UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

**Estudo dinâmico conformacional da proteína calgranulina C
(S100A12) mediante interação com íons e receptor RAGE**

Renata Almeida Garcia Reis

Ribeirão Preto
2012

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

**Estudo dinâmico conformacional da proteína calgranulina C
(S100A12) mediante interação com íons e receptor RAGE**

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Ciências
Farmacêuticas para obtenção do Título de
Mestre em Ciências

Área de Concentração: Química e Física
Biológica

Orientada: Renata Almeida Garcia Reis

Orientador: Antonio Caliri

Versão corrigida da Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em 03/05/2012. A versão original encontra-se disponível na Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP.

Ribeirão Preto
2012

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.

Reis, Renata Almeida Garcia

Estudo dinâmico conformacional da proteína calgranulina C (S100A12) mediante interação com íons e receptor RAGE. Ribeirão Preto, 2012.
122 p.; 30cm.

Dissertação de Mestrado, apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Química e Física Biológica.

Orientador: Caliri, Antonio.

1. Calgranulina C. 2. Dinâmica Molecular. 3. Conformação Estrutural.

FOLHA DE APROVAÇÃO

Nome do aluno: Renata Almeida Garcia Reis

Título do trabalho: Estudo dinâmico conformacional da proteína calgranulina C (S100A12) mediante interação com íons e receptor RAGE

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Ciências
Farmacêuticas para obtenção do Título de
Mestre em Ciências

Área de Concentração: Química e Física
Biológica.

Orientador: Antonio Caliri

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Dedicatória

Aos meus pais, José Cláudio e Maria Izilda

Agradecimento

Aos meus pais, José Cláudio e Maria Izilda, pessoas tão admiráveis, agradeço pelo apoio e pelo incentivo incondicional, e por terem me ensinado a enxergar a beleza no aprender.

Ao meu orientador, Prof. Dr. Antonio Caliri, agradeço pela orientação neste trabalho, pela maneira generosa como me recebeu no grupo de Física Biológica desde a iniciação científica, e pela disposição em discutir e ensinar uma farmacêutica.

À Universidade de São Paulo, e à Faculdade de Ciências Farmacêuticas de Ribeirão Preto, agradeço pela infra-estrutura, pelo apoio e por permitir que eu realize as primeiras etapas no sonho de tornar-me professora e pesquisadora.

Ao meu irmão Vitor agradeço pela amizade, companheirismo e compreensão das minhas constantes ausências.

Ao meu sobrinho Caio agradeço por nos trazer tanta alegria, e por ter sempre perguntas novas para nos desafiar, e à minha afilhada Júlia, que em breve chegará, embora já nos inspire com os ares de uma nova vida.

Aos meus avós agradeço pelo carinho acolhedor.

Aos amigos do H03, Carol, Diego, Elaine, Moisés, Pamela, Sabrina e Vanessa, por terem compartilhado comigo a experiência de viver no Porto e de conhecer tantos lugares novos. Agradeço pela amizade de hoje, pela presença constante, ainda que exista distância.

Aos amigos da 79ª turma da Faculdade de Ciências Farmacêuticas de Ribeirão Preto, agradeço pelos anos da graduação, pelo entusiasmo, pelas discussões e por todas as maravilhosas diferenças que nos fizeram pessoas melhores.

Aos amigos, Carol, Daiane, Henrique, Juliana M., Juliana V., Livia e Paula, pelo companheirismo do dia a dia.

Aos amigos do laboratório, Flávio, João, Leandro, Ricardo e Rodrigo pelo ambiente tão harmonioso e pelos conhecimentos compartilhados.

A todos os meus professores, desde os anos do colégio, agradeço por terem me revelado suas perspectivas acerca de tantos assuntos.

Aos desenvolvedores do Pacote de Simulação GROMACS e de todos os programas que foram utilizados no presente trabalho, por disponibilizarem gratuitamente os resultados dos seus esforços.

Ao programa de Pós Graduação em Ciências Farmacêuticas e à FAPESP pelo apoio financeiro.

Epígrafe

*"If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the **atomic hypothesis** (or the atomic fact, or whatever you wish to call it) **that all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another.** In that one sentence, you will see, there is an enormous amount of information about the world, if just a little imagination and thinking are applied".*

Richard P. Feynman (The Feynman Lectures on Physics)

RESUMO

REIS, R. A. G. **Estudo dinâmico conformacional da proteína calgranulica C (S100A12) mediante interação com íons e receptor RAGE**. 2012. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2012.

Calgranulina C (S100A12) é membro da família das proteínas S100 “EF-hands” que complexam cálcio. A S100A12 humana é expressa predominantemente por granulócitos e é superexpressa em compartimentos inflamatórios. Níveis séricos elevados de S100A12 são encontrados em pacientes acometidos por distúrbios inflamatórios, neurodegenerativos, metabólicos e neoplasias. A S100A12 intracelular existe como um homodímero anti-paralelo. Cada monômero é composto por um “EF-hand” clássico, C-terminal ($H_I - L_I - H_{II}$), e um “EF-hand” N-terminal, o pseudo “EF-hand” ($H_{III} - L_{III} - H_{IV}$). Os “motifs” são conectados pela região do “hinge” (L_{II}). A Calgranulina C também liga íons zinco e cobre em uma região formada pelas duas subunidades do dímero. Mudanças nas concentrações citosólicas de íons regulam uma grande variedade de processos celulares, e as proteínas que complexam íons são moléculas importantes na transdução do sinal, diferenciação e controle do ciclo celular. O mecanismo pelo qual a Calgranulina C modula o curso do processo inflamatório está relacionado à interação com o receptor para produtos finais de glicosilação (RAGE). Para obter detalhes sobre os mecanismos envolvidos nas etapas de sinalização celular das quais a S100A12 participa, nosso objetivo foi qualificar e quantificar a atividade conformacional dos domínios da S100A12 induzida por variações de parâmetros termodinâmicos intensivos, como mudanças nas concentrações de íons. Além disso, nós investigamos os detalhes da interação entre S100A12 e RAGE para elucidar a região do receptor com a qual a S100A12 interage e quais são os resíduos envolvidos nesta interação. Para os estudos da influência da presença de íons na dinâmica conformacional da S100A12, simulações de dinâmica molecular foram realizadas usando o pacote de simulação GROMACS com o campo de força OPLS-AA, no “ensemble” NVT. As estruturas iniciais usadas foram as estruturas cristalográficas da S100A12 (PDB ID: 2WCE e 1E8A). Estas foram submetidas a diferentes concentrações de cloreto de sódio, cálcio e zinco em sistemas separados e solvatados com o modelo de água “SPC”. Nossos resultados sugerem que em baixas concentrações de Ca^{2+} , o L_I permanece ocupado pelo Na^+ . No período entre ondas de Ca^{2+} , este íon tem acesso à proteína exclusivamente pelo L_{III} (no EF-2). A medida em que há presença de Zn^{2+} , esse contribui para a saída do Na^+ do L_I , evento que envolve a participação do resíduo Asp25, permitindo que o L_I se abra e descomplexe o Na^+ . Além disso, devido a alta deformabilidade estrutural, a H_{III} é muito influenciada pelos íons Na^+ e Ca^{2+} , sendo que em determinadas concentrações, ambos levam a perda parcial desta hélice e da H_{IIa} (“Hinge-Region”) e ao aumento da flexibilidade desta região, embora apenas o Ca^{2+} seja capaz de se complexar, via H_{III} , a região próxima ao L_{III} . Com relação aos estudos com o RAGE, foram realizados estudos de “docking” molecular e simulações de SMD (“Steered Molecular Dynamic”). A análise dos nossos resultados, sugere que a interação da S100A12 com o RAGE ocorre tanto no domínio V, quanto no domínio C1 do RAGE e depende da região de conexão entre estes domínios. Também, observamos que estados oligoméricos maiores, por exemplo, hexâmeros de S100A12 (PDB ID: 1GQM), têm possibilidades maiores de interação com RAGE e que nestes casos, as regiões relevante da interação envolvem, de acordo com nossos resultados, porções N e C-terminal da H_I e C-terminal da H_{IV} da S100A12.

Palavras-chave: Calgranulina C, Dinâmica Molecular, Conformação Estrutural

ABSTRACT

REIS, R. A. G. **Dynamic and conformational study of protein calgranulin C (S100A12) in interaction with ions and RAGE receptor**. 2012. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2012.

Calgranulin C (S100A12) is a member of the S100 family of EF-hand calcium-binding proteins. Human S100A12 is predominantly expressed by granulocytes and is markedly overexpressed in inflammatory compartments. Elevated serum levels of S100A12 are found in patients suffering from various inflammatory, neurodegenerative, metabolic, and neoplastic disorders. Intracellular S100A12 exists as an anti-parallel homodimer. Each monomer is composed of a C-terminal, classic EF-hand ($H_I - L_I - H_{II}$), an N-terminal, pseudo EF-hand ($H_{III} - L_{III} - H_{IV}$). The motifs are linked by the hinge-region. Calgranulin C also binds zinc and copper ions in a site formed by both subunits of dimer. Changes in cytosolic ions concentrations regulate a wide variety of cellular process, and ions-binding proteins are the key molecules in signal transduction, differentiation, and cell cycle control. The mechanism by which calgranulin C modulates the course of inflammatory process is related to its interaction with the receptor for advanced glycated products (RAGE). In order to obtain details about the mechanism involved in cell signaling steps in which S100A12 participates, our goal was to qualify and quantify the activity conformational of S100A12 domains, induced by variations of intensive thermodynamic parameters, as changes in the concentration of ions. Furthermore we investigated the details of the interaction between S100A12 and RAGE in order to elucidate the region of the receptor which interacts with S100A12 and what are the residues involved in this interaction. In order to access the influence of the presence of ions over the conformational dynamics of S100A12, molecular dynamics simulations were performed using the GROMACS suite with the OPLS-AA force field and NVT ensemble. The initial structures used were experimentally determined by X-ray crystallography (PDB ID: 2WCE and 1E8A). They were separately submitted to different concentrations of sodium, calcium and zinc chloride and solvated with the SPC water model. Our results suggest that at low concentrations of Ca^{2+} , L_I remains occupied by Na^+ . During calcium-waves, it can reach the protein exclusively through L_{III} (in EF-2). As the Zn^{2+} concentration rises, it contributes to the Na^+ unbinding from L_I , an event that involves the residue ASP-25, which allows L_I to open and the Na^+ to unbind. Furthermore, because of its high structural deformability, H_{III} is strongly influenced by both Na^+ and Ca^{2+} ions which, in certain concentrations, leads to partial loss of this helix and of H_{IIa} (Hinge-Region) and increases in the flexibility of this region, although only Ca^{2+} is able to bind, through H_{III} , to the region near L_{III} . Regarding the RAGE studies, molecular docking essays and SMD (Steered Molecular Dynamics) simulations were performed. Our data analysis suggests that the interaction between S100A12 and RAGE takes place through both V and C1 RAGE domains and depends upon the interdomain region. Additionally, we observed that higher oligomeric states, e.g. S100A12 hexamers (PDB ID: 1GQM), have more interaction possibilities with RAGE and that, according to our results, in this case the interacting region of S100A12 comprises the N- and C-terminal portions of H_I and C-terminal of H_{IV} .

Keywords: Calgranulin C, Molecular-Dynamics, Structural conformation

LISTA DE FIGURAS

Figura 1. Representação esquemática do “motif” “EF-hand” [7].....	4
Figura 2. A e B representam a geometria de ligação do íon cálcio no domínio “EF-hand”, sendo que o íon ligado fica no centro da figura. Nas posições X e Y, encontram-se normalmente cadeias laterais de ácido aspártico ou asparagina; na posição Z encontram-se as cadeias laterais, também, de ácido aspártico ou asparagina, ou ainda, serina; na posição –Y encontra-se o oxigênio da carbonila do peptídeo. –X normalmente é uma molécula de água e –Z é a cadeia lateral de um resíduo como ácido glutâmico ou ácido aspártico. Adaptado de [8].....	5
Figura 3. Representação estrutural completa de uma proteína da família S100.....	7
Figura 4. Representação esquemática do RAGE, mostrando os domínios do receptor. A parte extracelular do RAGE é chamada RAGE-solúvel (sRAGE), enquanto o conjunto formado pela parte extracelular mais o domínio transmembrana é chamado DN-RAGE. As porções extracelular, transmembrana e intracelular formam o RAGE completo. Adaptado de [26].....	10
Figura 5. Representação esquemática do papel do RAGE e seus ligantes (especialmente calgranulinas S100) no sítio da resposta imune inflamatória.....	13
Figura 6. S100A12 é liberada tanto no tecido inflamado quanto na corrente sanguínea por neutrófilos ativados. S100A12 secretada liga-se ao RAGE em células endoteliais. Isto leva a maior expressão de ICAM-1 e VCAM-1. As propriedades quimiotáticas adicionais da S100A12 contribuem para aumentar o extravasamento de granulócitos e monócitos para locais da inflamação [20].....	14
Figura 7. RAGE converte a ativação celular aguda em uma resposta celular sustentada. Ativação de NF- κ B por uma variedade de estímulos resulta no aumento da expressão de genes regulados por este fator de transcrição, por exemplo diferentes genes envolvidos na resposta imune e inflamatória incluindo IL-1 β , TNF- α , IL-2, IL-6, IL-8 e IL-12 [31]. Também, o promotor do RAGE é ativado pelo próprio NF- κ B, de forma que a ativação do RAGE leva ao aumento da expressão deste receptor, em um ciclo de “feedback” positivo, fornecendo um número crescente de sítios para a interação de ligantes com o RAGE. A interação destes ligantes, incluindo as calgranulinas, leva à perpetuação da resposta inflamatória.....	15

Figura 8. Ativação dos neutrófilos por quimioatraentes como FMLP (Peptídeo N-formilmetionil- leucil-fenilalanina), PAF (Fator Ativador de Plaquetas), C5a (Componente C5a do Sistema Complemento) e LTB4 (Leucotrieno B4) aumentam a concentração de Ca^{2+} citosólico com a resultante geração de espécies reativas de oxigênio (ROS) e liberação de proteases, metaloproteinases (MMPs) e mediadores lipídicos. A injúria tissular que pode ser associada com a liberação dessas moléculas altamente danosas no tecido contribui na patogênese de um número grande de doenças [35].....	18
Figura 9. Dímero da S100A12 humana (PDB ID: 1E8A) resolvida por Moroz e colaboradores, 2001 [18]. Visualização das hélices que formam o “EF-hand” pseudo-canônico em azul escuro (H_I) e azul claro (H_{II}), e das hélices que formam o “EF-hand” canônico em laranja (H_{III}) e amarelo (H_{IV}). São mostrados também os dois íons de cálcio complexados por monômero (esferas brancas).....	19
Figura 10. Em destaque hélices que formam o “EF-hand” pseudo-canônico em azul escuro (H_I) e azul claro (H_{II}) e hélices que formam o “EF-hand” canônico em laranja (H_{III}) e amarelo (H_{IV})(PDB ID: 1E8A). B. Sequência dos resíduos de aminoácidos que constituem o monômero da S100A12, as linhas pontilhadas indicam os dois “motifs” “EF-hands”.....	20
Figura 11. Detalhe do sítio de complexação do Ca^{2+} (esfera branca) no L_I do EF-1, obtido de uma estrutura resolvida na presença de Ca^{2+} (PDB ID: 1E8A). Os valores de distância de ligação são dados em angstrom.....	20
Figura 12. Detalhe do sítio de complexação do Ca^{2+} (esfera branca) no L_{III} do EF-2, obtido de uma estrutura resolvida na presença de Ca^{2+} (PDB ID: 1E8A). Os valores de distância de ligação são dados em angstrom.....	21
Figura 13. Detalhe do sítio de complexação do Zn^{2+} (esfera cinza) no “motif” His-X-X-X-His entre os monômeros, obtido de uma estrutura resolvida na presença de Zn^{2+} (PDB ID: 2WCB). Os valores de distância de ligação são dados em angstrom.....	21
Figura 14. Detalhe do sítio de complexação do Cu^{2+} (esfera laranja) no “motif” His-X-X-X-His entre os monômeros, obtido de uma estrutura resolvida na presença de Ca^{2+} e Cu^{2+} (PDB ID: 1ODB). Os valores de distância de ligação são dados em angstrom.....	22
Figura 15. Detalhe do sítio de complexação do Na^+ (esfera violeta) no L_I do EF-1, obtido de uma estrutura resolvida na ausência de Ca^{2+} (PDB ID: 2WCE). Os valores de distância de ligação são dados em angstrom.....	23

<p>Figura 16. Algumas das estruturas disponíveis da S100A12 no PDB. A. PDB ID: 1E8A, estrutura resolvida na presença de Ca^{2+}, que complexa dois Ca^{2+} por monômero (esferas brancas). B. PDB ID: 2WCE, estrutura resolvida na ausência de Ca^{2+}, que complexa dois Na^+ por dímero no L_I dos EF-1 (esferas violetas). C. PDB ID: 1ODB, estrutura resolvida na presença de Ca^{2+} e Cu^{2+}, que complexa dois Ca^{2+} por monômero (esferas brancas) e dois Cu^{2+} por dímero na sequência His-X-X-X-His (esferas laranjas). D. PDB ID: 1GQM, estrutura resolvida na presença de Ca^{2+}, observa-se a formação de uma estrutura de hexâmero e a complexação de Ca^{2+} nos <i>loops</i> L_I e L_{III} e numa região de interface entre os dímeros (esferas brancas). E. PDB ID: 2WCB, estrutura resolvida na ausência de Ca^{2+}, que complexa dois Zn^{2+} por dímero na sequência His-X-X-X-His (esferas cinzas) e dois Na^+ por dímero no L_I dos EF-1 (esferas violetas).....</p>	24
<p>Figura 17. Alinhamento das estruturas da 1E8A (em preto) e 2WCE (em cinza). Podemos observar que na presença de Ca^{2+} (1E8A) as hélices do EF-2 estão em uma conformação aberta, enquanto na ausência de Ca^{2+} (2WCE) estão em uma conformação dita fechada...</p>	25
<p>Figura 18. Entrada e saída de simulações moleculares.....</p>	29
<p>Figura 19. Condições periódicas de contorno em duas dimensões [50].....</p>	36
<p>Figura 20. Esquema geral dos sistemas simulados nas baterias 1A e 1B. Foram realizadas simulações da 2WCE (S100A12 resolvida na ausência de Ca^{2+}, que complexa dois íons Na^+ nos EF-1 de cada monômero), submetida a concentrações 100, 75, 50 e 25mM de NaCl, CaCl_2 e ZnCl_2 em sistemas separados.....</p>	46
<p>Figura 21. Deformação estrutural da 2WCE na presença de íons, de acordo com os valores de RMSD. Em destaque na figura, é mostrado o RMSD médio no intervalo entre 50 e 70ns de simulação. As barras de erros correspondem aos desvios padrão dos respectivos valores médios.....</p>	47
<p>Figura 22. RMSF para as cadeias A e B. Somente as regiões N e C-terminais apresentam comportamento similar. No que se refere a cadeia A, gráfico à esquerda, destacam-se os fatos de que na presença de CaCl_2, linha vermelha, observamos flutuações menores, quando comparadas às flutuações na presença de ZnCl_2, linha verde, enquanto as flutuações provocadas por 100mM NaCl são intermediárias.....</p>	48
<p>Figura 23. Deformação estrutural das hélices e L_{II} ("hinge" - região que conecta os "EF-hands") para cadeias A (coluna à esquerda) e B (coluna à direita) da proteína 2WCE na presença de íons.....</p>	49

Figura 24. Superfície hidrofóbica acessível ao solvente ao longo das simulações a 100mM.....	51
Figura 25. Densidade de visitação dos íons mostrada na estrutura final após 80ns de simulação. Sistemas com 100mM Na ⁺ (esferas vermelhas), Ca ²⁺ (esferas brancas) e Zn ²⁺ (esferas cinzas).....	52
Figura 26. Superfície acessível ao solvente colorida de acordo com o potencial eletrostático, calculado utilizando a ferramenta APBS para o início da simulação com 100mM CaCl ₂ . Apresentamos a visão dos EF-1 das cadeias, à esquerda, enquanto à direita temos uma visão dos EF-2.....	53
Figura 27. Deformação estrutural das hélices e L _{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da proteína 2WCE na presença de diferentes concentrações de NaCl.....	55
Figura 28. Deformação estrutural das hélices e L _{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da 2WCE na presença de diferentes concentrações de CaCl ₂	56
Figura 29. Deformação estrutural das hélices e L _{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da proteína 2WCE na presença de diferentes concentrações de ZnCl ₂	57
Figura 30. RMSF para as cadeias A e B do dímero da 2WCE para diferentes concentrações de Na ⁺ , Ca ²⁺ e Zn ²⁺	59
Figura 31. Conteúdo de estrutura secundária em termos de porcentagem média de alfa-hélices (entre 50 e 70ns das simulações). As barras de erros correspondem aos respectivos valores dos desvios padrão dos valores médios. Entre 50 e 70ns não há variação significativa na porcentagem de hélices, para todos os casos, possibilitando ter seus valores médios comparados entre as simulações realizadas. O cálculo da composição de estrutura secundária da proteína ao longo das simulações foi feito em intervalos de 20ps, usando o algoritmo STRIDE. A composição de estrutura secundária correspondente as folhas-beta é mantida constante em 5% (dado não apresentado). A perda de hélice resulta na formação de “coil”. A barra mais escura indica a porcentagem de composição de estrutura secundária do tipo alfa-hélice para a estrutura cristalográfica (PDB ID: 2WCE).....	60

Figura 32. Densidade de visitação dos íons mostrada na estrutura final após 100ns de simulação a 75mM NaCl (Na^+ esferas vermelhas), 50mM CaCl_2 (Ca^{2+} esferas cinzas). Podemos observar que os íons Na^+ , embora aproximem-se da proteína e perturbem a H_{III} na cadeia A, não permanecem ligados à proteína (distribuição homogênea dos íons). Por sua vez o íon Ca^{2+} , uma vez que se aproxima das regiões do L_{III} (do EF-2) por meio da H_{III} na cadeia A, é complexado e permanece ligado ao longo da simulação, embora também perturbe a H_{III} . Notamos novamente também, que a aproximação do Ca^{2+} é realizada pela face dos “loops” L_{III} (EF-2). Destacamos ainda o fato do Ca^{2+} além de permanecer interagindo na H_{III} da cadeia A, também mantém-se complexado à região do L_{II} , o “Hinge-Region”, na cadeia B.....

61

Figura 33. Análise local comparativa da H_{III} e L_{II} em diferentes circunstâncias. **A.** Estrutura cristalográfica da S100A12 (PDB ID: 2WCE) antes de qualquer etapa da simulação. **B.** Visão local da densidade de visitação dos íons Ca^{2+} (esferas brancas) na simulação com 50mM de CaCl_2 e **C.** Visão local da densidade de visitação dos íons Na^+ (esferas vermelhas) na simulação com 75mM de NaCl.....

62

Figura 34. Visão do L_{II} , conectando as hélices H_{III} (à direita) e H_{II} (à esquerda) em diversas condições. Estrutura cristalográfica da S100A12 resolvida na ausência de Ca^{2+} , 2WCE. Estrutura 2WCE após 100ns de simulação na presença de 50mM CaCl_2 . Estrutura cristalográfica da S100A12 resolvida na presença de Ca^{2+} , 1E8A.....

63

Figura 35. Mapa de existência da complexação do Na^+ ao L_I para as cadeias A (indicadas por ChA) e cadeias B (indicadas por ChB) para as simulações com variações de concentrações de NaCl, CaCl_2 e ZnCl_2 . O comprimento da barra horizontal marca o instante em que se dá a descomplexação. A complexação foi analisada por meio do cálculo das distâncias do íon sódio inicialmente complexado em relação ao centro de massa do L_I (levando em conta os resíduos que complexam o sódio inicialmente). Ressaltamos que em nenhum dos casos onde o Na^+ foi descomplexado houve recomplexação do íon.....

64

Figura 36. Configurações em tempos distintos do L_I na cadeia B da simulação com 50mM ZnCl_2 . Embora a saída do Na^+ (esfera azul) ocorra aos 16ns (item F) a abertura do L_I ocorre logo nos primeiros picossegundos de simulação. A presença do Zn^{2+} (esfera cinza) interagindo com o resíduo Asp25, que faz parte do L_I , faz com que este loop se abra, já aos 2ns (item D).....

