

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

VINICIUS ALVES FERNANDES

**AÇÃO DO COLESTEROL SOBRE A ARQUITETURA DA CROMATINA:
DOCKING MOLECULAR DO COLESTEROL NO NUCLEOSSOMO**

BRASÍLIA
2016

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

VINICIUS ALVES FERNANDES

**AÇÃO DO COLESTEROL SOBRE A ARQUITETURA DA CROMATINA:
DOCKING MOLECULAR DO COLESTEROL AO NUCLEOSSOMO**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ciências Farmacêuticas da Faculdade de Ciências da Saúde, Universidade de Brasília, como requisito parcial à obtenção do título de Mestre em Ciências Farmacêuticas.

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

BRASÍLIA

2016

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

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

AAç999a Alves Fernandes, Vinícius
Ação do colesterol sobre a arquitetura da cromatina: docking molecular do colesterol ao nucleossomo / Vinícius Alves Fernandes; orientador Guilherme Santos; co-orientador Werner Treptow. -- Brasília, 2016.
83 p.

Dissertação (Mestrado - Mestrado em Ciências Farmacêuticas) -- Universidade de Brasília, 2016.

1. Biologia Molecular. 2. Biologia Computacional. 3. Reguladores da cromatina. I. Santos, Guilherme, orient. II. Treptow, Werner, co-orient. III. Título.

Vinícius Alves Fernandes

Ação do colesterol sobre a arquitetura da cromatina: *docking* molecular do colesterol ao nucleossomo

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas, Faculdade de Ciências da Saúde, Universidade de Brasília, como requisito parcial à obtenção do título de Mestre em Ciências Farmacêuticas.

Aprovada em 01 de julho de 2016.

Banca Examinadora

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

Prof. Dr. João Alexandre Ribeiro Gonçalves Barbosa
Universidade de Brasília

Prof. Dr. Ângelo Henrique de Lira Machado -
Universidade de Brasília

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

AGRADECIMENTOS

Inicialmente, gostaria de agradecer a 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 prof. Guilherme Santos, pela oportunidade recebida e pelo horizonte científico, por ele, me revelado. A sua motivação e entusiasmo demonstrados em cada resultado são extraordinariamente revigoradores do espírito científico. Agradeço também ao meu orientador prof. Werner Treptow, pelo incentivo nas horas mais árduas da jornada acadêmica 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 pela atenção a mim dispensada e por guiarem meus passos nessa jornada.

Agradeço aos meus colegas do Laboratório de Biologia Teórica e Computacional (LBTC) Caio, Letícia, Juliana, Camila, Alessandra, Mônica e Leonardo pela prontidão em auxiliar em todos os momentos do desenvolvimento desse trabalho. Também agradeço pelos momentos de descontração e por tornarem o ambiente do laboratório aconchegante.

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

Agradeço à minha namorada, 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 também a CAPES pelo apoio financeiro.

RESUMO

FERNANDES, Vinícius Alves. **Ação do colesterol sobre a arquitetura da cromatina: *docking* molecular do colesterol ao nucleossomo.** Brasília, 2016. Dissertação (Mestrado em Ciências Farmacêuticas) – Faculdade de Ciências de Saúde, Universidade de Brasília, Brasília, 2016.

A cromatina é essencial para manutenção da fisiologia celular por desempenhar papel-chave na regulação da transcrição, reparo e replicação do DNA. Para tanto, esta deve ser altamente dinâmica, permitindo alterações entre estados compactados e relaxados. O primeiro nível de compactação da cromatina é o nucleossomo, formado pelo DNA e um octâmero de histonas. Moléculas que se ligam ao nucleossomo tem potencial para modular a cromatina. Diversas evidências apontam para o colesterol, lipídio importante para o funcionamento basal das células, como uma molécula ligante de cromatina. Recentemente, foi demonstrado que o colesterol pode modular a estrutura das fibras de cromatina reconstituídas *in vitro*. Juntos, estes dados sugerem que o colesterol possui um efeito modulador da arquitetura da cromatina por se ligar diretamente ao nucleossomo. Entretanto, os detalhes moleculares da ação do colesterol sobre a estrutura da cromatina ainda não foram esclarecidos. O objetivo deste estudo foi investigar, *in silico*, a ligação do colesterol ao nucleossomo e observar seu impacto ao nível atômico. Logo, metodologias capazes de tratar interações moleculares com resolução atômica são fundamentais, como a dinâmica molecular (DM) e *docking* molecular. Aqui identificamos vários sítios de ligação do colesterol distribuídos pelo nucleossomo. Após 200 ns de DM, seis moléculas de colesterol amostradas permaneceram ligadas ao nucleossomo. Para cada sítio investigado, a energia livre de ligação, estimada pelo método LIE (*Linear Interactions Energy*), indica o favorecimento do complexo entre o colesterol e o nucleossomo. Foi observado um impacto nas flutuações atômicas de vários aminoácidos formadores dos sítios de ligação. Dentre estes, alguns são conhecidamente importantes para modulação da estrutura da cromatina. Além disso, houve redução significativa do número de moléculas de água nas interfaces de interação colesterol-nucleossomo, especialmente em regiões hidrofóbicas essenciais para a manutenção estrutural do nucleossomo. Esses resultados sugerem que o colesterol pode afetar a arquitetura da cromatina através da interação direta com o nucleossomo nos múltiplos sítios de ligação.

Palavras-chave: cromatina, nucleossomo, colesterol, dinâmica molecular, *docking* molecular, energia livre de ligação.

ABSTRACT

FERNANDES, Vinícius Alves. **Ação do colesterol sobre a arquitetura da cromatina: *docking* molecular do colesterol ao nucleossomo.** Brasília, 2016. Dissertação (Mestrado em Ciências Farmacêuticas) – Faculdade de Ciências de Saúde, Universidade de Brasília, Brasília, 2016.

Chromatin is essential for maintenance of cell physiology playing a key role in transcription regulation, DNA repair and replication. Therefore, it must be highly dynamic, allowing changes between compacted and relaxed states. The first level of chromatin compaction is the nucleosome, formed by DNA and histone octamer. Molecules that bind to the nucleosome have the potential to modulate chromatin. Several evidences indicate that cholesterol, important lipid for cell function, may bind to chromatin. Recently, it was demonstrated that *in vitro* cholesterol may modulate the structure of reconstituted chromatin fibers. Together, these data suggest that cholesterol may modulate chromatin architecture by direct binding to the nucleosome. However, the molecular details of cholesterol's action on chromatin structure remain unclear. The aim of this study was an *in silico* investigation of the cholesterol binding to the nucleosome and its impact at the atomic level. Therefore, methodologies capable of handling molecular interactions with atomistic resolution are critical, such as molecular dynamics (MD) and molecular docking. Here, we identified several cholesterol binding sites distributed over the nucleosome. After 200 ns of MD, six sampled cholesterol molecules remained bound to the nucleosome. For every site, the estimated absolute free energies obtained by the LIE method, favour cholesterol binding thus explaining the stability of the molecular complex. Cholesterol impacts the atomic fluctuations of several amino acids forming the binding sites. Among these, some are well-known to modulate the chromatin structure. Furthermore, a significant reduction in the number of water molecules in cholesterol-nucleosome interfaces was observed, especially at key hydrophobic regions for nucleosome structural maintenance. These results suggest that cholesterol may affect the chromatin's architecture via direct interaction with the nucleosome through multiple binding sites.

Keywords: chromatin, nucleosome, cholesterol, molecular dynamics, docking, binding free energy.

LISTA DE ILUSTRAÇÕES

- Figura 1. Ilustração da compactação do DNA. Adaptado de (PIERCE, 2007)..... 15
- Figura 2. Representação do NCP. a) Coordenadas obtidas no PDB ID 1KX5. DNA (prata e branco); Histonas H3 (azul), H4 (verde), H2A (amarelo) e H2B (rosa). b) Representação do tetrâmero H3:H4 e do dímero H2A:H2B com as regiões N e C – terminais indicadas. c) Esquema mostrando os domínios estruturais de cada histona. Esse esquema foi adaptado de (MCGINTY; TAN, 2014)..... 17
- Figura 3. Regiões específicas presentes no nucleossomo. a-b) NCP destacando os 4-helix bundles (H3:H3 em azul escuro, H2B vermelho escuro e H4 verde escuro). c) Extensões do domínio globular central da H3, hélice N-terminal (α -N), e da H2A, domínio de docking. d) NCP representado como na Figura 2a e destacando o patch ácido. Painel reconstruído similar a (MCGINTY; TAN, 2014)..... 19
- Figura 4. Proteínas ligantes do nucleossomo (NBP). A superfície carregada do NCP está representada, destaca-se a região negativa (patch ácido) onde as NBPs resolvidas por cristalografia de raios X ou NMR estão interagindo. Adaptado de (CABRAL; MACHADO; SANTOS, 2016)..... 24
- Figura 5. Estrutura química e representação 3D do colesterol..... 26
- Figura 6. Resumo das etapas seguidas nesse trabalho. As setas sem preenchimento significam etapas que foram seguidas após as respectivas análises (setas pretas)..... 30
- Figura 7. 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..... 32
- Figura 8. Esquema que ilustra as condições periódicas de contorno (PBC). Fonte: <http://isaacs.sourceforge.net/phys/psc.html>..... 36
- Figura 9. Perfil do RMSD para o sistema do nucleossomo simulado em água por 20 ns. a) Prováveis seleções atômicas para a sobreposição estrutural entre os frames. Em todas as seleções apenas os átomos da cadeia principal foram considerados. b) Verificação da estabilidade do sistema em busca de um ensemble de equilíbrio (destacado por um quadrado) para as análises de docking molecular. Foram selecionados apenas os átomos pesados, ou seja, exclui-se os átomos de hidrogênio. 4-HB \rightarrow 4-helix bundle, DD \rightarrow domínio de docking..... 48
- Figura 10. Resultados do docking após tratamento das soluções encontradas. Está representado o DNA (prata e branco) e as histonas H3 (azul), H4 (verde), H2A (amarelo) e H2B (rosa), com as 16 nuvens de soluções encontradas pelo docking..... 49
- Figura 11. Energias de interação de van der Waals (preto) e eletrostática (vermelho) para cada molécula de colesterol ligada durante a simulação..... 52
- Figura 12. Moléculas de colesterol nos sítios de ligação ao final da simulação (~200 ns). a - c) As seis moléculas de colesterol nos sítios 1 ao 6 (S1 a S6). d- f) Detalhes da interação do colesterol nos sítios simulados. Em destaque estão regiões importantes do nucleossomo, como o 4-helix bundle H3:H3 (d), o domínio de docking (d) e as caudas N-terminais da H4 (e, f). Também estão destacados os aminoácidos do patch ácido (E91 e E92 da H2A), da cauda da H4 (H18, R19, K20, L22, R23) e os aminoácidos F78, E133 da H3 e R96, K105 da H2B, que foram relatados anteriormente interagirem com as NBPs. DNA (roxo) e histonas H3 (azul), H4 (verde), H2A (amarelo) e H2B (rosa) também estão representados..... 54
- Figura 13. RMSD dos átomos pesados do nucleossomo. a) Azul \rightarrow átomos do NCP na presença do colesterol; Verde \rightarrow átomos do NCP na ausência de colesterol; Vermelho \rightarrow átomos do NCP excluindo as caudas das histonas na presença do colesterol; Preto átomos do NCP excluindo as caudas das histonas na ausência de colesterol. Os valores com colesterol representam a média entre os sistemas PD1 e PD2. b) RMSD dos sítios de ligação do colesterol da simulação controle (preto) e da simulação com o ligante (vermelho)..... 57

Figura 14. Curvas do RMSF das histonas na presença (vermelho) e ausência (preto) de colesterol. Apenas o carbono alfa foi incluído no cálculo. Regiões destacadas representam o domínio de docking e a alça L1 da H2B.....	58
Figura 15. RMSF dos aminoácidos formadores dos sítios de ligação do colesterol. Os aminoácidos estão representados apenas pelo número na sequência, porém a ordem de apresentação é a mesma da Tabela 3. O índice -12 no sítio 1 representa um nucleotídeo (Timina 12). Apenas o carbono alfa foi incluído no cálculo.	59
Figura 16. Posição dependente do tempo (centro de massa) do colesterol em cada sítio de ligação.....	60
Figura 17. Distância dependente do tempo entre o colesterol e o respectivo sítio de ligação (centro de massa).....	61
Figura 18. Distância dependente do tempo entre a cauda de hidrocarbonetos do colesterol e o respectivo sítio de ligação (centro de massa).....	62
Figura 19. Distância dependente do tempo entre o grupo polar do colesterol e o respectivo sítio de ligação (centro de massa).....	62
Figura 20. Propriedades de hidratação do nucleossomo. a) Mudanças na quantidade líquida de moléculas de água a 3,5 Å de resíduos do nucleossomo. b) O número médio de moléculas de água que ocupam cada um dos sítios de ligação a 10 Å.....	64

LISTA DE TABELAS

Tabela 1. Energia livre de ligação média calculadas pela scoring function para cada sítio de ligação do colesterol.....	50
Tabela 2. Resultados obtidos para as energias de interação do colesterol no nucleossomo através do LIE em cada sítio. São apresentadas as energias livres de ligação absoluta, a constante de ligação e as médias dos potenciais de van der Waals e eletrostática.....	52
Tabela 3. Aminoácidos formadores das cavidades do colesterol. * Em negrito estão os aminoácidos envolvidos em interação com as NBPs . ° Aminoácidos 80 a 119 fazem parte do domínio de docking.....	55
Tabela 4. Distância média, em angströms (Å), entre as moléculas de colesterol e o respectivo sítio de interação. Valores médios e desvios das distâncias apresentadas nas figuras 17, 18 e 19.....	71

LISTA DE ABREVIATURAS E SIGLAS

CENP-A: *Centrome Protein A*

CHARMM: *Chemistry at Harvard Macromolecular Mechanics.*

HO: *Histone Octamer*

NAMD: *Nanoscale Molecular Dynamics program.*

NBP: *Nucleosome Binding Proteins.*

NBM: *Nucleosome Binding Molecules.*

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*

RMSD: *Root Mean Squared Deviation.*

RMSF: *Root Mean Squared Fluctuation.*

VdW: *van der Waals.*

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 Cromatina.....	14
1.1.1 Partícula Central do Nucleossomo (NCP).....	16
1.1.1.1 Domínio central globular.....	18
1.1.1.2 Extensões N e C terminais (caudas) das histonas.....	20
1.2 Moléculas ligantes de nucleossomos (NBMs).....	22
1.2.1 NBPs.....	22
1.2.2 Íons e lipídios.....	24
1.2.3 Colesterol.....	25
1.3 Ferramentas computacionais.....	27
2 OBJETIVOS.....	29
2.1 Objetivo geral.....	29
2.2 Objetivos específicos.....	29
3 MATERIAIS E MÉTODOS.....	30
3.1 Simulações de Dinâmica Molecular.....	31
3.1.1 Campo de força.....	32
3.1.2 Tratamento das Interações.....	34
3.1.2.1 Condições Periódicas de Contorno.....	36
3.1.3 Equações de movimento.....	36
3.1.4 Ensemble NPT.....	38
3.1.5 RMSD e RMSF.....	39
3.2 Docking molecular.....	40
3.3 Avaliação da Energia de ligação – LIE (Linear Interaction Energy).....	41
3.4 Protocolos.....	42
3.4.1 Protocolo de simulações de DM.....	42
3.4.1.1 Controle.....	44
3.4.1.2 Colesterol livre em solução.....	44

3.4.1.3 Pós docking (PD).....	44
3.4.2 Protocolo de Docking Molecular.....	45
3.4.3 Protocolo para o cálculo do LIE:.....	46
4 RESULTADOS.....	47
4.1 Estabilidade estrutural do ensemble de equilíbrio.....	47
4.2 Sítios de interação do colesterol no nucleossomo.....	49
4.2.1 Energia livre de ligação.....	51
4.3 Dinâmica da interação colesterol-nucleossomo.....	53
4.3.1 Mapa de contatos.....	53
4.3.2 Impacto estrutural do colesterol no nucleossomo.....	56
4.3.3 Dinâmica do colesterol no sítio.....	59
4.3.4 Perfil de hidratação dos sítios.....	63
5 DISCUSSÃO.....	65
6 CONCLUSÕES.....	74
7 REFERÊNCIAS.....	75

1 INTRODUÇÃO

1.1 Cromatina

O DNA é essencial para os processos biológicos, contendo o código genético que define o fenótipo de um indivíduo. No *homo sapiens*, cada célula armazena por volta de 6 bilhões de pares de bases. Se esticado, seriam aproximadamente 2 metros de DNA, compactado dentro do microscópico núcleo eucariótico, cujo diâmetro médio é menor do que 10 μm . Para atingir tal compactação, espécies que vão desde leveduras a humanos acomodam seu genoma em um complexo polimérico conhecido como cromatina, que possui o nucleossomo como uma unidade repetitiva (MCGINTY; TAN, 2014). Os nucleossomos são os blocos primários de formação da cromatina, organizando aproximadamente 147 pares de bases de DNA. Para isto, proteínas empacotadoras de DNA denominadas histonas são essenciais.

Além de sua função estrutural, a cromatina também está relacionada a manutenção do genoma e regulação da expressão gênica, desempenhando papel repressor crucial para o funcionamento celular, regulando o acesso de proteínas, como fatores de transcrição, ao genoma. Entretanto, a transcrição gênica deve ocorrer em várias etapas da vida celular, assim a cromatina é altamente dinâmica permitindo modificações na sua organização e compactação (HORN; PETERSON, 2002). Tais modificações podem ser iniciadas por alguns processos já bem caracterizados, entre os quais destacam-se: as modificações pós-traducionais de histonas, incorporação de variantes de histonas ao nucleossomo e interação com proteínas remodeladoras da cromatina (Revisado em: KALASHNIKOVA *et al.*, 2013; MUSSELMAN *et al.*, 2012; WEBER; HENIKOFF, 2014).

Didaticamente, podemos dividir a cromatina em dois estados organizacionais extremos: um estado compactado, conhecido como heterocromatina, geralmente associado ao silenciamento gênico e baixa atividade transcricional e outro estado, denominado eucromatina, no qual a cromatina encontra-se mais relaxada e, conseqüentemente, o DNA está mais acessível a maquinaria transcricional basal (GREWAL; ELGIN, 2007; HSU, 1962). A cromatina também pode ser classificada de acordo com os níveis sucessivos de condensação do DNA. A repetição de nucleossomos conectados por segmentos de DNA – nomeados *linker* DNA, variando de 15 a 100 pares de bases (KORNBERG, 1977) – forma a fibra de 10 nm, conhecida como “colar de contas”. A *linker* histona (H1 ou H5) pode associar-se a fibra de 10 nm gerando uma fibra mais compacta, com aproximadamente 30 nm (LI; ZHU, 2015). Porém, a dobragem específica da fibra de 30 nm é bastante controversa, pois, apesar de a sua ocorrência *in vitro* ser suportada por vários autores (ROBINSON *et al.*, 2006), ainda não foi possível visualizá-las no contexto celular (FUSSNER; CHING; BAZETT-JONES, 2011). A compactação da cromatina continua aumentando a densidade de DNA empacotado até seu nível máximo, o cromossomo (Figura 1).

