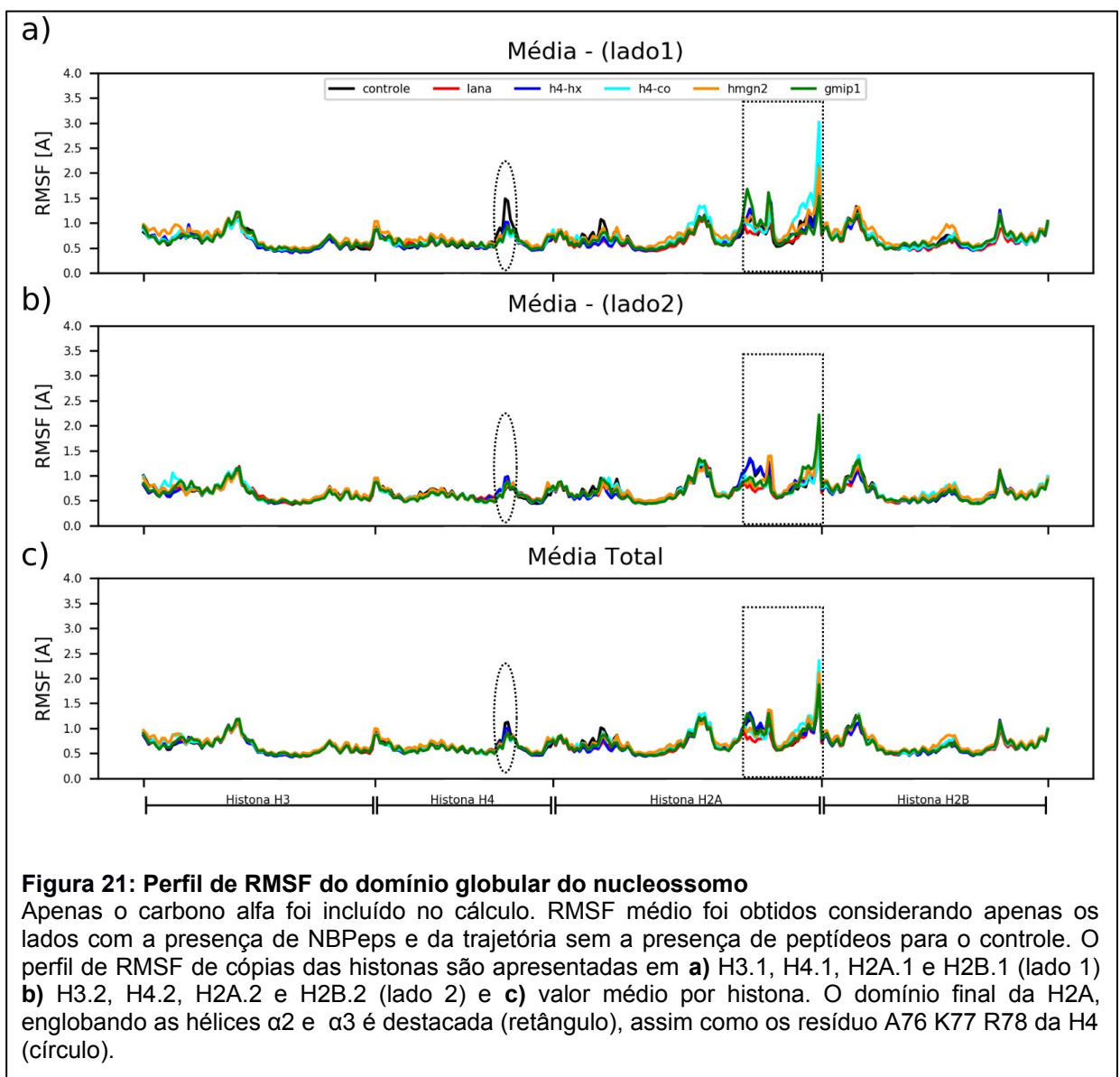
Figura 20: Alteração na hidratação do *patch* ácido.

A quantidade de moléculas de água foi determinada utilizando esferas como limite de contagem. **a)** Representação das esferas criadas de 14Å e 20Å (roxo) em torno do *patch* ácido. Nucleossomo (branco), histonas H2A (amarelo), H2B (rosa) e aminoácidos do *patch* ácido (vermelho) **b)** Foi calculada a alteração proporcional para cada NBPeP em relação a ao sistema controle. Barras indicam o erro padrão.

II.4.6 Impacto estrutural

Para avaliar o impacto estrutural causado pelos peptídeos foi computado o RMSF de todos sistemas após o tempo de equilíbrio (últimos 100 ns). O RMSF é

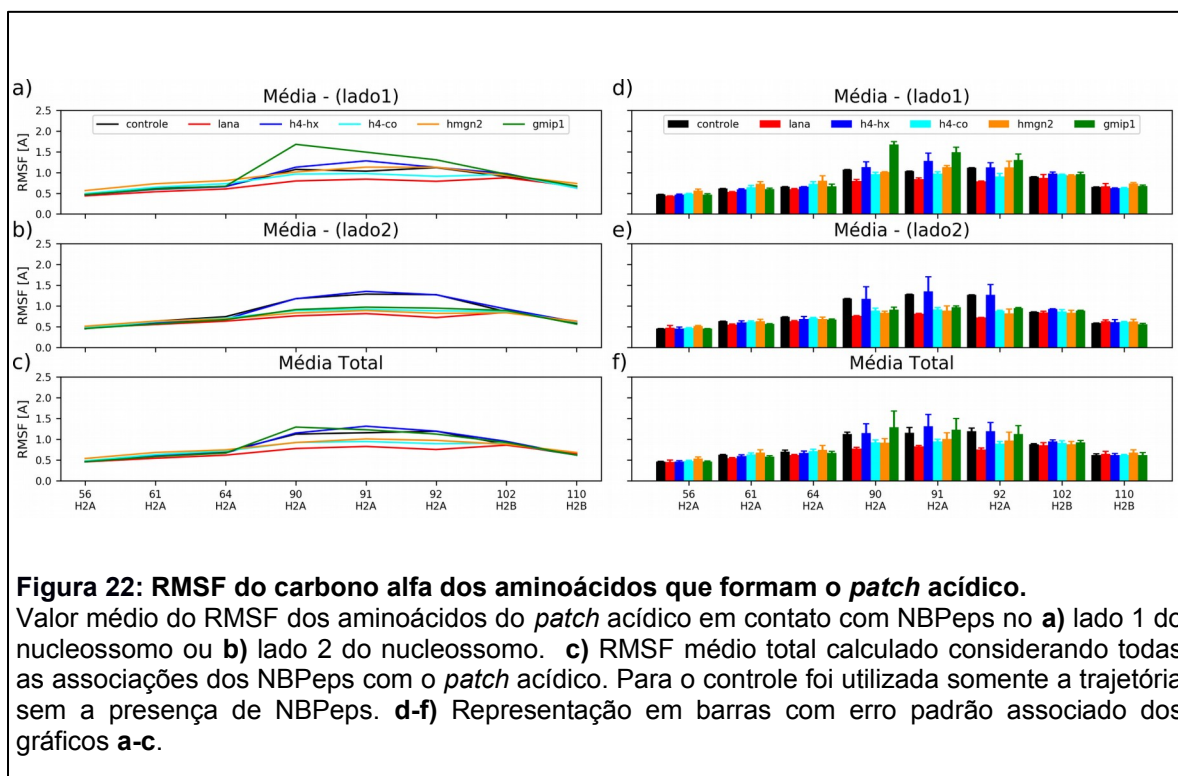
um desvio que permite quantificar a movimentação dos átomos do sistema. Assim consegue identificar regiões flexíveis e, no caso da comparação dos valores na presença e ausência dos NBPEps, descrever impactos estruturais locais. Visto que essa análise se dá por átomo, apenas o carbono alfa das histonas e o fosfato do DNA foi utilizado nos cálculos. Além disso, foram considerados somente os átomos do domínio central globular das histonas, pois parte das caudas foram removidas do nucleossomo e, mesmo os segmentos da cauda que permaneceram nos sistemas, normalmente apresentam alta flexibilidade. Foi computado o RMSF médio para ambos os lados do nucleossomo, assim como o valor médio total (Figura 21).



Observando o perfil do RMSF dos lados 1 e 2 do nucleossomo (Figura 21a-b), as histonas H3 e H4 não sofreram alterações significativas, exceto os resíduos A76 K77 R78 da H4. Esses resíduos estão na interface de interação entre a proteína e o DNA e apresentaram flutuações reduzidas na presença de NBPEps. Apesar dessa redução ser observada apenas no lado 1 do nucleossomo. O perfil do RMSF médio, considerando ambos lados do nucleossomo e as três condições de simulação, indica pequenas alterações nas flutuações atômicas locais (Figura 21c). Aparentemente, H4-coil foi o NBPEp que mais afetou a histona H2A e o HMGN2pep a histona H2B. Também foram observadas variações na região final da H2A nas duas cópias para quase todos os NBPEps, exceto H4-hx (Figura 21c). Nessa região está localizada a maior parte dos aminoácidos do *patch* ácido. Em vista disso, foi analisado o RMSF apenas do aminoácido do *patch* ácido (Figura 22).

Primeiramente, é nítido que o *patch* ácido na ausência de ligantes apresenta diferenças de flutuações atômicas entre os lados do nucleossomo. Os aminoácidos E90, E91 e E92 apresentam uma movimentação maior no lado 2 do nucleossomo. Também é possível perceber que esses três aminoácidos foram os mais afetados pelos NBPEps (Figura 22c,f), e assim as análises focaram nesses resíduos.

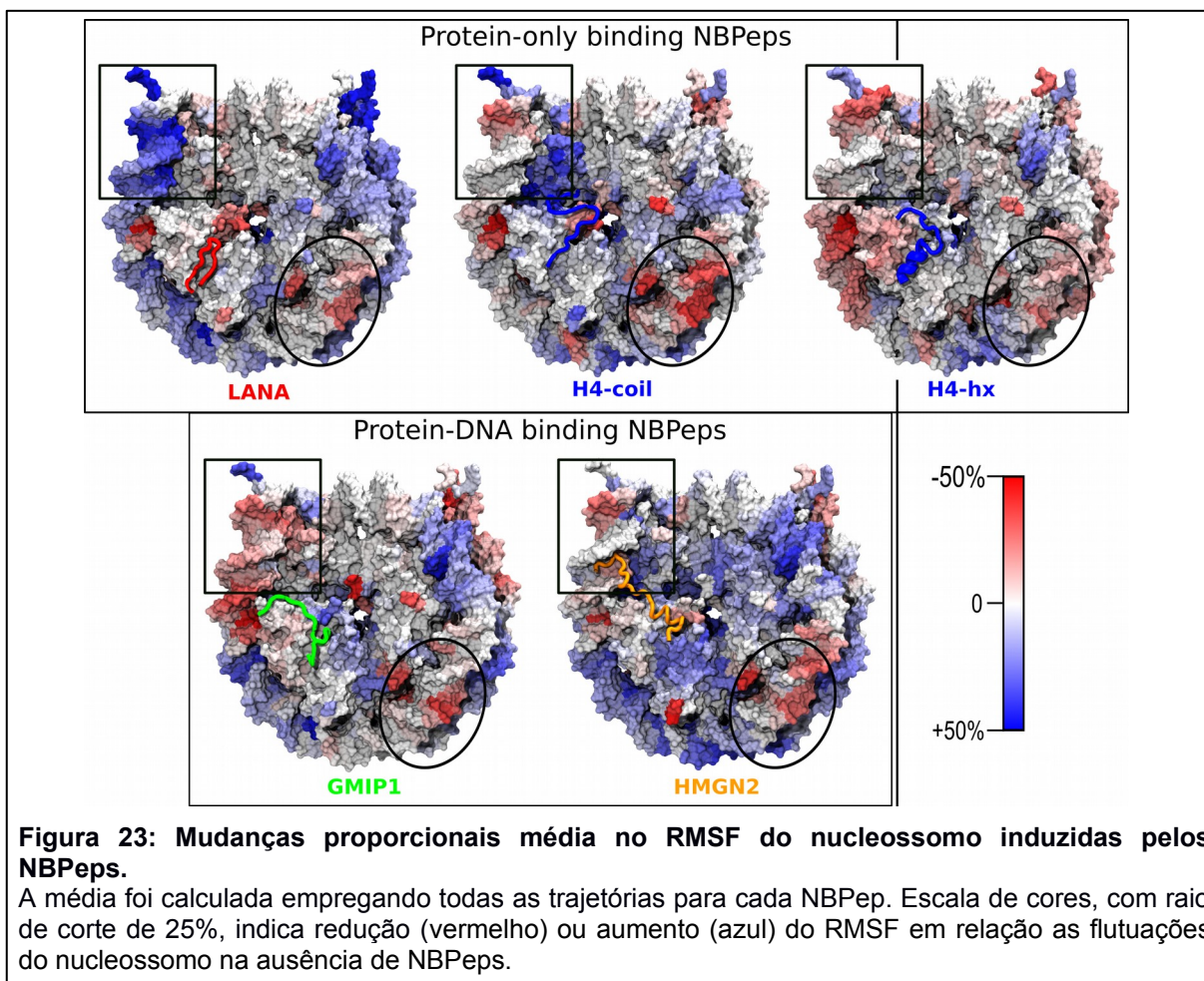
O GMIP1 exibiu resultados mais divergentes entre os lados do nucleossomo, pois no lado 1 foi observado aumento do RMSF e lado 2 redução. No entanto, para LANA os resultados foram uniformes, revelando redução das flutuações atômicas tanto nas situações locais (lado 1 e lado 2, Figura 22 a-b,d-e) quanto observando o impacto geral (Figura 22c,f). Os NBPEps H4-coil e HMGN2pep também limitaram a movimentação dos aminoácidos E90, E91 e E92 uniformemente, porém em menor intensidade do que o LANA. O peptídeo da cauda da H4 na conformação em alfa-hélice (H4-hx) apresentou tendência de aumento das flutuações atômicas. Entretanto, o ruído demonstrado pelas barras de erro foi grande em ambos os lados do nucleossomo.



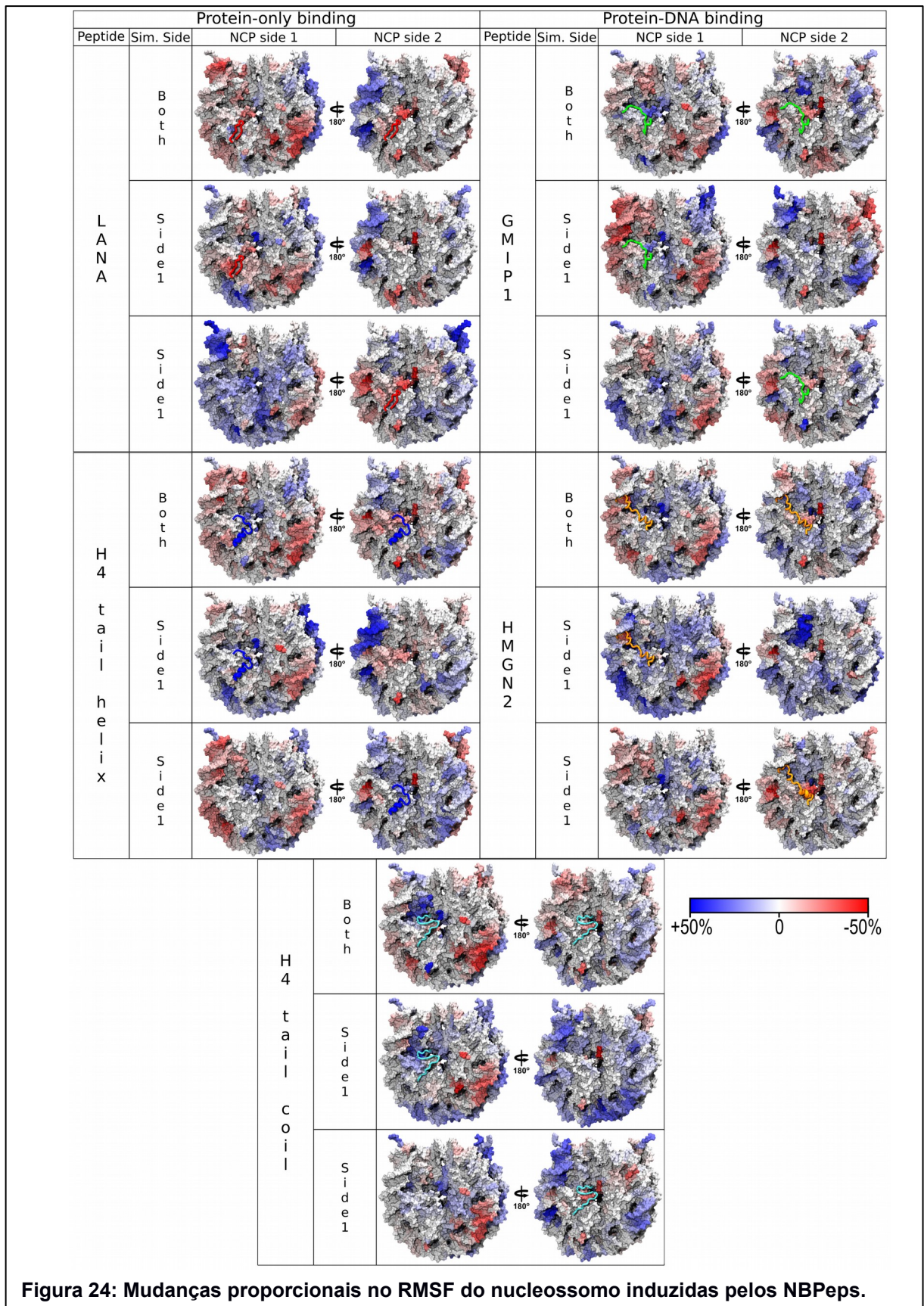
A fim de obter um panorama mais geral sobre o impacto dos NBPEps nas flutuações atômicas do nucleossomo foi criado um *heatmap*, por aminoácido, das alterações nos valores de RMSF do nucleossomo (Figura 23,24). Para isso, foi computada a diferença proporcional entre os valores de RMSF na ausência e na presença de cada NBPEp. Logo, para cada átomo foi calculado um valor, em porcentagem, de alteração do RMSF em relação ao controle. Esses dados foram representados no nucleossomo com um gradiente das cores vermelho, branco e azul. Valores negativos, exibidos em tons de vermelho, refletem uma redução da medida do RMSF causado pelos NBPEps. Valores positivos, representados em tons de azul, indicam um incremento nas flutuações provocado pelos NBPEps. Para facilitar a visualização, foi usado um raio de corte máximo para as alterações de 25% (Figura 23) ou 50% (Figura 24).

O mapa é bem variado entre os NBPEps, portanto, é difícil encontrar padrões mesmo para um único NBPEp simulado em condições diferentes. No entanto, duas regiões merecem destaques. A primeira é a interface de interação entre o DNA e as histonas, destacados com um círculo na figura, onde os aminoácidos A76 K77 R78

da H4 já mencionados anteriormente estão localizados. Essa região apresenta redução das movimentações atômicas no lado 1 do nucleossomo. É possível perceber que o DNA nesse local também foi afetado pela ligação de todos NBPeps, no entanto, com intensidade díspares.



A H4-coil foi o NBPep que apresentou maior impacto nessa região e o GMIP1 foi o mais brando. A segunda região que pode ser destacada é o final do DNA no lado 2 do nucleossomo (representada por um retângulo). Essa região tem sua movimentação atômica levemente aumentada pelos NBPeps que interagem somente com o *patch* ácido (LANA e H4). Como esperado GMIP1 e HMGN2, que apresentam pontos de contato no DNA também, reduziram as flutuações atômicas dessa região do DNA em ao menos uma simulação. No entanto, a HMGN2, quando simulado apenas no lado 1 teve o efeito contrário. As flutuações do DNA próximo ao peptídeo aumentaram na presença desse NBPep.

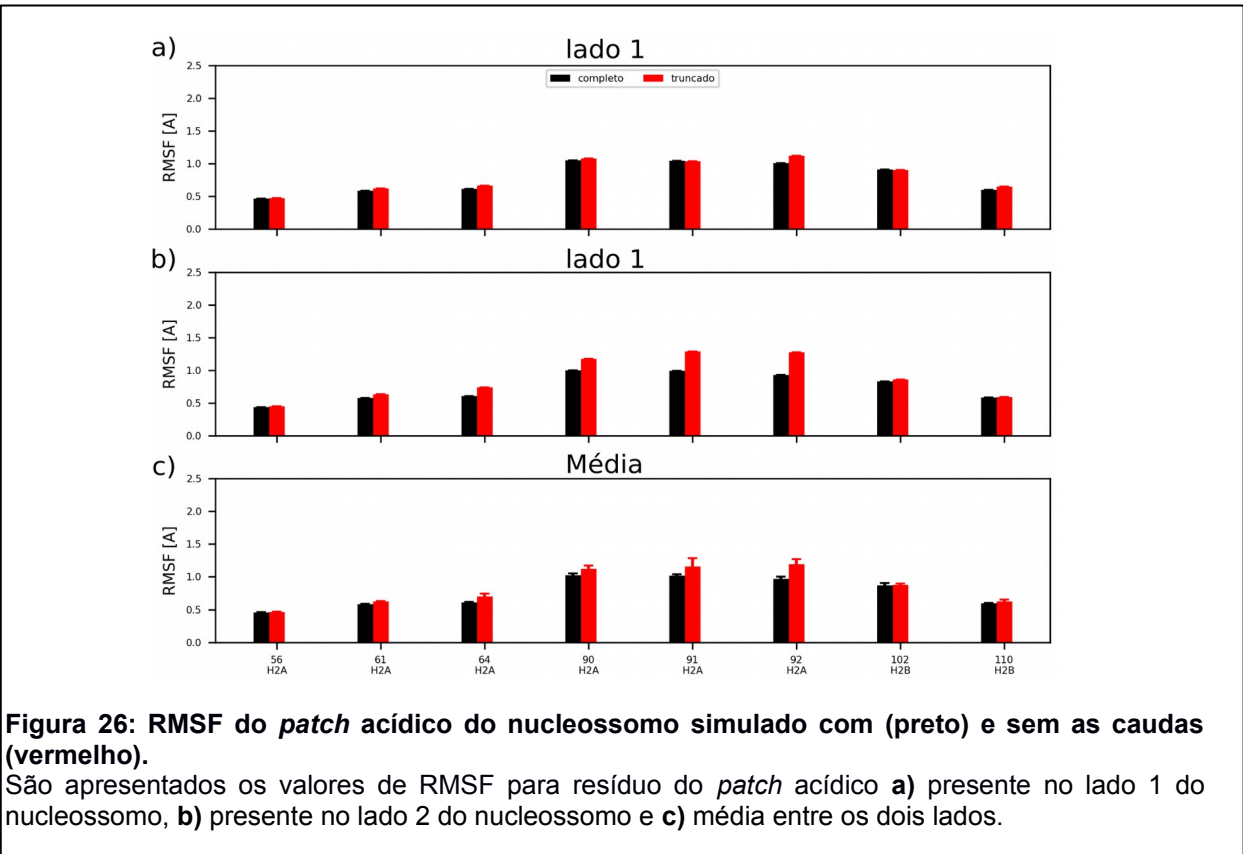
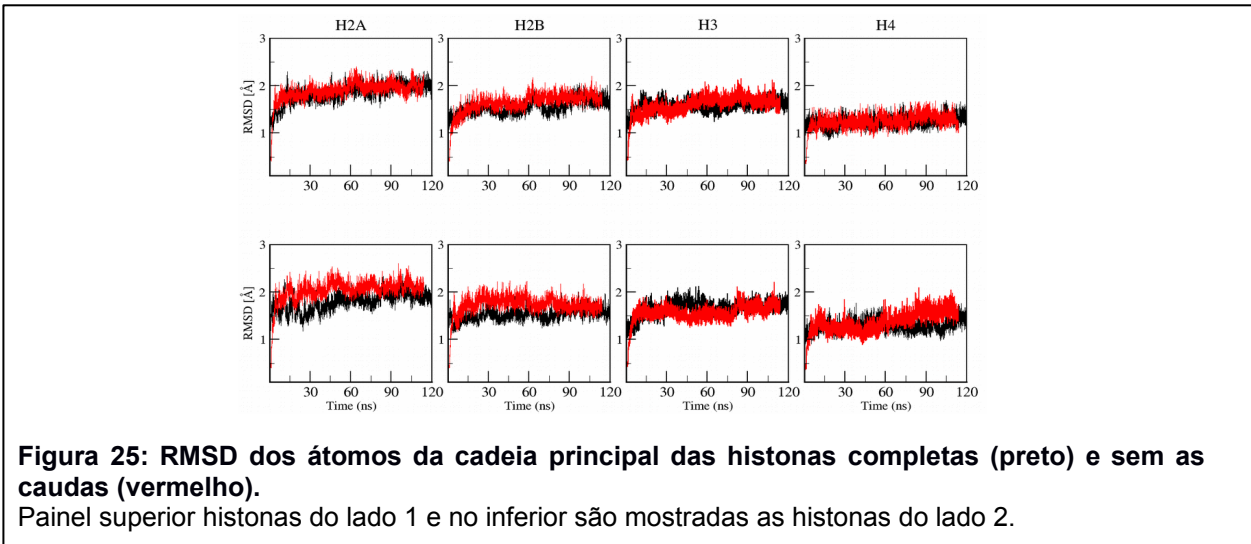


II.4.7 Considerações gerais e discussão

Moléculas ligantes do nucleossomo são amplamente estudadas no contexto da definição dos estados da cromatina. Como já mencionado, várias NBPs já foram descritas (MCGINTY; TAN, 2016; ZHOU; GAULLIER; LUGER, 2019) destacando a superfície do nucleossomo como uma ancora para a ligação de fatores nucleares. Tais fatores regulam a arquitetura da cromatina e conseqüentemente, o acesso de proteínas ao DNA. Dessa forma, no que diz respeito aos peptídeos NBPeps, foram identificados alguns aspectos sobre a dinâmica da interação desses peptídeos com o nucleossomo. Para que esses resultados fossem obtidos várias abordagens computacionais foram empregadas, as quais vão desde a montagem dos sistemas iniciais, passando pela simulação de DM de tais sistemas e também as análises das trajetórias.

Como descrito na metodologia, as caudas do nucleossomo utilizado como receptor dos NBPeps foram removidas. A remoção dessas regiões possibilitou um redução de aproximadamente 100.000 átomos para cada sistema. Essa diferença do número total de átomos é significativa na redução do tempo necessário para que as trajetórias fossem obtidas. Visto que cada NBPep, foi simulado em três condições e que no total seis NBPeps foram estudados, o ganho de tempo computacional foi imenso. Provavelmente, o nucleossomo com as caudas demandaria por volta de um mês a mais de tempo computacional por NBPep. Entretanto, como demonstrado Biswas em 2011 a remoção das caudas pode impactar a estrutura do nucleossomo (BISWAS et al., 2011). Assim, a Figura 25 apresenta a comparação do RMSD individual da região globular das histonas com e sem as caudas. As estruturas do nucleossomo com caudas foram adquiridas de uma longa trajetória de equilíbrio (220 ns) produzidas previamente (FERNANDES, 2016; SILVA et al., 2017). O perfil para todas as histonas indica que a ausência das caudas não impediu que o perfil associado ao equilíbrio estrutural fosse alcançado. Somente a H2A do lado 2 apresentou variações, mesmo que pequenas no RMSD em comparação com o sistema com caudas. Dado que, a maioria dos aminoácidos que compõem o *patch* ácido integram a histona H2A, foi comparado também o RMSF (Figura 26) dos resíduos do *patch* ácido com e sem caudas. O nucleossomo simulado com todas

as caudas apresenta perfil levemente mais homogêneo do que o simulado sem as caudas (Figura 26c). A alteração no entanto é pequena, mas é expressa nos aminoácidos mais afetados pelas NBPePs, E90, E91 e E92 da histona H2A, embora perceptível somente no lado 2 do nucleossomo. Isso sugere que esses aminoácidos são mais sensíveis a alterações no nucleossomo.



Um desafio para a modelagem molecular de interações sem modelos atômicos é obter a interface de interação entre as moléculas. Diante disso, com intuito de minimizar possíveis erros na montagem dos sistemas foram buscados dados referenciais que validassem as configurações iniciais dos complexos. Os NBPePs que dispunham de modelos atômicos no PDB (Protein Data Bank) foram o LANA e a cauda da H4-coil. A superfície de contato da HMGN2 com o nucleossomo foi determinada por técnicas baseadas em NMR e análises mutacionais por Kato e colaboradores (KATO et al., 2011). O modelo criado não estava disponível no PDB, contudo, os autores foram solícitos em enviá-lo. Os demais NBPePs, dependeram de dados indiretos. Naturalmente, para o GMIP1, o modo de interação do RCC1 foi crucial para a identificação da configuração inicial a ser simulada. O peptídeo da cauda da H4 na conformação em alfa-hélice foi proposto por Yang e Arya (2011), assim como o modo de interação com o nucleossomo (YANG; ARYA, 2011).

Apenas o peptídeo CtrlNeg não dispunha de dados na literatura acerca da interface de interação com o nucleossomo, pois o mesmo foi desenhado para esse trabalho com o objetivo de não se ligar ao *patch* ácido. Esse peptídeo serve de uma validação crucial para a metodologia utilizada nesse projeto, conferindo maior robustez aos resultados encontrados. Nas análises de *docking* já ficou evidente a tendência desse peptídeo de não interagir com o nucleossomo. Visto que a nuvem de soluções mais próxima do *patch* ácido ficou por volta de 18Å, e mais duas foram encontradas aumentando o raio para 25Å. Esses valores são cerca de duas vezes maiores do que os apresentados para o o demais NBPePs que passaram pela etapa de *docking* (aproximadamente 11Å). Como previsto, o peptídeo CtrlNeg se afasta do *patch* ácido em duas trajetórias (Figura 18). Provavelmente, nas outras duas situações esse peptídeo também se dissociaria do nucleossomo caso as trajetórias fossem estendidas.