65

Figura 37. Esquema geral dos sistemas simulados na bateria 2. Foram realizadas simulações da 1E8A (S100A12 resolvida na presença de Ca^{2+} , que complexa dois Ca^{2+} por monômero, nos “ <i>loops</i> ” L_I e L_{III}) submetida a concentrações de 50mM de NaCl, CaCl_2 e ZnCl_2 em sistemas separados.....	67
Figura 38. Deformação estrutural das hélices e L_{II} (“ <i>hinge</i> ” - região que conecta os “EF- <i>hands</i> ”) para cadeias A e B da 1E8A na presença de íons.....	68
Figura 39. RMSF para as cadeias A e B das simulações realizadas.....	69
Figura 40. Conteúdo de estrutura secundária em termos de porcentagem média de alfa-hélices entre 50 e 70ns das simulações. As barras de erros correspondem aos desvios padrão dos valores médios correspondentes. O cálculo da composição de estrutura secundária da proteína ao longo das simulações foi feito em intervalos de 20ps, usando o algoritmo STRIDE. A composição de estrutura secundária correspondente a folhas-beta é mantida constante em 3,5% (dado não apresentado). A perda de hélice resulta na formação de “ <i>coil</i> ”. A barra mais escura indica a porcentagem de composição de estrutura secundária do tipo alfa-hélice para a estrutura cristalográfica (PDB ID: 1E8A).....	70
Figura 41. Alinhamento das estruturas da 1E8A antes e depois de 100ns de simulação na presença de 50mM de NaCl, em cinza e azul, respectivamente. Em B, podemos observar o início da formação de uma hélice (a H_{IIa}) na região “ <i>Hinge</i> ” após 100ns de simulação.....	71
Figura 42. Alinhamento das estruturas da 1E8A antes e depois de 100ns de simulação na presença de 50mM de NaCl_2 , e da 2WCE (S100A12 antes da complexação com o Ca^{2+}) em amarelo, violeta e azul, respectivamente.....	71
Figura 43. Visão do EF-2 composto pelas H_{III} e H_{IV} , na estrutura da 1E8A antes e após 100ns de DM na presença de 50mM de NaCl. No início é possível observar um núcleo hidrofóbico, formado por resíduos da H_{III} , do “ <i>Hinge</i> ” e da H_{IV} . Em A temos as representações das cadeias laterais dos resíduos que formam o núcleo hidrofóbico (Leu35, Leu36, Val40, Ile44, Ile47, Val52, Ile53, Ile56, Ile60, Leu76 e Ile79). Em B temos a representação em “ <i>dots</i> ” das cadeias laterais destes mesmos resíduos. Em C temos também em destaque os resíduos carregado da H_{III} (Lys50, Asp54, Glu55 e Gln58).....	73
Figura 44. Esquema utilizado para o estudo da complexação S100A12-RAGE.....	78
Figura 45. Reclassificação das respostas geradas pelo “ <i>docking</i> ” após minimização de energia no GROMACS em termos de energia potencial de interação proteína-proteína.....	80

Figura 46. Alinhando as estruturas dos 3 melhores complexos disponíveis em termos de energia potencial de interação para cada sistema (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). O RAGE está representado em cinza, enquanto 2WCE, 1E8A e 1GQM estão representadas em verde, azul e rosa (sendo essa representação equivalente as configurações de menores energias de interação de acordo com nossa reclassificação)..... 81

Figura 47. Sistemas simulados (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). Os sistemas são compostos pelas diferentes formas de S100A12, RAGE, moléculas de água (modelo “SPC”) e íons Na^+ e Cl^- , na concentração de 50mM. Os sistemas RAGE+2WCE e RAGE+1E8A são formados por pouco mais de 69000 átomos dos quais 6192 átomos são das proteínas, enquanto 63000 átomos (21000 moléculas) são do solvente. O sistema RAGE+1GQM é formado por cerca de 99000 átomos dos quais 11714 átomos são das proteínas, enquanto 87000 átomos (29000 moléculas) são do solvente..... 82

Figura 48. Deformação estrutural analisada de acordo com os valores de RMSD para os complexos RAGE+2WCE, RAGE+1E8A, RAGE+1GQM..... 83

Figura 49. Exemplos de sistemas submetidos a simulações de SMD (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). Os sistemas são compostos pelas diferentes formas de S100A12, RAGE, moléculas de água (modelo “SPC”) e íons Na^+ e Cl^- , na concentração de 50mM. Os sistemas RAGE+2WCE e RAGE+1E8A são formados por pouco mais de 231000 átomos dos quais 6192 átomos são das proteínas, enquanto 225000 átomos (75000 moléculas) são do solvente. O sistema RAGE+1GQM é formado por cerca de 300000 átomos dos quais 11714 átomos são das proteínas, enquanto 288000 átomos (96000 moléculas) são do solvente..... 85

Figura 50. Imagens de uma das 100 simulações de SMD para os diferentes sistemas (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM) no início, aos 100 e 200ps de simulação..... 86

Figura 51. À esquerda observamos a coleção de perfis de trabalho ao longo da coordenada de reação para as simulações de SMD (estando o trabalho em kJ/mol – como é dado o resultado no GROMACS), enquanto à direita temos a distribuição de frequência normalizada dos trabalhos (já em kcal/mol) ao longo das simulações de SMD..... 87

Figura 52. Interface de complexação mostrada separadamente para os complexos, a saber 1E8A e RAGE e 2WCE e RAGE. Podemos observar que os resíduos envolvidos na complexação do RAGE com a 1E8A (resíduos em laranja) localizam-se principalmente no L_{II} , H_{III} e H_{IV} , enquanto os resíduos envolvidos na complexação do RAGE com a 2WCE (resíduos em laranja) localizam-se em uma porção da H_{III} e H_{IV} . Com relação aos resíduos do RAGE envolvidos na complexação (resíduos em magenta), estão destacados nas

figuras. Podemos observar que a complexação com o RAGE envolve resíduos pertencentes tanto ao domínio V, quanto ao domínio C1. O critério sobre os resíduos envolvidos na complexação baseou-se nas distâncias de interações entre resíduos do RAGE-S100A12 menores que 4Å.....

89

Figura 53. Interface de complexação mostrada separadamente. Podemos observar que a complexação da 1GQM com o RAGE envolve resíduos que formam uma interface entre dois dímeros do hexâmero. Esses resíduos (destacados em laranja) localizam-se principalmente na H_I, L_I, L_{II} e H_{IV}. Os resíduos do RAGE envolvidos na complexação (em magenta), também estão destacados na figura. Podemos observar que a complexação com o RAGE envolve resíduos pertencentes principalmente ao domínio V, e a uma região de “loop” do domínio C1 que também auxilia na complexação do hexâmero da S100A12. O critério sobre os resíduos envolvidos na complexação baseou-se nas distâncias de interações entre resíduos do RAGE-S100A12 menores que 4Å.....

90

Figura 54. Interface de complexação RAGE+1QWM. À esquerda, vemos detalhes do RAGE complexando resíduos da 1GQM (representação de superfície em verde). Ao centro, podemos observar a 1GQM (também em representação em superfície) complexada ao RAGE, estando colorido em laranja as outras duas faces da região entre os dímeros, equivalente a face em verde. E à direita, vemos a 1GQM (também em representação em superfície) complexada ao RAGE, estando colorido em laranja as três faces da região entre os dímeros que são regiões as quais o RAGE pode se complexar.....

91

LISTA DE TABELAS

Tabela 1 – Propriedades dos elementos sódio, cálcio, zinco e cobre.....	23
Tabela 2 – Comparação entre a classificação em termos de menores valores energia potencial de interação proteína-proteína (10 melhores modelos de cada sistema) e a posição na classificação dos mesmos modelos nas respostas fornecidas pelo GRAMM-X.....	81

LISTA DE ABREVIATURAS E SIGLAS

v

AGEs	Produtos Finais de Glicosilação
APBS	<i>“Adaptive Poisson-Boltzmann Solver”</i>
C5a	Componente C5a do Sistema Complemento
CaBP	Proteínas Ligantes de Cálcio
ChA	Cadeia A
ChB	Cadeia B
DAMPs	Padrões Moleculares Associados ao Dano
DM	Dinâmica Molecular
DNA	Ácido desoxirribonucleico
DSSP	<i>“Define Secondary Structure of Proteins”</i>
EF	<i>“EF-hand”</i>
FFT	Transformada Rápida de Fourier
FMLP	Peptídeo N-formilmetionil- leucil-fenilalanina
GRAMM-X	<i>“Global Range Molecular Matching”</i>
GROMACS	<i>“GRONingen MACHine for Chemical Simulation”</i>
HMGB1	Proteínas do Grupo de Alta Mobilidade Box 1
ICAM-1	Moléculas de Adesão Intercelular
IL	Interleucina
IP ₃	Inositol trifosfato
LTB4	Leucotrieno B4
MAPK	Proteína Quinase Ativada por Mitógeno
MBP	Proteína de Ligação à Maltose
m _e	massa elétron
MHC	Complexo Principal de Histocompatibilidade

MMPs	Metaloproteinases
m_p	massa próton
MT	Metalotioneína
NADPH	Nicotinamida Adenina Dinucleotídeo Fosfato reduzida
NF-kappa B	Fator de Transcrição Nuclear kappa B
OPSL-AA	<i>“Optimized Potentials for Liquid Simulations - All Atom”</i>
PAF	Fator Ativador de Plaquetas
PAMPs	Padrões Moleculares Associados aos Patógenos
PBC	Condições Periódicas de Contorno
PDB	Banco de Dados de Proteínas
PME	<i>“Particle Mesh Ewald”</i>
PRRs	Receptores de Reconhecimento de Padrões
RAGE	Receptor para Produtos Finais de Glicosilação
RMN	Ressonância Magnética Nuclear
RMSD	Desvio da Raíz Média Quadrática
RMSF	Flutuação da Raíz Média Quadrática
ROS	Espécies Reativas de Oxigênio
SAS	Superfície Acessível ao Solvente
SMD	<i>“Steered Molecular Dynamics”</i>
SOD	Superóxido desmutase
SPC	<i>“Simple Point Charge”</i>
STRIDE	<i>“STRuctural IDentification”</i>
TLRs	Receptores do tipo Toll
TNF- α	Fator de Necrose Tumoral – α
VCAM-1	Moléculas de Adesão da Célula Vascular

LISTA DE SÍMBOLOS

α	letra grega alfa, encontrada em “ α -hélice”
β	letra grega beta, encontrada em “folha β ”
γ	letra grega gama, representa fase do ângulo diedro
Δ	letra grega delta, usualmente utilizada como símbolo de variação
ε	letra grega epsilon, representando um parâmetro da função de Lennard Jonnes
ζ	letra grega zeta, indicando tempo de simulação
σ	letra grega sigma, representando um parâmetro da função de Lennard Jonnes
ϕ	letra grega phi, representa ângulo diedro
C_{α}	carbono alfa
$\text{\AA}$	Angstrom

SUMÁRIO

Resumo	i
Abstract	ii
Lista de figuras	iii
Lista de tabelas	iv
Lista de abreviaturas e siglas	v
Lista de símbolos	vi
 CAPÍTULO 1	
Contexto fisiopatológico e detalhes estruturais da proteína Calgranulina C.....	1
1.0 Cálcio como mensageiro secundário em transduções de sinais.....	2
1.1 Proteínas ligantes de cálcio (CaBP – “calcium binding protein”).....	3
1.1.1 O “motif” “EF-hand”	4
1.1.2 Proteínas da família S100.....	6
1.2 Calgranulina C (S100A12).....	9
1.2.1 Um receptor para a S100A12.....	9
1.3 Considerações relevantes sobre a resposta imune mediada por células.....	11
1.3.1 Considerações relevantes sobre a resposta imune mediada por células..	11
1.3.2 Consequências patológicas da interação S100A12-RAGE.....	13
1.4 Detalhes moleculares da interação S100A12-RAGE.....	16
1.5 O microambiente fisiopatológico da S100A12.....	16
1.6 Detalhes estruturais da S100A12.....	19
1.6.1 A biblioteca de estruturas disponíveis.....	24
 CAPÍTULO 2	
Uma abordagem utilizando simulação computacional: dinâmica molecular....	27
2.0 Uma abordagem utilizando simulação molecular computacional.....	28
2.1 Dinâmica molecular.....	29
2.1.1 Campo de força.....	31
2.1.2 Condições iniciais.....	33
2.1.3 Otimização da geometria molecular.....	33
2.1.4 Integração das equações de movimento.....	34
2.1.5 Aproximações para representação do estado termodinâmico.....	35
2.1.6 Análises das trajetórias.....	38
2.2 GROMACS – O pacote de simulação.....	41
2.2.1 Etapas da simulação de dinâmica molecular.....	41
 CAPÍTULO 3	
Dinâmica conformacional da proteína S100A12 mediante interação com íons	42
3.1 Conjunto de parâmetros definidos para simulações.....	43
3.2 Bateria 1.....	44

3.2.1 BATERIA 1A – Comportamento da estrutura da S100A12 resolvida na ausência de Ca^{2+} (PDB ID: 2WCE) mediante presença de 100mM de íons Na^+ , Ca^{2+} , Zn^{2+} tendo como contra-íon o Cl^-	45
3.2.2 BATERIA 1B – Comportamento da estrutura da S100A12 resolvida na ausência de Ca^{2+} (PDB ID: 2WCE), mediante presença de diferentes concentrações de íons - análise da Influência da variação da concentração de íons (25mM, 50mM e 75mM de íons Na^+ , Ca^{2+} , Zn^{2+} tendo como contra-íon o Cl^-).....	45
3.2.3 Resultados e análises das simulações: Bateria 1.....	46
3.3 Bateria 2.....	67
3.3.1 BATERIA 2 – Análise do comportamento da estrutura da S100A12 resolvida na presença de Ca^{2+} (PDB ID: 1E8A) mediante presença de 50mM de íons Na^+ , Ca^{2+} , Zn^{2+} tendo como contra-íon o Cl^-	67
3.3.2 Resultados e análises das simulações: Bateria 2.....	67
3.4 Conclusões acerca da dinâmica conformacional da proteína Calgranulina C (S100A12) mediante interações com íons.....	74
 CAPÍTULO 4	
Estudo da interação da proteína Calgranulina C (S100A12) com o receptor RAGE.....	76
4.1 Detalhes sobre a interação da S100A12 com o receptor RAGE.....	77
4.2 Estratégia para o estudo da interação RAGE-S100A12.....	78
4.2.1 “Docking” molecular para gerar modelos dos complexos.....	78
4.2.2 Reclassificação e seleção dos melhores modelos.....	79
4.2.3 Simulações de DM dos melhores modelos.....	82
4.2.4 Simulações de SMD dos melhores modelos para obtenção de detalhes da complexação.....	83
4.3 Conclusões sobre a interação RAGE-S100A12.....	91
 REFERÊNCIAS BIBLIOGRÁFICAS.....	 93

CAPÍTULO 1

Contexto Fisiopatológico e
Detalhes Estruturais da Proteína Calgranulina C (S100A12)

Comunicação Celular

Em um organismo multicelular, as células precisam coordenar suas funções de várias maneiras. Um aspecto importante desta coordenação consiste nos mecanismos relacionados à comunicação celular. Evidente que se esta comunicação celular falhar em alguma instância, há comprometimento do processo bioquímico envolvido, o que pode desencadear eventos patológicos graves e variados.

A recepção de um sinal inicia no ponto onde o agonista, eventualmente, proveniente de fora da célula, encontra a molécula-alvo que pertence à própria célula. Em praticamente todos os casos, a molécula-alvo é um receptor protéico, normalmente ativado por um tipo de sinalizador. O receptor protéico desempenha o primeiro passo na transdução: ele recebe o sinal externo e gera um novo sinal intracelular como resposta. Via de regra, este é somente o primeiro evento numa cadeia subsequente de processos intracelulares de transdução de sinal. Nestes, a mensagem é passada de um conjunto de moléculas de sinalização intracelulares a outro, cada uma provocando a produção da próxima até, por exemplo, que uma enzima metabólica seja ativada, ou a expressão de um gene seja desencadeada ou que o citoesqueleto seja posto em ação. Este resultado final é a resposta da célula e estas cadeias de propagação de moléculas de sinalização intracelulares, ou cascatas de sinalização, têm várias funções cruciais [1]. Embora o número de sinais biológicos seja enorme, como o é a variedade de respostas biológicas a esses sinais, os organismos usam apenas alguns poucos mecanismos evolucionários para detectar sinais extracelulares e transformá-los em alterações intracelulares. Um desses mecanismos envolve a participação de íons no complexo papel de mensageiros intracelulares, especialmente o cálcio (Ca^{2+}), executando a tarefa fundamental de transportar sinais de um grande número de atividades bioquímicas nos diferentes compartimentos subcelulares.

1.0 Cálcio como mensageiro secundário em transduções de sinais

Em muitas células que respondem a sinais extracelulares, o Ca^{2+} funciona como um mensageiro secundário que desencadeia respostas intracelulares como a exocitose nos neurônios e nas células endócrinas, a contração no músculo ou os rearranjos no citoesqueleto durante o movimento amebóide.

Normalmente, a concentração citosólica de Ca^{2+} é mantida muito baixa ($< 10^{-7}\text{M}$) pela ação das bombas de Ca^{2+} no retículo endoplasmático, na mitocôndria e na membrana plasmática. Determinados estímulos levam ao influxo de Ca^{2+} para dentro da célula por meio de canais de Ca^{2+} específicos na membrana plasmática ou uma liberação do Ca^{2+} sequestrado no retículo

endoplasmático ou na mitocôndria, em ambos os casos elevando a concentração Ca^{2+} citosólica e desencadeando uma resposta celular.

Uma importante diferença com relação à concentração de Ca^{2+} nos meios extra e intracelulares, refere-se à porção de cálcio ionizado. Enquanto no meio extracelular a porção ionizada corresponde à metade da quantidade total, no meio intracelular, esta mesma porção corresponde a somente 0,1% ou menos do total [2].

Esta diferença marcante na concentração de cálcio ionizado é facilmente justificada quando imaginamos que, ao utilizar este íon como mensageiro intracelular, tornou-se necessário que as células desenvolvessem meios para manter a concentração intracelular de Ca^{2+} muito baixa, para que variações pequenas na concentração pudessem desencadear a sinalização necessária para os eventos bioquímicos celulares [3].

A concentração de outros mensageiros secundários como nucleotídeos cíclicos, inositol trifosfato, ou diacilglicerol, pode ser facilmente modificada por síntese e degradação metabólica. Como isso não é possível para o cálcio, mecanismos diferentes foram desenvolvidos para controlá-lo dentro das células. Sabe-se atualmente, que as células efetuam essa regulação por meio da formação de complexos reversíveis de cálcio com ligantes específicos. Considerando a alta especificidade requerida, ligantes complexos são necessários. Como é regra para situações biológicas que exigem um alto nível de complexidade estrutural, moléculas protéicas são utilizadas neste caso. Assim, alterações na concentração Ca^{2+} intracelular são detectados por proteínas que se ligam ao cálcio e regulam uma variedade de outras proteínas [3].

1.1 Proteínas ligantes de cálcio (CaBP – “*calcium binding protein*”)

As proteínas que se ligam ao Ca^{2+} e estão envolvidas na regulação e na transferência do sinal podem ser divididas em dois grupos, as proteínas solúveis (globulares), e as proteínas ancoradas na membrana. Ambos os tipos contribuem para o tamponamento de Ca^{2+} na célula, mas é evidente que a quantidade de cálcio que pode ser armazenada pelas proteínas solúveis no citosol é limitada. Por outro lado, esta limitação não se aplica às proteínas transmembranares, as quais ligam cálcio e o transportam para o interior da célula, retornando a um estado no qual podem ligar outro íon novamente, caracterizando um transporte contínuo.

As proteínas ancoradas na membrana constituem a família das anexinas, que interagem com fosfolipídios nas membranas celulares, de forma dependente do cálcio. Estas proteínas são conhecidas como lipocortinas, calcimedinas ou calpactinas [4].

Com relação às proteínas solúveis, estas formam complexos com íon cálcio, que têm como papel principal a transferência do sinal, para que este seja traduzido em respostas mecânicas e

metabólicas, via interação com efetores celulares em um modelo dependente de Ca^{2+} . O ingrediente essencial nesta transferência do sinal é a mudança de conformação protéica, por meio do aumento da exposição de sítios hidrofóbicos e variação do conteúdo de α -hélice, mediante a ligação com o Ca^{2+} . Nessa nova conformação, as proteínas podem também interagir com enzimas alvo no interior da célula [5]. As proteínas desse grupo apresentam propriedades comuns: (i)- têm pequenas massas; (ii)- são carregadas negativamente em pH neutro; (iii)- contêm normalmente alto teor de aminoácidos ácidos; e (iv)- ligam de dois a quatro Ca^{2+} com alta afinidade.

Dentre as proteínas moduladas pelo cálcio incluem-se a calmodulina, troponina-C, parvalbumina, oncomodulina e proteínas da família S100. Elas exibem uma estrutura principal, o “motif” “EF-hand”, que será discutido adiante.

1.1.1 O “motif” “EF-hand”

O termo “EF-hand” foi criado por R.H Kretsinger, mediante estudo da parvalbumina, uma pequena proteína ligante de cálcio isolada do músculo de carpa [6]. Esta denominação designa um “motif” do tipo hélice-“loop”-hélice, que apresenta alta afinidade pelo Ca^{2+} , sendo que as letras E e F referem-se as hélices E e F da parvalbumina. O clássico “motif” “EF-hand” é geralmente formado por um “loop” de 10-12 resíduos de aminoácidos, ladeados por duas hélices apontando para direções opostas e perpendiculares entre si (Figura 1).

Figura 1. Representação esquemática do “motif” “EF-hand” [7].

O “loop” é o local adequado para ligação com o cálcio. Este sítio pode ser representado como: $\text{X}\bullet\text{Y}\bullet\text{Z}\bullet\text{Y}\bullet\text{X}\bullet\bullet\text{Z}$, onde X, Y, Z, $-\text{X}$, $-\text{Y}$, e $-\text{Z}$ são os ligantes que participam na coordenação do metal e os pontos ($\bullet$) representam os resíduos entre eles (Figura 2) [8].

A sequência de resíduos que compõem o “EF-hand” forma um “loop” que pode acomodar ligantes cálcio ou magnésio, diferindo apenas na formação geométrica, que depende do íon presente. De forma geral, sete ligantes nos vértices de uma pirâmide pentagonal coordenam o cálcio, como visto adiante (Figura 2).

Figura 2. A e B representam a geometria de ligação do íon cálcio no domínio “EF-hand”, sendo que o íon ligado fica no centro da figura. Nas posições X e Y, encontram-se normalmente as cadeias laterais de ácido aspártico ou asparagina; na posição Z encontram-se as cadeias laterais, também, de ácido aspártico ou asparagina, ou ainda, serina; na posição -Y encontra-se o oxigênio da carbonila do peptídeo. -X normalmente é uma molécula de água e -Z é a cadeia lateral de um resíduo como ácido glutâmico ou ácido aspártico. Adaptado de [8].

Os “motifs” “EF-hand” normalmente ocorrem em pares na mesma cadeia, formando um domínio de quatro hélices; o pareamento desses “motifs” ativa cooperativamente a ligação de Ca^{2+} , sendo essencial para a geração de uma resposta clara mesmo sob pequena variação na concentração de cálcio durante a sinalização [9]. As α -hélices são de fundamental importância para a estabilização da proteína e para a interação com moléculas alvo, pois atuam como âncoras sobre o “loop”, mantendo-o na conformação correta. Tal conformação estrutural permite a transmissão de eventuais mudanças conformacionais geradas por meio da ligação dos íons para as demais regiões da proteína. [10].

As proteínas do tipo “EF-hand” podem apresentar de dois a oito “motifs” potenciais, embora nem todos sejam funcionais. Além disso, são divididas em aproximadamente 66 subfamílias, apresentando funções tão diversas que incluem desde o tamponamento do cálcio no citosol e a transmissão de sinal entre os compartimentos celulares até a contração muscular [11].

Estudos estruturais têm revelado diferenças substanciais nas mudanças conformacionais induzidas por Ca^{2+} entre os domínios “EF-hand”. Em proteínas pertencentes à superfamília “EF-

hand”, os “*motifs*” podem ser Ca^{2+} - específicos ou ligar Ca^{2+} e também Mg^{2+} ; os sítios estruturais e de tamponamento ligam tanto Ca^{2+} quanto Mg^{2+} [12]. Variações estruturais na forma de organização dos domínios “EF-*hand*” contribuem significativamente para a seletividade de moléculas alvo e para diversidade funcional das proteínas ligantes de cálcio.

Com relação aos genes envolvidos na síntese destas proteínas, acredita-se que todas as proteínas “EF-*hand*” estejam envolvidas com um ancestral único do “*motif*” “EF-*hand*”. Contudo, durante a evolução, reiteraões múltiplas desse gene principal deram origem a genes que codificam proteínas com múltiplos “EF-*hands*”. A posterior duplicação gênica levou a uma grande família de proteínas. Juntamente com o crescimento do número, funções e padrões de expressão também divergiram. Por meio de mutações no “*loop*” “EF-*hand*”, alguns domínios perderam a capacidade de se ligarem ao cálcio. A aquisição de promotores diferentes permitiu que outros domínios pudessem ser expressos de maneira ubíqua (por exemplo, a calmodulina), enquanto a maioria dos demais tecidos mostra um padrão específico de expressão. Finalmente, foram criadas proteínas híbridas, contendo “EF-*hand*” e outros domínios funcionais [4].

Com relação à distribuição destas proteínas, há imensa variedade. A calmodulina, por exemplo, está presente em todas as células, enquanto outras proteínas estão presentes em células de determinados tecidos [13].

1.1.2 Proteínas da família S100

A família das proteínas S100 consiste no maior subgrupo de proteínas “EF-*hand*” que ligam Ca^{2+} . Foram primeiramente identificadas por B. W. Moore em 1965 e são chamadas S100 por serem solúveis em solução de sulfato de amônio 100% saturada, em pH neutro [14].

As proteínas S100 foram inicialmente purificadas do cérebro de bovinos e subsequentemente foi demonstrado que elas são dímeros de subunidades α e β . A sequência primária de ambas as subunidades contém dois “EF-*hands*” e compartilham de identidade superior à 50% na constituição de aminoácidos.

Foram identificadas aproximadamente 25 membros nesta família em humanos, sendo que os genes das S100 estão ausentes em invertebrados. Os membros da família das S100 estão envolvidos com múltiplas funções e são expressos em ampla faixa de tecidos diferentes. São associadas a diversas respostas fisiológicas, através de moléculas alvos; podendo promover o crescimento e diferenciação de células, regulação do ciclo celular, transcrição, organização estrutural de membranas, proteção contra dano oxidativo celular, atividade de receptores de superfície celular, fosforilação protéica e secreção. Essa variedade de funções ocorre pelos seguintes fatores:

- Ampla diversificação dos diferentes membros desta família (25 em humanos);
- Diferentes propriedades de ligação de íons metálicos nas proteínas S100;
- Distribuição específica nos espaços extra e intracelulares, e;
- Capacidade de formação de homo- e heterodímeros não covalentes, permitindo a mudança da dinâmica das subunidades das proteínas S100.

Do ponto de vista fisiológico, as proteínas S100 não têm atividade catalítica intrínseca, elas são elaboradas para serem proteínas sensores de cálcio, como ocorre com a calmodulina e a troponina C, sofrendo mudanças conformacionais e modulando atividade biológica via ligação com o cálcio [15].

Do ponto de vista estrutural, as proteínas S100 são pequenas, apresentando de 10 a 12kDa e contendo duas regiões distintas de “EF-hands”. O “motif” “EF-hand” N-terminal (EF-1) é composto por duas α -hélices (H_I e H_{II}) conectadas por um “loop” de ligação ao Ca^{2+} (L_I); o segundo “motif” “EF-hand” (EF-2) é constituído pelas hélices H_{III} e H_{IV} , conectadas por um “loop” de ligação ao Ca^{2+} (L_{III}). Os dois “EF-hands” são conectados por uma região de comprimento variável, chamado “Hinge Region” ou “loop” L_{II} [16] (Figura 3).