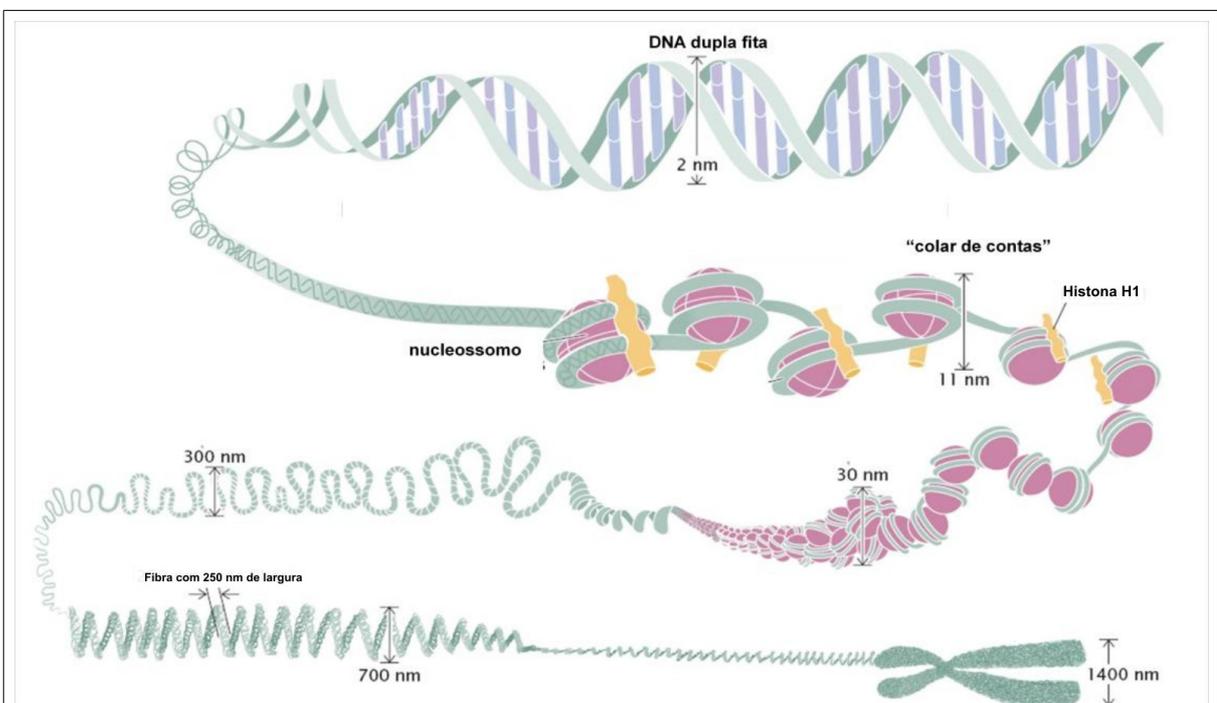


Figura 1. Ilustração da compactação do DNA. Adaptado de (PIERCE, 2007)

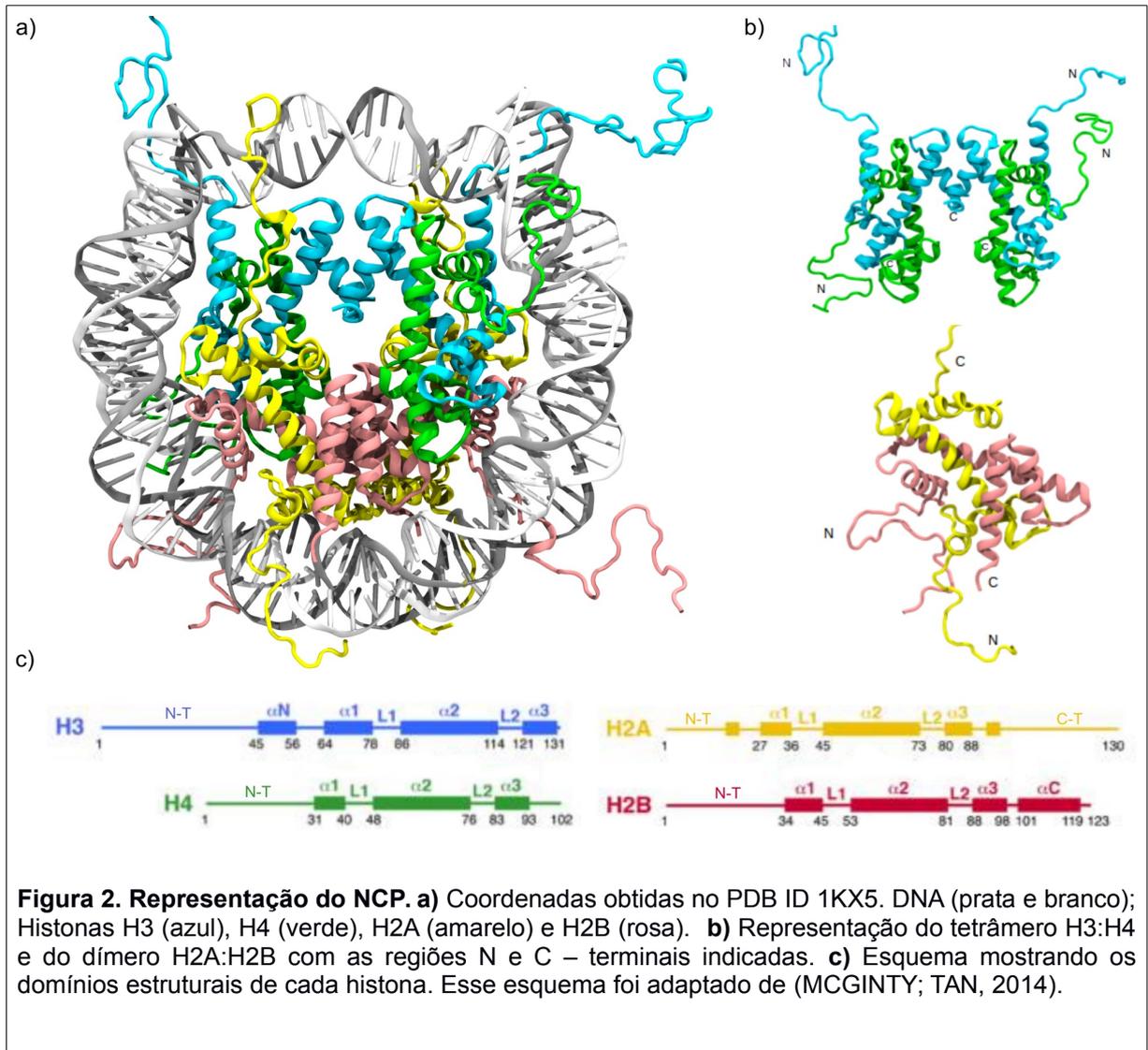
1.1.1 Partícula Central do Nucleossomo (NCP)

Cada nucleossomo, unidade repetitiva da cromatina, contém uma partícula central do nucleossomo (NCP, **Nucleosome Core Particle**) constituída por 145 a 147 pares de bases de DNA enrolados em um complexo octamérico de histonas. Ainda, conhecemos o cromatossomo, que é formado a partir da junção entre a *linker* histona, histona que não pertence ao NCP, e a partícula central do nucleossomo acrescentada por ~20 pares de bases (SIMPSON, 1978).

Os primeiros estudos estruturais sobre o nucleossomo aconteceram há quase 40 anos (FINCH *et al.*, 1977; FINCH; KLUG, 1978), onde foi possível deduzir apenas a forma de disco do NCP e a supertorção negativa para esquerda do DNA em torno do octâmero. Dados que foram confirmados posteriormente por um experimento de cristalografia de raios X com baixa resolução, 7 Å (RICHMOND *et al.*, 1984). Apenas em 1997, a primeira estrutura cristalográfica em alta resolução (2,8 Å) foi publicada, revelando detalhes atômicos do nucleossomo, especialmente as interações entre histonas e histonas-DNA (LUGER *et al.*, 1997). Esse cristal mostra 146 pares de bases de DNA alfa-satélite enrolados 1,65 vezes em torno do octâmero de histonas (HO), formado por duas cópias das histonas H2A, H2B, H3 e H4. O nucleossomo apresenta simetria bilateral 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.

A associação do DNA ao octâmero de histonas é forte o suficiente para protegê-lo de digestões enzimáticas por endonucleases, como a nuclease microcócica (VAN HOLDE, 1989). A neutralização das cargas das histonas proporcionada pelo DNA, é essencial para que o octâmero se mantenha estável em solução fisiológica. Na ausência de DNA, os oligômeros de histonas formam complexos estáveis, como os dímeros das histonas H2A e H2B e o tetrâmero, formado entre as histonas H3 e H4. Durante o processo de construção do nucleossomo, o tetrâmero (H3)₂-(H4)₂ forma um complexo estável com ~ 120 pares de base DNA, conhecido como tetrassomo. Posteriormente, os dímeros (H2A-H2B) se ligam a cada um dos lados do tetrâmero atingindo ~ 147 pares de base de DNA

enrolado (KARANTZA *et al.*, 1995; KARANTZA; FREIRE; MOUDRIANAKIS, 1996)
Figura 2.



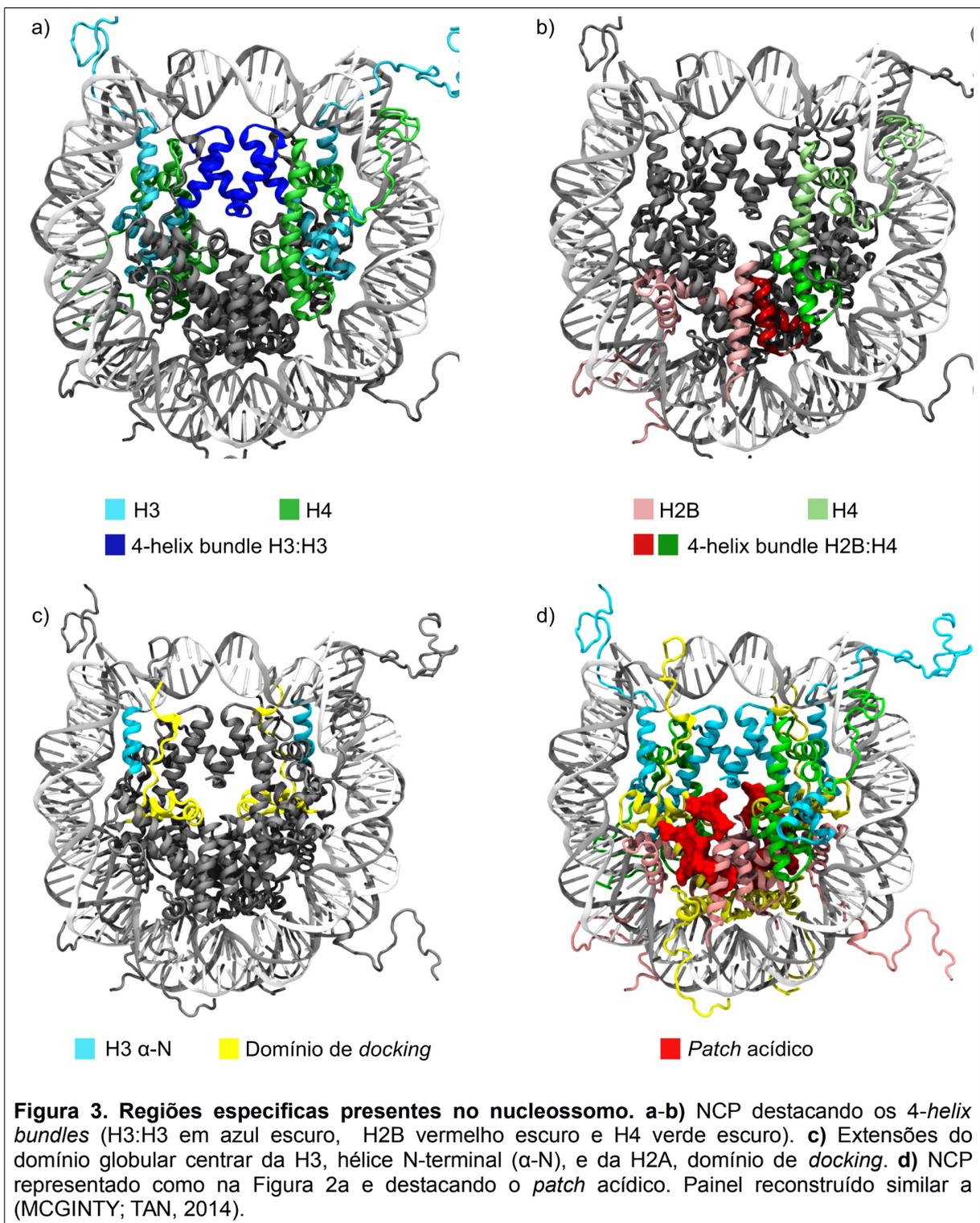
Além das histonas canônicas (H3, H4, H2A e H2B), existem também variantes de histonas que podem ser incorporadas ao nucleossomo, alterando a sua estrutura, dinâmica, estabilidade e, conseqüentemente, a arquitetura da cromatina. As variantes de histonas exercem funções específicas em vários processos celulares, incluindo início e término da transcrição, reparo do DNA e condensação de

cromossomos sexuais. Maiores detalhes podem ser encontrados nas revisões TALBERT; HENIKOFF, 2010 e WEBER; HENIKOFF, 2014.

As histonas são bastante conservadas evolutivamente, apresentando dois domínios estruturais distintos: (1) um domínio globular central, englobando $\sim \frac{3}{4}$ da massa total da histona, responsável pelas interações entre histonas no HO e (2) a massa restante consiste do domínio formado pela extensão N-terminal, também conhecido como cauda (ARENTS *et al.*, 1991), que são regiões desestruturadas e flexíveis.

1.1.1.1 Domínio central globular

O núcleo proteico do nucleossomo é constituído predominantemente por alfa-hélices, formando o arcabouço octamérico, em que o DNA se enrola. Nesse domínio central, as quatro histonas compartilham uma região estruturalmente muito similar, constituída por três alfa-hélices (α_1 , α_2 , α_3) conectadas por duas alças (L1 e L2), conhecida como motivo de dobra da histona. Cada motivo de dobra é unido no sentido anti-paralelo a outro motivo complementar, formando heterodímeros onde, a H3 interage com a H4 e a H2A com a H2B. Estes dímeros interagem por meio de ligações de hidrogênio e várias interações hidrofóbicas em uma interface conhecida como feixe de quatro hélices, “*4-helix bundle*”, formado pelas hélices α_2 e α_3 de cada motivo de dobra. Os dímeros H3-H4 tem baixa estabilidade em solução fisiológica e rapidamente formam um tetrâmero através de um *4-helix bundle* entre as histonas H3. Cada um dos dímeros H2A-H2B se associam a um dos lados do tetrâmero também através de um *4-helix bundle* entre a H2B e H4 (LUGER *et al.*, 1997). Essa região formada por quatro hélices é fundamental para a estabilização estrutural do HO, e conseqüentemente do nucleossomo (Figura 3).



Além do motivo de dobra, as histonas H3, H2B e H2A apresentam regiões estruturadas adicionais. A H2B possui uma hélice C-terminal que define o limite

longitudinal do octâmero de histonas. A H3 contém uma hélice N-terminal (α -N) que auxilia no empacotamento de aproximadamente 13 pares de base DNA na região da *dyad* (Figura 3). Essa hélice é guiada para a área de entrada e saída do DNA pela extensão C-terminal da H2A, conhecida como domínio de *docking* (aminoácidos 80-119 H2A). Além disso, a pequena hélice presente nesse domínio (~ 4 aminoácidos) interage com a extensão C-terminal da H4. Essas características sugerem que o domínio de *docking* seja importante para a formação do nucleossomo, auxiliando a ligação entre os dímeros (H2A-H2B) e o tetrâmero (H3-H4) (LUGER *et al.*, 1997). Interessantemente, foi observado que a deleção no domínio de *docking* impediu a associação da *linker* histona (H1) ao nucleossomo, pois o ângulo na região de entrada e saída do DNA essencial para ligação da H1 foi alterado (SHUKLA *et al.*, 2010; VOGLER *et al.*, 2010).

Na superfície do nucleossomo, destaca-se uma região onde oito aminoácidos carregados negativamente (E56, E61, E64, D90, E91, E92 da H2A e E102, E110 da H2B) formam uma região conhecida como *patch* ácido. Essa região é responsável por interações internucleossomais, que pode acontecer após a ligação da cauda da H4 do nucleossomo adjacente, colaborando para que estados mais compactados da cromatina sejam alcançados (DAVEY *et al.*, 2002; LUGER *et al.*, 1997). Dentre as proteínas ligantes de nucleossomo (NBP *Nucleosome Binding Proteins*) já investigadas por técnicas com resolução atômica, todas utilizam essa região como sítio de ligação no nucleossomo. Essa interação é capaz de desencadear a abertura ou fechamento da cromatina dependendo da proteína (SILVA; OLIVEIRA; SANTOS, 2015).

1.1.1.2 Extensões N e C terminais (caudas) das histonas

As extensões do domínio central estão localizadas nas regiões N-terminais de todas as histonas do HO e na região C-terminal da H2A, particularmente esta histona possui duas caudas totalizando dez no nucleossomo. Estas porções conhecidas como 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 são formadas, em sua maioria, por resíduos de arginina e lisina que lhes conferem carga positiva, dessa forma, possuem atração pelo grupo fosfato da cadeia principal do DNA. De fato, estruturas cristalográficas demonstraram que as caudas das histonas passam pelos sulcos menores do DNA em direção ao exterior do nucleossomo, auxiliando na organização local do DNA (DAVEY *et al.*, 2002; LUGER *et al.*, 1997). Além disso, a remoção das caudas intensifica o desenrolamento do DNA do nucleossomo aumentando a acessibilidade de fatores de transcrição ao DNA (LEE *et al.*, 1993; VETTESE-DADEY *et al.*, 1994; YANG *et al.*, 2005). Biswas e colaboradores realizaram uma série de simulações de DM com o truncamento de todas as caudas do nucleossomo, uma de cada vez. A retirada das caudas da H3 e H2A impactaram estruturalmente o domínio globular central das histonas, afetando a estabilidade do HO. Já a perda das caudas da H4 e H2B não apresentaram efeitos significativos sobre a estabilidade do HO (BISWAS *et al.*, 2011).

A cauda da H4 parece ser de extrema importância para a formação dos níveis mais compactados de cromatina, mediando interações internucleossomais. As primeiras evidências surgiram nas estruturas cristalográficas de 1997 e 2002. Foi observada a interação da cauda da H4 (resíduos 16 a 25) com o *patch* ácido (DAVEY *et al.*, 2002; LUGER *et al.*, 1997). Investigações posteriores, como a neutralização de cargas da lisina 16 da H4 (DORIGO *et al.*, 2003) e retirada de parte da cauda (ALLAHVERDI *et al.*, 2011; ROBINSON *et al.*, 2008) resultaram na formação incompleta da fibra de 30 nm *in vitro*. Yang e Arya, após extensivas simulações de DM, *docking* e cálculos de afinidade chegaram a conclusão que a cauda da H4 tem forte tendência a formar alfa-hélice e, nessa conformação, possui alta afinidade com o *patch* ácido (YANG; ARYA, 2011).

Uma das formas encontradas pela natureza para permitir o remodelamento constante e dinâmico da cromatina são as modificações pós-traducionais (PTM) das histonas, principalmente nas caudas. Tais marcas epigenéticas são desencadeadas por proteínas com função enzimática e podem agir indiretamente, servindo como marcadores para o reconhecimento de outras proteínas (MUSSELMAN *et al.*, 2012), ou diretamente, influenciando a interação entre nucleossomos. Por exemplo, a

acetilação na cauda da H4, na lisina 16, impede sua ligação ao *patch* ácido do nucleossomo adjacente, pois ocorre neutralização da carga positiva que mediava esse processo (SHOGREN-KNAAK *et al.*, 2006).

1.2 Moléculas ligantes de nucleossomos (NBMs)

A cromatina é um complexo extremamente dinâmico e sensível à ligação de proteínas e mudanças no ambiente nuclear. Várias proteínas e pequenas moléculas já foram identificadas como ligantes de nucleossomo. Entretanto, o perfil completo de NBMs capazes de regular a estrutura da cromatina ainda é alvo de intensa investigação.

1.2.1 NBPs

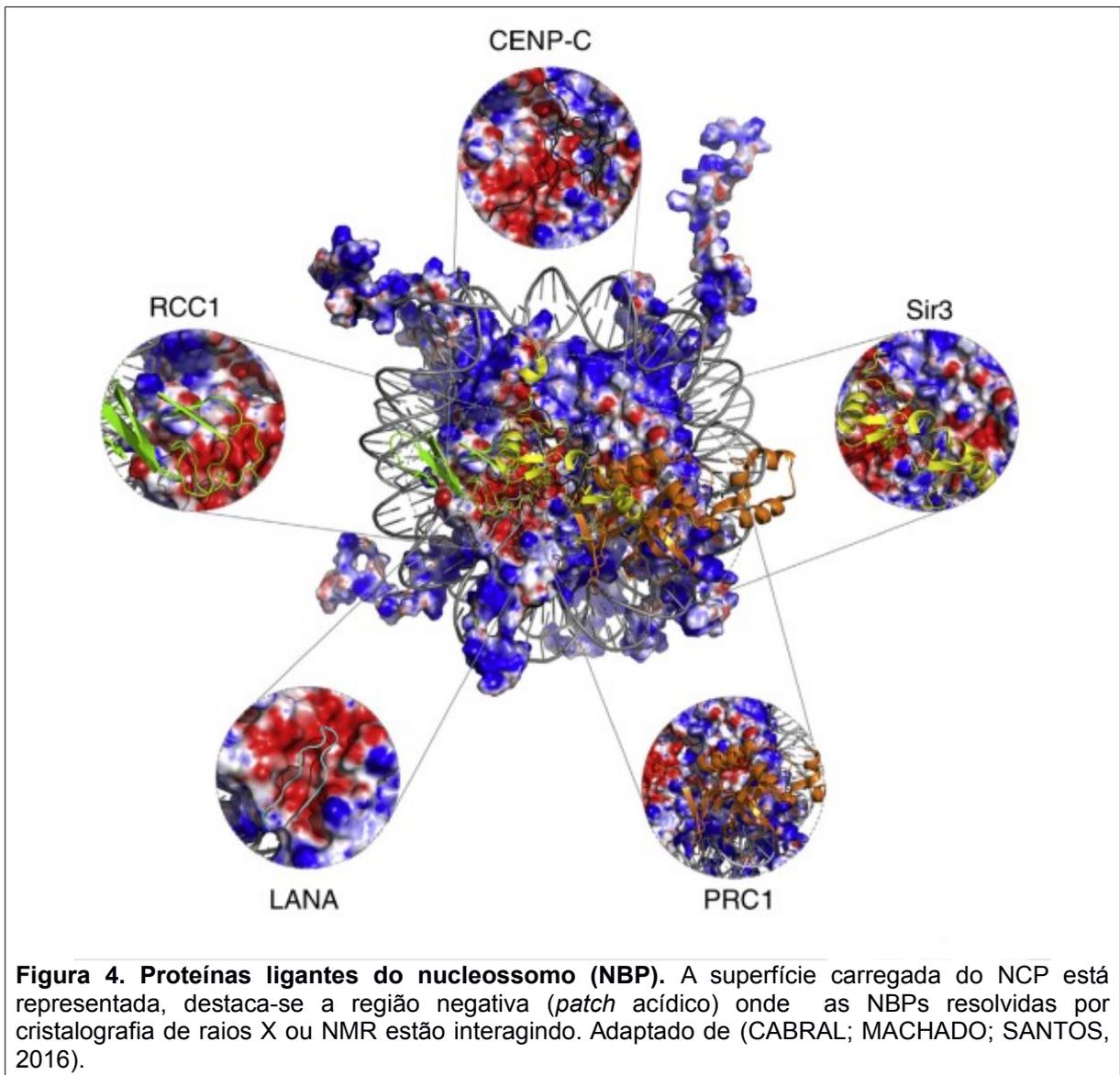
Até o momento, as NBPs (*Nucleosome Binding Proteins*) descritas com modelo estrutural atômico de ligação ao nucleossomo, disponíveis no PDB (*Protein Data Bank*), são as seguintes:

- LANA (*Latency-Associated Nuclear Antigen*) do KSHV (*Kaposi's Sarcoma-Associated Herpesvirus*). Foi o primeiro cristal publicado de uma molécula ligada ao nucleossomo. Um peptídeo formado por apenas 23 aminoácidos foi cristalizado, este se encaixa na cavidade do *patch* ácido formando um *hairpin* nessa região (BARBERA *et al.*, 2006);
- IL-33 (Interleucina **33**). Carriere e colaboradores demonstraram que a IL-33 está presente na cromatina e participa da regulação da transcrição (CARRIERE *et al.*, 2007). Além disso, foi identificada uma pequena porção da IL-33 (19 aminoácidos) que interagia com a cromatina. Os detalhes da ligação desse peptídeo ao NCP foram obtidos através de modelagem computacional, utilizando o cristal do LANA como guia. Similar ao LANA, o IL-33 forma um *hairpin* na região do *patch* ácido (ROUSSEL *et al.*, 2008).
- RCC1 (*Regulator of Chromosome Condensation 1*). Primeira proteína intacta cristalizada em complexo com o nucleossomo. O RCC1 é importante para o processo de mitose e transporte de macromoléculas entre o citoplasma e o

núcleo. No cristal, o RCC1 interage com o *patch* ácido nos dois lados do nucleossomo. Além disso, foi identificado que essa NBP também realiza contato com o DNA (MAKDE *et al.*, 2010, p. 1);

- SIR3 (*Silent Information Regulator 3*). A estrutura cristalográfica resolvida apresenta apenas o domínio BAH (*Bromo-Associated Homology*) do SIR3, região descrita como alvo de várias mutações que resultam no silenciamento transcricional. Foi possível observar interfaces adicionais de interação com a NCP, como as hélices terminais da H2B, α -3 e α -C, e a cauda da H4 (ARMACHE *et al.*, 2011);
- CENP-C (*Centromere Protein C*). Nesse cristal, apenas um peptídeo da CENP-C foi cristalizado com o nucleossomo contendo uma variante da histona H3, CENP-A. Os nucleossomos presentes na região centromérica contêm a histona variante CENP-A, a qual auxilia na segregação cromossômica. Na estrutura, o peptídeo interage com o *patch* ácido e com regiões específicas da CENP-A (KATO *et al.*, 2013);
- PRC1 (*Polycomb Repressive Complex 1*). O PRC1 é um módulo de ubiquitinação formado por três proteínas, que participa na repressão da transcrição. A estrutura resolvida mostra que as três proteínas juntas, interagem com todos os componentes do nucleossomo. Porém o *patch* ácido e alguns outros aminoácidos básicos são fundamentais para a formação do complexo PRC1-NCP (MCGINTY; HENRICI; TAN, 2014);
- 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, além disso participam de outros processos celulares, como no reparo do DNA e na extensão das PTMs de histonas. Os autores utilizaram ressonância magnética nuclear (NMR) para chegar a um modelo de interação, no qual um peptídeo da HMGN2 (24 aminoácidos) formava um “grampo” entre o DNA nucleossomal, ligando-se na

dyad, e o núcleo proteico do nucleossomo, ligando na região do *patch* ácido (KATO *et al.*, 2011).