Durante a análise de simulações de biomoléculas por DM é essencial que sejam estudadas estruturas estáveis do sistema criado (KARPLUS; MCCAMMON, 2002). A estrutura inicial obtida por cristalografia de raio-X revela a conformação do

sistema de estudo na conformação predominante no ambiente cristalino, a qual pode divergir da conformação em solução (TERADA; KIDERA, 2012). Além disso, a estabilidade das estruturas da trajetória é uma forma de avaliar se o campo de força não está promovendo alterações significativas na estrutura do sistema (CAO; WANG, 2010). Nesse sentido, a avaliação do perfil do RMSD é a alternativa comumente utilizada para julgar a estabilidade de um sistema simulado por DM. O perfil do RMSD, próximo a 2 Å, para o conjunto de átomos pesados do nucleossomo, excluindo as caudas das histonas, é condizente com flutuações térmicas normais de uma proteína. Além disso, também está concordante com os fatores de temperatura cristalográficos encontrados na estrutura utilizada nesse trabalho (DAVEY et al., 2002). A sobreposição estrutural dos átomos (etapa que precede o computo do RMSD e RMSF), é crucial para o resultado final e deve interferir o mínimo nos valores calculados. Caso esta inclua regiões intensamente dinâmicas, o RMSD e o RMSF aumentaria para todos os átomos, resultando numa avaliação equivocada da estabilidade do sistema. Várias estratégias foram desenvolvidas para aperfeiçoar a sobreposição entre estruturas, incluindo o tratamento diferenciado para os átomos com menor desvio, ou seja, aqueles que apresentam menores flutuações (MARTÍNEZ, 2015). Assim utilizar como referência para a rotação e translação do sistema apenas os átomos da cadeia principal de regiões estáveis ao longo da trajetória é uma alternativa simples e eficaz. A Figura 27 exhibe o perfil de RMSD para várias seleções de átomos, fica evidente que o conjunto de átomos pesados das hélices do nucleossomo é uma ótima alternativa.

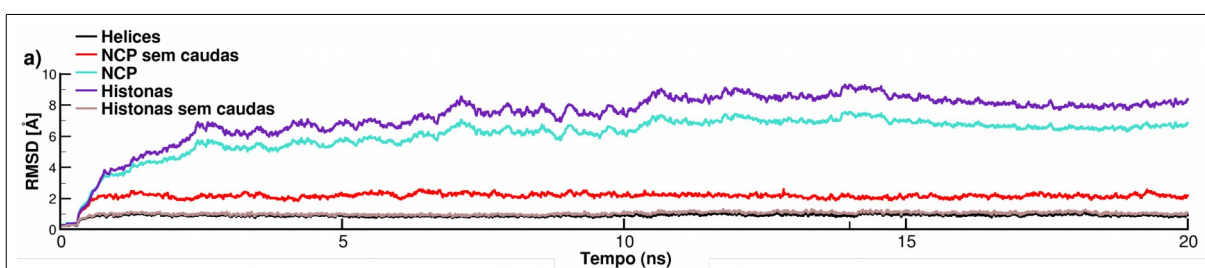


Figura 27: Perfil do RMSD de diversas seleções de átomos do nucleossomo.

Prováveis seleções atômicas para a sobreposição estrutural da trajetória. Em todas as seleções apenas os átomos da cadeia principal foram considerados.

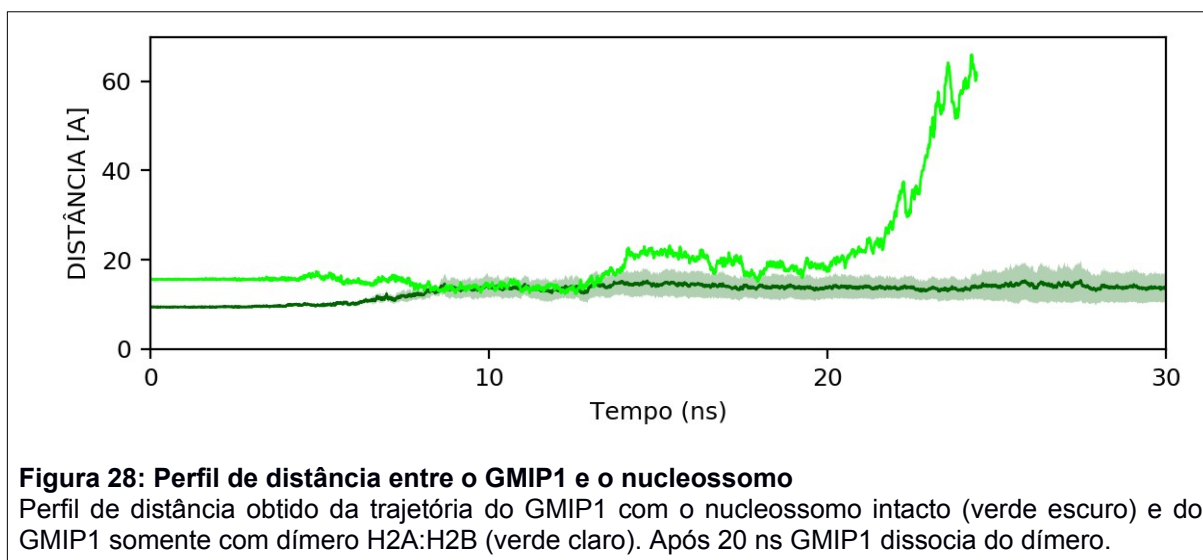
Em suma, os resultados encontrados nesse trabalho sugerem que os NBPeps interagem com o nucleossomo desencadeando mudanças estruturais. Todos os NBPeps se ligaram ao nucleossomo e mantiveram essa interação por 100 ns de trajetória de equilíbrio. LANA e H4-coil foram os NBPeps que permaneceram mais próximos do patch ácido, conseqüentemente deslocaram mais moléculas de água e foram os que mais reduziram as flutuações atômicas dessa região. Porém, esses NBPeps demonstraram padrão molecular de interações distintos. Enquanto que o LANA, em todas as simulações, reteve contato com os mesmos aminoácidos do *patch* ácido, H4-coil diversificou os resíduos com os quais interagiu.

O peptídeo da cauda da H4 foi investigado em duas conformações, em alfa-hélice como proposto por Yang e Arya (2011) e em uma configuração aberta, evidenciado na primeira estrutura de alta resolução do nucleossomo (LUGER et al., 1997). Os resultados desse trabalho indicam que a conformação mais flexível (H4-coil) promove alterações mais significativas no nucleossomo do que a conformação em hélice. Em virtude dos contatos com um maior número de aminoácidos do *patch* ácido, além de redução da hidratação e alteração das flutuações dos aminoácidos dessa interface.

Estudos computacionais relacionaram a acetilação da lisina 16 da cauda da H4 com aumento da probabilidade de regiões estruturadas, como alfa-hélice (WINOGRADOFF et al., 2015). Essa maior probabilidade também foi demonstrada por dicroísmo circular (WANG et al., 2000). Por sua vez, a acetilação da K16 da H4 está associada ao aumento da presença de regiões na cromatina mais acessíveis a maquinaria transcricional (eucromatina) (KUNDU et al., 2000; VERDONE; CASERTA; DI MAURO, 2005). Também é observada uma redução na interação da cauda acetilada com o *patch* ácido do que a cauda não acetilada na conformação aberta (ZHANG; ERLER; LANGOWSKI, 2017). Assim, a diferença observada nos efeitos e na interação com o nucleossomo entre as conformações do peptídeo da cauda da H4 está em concordância com os dados da literatura. Além do mais, ensaios bioquímicos mostraram que o H4pep, mas não HMGN2pep, GMIP1 e LANA, é altamente específico para atingir o nucleossomo, com importantes efeitos na estrutura final do nucleossomo e efeitos robustos *in vivo* (TELES et al., 2020). Ainda,

análise de dicroísmo circular indicou a prevalência de conformações mais flexíveis para o peptídeo da cauda da H4 (TELES, 2020).

Foi observado contato do H4-coil com o DNA, diferente de sua outra conformação. Entretanto os NBPEps que mantiveram o maior número de contato com o DNA foram os inicialmente descritos como ligantes do *patch* ácido e DNA, HMGN2pep e GMIP1. O DNA inclusive parece ser necessário para a ligação do GMIP1 ao nucleossomo. O ensaio de ligação *in vitro* (NMR) do GMIP1 não indicou a ligação do GMIP1 ao dímero H2A:H2B (TELES, 2020). Simulações de DM nas mesmas condições do experimento, i.e. GMIP1 e dímero H2A:H2B em solução, mostrou que GMIP1 não mantém a interação com as proteínas (Figura 28). Assim, uma possível explicação para o resultado negativo do ensaio de ligação é a ausência do nucleossomo completo, principalmente a inexistência do DNA.



II.4.8 Limitações e perspectivas

Entre as limitações desse trabalho pode-se destacar a estrutura inicial do complexo GMIP1:nucleossomo. Apesar desse peptídeo ter sido baseado no modo de ligação do RCC1, é plausível imaginar que esse NBPEp possa interagir de modo diferente deste, pois trata-se de uma nova molécula. Trajetórias mais longas com tempo de equilíbrio de distanciamento estendidos poderiam elucidar melhor os

efeitos dos NBPEps no nucleossomo. Ainda, a análise de distância, embora informativa do ponto de vista inicial de uma interação, revela pouco sobre as orientações e conformações visitadas pelos peptídeos no sítio de ligação. E a investigação sobre a energética da interação entre os NBPEps e o nucleossomo seria relevante para uma descrição mais completa do complexo.

De forma geral, os resultados da interação de NBPEps com o nucleossomo são positivos e indicaram que peptídeos apresentam um potencial de ligação e impacto na estrutura do nucleossomo. Um panorama para estudos futuros inclui o desenho de novos peptídeos com melhor afinidade para o nucleossomo, seguido de investigações acerca do impacto que variantes de histonas desempenham na interação. Essas informações serão cruciais para melhor compreensão da modulação da estrutura do nucleossomo e possivelmente para o desenvolvimento de abordagens terapêuticas.

III . CAPÍTULO III. LIPÍDIOS

III.1 INTRODUÇÃO

Os lipídios estão entre os principais compostos encontrados em tecidos vivos e que apresentam como característica marcante a solubilidade em solventes não polares (CAMMACK et al., 2006). Diversas moléculas como ácidos graxos, fosfolipídeos, hormônios esteroides e outras moléculas se enquadram na definição de lipídios e realizam diversas funções celulares. Por exemplo, o colesterol é o principal esteroide sintetizado pelo organismo humano, e auxilia na manutenção da integridade da membrana celular, assim como atua como precursor dos hormônios esteroides, ácidos biliares e vitamina D (HANUKOGLU, 1992).

III.1.1 Lipídios no núcleo

Ao longo dos anos diversos estudos revelaram a presença de lipídios no compartimento nuclear, como a presença de fosfolipídios no núcleo de células do fígado (CHAYEN; GAHAN, 1958) ou a associação de colesterol a fibras de cromatina em células de ratos (ERICKSON; DAVISON; GOULD, 1975). Posteriormente, esforços foram realizados a fim de descrever as funções de moléculas lipídicas no núcleo (ALBI; MAGNI, 2004) e sua organização no compartimento nuclear, como a formação de gotas lipídicas nucleares (LAYERENZA et al., 2013). Formadas por lipídios neutros que aparecem como pequenos pontos na visualização no microscópio, as gotas lipídicas participam de vários processos fisiológicos e também podem atuar como repositórios de lipídios para uso sob demanda da célula.

Após os dados iniciais que indicaram a presença de lipídios no núcleo, a comunidade científica vem realizando esforços para quantificar os lipídios no ambiente nuclear. Recentemente esses dados foram compilados em uma revisão (FERNANDES et al., 2018) mostrando que as concentrações de lipídeos no núcleo celular variam de acordo com o tipo e compartimento nuclear. Os fosfolipídios foram

a classe de lipídio mais abundante independentemente do tipo ou compartimento nuclear.

Tabela 4. Receptores nucleares e seus ligantes lipídicos.

Receptor Nuclear	Ligante lipídico	Função
PPAR α	oxPC e POPC	metabolismo lipídico hepático, resposta inflamatória e metabolismo da glicose
PPAR γ	ácido graxo oxidados e oxPC	lipogênese e sensibilidade à insulina no tecido adiposo
LXR	Oxisterois e derivados do colesterol	metabolismo do colesterol
FXR	Ácidos biliares	metabolismo de lipoproteínas
ER	Hormônios esteroides	regulação de genes envolvidos no crescimento celular e resposta endócrina

PPAR (Receptor Ativado por Proliferador de Peroxissomo), LXR (Receptor de Fígado X), FXR (Receptor de Farnesóide X), ER (Receptor de Estrogênio). oxPC (fosfatidilcolina oxidada), POPC (1-palmitoil-2-oleoil-fosfatidilcolina).

No ambiente nuclear, as gotas lipídicas interagem com numerosas proteínas e podem participar especificamente da regulação de eventos nucleares (WELTE, 2015). Não obstante, evidências apontam que diferentes lipídios como ácidos graxos e metabólitos de colesterol (ITOH et al., 2008; LIBERATO et al., 2012; WARNER; HUANG; GUSTAFSSON, 2017), assim como fosfolípidos (CROWDER; SEACRIST; BLIND, 2017), atuam diretamente como ligantes de receptores nucleares. Na tabela 4 estão listados alguns receptores nucleares com suas respectivas funções e ligantes lipídicos.

Assim, fica evidente que lipídios podem regular a expressão gênica por meio da interação com esses fatores de transcrição presentes no núcleo. Ademais, é plausível considerar que lipídios possam interagir com regiões hidrofóbicas da cromatina, já que necessitam permear o ambiente nuclear para se ligar aos receptores nucleares. Recentemente foi mostrado que o colesterol, um lipídio extremamente importante clínico e fisiologicamente, é capaz de se ligar ao nucleossomo e a fibra de cromatina, facilitar a formação *in vitro* de ambos e induzir a compactação da fibra de 30nm. Experimentos *in silico* revelaram seis sítios de

ligação do colesterol na superfície do nucleossomo. Além disso, análises computacionais sugeriram duas explicações para os efeitos observados *in vitro* do colesterol sobre cromatina. (1) o fortalecimento das interações entre aminoácidos hidrofóbicos devido a desidratação da superfície do nucleossomo causada pelo colesterol e (2) a presença sítios de ligação na cauda da H4, conhecida pelo seu papel fundamental nas interações internucleossomais, (SILVA et al., 2017). Todavia, ainda assim pouco se sabe sobre o impacto da interação de outros lipídios com a cromatina.

Desta forma, esse capítulo está voltado ao estudo da interação de moléculas lipídicas com o nucleossomo e uma discussão sobre o impacto de tais moléculas na cromatina e nos processos por ela regulado.

III.2 OBJETIVOS

III.2.1 Objetivo geral

Investigar as características da interação de moléculas ligantes de nucleossomo (lipídios e peptídeos), avaliando os efeitos dessa interação em nível atômico.

III.2.1 Objetivos específicos

b) Lipídios e cromatina

- Buscar os sítios de ligação de vários lipídios no nucleossomo;
- Comparar os resultados obtidos com a análise realizada com colesterol;
- Analisar os possíveis efeitos da presença de lipídios na cromatina.

III.3 PROTOCOLOS

Foram realizados experimentos de *docking* molecular de onze lipídios no nucleossomo para verificar a existência de sítios compartilhados com o colesterol. Posteriormente, os resultados foram comparados aos obtidos para o colesterol.

III.3.1 Protocolo de *docking* molecular

Os experimentos de *docking* foram realizados com o programa Auto Dock Vina, em um O *grid* de 110 x 110 x 110 Å³, suficiente para envolver o nucleossomo e as caudas das histonas próximas ao DNA. Para considerar a flexibilidade do nucleossomo foram utilizadas 120 estruturas aleatórias do nucleossomo. Estas estruturas foram obtidas da trajetória de equilíbrio do nucleossomo livre em solução produzida durante os estudos com o colesterol (SILVA et al., 2017). As coordenadas atômicas dos ligantes foram coletadas no banco de dados do PDB com os códigos de acesso: 4N16 (ácido cólico); 5DXE (estradiol); 2VDY (cortisol); 1T20 (pregnenolona); 2AA2 (aldosterona); 2VV4 (6-oxoOTE), 2VV3(4-oxoDHA), 2VV2(5-HEPA), 2VV0(DHA), 2VSR (9-HODE). O Auto Dock Vina considerou o espaço conformacional de todos os ligantes, exceto o POPC, por apresentar 40 pontos de torções. A flexibilidade do POPC foi resolvida usando 20 configurações aleatórias advindas de simulação com uma membrana de POPC totalmente hidratada e equilibrada. Esta membrana foi obtida pelo *plugin Membrane Builder* do VMD e simulada com campo de força CHARMM36 no software NAMD por 50 ns.

Foram selecionadas as 50 melhores soluções em cada uma das 120 conformações do nucleossomo utilizadas como receptor do *docking*. Como resultado, 6.000 formas e interfaces de encaixe de cada lipídeo ao nucleossomo foram encontradas. No caso do POPC, foram 120000 soluções. Os resultados foram agrupados de acordo com a proximidade entre si, em nuvens de soluções. O critério de agrupamento foi um *cutoff* de 10 Å de distância entre os resultados. Exceto para o POPC, que por apresentar aproximadamente o dobro do comprimento dos demais lipídios, teve um *cutoff* de 15Å. Para fins comparativos, as análises de *docking* com o colesterol foram refeitas nas mesmas condições dos experimentos com as demais

moléculas lipídicas. Em todos experimento os parâmetros foram deixados no padrão do programa, com exceção apenas da exaustividade que foi aumentada para 200.

III.4 RESULTADOS

III.4.1 Interface de interação dos lipídios no nucleossomo

Recentemente, o grupo de pesquisa do professor Guilherme Santos investigou a interação do colesterol com o nucleossomo. Nesse estudo foi observado que o colesterol induziu a compactação das fibras de cromatina reconstituídas *in vitro*. Além disso, por métodos computacionais foi estimado a afinidade do colesterol pelo nucleossomo, a qual foi baixa, da ordem de 0.01 a 0.2 mM^{-1} (SILVA et al., 2017). Dessa forma, é razoável imaginar que outras moléculas lipídicas possam interagir com o nucleossomo, mesmo que seja com baixa afinidade, e afetar a arquitetura da cromatina.

Baseado nessa premissa, cálculos de *docking* exploratórios entre lipídios, além do colesterol, e o nucleossomo foram realizados. Os lipídios investigados foram aqueles descritos como ligantes de receptores nucleares, ou que já tiveram sua presença no núcleo celular constatada. No total, foi averiguada a interação de onze moléculas lipídicas com o nucleossomo. Estas análises incluíram os hormônios esteroides: aldosterona, estradiol e pregnenolona; um ácido biliar: ácido cólico; e um fosfolípido: 1-palmitoil-2-oleoil-fosfatidilcolina (Figura 29). Também foram incluídos ácidos graxos propostos como ligantes naturais do PPAR γ (ITOH et al., 2008): ácido hidroxi-octadecadienóico (9-HODE), ácido docosaheptaenóico (DHA), ácido 5-hidroxi-eicosapentaenóico (5-HEPA), ácido 4-oxo-docosaheptaenóico (4-oxoDHA) e ácido 6-oxo-octadecatrienoico (6-oxoOTE).

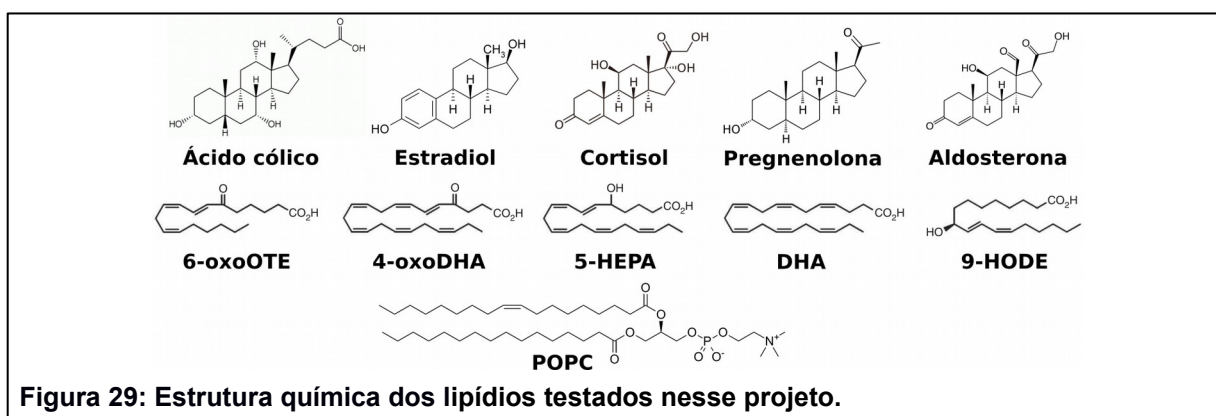


Figura 29: Estrutura química dos lipídios testados nesse projeto.

A figura 30 exibe as nuvens de soluções do *docking* dos lipídios (ver descrição no tópico III.3.1). Para todos os lipídios é possível observar uma maior densidade de soluções próximas ao DNA, preferencialmente na interface entre o octâmero de histonas e o DNA. Algumas soluções também foram encontradas na superfície proteica do nucleossomo, entretanto em número menor em comparação com as soluções em contato com o DNA. Tal resultado poderia indicar uma maior afinidade desses lipídios pelo DNA do que pelas histonas. Entretanto, um padrão semelhante de soluções de *docking* foi observado nas análises com o colesterol (FERNANDES, 2016). E no caso do colesterol, não foi verificadas diferenças significativas nas estimativas de constantes de ligação entre sítios presentes unicamente na superfície proteica ou em contato com DNA. Dessa forma, seguiram análises comparativas entre os resultados do *docking* aqui apresentados e os resultados obtidos para o colesterol.

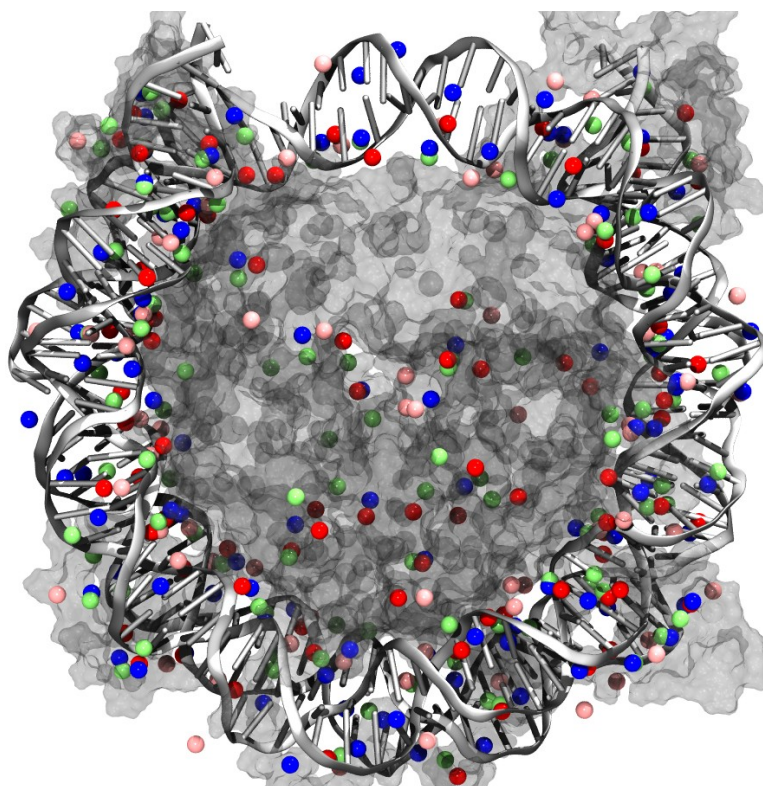


Figura 30: Possíveis locais de interação de algumas moléculas lipídicas com o nucleossomo. Centro de massa das nuvens de solução encontradas pelo procedimento de docking. Para facilitar a visualização, apenas um lipídio de cada classe está sendo representado. São eles: Estradiol (verde), Ácido cólico (azul), 6-oxo-OTE (vermelho) e POPC (rosa).

III.4.2 Comparação de locais de interação entre lipídios

Foi analisada a sobreposição dos locais de interação dos lipídios investigados nesse trabalho com os locais observados para o colesterol, incluindo os sítios de interação s1 a s6 investigados por meio de simulações de DM Tabela 5. Como destacado no protocolo de *docking*, as análises para o colesterol foram refeitas nas mesmas condições dos outros lipídios. Assim, para todas as moléculas lipídicas 45000 soluções de *docking* foram computadas. É importante destacar que o número de nuvens de solução em um local não é necessariamente correlacionado com o número de soluções no sítio. Dado que, caso várias soluções de *docking* sejam similares, estas serão agrupadas em uma mesma nuvem de solução.

Tabela 5. Comparação dos sítios de ligação do colesterol com os demais lipídios

	Sítio s1		Sítio s2		Sítio s3		Sítio s4		Sítio s5		Sítio s6		Outros (HO)		DNA		Total (n)
	n	Energia	n	Energia	n	Energia	n	Energia	n	Energia	n	Energia	n	Energia	n	Energia	
Colesterol	2	-7.61	1	-7.48	1	-7.43	1	-6.80	1	-7.22	1	-7.26	13	-7.49	85	-7.42	105
Aldosterona	2	-8.43	2	-8.20	---	---	---	---	1	-8.16	1	-8.00	14	-8.14	81	-8.28	101
Cortisol	2	-8.24	2	-8.03	---	---	---	---	1	-8.07	1	-7.78	12	-7.98	88	-8.10	106
Estradiol	2	-7.73	1	-7.51	1	-7.33	---	---	1	-8.15	1	-7.36	19	-7.55	83	-7.58	108
Pregнено.	2	-7.82	1	-7.59	---	---	---	---	1	-7.57	---	---	17	-7.52	86	-7.60	107
Ácido c.	1	-7.95	2	-7.85	---	---	---	---	1	-7.82	1	-7.77	12	-7.93	87	-8.04	104
4-oxo-DHA	2	-5.77	2	-5.60	1	-5.66	1	-5.73	1	-5.78	1	-5.68	13	-5.71	65	-5.64	86
5-HEPA	2	-5.81	1	-5.65	1	-5.63	1	-5.70	1	-5.76	1	-5.40	12	-5.81	68	-5.75	87
6-oxo-OTE	1	-5.34	2	-5.30	1	-5.20	1	-5.00	1	-5.36	1	-5.30	14	-5.42	65	-5.33	86
9-HODE	2	-5.41	1	-5.31	1	-5.33	1	-4.80	1	-5.40	1	-5.36	11	-5.39	66	-5.37	84
DHA	2	-5.76	2	-5.62	1	-5.64	1	-5.70	1	-5.79	1	-5.46	11	-5.71	65	-5.67	84
POPC	1	-10.50	2	-11.66	---	---	1	---	1	-11.47	2	-11.47	11	-11.60	55	-11.93	73

n = número de nuvens de solução; Energia (ΔG^0) em kcal/mol obtida pelo Autodock Vina.

Inicialmente, é possível observar que apenas os sítios s3 e s4 não são comuns para todos os lipídios. Somente os ácidos graxos (4-oxo DHA, 5-HEPA, 6-

oxo OTE, 9-HODE e DHA) compartilharam todos os seis sítios simulados do colesterol. Porém, apresentaram valores de energia livre de ligação menos favoráveis do que os demais lipídios. Embora os valores de energia livre de ligação preditos pelo programa de *docking* não sejam os mais acurados, pode-se extrair tendências desses valores. Como no caso do POPC, que apresentou resultados indicativos de maior afinidade do que todos os outros lipídios testados, inclusive o colesterol. Interessante perceber também que, apesar de ser observado um número consideravelmente maior de soluções no DNA, os valores de energia livre de ligação pouco diferem das soluções presentes na interface proteica do DNA.

III.4.3 Considerações gerais e discussão

Há várias décadas, Linus Pauling declarou que o reconhecimento molecular é o segredo da vida. Realmente, vários processos biológicos cruciais são regulados por interações não covalentes e associações moleculares são alvo de extensivos estudos (BARON; MCCAMMON, 2013). O *docking* é uma ferramenta computacional que visa a geração de modelos de interação entre um ligante e um receptor, ou seja, busca o reconhecimento molecular entre essas moléculas. É uma técnica mais rápida e econômica do que técnicas experimentais nos primeiros passos de uma investigação acerca de interações não covalentes entre biomoléculas (SOUSA et al., 2013). Entretanto, tais vantagens só são possíveis pelas simplificações dos modelos. A rigidez imposta a estruturas com número de átomos elevado é particularmente uma notável limitação. Nesse sentido, o *ensemble docking* é uma excelente alternativa para contornar esse obstáculo. Ao utilizar um conjunto de estruturas provenientes de simulações de DM, as flutuações das moléculas envolvidas no *docking* são levadas em consideração (AMARO et al., 2018). Tais flutuações, como o rearranjo das cadeias laterais, fazem parte do reconhecimento molecular e podem alterar a acessibilidade de sítios de ligação (GUEDES; DE MAGALHÃES; DARDENNE, 2014). Nos cálculos de *docking* com os lipídios foram utilizadas um total de 120 estruturas do nucleossomo coletadas de um intervalo de 6ns de trajetória de equilíbrio. Valores razoáveis quando comparados a literatura

para receptores pouco flexíveis como o nucleossomo (YESYLEVSKYY et al., 2016).