Figura 3. Representação estrutural completa de uma proteína da família S100.

O “motif” “EF-hand” N-terminal (EF-1), também denominado “EF-hand” pseudo-canônico, é característico das proteínas S100, sendo constituído por 31 resíduos, dos quais 14 compreendem a região do L_I , que constitui o local de complexação do Ca^{2+} . Nesse “motif”, o Ca^{2+} é complexado principalmente pelos átomos de oxigênio da cadeia principal protéica, embora saiba-se que o EF-1 tem baixa afinidade pelo cálcio quando comparado à calmodulina e à troponina C e também, quando comparado ao EF-2 das próprias proteínas S100. Por outro lado, o “motif” “EF-hand” C-terminal, “EF-hand” canônico, é muito similar aos “EF-hands” encontrados na calmodulina, troponina C e parvalbumina. O EF-2 é constituído por 29 resíduos, dos quais 12 compreendem a região do L_{III} , que constitui o local de complexação do Ca^{2+} . Nesse “motif”, o Ca^{2+} é complexado

principalmente por átomos da cadeia lateral protéica, e sabe-se que o EF-2 tem alta afinidade pelo Ca^{2+} . As proteínas S100 compartilham de grande similaridade com relação à sequência de resíduos, sendo que os segmentos menos conservados constituem a região C-terminal e a região que conecta os dois “*motifs*” “EF-hands” (L_{II}). Assume-se que estas regiões possam ser responsáveis pela especificidade da função de cada proteína do tipo S100 [17].

A proposta preponderante atualmente é a de que as proteínas S100 comportem-se de forma análoga à calmodulina no que se refere aos eventos que envolvem a complexação do Ca^{2+} . Para a calmodulina, a complexação do Ca^{2+} leva à exposição de uma região hidrofóbica, que consiste numa superfície de reconhecimento. Estudos de Ressonância Magnética Nuclear (RMN), que investigam detalhes estruturais, fornecem evidências de que essa mesma hipótese da calmodulina possa ser aplicada às proteínas S100, embora se saiba que a natureza das mudanças conformacionais e a localização dessa superfície de reconhecimento sejam significativamente diferentes do que têm sido determinado para a calmodulina.

Além disso, apesar do fato do EF-1 apresentar menos afinidade para o Ca^{2+} , um modelo para as alterações estruturais promovidas pela ligação do Ca^{2+} às proteínas S100 sugere que o Ca^{2+} primeiramente liga-se ao L_{I} do EF-1, o qual em termos de conformação é mais próximo do estado ligado (maior acessibilidade) ao Ca^{2+} desde o início, não apresentando alterações significativas com a chegada deste íon. O L_{III} do EF-2 no estado em que a proteína não está complexada ao Ca^{2+} apresenta-se muito mais desordenado. O modelo em questão sugere que a complexação do Ca^{2+} ao L_{I} (EF-1) leva a uma comunicação entre os “EF-hands” que causa a reorganização do L_{III} (EF-2). Quando o Ca^{2+} é complexado, o EF-2 perde sua flexibilidade e isso resulta em uma reorientação da hélice H_{III} que é fixada em uma conformação que expõe uma região hidrofóbica, mas que também apresenta algumas regiões carregadas negativamente. De forma geral, a complexação do Ca^{2+} faz com que ocorra um rearranjo das hélices das proteínas S100, revelando uma abertura, que constitui o alvo de ligação da proteína [18].

Uma característica adicional apresentada pelas proteínas S100 é a afinidade que podem apresentar por outros metais bivalentes como os íons Zn^{2+} e Cu^{2+} , sendo que os sítios de ligação para estes são distintos dos sítios ligantes de Ca^{2+} . A ligação com alta afinidade desses íons bivalentes pode desempenhar um papel importante na quimiotaxia e homeostase de íons metálicos tóxicos para células. Algumas proteínas da família S100 são influenciadas pelo íon Zn^{2+} , apresentando sua afinidade pelo Ca^{2+} afetada [19].

Os detalhes dos modelos existentes sobre alterações estruturais promovidas pela complexação de íons às proteínas S100 constituem um ponto relevante no presente trabalho, uma

vez que nosso objetivo é contribuir para a elucidação do panorama de alterações conformacionais promovidas por esses íons que se complexam a uma proteína do tipo S100. Deve-se ressaltar ainda que, devido à ampla ocorrência das proteínas S100 em tecidos diversos, bem como à alteração da expressão desta família em condições patológicas variadas e relevantes, torna-se muito importante o estudo molecular dos mecanismos de ativação destas proteínas mediante interação com íons, a fim de aumentar a compreensão sobre os eventos celulares nos quais estão envolvidas.

1.2 Calgranulina C (S100A12)

No presente trabalho, o estudo proposto tem como alvo uma proteína da família das S100, chamada Calgranulina C (S100A12). A S100A12, juntamente com a S100A8 (Calgranulina A) e S100A9 (Calgranulina B), formam um subgrupo de proteínas S100, denominado subfamília das Calgranulinas, sendo expressas principalmente em granulócitos e sendo relativamente abundantes neste tipo de leucócitos [20]. Estas proteínas estão envolvidas com o papel destas células na resposta imune relacionada à inflamação em mamíferos. A S100A12 humana compartilha 40 e 46% de identidade de sequência com S100A8 e S100A9 humanas, respectivamente [21].

A S100A12 humana, primeiramente caracterizada por Guinard e colaboradores em 1995, foi descoberta em células do sangue e foi estimado que constitui cerca de 5% do conteúdo proteico total do citosol de neutrófilos. Essa proteína também tem sido encontrada em pequenas quantidades em lisados de monócitos e linfócitos [22]. A S100A12 tem homólogos em granulócitos porcinos e no fluido amniótico de bovinos, com os quais compartilha uma identidade de sequência de 66 e 70%, respectivamente [23].

Como a maioria das proteínas S100, a S100A12 é um dímero em solução (detalhes estruturais serão discutidos adiante), e apresenta funções nos ambientes intra e extracelulares durante a infecção e a inflamação. No meio intracelular a calgranulina C exerce função sinalizadora, mas também pode ocorrer no espaço extracelular, onde atua como citocina. Assim, foi identificado um receptor de superfície celular para a S100A12, o RAGE (*“Receptor for Advanced Glycation Endproducts”* - Receptor para Produtos Finais de Glicosilação) [24].

1.2.1 Um receptor para a S100A12

O receptor de superfície identificado para S100A12, quando esta encontra-se no ambiente extracelular, é o RAGE, sendo o mesmo indicado também como um receptor putativo para outras

proteínas S100. O RAGE é um membro da superfamília das imunoglobulinas de peso molecular de 35kDa, que age como mediador de sinalização intracelular e do metabolismo celular. Foi identificado pela primeira vez no tecido pulmonar, onde está localizado nas membranas basolaterais das células do epitélio alveolar. O gene para o RAGE está localizado no cromossomo 6, próximo ao gene do complexo principal de histocompatibilidade (MHC) em humanos e camundongos, e na proximidade dos genes que codificam TNF (*"Tumor Necrosis Factor"*- Fator de Necrose Tumoral) e linfotóxina [25].

A molécula foi nomeada RAGE por ter sido originalmente descrita como um receptor para produtos finais da glicosilação, os AGEs (*"Advanced Glycation Endproducts"* - Produtos Finais de Glicosilação). AGEs são produtos de glicosilações não enzimáticas e oxidação de proteínas, lipídios e outras macromoléculas que aparecem especialmente em condições nas quais há maior disponibilidade de açúcares redutores, ou em estado de estresse oxidativo avançado.

O RAGE é composto por uma porção extracelular (314 resíduos de aminoácidos), uma porção transmembranar (27 resíduos de aminoácidos) e um curto domínio citosólico (41 resíduos de aminoácidos). A parte extracelular do RAGE contém um domínio variável – V e dois domínios constantes – C1 e C2 (Figura 4). O curto domínio citosólico, chamado cauda citosólica, não tem *"motifs"* conhecidos de sinalização, como sítios de fosforilação e domínios quinase. Foi evidenciado que a cauda citosólica é essencial para a transdução de sinal do RAGE, porque uma forma do RAGE na qual se tenha deletado a cauda citosólica é capaz de interagir com ligantes, como o tipo selvagem, mas não é capaz de mediar qualquer ativação celular [26].

Figura 4. Representação esquemática do RAGE, mostrando os domínios do receptor. A parte extracelular do RAGE é chamada RAGE-solúvel (sRAGE), enquanto o conjunto formado pela parte extracelular mais o domínio transmembranar é chamado DN-RAGE. As porções extracelular, transmembranar e intracelular formam o RAGE completo. Adaptado de [26].

O RAGE é expresso tanto constitutivamente, quanto pode ser induzido dependendo do tipo e do estágio de desenvolvimento da célula. O RAGE é constitutivamente expresso durante o

desenvolvimento embrionário, mas tem expressão diminuída na vida adulta. Exceções conhecidas para esse fato incluem a pele e os pulmões, que constitutivamente expressam RAGE ao longo da vida. A maioria das outras células estudadas até agora, incluindo monócitos/macrófagos, células endoteliais, células musculares lisas, fibroblastos e células neuronais não expressam quantidades significativas de RAGE em condições fisiológicas, mas podem ser induzidas a expressar RAGE em situações onde mediadores inflamatórios e ligantes se acumulam.

Embora não se compreenda por completo o papel fisiológico do RAGE, é possível que esse receptor se encaixe no conceito de pleiotropia antagonista. Esse conceito postula que genes benéficos durante a fase reprodutiva da vida tornam-se prejudiciais com o passar do tempo e podem estar relacionados ao desenvolvimento de doenças relacionadas à idade [27]. O RAGE tem papel principalmente em doenças crônicas, em particular pelo fato de ter sua expressão aumentada num conjunto de doenças inflamatórias crônicas, como artrite reumatóide, doença renal inflamatória, aterosclerose, doença inflamatória do intestino, entre outras.

O RAGE interage com diversos ligantes, incluindo AGEs, como já foi dito, mas também interage com moduladores de resposta inflamatória, como as S100 calgranulinas e anfoterinas. A ligação ao RAGE converte sinais rápidos de ativação celular em mecanismos duradouros de disfunção celular e destruição tissular, como será detalhado adiante [26].

1.3 Eixo S100A12-RAGE na resposta imune mediada por células

Animais multicelulares precisam distinguir a respeito da integridade de suas células e devem ser capazes de detectar quando microorganismos estão presentes, o que é realizado por meio de mecanismos de vigilância, defesa e reparo, levando à finalização da resposta. Evidentemente, um quadro completo da ativação e regulação desses mecanismos ainda não está definido, embora seja de grande relevância clínica o estudo deste cenário da resposta imune.

O eixo S100A12-RAGE insere-se nesse contexto de defesa, por estar relacionado à amplificação da resposta inflamatória [28]. Adiante veremos um quadro geral da resposta imune mediada por células, antes de estudarmos mais detalhadamente o eixo S100A12-RAGE.

1.3.1 Considerações relevantes sobre a resposta imune mediada por células

Um sistema completo para detecção, contenção e reparo dos danos causados às células no organismo pressupõe a existência de sinais de advertência, fornecidos por células capazes de responder via receptores e caminhos de sinalização, desencadeando uma resposta fisiológica. Classicamente um subconjunto desse sistema é conhecido como PAMPs (*“Pathogen-Associated*

Molecular Patterns”- Padrões Moleculares Associados aos Patógenos) e se constituem em um conjunto diversificado de moléculas microbianas que compartilham um número de características bioquímicas reconhecíveis, que podem ser moléculas inteiras ou, mais frequentemente, parte de moléculas, responsáveis por alertar o organismo para que este possa intervir nos patógenos. PAMPs são reconhecidos pelas células da resposta imune inata e adquirida, e os receptores de PAMPs são chamados PRRs (*“Pattern Recognition Receptor”* - Receptores de Reconhecimento de Padrões). Um PRR particular pode reconhecer um padrão molecular que deve estar presente em uma quantidade de patógenos diferentes permitindo ao receptor reconhecer uma variedade de patógenos diferentes.

Como resultado, algumas células são ativadas para destruir os patógenos e/ou células infectadas pelos patógenos, e a resposta imune é acionada, no sentido de produzir e selecionar receptores específicos de célula T e anticorpos que sejam mais adequados para reconhecer o patógeno em uma ocasião futura. A maioria das respostas desencadeadas por PAMPs enquadra-se nas categorias gerais de inflamação e imunidade [29].

Entretanto, deve-se considerar que patógenos não são os únicos agentes causadores de danos aos tecidos e células. Tecidos podem ser acometidos por traumas, podem ser danificados por queimaduras, torções, substâncias químicas, radiação, de forma que a evolução incorporou em nossa genética maneiras de lidar com esses danos, que embora não sejam causados por patógenos, são bastante comuns. Dessa forma, foi estabelecida uma nomenclatura para diferenciar as moléculas endógenas liberadas por diferentes células quando acometidas por diferentes tipos de estresse e processos necróticos, sendo essas moléculas chamadas alarminas. Elas são capazes de ativar receptores do sistema imune de forma similar a padrões bacterianos.

Alarminas e PAMPs constituem o que se conhece de maneira geral como DAMPs (*“Damage-Associated Molecular Patterns”* - Padrões Moleculares Associados ao Dano), sendo as alarminas equivalentes aos PAMPs, mas com a peculiaridade de serem moléculas endógenas. Alarminas e PAMPs apresentam diversas características, das quais destacam-se: (i)- o fato de serem rapidamente liberadas após a morte celular não programada mas não serem liberadas por células apoptóticas; (ii)- serem produzidas e liberadas também por células do sistema imune sem que estas estejam em processo de morte celular, geralmente utilizando sistemas de secreção especializada ou pela via de secreção do retículo endoplasmático-golgi; (iii)- recrutarem e ativarem células que expressam receptores do sistema imune inato, incluindo células dendríticas, e também direta ou indiretamente promoverem a resposta imune adaptativa; (iv)- finalmente, restaurarem a homeostase através da promoção da reconstrução do tecido que foi destruído tanto pelos danos diretos quanto pelos efeitos secundários do processo inflamatório [30].

Nesse contexto, as Calgranulinas (Proteínas S100 fagocíticas), são um exemplo destas substâncias que atuam como fatores pró-inflamatórios da resposta imune inata (DAMPs). Há evidências de que as proteínas S100 são ativamente secretadas através da via alternativa, ignorando a rota clássica do Golgi. Esse modo de secreção é típico para os DAMPs que têm papel na homeostase da célula como moléculas intra-celulares, mas se transformam em sinais pró-inflamatórios após liberação no compartimento extracelular, devido a danos celulares, infecções, ou inflamação [20].

Um conceito emergente de padrões de reconhecimento envolve o receptor de multiligantes para produtos de glicosilação (RAGE), que é o receptor identificado para proteína S100A12, e os TLRs ("Toll-Like Receptors") na detecção de DAMPs. Nesse sentido, após serem liberadas para o espaço extracelular, as proteínas S100 podem interagir com o RAGE, expresso na membrana de determinados tipos de células, desencadeando uma série de eventos chave na resposta de defesa. A resposta imune inata tem um papel fundamental na defesa do hospedeiro e na inicialização do processo inflamatório. Alvos farmacêuticos envolvidos nos mecanismos da resposta inata têm se mostrado eficazes na maioria dos relevantes distúrbios autoinflamatórios [20].

1.3.2 Consequências patológicas da interação S100A12-RAGE

A proteína S100A12 é também chamada proteína extracelular de ligação ao RAGE recentemente identificada (EN-RAGE). Trata-se de um dos ligantes do RAGE que ativa células endoteliais, macrófagos, células da musculatura lisa e células mononucleares do sangue periférico incluindo células T via RAGE, e então desencadeia a ativação de cascatas de sinalização e geração de citocinas e moléculas de adesão pró-inflamatórias [26] (figura 5).

Figura 5. Representação esquemática do papel do RAGE e seus ligantes (especialmente calgranulinas S100) no sítio da resposta imune inflamatória. A hipótese proposta [26] é a de que a interação S100-RAGE leva à amplificação da resposta inflamatória por mediar a ativação de células que expressam RAGE, incluindo fagócitos mononucleares, linfócitos e células do endotélio vascular. Embora o RAGE e seus ligantes pareçam não estar envolvidos no início da resposta do hospedeiro, a qual está relacionada à resposta primária, estímulo inflamatório, ou infeccioso, o eixo S100-RAGE está diretamente relacionado à perpetuação do estímulo inflamatório e portanto às doenças relacionadas à inflamação crônica.

A interação com o receptor RAGE representa uma função crucial em inflamações relacionadas à S100A12. É muito provável que a S100A12, bem como algumas outras proteínas S100, pertença a uma nova classe de quimiocinas e contribua para a migração celular. Sabe-se que a interação da S100A12 com RAGE resulta na secreção de mediadores próinflamatórios.

S100A12 tem sido descrita como responsável pelo aumento da expressão de moléculas de adesão intercelular (ICAM-1) e moléculas de adesão da célula vascular (VCAM-1) no endotélio. Em humanos, S100A12 pode atuar como um mediador próinflamatório ligando-se e ativando células que expressam RAGE [20] (Figura 6).

Figura 6. S100A12 é liberada tanto no tecido inflamado quanto na corrente sanguínea por neutrófilos ativados. S100A12 secretada liga-se ao RAGE em células endoteliais. Isto leva a maior expressão de ICAM-1 e VCAM-1. As propriedades quimiotáticas adicionais da S100A12 contribuem para aumentar o extravasamento de granulócitos e monócitos para locais da inflamação [20].

A ativação do RAGE leva a inicialização de NF-kappa B (*Factor Nuclear Kappa B* - Fator de Transcrição Nuclear kappa B) e ativação das vias que envolvem a proteína MAPK (*Mitogen-Activated Protein Kinase* - Proteína Quinases Ativada por Mitógeno). Em contraste com outros receptores, a ativação celular mediada por RAGE inclui o aumento da expressão do próprio RAGE. Esse ciclo de *“feedback”* positivo caracterizado pela interação do receptor com o ligante seguido pelo aumento da expressão do próprio receptor, sugere para o RAGE o papel de fator de propagação e perpetuação do processo inflamatório [31] (Figura 7).

Figura 7. RAGE converte a ativação celular aguda em uma resposta celular sustentada. Ativação de NF-κB por uma variedade de estímulos resulta no aumento da expressão de genes regulados por este fator de transcrição, por exemplo diferentes genes envolvidos na resposta imune e inflamatória incluindo IL-1β, TNF-α, IL-2, IL-6, IL-8 e IL-12 [31]. Também, o promotor do RAGE é ativado pelo próprio NF-κB, de forma que a ativação do RAGE leva ao aumento da expressão deste receptor, em um ciclo de “*feedback*” positivo, fornecendo um número crescente de sítios para a interação de ligantes com o RAGE. A interação destes ligantes, incluindo as calgranulinas, leva à perpetuação da resposta inflamatória.

Baseado no que foi discutido nesse item, de forma geral, a participação do eixo S100A12-RAGE na resposta do hospedeiro se inicia com a liberação da S100A12 tanto no tecido inflamado quanto na corrente sanguínea após a existência de determinado estímulo inflamatório ou infeccioso por neutrófilos ativados. Nestas circunstâncias, algumas células, como células endoteliais, macrófagos, células da musculatura lisa e células mononucleares do sangue periférico incluindo células T, aumentam a expressão de RAGE. A interação S100A12-RAGE desencadeia uma cascata de eventos, anteriormente detalhada, responsável pela perpetuação do estímulo inflamatório, que inclui inclusive o aumento da expressão do próprio RAGE, em um ciclo de “*feedback*” positivo. Além disso, a S100A12 leva ao aumento da expressão de moléculas de adesão (ICAM-1 e VCAM-1) no endotélio, contribuindo diretamente para o aumento do extravasamento de granulócitos e monócitos para locais da inflamação, de forma a reforçar a resposta de defesa. O recrutamento e ativação excessiva destas células é o mecanismo patofisiológico de muitas doenças, como artrite reumatóide, doença inflamatória intestinal crônica, arteriosclerose, glomerulonefrites e vasculites [31].

Análise minuciosa das bases moleculares dos efeitos específicos da interação entre S100A12 e os íons por ela complexados, bem como da interação S100A12-RAGE, podem nos orientar no esclarecimento de importantes mecanismos da imunidade inata. As proteínas S100, como novos DAMPs, podem servir como biomarcadores da inflamação dirigida por células fagocíticas, e podem eventualmente, serem alvos atraentes para novas terapias anti-inflamatórias.

1.4 Detalhes moleculares da interação S100A12-RAGE

Estudos sugeriram que os resíduos de 3 a 39 da S100A12 representam o sítio de ligação ao RAGE, uma vez que esse trecho compartilha alguma homologia sequencial com a região C-terminal de outro ligante muito conhecido do RAGE, o HMGB1 (*“High-Mobility Group Box 1 Protein”* - Proteínas do Grupo de Alta Mobilidade Box 1). Essa região serve como *"motif"* de ligação ao RAGE para o HMGB1. Entretanto a natureza exata da interação S100A12-RAGE ainda não está definida [20].

A afinidade pelo Ca^{2+} e por outros íons bivalentes (Zn^{2+} e Cu^{2+}) influencia também a atividade exercida pelas proteínas S100 no espaço extracelular. Recentemente Xie e colaboradores, sugeriram, a partir de estudos utilizando fluorescência e espectroscopia RMN, que a S100A12 liga-se especificamente ao domínio C do receptor RAGE, tanto na presença, quanto na ausência de cálcio. Contudo a interação na ausência de Ca^{2+} , isto é, a interação apenas da porção protéica (apo-S100A12) com RAGE, é 100 vezes mais fraca do que a interação na presença de cálcio. Este estudo sugere ainda que hexâmeros de S100A12 estão complexados com tetrâmeros RAGE [32].

Contudo, outro estudo sugere a ligação de tetrâmeros de S100A12 no domínio V do RAGE [33]. Desta forma, maiores detalhes moleculares a respeito da oligomerização da S100A12 e de sua interação com receptor RAGE são necessários para compreender as discrepâncias entre os resultados obtidos atualmente.

1.5 O microambiente fisiopatológico da S100A12

Observamos anteriormente a ampla participação do eixo S100-RAGE no processo inflamatório. Uma questão relevante refere-se às condições no interior das células quando ocorre a liberação da S100A12 no tecido inflamado e na corrente sanguínea. Como vimos, a liberação da S100A12 é efetuada por neutrófilos ativados, que participam na regulação de uma grande variedade de processos fisiológicos e patológicos.

É conhecido o fato de que durante o processo de ativação destas células efectoras do sistema imune, há aumento de Ca^{2+} intracelular. Em monócitos e granulócitos, o Ca^{2+} intracelular regula várias atividades de resposta aguda como o “burst respiratório”, fagocitose, degranulação e a liberação de enzimas. O mecanismo molecular que envolve a mobilização de Ca^{2+} após a ativação por receptores de quimioatraentes e quimiocinas, é bem caracterizado.

A interação de receptores de membrana dos neutrófilos com quimiocinas, partículas opsonizadas, ou a adesão de moléculas do endotélio, resulta em abrupto e transitório aumento de Ca^{2+} citosólico, que precede e é um pré-requisito para o início das atividades pró-inflamatórias dessa célula. Estes receptores possuem sete domínios transmembranar acoplados a proteína G, resultando na ativação de fosfolipase C, a qual medeia a produção de IP_3 (inositol trifosfato), por hidrólise de fosfolípidos de membrana. O IP_3 interage com receptores que induzem a mobilização do cálcio das vesículas intracelulares, resultando numa descarga de cálcio no citosol [34].

O Ca^{2+} ativa funções inflamatórias que incluem a geração de superóxido (O_2^-) por um complexo enzimático associado à membrana citoplasmática dos neutrófilos e à membrana dos grânulos específicos, a NADPH oxidase; além de promover adesão ao endotélio vascular, degranulação, ativação da fosfolipase A2 citosólica e da 5-lipoxigenase, bem como síntese de IL-8, uma quimiocina para leucócitos específicos. A ativação da NADPH oxidase é a responsável pelo “burst” respiratório em neutrófilos e muitas espécies reativas de oxigênio, quando potente ação oxidante sobre estruturas e compostos celulares, são geradas. Essas espécies são capazes de oxidar membranas, inativar enzimas envolvidas no metabolismo bacteriano, quebrar cadeias polinucleotídicas, bem como degradar bases púricas e pirimidícas, nos ácidos nucleicos. Em processos nos quais ocorre liberação maciça destes reativos (O_2^- , H_2O_2 e $\cdot\text{OH}$) e de proteinases sobre os tecidos, por exemplo, em processos inflamatórios crônicos, essas células tornam-se responsáveis pela destruição dos tecidos do hospedeiro (Figura 8). Pode-se observar que há uma dependência crítica do Ca^{2+} para ativação das atividades pro-inflamatórias dos neutrófilos e está claro que um caminho molecular da transdução de sinal de Ca^{2+} envolve as proteínas que ligam cálcio da família S100, entre elas as calgranulinas [35].

Figura 8. Ativação dos neutrófilos por quimioatraentes como FMLP (Peptídeo N-formilmetionil- leucil-fenilalanina), PAF (Fator Ativador de Plaquetas), C5a (Componente C5a do Sistema Complemento) e LTB₄ (Leucotrieno B₄) aumentam a concentração de Ca²⁺ citosólico com a resultante geração de espécies reativas de oxigênio (ROS) e liberação de proteases, metaloproteinases (MMPs) e mediadores lipídicos. A injúria tissular que pode ser associada com a liberação dessas moléculas altamente danosas no tecido contribui na patogênese de um número grande de doenças [35].