1.2.2 Íons e lipídios

Além das proteínas, já foram caracterizados outros importantes moduladores da estrutura da cromatina. O magnésio (Mg^{2+}), por exemplo, é capaz de desencadear a condensação máxima das fibras em uma concentração pequena, apenas 1 mM (KOROLEV *et al.*, 2010). Ademais, vários estudos sinalizaram que os

lipídios, como o colesterol e ácidos graxos de cadeia curta, funcionam como componentes reguladores da estrutura da cromatina. Estes podem ativar receptores de superfície celular ou interagir com fatores regulatórios citoplasmáticos que estimulam ou reprimem vias específicas de transdução de sinal (Revisado em ZAINA *et al.*, 2005). Apesar de ter sido observado no núcleo da célula e presente na cromatina (ALBI; MAGNI, 2004), a ligação de nenhum lipídeo ao nucleossomo foi estudado ao nível atômico.

1.2.3 Colesterol

Neste trabalho, focamos no colesterol, um lipídio celular ubíquo, como uma NBM. O colesterol é o principal esterol sintetizado pelo organismo humano, atuando como precursor dos hormônios esteroides (progesterona, estrógeno, testosterona, glicocorticoides e mineralocorticoides), ácidos biliares e vitamina D (HANUKOGLU, 1992). Além de ser primordial para manter a integridade da membrana celular, o colesterol associado com esfingolipídios e proteínas específicas formam domínios membranares conhecidos como “*lipid rafts*”. Tais domínios estão envolvidos na transdução de sinais, reconhecimento molecular, endocitose, dentre outros processos celulares (SIMONS; TOOMRE, 2000). Diversas patologias estão associadas com alterações na homeostase do colesterol no organismo, sendo a aterosclerose mais comum (SIMONS; IKONEN, 2000). Quimicamente, o colesterol é um álcool de cadeia longa e contém uma estrutura de anéis esteroidais com uma cabeça polar e uma cauda ramificada de hidrocarbonetos (apolar), portanto é uma molécula anfipática, ou seja, possui uma pequena região hidrossolúvel e outra lipossolúvel (ALBERTS *et al.*, 1994) (Figura 5). De forma geral, os lipídios são formados majoritariamente por regiões apolares, sendo assim insolúveis em água e conseqüentemente no sangue. Portanto, o colesterol é transportado por apolipoproteínas (Apo), que se ligam às moléculas de colesterol e formam as lipoproteínas (GRUMMER; CARROLL, 1988).

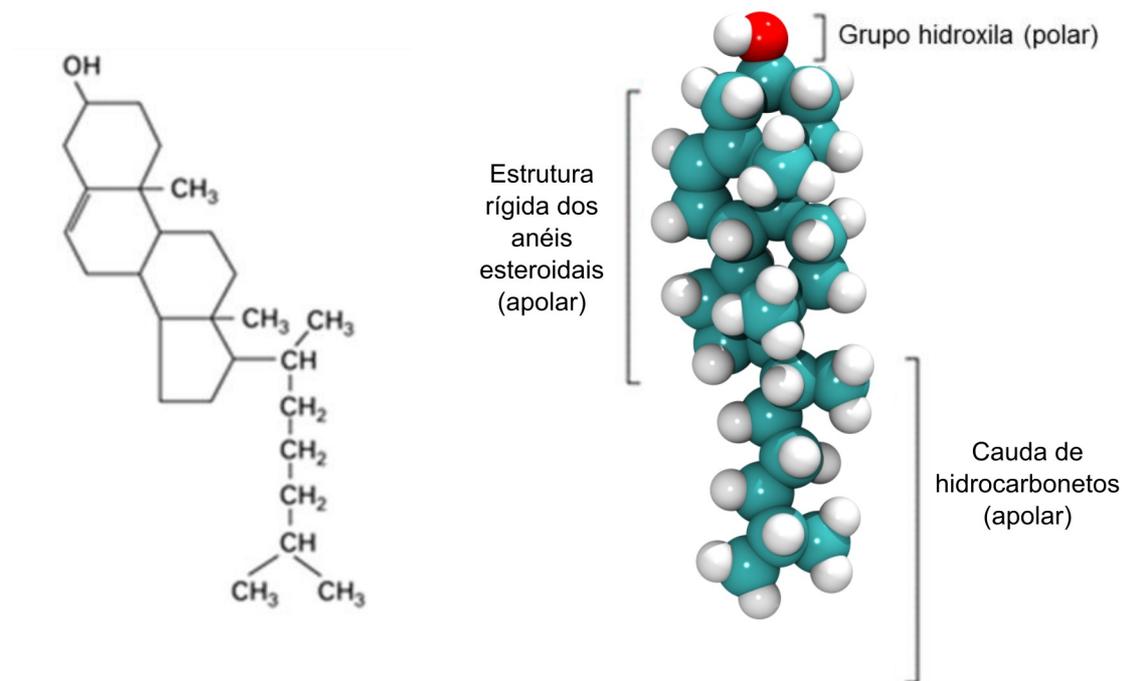


Figura 5. Estrutura química e representação 3D do colesterol

Os primeiros indicativos da presença de colesterol no ambiente nuclear surgiram há mais de 30 anos. O colesterol marcado administrado em ratos por gavagem, foi observado, após 2 horas, no núcleo celular e em associação com as fibras de cromatina (ERICKSON; DAVISON; GOULD, 1975). Ademais, células de fígado de camundongo incubadas com colesterol marcado, em intervalos de tempo e concentrações diferentes, mostrou que o colesterol se liga à cromatina, possivelmente através de uma ou mais proteínas. (REGENASS-KLOTZ; HEINIGER, 1984).

Alguns anos mais tarde, um sinal da ligação do colesterol à cromatina foi apresentado por cromatografia de hepatócitos de camundongos. Os autores observaram que na fração lipídica isolada da cromatina, o colesterol encontrava-se complexado à esfingomiélna. Além disso, durante a regeneração do fígado dos camundongos, foi constatado um aumento nos níveis de colesterol na cromatina,

indicando que a presença de colesterol é importante para a fase de proliferação celular (ALBI; MAGNI, 2002).

Em 2013, o grupo do professor Guilherme Santos avaliou o efeito do colesterol sobre fibras de cromatina de 10 nm, reconstituídas *in vitro*, submetidas ao aumento gradual da temperatura. Neste trabalho, o colesterol proporcionou uma perda gradual de nucleossomo das fibras, evitando a perda súbita que ocorre sem colesterol. Também foi verificado que o colesterol antecipa o ponto de saturação da fibra de 10 nm, requerendo uma menor concentração de octâmero de histonas para a formação da fibra de 10 nm (SILVA, 2013). Recentemente, foi demonstrado que a forma compactada da cromatina, conhecida como fibra de 30 nm, foi também favorecida pela presença de colesterol. Além disto, testes de estabilidade térmica do nucleossomo em presença de baixas concentrações de colesterol mostraram não afetar a estrutura do nucleossomo, mas altas concentrações antecipam a dissociação do complexo H2A-H2B e posteriormente do tetrassomo de H3-H4. Estes dados permitem considerar que o colesterol pode alterar a estrutura do nucleossomo e cromatina, possivelmente afetando processos nucleares importantes para a expressão gênica e manutenção do genoma (artigo submetido a publicação).

Assim, as evidências bioquímicas encontradas até o momento sugerem um potencial efeito modelador do colesterol sobre as fibras de cromatina, porém os detalhes atômicos moleculares e os modos de interação entre este lipídio e o nucleossomo ainda não estão claros.

1.3 Ferramentas computacionais

Técnicas experimentais como a cristalografia de raios X, podem fornecer detalhes atômicos importantes da estrutura de biomoléculas. Entretanto, tais informações ainda são insuficientes para entendermos por completo a atividade biológica dessas moléculas. Uma alternativa que permite a investigação da dinâmica de sistemas moleculares são as simulações computacionais. A partir dessas simulações, escalas menores de tempo e tamanho são estudadas, uma espécie de “microscópio computacional”. Transporte de pequenas moléculas pela membrana

celular, ligação droga-receptor, estados conformacionais transitórios e enovelamento de proteínas são alguns exemplos que podem ser estudados por tais simulações. Entre as principais restrições desse tipo de experimento está o custo computacional e o desenvolvimento de modelos físicos apropriados (DROR *et al.*, 2012).

Entre as técnicas computacionais existentes, as simulações atômicas de dinâmica molecular (DM) são largamente empregadas em estudos estruturais e dinâmicos de um sistema. Essa técnica já permitiu observar detalhes atômicos de diferentes sistemas biológicos, numa escala de tempo de aproximadamente 200 ns (BISWAS; LANGOWSKI; BISHOP, 2013). Com relação à cromatina, Bowerman e Wereszczynski utilizaram DM para obter informações importantes sobre o possível mecanismo pelo qual variantes da H2A (macro-H2A, e H2A.Z) alteram a estabilidade do nucleossomo. Eles observaram que a macro-H2A propicia estados mais estáveis do nucleossomo alterando a dinâmica e a energética do nucleossomo, podendo assim conduzir ao silenciamento gênico (BOWERMAN; WERESZCZYNSKI, 2015). Outra pesquisa demonstrou que quando lisinas das caudas da H3 e H4 são acetiladas, há um aumento probabilístico de formação de estruturas secundárias nessas caudas. Os autores sugerem que a redução da mobilidade resultante seja responsável por prejudicar as interações internucleossomais, afetando a condensação da cromatina (COLLEPARDO-GUEVARA *et al.*, 2015).

2 OBJETIVOS

2.1 Objetivo geral

Investigar *in silico* as características da interação do complexo formado entre o nucleossomo e o colesterol e avaliar o impacto dessa interação ao nível atômico.

2.2 Objetivos específicos

- Construir e simular os sistemas atomísticos do nucleossomo e do colesterol;
- Buscar os sítios de ligação do colesterol no nucleossomo através de *docking* molecular;
- Avaliar, dentre os sítios encontrados, quais estão condizentes com a literatura e que possam ajudar a esclarecer o efeito do colesterol observado *in vitro*;
- Simular por DM o complexo nucleossomo-colesterol;
- Identificar os resíduos essenciais para a ligação do colesterol ao nucleossomo;
- Observar como a ligação do colesterol interfere na estrutura do nucleossomo;
- Calcular a energia livre de ligação do colesterol ao nucleossomo

3 MATERIAIS E MÉTODOS

Em função do nível de detalhamento almejado nesse estudo, metodologias capazes de tratar interações moleculares com resolução atômica são fundamentais. Assim, para que as metas aqui colocadas fossem alcançadas, métodos teóricos e computacionais foram empregados. No que diz respeito à dinâmica molecular (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 e protocolos utilizados nas simulações, bem como o desenho experimental seguido. Inicialmente, um ensemble de estruturas em equilíbrio do nucleossomo foi gerado abrangendo ~20 ns. Essa trajetória foi utilizada como receptor na busca dos possíveis sítios de interação do colesterol no nucleossomo. Após análise qualitativa e quantitativa (scoring function) dos resultados do docking, seis moléculas de colesterol foram simuladas por DM ligadas ao nucleossomo. A partir dessa nova trajetória, cálculos de energia livre de ligação foram realizados para resolver as afinidades específicas do colesterol em cada sítio amostrado. Além disso, foram realizadas análises acerca do impacto estrutural causado pelo colesterol no nucleossomo, bem como alterações superficiais do seu perfil de hidratação (Figura 6).

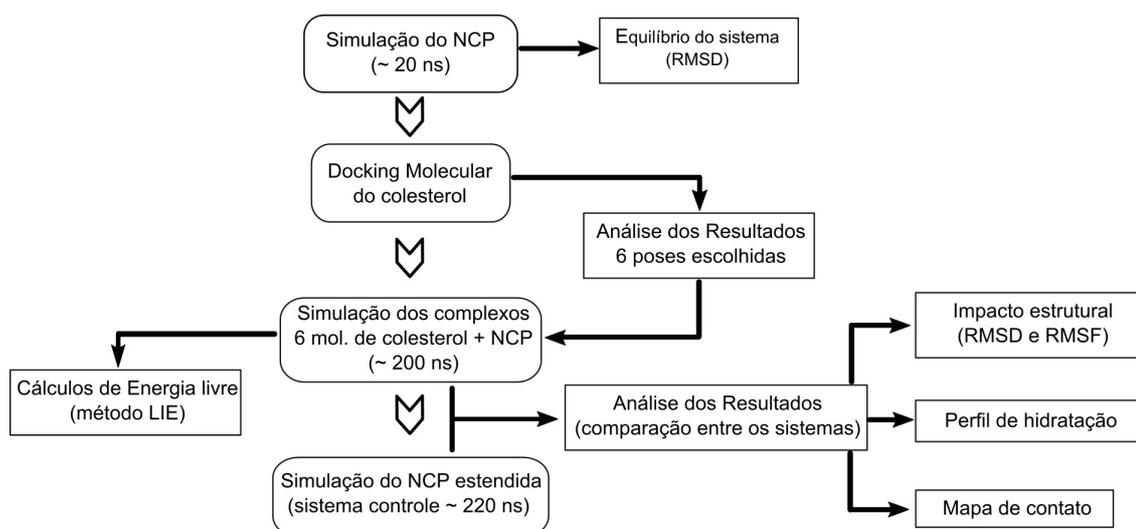


Figura 6. Resumo das etapas seguidas nesse trabalho. As setas sem preenchimento significam etapas que foram seguidas após as respectivas análises (setas pretas).

3.1 Simulações de Dinâmica Molecular

A simulação de DM é uma metodologia computacional capaz de produzir trajetórias para um sistema composto por um número finito de partículas (átomos). De acordo com a mecânica clássica, é obtida 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.

Porém, antes de calcularmos as forças entre os átomos e aplicarmos movimentos a estes, é imprescindível determinar as configurações iniciais do sistema que se busca estudar. Atualmente, existem duas formas principais de obter as coordenadas iniciais de sistemas biológicos complexos, como as proteínas e ácidos nucleicos, a partir de técnicas experimentais (cristalografia de raios X ou NMR) ou através de estudos de homologia. Tais estruturas podem ser obtidas no banco de dados *Protein Data Bank* (PDB), um repositório de coordenadas atômicas de estruturas moleculares obtidas via metodologias de alta resolução, como cristalografia de raios X, RMN e outras.

De modo geral, vários algoritmos são empregados na DM para garantir a precisão e controle do usuário sobre o sistema simulado.

Grande parte dos sistemas simulados em DM são ergódicos, ou seja há correspondência entre as médias temporais de uma trajetória e as médias por todo o espaço de fase. Assim, as médias temporais \bar{A}_t das trajetórias obtidas pela DM, podem ser utilizadas para o cálculo de propriedades termodinâmicas médias $\langle A \rangle$ de qualquer propriedade A do sistema (temperatura, pressão, energia livre, etc):

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

3.1.1 Campo de força

Para efetivamente movimentar um átomo i , primeiro é necessário realizar o cálculo das forças exercidas pelos demais átomos j sobre o átomo i . Para tal, funções que descrevem a energia potencial da interação entre pares de átomos são utilizadas. Estas funções, devem aliar precisão para gerar resultados fidedignos e eficiência, evitando alta demanda computacional para resolvê-las. Além disso, parâmetros moleculares que descrevem as constantes de equilíbrio e cargas parciais entre os diferentes átomos são empregados a fim de reproduzir propriedades do sistema em estudo. Esses dados são obtidos através de extensivos experimentos laboratoriais e cálculos quânticos, servindo como um calibrador para aperfeiçoar a precisão da função da energia potencial. O conjunto de parâmetros acompanhados de uma função de energia potencial é conhecido como campo de força. Existem vários campos de força, dentre os quais AMBER (WANG *et al.*, 2004), CHARMM (FOLOPPE; MACKERELL, 2000; MACKERELL *et al.*, 1998) e GROMOS (SCHULER; DAURA; VAN GUNSTEREN, 2001) são os mais utilizados em simulações de biomoléculas.

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. Por ele a energia potencial de uma partícula ($U(\mathbf{r})$) dependente da sua posição \mathbf{r} é resultado da soma de diferentes potenciais ($U = E_{\text{ligação}} + E_{\text{ângulo}} + E_{\text{diedro}} + E_{\text{Coulomb}} + E_{\text{vdw}}$). Tais potenciais podem ser divididos entre aqueles que descrevem as energias de interação entre os átomos

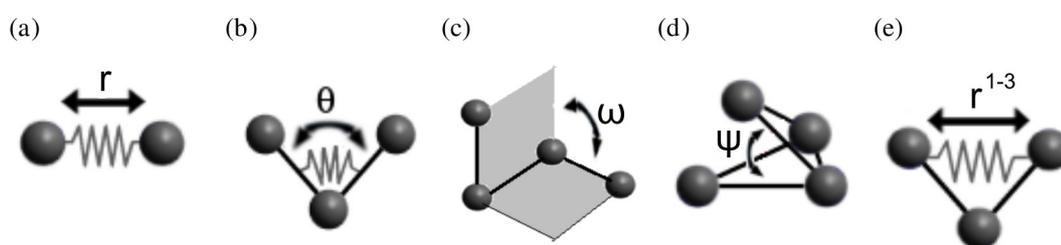


Figura 7. Ilustração dos potenciais entre átomos ligados empregados na DM. a) Energia de ligação; **b)** Energia do ângulo **c)** Energia do diedro **d)** Energia do diedro impróprio **e)** Energia Urey-Bradley.

ligados covalentemente e $(E_{\text{ligação}} + E_{\text{ângulo}} + E_{\text{diedro}})$ e não ligados $(E_{\text{Coulomb}} + E_{\text{vdw}})$ (Figura 7).

Especificamente, as contribuições para a energia potencial entre átomos ligados são descritas da seguinte forma:

$E_{\text{ligação}} = \sum k_r (r - r_0)^2$: potencial harmônico que representa a interação entre dois átomos separados por uma ligação covalente, a constante de força k_r e a distância de equilíbrio r_0 são parâmetros específicos para cada par de átomos ligados.

$E_{\text{ângulo}} = \sum k_\theta (\theta - \theta_0)^2$: potencial harmônico da torção do ângulo θ formado por três átomos em relação ao valor de equilíbrio θ_0 . k_θ é a constante de força angular dessa interação.

$E_{\text{diedro}} = \sum k_d [1 + \cos(n\omega - \omega_0)]$: energia potencial periódica da torção do ângulo diedral ω entre dois planos formado por quatro átomos ligados. k_d representa a barreira torsional ou constante de força diedral, n é a periodicidade e ω_0 o ângulo diedral de equilíbrio, ou fase.

Vale ressaltar que não existem termos referentes a formação ou quebra de ligações covalentes, pois na dinâmica molecular esses eventos não são modelados.

As energias de interação entre átomos não ligados, ou seja, dependentes apenas da distância entre esses átomos, são representadas por dois tipos de funções:

$E_{\text{Coulomb}} = \sum \frac{q_i q_j}{\epsilon r_{ij}}$: energia potencial eletrostática, ou seja energia das interações iônicas, entre as cargas elétricas parciais ou totais dos átomos i e j (q_i e q_j). Esse potencial é descrito de acordo com a Lei de Coulomb, onde ϵ é a constante dielétrica do ambiente e r_{ij} a distância entre os átomos.

$$E_{vdw} = \sum \varepsilon_{ij} \left[\left(\frac{R_{ij}^0}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^0}{r_{ij}} \right)^6 \right] : \text{energia potencial resultante das interações de van}$$

der Waals na forma do potencial de Lennard-Jones 6-12. Esse termo descreve a tendência que os átomos não ligados têm de se atraírem quando se aproximam da distância de equilíbrio R_{ij}^0 . E a forte repulsão quando estes estão muito próximos.

r_{ij} representa a distância entre os átomos e ε_{ij} corresponde à força de van der Waals (valor mínimo do potencial) para o par ij .

O CHARMM ainda fornece dois termos adicionais relacionados com restrições geométricas nas moléculas. Caso necessário esses termos são adicionados no cálculo da energia potencial U . São eles:

$E_{impróprio} = k_{\psi} (\psi - \psi_0)^2$: 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. k_{ψ} Representa a constante de força e ψ_0 o ângulo de equilíbrio da interação.

$E_{Urey-Bradley} = k_U (r - r_0)^2$: 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. k_U e r_0 correspondem a constante de força e à distância de equilíbrio, respectivamente.

3.1.2 Tratamento das Interações

O cálculo da energia potencial de cada átomo do sistema é a etapa mais dispendiosa da DM, pois são avaliados as contribuições de todas partículas que interagem com esse átomo. Para as interações ligadas, são considerados apenas os átomos distantes em até três ligações covalentes. Contudo, o leque de interações não ligadas é enorme, assim a determinação de um raio de corte (*cutoff*) aumenta a eficiência do cálculo.