No geral, os resultados do *docking* com os lipídios sugere que não há locais predominantes de interação desses ligantes no nucleossomo, pois foram observadas nuvens de soluções espalhadas pela superfície nucleossomal com energias livre de ligação similares. A maioria das soluções foram encontradas no DNA ou na interface entre o DNA e o HO. De fato, estudos *in vitro* demonstraram que lipídios neutros podem ser ligar especificamente ao DNA dupla hélice (ZHDANOV et al., 2015b). Assumindo que tais eventos ocorram no ambiente nuclear, a ligação de lipídios pode ter um impacto profundo na modulação da expressão gênica. De fato, vários estudos tentaram elucidar a função de lipídios no núcleo (Revisado em: ZHDANOV et al., 2015a). Já foi observado que o ácido gama-linoleico alterou a expressão gênica em células de gliomas em camundongos (BENADIBA; MIYAKE; COLQUHOUN, 2009). Alguns fosfolipídios, incluindo a fosfatidilcolina (POPC), apresentou inibição da síntese de RNA (MARALDI et al., 1984). Além disso, foi demonstrado um papel significativo dos lipídios na modulação da DNA polimerase e na estrutura do DNA, o que afetaria diretamente o processo transcricional (KAMATH-LOEB et al., 2014).

Porém, há de se considerar que apesar de o *docking* ser uma ferramenta eficaz na identificação de modos de encaixe de um ligante no receptor, ele não consegue fornecer os detalhes energéticos acurados do complexo, gerando muitas soluções falsamente positivas. Entretanto, mesmo que preliminares, estes dados sugerem que é possível a interação direta de lipídios com a cromatina. Consistente com essa ideia, é fascinante a possibilidade de que, mesmo por interações de baixa afinidade não específicas, moléculas lipídicas desempenham um papel no complexo mecanismo regulatório da cromatina. Além disso, recapitulando que vários lipídios já foram identificados como possíveis ligantes diretos de receptores nucleares, é razoável conceber que a cromatina também possa funcionar como um depósito lipídico para as proteínas nucleares.

Sob outra perspectiva, desconsiderando a ação direta de lipídios sobre a cromatina, é plausível ponderar que sob uma certa concentração uma molécula

hidrofóbica seria capaz de afetar a maneira pela qual macromoléculas, como o nucleossomo, interagirão entre si e, conseqüentemente, modulando seu estado. Já foi mostrado que membranas com diferentes misturas lipídicas têm a capacidade de formar fases separadas (ELSON et al., 2010). Nos últimos anos têm-se surgido evidências de processos de separação de fases em diversos compartimentos celulares (BOEYNAEMS et al., 2018), incluindo no contexto da cromatina (NARLIKAR, 2020). Na separação de fase uma solução homogênea de moléculas distintas se separa espontaneamente em fases líquidas coexistentes, onde uma fase está enriquecida por uma dessas moléculas e a outra não (ALBERTI, 2017). Em 2017 foi proposto que a formação de regiões de heterocromatina mediadas pela *linker* histona HP1 poderia ser explicada no contexto da separação de fases, um processo basicamente físico (LARSON et al., 2017; STROM et al., 2017). Assim, a separação de fases poderia ser responsável por mediar o silenciamento gênico induzido pela HP1 ao formar um compartimento composto por heterocromatina, que apesar de não possuir membrana ou outra barreira física, é isolado ao acesso de proteínas que possam induzir a descompactação da cromatina (LARSON et al., 2017). Mais recentemente, foi observado também que as caudas das histonas acompanhadas de marcações epigenéticas, como a acetilação, regula a separação de fases na cromatina (GIBSON et al., 2019).

Contudo, a natureza físico-químicas do ambiente da heterocromatina separada por fases ainda é obscuro. Pode-se então questionar se lipídios poderiam cooperar com a formação de hidrogéis baseados em proteínas ou mesmo influenciar as separações de fase na célula, modulando a arquitetura da cromatina. Extrapolando esse cenário molecular para a visão clínica, a ideia de que lipídios possa desencadear estados alterados da cromatina é extremamente relevante, posto que o entendimento de mudanças na dinâmica da cromatina podem desvendar mecanismos de doenças e promover o desenvolvimento de novos tratamentos (SILVA; OLIVEIRA; SANTOS, 2015).

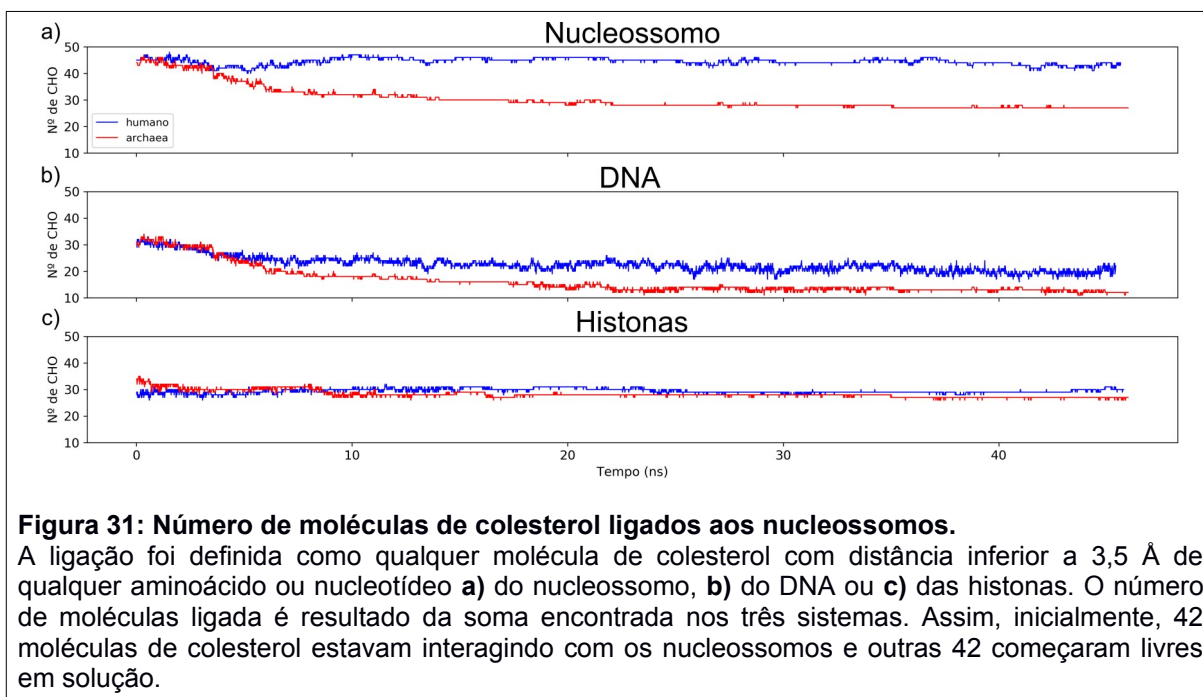
III.4.4 Limitações e Perspectivas

O Autodock Vina é um *software* de *docking* largamente utilizado e muito eficiente. Entretanto, sua *scoring function* é fundamentada principalmente em aprendizado de máquina, com a incorporação de dados não físicos no ranqueamento de soluções. Assim, uma investigação da interação dos lipídios com o nucleossomo por outras técnicas computacionais, como simulações de DM, serão de grande valia para uma melhor compreensão dos efeitos de lipídios no nucleossomo. Ademais, experimentos biofísicos *in vitro* a fim de verificar quais são os fenômenos causados pelos lipídios na cromatina também serão importantes.

Entre as perspectivas vale destacar a investigação sobre o papel de lipídios no nucleossomo ancestral eucariótico, o nucleossomo de archaea. Recentemente foi resolvida a estrutura tridimensional do nucleossomo de *Methanothermobacter thermautotrophicus* (do domínio Archaea) e foi observada uma grande semelhança com o nucleossomo eucariótico (MATTIROLI et al., 2017). De fato, de acordo com a árvore da vida vigente, o domínio Archaea é considerado o ancestral de todas as vidas eucarióticas (EME et al., 2017). Porém em contraste com a cromatina dos eucariotos, que formam arranjos de nucleossomo mais livres (do tipo “colar de contas”), os nucleossomo em Archaeas formam uma *superhelix* quase contínua de vários nucleossomo sobrepostos (MATTIROLI et al., 2017).

Esse trabalho já foi iniciado e inicialmente está sendo comparada a afinidade de moléculas de colesterol com o nucleossomo humano e de archaea. O colesterol foi o lipídio escolhido devido ao estudo prévio dos seus efeitos no nucleossomo e cromatina eucariótica (SILVA et al., 2017). Foram realizadas simulações de DM do tipo *flooding* com 10 mM de colesterol. Basicamente simulações de *flooding* são simulações de DM com presença explícita dos ligantes em alta concentração. Assim, é acelerada a amostragem da interação do ligante com o receptor permitindo observar os eventos com probabilidade mais acurada do que o *docking*, além de fornecer uma descrição com alta resolução espacial e temporal. No entanto, o grande poder descritivo dessas simulações é proporcional ao seu alto custo computacional.

Resumidamente, nesse projeto foram realizadas simulações de *flooding* do colesterol no nucleossomo humano e de archaea. Para acelerar a amostragem foram criados três sistemas dos prováveis estados desse complexo: (i) todos os ligantes (colesterol) em solução, não ligados ao nucleossomo (distante em pelo menos 10Å); (ii) metade das moléculas de colesterol ligadas ao nucleossomo e metade em solução, livre; (iii) todas as moléculas de colesterol ligadas ao nucleossomo. As posições iniciais do colesterol em ambos nucleossomos foram obtidas por *docking*. Os sistemas foram inseridos em uma caixa de simulação, a qual foi preenchida com solução aquosa com 150 mM de KCl. O volume dessas caixas de simulações diferiu entre os sistemas, afim de preservar a fração molar (número de ligantes dividido pelo número de moléculas de água) entre os sistemas uniforme.



A figura 31 exibe a quantidade total de moléculas de colesterol que permaneceram ligadas ao nucleossomo (31a), DNA (31b) e histonas (31c).

Podemos observar que o número de moléculas de colesterol no nucleossomo humano é maior do que no nucleossomo de archaea. Esse resultado não é surpreendente, visto que já existem dados que indicaram a ligação do colesterol ao

nucleossomo humano (SILVA et al., 2017). Ademais, esse resultado é promissor visto que foi possível verificar, mesmo que em uma trajetória curta, que essa metodologia é capaz de diferenciar a interação do colesterol entre os nucleossomos.

Outra possível explicação para a aparente menor afinidade do colesterol pelo nucleossomo de archaea são as diferenças bioquímicas entre os lipídios presentes em archaea e em eucariotos. Nos eucariotos são encontrados ácidos graxos com ligação de ésteres nos carbonos de glicerol (ŘEZANKA et al., 2018). Já em archaea, devido aos ambientes não convencionais e extremos em que se encontram, sua bioquímica lipídica é incomum, consistindo em cadeias de isoprenóides ligadas por ligações de éter a carbonos de glicerol (JACOBSEN et al., 2017). Essa diferença bioquímica é fundamental para a sobrevivência das arqueas. Conferindo uma membrana celular mais robusta com características importantes para um organismo extremófilo (PEARSON, 2019).

Assim experimentos envolvendo lipídios com características semelhantes às observadas em arqueas serão realizados de forma semelhante aos aqui apresentados. Também serão investigadas concentrações diferentes desses lipídios e trajetórias mais longas.

IV . CONCLUSÕES

Entender a dinâmica completa das interações de todos os ligantes de nucleossomo no contexto celular é um desafio enorme para os cientistas. Recentemente inúmeros dados sobre a dinâmica da cromatina nos leva a reinterpretar o cenário simplista sobre a sua regulação, no qual os complexos de proteínas definem os estados da cromatina.

Vários NBPs foram bem caracterizadas em nível atômico e estas estruturas revelaram os detalhes da interação com o nucleossomo. O *patch* ácido destaca-se como a principal região de ancoragem dessas proteínas. No entanto, esta região não parece ser a única interface de contato para moléculas externas do nucleossomo. Assim como as NBPs não parecem ser as únicas moléculas capazes de interagirem diretamente com a cromatina. Visto que foram encontraram vários locais de ligação para o colesterol no nucleossomo e possivelmente para outras moléculas lipídicas em regiões próximas e distantes ao *patch* ácido.

Nesse trabalho, foram apresentados dados preliminares sobre a interação de moléculas lipídicas com o nucleossomo. Assim como evidências, ao nível atômico, sobre o impacto de diferentes peptídeos NBPePs na estrutura do nucleossomo. Ambos resultados são promissores e sugerem um papel dessas moléculas na regulação da cromatina.

V . REFERÊNCIAS

- ABRAMYAN, T. M. et al. Cluster Analysis of Molecular Simulation Trajectories for Systems where Both Conformation and Orientation of the Sampled States are Important. **Journal of computational chemistry**, v. 37, n. 21, p. 1973–1982, 5 ago. 2016.
- ALBERTI, S. Phase separation in biology. **Current Biology**, v. 27, n. 20, p. R1097–R1102, 23 out. 2017.
- ALBI, E.; MAGNI, M. P. V. The role of intranuclear lipids. **Biology of the Cell**, v. 96, n. 8, p. 657–667, 1 out. 2004.
- ALLAHVERDI, A. et al. The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association. **Nucleic Acids Research**, v. 39, n. 5, p. 1680–1691, mar. 2011.
- AMARO, R. E. et al. Ensemble Docking in Drug Discovery. **Biophysical Journal**, v. 114, n. 10, p. 2271–2278, 22 maio 2018.
- ARENTS, G. et al. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. **Proceedings of the National Academy of Sciences of the United States of America**, v. 88, n. 22, p. 10148–10152, 15 nov. 1991.
- ARMACHE, K.-J. et al. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. **Science (New York, N.Y.)**, v. 334, n. 6058, p. 977–982, 18 nov. 2011.
- BARBERA, A. J. et al. The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus LANA. **Science**, v. 311, n. 5762, p. 856–861, 10 fev. 2006.
- BARON, R.; MCCAMMON, J. A. Molecular recognition and ligand association. **Annual Review of Physical Chemistry**, v. 64, p. 151–175, 2013.
- BARTMAN, C. R.; BLOBEL, G. A. Perturbing Chromatin Structure to Understand Mechanisms of Gene Expression. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 80, p. 207–212, 2015.
- BEAUCHEMIN, C. et al. ASSAY DEVELOPMENT AND HIGH THROUGHPUT SCREENING FOR INHIBITORS OF KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS N-TERMINAL LANA BINDING TO NUCLEOSOMES. **Journal of biomolecular screening**, v. 19, n. 6, p. 947–958, jul. 2014.
- BEDNAR, J. et al. Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. **Molecular Cell**, v. 66, n. 3, p. 384–397.e8, 4 maio 2017.

BENADIBA, M.; MIYAKE, J. A.; COLQUHOUN, A. Gamma-linolenic acid alters Ku80, E2F1, and bax expression and induces micronucleus formation in C6 glioma cells in vitro. **IUBMB Life**, v. 61, n. 3, p. 244–251, 2009.

BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81, n. 8, p. 3684–3690, 15 out. 1984.

BISWAS, M. et al. Role of Histone Tails in Structural Stability of the Nucleosome. **PLoS Comput Biol**, v. 7, n. 12, p. e1002279, 15 dez. 2011.

BOEYNAEMS, S. et al. Protein Phase Separation: A New Phase in Cell Biology. **Trends in Cell Biology**, v. 28, n. 6, p. 420–435, 1 jun. 2018.

CABRAL, W. F.; MACHADO, A. H. L.; SANTOS, G. M. Exogenous nucleosome-binding molecules: a potential new class of therapeutic drugs. **Drug Discovery Today**, v. 21, n. 5, p. 707–711, maio 2016.

CAMMACK, R. C. et al. lipid. In: **Oxford Dictionary of Biochemistry and Molecular Biology**. [s.l.] Oxford University Press, 2006.

CAO, Z.; WANG, J. A Comparative Study of Two Different Force Fields on Structural and Thermodynamics Character of H1 Peptide via Molecular Dynamics Simulations. **Journal of Biomolecular Structure and Dynamics**, v. 27, n. 5, p. 651–661, 1 abr. 2010.

CHAYEN, J; GAHAN, P. Lipid components in nucleohistone. **Biochemical Journal**, Portland Press 59 Portland Place, London W1n 3aj, England (1958). De Sandre-Giovannoli, A., R. Bernard, P. Cau, C. Navarro, J. Amiel, I. Boccaccio, S. Lyonnet, C. L. Stewart, A. Munnich, M. Le Merrer and N. Levy (2003). "Lamin a truncation in Hutchinson-Gilford progeria." **Science** 300(5628): 2055.

CHEN, Q. et al. Regulation of Nucleosome Stacking and Chromatin Compaction by the Histone H4 N-Terminal Tail–H2A Acidic Patch Interaction. **Journal of Molecular Biology**, Deciphering Histone Modifications in Development and Disease. v. 429, n. 13, p. 2075–2092, 30 jun. 2017.

COLLEPARDO-GUEVARA, R.; SCHLICK, T. Chromatin fiber polymorphism triggered by variations of DNA linker lengths. **Proceedings of the National Academy of Sciences**, v. 111, n. 22, p. 8061–8066, 3 jun. 2014.

CROWDER, M. K.; SEACRIST, C. D.; BLIND, R. D. Phospholipid Regulation of the Nuclear Receptor Superfamily. **Advances in biological regulation**, v. 63, p. 6–14, jan. 2017.

CUTTER, A. R.; HAYES, J. J. A brief review of nucleosome structure. **FEBS Letters**, v. 589, n. 20PartA, p. 2914–2922, 7 out. 2015.

- DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald - An Nlog(N) method for Ewald sums in large systems. **J. Chem. Phys**, v. 98, n. 12, p. 10089–10092, 1993.
- DAURA, X. et al. Peptide Folding: When Simulation Meets Experiment. **Angewandte Chemie International Edition**, v. 38, n. 1□2, p. 236–240, 15 jan. 1999.
- DAVEY, C. A. et al. Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution. **Journal of Molecular Biology**, v. 319, n. 5, p. 1097–1113, 21 jun. 2002.
- DAVEY, G. E. et al. Nucleosome acidic patch-targeting binuclear ruthenium compounds induce aberrant chromatin condensation. **Nature Communications**, v. 8, n. 1, p. 1575, 17 nov. 2017.
- DHANIK, A.; KAVRAKI, L. E. Protein-Ligand Interactions: Computational Docking. In: JOHN WILEY & SONS, LTD (Ed.). . **eLS**. Chichester, UK: John Wiley & Sons, Ltd, 2012.
- DORIGO, B. et al. Chromatin Fiber Folding: Requirement for the Histone H4 N-terminal Tail. **Journal of Molecular Biology**, v. 327, n. 1, p. 85–96, 14 mar. 2003.
- ELSON, E. L. et al. Phase separation in biological membranes: integration of theory and experiment. **Annual Review of Biophysics**, v. 39, p. 207–226, 2010.
- EME, L. et al. Archaea and the origin of eukaryotes. **Nature Reviews. Microbiology**, v. 15, n. 12, p. 711–723, 10 nov. 2017.
- ERICKSON, S. K.; DAVISON, A. M.; GOULD, R. G. Correlation of rat liver chromatin-bound free and esterified cholesterol with the circadian rhythm of cholesterol biosynthesis in the rat. **Biochimica Et Biophysica Acta**, v. 409, n. 1, p. 59–67, 21 out. 1975.
- FERNANDES, V. et al. Fat nucleosome: Role of lipids on chromatin. **Progress in Lipid Research**, 17 abr. 2018.
- FERNANDES, V. A. Ação do colesterol sobre a arquitetura da cromatina: docking molecular do colesterol ao nucleossomo. 20 set. 2016.
- FINCH, J. T. et al. Structure of nucleosome core particles of chromatin. **Nature**, v. 269, n. 5623, p. 29–36, 1 set. 1977.
- FOLOPPE, N.; MACKERELL, JR., Alexander D. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. **Journal of Computational Chemistry**, v. 21, n. 2, p. 86–104, 30 jan. 2000.

FRENKEL, D.; SMIT, B. **Understanding Molecular Simulation, Second Edition: From Algorithms to Applications**. 2 edition ed. San Diego: Academic Press, 2001.

FUSSNER, E.; CHING, R. W.; BAZETT-JONES, D. P. Living without 30nm chromatin fibers. **Trends in Biochemical Sciences**, v. 36, n. 1, p. 1–6, jan. 2011.

FYODOROV, D. V. et al. Emerging roles of linker histones in regulating chromatin structure and function. **Nature reviews. Molecular cell biology**, v. 19, n. 3, p. 192–206, mar. 2018.

GIBSON, B. A. et al. Organization of Chromatin by Intrinsic and Regulated Phase Separation. **Cell**, v. 179, n. 2, p. 470–484.e21, 3 out. 2019.

GUEDES, I. A.; DE MAGALHÃES, C. S.; DARDENNE, L. E. Receptor–ligand molecular docking. **Biophysical Reviews**, v. 6, n. 1, p. 75–87, mar. 2014.

HANUKOGLU, I. Steroidogenic enzymes: structure, function, and role in regulation of steroid hormone biosynthesis. **The Journal of Steroid Biochemistry and Molecular Biology**, v. 43, n. 8, p. 779–804, dez. 1992.

HOOVER, NULL. Canonical dynamics: Equilibrium phase-space distributions. **Physical Review. A, General Physics**, v. 31, n. 3, p. 1695–1697, mar. 1985.

HSU, T. C. Differential rate in RNA synthesis between euchromatin and heterochromatin. **Experimental Cell Research**, v. 27, n. 2, p. 332–334, 1 ago. 1962.

HUANG, J.; MACKERELL, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. **Journal of computational chemistry**, v. 34, n. 25, p. 2135–2145, 30 set. 2013.

HUANG, Y.-C. et al. The effect of linker DNA on the structure and interaction of nucleosome core particles. **Soft Matter**, v. 14, n. 45, p. 9096–9106, 21 nov. 2018.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33–38, 1 fev. 1996.

ITOH, T. et al. Structural basis for the activation of PPAR γ by oxidized fatty acids. **Nature Structural & Molecular Biology**, v. 15, n. 9, p. 924–931, set. 2008.

JACOBSEN, A.-C. et al. Archaeal lipids in oral delivery of therapeutic peptides. **European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences**, v. 108, p. 101–110, 15 out. 2017.

JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. **J. Chem. Phys**, v. 79, n. 2, p. 926–935, 1983.

KALASHNIKOVA, A. A. et al. The role of the nucleosome acidic patch in modulating

higher order chromatin structure. **Journal of the Royal Society Interface**, v. 10, n. 82, 6 maio 2013.

KAMATH-LOEB, A. S. et al. Sphingosine, a Modulator of Human Translesion DNA Polymerase Activity. **Journal of Biological Chemistry**, v. 289, n. 31, p. 21663–21672, 1 ago. 2014.

KARANTZA, V. et al. Thermodynamic studies of the core histones: Ionic strength and pH dependence of H2A-H2B dimer stability. **Biochemistry**, v. 34, n. 17, p. 5988–5996, 1 maio 1995.

KARANTZA, V.; FREIRE, E.; MOUDRIANAKIS, E. N. Thermodynamic studies of the core histones: pH and ionic strength effects on the stability of the (H3-H4)/(H3-H4)₂ system. **Biochemistry**, v. 35, n. 6, p. 2037–2046, 13 fev. 1996.

KARPLUS, M.; MCCAMMON, J. A. Molecular dynamics simulations of biomolecules. **Nature Structural & Molecular Biology**, v. 9, n. 9, p. 646–652, set. 2002.

KATO, H. et al. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, n. 30, p. 12283–12288, 26 jul. 2011.

KORNBERG, R. D. Structure of Chromatin. **Annual Review of Biochemistry**, v. 46, n. 1, p. 931–954, 1977.

KOROLEV, N. et al. Electrostatic Origin of Salt-Induced Nucleosome Array Compaction. **Biophysical Journal**, v. 99, n. 6, p. 1896–1905, 22 set. 2010.

KOYAMA, M.; KURUMIZAKA, H. Structural diversity of the nucleosome. **The Journal of Biochemistry**, v. 163, n. 2, p. 85–95, 1 fev. 2018.

KUNDU, T. K. et al. Activator-dependent transcription from chromatin in vitro involving targeted histone acetylation by p300. **Molecular Cell**, v. 6, n. 3, p. 551–561, set. 2000.

LAI, W. K. M.; PUGH, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. **Nature Reviews. Molecular Cell Biology**, v. 18, n. 9, p. 548–562, set. 2017.

LARSON, A. G. et al. Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. **Nature**, v. 547, n. 7662, p. 236–240, jul. 2017.

LAWRENCE, M.; DAUJAT, S.; SCHNEIDER, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. **Trends in genetics: TIG**, v. 32, n. 1, p. 42–56, jan. 2016.

- LAYERENZA, J. P. et al. Nuclear lipid droplets: a novel nuclear domain. **Biochimica Et Biophysica Acta**, v. 1831, n. 2, p. 327–340, fev. 2013.
- LEE, D. Y. et al. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. **Cell**, v. 72, n. 1, p. 73–84, 15 jan. 1993.
- LESBATS, P. et al. Structural basis for spumavirus GAG tethering to chromatin. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, n. 21, p. 5509–5514, 23 2017.
- LI, G.; ZHU, P. Structure and organization of chromatin fiber in the nucleus. **FEBS Letters**, v. 589, n. 20PartA, p. 2893–2904, 7 out. 2015.
- LIBERATO, M. V. et al. Medium Chain Fatty Acids Are Selective Peroxisome Proliferator Activated Receptor (PPAR) γ Activators and Pan-PPAR Partial Agonists. **PLOS ONE**, v. 7, n. 5, p. e36297, 23 maio 2012.
- LIM, J. B.; ROGASKI, B.; KLAUDA, J. B. Update of the Cholesterol Force Field Parameters in CHARMM. **The Journal of Physical Chemistry B**, v. 116, n. 1, p. 203–210, 12 jan. 2012.
- LUGER, K. et al. Crystal structure of the nucleosome core particle at 2.8 Å resolution. **Nature**, v. 389, n. 6648, p. 251–260, 18 set. 1997.
- LUGER, K.; DECHASSA, M. L.; TREMETHICK, D. J. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? **Nature Reviews Molecular Cell Biology**, v. 13, n. 7, p. 436–447, jul. 2012.
- MACKERELL, A. D. et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. **The Journal of Physical Chemistry B**, v. 102, n. 18, p. 3586–3616, 1 abr. 1998.
- MACKERELL, A. D., Jr; FEIG, M.; BROOKS, C. L., 3rd. Improved treatment of the protein backbone in empirical force fields. **Journal of the American Chemical Society**, v. 126, n. 3, p. 698–699, 28 jan. 2004.
- MAKDE, R. D. et al. Structure of RCC1 chromatin factor bound to the nucleosome core particle. **Nature**, v. 467, n. 7315, p. 562–566, 30 set. 2010.
- MARALDI, N. M. et al. Conformational changes of nuclear chromatin related to phospholipid induced modifications of the template availability. **Advances in Enzyme Regulation**, v. 22, p. 447–464, 1 jan. 1984.
- MARTÍNEZ, L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. **PLoS ONE**, v. 10, n. 3, 27 mar. 2015.

MARTYNA, G. J.; TOBIAS, D. J.; KLEIN, M. L. Constant pressure molecular dynamics algorithms. **The Journal of Chemical Physics**, v. 101, n. 5, p. 4177–4189, 1 set. 1994.

MATTIROLI, F. et al. Structure of histone-based chromatin in Archaea. **Science**, v. 357, n. 6351, p. 609–612, 11 ago. 2017.

MCGINTY, R. K.; TAN, S. Recognition of the nucleosome by chromatin factors and enzymes. **Current Opinion in Structural Biology**, v. 37, p. 54–61, abr. 2016.

MORGAN, M. T. et al. Structural basis for histone H2B deubiquitination by the SAGA DUB module. **Science (New York, N.Y.)**, v. 351, n. 6274, p. 725–728, 12 fev. 2016.

NARLIKAR, G. J. Phase-separation in chromatin organization. **Journal of Biosciences**, v. 45, n. 1, p. 5, 20 jan. 2020.

NOSÉ, S. A molecular dynamics method for simulations in the canonical ensemble. **Molecular Physics**, v. 52, n. 2, p. 255–268, 10 jun. 1984.

OLINS, A. L.; OLINS, D. E. Spheroid Chromatin Units (v Bodies). **Science**, v. 183, n. 4122, p. 330–332, 25 jan. 1974.

OLIVEIRA, S. H. P. et al. KVFinder: steered identification of protein cavities as a PyMOL plugin. **BMC bioinformatics**, v. 15, p. 197, 17 jun. 2014.

PEARSON, A. Resolving a piece of the archaeal lipid puzzle. **Proceedings of the National Academy of Sciences of the United States of America**, v. 116, n. 45, p. 22423–22425, 5 nov. 2019.

PENG, J. et al. Molecular mechanism of histone variant H2A.B on stability and assembly of nucleosome and chromatin structures. **Epigenetics & Chromatin**, v. 13, 14 jul. 2020.

PENTAKOTA, S. et al. Decoding the centromeric nucleosome through CENP-N. **eLife**, v. 6, p. e33442, 27 dez. 2017.

PHILLIPS, J. C. et al. Scalable Molecular Dynamics with NAMD. **Journal of computational chemistry**, v. 26, n. 16, p. 1781–1802, dez. 2005.

QUIGLEY, D.; PROBERT, M. I. J. Langevin dynamics in constant pressure extended systems. **The Journal of Chemical Physics**, v. 120, n. 24, p. 11432, 2004.

ŘEZANKA, T. et al. Lipidomic Analysis: From Archaea to Mammals. **Lipids**, v. 53, n. 1, p. 5–25, 2018.