Por outro lado, o papel do Zn²⁺ e do Cu²⁺ na modulação desse estresse oxidativo tem sido cada vez mais reconhecido. Sabe-se que o Zn²⁺ age como inibidor da NADPH oxidase, reduzindo portanto a produção de superóxido a partir do oxigênio molecular. Sabe-se também, que o superóxido (O₂⁻) é reduzido a peróxido (H₂O₂) pela enzima superóxido desmutase (SOD), a qual contém Zn²⁺ e Cu²⁺. A deficiência de Zn²⁺ tem sido relacionada ao aumento de citocinas inflamatórias, como TNF-α (fator de necrose tumoral – α) e IL-1β [36]. Além disso, o Zn²⁺ também induz a produção de metalotioneína (MT), uma família de proteínas muito rica em resíduos de cisteína (Cys), que funcionam como excelentes sequestradores de íons, e em particular ·OH. As MTs são encontradas principalmente em associação com o Zn²⁺, mas associam-se também a uma série de outros íons. Essa é uma família de proteínas pequenas (de 500 a 15000Da), presentes em uma abrangente gama de grupos taxonômicos. Em isoformas presentes em humanos, possuem uma estrutura básica muito própria, na qual 4 átomos de zinco estão ligados à 11 resíduos de cisteína no domínio C-terminal (α-domínio), e 3 átomos de zinco estão ligados à 9 resíduos de cisteína no domínio N-terminal (β-domínio). Sabe-se que a concentração de Zn²⁺ nas células é muito baixa, e portanto, o movimento de zinco livre é finamente controlado. O postulado do transporte de Zn²⁺ dentro da célula pode ter na MT seu fundamento. Os átomos de enxofre das cisteínas podem ser oxidados ou reduzidos com a consequente liberação ou complexação de zinco e a formação ou quebra de ligações dissulfeto. Observa-se que com o estresse oxidativo, por exemplo, pode haver aumento das ligações dissulfeto e liberação de Zn²⁺ no meio intracelular. O

Zn^{2+} livre inibe a NADPH oxidase, como já foi dito, reduzindo portanto o processo do “burst” respiratório dos neutrófilos [37].

A análise de todos esses eventos nos permite inferir que existe uma dinâmica nas variações das concentrações desses íons que contribui para promoção, regulação e finalização dos processos inflamatórios. A relevância da participação dos íons na ativação e regulação da resposta imune mediada por células deve ser considerada e está relacionada a proteínas que têm capacidade de interagir com esses diferentes íons. Evidente que esses mecanismos são bastante complexos e que a elucidação desse quadro do ponto de vista das variações estruturais promovidas por esses íons em proteínas em diferentes momentos, pode contribuir para a ampliação do conhecimento a respeito das vias de sinalização intra e extracelulares e de como eventos diversos são concatenados durante o início e o perdurar do processo inflamatório.

1.6 Detalhes Estruturais da S100A12

A calgranulina C (S100A12) madura consiste em 91 resíduos de aminoácidos e massa molecular estimada de 10.444Da. Como a maioria das proteínas S100, a S100A12 intracelular consiste em um homodímero antiparalelo em solução, com a interface entre duas subunidades unidas por contatos hidrofóbicos. Contudo, a ausência de cisteínas na estrutura primária da calgranulina C indica que se existe sua forma dimérica *in vivo*, esta não é covalentemente associada [38].

Cada monômero é composto por um “EF-hand” canônico C-terminal (H_I - L_I - H_{II}), comum a todas as proteínas “EF-hand”; e um “EF-hand” pseudo-canônico N-terminal (H_{III} - L_{III} - H_{IV}), o qual é exclusivamente encontrado em proteínas S100. Cada um dos “loops” de cada “motif” complexa um íon cálcio, e portanto quatro Ca^{2+} por dímero (Figuras 9 e 10).

Figura 9. Dímero da S100A12 humana (PDB ID: 1E8A) resolvida por Moroz e colaboradores, 2001 [18]. Visualização das hélices que formam o “EF-hand” pseudo-canônico em azul escuro (H_I) e azul claro (H_{II}), e das hélices que formam o “EF-hand” canônico em laranja (H_{III}) e amarelo (H_{IV}). São mostrados também os dois íons de cálcio complexados por monômero (esferas brancas).

Figura 10. A. Em destaque hélices que formam o “EF-hand” pseudo-canônico em azul escuro (H_I) e azul claro (H_{II}) e hélices que formam o “EF-hand” canônico em laranja (H_{III}) e amarelo (H_{IV}) (PDB ID: 1E8A). **B.** Sequência dos resíduos de aminoácidos que constituem o monômero da S100A12, as linhas pontilhadas indicam os dois “motifs” “EF-hands”.

A hélice H_I é formada pelos resíduos 2-18, a hélice H_{II} pelos resíduos 29-39, a hélice H_{III} pelos resíduos 50-60 e a hélice H_{IV} pelos resíduos 70-86. No EF-1, o Ca^{2+} é coordenado pelos grupos carbonílicos da cadeia principal dos resíduos Ser18, Lys21, His23, Thr26, pelo grupo carboxilato bidentado da cadeia lateral do resíduo Glu31 e por uma molécula de água (Figura 11). No EF-2, o Ca^{2+} é coordenado pelos grupos carboxilatos monodentados da cadeia lateral dos resíduos Asp61, Asn63, Asp65, pelo grupo carbonílico da cadeia principal do resíduo Gln67, pelo grupo carboxilato bidentado da cadeia lateral do resíduo Glu72 e por uma molécula de água. Para ambos os “EF-hands” a geometria dos ligantes é uma bipirâmide pentagonal [18] (Figura 12).

Figura 11. Detalhe do sítio de complexação do Ca^{2+} (esfera branca) no L_I do EF-1, obtido de uma estrutura resolvida na presença de Ca^{2+} (PDB ID: 1E8A). Os valores de distância de ligação são dados em angstrom.

Figura 12. Detalhe do sítio de complexação do Ca^{2+} (esfera branca) no L_{III} do EF-2, obtido de uma estrutura resolvida na presença de Ca^{2+} (PDB ID: 1E8A). Os valores de distância de ligação são dados em angstrom.

Calgranulina C também apresenta afinidade por íons Zn^{2+} and Cu^{2+} , em um sítio diferente do que complexa Ca^{2+} , formado por ambos os monômeros (cadeias A e B) e localizado na região de interface do dímero (Figuras 13 e 14). A S100A12 mostra afinidade por Zn^{2+} e Cu^{2+} na região C-terminal que possui uma sequência His-X-X-X-His compreendendo os resíduos His85 a His89. Estas duas histidinas fazem parte de uma hélice em uma conformação tal que favorece a ligação a esses íons [39].

Figura 13. Detalhe do sítio de complexação do Zn^{2+} (esfera cinza) no “motif” His-X-X-X-His entre os monômeros, obtido de uma estrutura resolvida na presença de Zn^{2+} (PDB ID: 2WCB). Os valores de distância de ligação são dados em angstrom.

Figura 14. Detalhe do sítio de complexação do Cu^{2+} (esfera laranja) no “motif” His-X-X-X-His entre os monômeros, obtido de uma estrutura resolvida na presença de Ca^{2+} e Cu^{2+} (PDB ID: 1ODB). Os valores de distância de ligação são dados em angstrom.

Quanto às propriedades de ligação ao Ca^{2+} , experimentos indicam que Zn^{2+} e Ca^{2+} se ligam a calgranulina C, e que a ligação do Zn^{2+} induz uma profunda variação nas propriedades da proteína ao se ligar ao Ca^{2+} [39]. Na ausência de Zn^{2+} , a S100A12 se liga a um íon Ca^{2+} por monômero, enquanto que na presença de Zn^{2+} esta mesma proteína liga-se a dois íons Ca^{2+} por monômero, sendo este último caso consistente com a presença dos dois “motifs” “EF-hands”, observados através da análise da sequência de resíduos de aminoácidos. Portanto, existe um comportamento dinâmico da variação da concentração desses íons que interfere nas alterações conformacionais da proteína em questão.

Além das conhecidas complexações de Ca^{2+} , Zn^{2+} e Cu^{2+} , em uma das estruturas da S100A12 obtida por estudo cristalográfico, realizado na ausência de Ca^{2+} (PDB ID: 2WCE), é possível observamos a complexação de íon sódio (Na^+) nos “loops” L_1 do EF-1 (pseudo-canônico) de ambos os monômeros, totalizando dois Na^+ complexados por dímero. Nesse caso, os resíduos que complexam o Na^+ são os mesmos que complexam o Ca^{2+} nesse mesmo “loop”, com exceção do Glu31 (Figura 15).

Figura 15. Detalhe do sítio de complexação do Na^+ (esfera violeta) no L_1 do EF-1, obtido de uma estrutura resolvida na ausência de Ca^{2+} (PDB ID: 2WCE). Os valores de distância de ligação são dados em angstrom.

Quando analisamos as propriedades destes íons que são complexados à S100A12, podemos observar que há grande semelhança no que se refere aos raios iônicos do Ca^{2+} e Na^+ , o que poderia justificar o fato do Na^+ estar complexado ao mesmo “loop” que complexa Ca^{2+} . Por outro lado, há também semelhança com relação aos raios iônicos de Zn^{2+} e Cu^{2+} , que possuem o mesmo sítio de complexação que é, por sua vez, distinto do sítio de complexação do Ca^{2+} (Tabela 1).

Elemento	Núm. Atômico	Massa Atômica (g/mol)	Densidade (g/cm ³)	Est. Oxidação	Eletro-negatividade	Raio Atômico (pm)	Raio Iônico (pm)	Raio de VdW (pm)
Sódio	11	22,99	0,968	+1, 0, -1	0,93	186	Na^+ 102	227
Cálcio	20	40,08	1,55	+2, +1	1,00	197	Ca^{2+} 100	231
Zinco	30	65,38	7,14	+2, +1, 0	1,65	134	Zn^{2+} 74	139
Cobre	29	63,55	8,94	+4, +3, +2, +1	1,90	128	Cu^{2+} 73	140

Tabela 1. Propriedades dos elementos sódio, cálcio, zinco e cobre.

A partir do que foi dito, tendo em vista as estruturas disponíveis desta proteína, nosso trabalho consiste em examinar a dinâmica conformacional da S100A12 na presença destes diferentes íons.

1.6.1 A biblioteca de estruturas disponíveis

Existe, diversas estruturas da Calgranulina C disponíveis no Banco de Dados de Proteínas, o PDB (*Protein Data Bank*). Todas essas estruturas são resultado de estudos cristalográficos, nos quais a Calgranulina C foi submetida a diferentes condições, que vão desde a ausência de íons até variações nas concentrações dos mesmos. Mostramos adiante as estruturas que estão disponíveis para nosso estudo (Figura 16).

Figura 16. Algumas das estruturas disponíveis da S100A12 no PDB. **A.** PDB ID: 1E8A, estrutura resolvida na presença de Ca^{2+} , que complexa dois íons Ca^{2+} por monômero (esferas brancas). **B.** PDB ID: 2WCE, estrutura resolvida na ausência de Ca^{2+} , que complexa dois íons Na^+ por dímero no L_I dos EF-1 (esferas violetas). **C.** PDB ID: 1ODB, estrutura resolvida na presença de Ca^{2+} e Cu^{2+} , que complexa dois íons Ca^{2+} por monômero (esferas brancas) e dois íons Cu^{2+} por dímero na sequência His-X-X-X-His (esferas laranjas). **D.** PDB ID: 1GQM, estrutura resolvida na presença de Ca^{2+} , observa-se a formação de uma estrutura de hexâmero e a complexação de íons Ca^{2+} nos *loops* L_I e L_{III} e numa região de interface entre os dímeros (esferas brancas). **E.** PDB ID: 2WCB, estrutura resolvida na ausência de Ca^{2+} , que complexa dois íons Zn^{2+} por dímero na sequência His-X-X-X-His (esferas cinzas) e dois íons Na^+ por dímero no L_I dos EF-1 (esferas violetas).

Existem diferenças estruturais consideráveis, sobretudo no que se refere às estruturas resolvidas na ausência ou na presença de Ca^{2+} . De forma geral, na presença de Ca^{2+} as hélices H_{III}

e H_{IV} do EF-2 estão em uma conformação dita aberta, enquanto na ausência de Ca²⁺ elas estão em uma conformação fechada [40] (Figura 17).

Figura 17. Alinhamento das estruturas da 1E8A (em preto) e 2WCE (em cinza). Podemos observar que na presença de Ca²⁺ (1E8A) as hélices do EF-2 estão em uma conformação aberta, enquanto na ausência de Ca²⁺ (2WCE) estão em uma conformação dita fechada.

Embora seja possível perceber que uma quantidade significativa de dados já se tenha acumulado a respeito da proteína S100A12, no que se refere a sua estrutura, interação com íons bivalentes e com o receptor RAGE, pouco se conhece sobre o comportamento dinâmico dos diferentes domínios da mesma, e sobre os mecanismos e parâmetros moleculares que controlam suas atividades e suas especificidades.

A elucidação de detalhes energéticos e dinâmico-estruturais da S100A12 se justifica não apenas por promover maior compreensão a respeito dos eventos celulares nos quais ela está inserida, mas também, por permitir ampliar o conhecimento a respeito das circunstâncias patológicas que envolvem alteração em sua expressão.

Há uma série de questões a serem elucidadas sobre a S100A12, que envolvem os detalhes a respeito de sua variação conformacional na presença de íons cálcio e zinco, ocorrência de mecanismos de oligomerização *in vivo* e minúcias moleculares da interação com o receptor RAGE. O que se propõe no presente trabalho é uma abordagem ampla destas questões relacionadas à proteína S100A12, que pretende aliar resultados obtidos a partir de estudos baseados em modelos computacionais, tendo em vista a relevância deste potencial alvo terapêutico nos eventos da inflamação.

As hipóteses existentes sobre os mecanismos subjacentes às mudanças estruturais baseiam-se, principalmente, em dados experimentais (cristalográficos). Portanto, métodos e técnicas de simulação computacional (basicamente Dinâmica Molecular), que foram empregados

de forma intensiva neste trabalho, constituem processos de fundamental importância para se investigar as possíveis trajetórias dinâmicas, completando assim o cenário das evoluções conformacionais induzidas por alterações nas condições fisiológicas (grandezas termodinâmicas intensivas).

CAPÍTULO 2

Uma abordagem utilizando simulação computacional:
dinâmica molecular

2.0 Uma abordagem utilizando simulação molecular computacional

De uma forma geral, simulação molecular computacional pode ser definida como o uso de métodos computacionais para descrever o comportamento da matéria em nível atômico ou molecular. A hierarquia de escalas de tempo e complexidade que existe entre o contínuo e o atômico e o desenvolvimento de modelos de multi-escalas, constituem ferramentas que conectam o contínuo e o atômico nesta área de pesquisa muito ativa.

Dentro de uma ampla faixa de técnicas, podemos distinguir entre métodos computacionais químico-quânticos; métodos de simulação molecular, como dinâmica molecular e monte-carlo; e métodos de modelagem “*coarse-grained*”, há também métodos híbridos, que consistem em misturas desses métodos. Os métodos computacionais químico-quânticos são baseados em física quântica, e levam em conta a estrutura eletrônica de átomos ou moléculas. As simulações envolvendo mecânica molecular são baseadas na física clássica e as interações entre os átomos são tratadas com funções potenciais empíricas. Como resultado, escalas de tempo maiores podem ser acessadas com simulações moleculares, embora ocorram algumas perdas do realismo químico. Métodos “*coarse-grained*” permitem ainda que maiores escalas de tempo e complexidade sejam investigadas por meio da substituição de alguns átomos por uma esfera que os represente, ou por uma parte da coleção de átomos dentro de sítios efetivos [41].

É inegável o fato de que a simulação molecular é uma área em ativo crescimento, sobretudo guiada pelos avanços computacionais no que se refere à velocidade na realização de cálculos pelos computadores, estratégias de paralelização e avanços metodológicos, podendo ser útil na predição de novos fenômenos, na explicação de fenômenos pouco compreendidos e na determinação de propriedades de materiais. Sabe-se que as simulações computacionais tem se tornado uma importante ferramenta tanto no ambiente da pesquisa acadêmica quanto na indústria, sobretudo na indústria química e farmacêutica, sendo que está última é a que mais investe em simulação computacional em todo o mundo [41].

Uma questão recorrente refere-se à discussão quanto ao tratamento das simulações computacionais como estudos teóricos ou experimentais. Devemos considerar que simulações computacionais são similares a experimentos porque requerem uso de máquinas, têm os mesmos erros estatísticos e problemas com reprodutibilidade encontrados nos experimentos. Ainda assim, todos os cálculos envolvidos nas simulações não devem ser considerados como puramente experimentais, uma vez que se tratam de cálculos puros, nos quais não são realizadas medidas de sistemas reais. É comum esquivarmo-nos deste questionamento tratando as simulações como uma área separada, em algum lugar entre teoria e experimento. Uma boa maneira de abordar essa questão é considerar as simulações computacionais como experimentos computacionais que

conectam teoria e experimento. Simulações moleculares requerem como entrada um modelo molecular caracterizado e permitem obter como resultado propriedades macroscópicas e informações microscópicas (Figura 18).

Figura 18. Entrada e saída de simulações moleculares.

Embora a conexão entre teoria, simulação e experimentos sugira a simulação computacional como um novo campo bastante promissor, e faça da dinâmica molecular uma ferramenta bastante útil no estudo, por exemplo, de moléculas biológicas, não devemos desconsiderar o fato de que simulações são soluções de modelos que se aproximam da realidade, de forma que a análise dos resultados obtidos com o emprego de simulações moleculares e a conexão destes com a realidade deve ser realizada de maneira absolutamente criteriosa.

2.1 Dinâmica molecular

As técnicas de simulações computacionais por Dinâmica Molecular (DM) baseiam-se na determinação do movimento dos átomos constituintes do sistema de interesse a partir do conhecimento do potencial de interação entre eles, das equações que regem o movimento e de condições impostas para representação do sistema no estado termodinâmico de interesse. Além disso, a integração numérica das equações de movimento exige definição de um conjunto de coordenadas e velocidades iniciais. Estes requisitos serão abordados, ainda que de maneira introdutória, nos próximos itens, embora ressalte-se a existência de uma ampla descrição na literatura do método em questão.

O método de dinâmica molecular foi primeiramente introduzido por Alder e Wainwright na década de 1950 com o estudo de esferas rígidas, sendo que muitas das importantes descobertas a respeito do comportamento de líquidos simples vêm destes estudos. Um avanço ocorreu em 1964, quando Rahman realizou a primeira simulação utilizando potencial realístico para argônio líquido. A primeira dinâmica molecular com um sistema realístico foi feita por Rahman e

Stillinger, quando foi simulada água líquida em 1974. A primeira simulação de proteína apareceu em 1977, com a simulação do inibidor de tripsina pancreática bovina [42].

Desde a realização desta primeira simulação de uma proteína, dinâmica molecular tem se tornado uma ferramenta muito útil no estudo de biomoléculas, complementando técnicas experimentais. Atualmente, é comum encontrar na literatura trabalhos que empregam simulações por dinâmica molecular de proteínas solvatadas, complexos DNA-proteínas, sistemas de membranas lipídicas, estudos de interação entre receptores e ligantes e “*folding*” de proteínas.

Um sistema é descrito mais exatamente pela resolução quanto-mecânica para cada geometria molecular. Porém, quando trabalha-se com macromoléculas biológicas, o elevado número de átomos requer uma representação física através de uma função potencial relativamente simples, descrevendo de forma clássica a interação entre as unidades atômicas. Problemas como estes são tratados pela substituição da abordagem *ab initio* por modelos físicos de função potencial mais simples.

Genericamente, atribui-se a denominação de campo de força a descrição de um sistema de muitas partículas pela sobreposição de termos simples, que descrevem a interação entre pares. Para o tratamento de centenas ou milhares de átomos como proteínas e ácidos nucleicos são introduzidas funções potenciais empíricas, calibradas por informações experimentais e cálculos quânticos sobre pequenas moléculas. Frequentemente se refere ao conjunto desses parâmetros de ajuste, juntamente com as respectivas funções, como “campo de força molecular” [43].

Potenciais podem ser definidos em vários níveis de precisão, e os mais comumente usados no tratamento de macromoléculas são baseados em mecânica molecular. O fato de a mecânica molecular produzir resultados satisfatórios mesmo desprezando eventuais efeitos quânticos se deve a vários fatores, e a portabilidade dos campos de força é uma delas, isto é: parâmetros adequados para moléculas pequenas podem ser transpostos para grandes polímeros, da mesma forma que um pequeno número de casos é suficiente para descrever uma grande quantidade de sistemas diferentes.

Outra importante propriedade da mecânica molecular clássica é que as energias potenciais são definidas em função das coordenadas moleculares, o que só é possível graças à aproximação de Born-Oppenheimer. Essa aproximação se baseia na grande diferença entre massa m_e de um elétron e a massa m_p de um próton ($m_p = 1836 m_e$), o que significa que os elétrons podem ajustar-se quase que instantaneamente aos prótons, sem influenciá-los, de forma que a dinâmica dos núcleos é desacoplada da dinâmica dos elétrons. Assim, os movimentos eletrônicos não são considerados: considera-se que os elétrons ajustam suas dinâmicas instantaneamente

quando ocorrem mudanças nas posições atômicas, de acordo com a aproximação de Born-Oppenheimer, e permanecem em seu estado fundamental [44].

2.1.1 Campo de força

Uma das maiores questões para a realização de simulações de DM é a parametrização dos campos de força, sobretudo no caso de moléculas como as proteínas pois a parametrização é complexa. A quantidade de dados experimentais termodinâmicos disponíveis para as proteínas é bastante limitada, e por esta razão, a parametrização de um campo de forças para proteínas geralmente envolve um trabalho sistemático que busca reproduzir dados energéticos obtidos a partir de cálculos *ab-initio*.

Existem diversos potenciais efetivos clássicos desenvolvidos especificamente para proteínas. Esses campos de força descrevem as interações entre diferentes sítios pela soma de interação de curto alcance, do tipo Lennard-Jones, e termos de interação eletrostática coulombianos de longo alcance. Estes potenciais são dados em termos de interações aditivas entre pares atômicos, de dois tipos:

$$E = E^{\text{ligadas}} + E^{\text{não-ligadas}} \quad (\text{Eq. 1})$$

Os conjuntos de parâmetros de cada um destes potenciais são continuamente ajustados e testados em diferentes condições de simulação tornando-os representações cada vez mais realistas dos potenciais reais.

As forças que agem sobre os átomos são obtidas calculando-se a primeira derivada do potencial em relação às coordenadas atômicas. Em seguida, tais forças são usadas para a solução das equações do movimento de Newton, como veremos nos próximos itens [45].

O campo de força utilizado neste trabalho é o OPSL-AA “*Force Field*” (“*Optimized Potentials for Liquid Simulations - All Atom*”) [46]. Esse campo de força é otimizado para simulações de sistemas líquidos e possui parâmetros desenvolvidos sob metodologia confiável em uma sistemática auto-consistente baseada apenas em cálculos quânticos. Outro aspecto, refere-se ao fato deste campo de força seguir o padrão “*All Atoms*”, no qual todos os átomos do sistema são modelados [47].

Nesse campo de força, os potenciais devido às interações não ligadas (E_{nl}) são compostos pelas interações eletrostáticas dadas por um potencial coulômbico clássico entre as

cargas parciais atribuídas a cada sítio e as interações de van der Waals são descritas pela função de Lennard-Jones, como está mostrado adiante:

$$V_{ij} = \frac{q_i q_j}{4 \pi \epsilon_o r_{ij}} + 4 \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (\text{Eq. 2})$$

Onde os índices i e j se referem a dois átomos interagentes, r_{ij} é a distância entre eles, q_i e q_j são as cargas parciais atribuídas aos átomos i e j , e ϵ_{ij} e σ_{ij} são os parâmetros da função de Lennard-Jones e ϵ_o refere-se a permissividade dielétrica no vácuo.

Com relação às ligações covalentes (interações ligadas), os potenciais devido ao estiramento das ligações entre dois átomos e distorções dos ângulos formados por três átomos, são representadas por potenciais harmônicos, respectivamente, V_{lig} e V_{ang} :

$$V_{lig} = K_\beta (r - r_{eq})^2 \quad \text{e} \quad V_{ang} = K_\theta (\theta - \theta_{eq})^2 \quad (\text{Eq. 3})$$

Onde K_β e K_θ são constantes harmônicas e r_{eq} e θ_{eq} são as posições de equilíbrio das ligações e ângulos.

E finalmente, os termos energéticos intracadeia referentes às torções do ângulo diedro, (ϕ), formado por quatro átomos, resultam de uma expansão em funções periódicas:

$$V_{die} = \sum_n \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \quad (\text{Eq. 4})$$

Onde n é a ordem das funções periódicas utilizadas na expansão, V_n é a amplitude de cada função, ϕ é o ângulo diedro e γ a sua fase.

Como foi dito, os potenciais do campo de força são dados em termos de interações aditivas entre pares atômicos, isso significa que todas as forças não ligadas resultam de uma soma das interações de pares não ligados. Interações de pares não aditivos, como por exemplo a polarizabilidade atômica, são representadas por potenciais de pares efetivos. Somente a

contribuição média de pares não aditivos são incorporadas. Isso significa também que as interações de pares não são puras, ou seja, não são válidas para pares isolados ou para situações que diferem significativamente a partir do sistema teste em que os modelos foram parametrizados. Embora na prática os potenciais de pares efetivos não sejam ruins, deve-se considerar que a omissão da polarizabilidade consiste em uma limitação do modelo.

2.1.2 Condições iniciais

Escolhido o campo de força, uma das primeiras etapas para simular um sistema de partículas através de DM é a geração das configurações iniciais que estejam o mais próximo possível de representações de um microestado do estado termodinâmico de interesse. Em simulações de dinâmica molecular, as condições iniciais consistem nas posições e velocidades de todas as partículas. Em simulações de proteínas as estruturas terciárias podem ser obtidas de experimentos de cristalografia de raio-X, espectroscopia de RMN e estudos computacionais por homologia. Evidente que além da estrutura terciária da proteína, são adicionadas ao sistema moléculas do solvente e contra-íons, necessários para neutralizar o sistema.

Além das configurações iniciais é necessário, para se iniciar uma simulação de DM, que sejam obtidas as velocidades iniciais para todos os átomos do sistema a ser simulado. Velocidades iniciais são geralmente atribuídas de forma aleatória aos átomos do sistema de acordo com a distribuição de Maxwell-Boltzmann para a temperatura desejada [48]. Desta forma, a energia cinética do sistema corresponderá à sua temperatura termodinâmica. Assim, condições iniciais para o sistema consiste em definir posições e velocidades iniciais para todos os átomos. E como todas as componentes das interações interatômicas dependem apenas das posições dos átomos, uma vez dadas as posições de todos os átomos do sistema, as forças atuando sobre cada átomo podem ser calculadas.