Podemos definir um limite, geralmente entre 10 Å e 15 Å, para o cálculo das interações definidas pelo potencial de Lennard-Jones. Pois sabe-se que, conforme a distância entre dois átomos se estende esse potencial tende a zero rapidamente. Entretanto, para preservar a acurácia e a conservação de energia da simulação, o corte repentino do potencial nas fronteiras do *cutoff* não é aplicado. Em vez disso, é utilizada uma função (*switching function*) que, em um intervalo de distâncias predeterminado, modifica o potencial de Lennard-Jones permitindo que este se aproxime de valores nulos de forma gradual. Dessa maneira, a energia da interação de van der Waals torna-se desprezível a intervalos maiores do que o raio de corte escolhido.

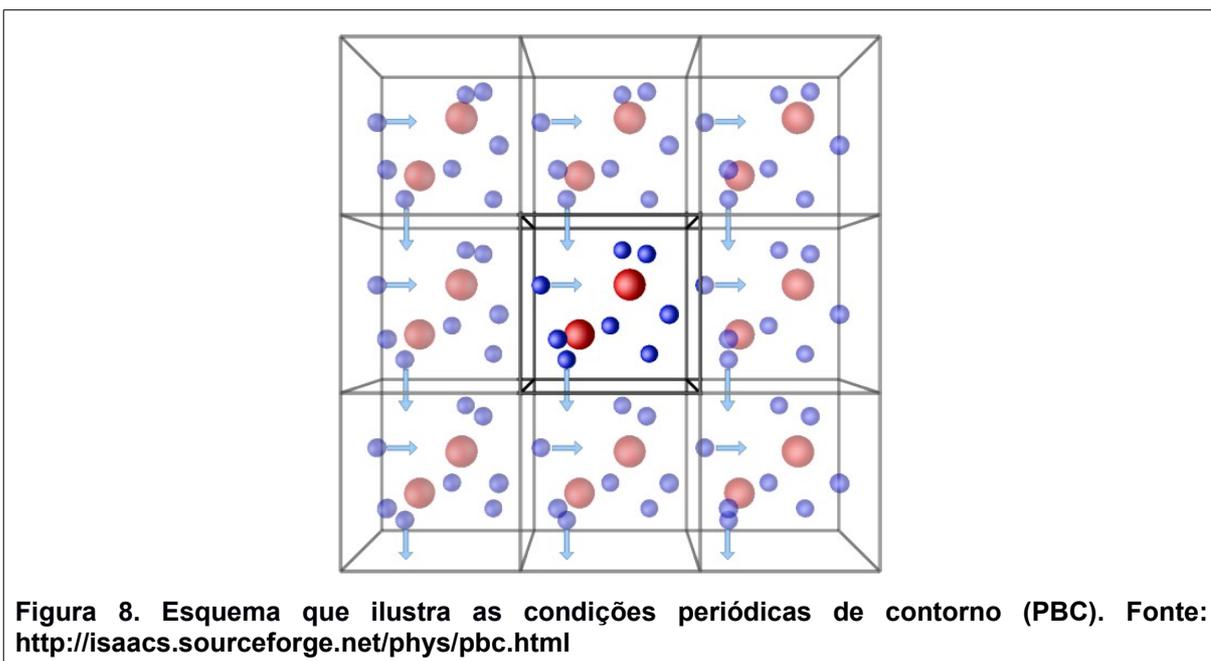
Todavia, as interações eletrostáticas não seguem o mesmo perfil do potencial de Lennard-Jones, pois para certos pares de átomos a energia eletrostática, mesmo a grandes distâncias, deve ser avaliada. Por conseguinte, as interações de Coulomb são tratadas decompondo a energia potencial eletrostática φ em dois componentes (φ_{curto} e φ_{longo}). O componente de curto alcance φ_{curto} é calculado explicitamente, considerando apenas partículas vizinhas. Já para o cálculo dos componentes de longo alcance φ_{curto} , somas de Ewald são empregada por meio do método conhecido como PME (***Particle Mesh Ewald***). Assim, a energia potencial eletrostática pode ser reescrita como:

$$E_{Coulomb} = \sum_{i,j} \varphi_{curto}(r_j - r_i) + \sum_k \varphi_{longo} |\rho(k)|^2 \quad (2)$$

onde φ_{longo} é a transformada de Fourier do potencial eletrostático e $\rho(k)$ a quantidade de cargas por volume, ou seja, a densidade de cargas. Com o emprego dessa compartimentalização das energias em de curto e longo alcance, o cálculo do potencial eletrostático fica mais rápido, já que ambos os termos convergem rapidamente (BATCHO; CASE; SCHLICK, 2001; TOUKMAJI; BOARD, 1996).

3.1.2.1 Condições Periódicas de Contorno

Geralmente na DM não há variação no número de partículas presente no sistema simulados, dessa maneira, devido a característica finita da simulação, uma parcela significativa de átomos encontra-se nos limites do sistema e podem sofrer efeitos de borda. Com o intuito de tratar tais efeitos, condições periódicas de contorno (PBC, *Periodic Boundary Conditions*) são frequentemente utilizadas em simulações de DM. Utilizando esse esquema, réplicas do sistema original são criadas em todos os eixos ($\pm x$; $\pm y$; $\pm z$), criando 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 8). As interações interatômicas são calculadas de acordo com a distância mínima entre os átomos, descontando as dimensões da caixa.



3.1.3 Equações de movimento

A trajetória de cada átomo é obtida a partir da integração das equações newtonianas de movimento. Geralmente, 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 *Velocity Verlet* (SWOPE *et al.*, 1982; 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 fórmula:

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

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

onde $U(r_i)$ é a função da energia potencial da partícula i , m_i sua massa e $O(dt^4)$ é o erro de 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 velocidades dos átomos são consideradas constantes e caso este seja muito grande, levará ao aumento do erro associado, além disso poderão surgir instabilidades na simulação. 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, que se aproxima a frequência da vibração dos átomos da molécula de água (FRENKEL; SMIT, 2001).

3.1.4 Ensemble NPT

Simulações de DM, por padrão, geram trajetórias em uma amostra 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) 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)

Como a temperatura de um sistema é correlacionada com a energia cinética de suas partículas, pode-se controlar a temperatura modificando a energia cinética de tais partículas, similar ao acoplamento do sistema a um banho térmico. Desta maneira o termostato de Langevin possibilita a geração de *ensembles* isotérmicos. Para que a temperatura escolhida seja mantida são adicionados um termo dissipador γp_i que retira energia cinética de cada átomo i e um termo que fornece energia cinética aos átomos. Esse último é uma força restauradora aleatória R_i equivalente a colisões com os átomos do banho. Assim as equações de movimento de cada átomo são modificadas e descritas por:

$$\ddot{r}_i = \frac{\dot{p}_i}{m_i} \quad (6)$$

$$\dot{p}_i = F_i - \gamma p_i + R_i \quad (7)$$

onde γ é o coeficiente de fricção de acordo com a viscosidade do meio (banho térmico).

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).

3.1.5 RMSD e RMSF

O cálculo do RMSD (**Root Mean Square Deviation**) no tempo é uma forma de avaliar a estabilidade estrutural de macromoléculas e obter um indicativo da qualidade da trajetória obtida (KUZMANIC; ZAGROVIC, 2010). Através desse método é calculado o desvio médio dos átomos em um instante da simulação $t > 0$ relativo a uma 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 através da equação:

$$RMSD(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N \|x_i(t) - x_i^{ref}\|^2} \quad (8)$$

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 flutua em torno de uma média conformacional estável, ou seja, o sistema está em equilíbrio, pode-se avaliar as flutuações atômicas em relação a posição atômica média, o RMSF (do inglês, **Root Mean Square Fluctuation**). Dessa forma é possível obter uma melhor representação quantitativa das flutuações locais do sistema. O RMSF é calculado pela seguinte equação:

$$RMSF = \sqrt{\frac{1}{T} \sum_{t_j=1}^T (x_i(t_j) - \bar{x}_i)^2} \quad (9)$$

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 .

3.2 Docking molecular

O *docking* ou ancoragem molecular é uma técnica de modelagem molecular utilizada para prever a estrutura mais favorável do complexo formado por duas ou mais moléculas. Duas etapas podem ser caracterizadas nos estudos de *docking*: (1) busca exaustiva das orientações e conformações possíveis do ligante 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 na busca está correlacionado com a sua demanda computacional. (2) Classificação das soluções encontradas de acordo com a estimativa da afinidade calculada pela *scoring function* do software. Existem diversos métodos para estimar a afinidade, alguns baseados em campos de força, com conceitos físicos, e outros utilizam parâmetros experimentais em bancos de dados para pesar termos da equação utilizada, conhecidas como *knowledge-based* e *empirical-based*. Porém a precisão da estimativa da afinidade ainda é um dos grandes desafios das *scoring functions* (DHANIK; KAVRAKI, 2012)

O *Auto Dock Vina* (TROTT; OLSON, 2010) é o sucessor do *AutoDock 4*, e embora tenha algumas das ideias do seu antecessor, incorporou uma grande variedade de elementos e aprimoramentos que o tornaram um dos programas mais citados de *docking*. Além de ser paralelizável, o que o transformou em aproximadamente 62 vezes mais rápido do que o *AutoDock 4*, vários fatores contribuíram para seu sucesso, como uma evolução na precisão dos modos de ligação, facilidade de uso, disponibilidade gratuita e alta performance (SOUSA *et al.*, 2013).

O algoritmo de busca utilizado pelo *Vina* é o *Iterated Local Search global optimizer*, no qual são realizadas mutações e otimizações locais diversas vezes na procura da melhor orientação e conformação do ligante. Já a *scoring function*

empregada no programa alia potenciais *knowledge-based* e funções empíricas para estimar as afinidades, porém a contribuição eletrostática não é contemplada. A pontuação das poses S encontradas pelo algoritmo de busca é dependente das interações intermoleculares c , as quais são calculadas entre cada par atômico formado pelo complexo ligante e receptor. São aplicados pesos no cálculo de c baseados em parâmetros empíricos, como o choque estérico entre o par e, quando aplicável, interações hidrofóbicas e ligações de hidrogênio também são consideradas. Assim

$$S = \frac{c}{1 + \omega N_{rot}} \quad (10)$$

na qual ω é o peso associado ao número de torções N_{rot} entre os átomos pesados.

O *Ensemble docking* é uma abordagem desenvolvida para lidar com a flexibilidade do receptor, na qual a busca de 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). Deste modo, a análise de *docking* se torna mais fiel a realidade biológica.

3.3 Avaliação da Energia de ligação – LIE (*Linear Interaction Energy*)

O método LIE (AQVIST; MEDINA; SAMUELSSON, 1994) assume que a energia livre de ligação entre moléculas pode ser dividida em contribuições polares e não polares, as quais escalam linearmente de acordo com as médias termodinâmicas, obtidas pelas simulações de DM, das interações eletrostáticas e de van der Waals. Para o cálculo das médias termodinâmicas, considera-se apenas as simulações de DM do ligante complexado com a proteína no sítio de ligação estudado, e do ligante livre em solução. Assim através da diferença nos potenciais médios das interações eletrostáticas e de van der Waals do ligante, nos estados ligado e livre, a energia de ligação $\Delta G_{ligação}^o$ é estimada:

$$\Delta G_{\text{ligação}}^{\circ} = \alpha \left(\langle V^{\text{vdW}} \rangle_{\text{complexo}} - \langle V^{\text{vdW}} \rangle_{\text{livre}} \right) + \beta \left(\langle V^{\text{elect}} \rangle_{\text{complexo}} - \langle V^{\text{elect}} \rangle_{\text{livre}} \right) + \gamma \quad (11)$$

onde, $\langle V^{\text{vdW}} \rangle$ e $\langle V^{\text{elect}} \rangle$ são médias termodinâmicas dos potenciais de interações de van der Waals e eletrostáticas do ligante, quando este estava ligado com o receptor ou livre em solução.

Os potenciais de interações não ligadas V^X são calculados diretamente das configurações microscópicas dos sistemas gerados através das simulações de DM. Os parâmetros empíricos α e β pesam as contribuições polares e não polares respectivamente e γ é um parâmetro relacionado a solvatação, frequentemente utilizado para ajustar a equação de acordo com resultados de energia livre de ligação de moléculas similares obtidos de dados experimentais.

Uma vez encontrada a energia livre de ligação, a constante de ligação/constante de dissociação pode ser calculada como:

$$K = \frac{e^{\frac{-\Delta G}{k_b T}}}{C^{\circ}} \quad (12)$$

k_b , T e C° são respectivamente a constante de *Boltzmann*, temperatura e concentração padrão.

3.4 Protocolos

3.4.1 Protocolo de simulações de DM

As coordenadas iniciais do nucleossomo foram obtidas no PDB, código 1KX5, com resolução 1.94 Å, (DAVEY *et al.*, 2002). Foi utilizado o colesterol presente na estrutura cristalográfica do complexo entre o receptor nuclear hormonal órfão alfa (ROR alfa) e o colesterol, com resolução de 1,63 Å, código PDB 1N83 (KALLEN *et al.*, 2002).

As simulações de DM foram realizadas com o programa NAMD (versões 2.9 e 2.10), 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 possui parâmetros para proteínas e ácidos nucleicos (HUANG; MACKERELL, 2013), em conjunto com a atualização do parâmetro do colesterol (LIM; ROGASKI; KLAUDA, 2012).

Todos os sistemas foram simulados empregando amostras 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 consideradas rígidas pelo algoritmo SHAKE. Os sistemas foram solvatados em uma caixa contendo uma solução aquosa com 150 mM de NaCl. A água é descrita pelo modelo TIP3 (JORGENSEN *et al.*, 1983). O tamanho da caixa variou entre os sistemas, mas foi utilizada uma distância mínima de 10 Å entre as extremidades da proteína e da caixa. Considerou-se ainda condições periódicas de contorno (PBC) nos eixos x, y e z.

Com o intuito de evitar deformações estruturais no nucleossomo no início das simulações, seguiu-se um protocolo de *annealing*. Primeiramente, para que o banho de eletrólitos se equilibrassem, foi imposto potencial harmônico de 1 kcal/mol contrários ao movimentos dos átomos do nucleossomo. Posteriormente, os potências harmônicos foram removidos sucessiva e gradualmente (de 1 a 0,1 a 0,01 kcal/mol) da cadeia lateral até a cadeia principal dos sistemas. O tempo e o padrão de remoção do potencial harmônico diferiu entre os sistemas, os detalhes serão fornecidos abaixo.

As simulações pertinentes foram conduzidas em supercomputadores de médio porte, disponibilizados pelo Laboratório de Biologia Teórica e Computacional 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 200 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 são utilizados tanto códigos de análise gerados no LBTC, quanto o VMD (*Visual Molecular Dynamics*) (HUMPHREY; DALKE; SCHULTEN, 1996), um software de análise gráfica 3D. Este último, também foi utilizado para realizar a sobreposição estrutural antecedente ao cálculo do RMSD.

Segue os detalhes dos sistemas estudados:

3.4.1.1 Controle

Utilizando a estrutura cristalográfica do nucleossomo acima citado, foi construído um sistema composto pelo nucleossomo hidratado. Esse sistema foi solvatado em uma caixa $152 \times 225 \times 100 \text{ \AA}^3$, totalizando 229.882 átomos. Desses, 192 íons sobreposição estruturalcloro, 339 íons sódio e 68.088 moléculas de água. Foram aplicadas restrições harmônicas nos átomos da cadeia principal durante 0,3 ns e em seguida equilibrado por 220 ns.

3.4.1.2 Colesterol livre em solução

Para avaliar a dinâmica do colesterol em água exigida para o cálculo de energia livre pelo LIE, foi criado um sistema com apenas uma molécula de colesterol. Esse sistema foi solvatado em uma caixa $55 \times 48 \times 45 \text{ \AA}^3$ e possui 11.558 átomos com 3.868 moléculas de água. Foram aplicadas restrições harmônicas apenas no colesterol por 2,41 ns.

3.4.1.3 Pós docking (PD)

A partir dos resultados de *docking* e posterior análise dos sítios encontrados, foram criados dois sistemas independentes contendo colesterol e nucleossomo. Em cada um deles, três moléculas de colesterol estavam ligadas ao nucleossomo já equilibrado por 18,75 ns (sistema controle inicial). Esta divisão foi realizada para evitar interações entre as moléculas de colesterol simuladas, assim a menor distância inicial entre os ligantes foi de 31 Å. As coordenadas das moléculas de colesterol, foram obtidas a partir do centro da nuvem de soluções dos respectivos sítios. Esses sistemas foram denominados: pós docking 1 (PD1, com o colesterol nos sítios 2, 3 e

5) e pós docking 2 (PD2, com o colesterol nos sítios 1, 4 e 6). A fim de evitar sobreposição de átomos, as moléculas de água que estavam a 2 Å de cada molécula de colesterol foram removidas. Foram aplicadas restrições harmônicas, por aproximadamente 7 ns no colesterol e 5 ns no nucleossomo. O padrão de retirada dos potenciais harmônicos no nucleossomo foi: DNA → Aminoácidos a 8 Å de qualquer molécula de colesterol (cadeia lateral e depois a cadeia principal) → Aminoácidos a 3 Å de qualquer molécula de colesterol (cadeia lateral e depois a cadeia principal). Já no colesterol, após 5,5 ns de simulação o potencial harmônico começou a ser reduzido. Dessa forma, as cadeias laterais do nucleossomo tiveram mais de 1,5 ns para acomodar-se ao colesterol com a movimentação restringida.

3.4.2 Protocolo de *Docking* Molecular

A fim de encontrar potenciais sítios de ligação de colesterol ao nucleossomo, foi utilizado o programa Auto Dock Vina (TROTT; OLSON, 2010). Em geral, o *docking* molecular considera o receptor como uma estrutura fixa. Assim para se considerar a flexibilidade inerente de moléculas biológicas em solução, são fornecidas ao programa estruturas em diferentes conformações do receptor. Tais conformações foram obtidas aleatoriamente, em um espaço de 6 ns (12,75 - 18,75 ns), da simulação de equilíbrio do nucleossomo em solução. No total 120 configurações independentes (*frames*) do nucleossomo foram selecionadas, proporcionando uma amostragem mais adequada do espaço de configurações do receptor. Também foi permitida a flexibilidade das ligações do colesterol. Os átomos de hidrogênio que faltavam ao colesterol foram adicionados usando o módulo PSFGEN do NAMD. Os parâmetros foram deixados no padrão do programa, com exceção apenas da exaustividade que fora 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).

O espaço de busca de soluções no programa é chamado de *grid*. O *grid* utilizado foi uma caixa de 110 x 110 x 110 Å³, englobando o nucleossomo e apenas as caudas das histonas próximas ao DNA. Para cada estrutura utilizada como receptor do *docking*, as 20 melhores soluções classificadas pelo software foram seleciona-

das. Como resultado, 2.400 formas e interfaces de encaixe do colesterol no nucleossomo foram encontradas. Esses 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, outros valores foram testados, entretanto 10 Å foi o que melhor distribuiu as nuvens de soluções sem perda de sítios de ligação e sobreposição de resultados.

3.4.3 Protocolo para o cálculo do LIE:

Os parâmetros empíricos utilizados para o cálculo de energia livre de ligação pelo método LIE foram: $\alpha=0.18$, $\beta=0.5$. O termo corretivo γ não foi utilizado nos cálculos ($\gamma=0$). As médias termodinâmicas das interações de van der Waals $\langle V^{vdW} \rangle_{complexo}$ e interações eletrostáticas $\langle V^{elect} \rangle_{complexo}$ para o complexo foram computados da simulação com o colesterol ligado ao nucleossomo. Já as médias termodinâmicas do colesterol livre, $\langle V^{vdW} \rangle_{livre}$ e $\langle V^{elect} \rangle_{livre}$, foram obtidas através da simulação do colesterol sozinho em solução. Ambos os sistemas foram simulados por 200 ns.

4 RESULTADOS

4.1 Estabilidade estrutural do *ensemble* de equilíbrio

Foi criado um sistema formado pelo nucleossomo hidratado explicitamente por moléculas de água, conforme descrito na metodologia (ver tópico 3.4.1.1). A simulação de DM desse sistema compreendeu cerca de 20 ns e foi utilizada para considerar a mobilidade das cadeias laterais dos aminoácidos e ácidos nucleicos do nucleossomo na etapa do *docking*. Porém, antes de prosseguir para a investigação da interação do colesterol, a estabilidade conformacional das estruturas geradas foi avaliada.

A forma usual de avaliar a estabilidade de simulações de DM é através do RMSD (ver tópico 3.1.5). Antes de computar os desvios entre os átomos, as estruturas da trajetória devem ser sobrepostas, a fim de minimizar o erro associado ao resultado do RMSD. Tal sobreposição estrutural entre os *frames* deve ser realizada baseada em uma seleção de átomos que, ao longo do tempo, apresente flutuações mínimas. Entretanto, esta deve englobar uma quantidade suficiente de átomos para garantir que a matriz de rotação encontrada seja capaz de rotacionar corretamente todo o sistema. Por conseguinte, para avaliar qual seleção de átomos seria a melhor escolha para a sobreposição estrutural, foi computado o RMSD de cinco conjuntos de átomos diferentes. Foram selecionados somente os átomos da cadeia principal e estes foram sobrepostos na própria seleção. Os conjuntos atômicos investigados foram: (1) átomos do NCP; (2) átomos do NCP, exceto as caudas das histonas; (3) átomos das histonas; (4) átomos do domínio central globular das histonas; (5) átomos das hélices das histonas (Figura 9a).

Podemos observar que ao selecionar os átomos da cauda das histonas, os valores do RMSD aumentaram muito, o que era esperado, pois essa é uma região conhecidamente desestruturada e flexível. O DNA também contribuiu para o aumento das flutuações atômicas, como pôde ser constatado no perfil de RMSD do nucleossomo e das histonas, ambos sem as caudas das histonas. Dois conjuntos de átomos apresentaram perfil de RMSD pequeno, por volta de 1 Å. A desigualdade entre essas

seleções foi somente as alças das histonas, portanto esta é a responsável pela discreta diferença entre as curvas. Em vista disso, as hélices das histonas foram utilizadas para a sobreposição entre estruturas anteriores aos cálculos de RMSD e RMSF.