ROBINSON, P. J. J. et al. EM measurements define the dimensions of the “30-nm” chromatin fiber: Evidence for a compact, interdigitated structure. **Proceedings of the**

National Academy of Sciences of the United States of America, v. 103, n. 17, p. 6506–6511, 25 abr. 2006.

ROBINSON, P. J. J. et al. `30nm' chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. **Journal of molecular biology**, v. 381, n. 4, 12 set. 2008.

SAUER, P. V. et al. Mechanistic insights into histone deposition and nucleosome assembly by the chromatin assembly factor-1. **Nucleic Acids Research**, v. 46, n. 19, p. 9907–9917, 2 nov. 2018.

SAURABH, S. et al. Atomistic Simulation of Stacked Nucleosome Core Particles: Tail Bridging, the H4 Tail, and Effect of Hydrophobic Forces. **The Journal of Physical Chemistry B**, v. 120, n. 12, p. 3048–3060, 31 mar. 2016.

SCHLICK, T. **Molecular Modeling and Simulation: An Interdisciplinary Guide**. New York, NY: Springer New York, 2010. v. 21

SCHULER, L. D.; DAURA, X.; VAN GUNSTEREN, W. F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. **Journal of Computational Chemistry**, v. 22, n. 11, p. 1205–1218, 1 ago. 2001.

SHAO, J. et al. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. **Journal of Chemical Theory and Computation**, v. 3, n. 6, p. 2312–2334, 1 nov. 2007.

SHAYTAN, A. K. et al. Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions. **Journal of Molecular Biology**, v. 428, n. 1, p. 221–237, 16 jan. 2016.

SHUKLA, M. S. et al. The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. **Nucleic Acids Research**, p. gkq1174, 3 dez. 2010.

SILVA, I. T. G. DA; OLIVEIRA, P. S. L. DE; SANTOS, G. M. Featuring the nucleosome surface as a therapeutic target. **Trends in Pharmacological Sciences**, v. 36, n. 5, p. 263–269, maio 2015.

SILVA, I. T. G. et al. Biophysical studies of cholesterol effects on chromatin. **Journal of Lipid Research**, v. 58, n. 5, p. 934–940, maio 2017.

SOLOVEI, I.; THANISCH, K.; FEODOROVA, Y. How to rule the nucleus: divide et impera. **Current Opinion in Cell Biology**, Cell nucleus. v. 40, p. 47–59, 1 jun. 2016.

SONG, F. et al. Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. **Science**, v. 344, n. 6182, p. 376–380, 25 abr. 2014.

SOUSA, S. F. et al. **Protein-Ligand Docking in the New Millennium – A Retrospective of 10 Years in the Field**. Text. Disponível em: <<http://www.ingentaconnect-com.ez54.periodicos.capes.gov.br/content/ben/cmcl/2013/00000020/00000018/art00002>>. Acesso em: 18 mar. 2016.

STROM, A. R. et al. Phase separation drives heterochromatin domain formation. **Nature**, v. advance online publication, 21 jun. 2017.

TALBERT, P. B.; HENIKOFF, S. Histone variants on the move: substrates for chromatin dynamics. **Nature Reviews Molecular Cell Biology**, v. 18, n. 2, p. 115–126, fev. 2017.

TELES, K. et al. Nucleosome binding peptide presents laudable biophysical and in vivo effects. **Biomedicine & Pharmacotherapy**, v. 121, p. 109678, 1 jan. 2020.

TELES, K. A. Estudos biofísicos e in vivo de moléculas ligantes de nucleossomo. 15 jun. 2020.

TERADA, T.; KIDERA, A. Comparative molecular dynamics simulation study of crystal environment effect on protein structure. **The Journal of Physical Chemistry. B**, v. 116, n. 23, p. 6810–6818, 14 jun. 2012.

TOLSMA, T. O.; HANSEN, J. C. Post-translational modifications and chromatin dynamics. **Essays in Biochemistry**, v. 63, n. 1, p. 89–96, 23 abr. 2019.

TROTT, O.; OLSON, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. 455–461, 30 jan. 2010.

VAN HOLDE, K. E. **Chromatin**. New York, NY: Springer New York, 1989.

VERDONE, L.; CASERTA, M.; DI MAURO, E. Role of histone acetylation in the control of gene expression. **Biochemistry and Cell Biology = Biochimie Et Biologie Cellulaire**, v. 83, n. 3, p. 344–353, jun. 2005.

VERLET, L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. **Physical Review**, v. 159, n. 1, p. 98–103, 5 jul. 1967.

VETTESE-DADEY, M. et al. Role of the histone amino termini in facilitated binding of a transcription factor, GAL4-AH, to nucleosome cores. **Molecular and Cellular Biology**, v. 14, n. 2, p. 970–981, fev. 1994.

VOGLER, C. et al. Histone H2A C-Terminus Regulates Chromatin Dynamics, Remodeling, and Histone H1 Binding. **PLOS Genet**, v. 6, n. 12, p. e1001234, 9 dez. 2010.

WANG, J. et al. Development and testing of a general amber force field. **Journal of Computational Chemistry**, v. 25, n. 9, p. 1157–1174, 15 jul. 2004.

WANG, X. et al. Acetylation Increases the α -Helical Content of the Histone Tails of the Nucleosome. **Journal of Biological Chemistry**, v. 275, n. 45, p. 35013–35020, 10 nov. 2000.

WARNER, M.; HUANG, B.; GUSTAFSSON, J.-A. Estrogen Receptor β as a Pharmaceutical Target. **Trends in Pharmacological Sciences**, v. 38, n. 1, p. 92–99, 2017.

WELTE, M. A. Expanding roles for lipid droplets. **Current biology: CB**, v. 25, n. 11, p. R470–R481, 1 jun. 2015.

WILSON, M. D. et al. The structural basis of modified nucleosome recognition by 53BP1. **Nature**, v. 536, n. 7614, p. 100–103, ago. 2016.

WINOGRADOFF, D. et al. The Acetylation Landscape of the H4 Histone Tail: Disentangling the Interplay between the Specific and Cumulative Effects. **Journal of the American Chemical Society**, v. 137, n. 19, p. 6245–6253, 20 maio 2015.

YANG, D.; ARYA, G. Structure and binding of the H4 histone tail and the effects of lysine 16 acetylation. **Physical Chemistry Chemical Physics**, v. 13, n. 7, p. 2911–2921, 1 fev. 2011.

YANG, Z. et al. The Core Histone N-Terminal Tail Domains Negatively Regulate Binding of Transcription Factor IIIA to a Nucleosome Containing a 5S RNA Gene via a Novel Mechanism. **Molecular and Cellular Biology**, v. 25, n. 1, p. 241–249, jan. 2005.

YESYLEVSKYY, S. O. et al. Selective Inhibition of STAT3 with Respect to STAT1: Insights from Molecular Dynamics and Ensemble Docking Simulations. **Journal of Chemical Information and Modeling**, v. 56, n. 8, p. 1588–1596, 22 ago. 2016.

ZHANG, R.; ERLER, J.; LANGOWSKI, J. Histone Acetylation Regulates Chromatin Accessibility: Role of H4K16 in Inter-nucleosome Interaction. **Biophysical Journal**, v. 112, n. 3, p. 450–459, 7 fev. 2017.

ZHDANOV, R. et al. Lipids contribute to epigenetic control via chromatin structure and functions. **ScienceOpen Research**, v. 2015, n. 10, p. 1–11, 1 out. 2015a.

ZHDANOV, R. I. et al. Lipid and fatty acid profiles of *Pseudomonas aurantiaca* DNA-bound lipids determined by mass spectrometry. **Microbiology**, v. 84, n. 1, p. 34–40, 1 jan. 2015b.

ZHOU, B.-R. et al. Structural Mechanisms of Nucleosome Recognition by Linker Histones. **Molecular cell**, v. 59, n. 4, p. 628–638, 20 ago. 2015.

ZHOU, B.-R.; BAI, Y. Chromatin structures condensed by linker histones. **Essays in Biochemistry**, v. 63, n. 1, p. 75–87, 23 abr. 2019.

ZHOU, K.; GAULLIER, G.; LUGER, K. Nucleosome structure and dynamics are coming of age. **Nature Structural & Molecular Biology**, v. 26, n. 1, p. 3–13, jan. 2019.

ANEXOS



Biophysical studies of cholesterol effects on chromatin^S

Isabel T. G. Silva,^{1,*} Vinícius Fernandes,^{1,*†} Caio Souza,[†] Werner Treptow,^{2,†}
and Guilherme M. Santos^{2,*}

Laboratório de Farmacologia Molecular, Departamento de Farmácia,* and Laboratório de Biologia Teórica e Computacional, Departamento de Biologia Celular,[†] Universidade de Brasília, Brasília, Brazil

ORCID ID: 0000-0003-1489-2161 (G.M.S.)

Abstract Changes in chromatin structure regulate gene expression and genome maintenance. Molecules that bind to the nucleosome, the complex of DNA and histone proteins, are key modulators of chromatin structure. Previous work indicated that cholesterol, a ubiquitous cellular lipid, may bind to chromatin in vivo, suggesting a potential function for lipids in modulating chromatin architecture. However, the molecular mechanisms of cholesterol's action on chromatin structure have remained unclear. Here, we explored the biophysical impact of cholesterol on nucleosome and chromatin fibers reconstituted in vitro and characterized in silico the cholesterol binding to the nucleosome. **Our findings support that cholesterol assists 10 and 30 nm chromatin formation and induces folding of long chromatin fibers as a result of direct interaction of the cholesterol to six nucleosomal binding sites.**—Silva, I. T. G., V. Fernandes, C. Souza, W. Treptow, and G. M. Santos. **Biophysical studies of cholesterol effects on chromatin.** *J. Lipid Res.* 2017. 58: 934–940.

Supplementary key words lipids • physical biochemistry • molecular biology

Chromatin is a macromolecular complex composed of distinct molecules. The nucleosome is the repetitive basic unit of chromatin. Nucleosomes arise from the precise stoichiometric interaction of proteins and DNA. Positively charged histone proteins interact with the DNA and provide an optimal strategy to functionally pack the genetic code inside a cell.

Changes in chromatin architecture are decisive for regulating the access of transcription factors, coregulators, and the basic transcription machinery to specific enhancers in target genes [reviewed in (1)]. Thus, chromatin state defines the functional genomic regulation.

The influence of different ions, such as of divalent cations Mg²⁺ or Ca²⁺, on the nucleosome and chromatin structure has been well-characterized in vitro (2). Under low-salt conditions in vitro, chromatin adopts an open conformation known as “beads on a string” or the 10 nm fiber. In a physiological ionic environment, with the presence of Mg²⁺ and the correct linker histone stoichiometry, the chromatin can fold to form a compacted chromatin structure, the 30 nm fiber (3). However, the impact of chromatin folding states on gene expression and phenotypic outcomes is difficult to predict due to the dynamic movement of the chromatin fibers in the nuclear environment. Furthermore, the full complement of nuclear molecules that affect chromatin structure and affect fundamental processes in the cell, such as transcription and DNA repair, is still unknown (4).

Although chromatin structural and functional states are largely defined by nucleosome binding proteins (NBPs) and the state of the histone H4 tail (5), the nuclear environment comprises several small molecules that may directly bind to chromatin [e.g., magnesium (6) and lipids (7)]. Conceivably, all these substances may have a considerable impact on chromatin architecture.

The characterization of endonuclear lipids has motivated a growing number of studies to hypothesize that this class of molecules may have an important and functional role on chromatin structure and gene expression (8).

Cholesterol is a ubiquitous cellular lipid, which is essential for steroid hormone production, the structural integrity of cellular membranes, and cell growth. Seminal studies in rats have shown that labeled cholesterol administered by gavage was detected within 2 h in the nuclei, associated with chromatin fibers purified from rats (9). Moreover, binding assays suggested that cholesterol binds to chromatin through

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Grants 483968/2013-7, 305008/2015-3, 23038.010052/2013-95, and 14/2013.

The authors declare no competing financial interests.

***Author's Choice**—Final version free via Creative Commons CC-BY license.

Manuscript received 16 January 2017 and in revised form 20 March 2017.

Published, JLR Papers in Press, March 22, 2017

DOI <https://doi.org/10.1194/jlr.M074997>

Abbreviations: HO, histone octamer; MD, molecular dynamics; MNase, micrococcal nuclease; NBP, nucleosome binding protein; NCP, nucleosome core particle; TBE, 18 mM Tris-borate EDTA buffer (pH 8) and 0.4 mM EDTA.

¹I. T. G. Silva and V. Fernandes contributed equally to this work.

²To whom correspondence should be addressed.

e-mail: gsantos@unb.br (G.M.S.); treptow@unb.br (W.T.)

^SThe online version of this article (available at <http://www.jlr.org>) contains a supplement.

one or more proteins (10). Later studies increased the understanding of the mechanisms of interaction between chromatin and small lipids and suggested that small lipid molecules, such as cholesterol and short-chain fatty acids, can regulate important epigenetic modifications that lead to chromatin structure changes (11). It has also been observed that the phosphoinositide pathway and changes in its nuclear profile may have a direct role in modulating chromatin structure (12).

Herein, motivated by the biophysical and biochemical aspects of chromatin interactions, we investigated the mechanism of cholesterol action on chromatin structure. *In vitro*, low concentration of cholesterol assisted 10 and 30 nm fibers and mononucleosome formation. Moreover, micrococcal nuclease (MNase) digestion, which targets nucleosome-free regions of long chromatin fibers, was inhibited in the presence of cholesterol. Finally, we used docking and molecular dynamics (MD) simulations to show that cholesterol interacts with nucleosomes through multiple binding sites.

METHODS

***In vitro* chromatin and mononucleosome reconstitution**

The histone octamer (HO) and linker histone (H5) were purified from chicken erythrocyte nuclei as described in Ref. 13. To form mononucleosomes and long chromatin fibers, we used respectively, short DNA (167 bp) and long array (177 bp repeated 36 times) of 601 sequence (14). Titration of HO was performed to determine the saturation point, as described in Ref. 13. Competitor DNA was added only to form long chromatin fibers. Folding of nucleosomal arrays was carried out by dialysis into 1 mM MgCl₂. Soluble cholesterol (S5442; Sigma) was added to the binding buffer or after dialysis when fibers or mononucleosomes were already formed. The analyses of reconstitution and folding were verified by electrophoresis in native agarose gels (0.8%) for long chromatin fibers and native bis-acrylamide gels (6%) for mononucleosomes. Electrophoresis was carried out with 0.2× TBE [18 mM Tris-borate (pH 8) and 0.4 mM EDTA] electrophoresis buffer at 20 mA. For folded chromatin, the samples were first fixed with 0.1% (v/v) glutaraldehyde on ice for 20 min. Densitometry was performed using ImageJ (National Institutes of Health, Bethesda, MD) version 1.49.

MNase assay

Reconstituted long chromatin fibers (60 nM), in presence or absence of soluble cholesterol, were digested for 8 min with 0.08 u of MNase (New England Biolabs) at 37°C. Aliquots of samples were collected at 0, 1, 2, 4, and 8 min of digestion and stopped with 10 mM EDTA (pH 8.0). The digestion was analyzed by electrophoresis in native agarose gels (0.8%), carried out with 0.2× TBE electrophoresis buffer at 20 mA.

Compaction of long chromatin fibers induced by Mg²⁺

Chromatin compaction by Mg²⁺ was adapted from Rhodes laboratory protocol (13). Briefly, long chromatin fibers were incubated with increasing concentrations of soluble cholesterol (1, 10, and 100 μM final) or labeled cholesterol (0.5 and 1 μM final) [TopFluor® cholesterol, 23-(dipyrrometheneboron difluoride)-24-norcholesterol; Avanti Polar Lipids] on ice for 30 min (water was used as control). Next, 4.5 mM MgCl₂ were added and incubated

for 15 min on ice. The compacted fibers were pelleted by centrifugation at 15,493 *g* for 15 min (4°C). The supernatants and pellets were verified by electrophoresis in native 0.8% agarose gels, carried out with 0.2× TBE electrophoresis buffer at 20 mA. Typhoon 9200 was used with 532 nm green laser excitation to visualize cholesterol fluorescence.

Computational methods

Atomistic models were built using high-resolution X-ray structure of the nucleosome core particle (NCP), Protein Data Bank code 1KX5 (15), and cholesterol, Protein Data Bank code 1N83 (16). Following the strategy adopted in our previous studies on ligand binding to proteins (17), AutoDock Vina (18) was used to probe the binding of the cholesterol molecule against 120 NCP independent structures to account for receptor flexibility. These structures were extracted from the last 6 ns of ~20 ns equilibration simulation of the NCP system. All docking solutions were clustered into different groups based on distance. Six putative binding sites for the cholesterol molecule were selected based on the protein region they were bound to. Two systems were then simulated for 200 ns in the presence of three cholesterol molecules representing the most stable conformation found at each binding site. Sites s1, s4, and s6 were simulated in one system, and sites s2, s3, and s5 were simulated in another one. Energy averages were sampled from these simulations in order to compute the binding constants for each cholesterol molecule by the linear interaction energy method (19, 20) (supplemental Fig. S6). All simulations were performed with the program, Nanoscale Molecular Dynamics (NAMD) (21), using the CHARMM 36 force field (22, 23) and TIP3 water model (24). The particle-mesh Ewald (PME) method (25) was employed on the electrostatic calculations and nonbonded interactions were cut off at 11 Å. All systems were simulated at 300 K, 1 atm, and with a 2 fs time step.

All images created and computational analyses were performed using visual MD (26). More details are provided as supplemental material.

RESULTS

Cholesterol assists mononucleosome and chromatin fiber formation *in vitro*

To better understand the effects of cholesterol on chromatin, we reconstituted long chromatin fibers and mononucleosomes *in vitro* using HOs purified from chicken erythrocytes (supplemental Fig. S1A) on a 601 bp DNA array. As a quality control of the chromatin reconstitution system, we analyzed open and compacted fibers by electron microscopy showing that this system was successful (supplemental Fig. S1B).

We used soluble cholesterol (SintheChol® NS0 supplement; Sigma) with proven physiological relevance (27, 28). Initially, we titrated three different concentrations of cholesterol during mononucleosome reconstitution (1, 10, and 100 μM) (Fig. 1A). When cholesterol was present, the HO concentration needed for DNA saturation was smaller than that needed for reconstitution without cholesterol (Fig. 1A). We observed that cholesterol, at final concentration at 1 μM, led to earlier DNA saturation by HO binding (Fig. 1A), as observed by the formation of mononucleosomes and loss of naked DNA. Pronounced cholesterol effects were also observed at 100 μM final cholesterol concentration. Under

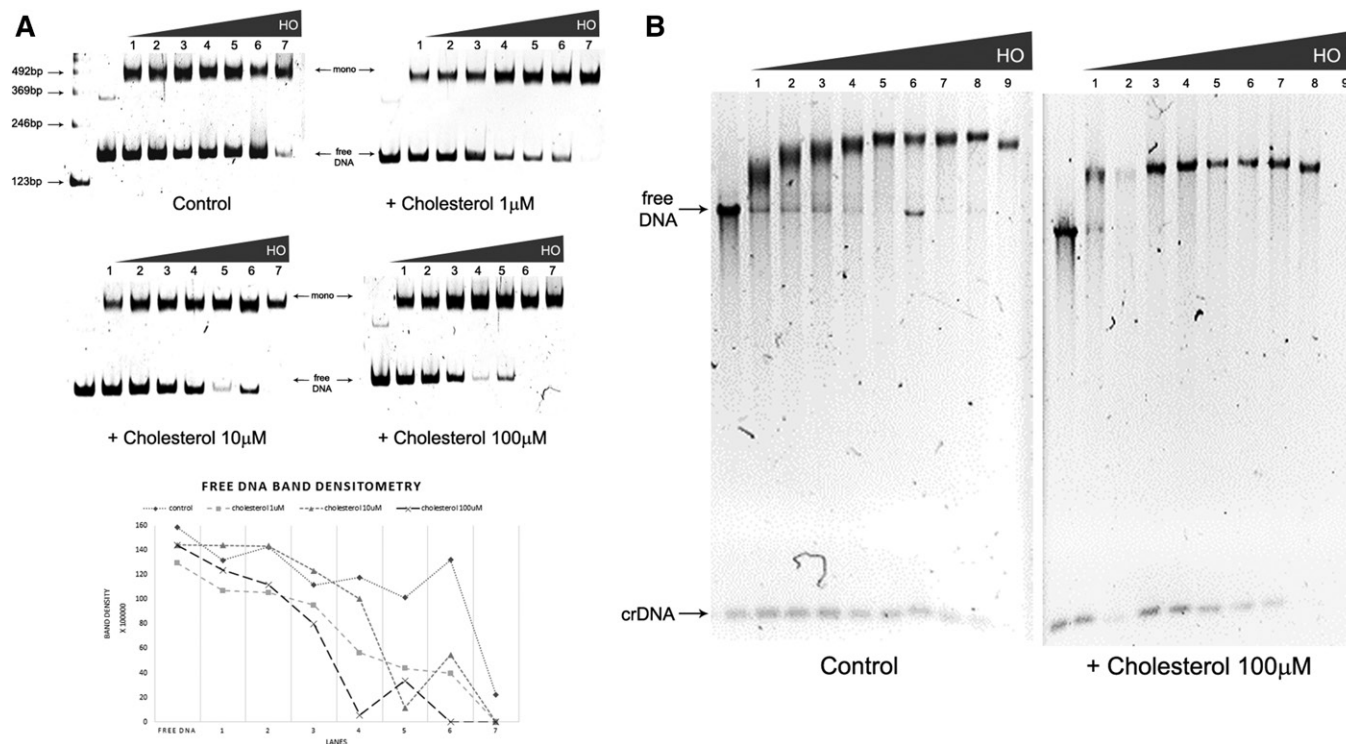


Fig. 1. Effect of cholesterol on nucleosomes and chromatin fibers reconstituted in vitro. A: Mononucleosome formation: Chicken HO titration (1 = 1.3 μM ; 2 = 1.5 μM ; 3 = 1.8 μM ; 4 = 2.0 μM ; 5 = 2.2 μM ; 6 = 2.5 μM ; 7 = 2.7 μM) in the absence (control) or presence of increasing cholesterol concentration (1, 10, and 100 μM) analyzed by EMSA in acrylamide gel 6%. DNA: 167.1. Gels were run under the same experimental conditions; free DNA band densitometry graph of upper gels; B: Chromatin fiber formation: Chicken HO titration (1 = 2.0 μM ; 2 = 2.2 μM ; 3 = 2.5 μM ; 4 = 2.7 μM ; 5 = 2.9 μM ; 6 = 3.1 μM ; 7 = 3.4 μM ; 8 = 4.0 μM ; 9 = 4.7 μM) in the absence (control) or presence of cholesterol (100 μM), analyzed by EMSA in agarose gel 0.8%. Array: 177.36. Gels run under the same experimental conditions. Legend: crDNA, competitor DNA. All gels are representative of three separate experiments.

those conditions, addition of 2.0 μM (Fig. 1A, lane 4) HO was sufficient for complete DNA saturation. This effect is graphically represented in the free DNA band densitometry, showing the faster decrease of the DNA band in the presence of cholesterol. The same effect is demonstrated in another representative gel, presented in the supplemental material (supplemental Fig. S1C).

A similar phenomenon was observed when cholesterol was present during the long fiber reconstitution in vitro. In this case, DNA saturation was reached at lower HO concentrations. In the presence of cholesterol, the array 177.36 was saturated at 2.0 μM HO (Fig. 1B, lane 1) compared with control samples that were saturated at 2.9 μM HO (Fig. 1B, lane 5). Detailed analysis of HO titration on DNA also revealed DNA saturation at lower HO concentration in the presence of cholesterol (supplemental Fig. S1D, lanes 1–3). This result is made clearer by the formation of an unstructured material (faster migration in the gel and precipitation) at 2.7–3.4 μM HO (supplemental Fig. S1D, lanes 11–13) compared with control samples at the same HO concentration.

Cholesterol favors chromatin folding in vitro

Because open chromatin fibers may behave differently from compacted fibers when exposed to cholesterol, we next evaluated the effect of cholesterol on in vitro reconstituted 30 nm fibers with purified linker histones (H5) from

chicken erythrocytes (supplemental Fig. S1A) in the presence of Mg^{2+} , which is a divalent ion with characterized properties to induce chromatin folding (3). Similarly to the previous results with 10 nm fibers, we observed that cholesterol also assisted the formation of the 30 nm fiber in vitro (Fig. 2A). Control samples showed diffuse bands in the native gel, indicating a high degree of heterogeneity in the system; the reconstitution performed in the presence of cholesterol showed sharper bands, suggesting a more homogenous compacted chromatin.

Because we observed that chromatin folding was facilitated by cholesterol, we wondered whether the effect of cholesterol on chromatin formation might have structural consequences. To evaluate this possibility, we performed MNase digestion of chromatin fibers formed in the presence or absence of cholesterol. We found that MNase digestion of long chromatin fibers was inhibited in the presence 100 μM of cholesterol (Fig. 2B). Importantly, the presence of cholesterol did not affect MNase digestion of free DNA (supplemental Fig. S1F); in fact, it improved the DNA digestion. These results are consistent with a model in which cholesterol affects chromatin formation by increasing chromatin compaction.

Next, we sought to verify whether cholesterol exerts its effects by directly binding to chromatin and inducing compaction of long chromatin fibers. It is well-described that chromatin compaction is induced by cations such as Mg^{2+} ,

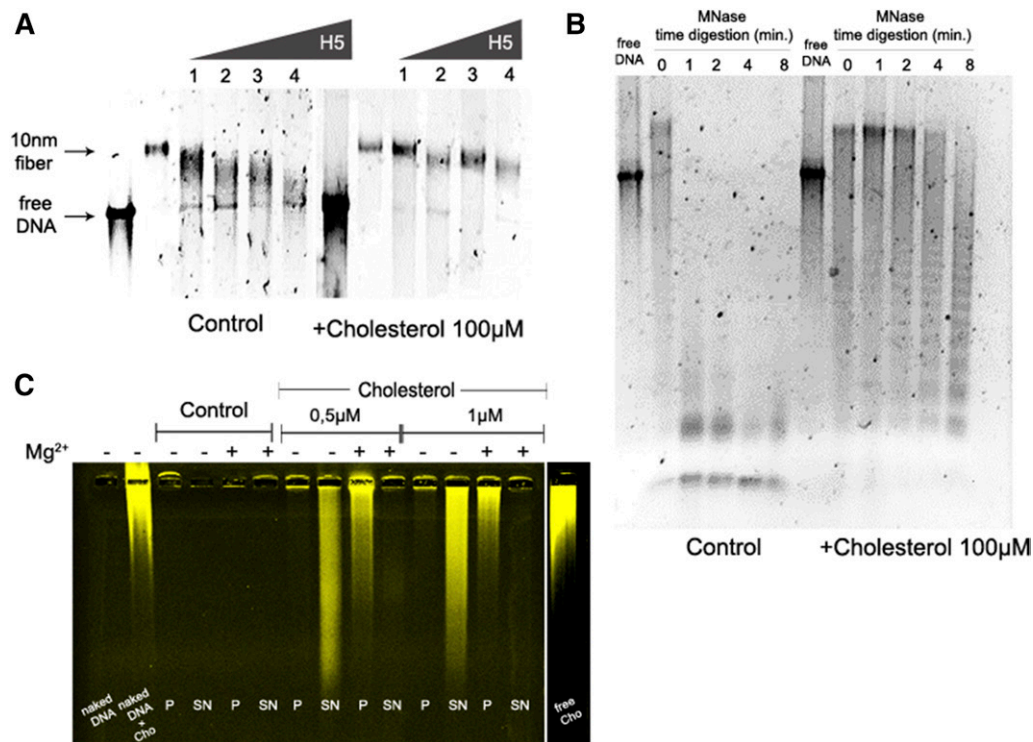


Fig. 2. Cholesterol favors chromatin compaction in vitro. **A:** Compacted chromatin fiber formation: Titration of H5 [linker histone (1 = 1.5 μM ; 2 = 3.0 μM ; 3 = 4.6 μM ; 4 = 6.2 μM)] on long chromatin fibers reconstituted in the absence (control) or presence of 100 μM cholesterol, analyzed by EMSA in agarose gel 0.8%. Array: 177.36. Gels were run under the same experimental conditions. **B:** Chromatin fiber digestion: MNase digestion of 10 nm fibers in the presence or absence of 100 μM cholesterol. Agarose gel 0.8%. Array: 177.36; **C:** Chromatin compaction assay: Compaction of 10 nm fibers induced by MgCl_2^{2+} in the absence (control) or presence of labeled cholesterol (0.5 or 1 μM), analyzed by EMSA in agarose gel 0.8%. Array: 177.36. Pellets (P), indicates compacted fibers; supernatant (SN), indicates open fibers; Cho, cholesterol. All gels are representative of three separate experiments.

which is known for promoting nucleosome-nucleosome interactions (6). Long chromatin fibers treated with 0.5–1.0 μM of labeled cholesterol were exposed to Mg^{2+} to induce compaction. Strikingly, cholesterol bound avidly to open or compacted chromatin fibers. In low concentrations (0.5 and 1 μM), labeled fibers moved from the supernatant to the pellet in the presence of Mg^{2+} (Fig. 2C). Labeled compacted fibers induced by Mg^{2+} showed decreased electrophoretic mobility when compared with open fibers, with the majority of compacted fibers being retained in the well. The presence of extensive smears in the gel was noticeable, which was due to the presence of free labeled cholesterol dissociated from the chromatin fibers. Due to technical limitations, we did not perform analytical ultracentrifugation with the long chromatin fibers in the presence of cholesterol to verify the precise degree of compaction.