2.1.3 Otimização da geometria molecular

As contribuições parciais dos potenciais geram uma hipersuperfície de energia a qual deve ser percorrida em busca de regiões de mínima energia. Essas regiões correspondem a pontos no espaço de configurações nos quais as forças que agem sobre os átomos estão equilibradas. As macromoléculas biológicas têm como característica um grande número de graus de liberdade, e assim cada região dessa hipersuperfície de energia define uma estrutura molecular particular, e uma correspondente energia estrutural [49].

A otimização da geometria molecular tem como objetivo encontrar um conjunto de coordenadas que minimizem a energia potencial do sistema considerado. Para tanto, realiza-se

uma busca sobre a superfície de potencial na direção em que a energia decresça de maneira que o sistema seja levado a um mínimo local próximo. De forma geral, a configuração final, após este processo, não difere em muito da estrutura experimental inicial, embora os ajustes nas posições atômicas relaxe as distorções nas ligações químicas, nos ângulos entre as ligações e nos contatos de van der Waals. A minimização ainda retira todas as componentes de energia cinética do sistema prevenindo ruídos térmicos.

Um algoritmo bastante empregado na minimização de energia é o “Steepest Descent”. Esse algoritmo não é o mais eficiente para a minimização da energia, porém é robusto e de fácil implementação, por isso é muito utilizado em sistemas que não apresentam maiores dificuldades de relaxação. Nesse método, após um primeiro cálculo das forças e energia potencial do sistema, novas posições são calculadas, conforme a expressão seguinte, para todos os átomos:

$$r_{n+1} = r_n + \frac{F_n}{\max|F_n|} h_n \quad (\text{Eq. 5})$$

Onde h_n é o deslocamento máximo e F_n é a força (gradiente negativo do potencial V). A notação $\max|F_n|$ significa o valor máximo absoluto do componente de força. A força e a energia são então novamente calculados nas novas posições:

Se $(V_{n+1} < V_n)$ as novas posições são aceitas e $h_{n+1} = 1.2h_n$

Se $(V_{n+1} \geq V_n)$ as novas posições são rejeitadas e $h_{n+1} = 0.2h_n$

Em vista do problema de convergência, usa-se algum critério para interromper a minimização, como por exemplo, um determinado número de passos ou uma certa diferença de energia entre dois passos sucessivos. Se a mudança de energia (ou força) calculada entre esses dois passos for menor que essa diferença estabelecida, encerra-se a minimização [50].

2.1.4 Integração das equações de movimento

Conhecendo como todas as partículas de um sistema interagem entre si, podemos ter as condições iniciais para poder prever a trajetória das partículas ao longo do tempo. Como vimos, ter condições iniciais para o sistema consiste em definir posições e velocidades iniciais para todos os átomos. Todas as componentes das interações interatômicas dependem apenas das posições dos átomos. Desta forma, dadas as posições de todos os átomos do sistema, as forças atuando sobre cada átomo podem ser calculadas. A força nada mais é do que a derivada do potencial:

$$\vec{F}(t) = -\nabla \vec{V}(t) \quad (\text{Eq. 6})$$

De acordo com a segunda lei de Newton, podemos calcular a aceleração de cada átomo, $\vec{a}(t)$, usando:

$$-\nabla \vec{V}(t) = m\vec{a}(t) \quad (\text{Eq. 7})$$

Sendo m a massa do átomo. O algoritmo mais simples para obter a evolução temporal do sistema consiste em calcular as velocidades no instante $t+dt$ por:

$$\vec{v}(t+dt) = \vec{v}(t) + \vec{a}(t) dt \quad (\text{Eq. 8})$$

E, usando a equação:

$$\vec{x}(t+dt) = \vec{x}(t) + \vec{v}(t) dt + \frac{\vec{a}(t) dt^2}{2} \quad (\text{Eq. 9})$$

obtem-se as posições no instante $t+dt$. Dadas as posições neste novo instante, os potenciais são recalculados, definindo novas forças, novas acelerações e, assim, permitindo novamente o cálculo das posições em um instante posterior. A repetição deste ciclo leva à obtenção da trajetória de todas as partículas do sistema ao longo do tempo.

A integração das equações de movimento é realizada através de algoritmos baseados nos métodos das diferenças finitas, nos quais a integração é dividida em pequenos intervalos de tempo (passos de integração), Δt , permitindo simular os movimentos de maior frequência do sistema, que muitas vezes são as vibrações das ligações.

Este método de integração das equações de movimento é o mais simples. Existem algoritmos diferentes que são ao mesmo tempo mais precisos e mais eficientes, mas cujo princípio fundamental é o mesmo [51].

2.1.5 Aproximações para representação do estado termodinâmico

Evidente que, devido a limitações computacionais, não podemos realizar simulações de DM de sistemas com número de partículas de mesma ordem do número de Avogadro ($6,023 \times 10^{23}$), durante um tempo necessário. Assim, existem limitações espaciais, no que se refere às dimensões da caixa de simulação, limitações temporais, no que se refere ao tamanho da trajetória obtida, e

aproximações para a representação do sistema no estado termodinâmico de interesse. Para simular um sistema através de DM, inicialmente é necessário criar uma caixa de simulação que compreenderá um número N de moléculas que varia entre 10^3 e 10^6 átomos, dependendo do sistema que se queira estudar.

Esse número N de partículas é insignificante quando comparado ao número de partículas contidas em uma porção macroscópica de matéria, e a razão entre o número de partículas da superfície e o número de partículas total seria muito maior do que na realidade, fazendo com que os efeitos de superfície parecessem muito mais importantes do que realmente são. Se considerássemos 1000 átomos arranjados em uma caixa cúbica de $10 \times 10 \times 10$, mais da metade dos átomos estariam nas faces externas e isto teria uma grande interferência sobre as propriedades medidas [52]. Uma solução para este problema é a utilização de condições periódicas de contorno (PBC - “*Periodic boundary conditions*”), que quando utilizadas, permitem que as partículas sejam colocadas em uma caixa e as imagens desta caixa sejam replicadas como imagens idênticas nas três direções cartesianas, preenchendo completamente o espaço, formando um sistema que, embora periódico, tenda ao limite termodinâmico [50] (figura 19).

Figura 19. Condições periódicas de contorno em duas dimensões [50]

As partículas nas réplicas se movimentam de maneira idêntica à célula original. A questão é que com as condições periódicas de contorno cada partícula pode interagir com outra partícula da caixa e também com suas imagens nas caixas próximas, eliminando virtualmente os efeitos de superfície no sistema. Nesta configuração periódica, as interações podem atravessar limites da caixa e isto implica em um número infinito de termos, o que seria de impossível resolução na prática. Contudo, se as forças envolvidas forem de suficiente curto alcance, com

relação às dimensões da caixa de simulação, pode ser aplicada a convenção da imagem mínima, isto é, considera-se que determinada molécula ocupa a posição central em uma região de mesmo tamanho e forma da caixa simulada. Esta molécula interage apenas com as $N-1$ demais moléculas situadas no interior desta região.

Para reduzir ainda mais o custo computacional, emprega-se também a aproximação do raio de corte esférico (chamado “*cutoff*”). Ou seja, o potencial decai praticamente a zero quando a distância entre os pares é maior do que o raio de corte estabelecido. É preciso que este seja grande o suficiente para não introduzir erros significativos, mas não maior do que metade do lado da caixa de simulação, por consistência com a convenção da imagem mínima.

No que se refere ao tratamento das interações de longo alcance, como as interações coulômbicas, a delimitação de um raio de corte é problemática, uma vez que o truncamento abrupto do potencial produz efeitos que, em muitos casos, comprometem severamente os resultados das simulações, de forma que se costuma empregar o método PME (“*Particle Mesh Ewald*”), que consiste no método de soma de Ewald mais eficiente, no que se refere ao ganho em desempenho, que a soma de Ewald clássica [53].

Outro ponto relevante relaciona-se ao fato de sistemas moleculares apresentarem movimentos que ocorrem em diferentes escalas temporais, considerando tanto as forças intermoleculares como as intramoleculares. Normalmente essas últimas são de ordem de magnitude superior às das primeiras, de modo que, por exemplo, os movimentos vibracionais são muito mais rápidos que os translacionais e rotacionais. As ligações com átomos leves podem vibrar em frequências cujo tempo característico é inferior ao passo de integração, ocorrendo assim erros na propagação das posições e na conservação da energia, alterando significativamente a estrutura do sistema [54]. No entanto, se tais movimentos mais rápidos forem congelados (subtraídos do cálculo das interações), restringindo os graus de liberdade correspondentes aos movimentos de maior frequência, pode-se conseguir uma maior eficiência através da utilização de um passo temporal maior na integração das equações do movimento, reduzindo consideravelmente o custo computacional. Essas ligações devem ser corrigidas usando “*constraints*” para eliminar o efeito espúrio dos movimentos rápidos. Existem vários métodos usados para realizar “*constraints*” moleculares [55]. Os mais utilizados são os algoritmos SHAKE, RATTLE e LINCS [56-58].

Devemos considerar também as condições termodinâmicas sob as quais as simulações serão executadas. Alguns parâmetros macroscópicos podem ser mantidos constantes, como temperatura (T), pressão (P), volume (V), número de partículas (N) e potencial químico

(μ). Sendo a simulação uma amostragem ampla de configurações do sistema, conjuntos desses parâmetros definem diferentes “ensembles”.

Por várias razões (flutuações durante a etapa de equilibração, flutuações no resultado do truncamento de forças e erros na integração, aquecimento devido a forças externas) faz-se necessário o controle da temperatura do sistema. Um método bastante utilizado acopla o sistema a um banho externo, sendo o acoplamento térmico é realizado através do uso de um termostato. Além de um termostato, pode-se acoplar também um barostato à simulação, se o objetivo for controlar a pressão do sistema, dependendo do “ensemble” utilizado [50]. Existem vários algoritmos para o controle de temperaturas, sendo os mais utilizados Berendsen [59], Nöse-Hoover [60,61], V-rescale [62].

2.1.6 Análises das trajetórias

Após a simulação ser realizada, dispomos de um conjunto de coordenadas que representam a evolução temporal do sistema, isto é, a posição e velocidade de cada átomo, nas condições termodinâmicas nas quais as simulações foram realizadas. Sendo o sistema em estudo ergódico, a média temporal de uma grandeza se iguala a média do “ensemble”, o que possibilita a utilização de resultados da Mecânica Estatística cujos princípios fundamentais permitem o cálculo de grandezas macroscópicas a partir de média e flutuações de propriedades microscópicas do “ensemble” [52].

Existe uma série de análises que podem ser feitas com base nas trajetórias geradas pelas simulações de DM, de forma que médias e flutuações de propriedades estruturais e dinâmicas podem ser calculadas.

Alguns exemplos de propriedades que serão muito úteis nos estudos propostos no presente trabalho, são detalhadas a seguir:

- RMSD (“root-mean-square deviation” - desvio da raiz média quadrática): calcula o desvio padrão médio quadrático para todos os átomos da proteína.

$$RMSD = \sqrt{\frac{1}{N} \sum_{j=i}^N [r_i(t) - r_i(0)]^2} \quad (\text{Eq. 10})$$

Onde $r_i(t)$ e $r_i(0)$ são as coordenadas do i -ésimo átomo no tempo t e tempo 0 , respectivamente, e N é o número de átomos no domínio de interesse, neste trabalho calculado

para os átomos de carbono-alfa (C_α). O RMSD é utilizado para estimar os desvios médios da estrutura da proteína em relação à estrutura referência ($t = 0$) ao longo da simulação. Dessa forma, podemos observar a evolução da mudança estrutural total da molécula ao longo da simulação e verificar questões relacionadas à estabilidade, equilibração e rigidez/flexibilidade do sistema. O RMSD fornece o desvio instantâneo de toda molécula ou parte da mesma.

- **Raio de Giro:** consiste na média das distâncias quadráticas de cada resíduo de aminoácido ao centro de gravidade da cadeia. Indica o grau de compactação da proteína e, de certo modo, sua estabilidade, podendo ser calculado pela equação:

$$R_g = \sqrt{\left(\sum (r_i)^2 m_i / \sum m_i \right)} \quad (\text{Eq. 11})$$

Onde m_i é a massa do átomo i , r_i é a posição do átomo i com relação ao centro de massa da molécula.

- **RMSF (“root-mean-square fluctuation” - flutuação da raiz média quadrática):** calcula o desvio padrão quadrático médio das posições atômicas ao longo da simulação com relação às posições dos átomos ou resíduos aminoácido de uma estrutura tomada como referência.

$$RMSF_i = \sqrt{\frac{1}{\zeta} \sum_{j=0}^{\zeta} [r_i(j) - \bar{r}_i]^2} \quad (\text{Eq. 12})$$

Onde $r_i(j)$ são as coordenadas do i -ésimo átomo no passo de tempo j , $\bar{r}_i$ suas posições médias e ζ é o tempo de simulações, expresso como o número total de passos de tempo (“time steps”) coletados. O RMSF é uma estratégia complementar ao RMSD, seu valor pode ser entendido como uma medida de flexibilidade do sistema analisado. O uso da análise do RMSF juntamente com a de RMSD permite identificar o grau de mudança conformacional, flexibilidade e equilibração para cada região da macromolécula em estudo.

- **Superfície acessível ao solvente (SAS):** considera-se a superfície hidrofílica, hidrofóbica e a área de superfície total disponível ao solvente. O cálculo de SAS é feito empregando o conceito de superfície de Conolly.

- **Ligações de Hidrogênio:** são determinadas usualmente pelo estabelecimento de um valor limite para o ângulo entre o aceptor de hidrogênio, o doador de hidrogênio e o hidrogênio, e uma determinada distância entre aceptor e doador de hidrogênio.

- **Acompanhamentos dos termos de energia do sistema:** o acompanhamento dos diferentes termos de energia do sistema (como por exemplo, Energia Total, Energia Cinética, Energia Potencial) são indicadores da estabilidade do sistema molecular durante a simulação.

- **Estrutura secundária:** Determinação do número de resíduos que formam uma estrutura secundária. Para análise de conteúdo de estrutura secundária neste trabalho foi utilizado o algoritmo STRIDE (*“STRuctural IDentification”*) [63]. Este algoritmo para atribuir elementos de estrutura secundária de uma proteína dadas suas coordenadas atômicas, tal como definido por Cristalografia e Difração de Raios-X, RMN, ou outros métodos de determinação de estrutura de proteína. Em adição ao critério de ligação de hidrogênio, utilizado pelo algoritmo mais popular, o DSSP (*“Define Secondary Structure of Proteins”*), o critério de atribuição do STRIDE também inclui potenciais de ângulo dihedral. Assim, seus critérios para definição de estruturas secundárias individuais são mais complexos do que os do DSSP. Os critérios para atribuição de elementos de estrutura secundária também contam com fatores de probabilidade estatística derivados de análises empíricas de estruturas resolvidas com atribuição visual de elementos de estruturas secundárias extraídas do PDB (*“Protein Data Bank”*) [49]. Embora o DSSP seja o método mais antigo e mais comumente usado, o STRIDE mostrou-se mais satisfatório na atribuição estrutural de 70% dos *“folds”* de 226 proteínas analisados por Frishman e colaboradores [63]. Principalmente, foi observado que o STRIDE corrige uma tendência do DSSP em atribuir estruturas secundárias mais curtas do que seria atribuído por um cristalógrafo especialista, geralmente devido a pequenas variações locais na estrutura que são mais comuns perto dos términos de elementos de estrutura secundária.

- **Análise de propriedades eletrostáticas** – O Software APBS (*“Adaptive Poisson-Boltzmann Solver”*) foi desenvolvido por Nathan Baker em colaboração com J. Andrew McCammon e Michael Holst, consiste em um programa disponível gratuitamente para cálculos eletrostáticos com macromoléculas (<http://www.poissonboltzmann.org/apbs>) [64]. É um software para modelagem de solvatação biomolecular por meio da resolução da equação de Poisson-

Boltzmann (PBE), em um dos mais populares modelos contínuos para descrever interações eletrostáticas entre solutos moleculares em meio aquoso salino.

2.2 GROMACS – O pacote de simulação

Existe uma série de pacotes de programas destinados a realização de simulações de DM disponíveis atualmente. Neste projeto optamos pelos softwares de simulação e análise contidos no pacote (livre) GROMACS [65-69], tendo em vista a natureza exploratória das simulações requeridas neste projeto, e a sua versatilidade. O nome GROMACS deriva das iniciais para “GRoningen MACHine for Chemical Simulation”, que resulta de um projeto conjunto do Departamento de Química e do Departamento de Ciências da Computação da Universidade de Groningen, na Holanda. O GROMACS é um programa versátil e eficiente para dinâmica molecular, especialmente de macromoléculas biológicas em ambiente aquoso ou de membrana, e é capaz de utilizar tanto processadores individuais, quanto processamento em paralelo. O pacote é gratuito e de código aberto, possibilitando atualmente, a participação de diversos desenvolvedores de partes distintas do mundo, de forma que diferentes versões são lançadas de tempos em tempos. No presente trabalho fez-se, de forma geral, uso da versão GROMACS v4.x.

GROMACS é um programa que suporta diferentes campos de força (como OPLS, utilizado no presente projeto, além do AMBER, CHARMM, entre outros [70]) e possui várias rotinas otimizadas que o torna um dos sistemas atuais mais rápidos. Adicionalmente, o pacote oferece uma série de ferramentas úteis para a análise das trajetórias, programas estes que foram amplamente utilizados neste estudo.

2.2.1 Etapas da simulação de dinâmica molecular

Os diversos sistemas simulados neste trabalho foram preparados de acordo com parâmetros que serão detlhados oportunamente nos próximos capítulos, com diferentes estruturas cristalográficas obtidas da proteína S100A12 em diferentes condições, sendo que todas estas estruturas estão disponíveis no banco de dados PDB. Todas as simulações foram realizadas utilizando o pacote GROMACS, sendo que as simulações foram divididas em baterias, que serão detalhas adiante.

Após a preparação do sistema, procedeu-se a etapa de minimização de energia e da dinâmica de restrição (simulação de equilíbrio), e em seguida foi realizada a simulação para aquisição de dados, a dinâmica molecular propriamente dita, e as consequentes análises das trajetórias geradas.

CAPÍTULO 3

Dinâmica conformacional da proteína S100A12 mediante
interação com íons

Como vimos no capítulo 1, existem importantes variações dinâmico-conformacionais da proteína Calgranulina (S100A12) induzidas pela presença de diferentes íons, dados estes que já foram constatados em estudos estruturais cristalográficos que constituem nosso ponto de partida neste trabalho.

Como também vimos, existem diversas situações fisiopatológicas no ambiente intra e extracelular que envolvem aumento de expressão de S100A12 e variação nas concentrações e nos tipos de íons presentes. Nosso objetivo neste ponto do trabalho é verificar como estas variações conformacionais ocorrem do ponto de vista dinâmico na presença destes íons, para que possamos obter respostas a respeito de que regiões desta proteína são importantes nas alterações conformacionais que envolvem a passagem do estado descomplexado ao estado complexado aos íons. A obtenção destas informações se justifica na medida em que identificando regiões relevantes do ponto de vista dinâmico-estrutural, poderemos atuar por exemplo, no planejamento de inibidores, que neste caso seriam importantes agentes antiinflamatórios, sobretudo no que se refere aos processos inflamatórios crônicos, nos quais, como sabemos, há aumento da expressão desta proteína.

3.1 Conjunto de parâmetros definidos para simulações

As simulações foram realizadas em um reservatório na forma de caixa-paralelepípedica, na qual a proteína foi solvatada por moléculas de água e o sistema submetido a um “*ensemble*” NVT (Número de partículas (N), Volume (V) e Temperatura (T) constantes - “*ensemble*” canônico), todos os átomos foram considerados explicitamente, usando o campo de forças OPLS-AA. O modelo de água utilizado foi SPC (“*Simple Point Charge*”) no qual, a molécula de água tem três centros de concentração de carga: com predominância de carga positiva sobre os dois átomos de hidrogênio e excesso de carga negativa sobre o átomo de oxigênio.

No “*ensemble*” canônico, o sistema de interesse é fechado e permanece em contato diatérmico com um reservatório de calor em determinada temperatura, mantida constante (para nossos experimentos, a temperatura utilizada foi de 309K). A ampla variação da pressão, que observamos em simulações teste, mesmo quando o sistema está acoplado a um barostato, nos fez optar pelo uso do “*ensemble*” NVT nestas simulações, ao invés do “*ensemble*” NPT.

Foram utilizados termostatos distintos para as etapas de restrição e da dinâmica molecular propriamente dita. De acordo com o manual do GROMACS v4, “V-rescale” é essencialmente o termostato de “Berendsen” com um termo adicional que assegura uma distribuição de energia cinética correta. Sabe-se que o acoplamento fraco de “Berendsen” é mais

eficiente para relaxar um sistema à uma dada temperatura, mas quando este sistema já está em equilíbrio, o mesmo não é o mais indicado para manter o “*ensemble*” canônico. Por outro lado, de acordo com o mesmo manual, o termostato de “Nose-Hoover” controla o período de flutuação de temperatura exatamente no equilíbrio. Dessa forma, nos pareceu apropriada a utilização do termostato “V-rescale” durante a dinâmica de restrição, e a utilização do termostato “Nose-Hoover” durante a dinâmica molecular propriamente dita [50].

Prezando pela uniformidade experimental, todos os experimentos neste trabalho compartilham de um mesmo conjunto básico de parâmetros, variando apenas, conforme nosso interesse, a estrutura inicial da proteína a ser simulada e o tipo/concentração de íons aos quais essa estrutura foi submetida. Detalhes a respeito dessas variações serão especificados oportunamente.

A fim de neutralizar as cargas implícitas geradas em pH 7,0 pelas porções N(+) e C(-) terminais, e pelos resíduos carregados que a proteína apresenta, foram inseridos íons ao sistema, em quantidades e qualidades apropriadas. As interações coulômbicas foram tratadas pelo método de somas de Ewald via PME (“*Particle Mesh Ewald*”) e foram aplicadas condições periódicas de contorno, e raios de corte de 1nm para as interações eletrostáticas e de van der Waals. O intervalo de tempo para a integração numérica do sistema foi de 2fs, sendo as coordenadas da simulação salvas a cada 1ps.

As análises das trajetórias foram feitas através de cálculos de propriedades estruturais e dinâmicas, descritas no item 2.1.6. As análises foram feitas com ferramentas do GROMACS e a utilização dos programas PYMOL [71], VMD [72], GRACE [73] e QTIPLLOT [74].

Nossas simulações foram feitas em baterias de realizações, sendo que cada bateria conta com um conjunto de simulações para verificar determinados aspectos do comportamento conformacional que serão descritos detalhadamente. Após a apresentação de cada bateria, seguem os resultados e discussões pertinentes ao conjunto das simulações correspondentes. Ressaltamos o fato de que embora todos os resultados e gráficos das propriedades descritas tenham sido gerados e analisados, optamos por apresentar apenas aquelas análises que nos forneceram indícios a respeito das alterações conformacionais da S100A12.

3.2 Bateria 1

Simulações de DM da S100A12 cuja estrutura cristalográfica foi obtida na ausência de íons Ca^{2+} (PDB ID: 2WCE).

3.2.1 BATERIA 1A – Comportamento da estrutura da S100A12 resolvida na ausência de Ca^{2+} (PDB ID: 2WCE) mediante presença de 100mM de íons Na^+ , Ca^{2+} , Zn^{2+} tendo como contra-íon o Cl^- .

Com o intuito de iniciar o estudo a respeito do comportamento da proteína S100A12 na presença de diversos íons, foram realizadas inicialmente quatro simulações de 80ns cada, na qual a estrutura 2WCE (S100A12 resolvida na ausência de Ca^{2+} , com dois íons Na^+ complexados aos L_1 nos EF-1) foi submetida à concentração de 100mM de íons Na^+ , Ca^{2+} , Zn^{2+} em sistemas separados. Para essa concentração de 100mM temos: 13 íons Na^+ , 13 íons Ca^{2+} e íons 16 Zn^{2+} nos sistemas simulados; em todos os casos os contra-íons adicionados para neutralizar o sistema foram íons Cl^- . (Figura 20).

3.2.2 BATERIA 1B – Comportamento da estrutura da S100A12 resolvida na ausência de Ca^{2+} (PDB ID: 2WCE), mediante presença de diferentes concentrações de íons – Análise da influência da variação da concentração de íons (25mM, 50mM e 75mM de íons Na^+ , Ca^{2+} , Zn^{2+} tendo como contra-íon o Cl^-).

Em seguida, iniciamos a investigação a respeito da interferência da variação na concentração de íons nas alterações conformacionais promovidas pelos diferentes íons na S100A12. Assim, a Bateria 1 foi complementada com simulações com concentrações de 75mM, 50mM e 25mM de NaCl , CaCl_2 e ZnCl_2 , sendo que cada simulação foi realizada para o tempo de 100ns. A essa altura do desenvolvimento do trabalho, já contávamos com mais recursos computacionais, que nos permitiram realizar um maior número de simulações e simulações mais longas. Ressaltamos que para a concentração de 75mM temos: 9 íons Na^+ , 9 íons Ca^{2+} e 12 íons Zn^{2+} nos sistemas simulados; para a concentração de 50mM temos: 6 íons Na^+ , 6 íons Ca^{2+} e 9 íons Zn^{2+} nos sistemas simulados e para a concentração de 25mM temos: 3 íons Na^+ , 3 íons Ca^{2+} e 6 íons Zn^{2+} nos sistemas simulados. Em todos os casos os contra-íons adicionados para neutralizar o sistema foram íons Cl^- .

Reunindo as simulações da Bateria 1A com as simulações da Bateria 1B, temos um quadro geral formado por 12 simulações, como podemos ver na Figura 20.

Figura 20. Esquema geral dos sistemas simulados nas baterias 1A e 1B. Foram realizadas simulações da 2WCE (S100A12 resolvida na ausência de Ca^{2+} , que complexa dois íons Na^+ nos EF-1 de cada monômero), submetida a concentrações 100, 75, 50 e 25mM de NaCl, CaCl_2 e ZnCl_2 em sistemas separados.