Empregando a sobreposição estrutural descrita acima, o RMSD dos átomos pesados do nucleossomo, excluindo as caudas das histonas, converge para um platô após 10 ns, com valores próximos a 2 Å, indicativo de estabilidade estrutural dessa trajetória (NCP sem caudas). Excluindo também o DNA e as alças das histonas (Hélices), o perfil do RMSD apresentado é ainda mais reduzido, um pouco acima de 1 Å. Também foram avaliadas outras regiões presentes no nucleossomo, como o domínio de *docking* e os 4-*helix bundle* H3:H3 e H2B:H4. É possível observar que nos 20 ns simulados, o perfil de RMSD dos dois 4-*helix bundles* flutua muito pouco, especialmente na região formada pelas histonas H2B e H4. O domínio de *docking* por sua vez sofre oscilações um pouco mais pronunciadas, mas mesmo assim não ultrapassa 3 Å de desvio. (Figura 9b).

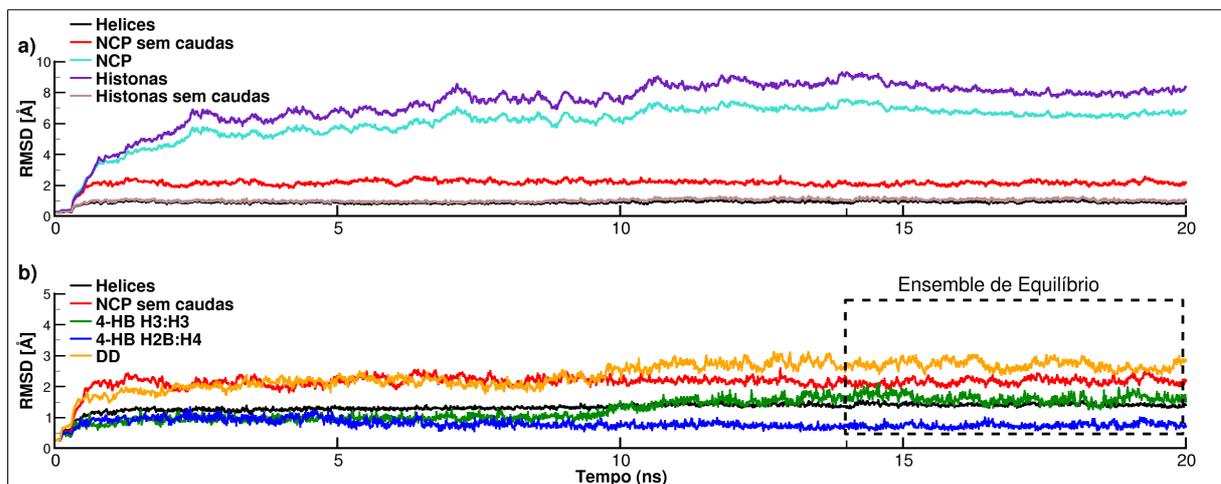
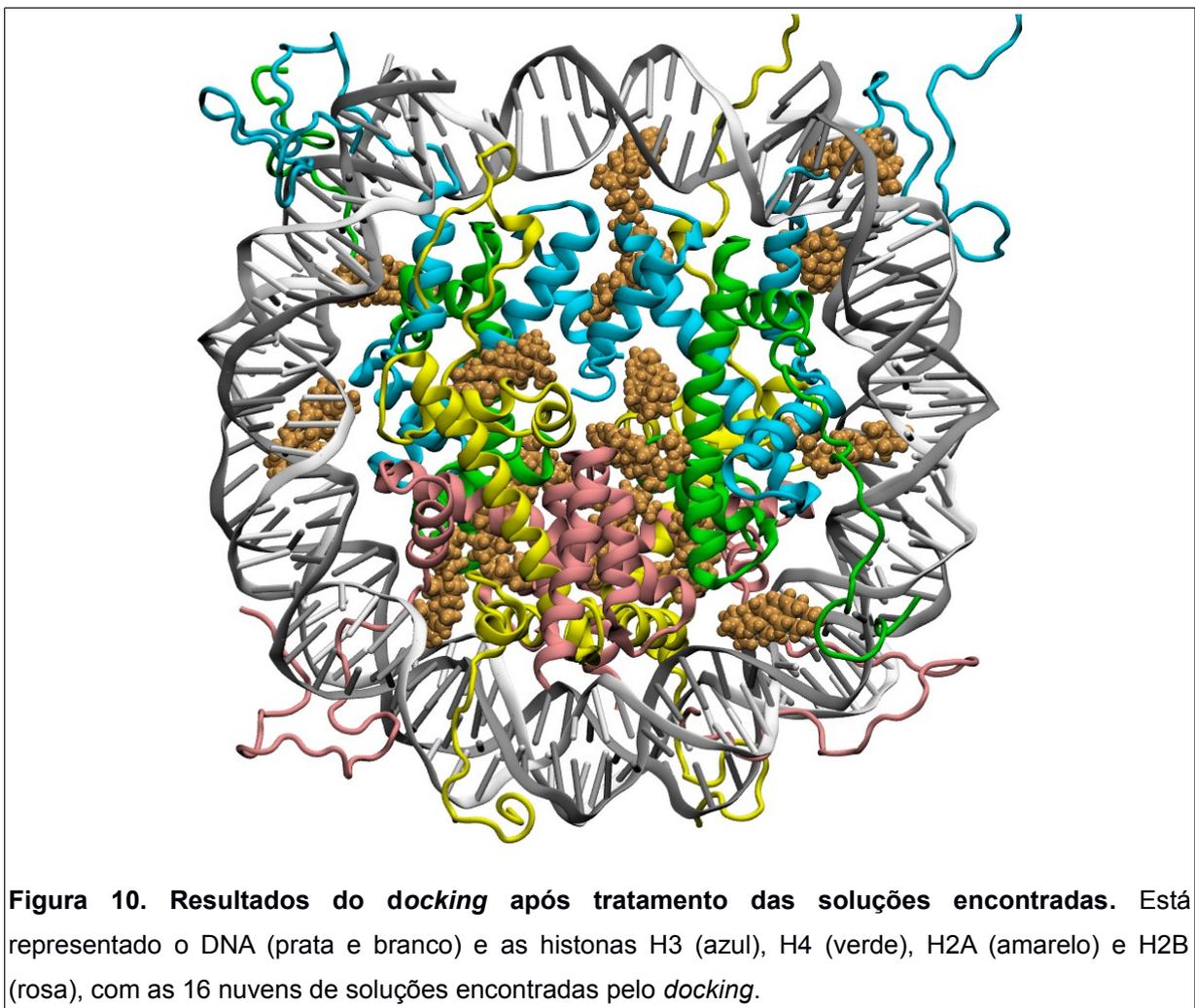


Figura 9. Perfil do RMSD para o sistema do nucleossomo simulado em água por 20 ns. a) Prováveis seleções atômicas para a sobreposição estrutural entre os *frames*. Em todas as seleções apenas os átomos da cadeia principal foram considerados. **b)** Verificação da estabilidade do sistema em busca de um *ensemble* de equilíbrio (destacado por um quadrado) para as análises de *docking* molecular. Foram selecionados apenas os átomos pesados, ou seja, exclui-se os átomos de hidrogênio. 4-HB → 4-*helix bundle*, DD → domínio de *docking*.

4.2 Sítios de interação do colesterol no nucleossomo

Após a constatação do equilíbrio estrutural atingido pelo *ensemble* do nucleossomo, foi possível avançar para a investigação da ligação do colesterol ao nucleossomo. Para este fim, foram realizados experimentos de *docking* molecular (ver tópico 3.4.2). Aproximadamente 2400 poses do colesterol foram encontradas e agrupadas em 47 nuvens de soluções do colesterol. Dentre estas, 31 nuvens estavam nos sulcos maiores e menores do DNA, essas soluções apresentaram apenas interações com o DNA e não foram analisadas nesse estudo. As 16 nuvens remanescentes formaram os possíveis sítios de ligação do colesterol na interface proteica do nucleossomo. A Figura 10 traz os sítios encontrados e a Tabela 1 traz a média da afinidade calculada pelo programa para cada sítio.



Nota-se que a superfície do nucleossomo exibe várias cavidades para a ligação do colesterol. Os dímeros apresentaram maior densidade de soluções do que o tetrâmero. Ademais, em alguns resultados do *docking* o colesterol foi detectado bem enterrado no nucleossomo ou próximo de regiões importantes para a sua estabilidade estrutural, como os *4-helix bundles* H3:H3 e H2B:H4, além do domínio de *docking*. Destacam-se ainda soluções encontradas nas caudas das histonas, principalmente a cauda da H4. Entretanto, não foram observadas soluções inseridas no *patch* ácido, apenas próximas dessa interface.

Tabela 1. Energia livre de ligação média calculadas pela *scoring function* para cada sítio de ligação do colesterol.

Sítios	Energia (kcal.mol ⁻¹)	Desvio (kcal.mol ⁻¹)
1	-8.7	0.1
2	-8.1	0.3
3	-7.0	0.1
4	-7.0	0.1
5	-7.9	0.2
6	-7.5	0.2
7	-7.8	0.3
8	-7.8	0.3
9	-8.0	0.4
10	-8.1	0.6
11	-7.6	0.4
12	-8.1	0.3
13	-8.0	0.4
14	-7.7	0.3
15	-7.7	0.3
16	-7.6	0.3

As energias livre de ligação estimadas pelo programa para cada sítio variaram entre -8.7 a -7.0 kcal/mol. Os desvios entre as soluções foram pequenos, não ultrapassando 0.6 kcal/mol. Como pode ser observado na Tabela 1, a diferença entre as afinidades das soluções encontradas não foi grande. Entretanto, essa estimativa é obtida por meio de várias simplificações e ajustes, sendo assim questionável. Para

aprimorar a predição de afinidades foi utilizado o método LIE (*Linear Interaction Energy*).

4.2.1 Energia livre de ligação

Para garantir uma boa amostragem das energias utilizadas pelo LIE, trajetórias de DM longas devem ser produzidas. Por conseguinte, as simulações com o colesterol ligado ao nucleossomo totalizaram cerca de 200 ns de trajetória de equilíbrio. Entretanto, apenas os 160 ns finais dessa trajetória foram empregados no cálculo, isso ocorreu devido a dinâmica observada pelo colesterol em alguns sítios (descrito posteriormente). Seis moléculas de colesterol que interagem com regiões sabidamente importantes para a estabilidade do nucleossomo, foram escolhidas para a amostragem do complexo. Os sítios escolhidos foram numeradas de fora do nucleossomo para dentro, sendo assim o colesterol 1 se encontra mais distante do centro do nucleossomo do que o colesterol 2. Os sítios selecionados abrangem os principais achados do *docking*, como a interação do colesterol com: a cauda N-terminal da H4 (sítios 1 e 2), os *4-helix bundles* (sítios 4, 3 e 6), o domínio de *docking* (sítio 5) e o *patch* ácido (sítios 3 e 6).

Os potenciais de interação eletrostática e de vdW obtidos através da simulação do colesterol com o nucleossomo podem ser observados na Figura 11. As médias termodinâmicas desses potenciais, necessárias para o cálculo de energia livre pelo método LIE, estão descritas na Tabela 2. Em todos os sítios, os valores de vdW foram mais negativos do que os valores obtidos para o potencial eletrostático, além disso, o potencial de vdW também apresentou maior homogeneidade entre os sítios, com variação proporcional de 12,55%, enquanto que o potencial eletrostático variou 16,80 % entre os sítios. Especificamente em relação ao potencial de interação de vdW, os sítios presentes na cauda da H4 (1 e 2) apresentaram valores mais negativos, enquanto que os sítios mais superficiais exibiram os potenciais menos negativos. Esse potencial para o colesterol livre em água representa 2/3 da média observada ($-50,31 \pm 2,67$) quando o colesterol está em complexo com o nucleossomo. As afinidades estimadas entre os ligantes são heterogêneas e variam

entre 0,01 – 0,2 nM⁻¹. Em especial, o ligante no sítio 5 apresentou valor mais favorável a formação do complexo.

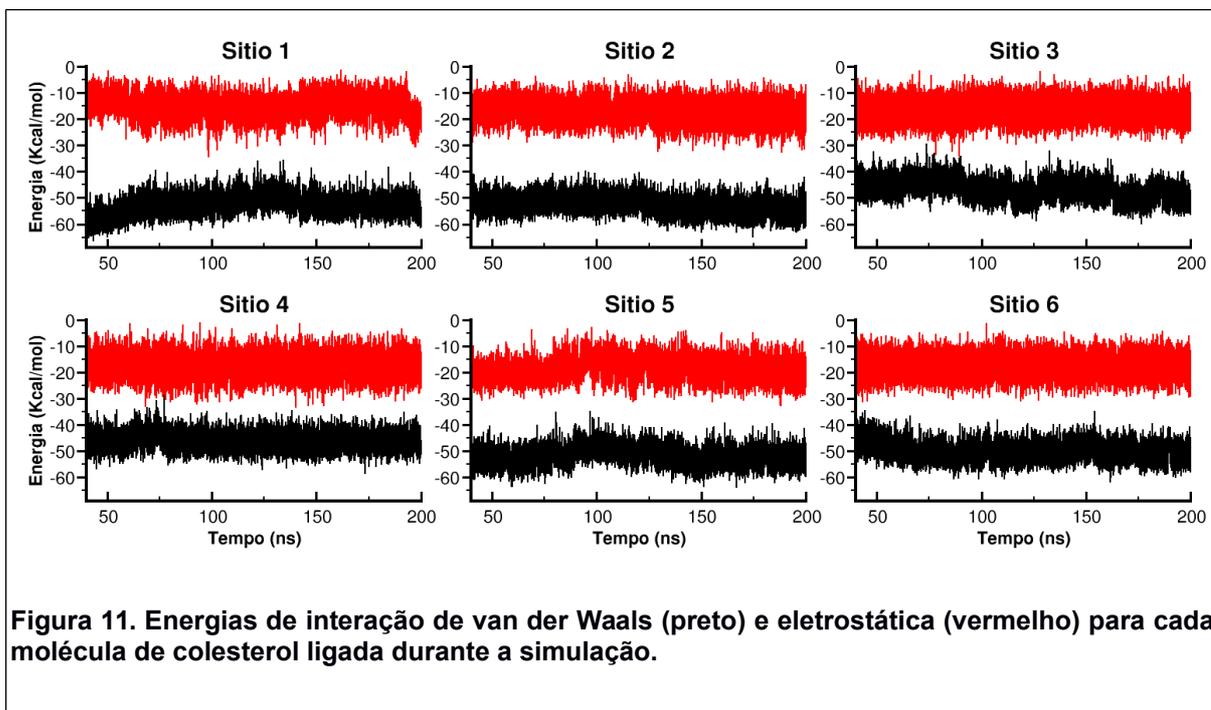


Tabela 2. Resultados obtidos para as energias de interação do colesterol no nucleossomo através do LIE em cada sítio. São apresentadas as energias livres de ligação absoluta $\Delta G_{ligação}^o$, a constante de ligação K e as médias dos potenciais de van der Waals $\langle V^{vdW} \rangle$ e eletrostática $\langle V^{elec} \rangle$.

Colesterol	ΔG^o (kcal.mol ⁻¹)	K (mM ⁻¹)	$\langle V^{elec} \rangle$ (kcal.mol ⁻¹)	$\langle V^{vdW} \rangle$ (kcal.mol ⁻¹)
Sítio 1	-1,75 ± 0,18	0,0188	-15,20 ± 4,26	- 52,97 ± 3,69
Sítio 2	-2,78 ± 0,16	0,1060	-17,24 ± 3,95	- 53,05 ± 3,34
Sítio 3	-1,29 ± 0,17	0,0087	-16,36 ± 4,06	- 47,20 ± 3,79
Sítio 4	-1,66 ± 0,17	0,0162	-17,38 ± 4,28	- 46,39 ± 3,21
Sítio 5	-3,13 ± 0,16	0,1906	-18,27 ± 4,02	- 52,07 ± 3,47
Sítio 6	-2,43 ± 0,16	0,0589	-17,57 ± 4,02	- 50,15 ± 3,39
Livre	_____	_____	-17,83 ± 4,11	- 35,94 ± 2,9

4.3 Dinâmica da interação colesterol-nucleossomo

As afinidades estimadas pelo LIE não condizem com uma interação forte entre o colesterol e o nucleossomo. Portanto, é provável que o colesterol apresente certa flexibilidade nos sítios, exceto naqueles com restrição estérica. Assim, a dinâmica do colesterol quando ligado ao nucleossomo foi investigada. Ainda se deram análises comparativas da estrutura do nucleossomo com e sem o ligante. Para isso, o sistema inicial apenas do nucleossomo em solução, equilibrado durante ~ 20 ns e utilizado no *docking*, foi estendido por aproximadamente 200 ns. Esse sistema se tornou o controle para as análises a seguir.

4.3.1 Mapa de contatos

Antes de prosseguir com as análises comparativas, foi definido quais eram os aminoácidos formadores dos sítios de interação de cada molécula de colesterol simulada. Para isto, foi criado um código de análise para definir os sítios de contato. Esse código calculou a probabilidade ao longo do tempo de cada átomo do nucleossomo estar a no máximo 3,5 Å de qualquer átomo do colesterol. Ao final dessa análise, foi obtida uma probabilidade média de contato com o colesterol para cada resíduo. Esses resultados permitiram a criação de um mapa de contato para cada molécula de colesterol. Foram considerados como pertencentes a determinado sítio, apenas aqueles resíduos com probabilidade maior do que 20 %, ou seja mais do que 40 ns dentro do raio de corte.

Com o mapa de contato completo, foram analisados quais aminoácidos do nucleossomo que formavam os sítios de interação com o colesterol já haviam sido citados na literatura por alguma particularidade. Especificamente buscou-se aminoácidos conhecidos por formarem regiões importantes para a estabilidade estrutural do nucleossomo e resíduos essenciais para a interação com proteínas ligantes de nucleossomo (NBP) (Tabela 3).

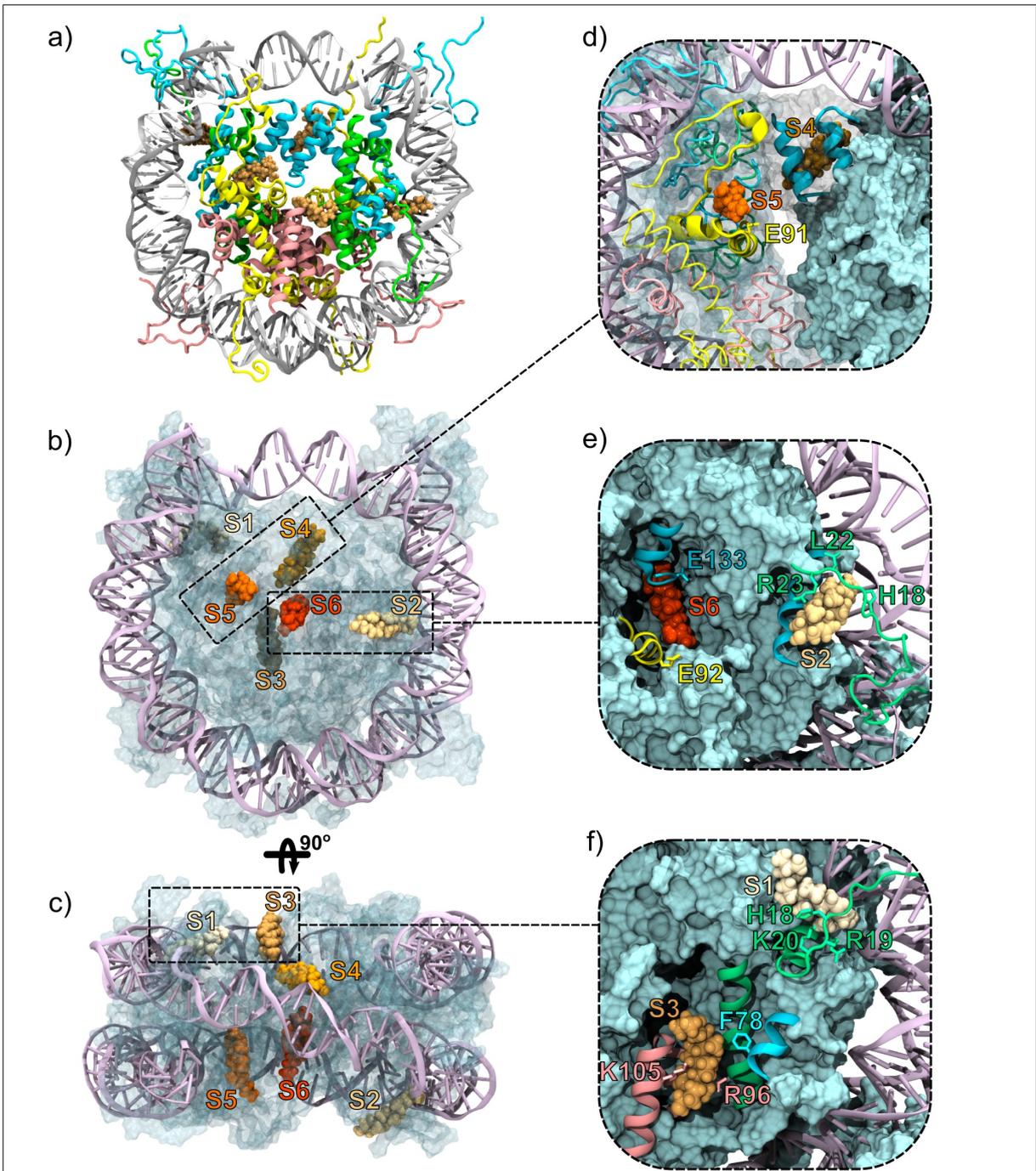


Figura 12. Moléculas de colesterol nos sítios de ligação ao final da simulação (~200 ns). a - c) As seis moléculas de colesterol nos sítios 1 ao 6 (S1 a S6). **d- f)** Detalhes da interação do colesterol nos sítios simulados. Em destaque estão regiões importantes do nucleossomo, como o 4-helix bundle H3:H3 (**d**), o domínio de docking (**d**) e as caudas N-terminais da H4 (**e**, **f**). Também estão destacados os aminoácidos do patch ácido (E91 e E92 da H2A), da cauda da H4 (H18, R19, K20, L22, R23) e os aminoácidos F78, E133 da H3 e R96, K105 da H2B, que foram relatados anteriormente interagirem com as NBPs. DNA (roxo) e histonas H3 (azul), H4 (verde), H2A (amarelo) e H2B (rosa) também estão representados.

Tabela 3. Aminoácidos formadores das cavidades do colesterol. * Em negrito estão os aminoácidos envolvidos em interação com as NBPs. ° Aminoácidos 80 a 119 fazem parte do domínio de *docking*.