It was not trivial to define the optimal concentration of cholesterol to be used in our in vitro system that would be physiologically relevant. The physiological concentration of cholesterol in the blood is approximately 5 mM. However, the apparent K_d of cholesterol bound to chromatin was estimated to be approximately 1 nM (10). Herein, our in vitro system rendered chromatin fibers at a final concentration of 60 nM and mononucleosomes at 0.9 μM , and the cholesterol used was at 100 μM . This concentration gave approximately 1,600 cholesterol molecules per chromatin

fiber with 36 nucleosomes, providing 45 cholesterol molecules per nucleosome. Considering that approximately 3,000 water molecules were found in the atomic structure of a nucleosome (29), with 121 molecules being important for HO interactions with the DNA, we believe that 45 cholesterol molecules per one nucleosome is stoichiometrically reasonable.

Cholesterol interacts with nucleosomes through multiple binding sites

Collectively, our findings indicate that cholesterol plays an important role in chromatin structure in vitro, a result that may rely on cholesterol binding to nucleosomes. To evaluate this hypothesis, we applied docking and MD-based free-energy calculations to study binding of cholesterol at the atomic level. Specifically, we docked the ligand against an ensemble of equilibrated structures of nucleosomes, generated from tens of nanoseconds of MD simulation of the X-ray structure of the construct (29). Starting from cholesterol-bound structures, we carried out free energy calculations to resolve site-specific affinities of the ligand against the nucleosomes. The linear interaction energy method was used for that purpose (see the supplemental material).

The docking results were clustered into 20 cholesterol groups at the protein-nucleosome surface (Fig. 3A).

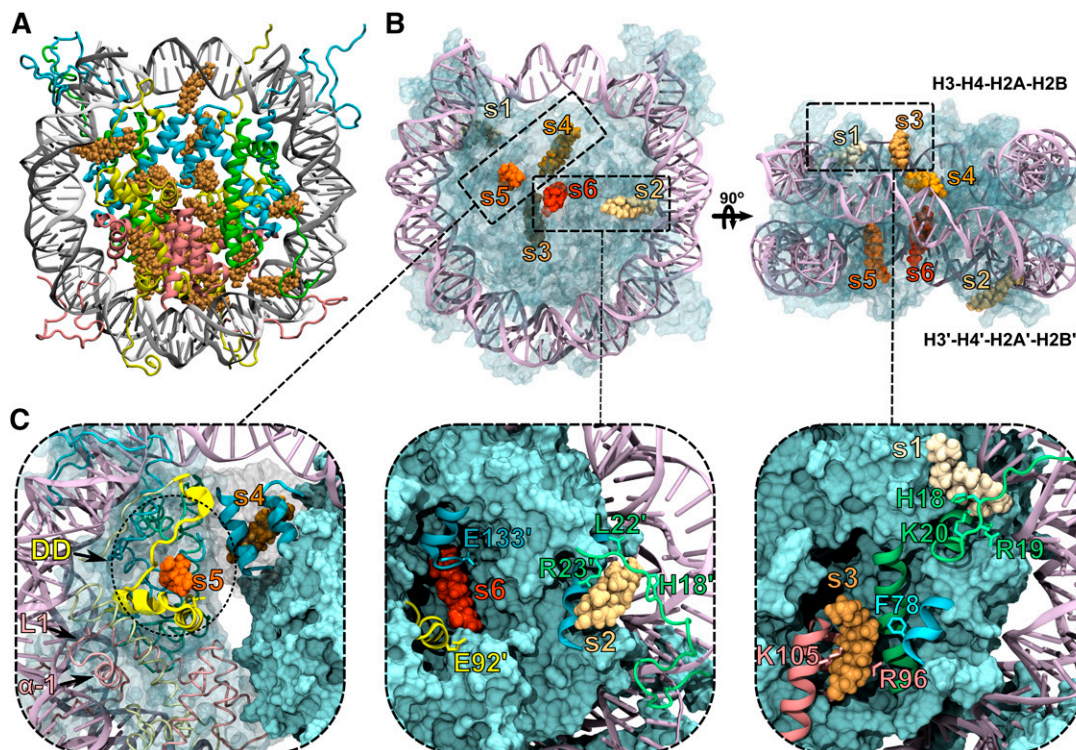


Fig. 3. Cholesterol binds the nucleosome through multiple sites. A: Docking solutions for cholesterol (orange) against the NCP. B, C: Bound cholesterol at sites s1 through s6 along each of the histone units H3 and H3' (blue), H4 and H4' (green), H2A and H2A' (yellow), H2B and H2B' (pink). The docking domain (DD) and H2B loop 1 (L1) and helix 1 (α -1). Highlighted are amino acids of the acidic patch (E91' and E92' from H2A'), of the H4/H4' tail (H18, R19, K20, L22, R23), and amino acids F78, E133' from H3 and R96, K105 from H2B, which were previously reported to bind NBP5 (5).

Interestingly, in six sites, the cholesterol lays very close to several amino acids, well-known to participate in the nucleosome structure and chromatin dynamics. Moreover, some cholesterol-contacting amino acids have also been correlated to nucleosome-nucleosome interactions or to nucleosome binding molecules. These interaction spots, hereafter called sites s1 through s6 (Fig. 3B, C), were selected for further MD and free energy calculations.

As detailed in supplemental Table S1, by binding to these multiple sites, cholesterol interacts with key nucleosome amino acids at the H4 tail, the acidic patch, and the docking domain in between the histone clusters H3-H4 and H2A-H2B (30).

After 200 ns of simulation, all six cholesterol molecules remained bound to the nucleosome under equilibrium conditions. As presented in supplemental Fig. S2, the nucleosome structure remained stable in its starting conformation throughout the simulation. In the presence of ligands, the root-mean-square deviation profile for the nucleosome converges to a plateau value of 2.0 Å, indicative of structural stability of the construct. Contrasting to the nucleosome structure, the bound ligand remains as flexible as in the free state in solution, which ensures a dynamic binding mode across the receptor binding sites (supplemental Fig. S3).

The close interaction between cholesterol and nucleosome suggests favorable binding affinities of the ligand to the receptor sites. To quantitatively evaluate this, we

proceeded with a series of independent free energy calculations to estimate the per-site affinity of cholesterol to nucleosome. As shown in **Table 1**, ligand affinities are heterogeneous and span a large range of values (0.01–0.2 mM^{-1}). For every site, the estimated absolute free energies favor cholesterol binding, rationalizing the stability of the molecular complex observed in the equilibrium MD simulations.

DISCUSSION

Cholesterol, an abundant cellular lipid, is among the molecules interacting with eukaryotic chromatin in vivo. However, very little is known about the mechanism and biological role of cholesterol-chromatin interactions.

Herein, we showed, in vitro and in silico, that cholesterol has an anticipated role in modulating chromatin structure

TABLE 1. Calculated values of binding free energies of cholesterol against the nucleosome

Sites	ΔG° (kcal·mol ⁻¹)	K (mM ⁻¹)
s1	-1.75 ± 0.18	0.0188
s2	-2.78 ± 0.16	0.1060
s3	-1.29 ± 0.17	0.0087
s4	-1.66 ± 0.17	0.0162
s5	-3.13 ± 0.16	0.1906
s6	-2.43 ± 0.16	0.0589

and provided atomistic insights into the impact of cholesterol on the nucleosome structure. Specifically, cholesterol appears to facilitate chromatin folding, verified by its effect on compacted chromatin fiber formation and by the inhibition of MNase digestion of long chromatin fibers. Complementary computational studies suggested direct molecular interaction of cholesterol with the nucleosome through six binding sites nearby important interacting regions for NBPs. Thus, our results support a model in which cholesterol induces chromatin formation and folding via direct interaction with nucleosomes through multiple binding sites.

Within this scenario, we speculate that the observed macroscopic effects on chromatin structure and stability might result from dewetting of the nucleosome, triggered by cholesterol binding. Specifically, dewetting transitions may impact core-to-core interactions at hydrophobic regions, including the docking domain and the H3-H3 four helix bundle, with global impact on chromatin. Consistent with this hypothesis, we found that cholesterol dehydrates the nucleosome surface in $\sim 17 (\pm 3)$ water molecules per binding site (Fig. 4B, C). As such, bound cholesterol reduces atom-to-atom position correlations at the docking domain and its neighboring regions, H2B loop 1 (L1) and helix 1 (α -1) (Fig. 4A), which likely influence critical packing interactions at the NCP. Note that in supplemental Fig. S4, this clear result also reflected on their increased root-mean-square fluctuations in the presence of bound

cholesterol. Besides dewetting and packing effects, we also speculate that additional mechanisms accounting for chromatin modulation might exist. Specifically, cholesterol binding to the H4 tail, well-known to assist chromatin compaction by making contacts with the acidic patch of adjacent nucleosome units, may also interfere with the final chromatin state. Consistent with this notion, note that in Fig. 4A, the bound cholesterol reduces coupling between the nucleosome and the H4 tail, enhancing its mobility freedom (root-mean-square fluctuations) up to $\sim 2 \text{ \AA}$, as compared with the ligand-free particle. Altogether, as shown in supplemental Fig., these views can be merged into a pictorial model that embodies the potential role of cholesterol in modulating particle-to-particle nucleosome interactions in chromatin.

It is known that some forms of transcriptionally silenced heterochromatin are localized adjacent to the nuclear lamina (31). Nevertheless, dynamically regulated genes have been found at the nuclear periphery when turned on (32). Despite that it is not clear how cholesterol under physiological conditions may interfere with chromatin structure inside the cell, together with previous evidence that it binds to chromatin (10), our findings suggest that cholesterol influences chromatin condensation by directly binding to the nucleosome.

Lipid microdomains, containing sphingomyelin and cholesterol in the inner nuclear membrane, have been suggested to provide a resting place for active chromatin and

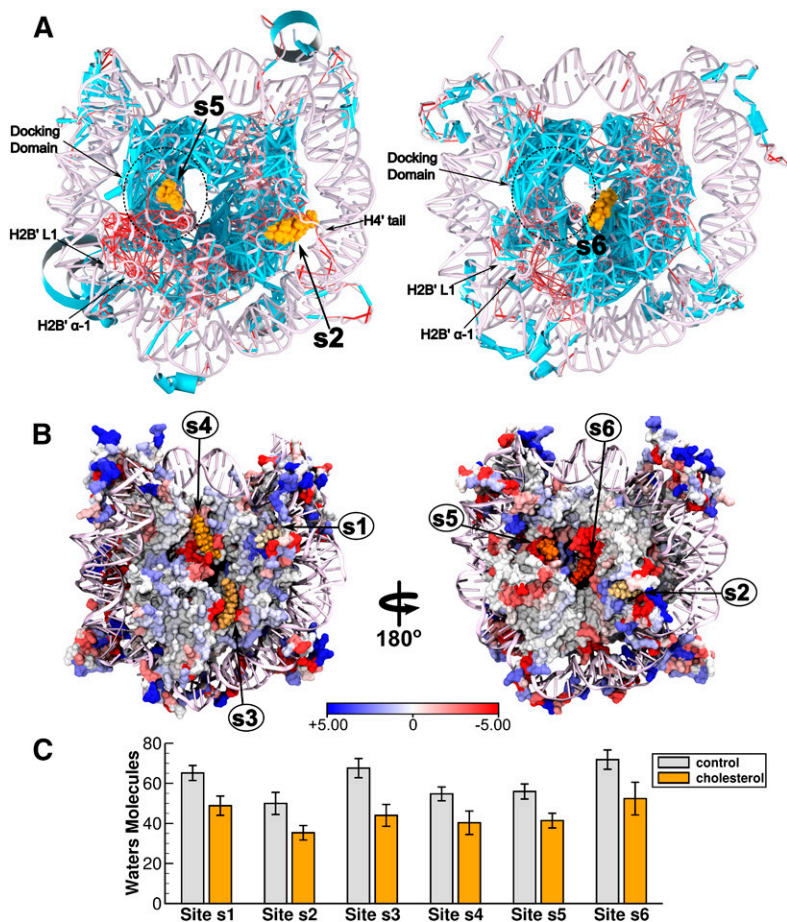



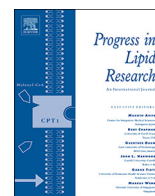
Fig. 4. Position correlation analysis and hydration properties of the nucleosome. A: Shown are the cholesterol-induced $C\alpha$ - $C\alpha$ position correlation changes of the nucleosome. Edge thickness distinguishes between enhanced (light blue) and weakened (red) correlation changes due to cholesterol binding. By binding at sites s2, s5, and s6, note that cholesterol reduces coupling of specific regions nearby the docking domain and the H4 tail in both simulation systems. Here, $C\alpha$ - $C\alpha$ correlation values were calculated with Carma (35). B: Cholesterol-induced changes on the net number of water molecules bound to NCP residues within a cut-off distance of 3.5 Å. C: Average number of water molecules occupying each of the binding sites within a cut-off distance of 10 Å. Averages and associated standard errors were computed from 200 ns of simulation.

transcription factors (33). In addition, the tethering of chromatin to the nuclear periphery has an important role in the chromatin flow by defining changes in nuclear shape (34). We could raise the hypothesis that cholesterol might trigger chromatin structural changes that may contribute to the intra-nuclear organization of chromatin. It would be very interesting to precisely follow the changes in nuclear cholesterol circadian rhythm following the chromatin structural changes and nuclear localization, correlating them with the transcriptional activity. 

The authors are grateful to LME/LNNano for the use of the electron microscopy facility. The authors thank Rodrigo Portugal for technical assistance, Daniela Rhodes for providing the DNA constructions, Carlos Pantoja for critical reading of the manuscript, and the anonymous reviewers for the important comments and suggestions on the manuscript.

REFERENCES

- Voss, T. C., and G. L. Hager. 2014. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.* **15**: 69–81.
- Korolev, N., A. Allahverdi, Y. Yang, Y. Fan, A. P. Lyubartsev, and L. Nordenskiöld. 2010. Electrostatic origin of salt-induced nucleosome array compaction. *Biophys. J.* **99**: 1896–1905.
- Routh, A., S. Sandin, and D. Rhodes. 2008. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. USA.* **105**: 8872–8877.
- Bartman, C. R., and G. A. Blobel. 2015. Perturbing chromatin structure to understand mechanisms of gene expression. *Cold Spring Harb. Symp. Quant. Biol.* **80**: 207–212.
- da Silva, I. T. G., P. S. L. de Oliveira, and G. M. Santos. 2015. Featuring the nucleosome surface as a therapeutic target. *Trends Pharmacol. Sci.* **36**: 263–269.
- Finch, J. T., and A. Klug. 1976. Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA.* **73**: 1897–1901.
- Albi, E., and M. P. Viola Magni. 2004. The role of intranuclear lipids. *Biol. Cell.* **96**: 657–667.
- Ledeer, R. W., and G. Wu. 2004. Nuclear lipids: key signaling effectors in the nervous system and other tissues. *J. Lipid Res.* **45**: 1–8.
- Erickson, S. K., A. M. Davison, and R. G. Gould. 1975. Correlation of rat liver chromatin-bound free and esterified cholesterol with the circadian rhythm of cholesterol biosynthesis in the rat. *Biochim. Biophys. Acta.* **409**: 59–67.
- Regenass-Klotz, M., and H. J. Heiniger. 1984. Specific binding of cholesterol to chromatin prepared from mouse spleen cells. *Can. J. Biochem. Cell Biol.* **62**: 94–99.
- Zaina, S., K. B. V. Døssing, M. Wickstro, and G. Lund. 2005. Chromatin modification by lipids and lipoprotein components: an initiating event in atherogenesis? *Curr. Opin. Lipidol.* **16**: 549–553.
- Jones, D. R., and N. Divecha. 2004. Linking lipids to chromatin. *Curr. Opin. Genet. Dev.* **14**: 196–202.
- Huynh, V. A. T., P. J. J. Robinson, and D. Rhodes. 2005. A method for the in vitro reconstitution of a defined “30 nm” chromatin fibre containing stoichiometric amounts of the linker histone. *J. Mol. Biol.* **345**: 957–968.
- Lowary, P. T., and J. Widom. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**: 19–42.
- Richmond, T. J., and C. A. Davey. 2003. The structure of DNA in the nucleosome core. *Nature.* **423**: 145–150.
- Kallen, J. A., J.-M. Schlaepfli, F. Bitsch, S. Geisse, M. Geiser, I. Delhon, and B. Fournier. 2002. X-ray structure of the hROR α LBD at 1.63 Å: structural and functional data that cholesterol or a cholesterol derivative is the natural ligand of ROR α . *Structure.* **10**: 1697–1707.
- Barber, A. F., Q. Liang, C. Amaral, W. Treptow, and M. Covarrubias. 2011. Molecular mapping of general anesthetic sites in a voltage-gated ion channel. *Biophys. J.* **101**: 1613–1622.
- Trott, O., and A. J. Olson. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**: 455–461.
- Åqvist, J., C. Medina, and J.-E. Samuelsson. 1994. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **7**: 385–391.
- Carlsson, J., and J. Åqvist. 2006. Calculations of solute and solvent entropies from molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **8**: 5385–5395.
- Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**: 1781–1802.
- Huang, J., and A. D. MacKerell, Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**: 2135–2145.
- Lim, J. B., B. Rogaski, and J. B. Klauda. 2012. Update of the cholesterol force field parameters in CHARMM. *J. Phys. Chem. B.* **116**: 203–210.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**: 926–935.
- Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**: 10089–10092.
- Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**: 33–38.
- Fonseca, A. C. R. G., T. Proença, R. Resende, C. R. Oliveira, and C. M. F. Pereira. 2009. Neuroprotective effects of statins in an in vitro model of Alzheimer’s disease. *J. Alzheimers Dis.* **17**: 503–517.
- Ding, J., D. Jiang, M. Kurczy, J. Nalepka, B. Dudley, E. I. Merkel, F. D. Porter, A. G. Ewing, N. Winograd, J. Burgess, et al. 2008. Inhibition of HMG CoA reductase reveals an unexpected role for cholesterol during PGC migration in the mouse. *BMC Dev. Biol.* **8**: 120.
- Davey, C. A., D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**: 1097–1113.
- Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* **389**: 251–260.
- Towbin, B. D., A. Gonzalez-Sandoval, and S. M. Gasser. 2013. Mechanisms of heterochromatin subnuclear localization. *Trends Biochem. Sci.* **38**: 356–363.
- Casolari, J. M., C. R. Brown, S. Komili, J. West, H. Hieronymus, and P. A. Silver. 2004. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell.* **117**: 427–439.
- Cascianelli, G., M. Villani, M. Tosti, F. Marini, E. Bartocchini, M. V. Magni, and E. Albi. 2008. Lipid microdomains in cell nucleus. *Mol. Biol. Cell.* **19**: 5289–5295.
- Schreiner, S. M., P. K. Koo, Y. Zhao, S. G. J. Mochrie, and M. C. King. 2015. The tethering of chromatin to the nuclear envelope supports nuclear mechanics. *Nat. Commun.* **6**: 7159.
- Glykos, N. M. 2006. Software news and updates. Carma: a molecular dynamics analysis program. *J. Comput. Chem.* **27**: 1765–1768.



Review

Fat nucleosome: Role of lipids on chromatin

Vinicius Fernandes^{a,b}, Kaian Teles^a, Camyla Ribeiro^a, Werner Treptow^b, Guilherme Santos^{a,*}^a Laboratório de Farmacologia Molecular, Departamento de Farmácia, Universidade de Brasília, Brasília 70919-970, Brazil^b Laboratório de Biologia Teórica e Computacional, Departamento de Biologia Celular, Universidade de Brasília, DF 70910-900, Brasília, Brazil

ARTICLE INFO

Keywords:

Lipids
Nucleosome
Chromatin

ABSTRACT

Structural changes in chromatin regulate gene expression and define phenotypic outcomes. Molecules that bind to the nucleosome, the complex of DNA and histone proteins, are key modulators of chromatin structure. Most recently, the formation of condensed chromatin regions based on phase-separation in the cell, a basic physical mechanism, was proposed. Increased understanding of the mechanisms of interaction between chromatin and lipids suggest that small lipid molecules, such as cholesterol and short-chain fatty acids, can regulate important nuclear functions. New biophysical data has suggested that cholesterol interacts with nucleosome through multiple binding sites and affects chromatin structure in vitro. Regardless of the mechanism of how lipids bind to chromatin, there is currently little awareness that lipids may be stored in chromatin and influence its state. Focusing on lipids that bind to nuclear receptors, clinically relevant transcription factors, we discuss the potential interactions of the nucleosome with steroid hormones, bile acids and fatty acids, which suggest that other lipid chemotypes may also impact chromatin structure through binding to common sites on the nucleosome. Herein, we review the main impacts of lipids on the nuclear environment, emphasizing its role on chromatin architecture. We postulate that lipids that bind to nucleosomes and affect chromatin states are likely to be worth investigating as tools to modify disease phenotypes at a molecular level.

1. Chromatin and regulators

Genetic information is structurally organized within chromatin, which dynamically regulates how, when and where gene expression occurs in the cell. Besides the regulation of gene expression, chromatin is also responsible for regulating genome maintenance [1]. Highly basic proteins and DNA combine to form the nucleosome core particle (NCP) thus generating the fundamental repetitive unit of chromatin. The NCP represents the first level of DNA compaction, followed by a cooperative nucleosome interaction to form higher-order chromatin structure. The current hierarchical model is that 11-nm fibers composed of DNA-nucleosome polymers may fold into 30-nm fibers that further fold into 120-nm, 300-nm to 700-nm chromatids and, ultimately, into mitotic chromosomes. However, it was recently demonstrated, using a new DNA-labeling method that chromatin is not as well organized as previously thought; it comprises a flexible and disordered 5–24-nm granular chain that is packed together, thus challenging the long-standing textbook model of higher-order chromatin structure [2].

In an open state, chromatin is more susceptible to loading with the proteins that trigger transcription initiation, DNA repair and processes that regulate genome stability. In contrast, the close state of chromatin, heterochromatin, may provide steric exclusion of transcription factors

and regulatory proteins such as nuclear receptors and RNA polymerases from target DNA, leading to gene silencing [3]. Indeed, the relaxed and the condensed chromatin structures provide the cell with a strategic plan to sharply regulate the access of proteins to DNA. Nevertheless, this is a simplistic picture, given that chromatin structure dynamics alone is not enough to produce predictive patterns of gene expression and phenotypic outcomes [4]. Moreover, the properties of the nuclear molecules that affect chromatin structure and fundamental processes, such as genome stability and transcription remain to be fully-characterized.

Chromatin dynamics are determined by a plethora of Nucleosome Binding Proteins, NBPs, several of which have already been characterized at the atomic level [5]. In addition to NBPs, the nuclear environment also comprises many small molecules with different chemical natures that can directly interact with nucleosomes, such as Mg²⁺ and lipids [6,7]. Thus, the molecular composition of the nuclear environment may have considerable impact on chromatin architecture, indeed, a phase diagram of the NCP published over a decade ago suggests that slight changes in ionic conditions may have significant impacts on nucleosome:nucleosome interactions [8].

More recently, an elegant hypothesis to explain the formation of heterochromatin regions was proposed by two different groups based

* Corresponding author.

E-mail address: gsantos@unb.br (G. Santos).

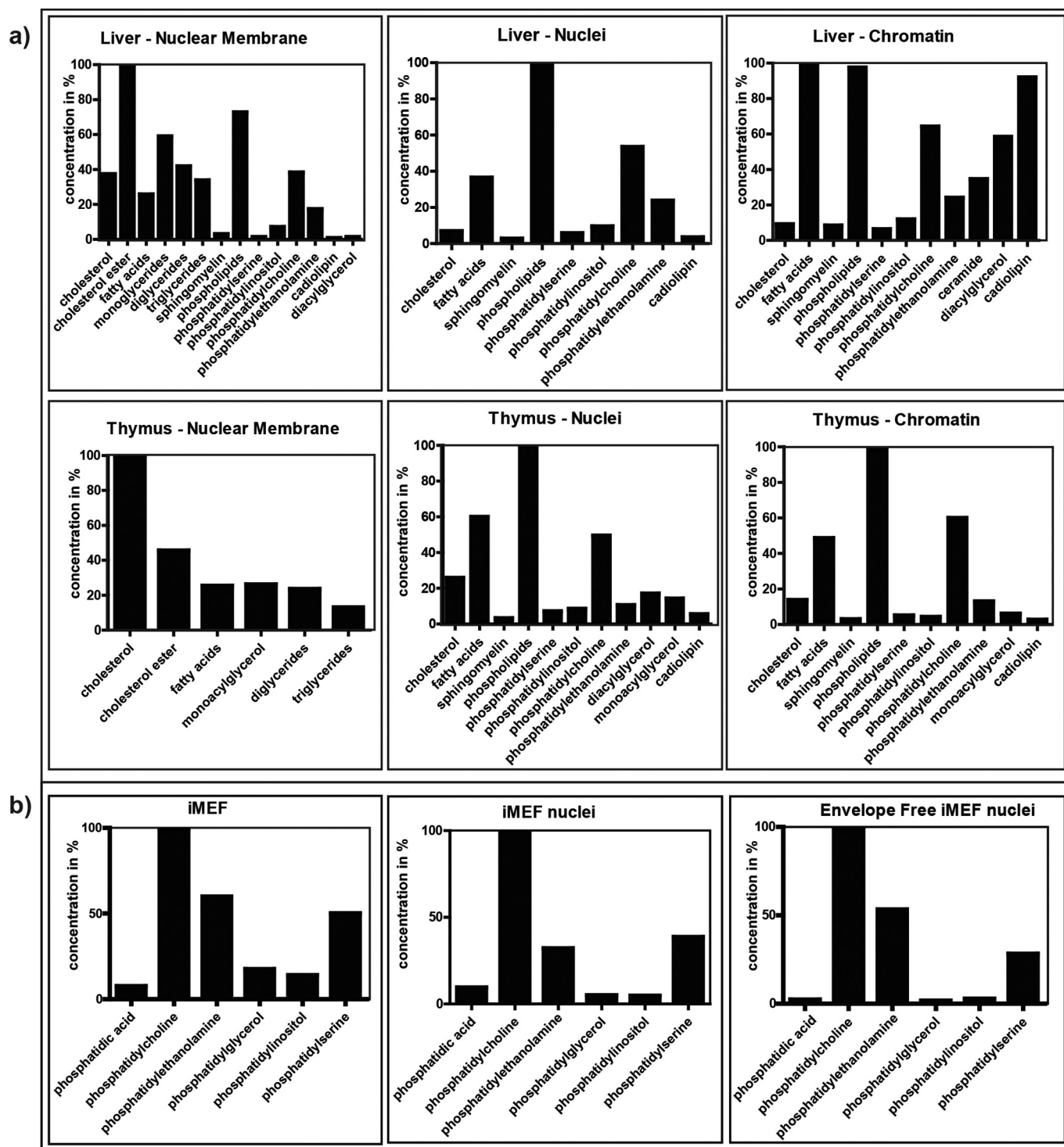


Fig. 1. Lipidic profile of nuclear compartments in liver, thymus and embryonic cells. A- The concentration of lipid is expressed in relation to protein:tissue, normalized in percentage related to the highest lipid concentration, from liver and thymus cells. Nucleus includes data from whole nuclear environment, including chromatin and nuclear membrane. References: Liver chromatin: [18,19,21] Liver nucleus: [18,19] Thymus chromatin: [20,22–24,27,28] Thymus nucleus: [22–25,27,28] Liver nuclear membrane: [17,19,20] Thymus nuclear membrane [20]. B- The concentration of phospholipids is expressed in percentage related to phosphatidylcholine from immortalized embryonic fibroblasts [26].

on phase-separation in the cell, a basic physical mechanism [9,10]. It was shown that a linker histone HP1, which is known as a key factor for inducing chromatin compaction, has the ability to form phase-separated droplets. Native, unmodified HP1 is soluble but phosphorylation of its N-terminal domain and binding to DNA induce the formation of phase-separated droplets. This process would potentially be responsible

for the formation of compartments in the cell that do not possess a membrane or other physical barrier, and, in turn mediate gene silence induced by HP1 heterochromatin formation. This mechanism of segregation is analogous to oil de-mixing from water. While it seems clear that HP1 droplets are able to physically sequester chromatin, inducing compaction, the precise nature of the physico-chemical environment of

the phase-separated heterochromatin remains obscure.

Nuclear lipid droplets were described in 2012 as potential sub-nuclear domains. These droplets which as well as lipid also contain proteins, and these could provide a repository for the storage of lipids in the nucleus [11].

Increased understanding of the mechanisms of interaction between chromatin and lipids suggest that small lipid molecules, such as cholesterol and short-chain fatty acids, can regulate important nuclear functions. The role of lipids on the nuclear environment is prominent in protein:lipids complexes, such as Nuclear Receptors and their ligands. However, there is no clear perception of whether lipids may impact the nucleosome and chromatin structure, regulating gene expression and genome maintenance. Herein, we review the concentration of several lipids in the nuclear environment of different tissues, highlighting the lipids bound to chromatin. We also briefly present some ligands of nuclear receptor, discussing how they could bind to the nucleosome and chromatin. Understanding the role of lipids on the nuclear environment may be critical to comprehend how chromatin states are determined and how these states may subsequently affect phenotypic outcomes.

2. Lipids in the nuclear environment: lipids and chromatin

Lipid droplets are recognized as repositories for the storage of lipids for use on demand and they also participate in several physiological processes. Although well characterized in the cytoplasm, the existence of nuclear lipid droplets (nLD) has been recently described, functioning as a modulator of lipid metabolism in the nucleus. Lipid droplets interact with numerous proteins and specifically participate in the regulation of nuclear events [12]. Indeed, histones interact with lipid droplets during oogenesis in *Drosophila* embryos, which causes histone sequestration outside chromatin during embryogenesis [13].