3.2.3 Resultados e análises das simulações: Bateria 1

Inicialmente, apresentamos os resultados a respeito do comportamento estrutural da S100A12 na presença de diferentes íons, para concentrações de 100mM de NaCl, CaCl_2 e ZnCl_2 . A figura 21 mostra a evolução do desvio estrutural (RMSD) para os três tipos de íons.

Figura 21. Deformação estrutural da 2WCE na presença de íons, de acordo com os valores de RMSD. Em destaque na figura, é mostrado o RMSD médio no intervalo entre 50 e 70ns de simulação. As barras de erros correspondem aos desvios padrão dos respectivos valores médios. Podemos observar que a presença de ZnCl₂ na concentração de 100mM no sistema simulado leva a alterações conformacionais globais um pouco acentuadas na 2WCE, principalmente quando comparamos à presença de 100mM de CaCl₂.

De forma geral, podemos observar, baseado nos dados de RMSD, que a presença de ZnCl₂ na concentração de 100mM no sistema simulado leva a alterações conformacionais globais mais significativas na 2WCE, quando comparamos à presença de 100mM de CaCl₂. Essa verificação é corroborada quando analisamos os resultados de RMSF destas simulações (Figura 22):

RMSF_cadeia A (2WCE) para 100 mM NaCl, CaCl₂, ZnCl₂ RMSF_cadeia B (2WCE) para 100mM NaCl, CaCl₂, ZnCl₂

Figura 22. RMSF para as cadeias A e B. Somente as regiões N e C-terminais apresentam comportamento similar. No que se refere a cadeia A, gráfico à esquerda, destacam-se os fatos de que na presença de CaCl₂, linha vermelha, observamos flutuações menores, quando comparadas às flutuações na presença de ZnCl₂, linha verde, enquanto as flutuações provocadas por 100mM NaCl são intermediárias. Observamos ainda, para a cadeia A, regiões de flexibilidade entre os resíduos 18 e 27 (que compreendem o L_I), e regiões de grande flexibilidade entre os resíduos 50 e 60 (que compreendem a região da hélice H_{III}) e entre os resíduos 60 e 67 (que compreendem a região do “loop” L_{III}). No que se refere a cadeia B, a região que demonstra maior flexibilidade para todas as simulações realizadas, situa-se entre os resíduos 110 e 118 (que compreendem a região do “loop” L_I), sobretudo na presença de NaCl, linha preta.

Analisando os gráficos de RMSF, podemos inferir que as cadeias A e B apresentam comportamentos distintos na presença de íons, embora trate-se de um homodímero de S100A12. Podemos observar também que na presença de 100mM de ZnCl₂ observamos, de forma geral, maiores flutuações estruturais do que quando comparamos com 100mM de CaCl₂ e NaCl.

Com relação às regiões de maior flexibilidade, destacamos os “loops” I e III e a região da hélice 3 (H_{III}). Como sabemos, regiões de “loop” tem naturalmente maior flexibilidade, no entanto a alta flexibilidade da H_{III} é surpreendente. Assim, acrescentamos gráficos de RMSD por hélices para as simulações (Figura 23), para compararmos com os resultados obtidos para o RMSF.

Figura 23. Deformação estrutural das hélices e L_{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da proteína 2WCE na presença de íons. De forma geral podemos observar que as maiores deformações estruturais, em todas as simulações, referem-se a região que compreende a hélice H_{III} (linha verde). As demais hélices apresentam poucas deformações ao longo do tempo de simulação, sobretudo a H_{IV} (linha azul) é bastante estável. Quando observamos em termos das diferenças promovidas pela presença dos diferentes íons, podemos observar que com 100mM de CaCl₂, o RMSD da H_{III} permanece constante ao longo da simulação. A presença de 100mM de NaCl e ZnCl₂ leva a oscilações nos valores de RMSD da H_{III} ao longo do tempo de simulação.

Os gráficos de RMSD por hélice nos mostram que, de forma geral, a hélice H_{III} é a que apresenta maiores deformações estruturais e que sua estrutura sofre oscilações na presença de 100mM de NaCl e ZnCl₂, embora permaneça estável quando na presença de 100mM de CaCl₂. Por outro lado, as demais hélices sofrem oscilações bem menores, mantendo-se estáveis nas simulações realizadas. Com base nestes resultados, é possível afirmar que a H_{III} é uma região importante no que se refere a mudanças conformacionais promovidas por íons na S100A12.

Estes resultados indicam que a 100mM de CaCl₂ alterações estruturais gerais menores ocorrem no dímero, quando comparadas com os sistemas a 100mM de ZnCl₂ e NaCl (Figura 21).

As médias temporais da flutuação estrutural por resíduo, RMSF (Figura 22) mostram que embora os diferentes íons afetem a flexibilidade da proteína nas mesmas regiões, tais regiões são diferentes nas cadeias (A e B) constituintes do dímero, indicando alguma assimetria na composição dimérica.

Finalmente, a figura 23 revela que a hélice H_{III} é o setor do dímero mais afetado pelos íons a 100mM: em todos os três casos, a H_{III} é substancialmente alterada em relação à sua conformação cristalográfica original. Contudo, enquanto para os sistemas com Na⁺ e Zn²⁺ as perturbações são instáveis, para o sistema com Ca²⁺, a H_{III} é alterada estruturalmente de forma estacionária.

Uma análise complementar refere-se à variação superfície hidrofóbica acessível ao solvente nestas simulações (Figura 24).

Figura 24. Superfície hidrofóbica acessível ao solvente ao longo das simulações a 100mM. Podemos observar que na presença de CaCl_2 (linha vermelha), houve redução da superfície hidrofóbica acessível ao solvente. Já na presença de ZnCl_2 , houve pontos de ligeiro aumento da superfície hidrofóbica acessível ao solvente (linha verde), enquanto que na presença de NaCl não houve variação significativa ao longo da simulação.

Podemos observar que na simulação com 100mM há uma redução sistemática da superfície hidrofóbica acessível ao solvente para todos os íons, indicando que o dímero como um todo reduz sua superfície de exposição, por meio de uma maior compactação estrutural em relação à configuração cristalográfica.

Para o sistema a 100mM, o sódio e sobretudo, o zinco, com superfícies expostas maiores, tem uma papel perturbador do dímero do ponto de vista estrutural, enquanto o cálcio, demonstra ter uma papel estabilizador do dímero.

Neste ponto, uma questão relevante refere-se a maneira como esses íons (Na^+ , Ca^{2+} e Zn^{2+}) distribuem-se nos sistemas simulados. Tendo como propósito obter uma visão geral da distribuição dos íons ao longo da simulação, foram feitas imagens da simulação (da seleção contendo proteína e cátions), de 10 em 10ns, e em seguida foi mantida a estrutura do último tempo, para a proteína e de todos os tempos para os cátions, para que pudéssemos obter um resultado qualitativo da densidade de visitação dos íons na proteína, ou seja, a sobreposição dos

íons nos locais em que estes se mantêm ligados ao longo da simulação, como mostrado na figura 25.

Figura 25. Densidade de visitação dos íons mostrada na estrutura final após 80ns de simulação. Sistemas com 100mM Na^+ (esferas vermelhas), Ca^{2+} (esferas brancas) e Zn^{2+} (esferas cinzas). Podemos observar que tanto o Na^+ quanto o Zn^{2+} embora aproximem-se da proteína, apresentem distribuição quase homogênea ao redor da proteína, o que sugere que não há complexação dos mesmos, ao passo que o Ca^{2+} é encontrado em certas regiões (L_{III} do EF-2), indicando que os mesmos permanecem ligados ao longo da simulação.

A análise das imagens qualificando a densidade de visitação dos íons, pode ser relacionada aos efeitos estruturais na proteína, resultantes da presença dos íons. Os íons Ca^{2+} ligam-se a regiões da proteína próximas aos “*loops*”; desta forma, eles têm sua mobilidade diminuída, enquanto os íons Na^+ e Zn^{2+} permanecem distribuídos pelo sistema e levam à alterações estruturais globais no dímero.

Outro ponto relevante na análise da densidade de visitação, refere-se à aproximação dos íons Ca^{2+} com relação à proteína. A análise da aproximação dos íons Ca^{2+} na figura 25, nos

permite observar que esta é feita via L_{III} , do EF-2 (canônico), tendo como resíduos importantes nesse aspecto Asp65 (em laranja) na cadeia A e Glu55 na cadeia B (em amarelo) na cadeia B.

Quando analisamos a superfície acessível ao solvente colorida de acordo com o potencial eletrostático, calculado utilizando a ferramenta APBS, para a proteína submetida a 100mM de $CaCl_2$, logo no início da simulação, obtemos uma imagem como mostrado adiante (figura 26).

Figura 26. Superfície acessível ao solvente colorida de acordo com o potencial eletrostático, calculado utilizando a ferramenta APBS para o início da simulação com 100mM $CaCl_2$. Apresentamos a visão dos EF-1 das cadeias, à esquerda, enquanto à direita temos uma visão dos EF-2. É possível notar que a região dos EF-2 possui carga negativa em toda a face, o que deve fazer com que o Ca^{2+} aproxime-se pelas hélices H_{III} e H_{IV} (que formam o EF-2) nas regiões dos L_{III} .

Além das simulações desta bateria, foi feita uma simulação controle para verificarmos a questão da região de chegada do Ca^{2+} . Embora tivéssemos indícios de que o Ca^{2+} aproximava-se pelas regiões do L_{III} , como já foi mostrado, fizemos uma simulação na qual a mesma estrutura da 2WCE foi submetida à presença de $CaCl_2$ ao longo de 100ns. No entanto, retiramos o Na^+ inicialmente complexado ao L_I , da estrutura obtida na cristalografia, para verificar se a aproximação do Ca^{2+} pelo L_{III} não era reflexo do fato do L_I estar complexando Na^+ . Também nesta simulação os resultados permanecem os mesmos quanto a ligação do Ca^{2+} à proteína e indicam que ainda que o L_I esteja livre desde o início da simulação, a aproximação do Ca^{2+} ocorre pela face dos “EF-hands” canônicos.

Nesse sentido, nossa hipótese contradiz o previsto pela literatura, já que o modelo para as alterações estruturais promovidas pela ligação do Ca^{2+} às proteínas S100, construído por analogia ao comportamento da calmodulina, sugere que o Ca^{2+} primeiramente liga-se ao L_I do EF-1, levando a uma comunicação entre os “EF-hands” que causa a reorganização do L_{III} (EF-2) [18].

Dois parâmetros muito importantes em nossas análises foram a variação do conteúdo de estrutura secundária (em termos de conteúdo de hélices) ao longo das simulações, e o comportamento do íon sódio, inicialmente complexado ao “*loop*” I em cada monômero da 2WCE. No entanto, para efeito de melhor comparação os gráficos que mostram estes dois parâmetros e suas implicações serão analisado considerando todas as simulações realizadas, incluindo as simulações que envolvem a variação de concentração de íons, como será mostrado mais adiante.

Tendo em vista a continuação do estudo, começamos a investigar a respeito da possibilidade da interferência da variação na concentração de íons nas alterações conformacionais da S100A12. Os resultados nessa bateria serão mostrados com adição dos dados das simulações de 100mM (bateria 1A) para efeito de comparação.

Em termos das deformações estruturais promovidas pelos íons, seguem os valores de RMSD para as simulações, sendo que serão mostrados os gráficos de RMSD por hélice para facilitar nossas análises posteriores.

- RMSD por hélices para variações na concentração de NaCl (Figura 27):
- RMSD por hélices para variações na concentração de CaCl₂ (Figura 28):
- RMSD por hélices para variações na concentração de ZnCl₂ (Figura 29):

Figura 27.

Figura 28.

Figura 29.

Figura 27. Deformação estrutural das hélices e L_{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da proteína 2WCE na presença de diferentes concentrações de NaCl. De forma geral podemos observar que os maiores valores de RMSD (maiores deformações estruturais) para todas as simulações referem-se a região que compreende a H_{III} (linha verde). As demais hélices apresentam poucas deformações ao longo do tempo de simulação, sobretudo as H_I (linha preta) e H_{IV} (linha azul) são bastante estáveis. Quando observamos em termos das diferenças promovidas pelas distintas concentrações dos íons, notamos que à concentração 50mM há uma maior estabilidade da hélice H_{III}, em ambas as cadeias do dímero, embora permaneça deformada em relação à estrutura cristalográfica de referência.

Figura 28. Deformação estrutural das hélices e L_{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da 2WCE na presença de diferentes concentrações de CaCl₂. De forma geral podemos observar que as maiores deformações e desestabilidade (valores de RMSD maiores e oscilantes) para todas as simulações referem-se a região que compreende a H_{III} (linha verde). As demais hélices apresentam poucas deformações ao longo do tempo de simulação, sobretudo as H_I (linha preta) e H_{IV} (linha azul) são bastante estáveis. Quando observamos em termos das diferenças promovidas pelas diferenças nas concentrações dos íons, podemos dizer que concentrações menores de CaCl₂ (de 50mM e 25mM, respectivamente) levam a maiores oscilações das deformações conformacionais da H_{III}, confirmando o papel de estabilização da presença de Ca²⁺ na solução.

Figura 29. Deformação estrutural das hélices e L_{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A (coluna à esquerda) e B (coluna à direita) da proteína 2WCE na presença de diferentes concentrações de ZnCl₂. Aqui também observamos que em geral as maiores deformações e flutuações estruturais (valores de RMSD maiores) para todas as simulações referem-se à hélice H_{III} (linha verde). Contudo, contrastando com os outros sistemas, tais deformações e flutuações são nitidamente crescentes com a concentração de Ca²⁺ na solução. Curiosamente, para a cadeia B, para o sistema com 75mM temos uma deformação estrutural também na H_{II} (linha em vermelho).

As análises conjuntas do RMSD por hélices, nas simulações com variação na concentração de íons, nos mostrou que nas concentrações inferiores (50mM e 25mM) de CaCl₂ há maiores oscilações das deformações conformacionais da H_{III}. Enquanto que para estas mesmas concentrações, as deformações promovidas pelo ZnCl₂ na H_{III} são menores. Desta forma, é possível afirmar que a proteína, e especificamente a H_{III} é sensível, do ponto de vista estrutural, a alterações nas concentrações de íons, respondendo de maneiras distintas à presença de cálcio e zinco. A seguir apresentamos os resultados de RMSF mostrando maiores detalhes a respeito dos efeitos da variação da concentração de íons na dinâmica estrutural da S100A12 (Figura 30):

Figura 30. RMSF para as cadeias A e B do dímero da 2WCE para diferentes concentrações de Na^+ , Ca^{2+} e Zn^{2+} . Os resultados mostram que de forma geral, as regiões de maior flexibilidade continuam sendo os “loops” L_I (resíduos 19 a 28 na cadeia A, e 111 a 120 na cadeia B), L_{III} (resíduos 61 a 69 na cadeia A e resíduos 153 a 161 na cadeia B), e a H_{III} (resíduos 50 a 60 na cadeia A, resíduos de 142 a 152 na cadeia B). A flexibilidade de cada região, em particular, é afetada por cada íon diferente a uma diferente concentração: para os sistema CaCl_2 , podemos observar que no que se refere à H_{III} (resíduos 50 a 60 na cadeia A, resíduos de 142 a 152 na cadeia B) a concentração de 50mM leva a um aumento de flexibilidade (linha em verde), tanto na cadeia A quanto na cadeia B. Nesta mesma região, para sistemas com 50mM de ZnCl_2 , observamos que há redução da flexibilidade da H_{III} (resíduos 50 a 60 na cadeia A, resíduos de 142 a 152 na cadeia B), enquanto que para concentração de 100mM a flexibilidade desta região aumenta (linhas em verde e preto, respectivamente). Outro ponto relevante das análises de RMSF, refere-se a grande flexibilidade da região do L_I na cadeia B das simulações com NaCl e nas cadeias A e B nas simulações na presença de diferentes concentrações de ZnCl_2 .

Nossas análises RMSF para as simulações com variações na concentração dos íons Na^+ , Ca^{2+} e Zn^{2+} nos permitiram confirmar que os diferentes íons afetam a flexibilidade de uma mesma região de diferentes formas. A concentração de 50mM de CaCl_2 a flexibilidade da H_{III} é a mais alta, enquanto que para esta mesma concentração, a flexibilidade desta região é a menor na presença de ZnCl_2 . Outro fato refere-se a alta flexibilidade do L_1 na presença de ZnCl_2 , à concentração de 75mM, tanto na cadeia A quanto na cadeia B.

Com o intuito de obter maiores detalhes a respeito do comportamento da S100A12 em nossas simulações, determinamos a variação do conteúdo de estrutura secundária, sobretudo no que se refere ao comportamento das alfa-hélices ao longo das simulações, uma vez que este é o padrão de estrutura secundária predominante nesta proteína. A composição de estrutura secundária foi calculada ao longo das simulações de DM, usando o algoritmo STRIDE (Figura 31).

Figura 31. Conteúdo de estrutura secundária em termos de porcentagem média de alfa-hélices (entre 50 e 70ns das simulações). As barras de erros correspondem aos respectivos valores dos desvios padrão dos valores médios. Entre 50 e 70ns não há variação significativa na porcentagem de hélices, para todos os casos, possibilitando ter seus valores médios comparados entre as simulações realizadas. O cálculo da composição de estrutura secundária da proteína ao longo das simulações foi feito em intervalos de 20ps, usando o algoritmo STRIDE. A composição de estrutura secundária correspondente as folhas-beta é mantida constante em 5% (dado não apresentado). A perda de hélice resulta na formação de “coil”. A barra mais escura indica a porcentagem de composição de estrutura secundária do tipo alfa-hélice para a estrutura cristalográfica (PDB ID: 2WCE). Em todos os sistemas, assim como em todas as concentrações de íons analisadas, há redução no conteúdo de alfa-hélices quando comparamos os valores médios com o conteúdo de alfa-hélices da estrutura cristalográfica. Os resultados mais significativos aqui mostrados referem-se a perda do conteúdo de alfa-hélices que ocorre nas simulações com 50mM de CaCl_2 e 75mM de NaCl .

Quando analisamos a variação do conteúdo de estrutura secundária em termos de porcentagem média de alfa-hélices amostrada entre 50 e 70ns de simulação, podemos observar que de forma geral há redução do conteúdo de alfa-hélices em todas as simulações quando comparamos os valores médios com o conteúdo de alfa-hélices da estrutura cristalográfica. Valores significativos de perda do conteúdo de alfa-hélices ocorrem nas simulações com 50mM de CaCl_2 e 75mM de NaCl . Tendo em vista este fato, nossa análise seguiu no sentido de detectar em qual região ocorria essa perda de hélice. Analisamos a localização da região de perda de hélice para as simulações de 50mM de CaCl_2 e 75mM de NaCl juntamente com a verificação da densidade de visitação dos íons na proteína. Foram feitas imagens, de 10 em 10ns e em seguida foi mantida a estrutura do último tempo (100ns), para a proteína e de todos os tempos para os íons, para que pudéssemos obter um resultado qualitativo da densidade de visitação dos íons na proteína, ou seja, observar a sobreposição dos íons nos locais em que esses se mantêm ligados ao longo destas duas simulações (figura 32).

Figura 32. Densidade de visitação dos íons mostrada na estrutura final após 100ns de simulação a 75mM NaCl (Na^+ esferas vermelhas), 50mM CaCl_2 (Ca^{2+} esferas cinzas). Podemos observar que os íons Na^+ , embora aproximem-se da proteína e perturbem a H_{III} na cadeia A, não permanecem ligados à proteína (distribuição homogênea dos íons). Por sua vez o íon Ca^{2+} , uma vez que se aproxima das regiões do L_{III} (do EF-2) por meio da H_{III} na cadeia A, é complexado e permanece ligado ao longo da simulação, embora também perturbe a H_{III} . Notamos novamente também, que a aproximação do Ca^{2+} é realizada pela face dos “loops” L_{III} (EF-2). Destacamos ainda o fato do Ca^{2+} além de permanecer interagindo na H_{III} da cadeia A, também mantém-se complexado à região do L_{II} , o “Hinge-Region”, na cadeia B.

Quando analisamos a estrutura da proteína após as simulações, podemos observar que a perda do conteúdo de estrutura secundária ocorre devido à perda parcial da H_{III}, que é perturbada principalmente nas simulações com 75mM de NaCl e 50mM de CaCl₂. Portanto, esta perda é localizada, embora no caso do Na⁺, este não se mantenha ligado ao longo da simulação a esta região, como acontece para o Ca²⁺ (Figura 33).

Figura 33. Análise local comparativa da H_{III} e L_{II} em diferentes circunstâncias. **A.** Estrutura cristalográfica da S100A12 (PDB ID: 2WCE) antes de qualquer etapa da simulação. **B.** Visão local da densidade de visitação dos íons Ca²⁺ (esferas brancas) na simulação com 50mM de CaCl₂ e **C.** Visão local da densidade de visitação dos íons Na⁺ (esferas vermelhas) na simulação com 75mM de NaCl. Podemos observar que nos itens B e C a perda de hélice é localizada na região da H_{III} e L_{II}. Embora a presença de 75mM de NaCl comprometa mais a H_{III}, os íons Na⁺ não permanecem complexados.

Quando analisamos em detalhes a região da H_{III} nestas duas simulações, podemos observar que a H_{III} é a região de maior flexibilidade nesta proteína, e que esta flexibilidade é dependente da concentração de íons no ambiente, de forma que acima ou abaixo desta concentração, esse comportamento de flexibilidade não é mantido. Podemos supor que esta hélice apresenta um comportamento dinâmico no sentido de forma-se e desformar-se em uma situação fisiológica na simples presença de determinada concentração de NaCl. No entanto, por apresentar capacidade de complexação ao íon cálcio, e não ao íon sódio, é apenas na presença de Ca²⁺ que se dá a mudança conformacional para o estado complexado, sem perturbar extremamente a hélice H_{III}.

A perda da hélice na região do “Hinge” é um indicativo bastante importante na transição da estrutura inicialmente descomplexada do Ca²⁺ (2WCE – dita conformação fechada) para o que

podemos observar na estrutura resolvida na presença de Ca^{2+} (1E8A – dita conformação aberta) [40]. No trabalho de Moroz e colaboradores; a autora mostra que a transição baseada na complexação de Ca^{2+} tem como ingrediente a perda da hélice no “*Hinge*”, a qual ela denomina H_{IIa} por ser uma extensão da hélice H_{II} . Em nossa simulação com 50mM de CaCl_2 , podemos observar essa perda da hélice H_{IIa} (figura 34).

Figura 34. Visão do L_{II} , conectando as hélices H_{III} (à direita) e H_{II} (à esquerda) em diversas condições. Estrutura cristalográfica da S100A12 resolvida na ausência de Ca^{2+} , 2WCE. Estrutura 2WCE após 100ns de simulação na presença de 50mM CaCl_2 . Estrutura cristalográfica da S100A12 resolvida na presença de Ca^{2+} , 1E8A. A transição de um estado de ausência de Ca^{2+} para um estado complexado ao Ca^{2+} , envolve a perda da hélice H_{IIa} e a perda parcial da hélice H_{III} , permitindo que estas regiões tornem-se mais flexíveis, o que pode ser observado após simulação na presença de 50mM CaCl_2 .

A perda da H_{IIa} , que passa a compor o L_{II} (“*Hinge-Region*” - que conecta os “*EF-hands*”) e a perda parcial da H_{III} , permite que estas regiões tornem-se mais flexíveis, e possam rotacionar com relação às hélices H_{II} e H_{IV} (que são as hélices mais estáveis) facilitando a entrada de Ca^{2+} no L_{III} . A ocorrência de todos esses eventos faz com que a proteína passe então do estado descomplexado ao estado complexado ao Ca^{2+} e isso é dependente da concentração de íons cálcio no ambiente.

Outro aspecto analisado em nossas simulações, refere-se ao comportamento do Na^+ inicialmente complexado ao L_1 do EF-1 na estrutura resolvida por cristalografia na ausência de cálcio (PDB ID 2WCE). Determinamos as distâncias do íon sódio inicialmente complexado em

relação ao centro de massa do L_I (levando em conta os resíduos que complexam o sódio inicialmente). A partir destas distâncias, podemos observar quando ocorre ou não a descomplexação do sódio e construímos um mapa de existência de complexação (Figura 35). Os gráficos de distância do Na^+ inicialmente complexado nas cadeias A e B, embora não estejam mostrados neste trabalho, nos forçaram a informação de que nos casos em que o Na^+ se descomplexa do L_I , ele não volta a ser complexado.

Figura 35. Mapa de existência da complexação do Na^+ ao L_I para as cadeias A (indicadas por ChA) e cadeias B (indicadas por ChB) para as simulações com variações de concentrações de $NaCl$, $CaCl_2$ e $ZnCl_2$. O comprimento da barra horizontal marca o instante em que se dá a descomplexação. A complexação foi analisada por meio do cálculo das distâncias do íon sódio inicialmente complexado em relação ao centro de massa do L_I (levando em conta os resíduos que complexam o sódio inicialmente). Ressaltamos que em nenhum dos casos onde o Na^+ foi descomplexado houve recomplexação do íon. De forma geral, a distância do Na^+ ao centro de massa do L_I mantém-se entre 1 e 2 Å enquanto o Na^+ está complexado e aumenta drasticamente quando da descomplexação. Podemos observar que na presença de $ZnCl_2$ há uma tendência a descomplexação do Na^+ inicialmente complexado ao L_I , sobretudo nas concentrações de 50mM e 75mM, o que não ocorre para $NaCl$ e $CaCl_2$.

Os dados do mapa de existência da complexação do Na^+ , inicialmente ligado ao L_I na estrutura da 2WCE após as simulações, e os fatos de que em concentrações de 25, 50 e 75mM CaCl_2 , o Na^+ permanece complexado ao L_I (exceto para a cadeia B da simulação de 50mM, na qual o sódio é descomplexado por volta dos 85ns da simulação) enquanto nessas mesmas concentrações, na presença de ZnCl_2 , o Na^+ é rapidamente descomplexado (exceto para a cadeia B da simulação de 25mM onde o sódio permanece complexado), sugerem a participação do Na^+ na manutenção da estrutura do L_I em períodos nos quais as concentrações de Ca^{2+} são inferiores. Essa hipótese também está presente no trabalho de Moroz e colaboradores, 2009; no qual a autora sugere ser possível que o Na^+ ocupe o L_I do EF-1 em períodos entre ondas de Ca^{2+} [40].