Sítios	H3*	H4*	H2A*°	H2B*
1	_____	H18, R19, K20, Q27	_____	_____
2	L65, Q68, R69, R72, Q76	H18, L22, R23	_____	_____
3	F78	R67	_____	R96, G101, A104, K105, V108
4	I112, H113, Q125, R128, R129, R134	_____	V114	_____
5	A98, V101, A102	_____	R88, E91, N94, G98, T101, G106, V107, L108, N110	_____
6	G132, E133	Q93, G94, R95	E92, K95, L96	P100

No sítio 1 os resíduos K16, H18, R19, K20, R23 – H4 além de fazerem parte da cauda da H4 que realiza contato com o *patch* ácido do nucleossomo adjacente, também foram descritos como importantes para a interação da NBP SIR3 com o nucleossomo. O sítio 2 possui três resíduos da cauda da H4 (H18, R19, K20 – H4), a L22 também faz parte da interface de interação do SIR3 com o nucleossomo. Fazem parte do sítio 3 aminoácidos que interagem com o SIR3 (F78 – H3; R96, S109 – H2B), além disso a K105 – H2B foi reportada como um dos resíduos que auxiliam na ligação da PRC1. No sítio 4, foram identificados L109 e H113 – H3 que compõem o “4 helix bundle” H3:H3. Na cavidade 5 temos vários resíduos do domínio de *docking*, Q84, R88, N94, K95, G98, T101, G106, V107 – H2A. O aminoácido E91 – H2A, faz parte do *patch* ácido, além de interagir com a CENP-C e HMG2. O E92 – H2A encontrado no sítio 6, por ser um dos aminoácidos formadores do *patch* ácido, está presente em várias interfaces de contato entre NBP e o nucleossomo, como o LANA, RCC1, SIR3, PRC1, HMG2 e IL-33. Nesse sítio também há o E133 – H3 que foi reportado como parte da região de interação com a CENP-C.

Analisando as características dos aminoácidos formadores das cavidades de interação do colesterol, foi observado que o sítio 5 é o mais hidrofóbico, 58 % dos resíduos são apolares, seguido pelos sítios 3 (57 %) e 6 (44 %). Os sítios 1 e 2, presentes na cauda das histonas, e o sítio 4, apresentaram um número menor de resíduos hidrofóbicos (0 %, 25 % e 29 %, respectivamente), porém, foram as regiões de interação com o colesterol que continham mais aminoácidos básicos, principalmente lisina e arginina.

4.3.2 Impacto estrutural do colesterol no nucleossomo

Após 200 ns de simulação, todas as moléculas de colesterol permanecem ligadas ao nucleossomo em estado de equilíbrio. Para avaliar a estabilidade e o impacto estrutural causado pelo colesterol, foram computados o RMSD e RMSF dos sistemas na ausência e presença do colesterol. Compuseram as seleções para o cálculo do RMSD apenas os átomos pesados do nucleossomo com e sem caudas.

O RMSD sem as caudas das histonas se manteve por volta de 2 Å, com e sem colesterol, a inclusão das caudas proporcionou um aumento esperado nos desvios (Figura 13a). A Figura 13b revela que nos sítios 1 e 2 a presença do colesterol impactou significativamente a movimentação atômica dos aminoácidos formadores dessas cavidades. Em especial, o sítio 1 nos primeiros 100 ns apresentou aproximadamente o triplo do desvio sem colesterol.

Outra análise empregada, o RMSF (*Root Mean Square Flutuaction*) permite identificar melhor regiões flexíveis, assim consegue descrever impactos estruturais mais especificamente. Visto que essa análise se dá por átomo, apenas o carbono alfa foi utilizado. Além disso, foram considerados somente os átomos do domínio central globular das histonas, pois as caudas são muito flexíveis e não foram simulados sítios exclusivos no DNA. Para simplificar a visualização dos resultados foi calculado o valor médio entre as cópias das histonas (ex: H2A.1 e H2A.2).

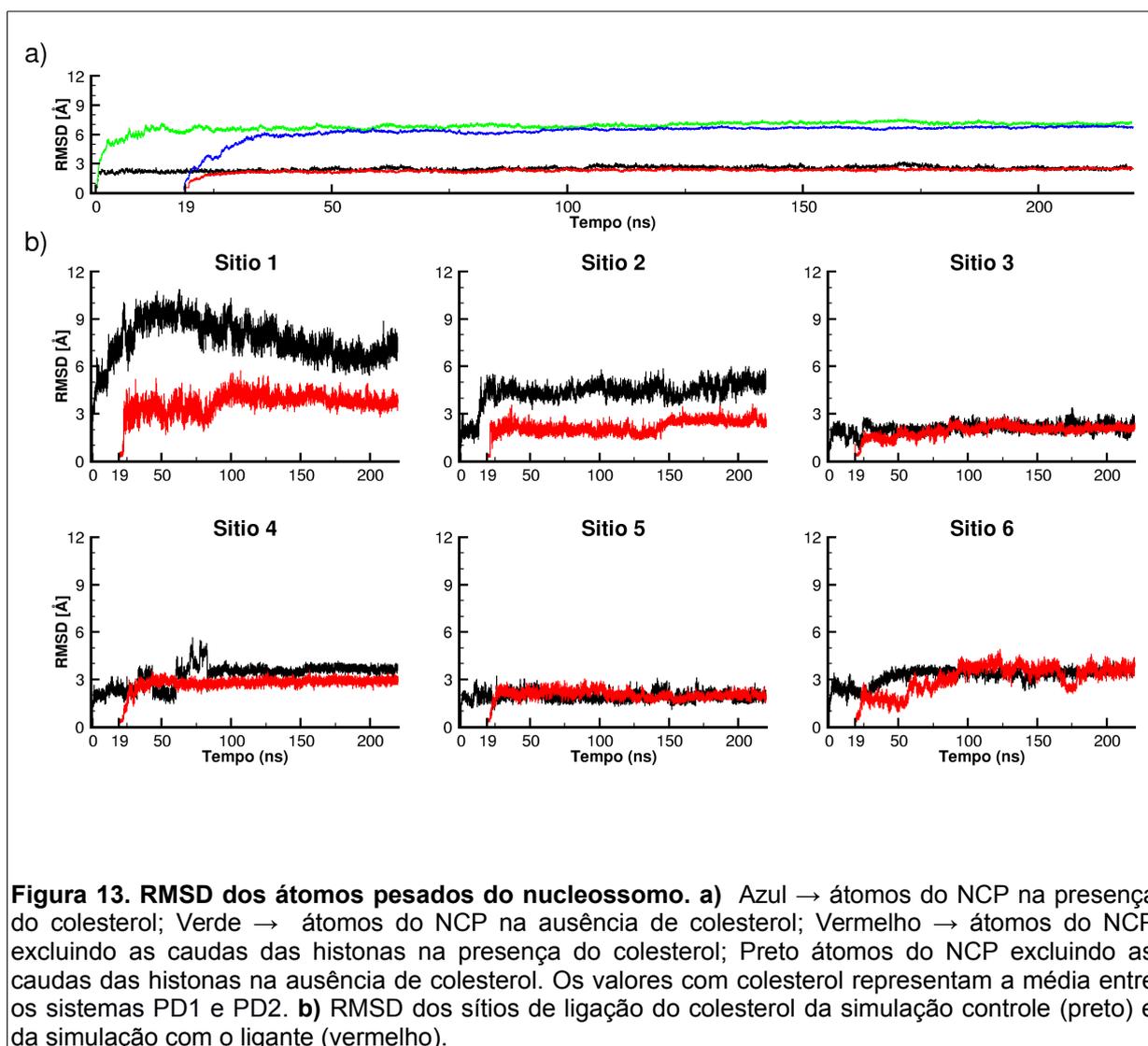


Figura 13. RMSD dos átomos pesados do nucleossomo. a) Azul → átomos do NCP na presença do colesterol; Verde → átomos do NCP na ausência de colesterol; Vermelho → átomos do NCP excluindo as caudas das histonas na presença do colesterol; Preto átomos do NCP excluindo as caudas das histonas na ausência de colesterol. Os valores com colesterol representam a média entre os sistemas PD1 e PD2. **b)** RMSD dos sítios de ligação do colesterol da simulação controle (preto) e da simulação com o ligante (vermelho).

Observando o perfil do RMSF (Figura 14), o domínio de *docking* sofre alterações notáveis. Tais alterações foram observadas nos dois sistemas com colesterol, porém o impacto maior se deu na simulação em que o colesterol estava presente no domínio de *docking* (sítio 5). Esse resultado sugere que a ação do colesterol sobre essa região não é sítio dependente. Com menor intensidade, houve também alterações nas flutuações na alça L1 da H2B, porém apenas na simulação com o colesterol nos sítios 2, 3 e 5. O sítio 5 encontra-se próximo dessa alça e o colesterol do sítio 3 está interagindo majoritariamente com a H2B, especificamente com as hélices α -C e α -3 - H2B. Provavelmente o colesterol em um desses sítios seja o responsável por essas alterações.

O RMSD (Figura 13b) e o RMSF Figura 15 dos sítios de ligação do colesterol revela que essa molécula possui um impacto nas flutuações dos aminoácidos formadores dos sítios. Especificamente, nos sítios 1 e 2, os resultados do RMSD e RMSF em conjunto sugerem que o colesterol teve um impacto maior sobre as cadeias laterais dos aminoácidos dessas regiões. Pois o RMSF, que apenas considera o carbono alfa, variou pouco em comparação com o RMSD, no qual todos os átomos pesados foram selecionados. O RMSD do sítio 4, até ~ 65 ns, demonstra oscilações observadas apenas na ausência de colesterol. Já o RMSF mostra que houve restrição pela metade da movimentação da R134 - H3. O RMSD do sítio 6 sinaliza que o colesterol estimulou uma pequena instabilidade nos aminoácidos dessa região até ~180 ns. O RMSF mostra que de fato houve alterações nas flutuações desses aminoácidos. Nos demais sítios, não foram observadas alterações significativas nas flutuações atômicas.

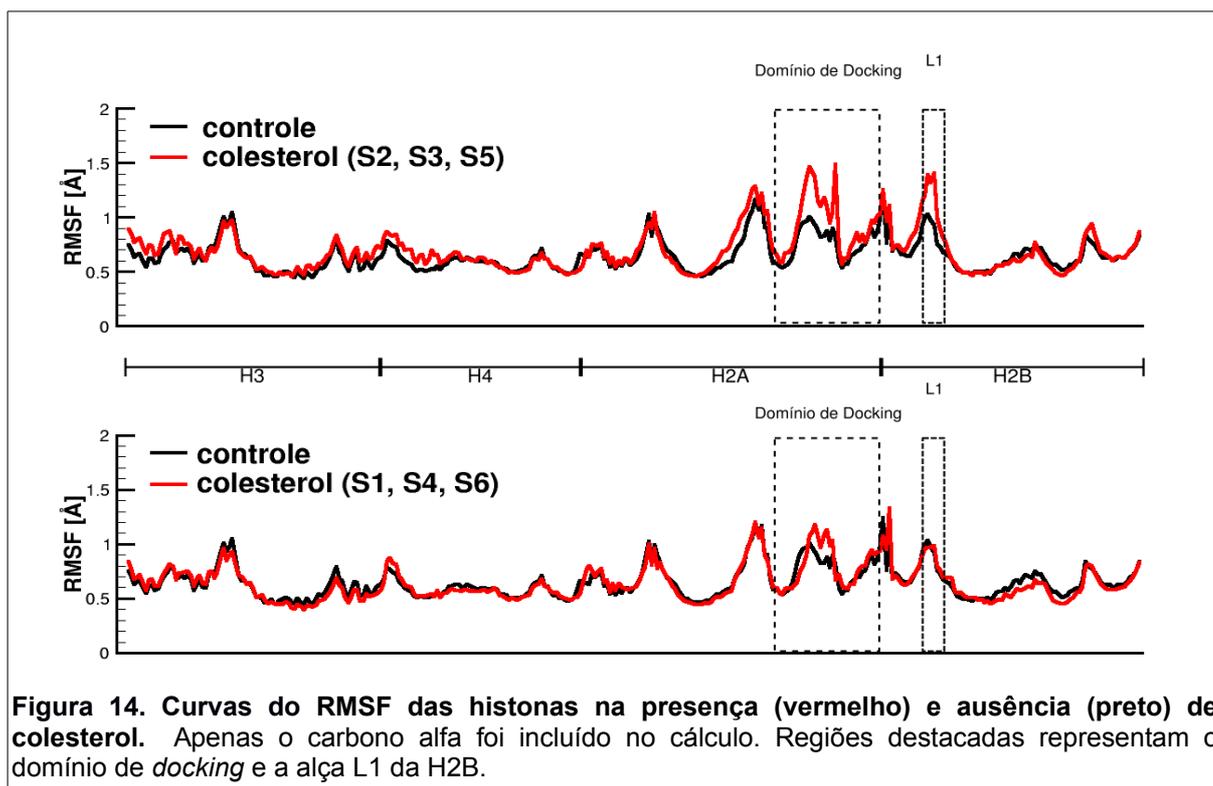


Figura 14. Curvas do RMSF das histonas na presença (vermelho) e ausência (preto) de colesterol. Apenas o carbono alfa foi incluído no cálculo. Regiões destacadas representam o domínio de *docking* e a alça L1 da H2B.

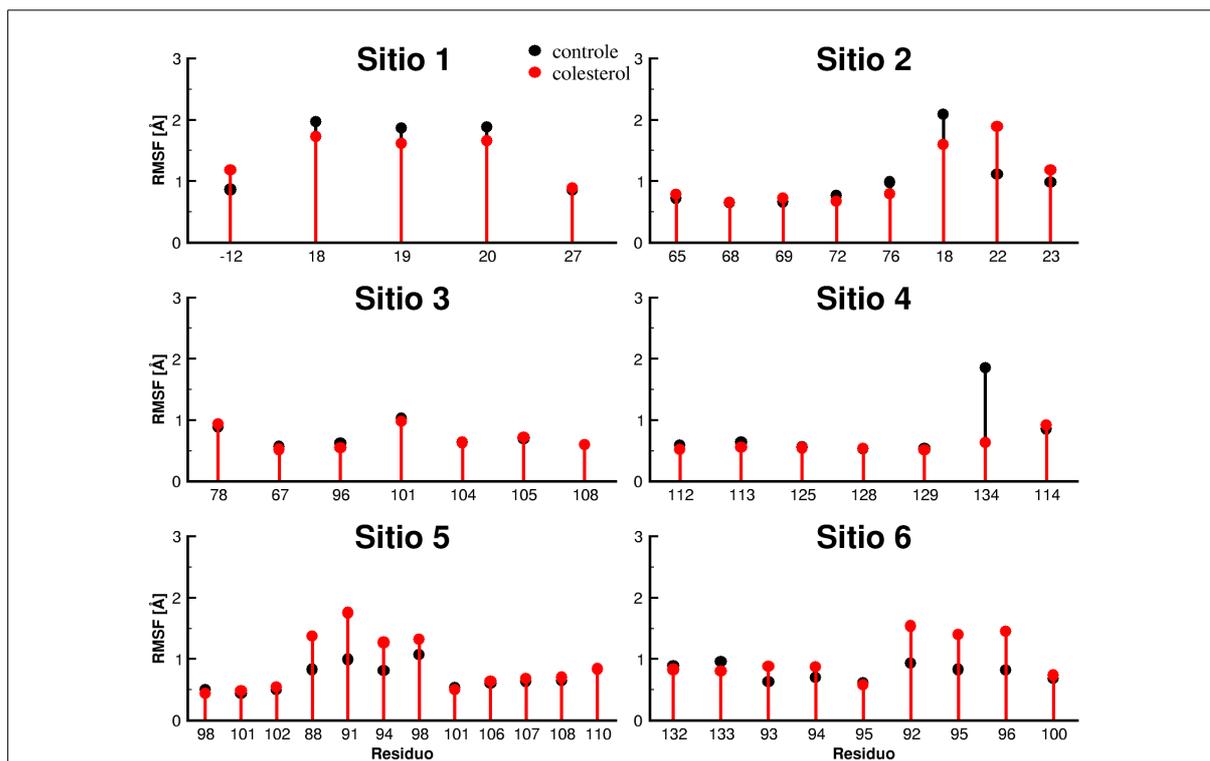


Figura 15. RMSF dos aminoácidos formadores dos sítios de ligação do colesterol. Os aminoácidos estão representados apenas pelo número na sequência, porém a ordem de apresentação é a mesma da Tabela 3. O índice -12 no sítio 1 representa um nucleotídeo (Timina 12). Apenas o carbono alfa foi incluído no cálculo.

4.3.3 Dinâmica do colesterol no sítio

Para observar o comportamento do colesterol ao longo da simulação, foi investigado a sua posição dependente do tempo em cada sítio de ligação. A posição do colesterol foi definida como o seu centro de massa. Assim, após a sobreposição entre as estruturas, foi computada a distância entre a posição inicial e a posição em cada instante da trajetória (Figura 16). Nesse gráfico, constata-se que após os primeiros 40 ns, as moléculas de colesterol presentes nos sítios 4, 5 e 6 experimentam apenas pequenas variações posicionais durante a simulação. Entretanto, foi interessante observar que o colesterol do sítio 6 apenas encontrou equilíbrio aproximadamente 10 Å distante de sua posição inicial. Analisando qualitativamente a trajetória, nesse sítio, o colesterol faz um caminho em direção ao centro do nucleossomo. O ligante nos sítios 1, 2 e 3 apresentou maiores flutuações ao longo da trajetória, porém, na primeira e na terceira cavidade, o colesterol retorna a valores estáveis após uma alta oscilação. O colesterol presente no sítio 2 flutua

aparentemente de uma posição estável para outra, entretanto, visualizando a trajetória, o ligante apenas deslocou-se um pouco para o lado no sítio. Como a molécula de colesterol que apresentou maior deslocamento (sítio 6) após 40 ns encontrava-se mais estável, o cálculo do LIE foi realizado empregando a trajetória posterior a esse período.

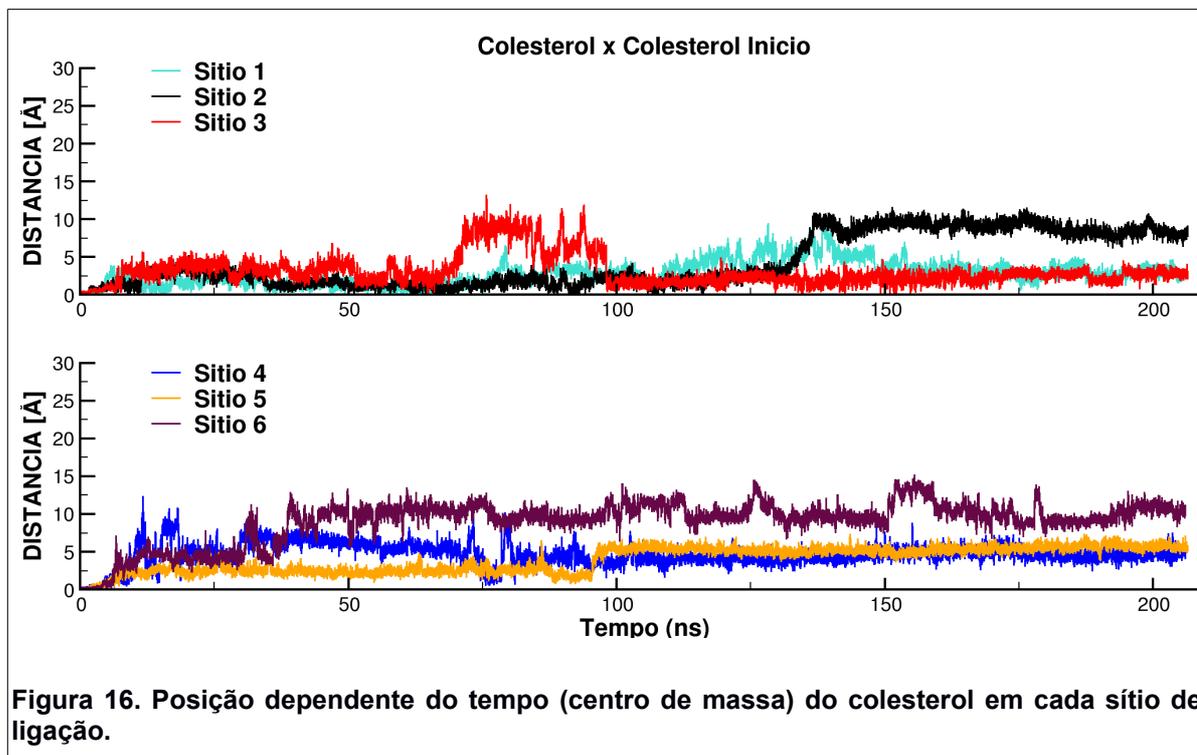


Figura 16. Posição dependente do tempo (centro de massa) do colesterol em cada sítio de ligação.

Observando os resultados da Figura 16 surge a seguinte dúvida: será que a movimentação apresentada pelo colesterol, durante a simulação, faz com que ele se afaste do nucleossomo, especificamente, dos sítios de interação? Para responder essa pergunta, foi analisada a distância ao longo do tempo entre o centro de massa do colesterol e o centro de massa do respectivo sítio. A Figura 17 demonstra que, em todos os casos, apesar de apresentar flutuações ao longo da trajetória, o colesterol não se afasta das cavidades. Como esperado, os ligantes presentes nos sítios superficiais e com menor impedimento estérico, sítios 3 e 4, apresentaram maiores oscilações na distância em relação ao centro da cavidade. O menor perfil de distância foi observado para o ligante presente no sítio 5, o mais enterrado no nucleossomo no início da simulação.