In 2013 [11], nuclei from rat hepatocyte were separated from cytosol using sucrose gradient ultracentrifugation, neutral lipids contained in nLDs were stained and visualized with confocal microscopy. These nLDs were organized around a hydrophobic core of triacylglycerol and cholesteryl esters with oleic acid, surrounded by further cholesteryl esters, polar lipids and proteins, but the protein composition of nLDs has not been fully characterized.

Cascianelli et al. [14] have shown that nuclear lipid microdomains, containing sphingomyelin, cholesterol and proteins in the inner nuclear membrane, may function as scaffolds for RNA synthesis. The same group hypothesized that nuclear lipid microdomains could act as sites for chromatin anchorage, which would have important functions during the cell cycle. Moreover, it was shown that the tethering of chromatin to the nuclear periphery plays a crucial function in the chromatin flow by changing the nuclear shape [15].

Seminal studies in rabbits and rats have shown the presence of different lipids, cholesterol, fatty acids and phospholipids in the nuclear environment and their concentrations have been determined. It has been proposed that the ratio of cholesterol to phospholipids would have a great structural impact in the nucleus, rather than just a metabolic function [16]. Data for the concentration range of lipids in the different compartments in the nuclear environment is available for different tissues, herein we focus on the liver and thymus cells that literature vastly covered, and these are summarized in Fig. 1 [17–26]. In addition, we also analysed data from murine embryonic fibroblast, which was one of the most thorough examinations, using electrospray ionization mass spectrometry (ESI-MS), to determine the intra-nuclear lipid content of cells (Fig. 1b).

The precise determination of lipid concentration in the cell nucleus is clearly challenging, with several studies having found different values for the same lipid in the same conditions [22,27–29]. Herein, the analyses were made based on the relation of protein:lipid or protein:tissue showed in the articles. However, in this way it was difficult to make a clear qualitative view of the distribution of the lipids in the cell. Therefore, to have a clearer picture of the lipids distribution, the y-

axis, which gives the concentration, was converted to percentage, considering 100% as the most prevalent lipid in each cell compartment (Fig. 1a) or in percentage related to phosphatidylcholine (Fig. 1b). This allowed us to compare the amount of different lipids in the same nuclear environment. In thymus and liver cells, predominately phospholipids are associated with chromatin. Indeed, phospholipids appeared to be the most abundant lipid, regardless of cell type and nuclear compartment (Fig. 1b) (detailed phospholipids characterization is found in [26]). Interestingly, cardiolipin, which was found to be almost exclusively in the inner mitochondrial membrane, is the third most abundant lipid associated with liver chromatin, in contrast to thymus that has very low cardiolipin concentration. Taking all together, it is noticeable that the distribution of lipids is cell type dependent.

The relevance of different lipid concentrations for intracellular signal during the cell cycle has been debated for a long time, suggesting a role for these molecules as structural parts of the chromatin [30,31].

Importantly, *in vitro* studies demonstrated that neutral lipids can bind specifically to the double helix DNA, but not triplex DNA. Spectroscopic and surface plasmon resonance studies indicate that linoleic acid is able to tightly bind to the minor groove of the double helix DNA [32]. If these findings are representative of the physiological reality in the nuclear environment, binding of lipids might have a profound impact on modulation of gene expression.

Following these findings, several more studies have attempted to elucidate the potential function of many lipids in the nucleus [29]. Mainly, gene expression was shown to be regulated by different lipids such as gamma linoleic acid and phosphatidylserine. Furthermore, a significant role for lipids on DNA polymerase modulation and DNA structure was demonstrated, and this would directly affect the transcriptional process.

2.1. Lipids and gene expression: steroids hormones, bile acid and fatty acids

It is well established that lipids regulate gene expression through transcription factors, such as nuclear receptors (NRs). NR ligands include lipids, such as cholesterol metabolites and fatty acids [33]. Moreover, phospholipids, one of the most abundant lipid molecules in the nucleus, may directly regulate NRs activity, evidencing an important action of these non-cholesterol-based molecules in modulating gene expression (reviewed in [34]).

Herein, we briefly point out some of the members of NR superfamily, which have lipids as ligands, to emphasize the biological and clinical relevance of its lipids in the nuclear environment, and provide references for more detailed information.

This superfamily of receptors is highly clinically relevant. The classical view of NR mechanism of transcription regulation is that NRs are activated by their ligands, bind to the promoter/enhancer region of target genes and recruit co-regulators to modulate both transcription activation and transcription repression [35]. We focus on three nuclear receptors that have lipids as ligands, peroxisome proliferator-activated receptors (PPARs), Liver X Receptor (LXR), Farnesoid X Receptor (FXR) and Estrogen Receptor (ER).

PPARs are involved in adipogenesis and lipid and carbohydrate metabolism, with high clinical relevance as insulin sensitizers [36]. PPAR α is a key regulator of hepatic lipid metabolism through the modulation of several genes related to mitochondrial and peroxisomal fatty acid β -oxidation. PPAR α is also involved in the inflammatory pathway and glucose metabolism. PPAR δ function is crucial for the glucose metabolism in skeletal, heart and pancreatic β -cells [37]. PPAR γ is important for cellular differentiation, lipogenesis and insulin sensitivity in adipose tissue [38]. In 2008, structures of PPAR γ bound to oxidized fatty acids were solved, suggesting that these fatty acids are natural ligands for this receptor. Interestingly, it was shown that PPAR γ can bind two fatty acids at the same time and couple covalently with conjugated oxo fatty acids, which would be particularly effective activators of PPAR γ [39]. Importantly, a component of Low Density

Lipoprotein (LDL) particles, Oxidized phosphatidylcholine (oxPC), is also able to bind and modulate PPAR α [40] and PPAR γ [41]. Although some phospholipids are found in the nuclear environment at lower levels, its importance on NRs activity, such as 1-Palmitoyl 2-Oleoyl PC (POPC) on PPAR α , has been demonstrated [42].

Another nuclear receptor, LXR (Liver X Receptor), which is highly expressed in the liver and adipose tissue, is also target for lipids such as oxysterols and cholesterol derivatives. LXR is able to target genes involved in cholesterol metabolism and is considered a sensor for intracellular cholesterol levels [43,44]. LXR activation may regulate target genes that lead to decrease cholesterol and increase HDL levels in the plasma [45,46].

FXR (Farnesoid X Receptor) has bile acids as ligands. This nuclear receptor may activate the VLDL receptor gene and other genes involved in the metabolism of lipoproteins [47]. It is expressed in the liver, kidney and intestines. FXR was described as a sensor for regulation of bile acids in the enterohepatic circulation [48,49].

Besides those NRs, the ER (Estrogen Receptor) is a key transcription factor that is regulated by steroid hormones, which are also classified as lipids. Selective ER ligands are highly promising agents for the treatment of the majority of breast cancers, as well as for cardiovascular, inflammatory, and neurodegenerative diseases. The ER is critical for the regulation of genes engaged in cell growth and endocrine response [33,50].

Although some of NRs ligands play a crucial role as second messengers, it is accepted that this extracellular molecules may also act as classic first messengers on NRs. From these examples amongst Nuclear Receptors briefly discussed here, it is clear that many lipidic ligands can regulate cell functioning, especially gene expression, via transcription factors present in the nucleus. Considering that these ligands may permeate the nuclear environment, it is plausible to believe that they could also bind the lipophilic regions of chromatin affecting its structure and, ultimately, gene expression. Moreover, it is probable that chromatin may store lipids that can act as NR ligands, which would then be available according to the NR's requirement.

3. Complexes of lipids with the nucleosome: cholesterol, steroids hormones, bile acid and fatty acids

In 1975, chromatin fibers purified from rats treated with labeled cholesterol revealed the association of this macromolecule with cholesterol [51]. Later, it was suggested that cholesterol binds to chromatin through one or more proteins rather than to nucleic acid [52].

Recently, some biophysical effects of cholesterol on chromatin were demonstrated [7], suggesting that cholesterol binds directly to chromatin and facilitates *in vitro* formation of mononucleosomes and chromatin fibers. Importantly, cholesterol induced the compaction of open chromatin fibers (Fig. 2). Molecular Dynamics experiments suggested six binding sites for cholesterol on the nucleosome. For all of these sites, the estimated binding free-energies indicated that cholesterol interacts with the nucleosome with low affinity. By binding at these locations cholesterol resided near amino acids from key nucleosome regions, such as H4 tail, the acidic patch (a negatively charged region on the nucleosome surface formed by six H2A and two H2B residues) and the docking domain in between the histone clusters H3-H4 and H2A-H2B (Fig. 3). Further computational comparative analysis in presence and absence of cholesterol led to the hypothesis that cholesterol effects on chromatin could be due to (i) dewetting of nucleosome surface that might enhance interactions between key structural hydrophobic amino acids and (ii) an additional mechanism involving specific binding locations of the H4 tail, known for internucleosomal interactions.

Moreover, as previously discussed, other small lipid molecules are present in the nucleus, and also are likely to interfere with chromatin structure [53]. Therefore, presumably other lipids types could affect chromatin structure through direct interaction with the nucleosome.

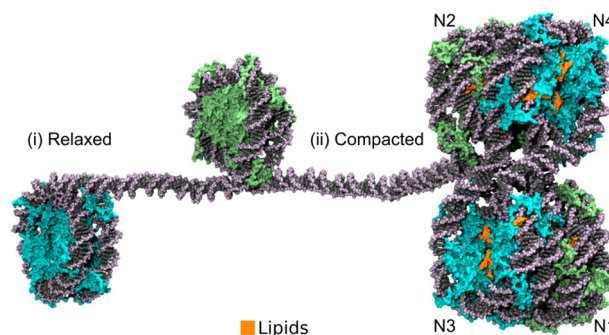


Fig. 2. Pictorial model for the potential role of cholesterol in modulating chromatin compaction. Two states of chromatin are shown, (i) relaxed fiber (left), in which the nucleosomes are more distant from each other; (ii) and compacted fiber (right), where the presence of cholesterol (orange) results in heterochromatin formation. The histone octamer are represented in blue and green to distinguish the nucleosomes. The tetranucleosome is numbered from N1 to N4 in the compacted state. The compacted fiber model was generated from the tetranucleosome structure revealed by Schallch and colleagues [62]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Consistent with that notion, exploratory docking calculations between lipid-chemotypes, other than cholesterol, and the nucleosome suggested that cholesterol and other lipids may bind at common sites on the nucleosome surface (unpublished data). Although preliminary, these calculations support that cholesterol binding sites can be promiscuous allowing binding of other lipids, including steroids hormones, bile acid, fatty acids and phospholipids. In the case of Nuclear Receptor ligands, they could act directly via chromatin, and not just on Nuclear Receptors. Furthermore, it may suggest that chromatin can function as a repository or depot, for these lipids, becoming available to NRs when they are needed.

Several other small-molecules have already been characterized as binding to chromatin and theoretical and molecular techniques have been very useful to show the interaction of ions and small peptides with chromatin. Fan and co-workers used Molecular Dynamics simulations of coarse-grained nucleosomes (simplified representation of chromatin compared to the classic atomistic models) to demonstrate that divalent (Mg^{2+}) and trivalent ($CoHex^{3+}$) ions resulted in aggregates of NCPs. For the trivalent ion, all 10 NCPs simulated could form aggregates, inducing phase-separation [54]. In addition, divalent metal ions bind to chromatin and may compete with histone tails to access the DNA grooves, which could alter the chromatin states [55]. Taken all together, this variety of molecules that interact with chromatin reinforce the view that chromatin could represent a depot for many different small molecules, in addition to the lipids which are the focus of this review.

4. Clinical outcome of lipids and chromatin interaction

Lipids play essential roles in the cell and a wide range of diseases has been associated with lipid metabolism. Plasma cholesterol and triglyceride levels are used as predictors of metabolic disorders and vascular disease risk and have been crucial for physicians over the past century [56]. However, it is still unknown how lipids with multiple functions such as cholesterol that is essential for steroid hormone production, the structural integrity of cellular membranes, and cell growth, may affect chromatin structure and, ultimately, clinical outcomes.

What we know so far is that cholesterol and other lipids may have a direct impact on the chromatin architecture *in vitro*, triggering changes that may contribute the intra-nuclear organization of chromatin [7]. It is also clear that the understanding of changes in chromatin dynamics may uncover mechanisms of diseases and shed light on potential treatment alternatives [57]. Several histone post-translational modifications responsible for changes in the chromatin structure have been

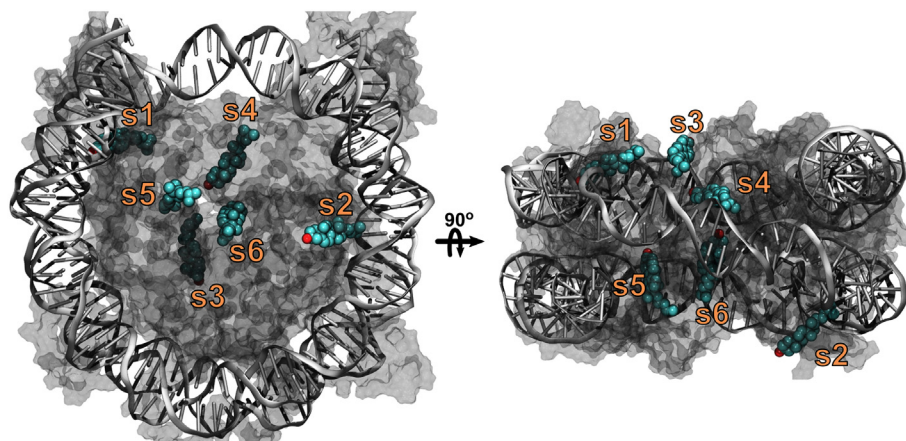


Fig. 3. Cholesterol binding on nucleosome surface. The nucleosome surface is shown (gray) with six cholesterol molecules (colored by atom type) bound at different regions (sites s1 to s6). Adapted from [7].

associated with cancer and other diseases, directly associating chromatin architectural changes with disease mechanisms. The epidrugs [58], an emergent class of drugs, target enzymes and other proteins that are able to control epigenetic changes and hence gene expression for disease treatment. Observing the great clinical importance of epidrugs, it is plausible to consider that any molecule that interacts with the nucleosome and affects chromatin dynamics would also be able to cause disease phenotypes. For the time being it is unclear if the actions of lipids on chromatin are potential therapeutical targets. The structural characterization, at atomic level, of the lipids binding mode to nucleosome could provide new insights in the development of hybrid molecules for modulating chromatin structure [59].

Moreover, detailed information about the nuclear environment under different physiological and pathological states is needed before this field can advance significantly toward clinical utility. It will be very important to precisely determine the variations in nuclear lipid concentrations and localization during the circadian rhythm, following the chromatin structural changes and nuclear localization and finally correlating them with transcriptional activity and phenotypic outcome. Understanding how lipids act on chromatin is a promising avenue for the development of new treatments.

5. Concluding remarks and future perspectives

The modern view of chromatin modulation by NBPs and post-translational modifications of histones must be updated. An avalanche of data concerning the manner of chromatin dynamics leads us to re-interpret the simplistic scenario, in which protein complexes dictate chromatin states.

The roles of several NBPs, key players in modulating chromatin structure, have been well characterized at atomic level. These structures revealed details of the interaction of NBPs with the nucleosome surface, highlighting the acidic patch as the principal protein-docking region. However, this region does not seem to be exclusive for NBPs. In silico experiments found multiple binding sites for cholesterol on the nucleosome, including the acidic patch.

Mounting evidences support that different lipids, such as cholesterol, fatty acids and phospholipids may interact directly with a number of nuclear protein structures to impact their function in a specific and concentration-dependent manner. Exemplars of nuclear protein are Nuclear Receptors superfamily, which was briefly discussed here. Differently, lipids appear here to interact with chromatin in a promiscuous non-specific way, through multiple low-affinity receptor sites. Binding energies as estimated from docking calculations support that lipid-chromatin interactions are low affinity, in the range of 10^{-5} mM (unpublished data). Although an entire new series of high-resolution

experimental and theoretical studies, as done with cholesterol, will be needed to shed light on the molecular nature of lipid-chromatin interactions. However, these preliminary observations thus support that lipid-chemotypes, may bind at common sites on the nucleosome surface. Still consistent with that notion is the fascinating possibility that via non-specific low-affinity interactions, chromatin may work as a lipid depot for nuclear proteins.

The 4D Nucleome Network [60] was recently presented as a potential third phase of the human genome project, aiming to develop and use a wide array of technologies to characterize the structure and dynamics of the human and mouse genomes. The main idea is to integrate several disciplines, from biology to mathematics, to decipher how the chromatin is organized and function inside the nucleus in space and time. It seems a fantastic and ambitious project, which will render many new insights that can ultimately be useful to understand how the human genome is packed in the cell in a functional way.

However, several questions regarding the molecules that compose the nuclear environment, and may affect chromatin architecture, remain to be answered. Firstly, getting a clear map of the contents in the nuclear environment in different conditions would provide us with the opportunity to make solid correlations with many events in the cell, including chromatin states, gene expression and DNA stability processes. At the same time, better characterization of substances, such as lipids, ions and a plethora of proteins that drive phase-separation properties on the chromatin structure is necessary. Independently whether the lipids retain or not a direct action on chromatin, we should at least consider that under a certain concentration a hydrophobic molecule would be able to affect the manner that macromolecules, such as the nucleosome, will interact with each other and, consequently, modulating its state.

At the same line of research, it was shown that protein domains of low-complexity, which don't comprise a balanced proportion of all 20 amino acids, may polymerize into cross- β structure. This, in turn, can be the driving force for hydrogel and liquid-like droplet formation [61]. We could then question whether lipids could cooperate with protein-based hydrogels formation and achieve the same role as the phase-separations in the cell, modulating the chromatin architecture.

Lately, by reasoning that any molecule that binds to the nucleosome and affects the chromatin state is able to tackle disease phenotypes at a molecular level, we anticipate that correlations of the outcome from the 4D nucleosome project with a map of nuclear molecules in certain space and time will advance rational drug design and development of new treatments for a wide range of diseases.

The rapid development of various technologies, from computational to structural biology, has occurred in the past decade to investigate the role of many substances on chromatin. Hopefully, the combination of

all of these areas will assist to overcome the arduous tasks of obtaining a clearer overview of cell function.

Acknowledgments

The authors are grateful to Tony Warne and Carlos Pantoja for critical reading of the manuscript.

References

- [1] W.K.M. Lai, B.F. Pugh, Understanding nucleosome dynamics and their links to gene expression and DNA replication, *Nat Rev Mol Cell Biol* (2017), <http://dx.doi.org/10.1038/nrm.2017.47>.
- [2] H.D. Ou, et al., ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells, *Science* 357 (2017) eaag0025.
- [3] T.C. Voss, G.L. Hager, Dynamic regulation of transcriptional states by chromatin and transcription factors, *Nat Rev Genet* 15 (2014) 69–81.
- [4] C.R. Bartman, G.A. Blobel, Perturbing chromatin structure to understand mechanisms of gene expression, *Cold Spring Harb Symp Quant Biol* 80 (2015) 207–212.
- [5] R.K. McGinty, S. Tan, Recognition of the nucleosome by chromatin factors and enzymes, *Curr Opin Struct Biol* 37 (2016) 54–61.
- [6] S. Tan, C.A. Davey, Nucleosome structural studies, *Curr Opin Struct Biol* 21 (2011) 128–136.
- [7] I.T.G. Silva, et al., Biophysical studies of cholesterol effects on chromatin, *J Lipid Res* 58 (2017) 934–940.
- [8] A. Bertin, et al., Structure and phase diagram of nucleosome core particles aggregated by multivalent cations, *Biophys J* 93 (2007) 3652–3663.
- [9] A.R. Strom, et al., Phase separation drives heterochromatin domain formation, *Nature* 547 (2017) 241–245.
- [10] A.G. Larson, et al., Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin, *Nature* 547 (2017) 236–240.
- [11] J.P. Layerenza, et al., Nuclear lipid droplets: a novel nuclear domain, *Biochim Biophys Acta* 1831 (2013) 327–340.
- [12] M.A. Welte, Expanding roles for lipid droplets, *Curr Biol* 25 (2015) 470–481.
- [13] Z. Li, et al., Drosophila lipid droplets buffer the H2Av supply to protect early embryonic development, *Curr Biol* 24 (2014) 1485–1491.
- [14] G. Cascianelli, et al., Lipid microdomains in cell nucleus, *Mol Biol Cell* 19 (2008) 5289–5295.
- [15] S.M. Schreiner, et al., The tethering of chromatin to the nuclear envelope supports nuclear mechanics, *Nat Commun* 6 (2015) 7159.
- [16] C.A. Stoneburg, Lipids of the cell nuclei, *J. Biol. Chem.* 129 (1939) 189–196.
- [17] E. Albi, et al., A possible role of cholesterol-sphingomyelin/phosphatidylcholine in nuclear matrix during rat liver regeneration, *J Hepatol* 38 (2003) 623–628.
- [18] I.K. Kolomyitseva, et al., A influence of β -carotene on liver nuclei lipids in normal and by the action of γ -radiation at a low dose-rate, *Dokl Akad Nauk SSSR* 363 (1998) 594–651.
- [19] E. Albi, et al., Reverse sphingomyelin-synthase in rat liver chromatin, *FEBS Lett* 549 (2003) 152–156.
- [20] E.S. Gevorgyan, et al., Cisplatin in vivo action on content of neutral lipids in rat liver and thymus nuclear membranes, *Biol J Armenia* 65 (2013) 99–103.
- [21] T.P. Kulagina, et al., The action of continuous γ -radiation at a low dose-rate and β -carotene on lipid metabolism of rats thymus nuclei, *Bull Exp Biol Med* 126 (1998) 311–313.
- [22] I.K. Kolomyitseva, et al., Nuclear and chromatin lipids: metabolism in normal and γ -irradiated rats, *Bioelectrochemistry* 58 (2002) 31–39.
- [23] T.P. Kulagina, Activation of lipid metabolism of nuclei and chromatin in thymocytes of rats subjected to chronic low dose intensity gamma-irradiation, *Biochem Mosc* 62 (1997) 1034–1038.
- [24] T.P. Kulagina, Phospholipids in rat thymus nuclei and chromatin at remote times following gamma-irradiation, *Radiobiologiya* 30 (1990) 745–748.
- [25] T.P. Kulagina, et al., Nonmonotonous changes of nuclei lipid metabolism under chronic action of γ -radiation at dose rate of a 3 cGy/day, *Dokl Akad Nauk SSSR* 370 (2000) 403–406.
- [26] E.K. Tribble, et al., Quantitative profiling of the endonuclear glycerophospholipidome of murine embryonic fibroblasts, *J. Lipid Res.* 57 (2016) 1492–1506.
- [27] T.P. Kulagina, Cholesterol and fatty acids of nuclei and chromatin of rat thymus at remote times following γ -irradiation, *Radiobiologiya* 30 (1990) 317–320.
- [28] T.P. Kulagina, et al., Difference in metabolic pools of cholesterol and free fatty acids of rat thymocyte chromatin and nuclear membranes, *Mol Biol (Mosk)* 28 (1994) 714–719.
- [29] A. Kagansky, et al., Lipids contribute to epigenetic control via chromatin structure and functions, *ScienceOpen Res* 10 (2015) 1–12.
- [30] H.G. Rose, J.H. Frenster, Composition and metabolism of lipids within repressed and active chromatin of interphase lymphocytes, *Biochim Biophys Acta Lipids Lipid Metab* 106 (1965) 577–591.
- [31] S. Furse, G.C. Shearman, Do lipids shape the eukaryotic cell cycle? *Biochim Biophys Acta* 1863 (2018) 9–19.
- [32] R.I. Zhdanov, et al., Lipid and fatty acid profiles of *Pseudomonas aurantiaca* DNA-bound lipids determined by mass spectrometry, *Microbiology* 84 (2015) 34–40.
- [33] M. Warner, et al., Estrogen receptor beta as a pharmaceutical target, *Trends Pharmacol Sci* 38 (2017) 92–99.
- [34] M.K. Crowder, et al., Phospholipid regulation of the nuclear receptor superfamily, *Adv Biol Regul* 63 (2017) 6–14.
- [35] G.M. Santos, et al., Negative regulation by nuclear receptors: a plethora of mechanisms, *Trends Endocrinol Metab* 22 (2011) 87–93.
- [36] A.I. Abushouk, et al., Peroxisome proliferator-activated receptors as therapeutic targets for heart failure, *Biomed Pharmacother* 95 (2017) 692–700.
- [37] L. Poulsen, et al., PPARs: fatty acid sensors controlling metabolism, *Semin Cell Dev Biol* 23 (2012) 631–639.
- [38] C. Janani, B.D. Ranjitha Kumari, PPAR gamma gene—a review, *Diabetes Metab Syndr* 9 (2015) 46–50.
- [39] T. Itoh, et al., Structural basis for the activation of PPAR γ by oxidized fatty acids, *Nat Struct Mol Biol* 15 (2008) 924–931.
- [40] H. Lee, et al., Role for peroxisome proliferator-activated receptor α in oxidized phospholipid-induced synthesis of monocyte chemoattractant Protein-1 and Interleukin-8 by endothelial cells, *Circ Res* 87 (2000) 516–521.
- [41] S.S. Davies, et al., Oxidized alkyl phospholipids are specific, high affinity peroxisome proliferator-activated receptor γ ligands and agonists, *J Biol Chem* 276 (2001) 16015–16023.
- [42] M.V. Chakravarthy, et al., Identification of a physiologically relevant endogenous ligand for PPAR α in liver, *Cell* 138 (2009) 476–488.
- [43] C. Zhao, K. Dahlman-Wright, Liver X receptor in cholesterol metabolism, *J Endocrinol* 204 (2010) 233–240.
- [44] C. Hong, P. Tontonoz, Liver X receptors in lipid metabolism: opportunities for drug discovery, *Nat Rev Drug Discov* 13 (2014) 433–444.
- [45] R. Komati, et al., Ligands of therapeutic utility for the liver X receptors, *Molecules* (2017) 22.
- [46] Z. Ma, et al., Liver X receptors and their agonists: targeting for cholesterol homeostasis and cardiovascular diseases, *Curr Issues Mol Biol* 22 (2017) 41–64.
- [47] F.J. Gonzalez, et al., Intestinal Farnesoid X receptor signaling modulates metabolic disease, *Dig Dis* 35 (2017) 178–184.
- [48] T.Q. de Aguiar Vallim, et al., Pleiotropic roles of bile acids in metabolism, *Cell Metab* 17 (2013) 657–669.
- [49] N. Tanaka, et al., Targeting nuclear receptors for the treatment of fatty liver disease, *Pharmacol Ther* (2017), <http://dx.doi.org/10.1016/j.pharmthera.2017.05.011>.
- [50] J.S. Carroll, et al., Genome-wide analysis of estrogen receptor binding sites, *Nat Genet* 38 (2006) 1289–1297.
- [51] S.K. Erickson, et al., Correlation of rat liver chromatin-bound free and esterified cholesterol with the circadian rhythm of cholesterol biosynthesis in the rat, *Biochim Biophys Acta* 409 (1975) 59–67.
- [52] M. Regenass-Klotz, H.J. Heiniger, Specific binding of cholesterol to chromatin prepared from mouse spleen cells, *Can J Biochem Cell Biol* 62 (1984) 94–99.
- [53] S. Zaina, et al., Chromatin modification by lipids and lipoprotein components: an initiating event in atherogenesis? *Curr Opin Lipidol* 16 (2005) 549–553.
- [54] Y. Fan, et al., An advanced coarse-grained nucleosome core particle model for computer simulations of nucleosome-nucleosome interactions under varying ionic conditions, *PLoS One* 8 (2013) e54228.
- [55] B. Wu, C.A. Davey, Using soft X-rays for a detailed picture of divalent metal binding in the nucleosome, *J Mol Biol* 398 (2010) 633–640.
- [56] W.V. Brown, et al., JCL roundtable: future of the lipid laboratory: using the laboratory to manage the patient (part 2), *J Clin Lipidol* 11 (2017) 846–854.
- [57] I.T.G. Silva, et al., Featuring the nucleosome surface as a therapeutic target, *Trends Pharmacol Sci* 36 (2015) 263–269.
- [58] M. Ivanov, et al., Epigenetic mechanisms of importance for drug treatment, *Trends Pharmacol Sci* 35 (2014) 384–396.
- [59] W.F. Cabral, et al., Exogenous nucleosome-binding molecules: a potential new class of therapeutic drugs, *Drug Discov Today* 21 (2016) 707–711.
- [60] J. Dekker, et al., The 4D nucleome project, *Nature* 549 (2017) 219–226.
- [61] S. Xiang, et al., The LC domain of hnRNP2A adopts similar conformations in hydrogel polymers, liquid-like droplets, and nuclei, *Cell* 163 (2015) 829–839.
- [62] T. Schalch, et al., X-ray structure of a tetranucleosome and its implications for the chromatin fibre, *Nature* 436 (2005) 138–141.



Nucleosome binding peptide presents laudable biophysical and *in vivo* effects



Kaian Teles^a, Vinicius Fernandes^{a,b}, Isabel Silva^a, Manuela Leite^a, Cesar Grisolia^c, Vincenzo R. Lobbia^d, Hugo van Ingen^d, Rodrigo Honorato^e, Paulo Lopes-de-Oliveira^e, Werner Treptow^b, Guilherme Santos^{a,*}

^a Laboratório de Farmacologia Molecular, Departamento de Farmácia, Universidade de Brasília, Brasília, 70919-970, Brazil

^b Laboratório de Biologia Teórica e Computacional, Departamento de Biologia Celular, Universidade de Brasília, DF, 70910-900, Brasília, Brazil

^c Laboratório de Genética Toxicológica, Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Brazil

^d NMR Spectroscopy Group, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH, Utrecht, the Netherlands

^e Laboratório Nacional de Biociências (LNBio), Campinas, SP, Brazil

ARTICLE INFO

Keywords:

Peptides
Nucleosome
Chromatin

ABSTRACT

Chromatin state is highly dependent on the nucleosome binding proteins. Herein, we used a multipronged approach employing biophysical and *in vivo* experiments to characterize the effects of Nucleosome Binding Peptides (NBPEs) on nucleosome and cell activity. We performed a series of structure-based calculations on the nucleosome surface interaction with GMIP1 (a novel NBPE generated *in silico*), and HMGN2 (nucleosome binding motif of HMGN2), which contains sites that bind DNA and the acid patch, and also LANA and H4pep (nucleosome binding motif of H4 histone tail) that only bind to the acidic patch. Biochemical assays shows that H4pep, but not HMGN2, GMIP1 and LANA, is highly specific for targeting the nucleosome, with important effects on the final nucleosome structure and robust *in vivo* effects. These findings suggest that NBPEs might have important therapeutic implications and relevance as tools for chromatin investigation.