Adicionalmente, observamos aqui, que esta hipótese pode ser extrapolada, pois mostramos também a participação do Zn^{2+} na saída de Na^+ do L_I . Como evidenciado na figura 35 acima, em determinadas concentrações de ZnCl_2 , a saída do Na^+ ocorre muito rapidamente.

Quando analisamos a saída do Na^+ em uma das simulações com ZnCl_2 , mais especificamente na cadeia B da simulação com 50mM ZnCl_2 , podemos observar um contato importante entre um Zn^{2+} e um resíduo que faz parte do L_I (Asp25). A aproximação do Zn^{2+} a esse resíduo perturba o L_I , fazendo com que ele se abra e permita a descomplexação do Na^+ , levando a saída deste íon (figura 36).

Figura 36. Configurações em tempos distintos do L_I na cadeia B da simulação com 50mM ZnCl_2 . Embora a saída do Na^+ (esfera azul) ocorra aos 16ns (item F) a abertura do L_I ocorre logo nos primeiros picossegundos de simulação. A presença do Zn^{2+} (esfera cinza) interagindo com o resíduo Asp25, que faz parte do L_I , faz com que este “loop” se abra, já aos 2ns (item D).

Essa interação do resíduo Asp25 com Zn^{2+} foi observada também nas outras simulações. Nesse ponto, devemos lembrar o fato de que esse resíduo de Asp25 é um dos que complexa o Zinco no domínio His-X-X-X-His, como foi mostrado na introdução deste trabalho.

Detalhe do sítio de complexação do Zn^{2+} (esfera cinza) no “motif” His-X-X-X-His entre os monômeros, obtido de uma estrutura resolvida na presença de Zn^{2+} (PDB ID: 2WCB). Os valores de distância de ligação são dados em angstrom (**Figura 13**, capítulo 1)

Isso nos fornece a informação de que o Asp25 pode ser o resíduo que permita uma comunicação entre o domínio de complexação do Zn^{2+} e o L_I . Ou seja, em determinadas concentrações, o Zn^{2+} pode facilitar a saída de Na^+ do L_I , permitindo que este fique livre para complexar o Ca^{2+} . Esse comportamento, constitui uma hipótese para o aumento de afinidade pelo Ca^{2+} verificada na presença de Zn^{2+} , uma vez que no trabalho de Dell'Angelica e colaboradores, 1994; é afirmado que o Zn^{2+} induz um aumento notável na afinidade da S100A12 pelo Ca^{2+} , cerca de 1500 vezes, e afirma-se ainda que na ausência de Zn^{2+} cada monômero complexa apenas um Ca^{2+} . Por outro lado, segundo o mesmo trabalho, na presença de Zn^{2+} cada monômero passa a complexar dois Ca^{2+} , conforme seria compatível, uma vez que cada monômero possui dois “EF-hands” [23].

3.3 Bateria 2

3.3.1 BATERIA 2 – Análise do comportamento da estrutura da S100A12 resolvida na presença de Ca^{2+} (PDB ID: 1E8A) mediante presença de 50mM de íons Na^+ , Ca^{2+} , Zn^{2+} tendo como contra-íon o Cl^- .

Como todo nosso estudo foi realizado para simulações que tinham como estrutura de partida a 2WCE (S100A12 resolvida na ausência de Ca^{2+}) decidimos realizar simulações também com a 1E8A (S100A12 resolvida por Cristalografia e Difração de Raios-X na presença de Ca^{2+}).

Considerando os recursos computacionais disponíveis e nossos resultados prévios sobre a influência da concentração de íons na dinâmica conformacional da S100A12, escolhemos uma concentração intermediária de 50mM e foram realizadas quatro simulações, de 100ns cada, nas quais a 1E8A foi submetida a 50mM de NaCl , CaCl_2 , ZnCl_2 (Figura 37).

Figura 37. Esquema geral dos sistemas simulados na bateria 2. Foram realizadas simulações da 1E8A (S100A12 resolvida na presença de Ca^{2+} , que complexa dois Ca^{2+} por monômero, nos “loops” L_I e L_{III}) submetida a concentrações de 50mM de NaCl , CaCl_2 e ZnCl_2 em sistemas separados.

3.3.2 Resultados e análises das Simulações: Bateria 2

Com relação aos valores de RMSD desta simulações, seguem as análises por hélices (figura 38):

Figura 38. Deformação estrutural das hélices e L_{II} (“hinge” - região que conecta os “EF-hands”) para cadeias A e B da 1E8A na presença de íons. De forma geral ao contrário do observado para as simulações com a 2WCE (S100A12 resolvida na ausência de Ca²⁺), a H_{III} nas simulações da 1E8A apresenta um comportamento bem mais estável estruturalmente para as simulações com 50mM de CaCl₂ e ZnCl₂. Contudo, para a simulação na presença de 50mM de NaCl, os valores de RMSD são maiores para a H_{III} e L_{II} sobretudo para a cadeia B do dímero, demonstrando maiores oscilações estruturais da H_{III}.

Quando analisamos os valores de RMSF, podemos obter informações a respeito das oscilações conformacionais nestas simulações (figura 39):

Figura 39. RMSF para as cadeias A e B das simulações realizadas. Podemos observar que além da natural flexibilidade na região C-terminal, em todas as simulações, podemos observar que de forma geral o L_I (resíduos 19 a 28 na cadeia A, e 111 a 120 na cadeia B), L_{III} (resíduos 61 a 69 na cadeia A e resíduos 153 a 161) e a H_{III} (resíduos 50 a 60 na cadeia A, resíduos de 142 a 152 na cadeia B) são mais estáveis do que nas simulações com a 2WCE na presença de 50mM de $CaCl_2$ e $ZnCl_2$. No entanto, na simulação com 50mM de $NaCl$, na cadeia B do dímero, há um aumento nos valores de RMSF, para as regiões do L_{II} e H_{III} .

A análise dos valores de RMSD e RMSF nos indicam que nas simulações com 50mM de $CaCl_2$ e $ZnCl_2$, a 1E8A permanece bastante estável, sem sofrer grandes deformações estruturais. No entanto, na simulação com 50mM de $NaCl$, há um aumento de flexibilidade na região que corresponde ao L_{II} e H_{III} . Antes de uma análise mais detalhada a respeito deste ponto, duas observações devem ser citadas. Em primeiro lugar, ressaltamos o fato de que na estrutura da 1E8A há dois Ca^{2+} complexados por monômero nos “loops” L_I e L_{III} , e que em nenhuma das simulações dessa bateria houve descomplexação desses Ca^{2+} . Outra nota refere-se à variação do conteúdo de estrutura secundária, em termos de porcentagem de alfa-hélices, ao longo das simulações (figura 40).

Figura 40. Conteúdo de estrutura secundária em termos de porcentagem média de alfa-hélices entre 50 e 70ns das simulações. As barras de erros correspondem aos desvios padrão dos valores médios correspondentes. O cálculo da composição de estrutura secundária da proteína ao longo das simulações foi feito em intervalos de 20ps, usando o algoritmo STRIDE. A composição de estrutura secundária correspondente a folhas-beta é mantida constante em 3,5% (dado não apresentado). A perda de hélice resulta na formação de “coil”. A barra mais escura indica a porcentagem de composição de estrutura secundária do tipo alfa-hélice para a estrutura cristalográfica (PDB ID: 1E8A). De forma geral, em todas as simulações há uma pequena redução no conteúdo de alfa-hélices quando comparamos os valores médios com o conteúdo de alfa-hélices da estrutura cristalográfica (PDB ID: 1E8A). Mas não notamos diferenças significativas quando comparamos as simulações com 50mM de NaCl, CaCl₂ e ZnCl₂.

Embora a alteração global no conteúdo de alfa-hélices não seja significativa, quando analisamos localmente as alterações estruturais que ocorrem na presença de 50mM de NaCl, podemos observar que nesta simulação, a hélice H_{IIa} da cadeia B foi refeita (figura 41). Apenas na simulação com NaCl, a região do “Hinge-Region” (L_{II}) recuperou a estrutura de hélice, típica dos casos em que não há complexação de cálcio.

Figura 41. Alinhamento das estruturas da 1E8A antes e depois de 100ns de simulação na presença de 50mM de NaCl, em cinza e azul, respectivamente. Em B, podemos observar o início da formação de uma hélice (a H_{IIa}) na região “Hinge” após 100ns de simulação.

Se compararmos a estrutura da 1E8A depois de 100ns de simulação de DM na presença de 50mM de NaCl com a estrutura da 2WCE, temos (figura 42):

Figura 42. Alinhamento das estruturas da 1E8A antes e depois de 100ns de simulação na presença de 50mM de NaCl₂, e da 2WCE (S100A12 antes da complexação com o Ca^{2+}) em amarelo, violeta e azul, respectivamente. Podemos observar o início da formação de uma hélice (a H_{IIa}) na região do “Hinge”, localizada na mesma região do que existe para a 2WCE.

Quando analisamos em detalhes a simulação na presença de 50mM de NaCl, podemos perceber que a H_{III} tem sua estrutura perturbada. Inicialmente, no estado complexado ao Ca²⁺ (PDB IB: 1E8A), há um núcleo hidrofóbico organizado, formado por resíduos da H_{III}, do “*Hinge*” e alguns resíduos da H_{IV}. No entanto, com o decorrer da simulação, os resíduos ácidos da H_{III} interagem com os íons Na⁺, embora não complexem esses íons. Da mesma forma como observamos para as simulações com a 2WCE, a presença de NaCl perturba do ponto de vista estrutural a estabilidade da H_{III}, embora os íons Na⁺ não se liguem de fato a nenhum resíduo específico, mas fiquem orbitando na região. A presença desses íons, faz com que o núcleo hidrofóbico se desorganize na simulação com 50mM de NaCl (figura 43) e aos poucos a H_{IIa} se forma, como vimos anteriormente.

Figura 43. Visão do EF-2 composto pelas H_{III} e H_{IV}, na estrutura da 1E8A antes e após 100ns de DM na presença de 50mM de NaCl. No início é possível observar um núcleo hidrofóbico, formado por resíduos da H_{III}, do “Hinge” e da H_{IV}. Em A temos as representações das cadeias laterais dos resíduos que formam o núcleo hidrofóbico (Leu35, Leu36, Val40, Ile44, Ile47, Val52, Ile53, Ile56, Ile60, Leu76 e Ile79). Em B temos a representação em “dots” das cadeias laterais destes mesmos resíduos. Em C temos também em destaque os resíduos carregados da H_{III} (Lys50, Asp54, Glu55 e Gln58). Podemos observar que antes da simulação na presença de 50mM de NaCl, o núcleo hidrofóbico está bem organizado. Mediante a presença dos íons Na⁺, os resíduos ácidos da H_{III} orientam-se para fora e o núcleo hidrofóbico fica desorganizado.

A orientação dos resíduos ácidos da H_{III} para fora e a desorganização do núcleo hidrofóbico estão provavelmente envolvidos na reorganização estrutural da proteína que envolve a formação da H_{IIa}.

3.4 Conclusões acerca da dinâmica conformacional da proteína Calgranulina C (S100A12) mediante interações com íons

Os resultados obtidos nas simulações de dinâmica molecular, em nossa contribuição acerca dos detalhes do mecanismo pelo qual os íons alteram a dinâmica conformacional da proteína S100A12, sugerem que em baixas concentrações de Ca²⁺, o L_I permanece ocupado pelo Na⁺. Essa hipótese também está presente no trabalho de Moroz, 2009; no qual a autora sugere ser possível que o Na⁺ ocupe o L_I do EF-1 em períodos entre ondas de Ca²⁺ [40].

Contudo, observamos aqui, que esta hipótese pode ser extrapolada, pois mostramos também a participação do Zn²⁺ na saída de Na⁺ do L_I. Como evidenciado em nossas simulações, em determinadas concentrações de ZnCl₂, a saída do Na⁺ ocorre muito rapidamente.

Dessa forma, no período entre ondas de Ca²⁺, este íon tem acesso à proteína exclusivamente pelo L_{III} (no EF-2). A medida em que há presença de Zn²⁺, esse contribui para a saída do Na⁺ do L_I, evento que envolve a participação do resíduo Asp25, permitindo que o L_I se abra e descomplexe o Na⁺. Assim, o L_I fica disponível para complexar Ca²⁺. Dessa forma, em situações fisiopatológicas nas quais há aumento de concentração intracelular de Ca²⁺ e Zn²⁺, como ocorre durante a resposta inflamatória, a presença do Zn²⁺ facilita a saída do Na⁺ facilitando a complexação do Ca²⁺ na S100A12. Esse comportamento, constitui uma hipótese para o aumento de afinidade pelo Ca²⁺ verificada na presença de Zn²⁺, uma vez que, no trabalho de Dell'Angelica, 1994 [23], a autora afirma que o Zn²⁺ induz um aumento notável na afinidade da S100A12 pelo Ca²⁺, cerca de 1500 vezes, e afirma ainda que na ausência de Zn²⁺ cada monômero complexa apenas um Ca²⁺.

Além disso, devido a alta deformabilidade estrutural, a H_{III} é muito influenciada pelos íons Na⁺ e Ca²⁺, sendo que em determinadas concentrações, ambos levam a perda parcial desta hélice e da H_{IIa} (“*Hinge-Region*”) e ao aumento da flexibilidade desta região, embora apenas o Ca²⁺ seja capaz de se complexar, via H_{III}, a região próxima ao L_{III}.

Além disso, mesmo após atingir a conformação complexada ao Ca²⁺, a H_{III} e a região do “*Hinge*” da S100A12 ainda são afetadas na presença de Na⁺, que leva a desorganização do núcleo hidrofóbico que se forma no estado complexado ao Ca²⁺. Assim, provavelmente a H_{III} e o “*Hinge*” apresentam uma dinâmica de formação com perda parcial do conteúdo de hélice na presença de

Na⁺, sendo que os eventos da perda e formação da hélice H_{IIa} facilitam o entendimento da complexação e descomplexação de Ca²⁺ pela S100A12.

No entanto, quando há aumento da concentração de Ca²⁺, este se aproxima do L_{III} via H_{III} e altera decisivamente a flexibilidade desta região. O aumento de flexibilidade permite que estas regiões possam rotacionar com relação às H_{II} e H_{IV}, facilitando a entrada de Ca²⁺ no L_{III}. A ocorrência de todos esse eventos faz com que a proteína passe então do estado não complexado ao estado complexado ao Ca²⁺.

Ressaltamos por fim, que a H_{IIa} e a H_{III} e seu movimento de abertura e fechamento constituem o ponto chave na mudança conformacional da S100A12 no que se refere a passagem do estado não complexado para o estado complexado ao Ca²⁺. Assim, essa é a região de maior relevância do ponto de vista dinâmico estrutural, quando pensamos em planejamento de inibidores, que neste caso seriam importantes agentes antiinflamatórios, sobretudo no que se refere aos processos inflamatórios crônicos, nos quais, como sabemos, há aumento da expressão desta proteína.

CAPÍTULO 4

Estudo da interação da
proteína Calgranulina C (S100A12) com receptor RAGE

Este capítulo aborda nossos estudos preliminares da interação intermolecular protéica. Nossos principais objetivos aqui visam desenvolver metodologias de estudo e dar os primeiros passos para o entendimento da interação da S100A12 com o receptor RAGE.

4.1 Detalhes sobre a interação da S100A12 com o receptor RAGE

Tendo em vista a sequência do estudo proposto neste trabalho, começamos a investigar a respeito da interação da proteína S100A12 com seu receptor extracelular, o RAGE. Como já foi dito, a interação com o receptor RAGE representa uma função crucial em inflamações relacionadas à S100A12, uma vez que a S100A12, bem como algumas outras proteínas S100, pertence a uma nova classe de quimiocinas e contribui para a migração celular. Sabe-se que a interação da S100A12 com RAGE resulta na secreção de mediadores pró-inflamatórios, sendo o eixo RAGE-S100A12 importante no que se refere ao prolongamento dos eventos inflamatórios. Entretanto a natureza exata da interação S100A12-RAGE ainda não está definida [28].

Não existem disponíveis estruturas resolvidas dos complexos S100A12-RAGE, embora existam estruturas resolvidas da S100A12 e do RAGE separadamente em diversas condições. Assim, optamos por desenvolver uma estratégia que envolve a utilização de ferramentas de “*Docking*” Molecular, Dinâmica Molecular e simulações de SMD (“*Steered Molecular Dynamics*” - que em português pode receber o nome de Dinâmica Molecular de Caminho Induzido) para estudar a natureza dessa interação.

4.2 Estratégia para o estudo da interação RAGE-S100A12

A primeira etapa do procedimento utilizada neste trabalho, foi submeter as estruturas cristalográficas da S100A12 e do RAGE a um servidor de “*Docking*” Global (GRAMM-X [75]), para que fossem gerados modelos de complexos protéicos. Em seguida, reclassificamos os modelos obtidos de acordo com critério energético, utilizando o GROMACS para a etapa de minimização de energia dos complexos. Após análises dos melhores modelos, realizamos simulações de DM para três complexos, e utilizamos um trecho de cada simulação correspondente de DM para gerar as estruturas iniciais dos complexos nos quais realizamos as simulações de SMD, obtendo as trajetórias das simulações de dissociação. Essa estratégia segue esquematizada na figura 44.

Figura 44. Esquema utilizado para o estudo da complexação S100A12-RAGE.

Adiante detalharemos, do ponto de vista metodológico, as etapas esquematizadas na figura 44, seguida pela apresentação dos resultados de cada etapa.

4.2.1 “*Docking*” Molecular para gerar modelos dos complexos

Como não existem disponíveis estruturas resolvidas dos complexos S100A12-RAGE, embora existam estruturas resolvidas da S100A12 e do RAGE separadamente, em diversas condições, optamos por utilizar um serviço de “*Docking*” Global, de acesso livre na internet [75], para gerar modelos de complexos que pudessem servir de ponto de partida para nosso estudo.

O “*Global Range Molecular Matching*” (GRAMM-X) foi o programa escolhido para o “*docking*” de proteínas. O GRAMM-X necessita das coordenadas atômicas de duas moléculas para prever possíveis estruturas do complexo e consiste em uma interface para Web de um software para “*Docking*” Molecular. O GRAMM-X difere do GRAMM original, mas ambos utilizam os pacotes FFT (“*Fast Fourier Transformation*” - Transformada rápida de Fourier) para a pesquisa global de conformações sem levar em conta a flexibilidade das macromoléculas. O GRAMM-X especificamente, é projetado exclusivamente para ancoragem de pares de moléculas de natureza protéica.

No método original do GRAMM, a energia potencial intermolecular é uma aproximação do potencial de Lennard-Jones. A suavização da superfície de energia intermolecular é alcançada pela redução do valor da parte repulsiva do potencial.

Na utilização desse serviço, é possível fornecer uma série de especificações prévias, por exemplo, a respeito de quais são os resíduos da interface do complexo, caso já se conheça algum detalhe sobre como se dá a complexação entre as proteínas. No caso do nosso estudo, como possuíamos indícios pouco claros, o que fizemos foi submeter as estruturas da S100A12 (PDB ID:

2WCE, 1E8A e 1GQM) e do RAGE (PDB ID: 3O3U [76]). Foram submetidos então três trabalhos no GRAMM-X:

→ **“Docking” A:** 2WCE (S100A12 resolvida na ausência de Ca^{2+}) e 3O3U (Domínios V e C1 do RAGE);

→ **“Docking” B:** 1E8A (S100A12 com Ca^{2+} complexados aos “EF-Hands”) e 3O3U;

→ **“Docking” C:** 1GQM (hexâmero de S100A12 que segundo referências [32] é o estado oligomérico que se complexa ao RAGE) e 3O3U.

Devemos ressaltar que a estrutura cristalográfica do RAGE utilizada (PDB ID: 3O3U) apresentava os domínios V e C1 do RAGE (domínios com os quais a S100A12 possivelmente interage) complexados a uma MBP (*“maltose-binding protein”* - proteína ligação de maltose), utilizada como modelo de busca para a determinação da estrutura por Cristalografia e Difração de Raios-X utilizando substituição molecular. Para submissão no GRAMM-X retiramos a porção da estrutura referente a MBP, de forma que restaram os domínios V e C1 do RAGE.

Foram gerados 300 modelos de complexação para cada um dos trabalhos submetidos, classificados inicialmente de acordo com uma função *“score”* própria do *“docking”*.

4.2.2 Reclassificação e seleção dos melhores modelos

Tendo em vista os modelos gerados pelo GRAMM-X e considerando o fato destes consistirem em resultados de *“Docking”* Global, onde nenhuma informação sobre a região de interação entre as proteínas foi fornecida, consideramos reclassificar os modelos de acordo com a energia potencial de interação proteína-proteína calculada após uma etapa de minimização de energia usando o GROMACS, e assim dispor de mais um critério para a seleção final.

Todos os 900 modelos, 300 para cada sistema (2WCE+RAGE, 1E8A+RAGE e 1GQM+RAGE) foram submetidos a um *“script”* que construía a caixa de simulação, colocava o complexo dentro da caixa, solvatava o sistema com moléculas de água, adicionava íons e efetuava uma minimização de energia, da mesma forma como foi descrito para as simulações de DM anteriormente (detalhes no capítulo 2). Contudo, neste caso, não era executada a DM propriamente dita, apenas a etapa de minimização. Após esta etapa, calculamos a energia potencial total da parte protéica do sistema e a energia potencial da seleção que continha a S100A12 e da seleção que continha o RAGE.

Assim, a energia potencial de interação do complexo foi considerada como sendo:

$$\text{Energia Potencial de Interação} = (\text{Energia Potencial Total}_{\text{PROTÉICA}}) - (\text{Energia Potencial}_{\text{S100A12}} + \text{Energia Potencial}_{\text{RAGE}})$$

Quando organizamos os complexos em termos de ordem crescente de energia potencial de interação (dada em kcal/mol) temos (figura 45):

Figura 45. Reclassificação das respostas geradas pelo “*docking*” após minimização de energia no GROMACS em termos de energia potencial de interação proteína-proteína.

De acordo com a figura anterior, observamos que em termos de energia potencial de interação, vários modelos de complexos para o sistema RAGE+1GQM, apresentam energias menores quando comparados aos complexos RAGE+2WCE e RAGE+1E8A.

Uma questão relevante nesse sentido, é que quando selecionados os melhores modelos em termos de energia potencial de interação, eles não condizem com as melhores respostas do “*docking*”. Os dez melhores modelos para cada sistema em termos de energia potencial entre moléculas, e a ordem original em termos do resultado gerado pelo “*docking*”, são mostrados na tabela 2.

Número do Modelo do "Docking"	Energia Potencial de Interação RAGE-2WCE (kcal/mol)	Número do Modelo do "Docking"	Energia Potencial de Interação RAGE-1E8A (kcal/mol)	Número do Modelo do "Docking"	Energia Potencial de Interação RAGE-1GQM (kcal/mol)
217	-1441,34	161	-1553,49	60	-2356,58
247	-1435,88	230	-1524,55	257	-2155,31
222	-1405,32	25	-1496,21	149	-2135,73
263	-1286,48	242	-1442,71	50	-2119,43
203	-1279,11	115	-1432,31	223	-2116,28
227	-1212,21	268	-1387,27	252	-2114,17
195	-1160,17	49	-1380,47	44	-2071,90
270	-1156,63	28	-1356,46	82	-1967,66
33	-1155,02	261	-1347,90	64	-1944,40
238	-1144,25	182	-1343,85	68	-1903,85

Tabela 2. Comparação entre a classificação em termos de menores valores energia potencial de interação proteína-proteína (10 melhores modelos de cada sistema) e a posição na classificação dos mesmos modelos nas respostas fornecidas pelo GRAMM-X.

Como nosso objetivo era estudar os detalhes da complexação para os três complexos (2WCE+RAGE, 1E8A+RAGE e 1GQM+RAGE) utilizando simulações de SMD e para tanto seria necessário realizar uma simulação prévia de DM dos complexos. Considerando a disponibilidade de recursos computacionais e o tamanho dos sistemas a serem simulados, foi necessário selecionar apenas um dos modelos gerados para cada sistema. Com cada um desses modelos foi realizada uma simulação de DM de 100ns que antecedeu a simulações de SMD.

Alinhando as estruturas dos 3 melhores complexos disponíveis em termos de energia potencial de interação para cada sistema, percebemos que as configurações dos melhores modelos são parecidas no que se refere à região de complexação da S100A12 ao RAGE (Figura 46), o que até melhores avaliações, parece ser um bom indicativo da relevância do critério energético utilizado na seleção dos modelos de complexação.

Figura 46. Alinhando as estruturas dos 3 melhores complexos disponíveis em termos de energia potencial de interação para cada sistema (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). O RAGE está representado em cinza, enquanto 2WCE, 1E8A e 1GQM estão representadas em verde, azul e rosa (sendo essa representação equivalente as configurações de menores energias de interação de acordo com nossa reclassificação).

Embora existam diferenças, de forma geral a 2WCE, a 1E8A e a 1GQM (hexâmero) interagem principalmente, de acordo com os melhores modelos gerados, na região do domínio V do RAGE e também em uma região entre os domínios V e C1 do RAGE. A interface de interação quando o complexo é formado pelo dímero (1E8A e 2WCE) é consideravelmente diferente da interface quando a interação se dá por meio do hexâmero (1GQM). Contudo, maiores detalhes acerca da região da interface do complexo serão mostradas e comparadas com o que já existe na literatura quando mostrarmos os resultados das simulações de SMD.

4.2.3 Simulações de DM dos melhores modelos

Considerando as semelhanças quanto à região de complexação ao RAGE dos modelos, optamos por escolher os modelos de menores energias de interação para cada sistema, para que fossem submetidos a uma simulação de DM de 100ns. Assim possuíamos 3 sistemas RAGE+2WCE, RAGE+1E8A e RAGE+1GQM que foram submetidos a simulação de DM com os mesmos parâmetros já descritos para as simulações da S100A12 com íons, no capítulo anterior. Adiante segue o esquema geral das simulações com os detalhes dos sistemas simulados (Figura 47).