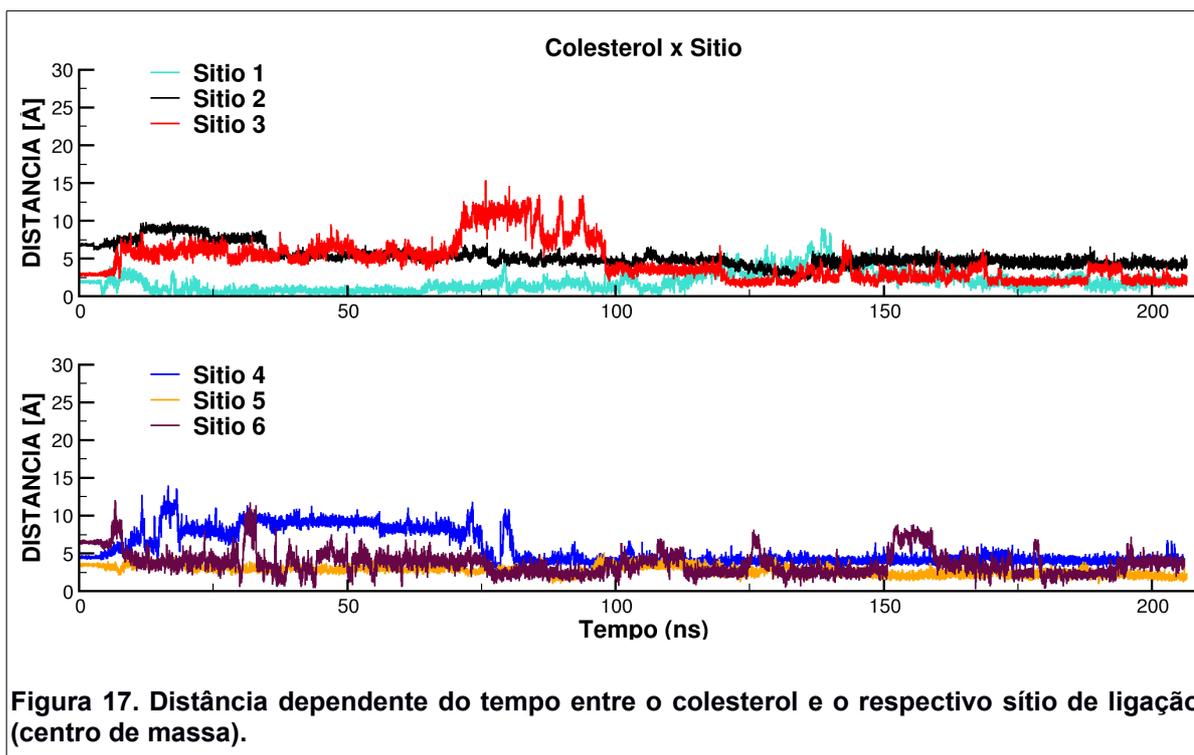


Figura 17. Distância dependente do tempo entre o colesterol e o respectivo sítio de ligação (centro de massa).

Foi investigado também, qual região do colesterol estava mais próxima das cavidades de interação. Assim, foi calculada a distância, ao longo da trajetória, do centro de massa dos sítios em relação à cauda apolar de hidrocarbonetos (Figura 18) e em relação à cabeça polar (Figura 19) do colesterol. Pelo gráfico constata-se que, na maioria dos casos, a cauda do colesterol está mais próxima do sítio do que o grupo hidroxila, indicando que sua interação com o nucleossomo seja predominantemente hidrofóbica, logo, é esperado que o ligante desloque moléculas de água nas cavidades de interação.

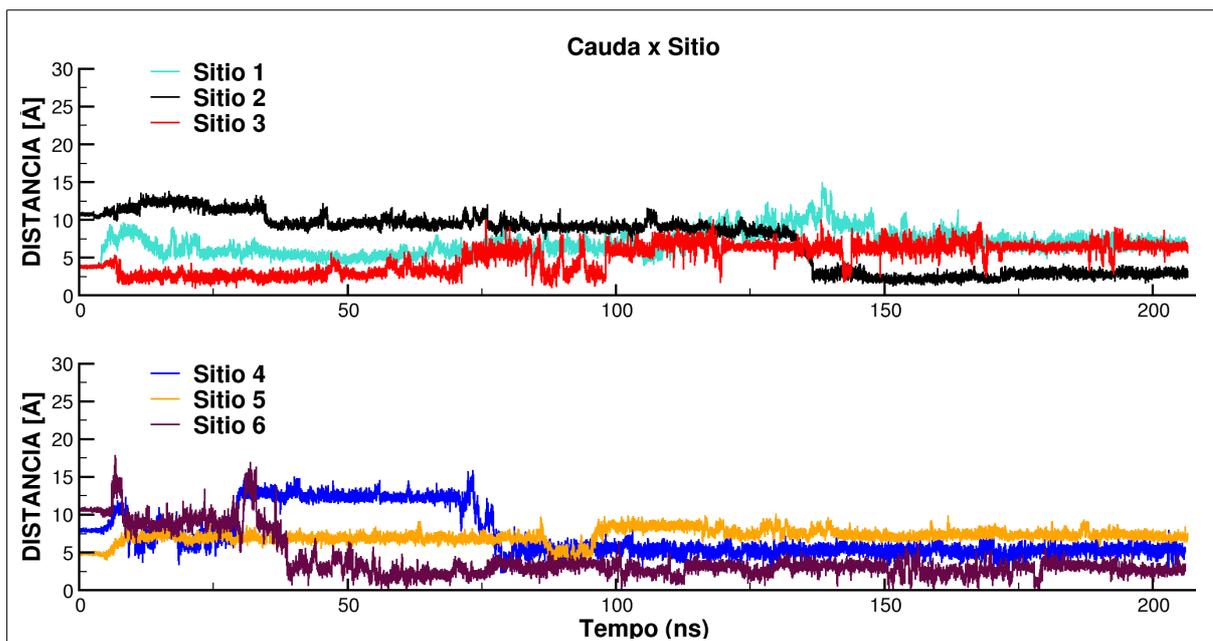


Figura 18. Distância dependente do tempo entre a cauda de hidrocarbonetos do colesterol e o respectivo sítio de ligação (centro de massa).

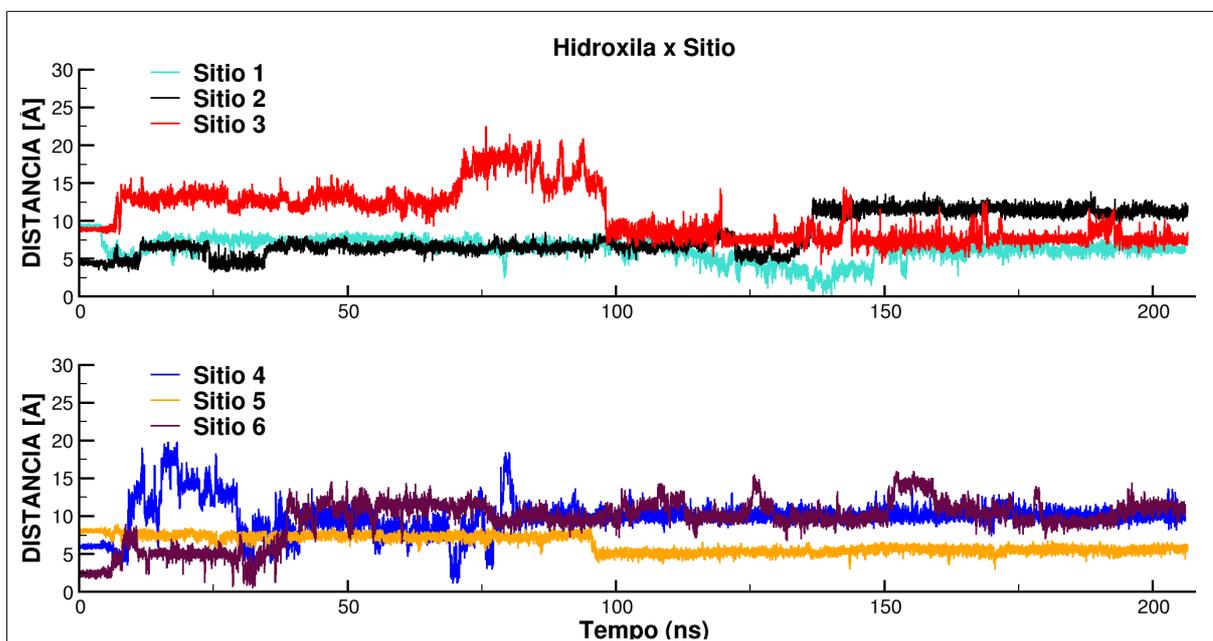


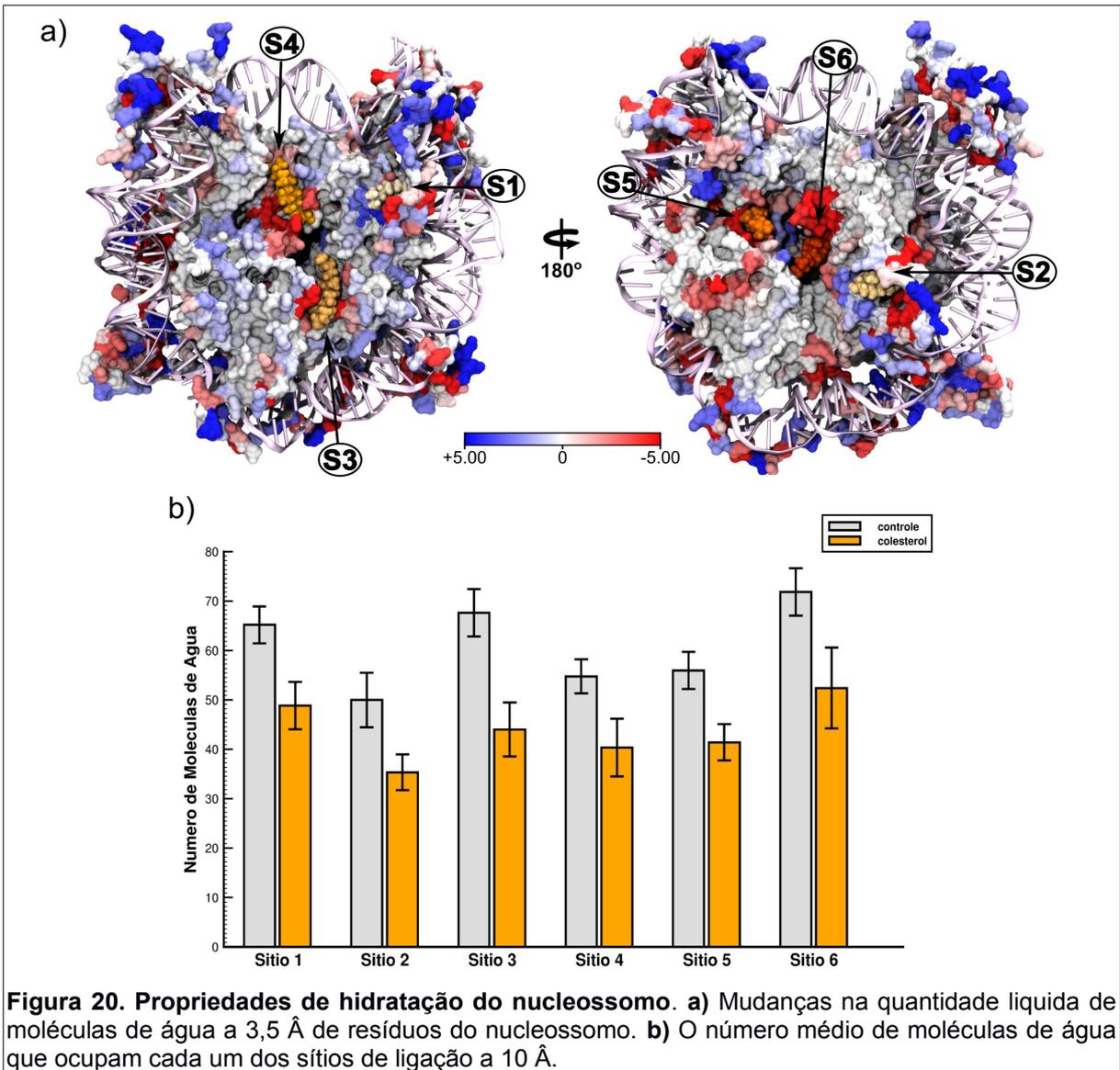
Figura 19. Distância dependente do tempo entre o grupo polar do colesterol e o respectivo sítio de ligação (centro de massa).

4.3.4 Perfil de hidratação dos sítios

Para avaliar o impacto causado pelo colesterol no perfil de hidratação do nucleossomo, foram realizadas duas análises. Na primeira uma esfera para a contagem de moléculas de água foi criada em cada sítio de interação do colesterol. O raio dessa esfera foi de 10 Å a partir do centro de cada sítio e o colesterol possui aproximadamente 17,5 Å de comprimento esticado x 6 Å de largura, portanto fica totalmente inserido na esfera criada. A seguir, calculou-se a quantidade média de moléculas de água presente em cada esfera durante toda a simulação, os resultados são apresentados na Figura 20b.

Também foi criado um mapa, por aminoácido, da quantidade de moléculas de água alteradas na presença do colesterol (Figura 20a). Para isso, foi computada a média de moléculas de água que, ao longo da trajetória, estavam a no máximo 3,5 Å de cada átomo do nucleossomo. Logo, cada átomo obteve um número médio de moléculas de água. Ao somar o número médio de todos os átomos de um mesmo resíduo, obteve-se o número total de moléculas de água para cada resíduo do nucleossomo. Posteriormente verificou-se a diferença entre a quantidade de água para cada resíduo na presença e ausência do colesterol (número de água do resíduo X com colesterol – número de água do resíduo X sem colesterol). Portanto valores abaixo de 0 simbolizam perda de água com colesterol e valores acima de 0 representam ganho de água.

É possível observar, tanto pelo mapa quanto pela análise por sítio, que o colesterol reduz significativamente a quantidade de moléculas de água na superfície do nucleossomo. Em média, o colesterol provoca uma redução de 17 ± 3 moléculas de água por esfera criada. Porém, no sítio 3 a redução chegou a 24 moléculas de água. O mapa de hidratação do nucleossomo revela que os aminoácidos próximos as moléculas de colesterol perderam água. Especialmente nos sítios 3, 4, 5 e 6, nos quais a redução do número de moléculas de água foi mais evidente do que nos sítios 1 e 2.



5 DISCUSSÃO

Moléculas ligantes do nucleossomo são de grande importância para a definição dos estados da cromatina. Como já mencionado, várias NBPs já foram descritas (SILVA; OLIVEIRA; SANTOS, 2015), 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. No ambiente celular existem várias moléculas que podem desencadear impacto sobre o nucleossomo e cromatina, entre elas o colesterol (REGENASS-KLOTZ; HEINIGER, 1984). Essa molécula participa de vários eventos celulares (PUCADYIL; CHATTOPADHYAY, 2006) e está associado a diversas doenças. Nesse trabalho, foram apresentadas evidências, ao nível atômico, sobre o impacto do colesterol na estrutura do nucleossomo, sugerindo um papel desse lipídio na regulação da cromatina.

Simulações de DM e *docking* molecular foram as metodologias principais utilizadas nesse trabalho. Baseado na estrutura cristalográfica de alta resolução publicada em 2002 (DAVEY *et al.*, 2002), foi criado um sistema atomístico do nucleossomo embebido em solução com moléculas de água. O PDB utilizado nesse trabalho é a base para a maioria das simulações de DM do nucleossomo (BISWAS; LANGOWSKI; BISHOP, 2013). O primeiro estudo do nucleossomo por simulação atomística de DM foi há mais de 10 anos (BISHOP, 2005), desde então vários estudos utilizaram dessa abordagem em suas investigações (BISWAS; LANGOWSKI; BISHOP, 2013). Assim a DM empregada em estudos do nucleossomo já foi extensivamente testada e validada como alternativa metodológica. O colesterol por sua vez, é alvo de extensivos estudos, baseados em DM, sobre sua interação com fosfolípidos (BERKOWITZ, 2009; RÓG; PASENKIEWICZ-GIERULA, 2001). Entretanto, a investigação desse lipídio com a cromatina, em especial o nucleossomo, é inédita.

Uma parte importante na análise de simulações de biomoléculas através da DM é a descrição das flutuações estruturais do sistema estudado (KARPLUS;

MCCAMMON, 2002). É vital que sejam analisadas estruturas estáveis do sistema criado. A estabilidade é uma forma de avaliar se o campo de força está descrevendo corretamente as interações entre os átomos (CAO; WANG, 2010), além disso a estrutura inicial da simulação representa uma conformação favorecida no ambiente cristalino (TERADA; KIDERA, 2012). Nesse sentido, a avaliação do perfil do RMSD é a alternativa comumente utilizada para julgar a estabilidade de um sistema simulado por DM. A sobreposição estrutural dos átomos (etapa que precede o computo do RMSD e RMSF), é crucial para o resultado final do RMSD e RMSF. Caso esta englobe partes da proteína altamente dinâmicas, o RMSD e o RMSF aumenta para todos os átomos, resultando numa avaliação errônea da estabilidade do sistema. Várias estratégias foram desenvolvidas para aperfeiçoar a sobreposição entre estruturas, porém, a maioria é baseada no tratamento diferenciado para os átomos com menor desvio, ou seja, aqueles que apresentam menores flutuações (MARTÍNEZ, 2015). Assim empregar 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.

Os resultados obtidos na busca da melhor seleção de átomos para a sobreposição estrutural, indicaram as hélices das histonas como a região com menores flutuações. Estruturas secundárias de proteínas, como alfa-hélices, apresentam alta estabilidade estrutural, a qual é aumentada quando em conjunto com outras estruturas secundárias (HOROVITZ; MATTHEWS; FERSHT, 1992). O perfil do RMSD para o conjunto de átomos pesados do nucleossomo, excluindo as caudas das histonas, se aproxima de 2 Å, similar ao valor encontrado em vários trabalhos realizados com esse tipo de sistema (MATERESE; SAVELYEV; PAPOIAN, 2009; ROCCATANO; BARTHEL; ZACHARIAS, 2007; RUSCIO; ONUFRIEV, 2006), que representa as flutuações térmicas normais de uma proteína. Além disso, também está condizente com os fatores de temperatura cristalográficos encontrados no cristal utilizado nesse trabalho (DAVEY *et al.*, 2002). Outras interfaces chave para a estabilidade do nucleossomo tiveram o perfil de RMSD avaliado. As três interfaces (domínio de *docking* e os *4-helix bundles* H3:H3 e H2B:H4) não ultrapassaram 3 Å de RMSD nos 20 ns analisados. No domínio de *docking* a presença de uma região

flexível, semelhante as caudas, provavelmente contribuiu para uma curva de RMSD maior do que as demais seleções. Os resultados do RMSD sugerem que existe um *ensemble* de estruturas em equilíbrio após 10 ns de simulação, deste modo esse *pool* de estruturas foram utilizadas nas investigações do *docking*.

Linus Pauling declarou, há várias décadas, que o reconhecimento molecular é o segredo da vida, de fato, vários processos biológicos cruciais são regulados por interações não ligadas e associações moleculares são alvo de extensivos estudos (BARON; MCCAMMON, 2013). O *docking* é uma técnica que busca o reconhecimento molecular entre um ligante, usualmente uma pequena molécula, e um receptor. Essa ferramenta proporciona uma análise relativamente rápida e principalmente mais econômica do que técnicas experimentais na investigação inicial de interações não covalentes entre biomoléculas (SOUSA *et al.*, 2013). O *ensemble* de estruturas em equilíbrio foi criado para a consideração das flutuações do nucleossomo nos cálculos de *docking*. Pois, hoje está claro que as proteínas em solução são sistemas dinâmicos, preexistindo em um conjunto de estados conformacionais que são essenciais para a sua função. Essas 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).

No geral, os resultados do *docking* sugerem que não existem sítios específicos de ligação do colesterol no nucleossomo, pois foram observadas nuvens de soluções espalhadas pelo receptor e as energias são favoráveis a interação de todas as moléculas de colesterol, mas com baixa afinidade. As nuvens de soluções encontradas somente no DNA não foram analisadas nesse estudo, pois foram buscadas interações com a porção proteica do nucleossomo. Tendo em vista análises futuras com mutações pontuais para confirmar a interação do colesterol, sítios que interajam com as histonas poderão ser testados e já sítios exclusivos do DNA, por baixa especificidade, não poderiam ser validados experimentalmente por essa abordagem. Observou-se uma maior densidade de soluções nos dímeros do que no tetrâmero. Talvez essa possível preferência do colesterol pelos dímeros, possa ajudar na compreensão dos resultados de ensaios de termoestabilidade do

mononucleossomo realizado pelo grupo do professor Guilherme Santos. Foi notado que a presença de colesterol em altas concentrações (5 – 10 μM) antecipa a evicção dos dímeros e posteriormente do tetrâmero (artigo submetido a publicação). Também foram encontrados sítios de interação do colesterol em regiões descritas na literatura como importantes para a compactação da cromatina. Por exemplo, a cauda da H4 (sítios 1 e 2), que interage com o *patch* ácido do nucleossomo adjacente para a formação de estruturas mais compactadas de cromatina (LUGER *et al.*, 1997). Soluções próximas ao *patch* ácido também foram encontradas (sítios 3 e 6), além do papel na condensação da cromatina, esta região é alvo de ligação de várias NBP (KALASHNIKOVA *et al.*, 2013). Portanto o colesterol poderia interferir nessas interações. Regiões chave para a formação do nucleossomo como o *4-helix bundle* H2B:H4 e o *4-helix bundle* H3:H3 (sítio 4) apresentaram alguns sítios de ligação do colesterol. Já no domínio de *docking* da H2A (sítio 5) foi verificada duas nuvens de soluções.

Apesar do *docking* auxiliar na identificação de modos de ligação do colesterol no nucleossomo, ele não consegue fornecer os detalhes energéticos do complexo, além disso, essa técnica pode gerar soluções falsamente positiva. Assim, com o intuito de aumentar a acurácia da estimativa da afinidade (energia livre de ligação) e validar os sítios encontrados pelo *docking*, foi buscado alternativas metodológicas para a predição de afinidades. Abordagens baseadas na termodinâmica estatística que requerem trajetórias do sistema de interesse, são as mais robustas para estimar a energia livre de ligação entre moléculas. Dentre as opções disponíveis baseadas nesse tratamento termodinâmico, as técnicas clássicas, como o FEP (*Free Energy Perturbation*) (MEZEI; SWAMINATHAN; BEVERIDGE, 1978; STJERNSCHANTZ *et al.*, 2006), são dispendiosas computacionalmente devido ao espaço amostral exigido. Como alternativa, foram desenvolvidos métodos mais aproximados para o cálculo da energia livre de ligação, conhecidos como métodos “*end-point*”. Tais métodos simplificam o cálculo de energia livre, pois não precisam amostrar os estados intermediários, físicos ou não físicos, deste modo são necessários apenas os estados finais (a proteína livre, o ligante livre e o complexo) (STEINBRECHER;

LABAHN, 2010). Arya e Yang aproveitaram de métodos *end-point* para estimar a afinidade de um peptídeo da cauda da H4 com o *patch* ácido (YANG; ARYA, 2011).