1. Introduction

Chromatin is a macromolecular complex composed of distinct molecules. Highly basic proteins (histone octamer) interact with DNA to form the nucleosome core particle (NCP), generating the fundamental repetitive unit of chromatin. The NCP represents the first level of DNA compaction, followed by a cooperative nucleosome interaction to form the higher-order chromatin structure (reviewed in [1]). Chromatin dynamics, which is controlled by a plethora of Nucleosome Binding Proteins (NBPs), is essential for genome integrity and gene expression regulation. From condensed to relaxed chromatin, NBPs may induce specific modifications of chromatin architecture dependent on their unique properties [2]. In addition to NBPs, the nuclear environment also comprises many small molecules with different chemical natures that can directly interact with nucleosomes, such as Mg^{2+} and lipids [3,4].

The nucleosome core particle surface contains the acidic patch, a highly negative region formed by six residues of H2A and two of H2B, that is responsible for nucleosome–nucleosome interactions and is a target for several NBPs [2]. The first structure of a nucleosome:peptide

complex showed at atomic level the binding mode of the viral peptide LANA to the acidic patch [5]. In the following years, other nucleosome:NBPs complex structures were solved, revealing the atomic details of the interaction of NBPs with the nucleosome surface, highlighting the acidic patch as the principal protein-docking region [6].

Unlike canonical drug targets such as enzymes or protein receptors, the nucleosome is a structural protein:DNA complex without typical druggable cavities, and this impedes research for new exogenous nucleosome binding molecules. Instead of focusing on small molecules for occupying the nucleosome surface, we try to understand and developed more complex molecules, such as the Nucleosome Binding Peptides (NBPEs).

Herein, we characterized the effects of NBPEs on nucleosome and chromatin structure. Firstly, we designed and generated *in silico* a novel NBPE, GMIP1, with nucleosome surface binding highly dependent on the DNA. Then, in order to understand how NBPEs with distinct nucleosome binding sites affect nucleosome structure, we performed a series of structure-based calculations on the nucleosome surface interaction to the NBPEs. We studied four NBPEs, GMIP1 and HMGN2 (the nucleosome binding motif from HMGN2), which contains sites that

* Corresponding author.

E-mail address: gsantos@unb.br (G. Santos).

<https://doi.org/10.1016/j.bioph.2019.109678>

Received 27 September 2019; Received in revised form 30 October 2019; Accepted 13 November 2019

0753-3322/ © 2019 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

bind DNA and the acid patch, and also LANA and H4pep (structure-based nucleosome binding motif derived from the N-terminal domain of histone H4) that only bind to the acidic patch. Interestingly, we observed that NBPePs induced specific atomic fluctuations of the nucleosome structure. *In vitro* studies corroborate the idea that NBPePs may affect the stability of the nucleosome structure, however only H4pep showed high nucleosome binding affinity and specific actions on the final nucleosome structure while HMGN2, LANA and GMIP1 pePs seems to be non-specific DNA interactors, based on oppositely charged residues. Cell-based assays showed that the four NBPePs penetrated the cell, localizing at the nuclear environment and, except GMIP1, interfered with tumor cell viability. However, beside the non specific properties on the nucleosome, apart from H4pep, the fish embryo toxicity (FET) test showed that the four NBPePs could cause tissue modifications, such as defects in pigmentation and induction of earlier hatching.

2. Methods

2.1. *In silico* NBPeP design

KV finder software with a PyMOL interface plugin (The PyMOL – Molecular Graphics System Version 1.3 Schrodinger, LLC) was used to define the cavities on the nucleosome:protein complex. YSARA software was used for the GMIP1 design and the optimized conformation by minimizing energy was performed using the force field YAMBER3.

2.2. NBPePs

All Peptides were bought from Biomatik with purity > 95 % and diluted in MilliQ H₂O. Fluorescent peptides were bought with TAMRA-(559/583 nm) in the N-terminus. The concentration was determined by spectrophotometric method as described in [7]. All peptides are described in the supplementary table 1.

2.3. Recombinant H2A-H2B for NMR

BL21(DE3)pLysS cells were used to express *Xenopus laevis* H2A or H2B in deuterated M9 medium containing 2 g/l (12C-2D) D-glucose (1,2,3,4,5,6,6-d7) and 0.5 g/L 15NHCl. One hour before induction with 1 mM IPTG, 60 mg/L α -ketobutyric acid (4-13 C,3,3-2d) and 80 mg/L α -ketoisovalerate (3-(methyl-d3),4-13 C,3-d) sodium salt were added to the medium. Histones were purified as described by Dyer et al. [8] To refold dimers, H2A and H2B were mixed in equimolar ratio in unfolding buffer (7 M urea, 150 mM NaCl, 50 mM NaPi, 1 mM EDTA, pH 7.5) and dialysed against high salt buffer (2 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5). Dimers were purified using size exclusion chromatography (Hiload superdex 200 16/600) and then dialysed against low salt buffer (20 mM NaPi, 0.01 % NaN₃, pH6.2).

2.4. *In vitro* chromatin fibers and nucleosome reconstitution

Histone octamers (HO) were purified from chicken erythrocyte nuclei as described in Huynh, V. A. T., P. J. J. Robinson, and D. Rhodes, 2005. 601 DNA Widom with 167 base pairs (bp) was used to reconstitute mononucleosomes and array 177.36 was used to reconstitute 10 nm chromatin fibers, both using the slow salt dialysis method as described in Huynh et al. (2005) [9]. The analyses of the reconstitution were verified by electrophoresis in native bis-acrylamide gels (6 %) or agarose gels (0.8 %).

2.5. Mononucleosome precipitation

Freshly reconstituted mononucleosomes (115 nM mononucleosome, Tris 10 mM pH 7.4, EDTA 1.5 mM NaCl 15 mM) were incubated with specified concentration of NBPePs for 30 min at room temperature. The

samples were centrifuged (Sigma centrifuge-2K15) at 15,493 x g for 20 min at 25 °C. The supernatant was transferred to another microcentrifuge tube and the pellet was resuspended in the same buffer as the mononucleosome. The samples were analyzed by electrophoresis in native 6 % bis-acrylamide gel carried out with 0.5 x TBE buffer at 15 mA. Densitometry was performed using ImageJ (National Institute of Health, Bethesda, MD, USA) version 1.49.

2.6. DNA binding assay

Widom 601 DNA fragments containing 167bp (30 nM DNA, 10 mM Tris pH 7.4, 135 mM NaCl) were incubated with specified concentrations of GMIP1 for 2 h at 37 °C and 100 RPM. The analysis was done in 0.8 % agarose gel in TBE 0.5 x . Samples were loaded with 30 % glycerol, to avoid interaction caused by phenol blue and GMIP1.

2.7. Nucleosome binding assay

Freshly reconstituted mononucleosomes (115 nM mononucleosome, Tris 10 mM pH 7.4, EDTA 1.5 mM NaCl 15 mM) were incubated with specified concentration of fluorescent NBPePs for 120 min at room temperature. Then samples were analyzed by electrophoresis in native 6 % bis-acrylamide gel carried out with 0.5 x TBE buffer at 15 mA. Gels were analyzed using Amersham Imager 600 (GE) with the RGB laser kit detection for 520 nm, to visualize the peptide, following incubation in ethidium bromide bath and analyzed with UV for ethidium bromide detection. For Kd determination, band densitometry was performed in the gel revealed with 520 nm laser, using ImageJ (National Institute of Health, Bethesda, MD, USA) version 1.49, followed by analysis in Prism 6 Graphpad software using Binding - saturation binding to total and non-specific template.

2.8. Chromatin compaction assay

Chromatin compaction by Mg²⁺ was adapted from Rhodes Lab protocol [10]. Briefly, assembled chromatin fibers were incubated with vehicle (10 mM Tris-HCl (pH 7.5) or peptides (150 μ M GMIP1 or H4pep) for 2 h in room temperature. Next, 3 mM MgCl₂ were added, incubated for 15 min on ice and centrifuged at 13,000 rpm for 15 min at 4 °C. The supernatants and pellets were verified by electrophoresis in native 0.8 % agarose gels, carried out with 0.2x TBE (18 mM Tris-borate, pH 8; 0.4 mM EDTA) electrophoresis buffer at 20 mA. Densitometry was performed using ImageJ (National Institute of Health, Bethesda, MD, USA) version 1.49.

2.9. Thermal shift assay

Thermalshift assay with NBPePs was adapted from Taguchi et al. [11]. Briefly, freshly reconstituted mononucleosomes (86 nM), in 10 mM Tris pH 7.4, 1.5 mM EDTA and 15 mM NaCl, were incubated with specified concentration of NBPePs for 30 min at room temperature. Next 1 mM dithiothreitol and 5X of SYPRO-Orange (SIGMA-ALDRICH) were added and incubated for 1 min at room temperature. Fluorescence was measured with a StepOnePlus Real-Time PCR unit (Applied Biosystems) with increases of 1 °C step from 25 °C to 95 °C. The fluorescence was detected at 570 nm. The Raw data was normalized using Graphpad Prism 6 software and for the determination of the temperature of melting (T_m) was obtained from the first derivative curve of the data.

2.10. MTT

For MTT assays, 8000 HeLa cells or ccd10595k cells were plated in 96-well culture plates and maintained at 37 °C and 5 % CO₂ in DMEM medium with 10 % fetal bovine serum, penicillin (100U/mL) and streptomycin (100ug/mL) for 24 h. Next, wells were washed 3 times

with PBS 1X and filled with 100 μ L of DMEM medium as described above containing the specified amount of NBPePs and incubated for 24 h in the same conditions. The 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) at 5 mg/mL was added to the wells (10 μ L) and incubated for 4 h at 37°C and 5 % CO₂. The wells were drained and the formazan crystals were solubilized in 100 μ L of acidic isopropanol solution (52 μ L of HCl 37 % to 12 mL of isopropanol) and agitated for 30 min at room temperature. Absorbance at 570 nm was determined with a plate spectrophotometer (DTX 800 Multimode Detector - Beckman Coulter) at 570 nm.

2.11. Flow cytometry

70.000 HeLa cells were plated in 12-wells culture plates for 16 h and maintained at 37 °C and 5 % CO₂ in DMEM medium with 10 % fetal bovine serum, penicillin (100U/mL) and streptomycin (100ug/mL). Prior to treatment with NBPePs, cells were incubated for 1 h at 37 or 4 °C. Next, cells were washed with 1X PBS and filled with DMEM medium with the specified amount of fluorescent NBPePs and incubated at 37 or 4 °C for the 1 or 3 h. Wells were washed 3 times with ice cold 1X PBS and filled with 500 μ L 1X PBS, cells were harvested with a cell scraper and analyzed by flow cytometry on FACSCalibur (BD biosciences). HeLa cells were gated to isolate the main population of living cells from cell debris. Data analysis was done using flowjo 8.7 software.

3. Fluorescence confocal microscopy analysis

i) **Peptide cell penetration:** confluent HeLa cells were treated with increased concentrations (0,1; 0,5; 1 μ M) of peptide GMIP1-TAMRA or vehicle (10 mM Tris–HCl) for 24 h. Next, plates were rinsed three times in PBS and fixed in formalin 3.7 % at room temperature for 10 min. Samples were rinsed three times in PBS and incubated with DAPI (300 ng/ml) for 5 min to stain nuclei. Finally, cells were rinsed three more times in PBS. Images were acquired using a laser scanning confocal microscope Leica TCS SP5. To visualize peptide localization, Z-planes of 0.20 mm thickness were acquired. The images were analyzed with LAS AF software (Leica Microsystems CMS GmbH).

ii) **Determination of peptide internalization:** two plates of 24 wells with confluent HeLa cells were rinsed with DMEM without fetal bovine serum twice and maintained in PBS for 15 min in a 5 % CO₂ humidified atmosphere at 37 °C. One plate was then maintained for 10 min at 4 °C. Next, the two plates were treated with GMIP1-TAMRA (1 μ M) of vehicle for 5, 15 and 30 min. After each treatment samples were fixed in formalin 3.7 % at room temperature for 1 h. Samples were rinsed three times in PBS and incubated with DAPI (300 ng/ml) for 5 min to stain nuclei. Finally, cells were rinsed three more times in PBS. Images were acquired using a laser scanning confocal microscope Leica TCS SP5. To visualize peptide localization, Z-planes of 0.20 mm thickness were acquired. The software LAS AF (Leica Microsystems CMS GmbH) was used for analyse images.

3.1. Zebrafish husbandry and embryo collection

Zebrafish (*Danio rerio*) were raised in an aquatic facility (ZebTec - Tecniplast, Italy) with a photoperiod cycle of 12:12 h (light:dark) at the University of Brasilia (Brazil). The water parameters were: temperature was maintained at 27.0 \pm 1 °C, conductivity at 650 \pm 100 μ S/cm, pH at 7.0 \pm 0.5 and dissolved oxygen \geq 95 % saturation. Zebrafish embryos were collected immediately after natural mating, rinsed in water, and checked under a stereomicroscope (Stereoscopic Zoom Microscope – Stemi 2000, Zeiss, Germany). The unfertilized eggs and those showing cleavage irregularities or injuries were discarded [12].

3.2. Fish embryo toxicity (FET)

FET was adapted from Morash et al. [13]. Briefly, Zebrafish embryos at 4, 28 and 52 h post fertilization (hpf) were used to evaluate the toxicity of NBPePs in 96-well plates. Each peptide was tested at 0.1, 1, 10 and 100 μ M in 100 μ L of water from aquarium system; pH in all conditions was tested using pH strips (92,120 – MACHEREY-NAGEL). Embryos were stored at 27 °C with 14h light 10h dark cycle and evaluated Stemi 508 (Carl Zeiss) microscope with 1 and 24 h of treatment. Embryos were assessed for pigmentation, development, hatching and lethality. 10 embryos were used for each condition, if the control group showed any alteration, the plate was discarded, alterations > 10 % were considered significant and were documented using Axiocam Erc 5 s (Carl Zeiss) and ZEN software (Carl Zeiss).

3.3. Fluorescence fish embryo

Zebrafish larvae with 80 hpf were incubated with fluorescent NBPePs with specified concentration for 3 h in 100 μ L in a 96-plate, larvae were washed 3 times in 100 mL to remove the excess of NBPePs, imaging was done using Axioskop 2 (Carl Zeiss) with HBO 100 lamps, Axiocam Erc 5 s (Carl Zeiss) and ZEN software (Carl Zeiss) with appropriate laser filter for TAMRA (filter 4).

3.4. Fluorescence blood smear

Adults Zebrafish at 2 years old were inject in the abdomen with 50 μ L, 1 mM of fluorescent NBPePs, and kept protected from light at 27 °C for 18 h. Blood was extracted from the fins using a pipet tip and heparin 250 IU to make the blood smear in a microscope slide. Images were acquired with Axioskop 2 (Carl Zeiss) with HBO 100 lamps, Axiocam Erc 5 s (Carl Zeiss) and ZEN software (Carl Zeiss) with appropriate laser filter fom TAMRA (filter 4).

3.5. NMR

All NMR experiments were carried out on a Bruker advance III HD 600 MHz. NMR spectra were processed in Bruker TopSpin [14] and analyzed using Sparky [15]. Dimer samples of [¹³C,¹⁵N]H2A-H2B at 100 μ M in 5 %D₂O/95 %H₂O; 25 mM NaPi + 100 mM NaCl pH6.2 + 0,01 % Na₃N + 1 mM 2-Mercaptoethano + PIC (complete EDTA-free Protease Inhibitor Cocktail (Roche)) were titrated against GMIP1 using 600 MHz Lamour frequency at 308 K. HSQC spectra were measured for free [¹³C,¹⁵N]H2A-H2B and after the addition of GMIP1 at 308 K. Titration consisting of 4 points in the range of 1:4.3 M ratio ([¹³C,¹⁵N]H2A-H2B:GMIP1) was performed.

3.6. Circular dichroism

Measurement of secondary structure of NBPePs was performed in Jasco j-815 spectropolarimeter in a 0,1 cm quartz cuvette in the range of 190–250 nm. Samples were diluted in MiliQ water in the concentration of 0.125 mg/mL for GMIP1, LANA, HMG2 and H4pep at 0.107 mg/mL at 25 °C. Data were plotted using BestSel data base (available at: <http://bestsel.elte.hu/>).

3.7. Computational studies NBPePs

Nucleosome Atomistic models were built using a high-resolution x-ray structure of the nucleosome (NCP) without histone tails, PDB code 3TU4 [16]. Peptides were simulated free in solution or in association with nucleosome at 1:1 and 2:1 stoichiometries. Table M&M provides information on the construction and initial atomic coordinates for the various systems under investigation. Analyses were performed with VMD version 1.9.3 [17].

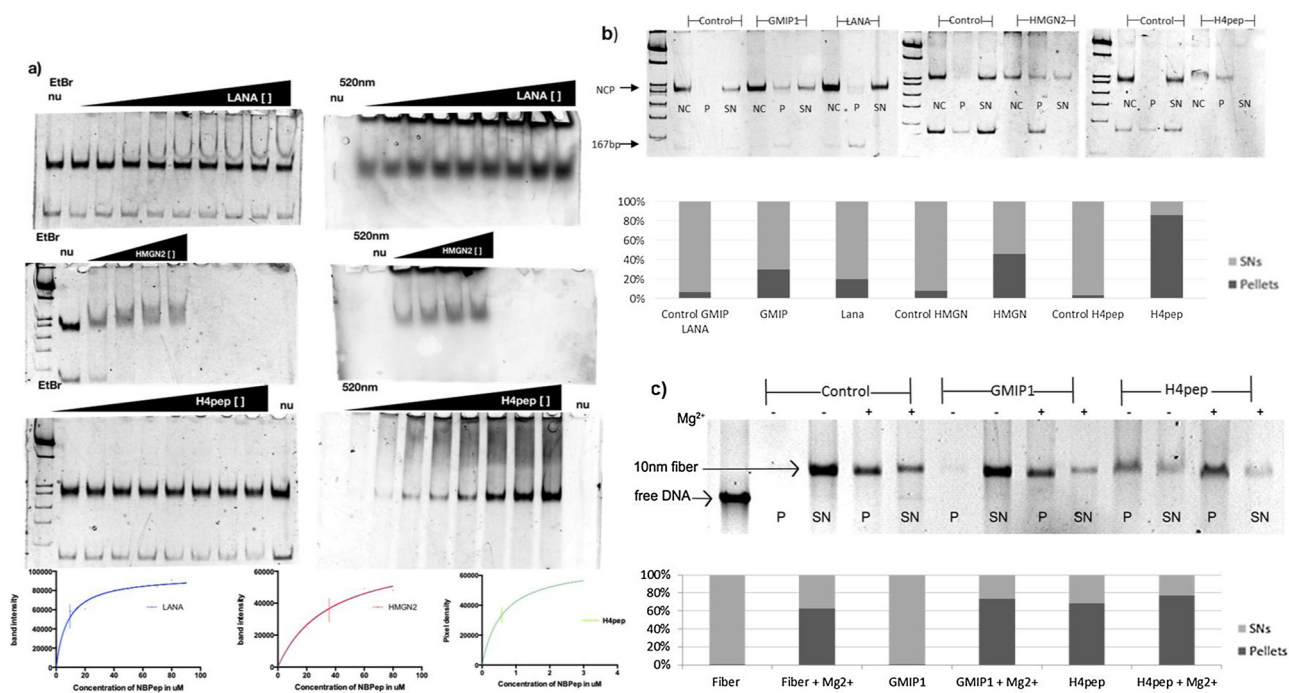


Fig. 1. NBPePs interaction assay a) Nucleosome binding assay with fluorescent NBPePs, nucleosome is incubated with LANA at 0, 10, 20, 30, 40, 50, 60, 70, 80, 90 μ M, with HMGN2pep at 0, 20, 40, 60, 80 μ M or with H4pep at 0, 200, 400, 600, 800, 1000, 1200, 1400, 1600 nM. It was then analyzed in acrylamide gel, following by detection of the fluorescent NBPeP and subsequently detection of DNA. Kd is represented by a vertical line in the densitometry graphs. b) Nucleosome precipitation assay with NBPePs, centrifuged nucleosome without NBPePs stays in the supernatant (SN); the addition of 50 μ M GMIP1, 10 μ M LANA, 10 μ M HMGN2 or 500 nM H4pep induce precipitation and pellet (P) formation, non-centrifuged (NC) samples were used as control. DNA band densitometry graph of upper gel. c) Chromatin compaction assay with Mg²⁺, 150 μ M H4pep induces 36-mer chromatin precipitation, but not 150 μ M GMIP1. DNA band densitometry graph of upper gel. DNA band densitometry graph of upper gel. These assays were performed at least 3 times, and the representative gel was presented.

3.8. Molecular docking

AutoDock Vina was used to resolve peptide binding to the acidic patch region of NCP [18]. To account for a larger ensemble of binding modes 20 independent structures randomly collected from the peptide-free MD equilibration was docked to the NCP. The exhaustiveness value was set to 200 and best solutions were gathered from each docking calculation, resulting in approximately 400 solutions per peptide. Solutions were clustered in 15–17 structural groups based on a maximum neighborhood criterion and the group with best fit to the acidic patch was chosen for further simulation.

3.9. Molecular dynamics simulations

The VMD software was used to solvate and neutralize with counter ions the NCP-peptides systems [17], resulting in simulations cells averaging $\sim 110 \text{ \AA} \times 145 \text{ \AA} \times 145 \text{ \AA}$ with ~ 210.000 atoms and 150 mM sodium chloride. All systems were simulated in a NPT ensemble at 300 K, 1 atm and with 2 fs time step for 115 ns with periodic boundary conditions (PBC) and enough water to avoid any interactions between the PBC images. Each system was thermalized and subsequently equilibrated for ~ 10 ns MD simulation. All simulations were run by NAMD version 2.10 [19] with CHARMM 36 force field [20] and TIP3 water model [21]. PME method [22] was employed on the electrostatic calculations and non-bonded interactions were cut-off at 11 \AA . No bias was needed to keep the peptides bound to NCP throughout the entire simulation.

4. Results

4.1. In silico development and analysis of structure-based nucleosome binding peptide

To design and develop a NBPeP that best-fitted on the nucleosome surface, we studied available nucleosome binding proteins structures as templates [5,16,23–25]. We used the software KvFinder to identify putative binding pockets and shallow crevices on the nucleosome surface to use these cavities as potential binding site to plan NBPePs [26]. We selected three crystallographic atomic structures to start our search, those of RCC1, Sir3 and LANA. The best-fitted molecule on the acidic patch was RCC1 (Supplementary Fig. 1a), which contains binding sites for both the acidic patch and the associated nucleosomal DNA [23]. Besides the deep anchor on the acidic patch, through R216 and R223, there is another region distal to the acid patch in which the residue T238 from the RCC1 may interact to T75 from H2A. Therefore, based on the structure of RCC1:Nucleosome complex, we generated the new NBPeP, GMIP1 (genetic modified inducible peptide 1), which included the two nucleosome binding sites to either DNA or the acidic patch, linked by three alanine residues (Supplementary Fig. 1 b–d).

4.2. Differential binding of NBPePs to the nucleosome and effect on chromatin conformation

We selected NBPePs based on their abilities to bind to the (i) acidic patch (LANA and H4) or (ii) both DNA and acidic patch (GMIP1 and HMGN2).

We already knew that the NBPePs LANA, H4 tail and HMGN2 could bind to the nucleosome [5,24,25], but we did not have biochemical information about GMIP1. Although GMIP1 was based on the nucleosome binding motif of RCC1, it has a new structure, with two independent binding regions linked by three alanine residues as

discussed earlier. Circular dichroism analysis of the NBPePs showed predominant random coil structure (Supplementary Fig. 2 and Supp. table 1 sequence NBPePs). NMR experiments with free GMIP1 indicated the peptide is in a random-coil state based on absence of medium or long-range NOEs and random coil ^{13}C chemical shifts (Santos GM and van Ingen H., personal communication).

To observe the binding and the impact of NBPePs to the nucleosome we performed nucleosome binding and precipitation assays (Fig. 1 a). First, we reconstituted nucleosomes *in vitro* with histone octamer from chicken and 167.1 (167 base pairs) DNA 601 and incubated it with labelled NBPePs-TAMRA. It demonstrated that H4pep binds to the mononucleosome at a K_d of 0,6 μM , LANA at K_d of 8 μM and HMGN2 at 358 μM respectively. GMIP1 induced nucleosome aggregation even in low concentration, which difficulted to determine the binding affinity constant. HMGN2pep promotes an electrophoretic mobility shift, suggesting that it is binding at more than one site on the nucleosome surface, further investigations will be need to explore this finding.

To confirm the binding assays results, non-labelled GMIP1, LANA, HMGN2 and H4pep were then incubated with mononucleosomes and then centrifuged. We observed that the nucleosome in absence of NBPePs does not precipitate, but all four NBPePs variably induced nucleosome precipitation since part of the complex was found in the pellet (Fig. 1 b). Notable, H4pep induced precipitation at 500 nM. At higher concentration, the NBPePs induce mononucleosome aggregation and did not migrate in the gel (data not shown).

To understand and validate the binding epitope of GMIP1 on the histone H2A:H2B surface, we performed an interaction study with *in vitro* reconstituted and isotope-labeled *xenopus laevis* H2A/H2B dimers by NMR. Binding of unlabeled GMIP1 should result in specific and clear changes in either peak position and/or intensities. Surprisingly, no significant spectral changes were observed even at high excess of GMIP1 (Supplementary Fig. 1 e), suggesting that this new NBPeP needs DNA binding to be stable on the nucleosome. Indeed, GMIP1 bound to the naked DNA *in vitro* (Supplementary Fig. 1 f) but at a K_d of 50 μM or weaker, which is probably in the range of non-specific binding of a charged peptide to DNA. Furthermore, GMIP1 showed to induce nucleosome aggregation which difficulted the K_d determination. Nucleosome (Supplementary Fig. 1 g). We then applied docking and Molecular Dynamics simulations to study binding of the GMIP1 at the atomic level and the conformational stability of the nucleosome. It showed that H2A/H2B dimer is not sufficient to hold GMIP1 on the nucleosome. Interestingly, the nucleosomal DNA stabilizes GMIP1 on both DNA and H2A/H2B dimer (Supplementary video and Supplementary Fig. 3).

It is already known that the non-acetylated H4 tail is able to induce chromatin condensation. [27]. In order to check whether GMIP1, which low affinity nucleosome binding is similar to LANA and HMGN2pep, could also affect the chromatin state such as described for non-acetylated H4 tail, we performed a chromatin compaction assay with Mg^{2+} (Fig. 1 c). For this, we reconstituted long chromatin fibers, 177.36 (36 nucleosomes), *in vitro* and incubated them with H4pep and GMIP1 and then with Mg^{2+} , following to centrifugation. Since the precipitated chromatin fiber recovered from the pellet was not cross-linked, it did not hold the compacted state and migrated in the gel similarly to the relaxed chromatin fiber, as observed in the gel. H4pep induced precipitation even in absence of Mg^{2+} , and precipitation increased when the divalent cation was added. However, GMIP1 did not show any effect alone, even when it was in presence of Mg^{2+} .

Taguchi et al. have established a technique to evaluate the physical properties of nucleosomes, developing a convenient assay of the thermal stability of nucleosomes *in vitro*. It was observed that increasing temperature induces nucleosome denaturation in two steps, firstly at 75 °C the eviction of H2A:H2B dimers occurs followed by the eviction of H3:H4 tetramers at and at 85 °C [11]. Therefore, herein thermal shift assays were performed to further check the hypothesis that NBPePs affect nucleosome stability (Fig. 2). GMIP1 showed a dose dependent

effect of the nucleosome stability, affecting first the H2A:H2B dimers, with a pronounced effect at higher concentrations, 150 μM . LANA has a notable effect on the H3:H4 tetramers, starting at 30 μM . HMGN2 induced earlier disruption of the H3:H4 tetramers at 3 μM . Clearly, H4pep induced H2A:H2B dimer stabilization at 100 nM concentration and did not affect H3:H4 tetramer. These findings suggested that NBPePs affect the nucleosome stability in different manners, however only H4pep presented a high specificity effect.

4.3. The interaction mode of NBPePs and nucleosome

In order to understand how the nucleosome structure senses the binding of NBPePs, MD simulations were performed to study the NBPePs based on their abilities to bind to the acidic patch or both DNA and acidic patch. More specifically, we studied peptide motifs from LANA and H4 tail that just bind to the H2A and H2B of the acidic patch and, for the second group, the peptide motif from HMGN2 and the new GMIP1 that bind to H2A and H2B and DNA. A scrambled peptide derived from GMIP1 was used as a negative control. All NBPePs tested here, but not the negative control, were able to bind and maintain the interaction with nucleosome over 115 ns of equilibrium trajectory (Fig. 3a).