O método adotado nesse trabalho para o cálculo de energia livre de ligação foi o LIE (*Linear Interaction Energy*) (AQVIST; MEDINA; SAMUELSSON, 1994), uma técnica “*end-point*” popular e entre as mais precisas (GENHEDEN; RYDE, 2012). Uma vantagem do LIE é que ele se concentra em simulação dos estados termodinâmicos relevantes. Entretanto a amostragem dos estados ligados e livre do ligante deve ser grande o suficiente para que o erro associado à média das interações eletrostáticas e de van der Waals seja pequeno. O valor de $\alpha=0,18$ empregado nesse estudo, foi parametrizado utilizando energias livres de ligação experimentais e, conseqüentemente, leva em conta todas as contribuições de van der Waals para a associação do ligante, por exemplo, entropias rotacionais e translacionais (CARLSSON; ÅQVIST, 2006). Originalmente o valor de β empregado foi de $\frac{1}{2}$, obtido a partir da aproximação da resposta linear para estimar o termo eletrostático da energia de solvatação (ÅQVIST; HANSSON, 1996; AQVIST; MEDINA; SAMUELSSON, 1994). Vários pesquisadores investigaram a variação desse parâmetro para melhor predição da energia livre de ligação. Valores inferiores ao originalmente descrito foram empregados conforme o aumento do potencial de formação de ligações de hidrogênio dos grupos -OH, (HANSSON; MARELIUS; ÅQVIST, 1998; MARELIUS; HANSSON; ÅQVIST, 1998). Entretanto para o colesterol, que possui apenas um grupo OH, essa variação não é aplicável, portanto foi escolhido $\beta=0,5$.

Devido a inviabilidade computacional para calcular a afinidade de todos os sítios encontrados pelo *docking*, seis moléculas de colesterol foram escolhidas e simuladas por DM ligadas ao nucleossomo. Os critérios de escolha dos sítios foram: (1) sítios que estejam interagindo em regiões funcionalmente importantes para a estrutura do nucleossomo e arquitetura da cromatina; (2) sítios mais superficiais, pois sítios muito enterrados podem ser falso positivo, devido ao acesso improvável. Por outro lado, mesmo que esses sítios não sejam falsos positivos, o tempo necessário para tirar conclusões satisfatórias deveria ser extenso, provavelmente muito difícil de atingir. Para evitar interferência no cálculo da energia livre de ligação

de cada sítio, as seis moléculas de colesterol foram divididas em duas simulações em que a distância mínima entre os ligantes foi de 31 Å. Apesar de apenas seis sítios terem sido escolhidos, os demais poderão ser considerados em estudos futuros.

Os sítios escolhidos envolvem as regiões principais encontradas pelo *docking* (especificados anteriormente entre parênteses). Todos os sítios avaliados apresentaram estimativa da energia livre de ligação favoráveis a ligação com o nucleossomo. Entretanto, como esperado, esses valores variaram de acordo com a cavidade em que o colesterol está inserido. Devido à característica apolar do colesterol (ALBERTS *et al.*, 1994), os valores de interação de vdW mais negativos que os valores do potencial eletrostático estão de acordo com o esperado. Os resultados sugerem que o colesterol presente no sítio 5 (domínio de *docking*) possui maior afinidade pelo nucleossomo. Este domínio exerce papel fundamental na interação entre o tetrâmero (H3-H4) e os dímeros (H2A-H2B) (LUGER *et al.*, 1997) e auxilia no empacotamento do DNA (VOGLER *et al.*, 2010). Assim, a interação nesse domínio pode eventualmente ocasionar em estados mais estáveis do nucleossomo. Como os resultados observados acerca da energética de interação do colesterol com o nucleossomo suportam a formação desse complexo, análises específicas sobre o impacto do ligante na estrutura do nucleossomo foram realizadas, como os perfis do RMSD e RMSF do nucleossomo na presença e ausência do colesterol, a análise da dinâmica das moléculas de colesterol e alterações na hidratação da superfície do nucleossomo desencadeadas pelo ligante.

O colesterol não influenciou a estabilidade do sistema, o perfil do RMSD permaneceu muito similar ao observado no sistema controle, dentro da faixa de equilíbrio descrita por outros trabalhos. Quando o RMSD do sistema está estável, a medida do RMSF se torna uma ferramenta poderosa para revelar informações específicas sobre os movimentos atômicos locais (KUZMANIC; ZAGROVIC, 2010; MARTÍNEZ, 2015). O RMSF calculado não considerou o DNA, pois os sítios de ligação do colesterol simulados, encontravam-se predominantemente nas histonas. A confirmação da estabilidade do sistema também permitiu que outras análises fossem realizadas.

Tabela 4. Distância média, em *angströms* (Å), entre as moléculas de colesterol e o respectivo sítio de interação. Valores médios e desvios das distâncias apresentadas nas figuras 17, 18 e 19.

Sítios	Colesterol	Cauda de hidrocarbonetos	Hidroxila
1	1,98 ± 1,30	6,98 ± 1,62	6,33 ± 1,47
2	5,28 ± 1,37	7,40 ± 3,55	8,05 ± 2,60
3	4,64 ± 2,60	5,02 ± 1,86	10,69 ± 3,40
4	5,65 ± 2,19	7,23 ± 3,03	9,91 ± 2,20
5	2,80 ± 0,60	7,20 ± 0,90	6,37 ± 1,07
6	3,71 ± 1,55	4,16 ± 2,85	9,56 ± 2,76

A flexibilidade apresentada pelo colesterol em conjunto com as energias livre de ligação estimadas, sugerem um modo de interação dinâmico do colesterol através dos sítios de ligação no nucleossomo. Foi verificado que havia uma correlação inversamente proporcional entre as flutuações apresentadas pelas moléculas de colesterol e as energias obtidas pelo LIE. Com exceção do sítio 1, nas cavidades em que o colesterol apresentou maior movimentação (maior desvio), a energia livre de ligação foi menor. Este dado indica que a adequação conformacional do colesterol no sítio de ligação é determinante para sua afinidade com o nucleossomo. A Tabela 4 apresenta os valores médios das distâncias dos sítios de interação em relação ao colesterol e suas regiões polares e apolares. O colesterol é uma molécula anfipática, porém, a região lipofílica é extensivamente maior do que a região hidrofílica. De fato, na maioria dos sítios de ligação, o colesterol interage predominantemente com a cauda de hidrocarbonetos (distância média entre a cauda do colesterol e o sítio menor do que a distância média entre a hidroxila e o sítio), apresentando menores flutuações (desvio) ao longo da trajetória nas cavidades mais hidrofóbicas, ou seja, sítios 5, 3 e 6. Entretanto, o ligante presente no sítio 1 é exceção as correlações observadas em relação as distâncias para o centro dos sítios, provavelmente isso ocorreu devido a interação do colesterol com o DNA nessa cavidade ser mais extensa do que nas demais cavidades, diminuindo a liberdade conformacional do ligante.

A mudança no perfil de hidratação do nucleossomo na presença de colesterol indica que essa molécula possa auxiliar os contatos em regiões hidrofóbicas. Tais regiões são importantes para as interações entre as histonas (formação do tetrâmero e do octâmero) (EICKBUSH; MOUDRIANAKIS, 1978) e foram nomeadas em 1997 (LUGER *et al.*, 1997) como *4-helix bundle*. Assim, o colesterol provavelmente aumenta a estabilidade do nucleossomo e facilita estados mais compactados da cromatina. De fato, experimentos realizados pelo grupo de pesquisa do Professor Guilherme Santos demonstram que o colesterol possui efeito empacotador das fibras de cromatina *in vitro* (resultados apresentados no artigo em anexo). Ainda sobre a hidratação do nucleossomo, estudo realizado pelo grupo de pesquisa liderado por Richmond, observou que moléculas de água contribuem para a formação do nucleossomo auxiliando de forma significativa a estabilidade da ligação entre o DNA e histonas. Além disso, as moléculas de água também desempenham um papel importante na promoção da mobilidade do nucleossomo, adaptando a superfície das histonas a variações na conformação do DNA. (DAVEY *et al.*, 2002). Visto os efeitos macroscópicos do colesterol em conjunto com sua ação removedora de moléculas de água, e ainda o fato da água ser importante para a interação entre o DNA e histonas, os sítios presentes no DNA provavelmente estariam relacionados a estados menos estáveis do nucleossomo e consequentemente da cromatina. Dessa forma, esses dados suportam a exclusão das soluções de *docking* presentes no DNA.

Dentre os sítios de ligação analisados, duas moléculas de colesterol apresentaram interações com aminoácidos do *patch* ácido (E91 e E92 – H2A) e reduziram as suas flutuações atômicas. Esse efeito sugere que o colesterol possivelmente interfira no remodelamento da cromatina mediado por NBPs (MCGINTY; TAN, 2016). Ademais, a ligação específica do colesterol a cauda da H4 também pode estar relacionada há um efeito modulador da cromatina. Dado que foi observado interação com o resíduo R23 – H4, o qual foi identificado como essencial para a ligação da cauda da H4 com o *patch* ácido (SONG *et al.*, 2014).

Em conjunto os resultados apresentados nesse estudo demonstram que o colesterol se liga ao nucleossomo e desencadeia mudanças estruturais locais e no

perfil de hidratação do nucleossomo. Tais alterações estão de acordo com os efeitos do colesterol observados pelo grupo de pesquisa do Professor Guilherme Santos.

Apesar de reveladas várias informações sobre a interação do colesterol com o nucleossomo, ainda existem algumas questões abertas. Por exemplo, como funciona a dinâmica da ligação do colesterol ao nucleossomo na presença de NBPs ou NBMs? Como o colesterol interage, ao nível atômico, com as fibras de cromatina?

6 CONCLUSÕES

A cromatina exerce papel fundamental no que diz respeito à manutenção e evolução da vida no nosso planeta. Portanto podemos considerá-la entre as mais importantes macromoléculas biológicas. Entender como a cromatina funciona representa um desafio imenso para os cientistas desse século. Dito isso, este estudo contribuiu para a compreensão da complexa interação entre colesterol e nucleossomo, moléculas tão importantes para o funcionamento celular.

Primeiramente foi verificado a ligação de seis moléculas de colesterol, durante 200 ns de simulação de equilíbrio, próximas a regiões essenciais para formação do nucleossomo e sua estabilidade estrutural, como o domínio de *docking*, o qual foi particularmente afetado pela ligação do colesterol. Essa interação possui uma baixa afinidade (estimada pelo método LIE). Além disso, alterações estruturais locais foram observadas, especificamente na cauda da H4 e em alguns aminoácidos do *patch* ácido, assim, o colesterol pode auxiliar a interação entre essas estruturas, resultando em estados mais compactados das fibras de cromatina, como observado *in vitro*. Também foi observada uma interação predominantemente hidrofóbica entre o colesterol e o nucleossomo e evidenciado a redução do número de moléculas de água nas cavidades de ligação com o colesterol. Tais resultados, em conjunto, sugerem que é possível que o colesterol aumente a estabilidade do nucleossomo através da desidratação da superfície nucleossomal, a qual pode fortalecer contatos hidrofóbicos em regiões fundamentais para a estrutura do nucleossomo.

Considerando todos os resultados aqui discutidos e em conjunto com resultados obtidos pelo grupo de pesquisa do Professor Guilherme, podemos afirmar que um esterol lipídico (colesterol) configura outra peça do complexo quebra-cabeça que a dinâmica da cromatina representa.

7 REFERÊNCIAS

- ALBERTS, B. *et al.* **Molecular Biology of the Cell, Third Edition**. 3rd edition ed. New York: Garland Science, 1994.
- 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.
- ALBI, E.; MAGNI, M. V. The presence and the role of chromatin cholesterol in rat liver regeneration. **Journal of Hepatology**, v. 36, n. 3, p. 395–400, mar. 2002.
- 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.
- ÅQVIST, J.; HANSSON, T. On the Validity of Electrostatic Linear Response in Polar Solvents. **The Journal of Physical Chemistry**, v. 100, n. 22, p. 9512–9521, 1 jan. 1996.
- AQVIST, J.; MEDINA, C.; SAMUELSSON, J. E. A new method for predicting binding affinity in computer-aided drug design. **Protein Engineering**, v. 7, n. 3, p. 385–391, mar. 1994.
- 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.
- BATCHO, P. F.; CASE, D. A.; SCHLICK, T. Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations. **The Journal of Chemical Physics**, v. 115, n. 9, p. 4003–4018, 1 set. 2001.
- 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.
- BERKOWITZ, M. L. Detailed molecular dynamics simulations of model biological membranes containing cholesterol. **Biochimica et Biophysica Acta (BBA)** -

Biomembranes, Lipid Interactions, Domain Formation, and Lateral Structure of Membranes. v. 1788, n. 1, p. 86–96, jan. 2009.

BISHOP, T. C. Molecular dynamics simulations of a nucleosome and free DNA. **Journal of Biomolecular Structure & Dynamics**, v. 22, n. 6, p. 673–686, jun. 2005.

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.

BISWAS, M.; LANGOWSKI, J.; BISHOP, T. C. Atomistic simulations of nucleosomes. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, v. 3, n. 4, p. 378–392, 1 jul. 2013.

BOWERMAN, S.; WERESZCZYNSKI, J. Effects of macroH2A and H2A.Z on nucleosome structure and dynamics as elucidated by molecular dynamics simulations. **arXiv:1507.05655 [physics, q-bio]**, 20 jul. 2015.

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.

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.

CARLSSON, J.; ÅQVIST, J. Calculations of solute and solvent entropies from molecular dynamics simulations. **Phys. Chem. Chem. Phys.**, v. 8, n. 46, p. 5385–5395, 2006.

CARRIERE, V. *et al.* IL-33, the IL-1-like cytokine ligand for ST2 receptor, is a chromatin-associated nuclear factor in vivo. **Proceedings of the National Academy of Sciences**, v. 104, n. 1, p. 282–287, 2 jan. 2007.

COLLEPARDO-GUEVARA, R. *et al.* Chromatin Unfolding by Epigenetic Modifications Explained by Dramatic Impairment of Internucleosome Interactions: A Multiscale Computational Study. **Journal of the American Chemical Society**, v. 137, n. 32, p. 10205–10215, 19 ago. 2015.

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.

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.

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.

DROR, R. O. *et al.* Biomolecular Simulation: A Computational Microscope for Molecular Biology. **Annual Review of Biophysics**, v. 41, n. 1, p. 429–452, 2012.

EICKBUSH, T. H.; MOUDRIANAKIS, E. N. The histone core complex: an octamer assembled by two sets of protein-protein interactions. **Biochemistry**, v. 17, n. 23, p. 4955–4964, nov. 1978.

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.

FINCH, J. T. *et al.* Structure of nucleosome core particles of chromatin. **Nature**, v. 269, n. 5623, p. 29–36, 1 set. 1977.

FINCH, J. T.; KLUG, A. X-ray and electron microscope analyses of crystals of nucleosome cores. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 42 Pt 1, p. 1–9, 1978.

FOLOPPE, N.; MACKERELL, J., 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.

GENHEDEN, S.; RYDE, U. Comparison of end-point continuum-solvation methods for the calculation of protein–ligand binding free energies. **Proteins: Structure, Function, and Bioinformatics**, v. 80, n. 5, p. 1326–1342, 1 maio 2012.

GREWAL, S. S.; ELGIN, S. C. R. Transcription and RNAi in the formation of heterochromatin. **Nature**, v. 447, n. 7143, p. 399–406, 24 maio 2007.

GRUMMER, R. R.; CARROLL, D. J. A review of lipoprotein cholesterol metabolism: importance to ovarian function. **Journal of Animal Science**, v. 66, n. 12, p. 3160–3173, dez. 1988.

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.

HANSSON, T.; MARELIUS, J.; ÅQVIST, J. Ligand binding affinity prediction by linear interaction energy methods. **Journal of Computer-Aided Molecular Design**, v. 12, n. 1, p. 27–35, jan. 1998.

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.

HORN, P. J.; PETERSON, C. L. Chromatin Higher Order Folding--Wrapping up Transcription. **Science**, v. 297, n. 5588, p. 1824–1827, 13 set. 2002.

HOROVITZ, A.; MATTHEWS, J. M.; FERSHT, A. R. α -Helix stability in proteins. **Journal of Molecular Biology**, v. 227, n. 2, p. 560–568, 20 set. 1992.

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.

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

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.

KALLEN, J. A. *et al.* 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**, v. 10, n. 12, p. 1697–1707, dez. 2002.

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.

KATO, H. *et al.* A Conserved Mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C. **Science (New York, N.Y.)**, v. 340, n. 6136, p. 1110–1113, 31 maio 2013.

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.

KUZMANIC, A.; ZAGROVIC, B. Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. **Biophysical Journal**, v. 98, n. 5, p. 861–871, 3 mar. 2010.

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.

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.

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.

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.

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.

MARELIUS, J.; HANSSON, T.; ÅQVIST, J. Calculation of ligand binding free energies from molecular dynamics simulations. **International Journal of Quantum Chemistry**, v. 69, n. 1, p. 77–88, 1 jan. 1998.

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.

MATERESE, C. K.; SAVELYEV, A.; PAPOIAN, G. A. Counterion Atmosphere and Hydration Patterns near a Nucleosome Core Particle. **Journal of the American Chemical Society**, v. 131, n. 41, p. 15005–15013, 21 out. 2009.

MCGINTY, R. K.; HENRICI, R. C.; TAN, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. **Nature**, v. 514, n. 7524, p. 591–596, 30 out. 2014.

MCGINTY, R. K.; TAN, S. Nucleosome Structure and Function. **Chemical Reviews**, 12 dez. 2014.

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.

MEZEI, M.; SWAMINATHAN, S.; BEVERIDGE, D. L. Ab initio calculation of the free energy of liquid water. **Journal of the American Chemical Society**, v. 100, n. 10, p. 3255–3256, 1 maio 1978.

MUSSELMAN, C. A. *et al.* Perceiving the epigenetic landscape through histone readers. **Nature structural & molecular biology**, v. 19, n. 12, p. 1218–1227, dez. 2012.

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

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

PIERCE, B. A. **Genetics: A Conceptual Approach 3rd Edition**. 3 edition ed. New York, NY: W. H. Freeman, 2007.

PUCADYIL, T. J.; CHATTOPADHYAY, A. Role of cholesterol in the function and organization of G-protein coupled receptors. **Progress in Lipid Research**, v. 45, n. 4, p. 295–333, jul. 2006.

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.

REGENASS-KLOTZ, M.; HEINIGER, H.-J. Specific binding of cholesterol to chromatin prepared from mouse spleen cells. **Biochemistry and Cell Biology**, v. 62, n. 2–3, p. 94–99, 1984.

RICHMOND, T. J. *et al.* Structure of the nucleosome core particle at 7 Å resolution. **Nature**, v. 311, n. 5986, p. 532–537, 11 out. 1984.

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.

ROCCATANO, D.; BARTHEL, A.; ZACHARIAS, M. Structural flexibility of the nucleosome core particle at atomic resolution studied by molecular dynamics simulation. **Biopolymers**, v. 85, n. 5–6, p. 407–421, 5 abr. 2007.

RÓG, T.; PASENKIEWICZ-GIERULA, M. Cholesterol effects on the phospholipid condensation and packing in the bilayer: a molecular simulation study. **FEBS Letters**, v. 502, n. 1–2, p. 68–71, 27 jul. 2001.

ROUSSEL, L. *et al.* Molecular mimicry between IL-33 and KSHV for attachment to chromatin through the H2A–H2B acidic pocket. **EMBO Reports**, v. 9, n. 10, p. 1006–1012, out. 2008.

RUSCIO, J. Z.; ONUFRIEV, A. A Computational Study of Nucleosomal DNA Flexibility. **Biophysical Journal**, v. 91, n. 11, p. 4121–4132, 1 dez. 2006.

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.

SHOGREN-KNAAK, M. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. **Science (New York, N.Y.)**, v. 311, n. 5762, p. 844–847, 10 fev. 2006.

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. **Estudos estruturais da cromatina: ação do colesterol e obtenção do complexo receptor nuclear: nucleossomo**. 2013. Dissertação. Universidade de Brasília, 2013. Disponível em: <<http://repositorio.unb.br/handle/10482/14345>>. Acesso em: 2 mar. 2016.

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.

SIMONS, K.; IKONEN, E. How cells handle cholesterol. **Science (New York, N.Y.)**, v. 290, n. 5497, p. 1721–1726, 1 dez. 2000.

SIMONS, K.; TOOMRE, D. Lipid rafts and signal transduction. **Nature Reviews Molecular Cell Biology**, v. 1, n. 1, p. 31–39, out. 2000.

SIMPSON, R. T. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. **Biochemistry**, v. 17, n. 25, p. 5524–5531, 12 dez. 1978.

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/cmc/2013/00000020/00000018/art00002>>. Acesso em: 18 mar. 2016.

STEINBRECHER, T.; LABAHN, A. Towards Accurate Free Energy Calculations in Ligand Protein-Binding Studies. **Current Medicinal Chemistry**, v. 17, n. 8, p. 767–785, 1 mar. 2010.

STJERNSCHANTZ, E. *et al.* Are Automated Molecular Dynamics Simulations and Binding Free Energy Calculations Realistic Tools in Lead Optimization? An Evaluation of the Linear Interaction Energy (LIE) Method. **Journal of Chemical Information and Modeling**, v. 46, n. 5, p. 1972–1983, 1 set. 2006.

SWOPE, W. C. *et al.* A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. **The Journal of Chemical Physics**, v. 76, n. 1, p. 637–649, 1 jan. 1982.

TALBERT, P. B.; HENIKOFF, S. Histone variants — ancient wrap artists of the epigenome. **Nature Reviews Molecular Cell Biology**, v. 11, n. 4, p. 264–275, 1 abr. 2010.

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.

TOUKMAJI, A. Y.; BOARD, J. A. Ewald summation techniques in perspective: a survey. **Computer Physics Communications**, v. 95, n. 2, p. 73–92, 1 jun. 1996.

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.
- 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.
- WEBER, C. M.; HENIKOFF, S. Histone variants: dynamic punctuation in transcription. **Genes & Development**, v. 28, n. 7, p. 672–682, 1 abr. 2014.
- 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.
- ZAINA, S. *et al.* Chromatin modification by lipids and lipoprotein components: an initiating event in atherogenesis? **Current Opinion in Lipidology**, v. 16, n. 5, p. 549–553, out. 2005.