It was observed that NBPePs induced specific atomic fluctuations of the nucleosome structure. Based on the differences of RSMF (root square mean fluctuation) of the nucleosomes, in presence and absence of the NBPePs, a map of the fluctuations was generated. There were several common fluctuation changes caused by the different NBPePs, however it is important to highlight the fluctuations of two distinct regions of the nucleosome (circles and arrows Fig. 3b). First, NBPePs that interact with just the acidic patch caused higher atomic movement at the nucleosome entryside-DNA region (arrows Fig. 3b). LANA significantly increase the DNA atomic fluctuations and H4pep slightly increase fluctuations of the H3 N-terminal helix. As expected, GMIP1, but not HMGN2, which also bind to the DNA, moderately reduced the DNA fluctuation. The second region affect by NBPePs binding was one of the known interactions points between the DNA and histones [25] (circles Fig. 3b). This interface, composed by H4 amino acids A76, K77, R78 and 8 nucleotides, showed narrow movement reduction caused by all NBPePs. Also, although rather small, HMGN2 was the peptide with greater impact (increasing) on the histones octamer atomic fluctuations.

Changes in hydration pattern of the nucleosome showed that the NBPePs reduced at least water content by 10 % on the acidic patch. GMIP1 removed fewer water molecules compared to the other NBPeP, even when compared to HMGN2 that also has two binding sites, DNA and acidic patch (Fig. 3c).

4.4. NBPePs cellular uptake and viability

Cell penetrating peptides, CPPs, are typically short peptides composed by high content of positively charged amino acids, such as lysine and arginine. Interestingly, NBPePs also contain several positively charged aminoacids [28]. In order to verify whether the four NBPePs could penetrate the cell environment, we performed flow cytometry assays in Hela cells. The four NBPePs penetrated the cell at 37 °C and GMIP1 had the highest uptake. H4pep in higher concentration (10 μM) was also taken up at 4 °C at a similarly rate to 37 °C (Fig. 4a).

To better understand and characterize whether GMIP1 that presented low affinity nucleosome binding could be uptaken and bind to the chromatin, we treated Hela cells with labeled GMIP1 and analyzed them with confocal microscopy. Images with orthogonal views demonstrated that peptide GMIP1 penetrates the cells even in the smallest concentration (0.1 μM). Confocal microscopy merged images of labeled GMIP1 and DNA stained with DAPI indicate the presence of GMIP1 associated with chromatin (Fig. 4b). Internalization of GMIP1 in Hela cells occurred only at 37 °C, suggesting an endocytosis transport

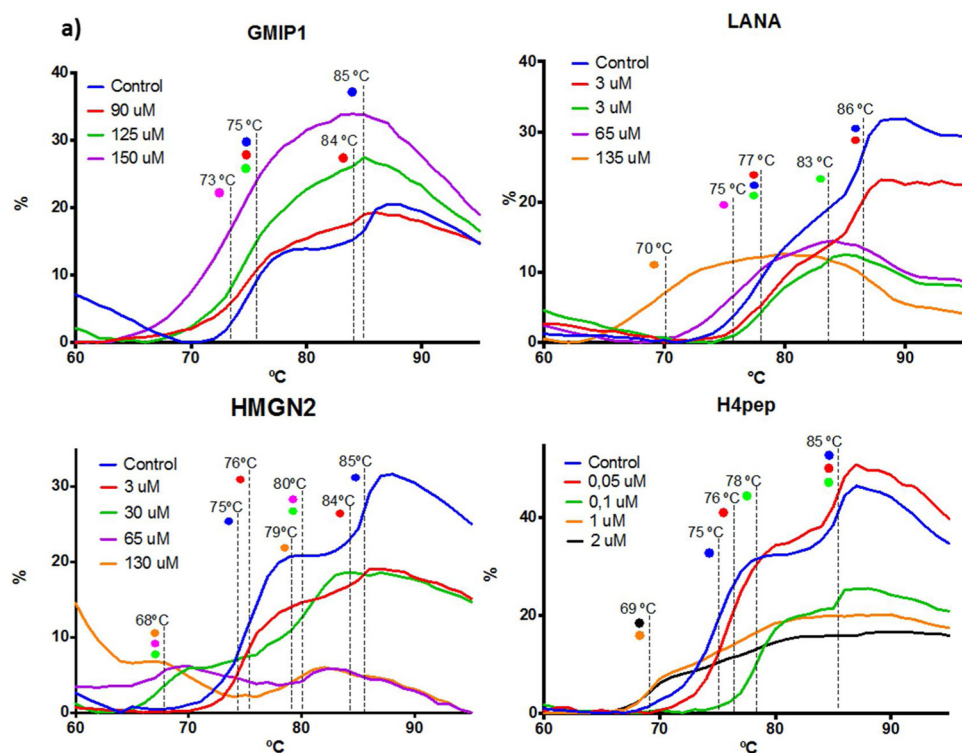


Fig. 2. Thermal Shift assay a) Thermal denaturation profiles of mononucleosomes in the presence of NBPePs. Mononucleosome denaturation happens in two steps, the first peak is the eviction of the dimers H2A/H2B and the second step if the eviction of the tetramers H3/H4 the formation of a peak at 68–70 °C appears when the mononucleosome is already disassembled. All thermal denaturation profiles are representative of three separate experiments.

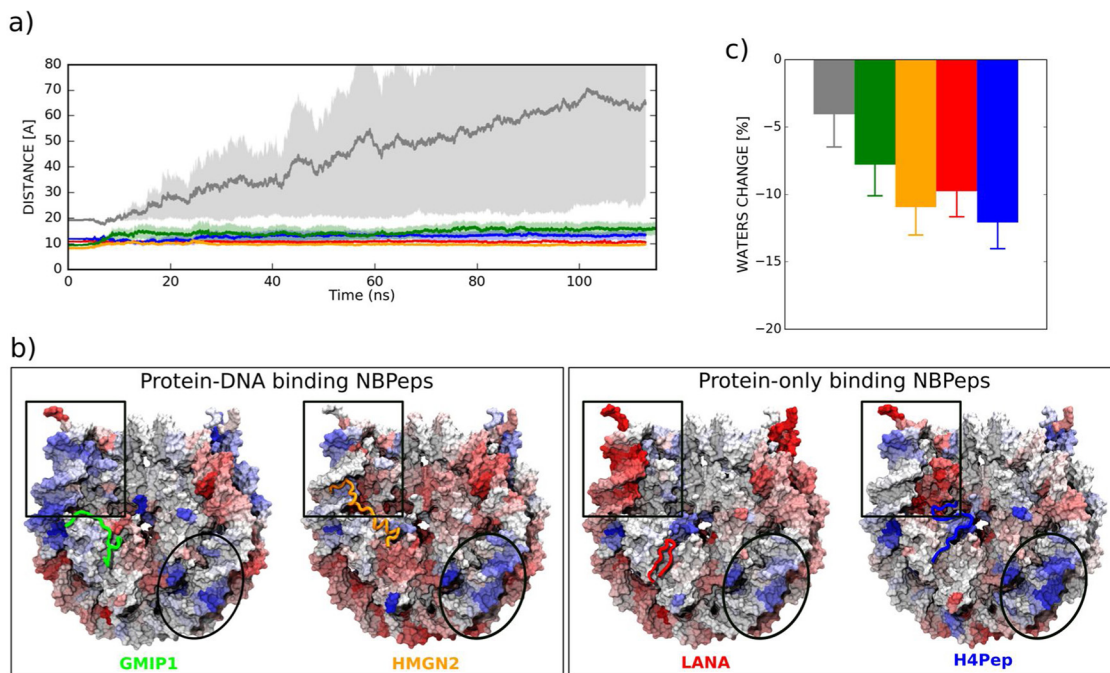


Fig. 3. NBPePs interactions with the nucleosome a) Time-dependent centroid distances profiles between the acidic patch region and NBPePs (Negative control – gray; GMIP1- green; HMGN2 – orange; LANA – red; H4pep - blue). Shaded regions indicate standard error between all simulations. b) NBPePs-induced proportional changes on the nucleosome (NCP) root-mean-square fluctuations (RMSF). The ratio was calculated after averaging all peptides association with the nucleosome stoichiometries. Color scale values distinguish between higher (red) and lower (blue) fluctuations compared to the RMSF of NBPeP free nucleosome, with a 25 % cut-off. Squares and circles indicate strongest fluctuation changes regions c) Average of proportional change in the number of water molecules occupying the acidic patch for each NBPeP. The change was calculated using the nucleosome free simulation as base value and within a cut-off distance of 20 Å. Averages and associated standard deviation were computed from 100 ns simulation. Colors code as a).

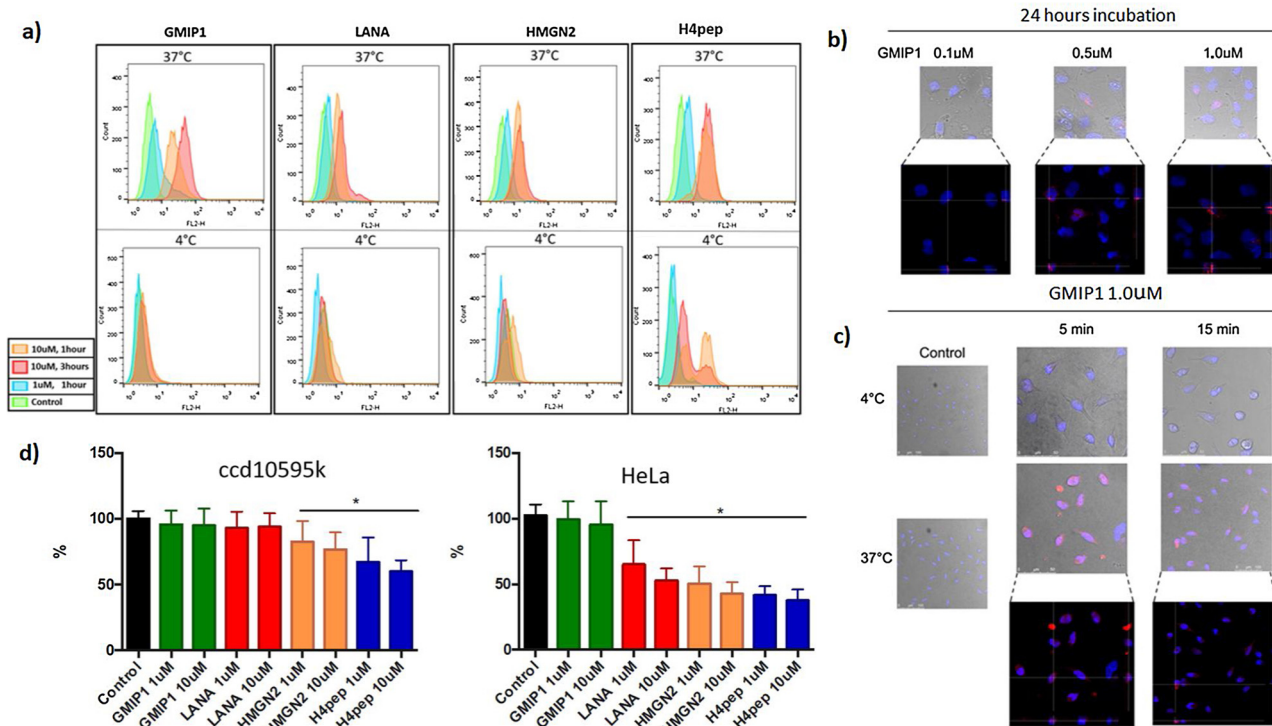


Fig. 4. NBPePs cell penetration a) Flow cytometry profile of HeLa cells uptake of fluorescent NBPePs (TAMRA), in histogram view with 1 or 3 h exposure and at 37 or 4 °C. Histogram is representative of three separate experiments.

b) GMIP1 bound to chromatin: confluent HeLa cells were treated with increased concentrations (0.1; 0.5; 1µM) of peptide GMIP1-TAMRA or vehicle (10 mM Tris–HCl) for 24 h. The images were acquired using a laser scanning confocal microscope Leica TCS SP5. c) GMIP1 internalization: Orthogonal analysis of GMIP1 overlapping with stained (DAPI) DNA. Cell penetration only occurs at 37 °C. The images were acquired using a laser scanning confocal microscope Leica TCS SP5. To visualize peptide localization, Z-planes of 0.20 µm thickness were acquired. The software LAS AF (Leica Microsystems CMS GmbH) was used for analyze images. All images are representative of three separate experiments. d) Relative cell viability compared to control over 24 h exposure to NBPePs in HeLa and CCD 10,595 K. HeLa cells showed a greater decreased in cell viability than CCD 10595K for every NBPeP tested with the exception of GMIP1. Data is shown as mean ± SD. * represent significant statistical difference (one-way ANOVA test) between the control and treated groups with $p < 0.05$ and $n = 2$.

(Fig. 4c).

4.5. Cell viability assay with NBPePs

In order to verify the potentiality of NBPePs for modulating cell physiology, we performed MTT assay using with different cell linages. In general, NBPePs, but not GMIP1, significantly affected the viability of HeLa cells, which support the evidences obtained *in vitro* that GMIP1 is acting non-specifically. Furthermore, NBPePs impacted the primary cell lineage ccd10595k much less, except H4pep that still showed high effect on the cell viability (Fig. 4d).

5. Fish embryo toxicity (FET) with NBPePs

To analyze *in vivo* the action of the NBPePs, we performed we performed Fish Embryo Toxicity (FET, OECD protocol n.236, 2013) test. After the exposure to NBPePs, at 3 different stages of development, at 4, 28 and 52 h post-fertilization (hpf) for 1 and 24 h, all NBPePs presented low or no toxicity (Supplementary table 2).

We were able to identify the labeled NBPePs in the nuclei of adult zebrafish erythrocyte (Fig. 5a). Moreover, NBPePs penetrated different tissues of fish larvae, after 3 hs incubation (Fig. 5b).

The respective outcome of the embryo's exposure to the NBPePs was highly dependent on the stage of development (Fig. 5c). The hatching rate of the zebra fish was largely affected by the presence of the NBPePs, with GMIP1 showing the strongest effect (Supplementary table 2).

GMIP1 presented a discrete pigmentation defect, in 26 % of embryos when at 28hpf and did not induce mortality at any concentration

tested.

LANA, at the highest concentration (100 uM), caused delay in the development of 100 % of the embryos when they were exposed at 4hpf. Also, LANA induced an increased hatching rate (Fig. 5c and Supp. table 2).

HMGN2 clearly affected pigment formation of 16 % of the embryos at 100 uM when at 56hpf.

Remarkably, H4pep were the only NBPePs tested that induced 100 % mortality at 4hpf, and the exposure for 1 h induce 20 % of mortality.

6. Discussion

Proteins that bind the nucleosome, the repetitive unit of chromatin, and the histone H4 tail are critical for establishing chromatin architecture and phenotypic outcomes [29]. A myriad of histone modifications, which strongly induce changes in chromatin structure, have been associated with malignancies and other diseases, highlighting the crucial role of chromatin architectural changes in disease mechanisms [30]. Thus, it is plausible that any nucleosome binding molecule might be able to interfere with the chromatin dynamics and modulate the access of proteins to the DNA.

At first glance, small molecules would be the best option for engaging the nucleosome surface, mainly the acidic patch, to control chromatin dynamics. An effort to identify specific inhibitors of Kaposi's sarcoma-associated herpesvirus (KSHV) focused on small molecules that could displace the LANA peptide from the acidic patch [31]. Investigators tested more than 350.000 small molecules, but could not find any compound able to displace LANA from the nucleosome surface, suggesting that more complex molecules, such as peptides, might be

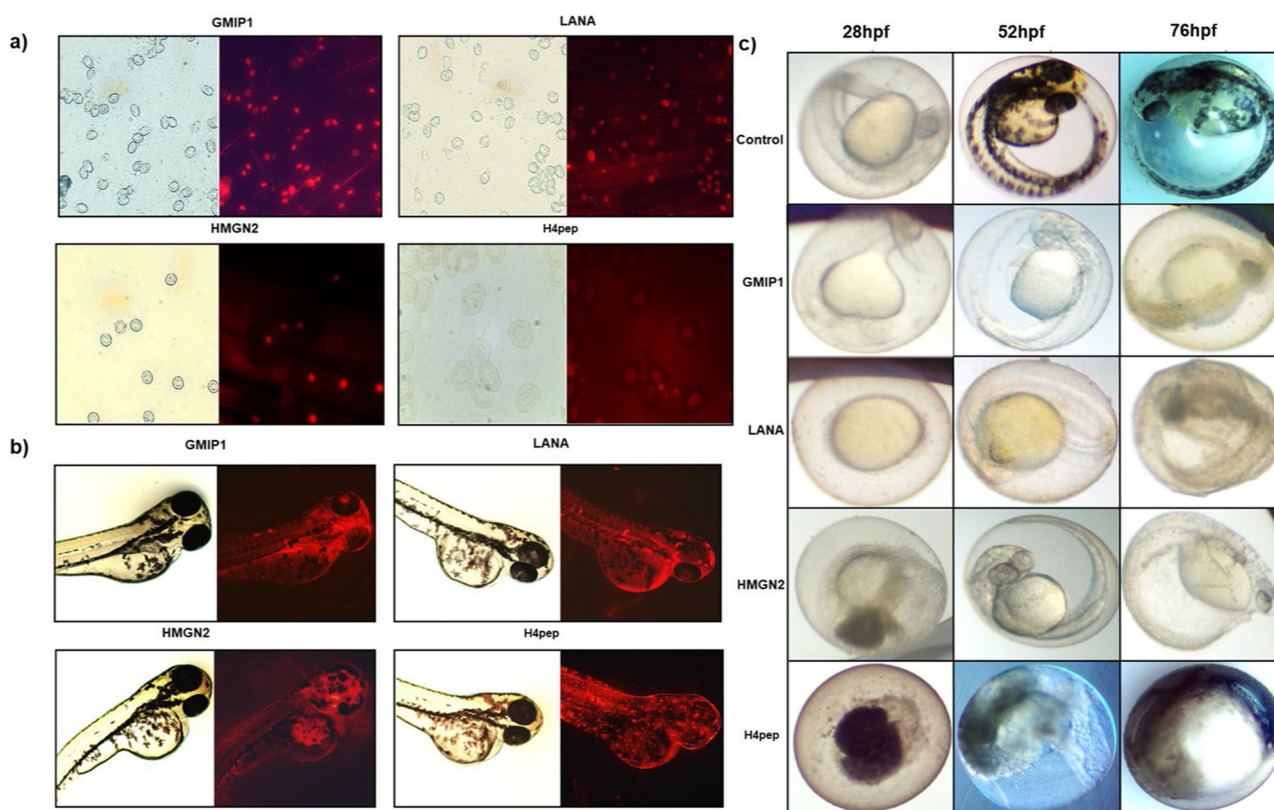


Fig. 5. NBPePs action *in vivo* a) Injection of fluorescent NBPePs accumulate in the nucleus erythrocytes of adult zebrafish, b) fluorescent NBPePs incubated for 3 h distributes heterogeneously over zebrafish larvae. At right panel, visualization using 520 nm laser. c) Fish Embryo Toxicity (FET) with NBPePs. Zebrafish embryos at 4, 28 and 52 h post fertilization (hpf) were incubated with NBPePs or vehicle for 24 h. Images are representative of three separate experiments.

required to bind in shallow crevices that mediate protein-protein interactions. Recently, binuclear ruthenium compounds was shown to be able to target the nucleosome surface and induce chromatin condensation [32].

Herein, we rationalized that NBPePs would be important candidates for occupying the nucleosome surface and to direct control the chromatin status and phenotypic outcomes. For that, NBPePs derived from binding motifs of structure-characterized nucleosome binding proteins, LANA, H4 tail, HMGN2 and the newly generated GMIP1, were tested to verify their ability in impacting the nucleosome stability, chromatin status, cell viability and fish embryo development. The four NBPePs bound to the nucleosome and induce precipitation at μM order, except H4pep that act at 500 nM.

We first searched for a novel NBPeP that could bind to and present specificity for the nucleosome, with high dependence on nucleosomal DNA for the nucleosome binding. This characteristic could potentially open an avenue for the design of NBPePs with high specificity for target genes. However, the biochemical data shows that GMIP1 has low nucleosome binding affinity. It is important to emphasize that all experiments performed were done with the Widom 601 DNA sequence, which is an artificial sequence with high specificity to the octamer.

In general, the NBPePs displaced at least 10 % of the water associated with the acidic patch. Interestingly, thermal shift assays suggest that GMIP1 and HMGN2pep, which also bind to the nucleosomal DNA, primarily disrupt H2A:H2B dimer. This is partially in agreement with the molecular dynamic experiments, in which the HMGN2pep increased histones atomic fluctuations and reduced DNA fluctuation changes at the nucleosome entry-side DNA region. LANA firstly disturb H3:H4 tetramer and H4pep induced H2A:H2B dimer stabilization but did not affect H3:H4. Together with the notion that DNA is needed for GMIP1 binding led us to consider the hypothesis that NBPeP with a DNA nucleosomal binding site might be by more effective for modifying

the common target on the nucleosome surface. Chromatin compaction was affected by the H4pep but not by GMIP1, supporting the idea that GMIP1 presents low affinity binding to the nucleosome. Cell-based assays demonstrated that all four NBPePs were able to penetrate the cell, with GMIP1 having the best cell uptake. The protein-derived NBPePs, but not the rationally designed GMIP1, significantly affected the viability of tumoral cells. However, NBPePs had weaker effects on the cellular viability of non-tumoral cell lineage, except for H4pep that still showed strong negative effects on cell viability.

Experiments performed in zebrafish showed that NBPePs have a differential interference in the embryo mortality, development, pigmentation and hatching. Labeled NBPePs penetrated different tissues of fish larvae and localized in the nuclei. These results are in agreement to the cell-based assays showing the NBPePs penetrate the cell. It is worthy of attention that GMIP1 did not induce fish embryos mortality, contrarily to H4pep. These data corroborate the results from cell-based assays MTT using non-tumoral ccd 10595k cells. Interestingly, all NBPePs induced earlier hatching, but GMIP1 and LANA induced development embryo delay.

Taken together the *in silico*, biochemical an *in vivo* data, however, it is too early to provide a straight correlation between NBPePs binding sites and the phenotypic outcome. Considering that three out of the four designed peptides are non-specific nucleosomal interactors, a question that is still to be explored is whether the NBPePs may engage different signaling events, besides the nucleosome binding, for example affecting function of HDAC.

Indeed, acetylated H4 N-terminal tail was explored as a molecular tool to establish and maintain the active state of p53 target genes *via* interaction with histone deacetylase 1 (HDAC1). It was suggested that it could be used as a novel strategy for anticancer therapy [33]. Herein, we further advance the idea that H4pep could affect tumoral cells, but through the direct binding to the nucleosome surface, since H4pep was

the unique high specific NBPEP, causing dramatic effect *in vivo*. Its effect emphasized the physiological role of H4 tail on chromatin condensation and transcriptional outcome.

In conclusion, we observed that NBPEPs with distinct nucleosome binding sites perturb the nucleosome structure in multiple ways. Despite sharing an apparently similar target, NBPEPs showed different roles in cell physiology, which is probably due to the non-specificity in targeting the nucleosome surface. However, new biophysical experiments in cell-based context should be performed to be able to straight correlate the *in silico* predictions with the *in vivo* findings.

Nevertheless, considering that DNA intercalators or damaging agents still have great importance in clinical oncology, the fact that NBPEPs do not present specific targeting would not preclude their use as therapeutic agents. Indeed, we believe that NBPEPs open novel opportunities to design hybrid molecules with higher specificity to regulate a plethora of cellular disorders.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors are grateful to Carlos Pantoja and Tony Warne for critical reading of the manuscript and anonymous reviewers for the important comments and suggestions on the manuscript. This work has been supported by CNPq for KT, VF and IT fellowships, Fap-DF grant number 04/2017 for GS, iNEXT grant number 6756, funded by the Horizon 2020 programme of the European Union, for HI.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.biopha.2019.109678>.

References

- [1] M.R. Hubner, M.A. Eckersley-Maslin, D.L. Spector, Chromatin organization and transcriptional regulation, *Curr. Opin. Genet. Dev.* 23 (2) (2013) 89–95.
- [2] A.A. Kalashnikova, M.E. Porter-Goff, U.M. Muthurajan, K. Luger, J.C. Hansen, The role of the nucleosome acidic patch in modulating higher order chromatin structure, *J. R. Soc. Interface* 10 (82) (2013) 20121022.
- [3] I.T.G. Silva, V. Fernandes, C. Souza, W. Treptow, G.M. Santos, Biophysical studies of cholesterol effects on chromatin, *J. Lipid Res.* (2017).
- [4] J.T. Finch, A. Klug, Solenoidal model for superstructure in chromatin, *Proc. Natl. Acad. Sci. U.S.A.* 73 (6) (1976) 1897–1901.
- [5] A.J. Barbera, J.V. Chodaparambil, B. Kelley-Clarke, V. Joukov, J.C. Walter, K. Luger, et al., The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA, *Science* 311 (5762) (2006) 856–861.
- [6] W.F. Cabral, A.H. Machado, G.M. Santos, Exogenous nucleosome-binding molecules: a potential new class of therapeutic drugs, *Drug Discov. Today* 21 (5) (2016) 707–711.
- [7] J.B. Murphy, M.W. Kies, Note on spectrophotometric determination of proteins in dilute solutions, *Biochim. Biophys. Acta* 45 (1960) 382–384.
- [8] P.N. Dyer, R.S. Edayathumangalam, C.L. White, Y. Bao, S. Chakravarthy, U.M. Muthurajan, et al., Reconstitution of Nucleosome Core Particles From Recombinant Histones and DNA. *Methods Enzymol.* 375, Academic Press, 2003, pp. 23–44.
- [9] V.A. Huynh, P.J. Robinson, D. Rhodes, A method for the *in vitro* reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone, *J. Mol. Biol.* 345 (5) (2005) 957–968.
- [10] B. Dorigo, T. Schalch, K. Bystricky, T.J. Richmond, Chromatin Fiber folding: requirement for the histone H4 N-terminal tail, *J. Mol. Biol.* 327 (1) (2003) 85–96.
- [11] H. Taguchi, N. Horikoshi, Y. Arimura, H. Kurumizaka, A method for evaluating nucleosome stability with a protein-binding fluorescent dye, *Methods.* 70 (2-3) (2014) 119–126.
- [12] OECD. Test No. 236: Fish Embryo Acute Toxicity (FET) Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, 2013.
- [13] M.G. Morash, S.E. Douglas, A. Robotham, C.M. Ridley, J.W. Gallant, K.H. Soanes, The zebrafish embryo as a tool for screening and characterizing pleurocidin host-defense peptides as anti-cancer agents, *Dis. Models Mech.* 4 (5) (2011) 622–633.
- [14] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6 (3) (1995) 277–293.
- [15] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: enhanced software for bio-molecular NMR spectroscopy, *Bioinformatics* 31 (8) (2015) 1325–1327.
- [16] K.J. Armache, J.D. Garlick, D. Canzio, G.J. Narlikar, R.E. Kingston, Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution, *Science* 334 (6058) (2011) 977–982.
- [17] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph. Model.* 14 (1) (1996) 27–28 33–8.
- [18] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2) (2010) 455–461.
- [19] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, et al., Scalable molecular dynamics with NAMD, *J. Comput. Chem.* 26 (16) (2005) 1781–1802.
- [20] J. Huang, A.D. MacKerell Jr, CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data, *J. Comput. Chem.* 34 (25) (2013) 2135–2145.
- [21] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, *J. Phys. Chem. A* 105 (43) (2001) 9954–9960.
- [22] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems, *J. Chem. Phys.* 98 (12) (1993) 10089–10092.
- [23] R.D. Makde, J.R. England, H.P. Yennawar, S. Tan, Structure of RCC1 chromatin factor bound to the nucleosome core particle, *Nature*. 467 (7315) (2010) 562–566.
- [24] H. Kato, H. van Ingen, B.R. Zhou, H. Feng, M. Bustin, L.E. Kay, et al., Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR, *Proc Natl Acad Sci U S A*. 108 (30) (2011) 12283–12288.
- [25] K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature* 389 (6648) (1997) 251–260.
- [26] S.H. Oliveira, F.A. Ferraz, R.V. Honorato, J. Xavier-Neto, T.J. Sobreira, P.S. de Oliveira, KVFfinder: steered identification of protein cavities as a PyMOL plugin, *BMC Bioinformatics* 15 (1) (2014) 197.
- [27] P.J. Robinson, W. An, A. Routh, F. Martino, L. Chapman, R.G. Roeder, et al., 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction, *J. Mol. Biol.* 381 (4) (2008) 816–825.
- [28] H. Derakhshankhah, S. Jafari, Cell penetrating peptides: a concise review with emphasis on biomedical applications, *Biomed. Pharmacother.* 108 (2018) 1090–1096.
- [29] C.L. Woodcock, Chromatin architecture, *Curr. Opin. Struct. Biol.* 16 (2) (2006) 213–220.
- [30] A.C. Mirabella, B.M. Foster, T. Bartke, Chromatin deregulation in disease, *Chromosoma* 125 (1) (2016) 75–93.
- [31] C. Beauchemin, N.J. Moerke, P. Faloony, K.M. Kaye, Assay development and high-throughput screening for inhibitors of Kaposi's sarcoma-associated herpesvirus N-Terminal latency-associated nuclear antigen binding to nucleosomes, *J. Biomol. Screen.* 19 (6) (2014) 947–958.
- [32] G.E. Davey, Z. Adhikrsan, Z. Ma, T. Riedel, D. Sharma, S. Padavattan, et al., Nucleosome acidic patch-targeting binuclear ruthenium compounds induce aberrant chromatin condensation, *Nat. Commun.* 8 (1) (2017) 1575.
- [33] K. Heo, J.S. Kim, K. Kim, H. Kim, J. Choi, K. Yang, et al., Cell-penetrating H4 tail peptides potentiate p53-mediated transactivation via inhibition of G9a and HDAC1, *Oncogene*. 32 (20) (2013) 2510–2520.