Figura 47. Sistemas simulados (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). Os sistemas são compostos pelas diferentes formas de S100A12, RAGE, moléculas de água (modelo “SPC”) e íons Na^+ e Cl^- , na concentração de 50mM. Os sistemas RAGE+2WCE e RAGE+1E8A são formados por pouco mais de 69000 átomos dos quais 6192 átomos são das proteínas, enquanto 63000 átomos (21000 moléculas) são do solvente. O sistema RAGE+1GQM é formado por cerca de 99000 átomos dos quais 11714 átomos são das proteínas, enquanto 87000 átomos (29000 moléculas) são do solvente.

Com base nas análises dos RMSD das simulações (figura 48), escolhemos trechos das simulações para amostrar 100 configurações de cada complexo que foram usadas nas simulações de SMD.

Figura 48. Deformação estrutural analisada de acordo com os valores de RMSD para os complexos RAGE+2WCE, RAGE+1E8A, RAGE+1GQM.

Para o sistema RAGE+2WCE escolhemos o trecho entre 40 a 80ns e obtivemos 100 configurações do complexo, cada uma a cada 400ps de simulação; para o sistema RAGE+1E8A escolhemos o trecho entre 20 a 100ns e obtivemos 100 configurações do complexo cada uma a cada 800ps de simulação; e para o sistema RAGE+1GQM escolhemos o trecho entre 60 a 100ns e obtivemos 100 configurações do complexo cada uma a cada 400ps de simulação, esses intervalos foram escolhidos uma vez que representam trechos nos quais não houve variações nos valores de RMSD para o complexo e para a S100A12 e RAGE separadamente.

4.2.4 Simulações de SMD dos melhores modelos para obtenção de detalhes da complexação

Nessa metodologia, uma força externa é aplicada a uma região definida, para induzir essa região a deslocar-se em um percurso relacionado ao fenômeno que se deseja estudar, no caso do

nosso estudo, foram feitas simulações de dissociação do complexo, para estudar os mecanismos de complexação. Os detalhes acerca das características do caminho não precisam ser todos conhecidos, para que essas simulações sejam realizadas [77].

Essa força que se aplica a determinada região deriva-se de um potencial harmônico, de forma que é como se prendêssemos o sítio escolhido a uma das extremidades de uma mola virtual e, na outra extremidade, há um sítio sendo puxado. Esse sítio é movido a velocidade constante e a força entre ambos é medida de acordo com:

$$\vec{F}_{ext}(t) = k(\vec{v}t - \Delta \vec{x}) \quad (\text{Eq.13})$$

Onde k é uma constante de força, $\vec{v}$ é o vetor velocidade do sítio sendo puxado e $\vec{x}$ é o deslocamento deste em relação à sua posição inicial, sendo que a direção em que esta força é aplicada é definida pelo vetor $\vec{x}$.

Um ponto fundamental nas simulações de SMD refere-se à aplicação da igualdade de Jarzynski [78]. É por meio dessa igualdade que podemos relacionar a variação de energia livre e o trabalho realizado em um processo fora do equilíbrio.

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta G} \quad (\text{Eq.14})$$

Onde $\beta = 1/k_B T$ e W é o trabalho externo, calculado pela integral da força ao longo de um caminho de reação escolhido.

Dessa forma, realiza-se um conjunto de simulações de SMD ao longo desta coordenada de reação e calcula-se os valores de trabalho W que entram no cálculo da média representada pela igualdade de Jarzynski [79-81].

Com relação aos detalhes metodológicos das simulações, foram obtidas 100 configurações iniciais para cada sistema (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM) a partir das simulações de DM dos complexos. Cada uma dessas configurações iniciais foi submetida a um “*scprit*” no qual o complexo era recolocado em uma caixa de simulação de dimensões suficientes para que pudesse ser feita a dissociação do complexo de acordo com a distância que queríamos separar. Os parâmetros das simulações de SMD foram os mesmos das simulações de DM para os complexos selecionados após o “*docking*”. O sítio escolhido para ser fixado foi o centro de massa do RAGE, enquanto o sítio puxado foi o centro de massa da S100A12 (2WCE, 1E8A e 1GQM). Exemplos dos

sistemas que foram submetidos as simulações de SMD, com os detalhes da simulação, seguem na figura 49.

Figura 49. Exemplos de sistemas submetidos a simulações de SMD (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). Os sistemas são compostos pelas diferentes formas de S100A12, RAGE, moléculas de água (modelo “SPC”) e íons Na^+ e Cl^- , na concentração de 50mM. Os sistemas RAGE+2WCE e RAGE+1E8A são formados por pouco mais de 231000 átomos dos quais 6192 átomos são das proteínas, enquanto 225000 átomos (75000 moléculas) são do solvente. O sistema RAGE+1GQM é formado por cerca de 300000 átomos dos quais 11714 átomos são das proteínas, enquanto 288000 átomos (96000 moléculas) são do solvente.

Os sistemas foram submetidos a simulação de restrição por 100ps, seguida da etapa do SMD propriamente dita, que foi realizada em 200ps. O valor do vetor velocidade do sítio sendo puxado, $\vec{v}$, foi $0,02 \text{ nm/ps}$, de forma que ao longo dos 200ps de SMD o deslocamento deste sítio em relação à sua posição inicial, $\vec{x}$, foi de 4nm. A constante de força, k , foi $10000 \text{ kJ.mol}^{-1} \text{ nm}^{-2}$. A definição inicial destes parâmetros nas simulações de SMD sucedeu uma bateria de testes, na qual foram experimentados diversos valores de k de forma que fosse efetuada a dissociação sem que houvesse prejuízo significativo da estrutura tridimensional das proteínas envolvidas no complexo. Ao todo foram produzidas 100 simulações de SMD para cada sistema (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM). Na figura adiante, temos um exemplo da aplicação do SMD para cada sistema, para que possamos visualizar de que forma essas simulações foram produzidas (Figura 50).

Figura 50. Imagens de uma das 100 simulações de SMD para os diferentes sistemas (RAGE+2WCE, RAGE+1E8A e RAGE+1GQM) no início, aos 100 e 200ps de simulação.

As curvas do perfil do trabalho em função do deslocamento da mola, bem como as distribuições de frequência normalizadas dos trabalhos ao longo das simulações são mostradas adiante (Figura 51).

Figura 51. À esquerda observamos a coleção de perfis de trabalho ao longo da coordenada de reação para as simulações de SMD (estando o trabalho em kJ/mol – como é dado o resultado no GROMACS), enquanto à direita temos a distribuição de frequência normalizada dos trabalhos (já em kcal/mol) ao longo das simulações de SMD.

Considerando a aplicação da igualdade de Jarzynski seria possível estimar os valores de energia livre relativa de complexação, a partir da média exponencial dos valores de trabalho, ou

de uma expansão cumulante. Contudo, tendo em vista o fato das nossas simulações de SMD terem sido realizadas com uma velocidade bastante grande ($\vec{v}$ empregado foi de $0,02\text{ nm/ps}$) nossos valores de trabalho estão superestimados, de forma que não seriam coerentes os valores de energia livre calculados a partir destes dados. Esse valor de velocidade foi escolhido uma vez que as simulações de SMD deveriam ter um custo computacional razoável para nossas possibilidades. Contudo, outros trabalhos que aplicam simulações de SMD utilizam velocidades $\vec{v}$ da ordem de 10^{-4} nm/ps . Dessa forma, como a separação do complexo deveria ocorrer em um tempo muito restrito, os valores da força ao longo das simulações foram superestimado. Para que o mesmo deslocamento dos sítios em relação à posição inicial, ($\vec{x} = 4\text{nm}$) fosse obtido em simulações com velocidade $\vec{v}$ da ordem de 10^{-4} nm/ps seriam necessários 40ns para cada simulação. Considerando fato de que um número mínimo de trajetórias devem ser geradas, nesse trabalho fizemos 100 simulações para cada sistema, haveria necessidade da existência de maiores recursos computacionais, ou mais tempo para as simulações.

Devemos ressaltar que essa porção do nosso estudo, refere-se a uma abordagem inicial, realizada sobretudo para que a utilização desta ferramenta fosse aprendida e pudesse fazer parte do conjunto de ferramentas com as quais somos capazes de trabalhar em nosso grupo de simulação computacional.

Contudo, embora os valores de trabalho estejam superestimados, é possível obter a informação, a partir da análise da figura 51, que o trabalho necessário para dissociar o complexo RAGE+1QGM é maior do que o trabalho necessário para dissociar o complexo RAGE+2WCE e RAGE+1E8A.

Com relação às regiões da interface dos complexos temos uma diferença quando comparamos os sistemas RAGE+2WCE e RAGE+1E8A com o sistema RAGE+1QGM (Figuras 52 e 53).

Figura 52. Interface de complexação mostrada separadamente para os complexos, a saber 1E8A e RAGE e 2WCE e RAGE. Podemos observar que os resíduos envolvidos na complexação do RAGE com a 1E8A (resíduos em laranja) localizam-se principalmente no L_{II}, H_{III} e e H_{IV}, enquanto os resíduos envolvidos na complexação do RAGE com a 2WCE (resíduos em laranja) localizam-se em uma porção da H_{III} e H_{IV}. Com relação aos resíduos do RAGE envolvidos na complexação (resíduos em magenta), estão destacados nas figuras. Podemos observar que a complexação com o RAGE envolve resíduos pertencentes tanto ao domínio V, quanto ao domínio C1. O critério sobre os resíduos envolvidos na complexação baseou-se nas distâncias de interações entre resíduos do RAGE-S100A12 menores que 4Å.

Figura 53. Interface de complexação mostrada separadamente. Podemos observar que a complexação da 1QGM com o RAGE envolve resíduos que formam uma interface entre dois dímeros do hexâmero. Esses resíduos (destacados em laranja) localizam-se principalmente na H_I , L_I , L_{II} e H_{IV} . Os resíduos do RAGE envolvidos na complexação (em magenta), também estão destacados na figura. Podemos observar que a complexação com o RAGE envolve resíduos pertencentes principalmente ao domínio V, e a uma região de “loop” do domínio C1 que também auxilia na complexação do hexâmero da S100A12. O critério sobre os resíduos envolvidos na complexação baseou-se nas distâncias de interações entre resíduos do RAGE-S100A12 menores que 4Å.

Considerando que a complexação com o RAGE por parte da 1QGM (hexâmero) envolve a participação de resíduos que formam uma interface entre dois dímeros do hexâmero, podemos observar que na estrutura completa do hexâmero existem outras interfaces iguais para complexação (Figura 54).

Figura 54. Interface de complexação RAGE+1GQM. À esquerda, vemos detalhes do RAGE complexando resíduos da 1GQM (representação de superfície em verde). Ao centro, podemos observar a 1GQM (também em representação em superfície) complexada ao RAGE, estando colorido em laranja as outras duas faces da região entre os dímeros, equivalente a face em verde. E à direita, vemos a 1GQM (também em representação em superfície) complexada ao RAGE, estando colorido em laranja as três faces da região entre os dímeros que são regiões as quais o RAGE pode se complexar.

As interfaces de interação da 1E8A e 2WCE com o RAGE são formadas principalmente pelo L_{II} , H_{III} e H_{IV} , isto é, na interface do “EF-hand” canônico, com base na análise dos nossos resultados do “*docking*” e das simulações de SMD. Quando observamos a interface da interação da 1GQM (hexâmero) com o RAGE, as regiões de interação estão localizadas entre dois dímeros do hexâmero, envolvendo porções da H_I , L_I , L_{II} e H_{IV} .

4.3 Conclusões sobre a interação RAGE-S100A12

Estudos detalhados da interação da S100A12 com RAGE foram realizados por Xie e colaboradores [32]. Nestes estudos, envolvendo resultados de fluorescência de TNS (2-p-toluidina naftaleno-6-sulfonato), filtração em gel e RMN, o autor utilizou uma região diferente do RAGE, os domínios C1C2, enquanto nós utilizamos os domínios VC1. A escolha dos domínios C1C2 no trabalho de Xie, resultou do fato da região VC1 apresentar baixa solubilidade dificultando os estudos. No entanto, a interação de outros AGEs ao RAGE, tem sido reportada na região V [28]. Nos estudos de Xie [32], o autor constatou que a interação RAGE+S100A12 depende fundamentalmente da região C1, uma vez que na porção isolada de C2 não há sinal de interação. Por outro lado, Xie realizou também experimentos com o domínio V isolado, e também não observou sinais de complexação.

A análise dos nossos resultados preliminares, sugere que a interação da S100A12 com o RAGE ocorre tanto no domínio V, quanto no domínio C1 e depende da região de conexão entre esse domínios.

Outro ponto relevante, refere-se ao fato de que estados oligoméricos maiores, por exemplo, hexâmeros, tem possibilidades maiores de interação com RAGE. Nos estudos de Xie, o autor relata que com o aumento da concentração de Ca^{2+} a tendência é de que ocorra a oligomerização da S100A12, que passaria a existir na forma predominante de hexâmero. Dessa forma, embora as formas 2WCE e 1E8A interajam com o RAGE, a interação predominante ocorre via hexâmero, uma vez que momentos celulares, nos quais há aumento da expressão de RAGE e de S100A12, envolvem aumento da concentração de íons Ca^{2+} como discutido anteriormente (Capítulo 1). Nesse sentido, nossos resultados referente ao trabalho necessário para separar os complexos, embora com valores fora de escala, mostram que a separação do complexo RAGE+1GQM envolve maior gasto energia do que a separação dos complexos RAGE+2WCE e RAGE+1E8A.

Mostramos adicionalmente que a interface de interação RAGE+1GQM (hexâmero) permite a interação de mais de um RAGE, e isso está de acordo com o proposto também pelos estudos de Xie [32], que relatam o fato de que a formação do hexâmero da S100A12 pode levar a formação de oligômeros de RAGE para que estes possam acessar a S100A12 hexamérica em mais de uma interface, e isso seria responsável por criar uma região de sinalização grande, ampliando a transdução de sinal intracelular de forma eficaz.

Com relação a essa interface de interação RAGE+hexâmero, as regiões relevante envolvem, de acordo com nossos resultados, porções N e C-terminal da H_I e C-terminal da H_{IV} . Nesse sentido, nossos resultados também estão de acordo com o proposto por Xie, uma vez que o autor conclui que estas mesmas regiões são as relevante no que se refere a interação S100A12+hexâmero [32]. Tendo isso em vista, esses dados fornecem pontos relevantes no que se refere a estudos para bloquear a interação S100A12+RAGE, uma vez que como sabemos, essa interação relaciona-se a manutenção do processo inflamatório nos tecidos.

Dessa forma, podemos concluir que embora deva existir interação RAGE+dímero de S100A12, essa interação é mais fraca e se dá por uma superfície de interação bastante distinta do que é observada para a interação RAGE+hexâmero, sendo esta última mais efetiva e permitindo a criação de outros pontos para interação com outros receptores RAGE, possibilitando amplificar a transdução de sinal intracelular, que como vimos, está relacionada ao perdurar da resposta inflamatória, nos processos de inflamação crônica.

Referências Bibliográficas

1. BRUCE, A. et al. Molecular biology of the cell. New York: Garland (1983). p.493-498.
2. HODGKIN, A. L.; KEYNES, R.D. Movements of labeled calcium in squid giant axons. **J. Physiol**, 1957. 138: p.253-281
3. CARAFOLI, E. Intracellular calcium homeostasis. **Ann. Rev. Biochem**, 1987. 56: p.395-433.
4. HEIZMANN, C. W.; HUNZIKER, W. Intracellular calcium-binding proteins: more sites than insights. **Trends Biochem. Sci.**, 1991. 16: p.98-103.
5. KRETSINGER, R. H.; NELSON, D.J. Calcium in biological systems. **Coord. Chem. Rev.**, 1976. 18 (1): p.29-124.
6. KRETSINGER, R. H.; NOCKOLDS, C. E. Carp muscle calcium-binding protein: II. Structure determination and general description. **J. Biol. Chem.**, 1973. 248: p.3313-3326.
7. STRYER, L.; BERG, J. M.; TYMOCZKO, J. L. Biochemistry. New York (2001). Cap.15.
8. LEWIT BENTLE, A.; RÉTY, S. EFhand calcium binding proteins. **Curr. Opin. Struc. Biol.**, 2000. 10: p.637643.
9. LINSE, S.; FORSEN, S. Determinants that govern highaffinity calcium binding. **Adv. Second Messenger Phosphoprot. Res.**, 1995. 30: p.89151.
10. FALKE, J. J.; DRAKE, S. K.; HAZARD, A. L; PEERSEN, O. B. Molecular tunini of ion binding to calvium signaling proteins. **Quarterly Reviews of Biophysics.**, 1994. 27: p.219-290.
11. NAKAYAMA, N.; KAWASAKI, H.; KRETSINGER, R. Evolution of “EF hand” proteins. **Topics Biol. Inor. Chem.**, 2000. 3: p.29-58.
12. HAPAK, R. C.; LAMMERS, P. J.; PALMISANO, W. A.; BIRNBAUM, E. R.; HENZL, M. T. Site-specific substitution of glutamate for aspartate at position 59 of rat oncomodulin. **J. Biol. Chem.**, 1989. 264: p. 18751-18760.
13. CELIO, M.R. Calbindin D-28k and parvalbumin in the rat nervous system. **Neuroscience**, 1990. 35(2): p.375-475.
14. MOORE, B. W. A soluble protein characteristic of the nervous system. **Biochem. Biophys. Res. Commun.**, 1965. 19: p.739-744.
15. SEDAGHAT, F.; NOTOPOULOS, A. S100 protein family and its application in clinical practice. **Hippokratio General Hospital of Thessaloniki**, 2008. 12 (4): p.198-204.

16. PIETZSCH, J.; HOPPMANN S. Human S100A12: a novel key player in inflammation? **Amino Acids**, 2009. 36: p.381-389.
17. KLIGMAN, D.; HILT, D. C. The S100 protein family. **Trends. Biochem. Sci.**, 1998. 13: p.437-443.
18. MOROZ, O. V. et al. The three-dimensional structure of human S100A12. **Acta Crystallogr. Sect. D Biol. Crystallogr**, 2001. 57: p.20-29.
19. BHATTACHARYA, S.; CHRISTOPHER, G.; BUNICK, W.; CHAIN, J. Target selectivity in EF-hand calcium binding proteins. **Biochimica et Biophysics Acta.**, 2004. 1742: p.69-79.
20. FOELL, D.; WITTKOWSKI, H.; VOGI, T. and ROTH, J. S100 proteins expressed in phagocytes: a novel group of damage-associated molecular pattern molecules. **Journal of Leukocyte Biology**, 2007. 81: p.28-37.
21. FOELL, D; ROTH, J. Proinflammatory S100 Proteins in Arthritis and Autoimmune Disease. **Arthritis & Rheumatism**, 2004. 50 (12): p.3762–3771
22. GUINARD, F.; MAUEL, J.; MARKET, M. Identification and characterization of a novel human neutrophil protein related to the S100 family. **Biochem. J.**, 1995. 309: p.395-401.
23. DELL'ANGELICA, E. C.; SCHLEICHER, C. H.; SANTOMÉ, J. A. Primary structure and binding properties of calgranulin C, a novel S100-like calcium-binding protein from pig granulocytes. **J. Biol. Chem.**, 1994. 269: p.28929-28936.
24. SPARVERO, L. V. et al. RAGE (Receptor for Advanced Glycation Endproducts), RAGE Ligands, and their role in Cancer and Inflammation. **Journal os Translational Medicine**, 2009. 7: p.17.
25. WEIGAND, M. A.; BOPP, C.; BÄTTIGER, B. W. The Emerging Role of RAGE in Sepsis. **Intensive care medicine**, 2006. p.14-22
26. SCHMIDT, A. M.; YAN, S. D.; YAN, S. F.; STERN, D. M. The multiligand receptor RAGE as a progression factor amplifying immune and inflammatory responses. **The Journal of Clinical Investigation**, 2001. 108 (7): p.949-955.
27. SIMM, A.; BARTLING, B.; SILBER, R. E., RAGE: a new pleiotropic antagonistic gene? **Ann N Y Acad Sci**, 2004. 1019: p. 228-231 .
28. SCHIMIDT, A. M. et al. RAGE Mediates a Novel Proinflammatory Axis: A Central Cell Surface Receptor for S100/Calgranulin Polypeptides.
29. MOGENSEN, T. H., Pathogen Recognition and Inflammatory Signaling in Innate Immune Defenses, 2009. 22 (2): p.240-273.

30. BIANCHI, M. E. DAMPs, PAMPs and alarmins: all we need to know about danger. **Journal of Leukocyte Biology**, 2007. 81(1): p.1-5
31. NAWROTH, P. P. et al. Understanding RAGE, the receptor for advanced glycation end products. **J Mol Med**, 2005. 83: p.876–886.
32. XIE, J. et al. Hexameric calgranulin C (S100A12) binds to the receptor for advanced glycosylated end products (RAGE) using symmetric hydrophobic target-binding patches. **J. Biol. Chem**, 2007. 282 (6): 4218-4231.
33. LECLERC, E.; FRITZ, G.; VETTER, S. W.; HEIZMANN, C. W. Binding of S100 proteins to RAGE: An update. **Biochimica et Biophysica Acta**, 2009. 1793: p.993-1007.
34. NATHAN, C. Neutrophils and immunity: challenges and opportunities. **Nature reviews immunology**, 2006. 6: p.173 182.
35. TINTINGER, G. R.; STEEL, H. C.; THERON, A. J.; ANDERSON, R. Pharmacological control of neutrophil-mediated inflammation: Strategies targeting calcium handling by activated polymorphonuclear leukocytes. **Drug, Design, Development and Therapy**, 2008. 2: p.95-104.
36. PRASAD, A. S. Zinc in Human Health: Effect of Zinc on Immune Cells. **Molecular Medicine**, 2008. 14: 353-357.
37. BELL, S. G.; VALLEE, B. L. The Metallothionein/Thionein System: An Oxidoreductive Metabolic Zinc Link. **ChemBioChem**, 2009. 10: p.5562.
38. NONATO, M. C., (1997) Estudos Cristalográficos em Macromoléculas biológicas: aplicações em Calgranulina C de granulócitos porcinos, Tripanotona redutase de Trypanosoma cruzi e Fosfolipase A2 extraída do veneno da serpente Bothrops moojeni. Tese (Doutorado) – Instituto de Física de São Carlos.
39. MOROZ, O. V. et al. The crystal structures of human S100A12 in apo form and in complex with Zinc: New insights into S100A12 oligomerisation. **J. Mol. Biol.**, 2009. 391: 536-551.
40. MOROZ, O. V. et al. Both Ca^{2+} and Zn^{2+} are essential for S100A12 protein oligomerization and function. **BMC Biochemistry**, 2009. 10 (11): p.1-18
41. MAGINN, E. J. From discovery to date: What must happen for molecular simulation to become a mainstream chemical engineering tool. **AIChE Journal**, 2009. 55: p.1304-1310.
42. ADCOCK, S. A.; MCCAMMON, J. A. Molecular Dynamics: Survey of methods for simulating the activity of proteins. **Chem. Rev.**, 2006. 106: 1589-1615.
43. PASCUTTI, P. G. Introdução à modelagem e dinâmica molecular, 2002. 1, p.1-38.

44. BARNETT, R. N.; LANDMAN, U. Born-Oppenheimer Molecular-Dynamics Simulations of Finite Systems - Structure and Dynamics of (H₂O)₂. **Physical Review B**, 2003. 48(4): p.2081-2097.
45. FRENKEL, D.; SMIT, B. Understanding Molecular Simulation. New York: Academic Press (1996).
46. KAMINSKI, G.A. et al. Evaluation and reparametrization of the OPLSAA force field for proteins via comparison with accurate quantum chemical calculations on peptides. **Journal of Physical Chemistry B**, 2001. 105(28): p. 6474-6487.
47. ARAUJO, A. S. (2006) Estudo do processo de complexação de calixarenos com íons metálicos e espécies neutras por simulações de dinâmica molecular. Tese (Doutorado) – Instituto de Física de São Carlos.
48. LEACH, A. R. Molecular Modelling: Principles and Applications. Pearson Education: Harlow (2001).
49. SANTOS, E.S. (2007) Estudo da Estabilidade Térmica da Redutase da Adenosina 5'-Fosfosulfato-reduzida em meio aquoso via Dinâmica Molecular . Dissertação (Mestrado) – Instituto de Física – Universidade Federal da Bahia.
50. VAN DER SPOEL, D. et al. Gromacs User Manual version 4.0, www.gromacs.org.
51. RAPAPORT, D. C. The Art of Molecular Dynamics Simulation. Cambridge University Press (2004).
52. SOUZA, P. C. T. (2009) Simulações de Dinâmica Molecular de Mutantes do Receptor do Hormônio Tireoideano. Dissertação (Mestrado). Departamento de Físico-Química - Universidade Estadual de Campinas .
53. ESSMANN, et al. A Smooth Particle Mesh Ewald Potential. **J. Chem. Phys.**, 1995, 103, pp. 8577.
54. ANDRADE, A. S. C (2008) Dinâmica molecular do fármaco antitumoral ecteinascidina e sua interação com o DNA. Dissertação (Mestrado). Departamento de Química - Universidade Federal do Rio Grande do Sul.
55. NAMBA, A. M.; SILVA, V. B.; SILVA, C. H. T. P. Dinâmica Molecular: teoria e aplicação em planejamento de fármacos. **Eclética Química**, 2008. 33 (4): p. 13-24.
56. MIYAMOTO, S.; KOLLMAN, P. A. SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. **J. Comp. Chem**, 1992. 13: p.952–962.
57. HESS, B.; BEKKER, H.; BERENDSEN, H. J. C.; FRAAIJE, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. **J. Comp. Chem.**, 1997. 18: p. 1463–1472.

58. RYCKAERT, J. P.; CICCOTTI, G.; BERENDSEN, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. **J. Comp. Phys.**, 1997. 23: p. 327-341.
59. BERENDSEN, H. J. C.; POSTMA, J. P. M., et al. Molecular-Dynamics with Coupling to an External Bath. **Journal of Chemical Physics**, 1984. 81(8): p. 3684-3690.
60. NOSE, S. A. Molecular dynamics method for simulations in the canonical ensemble. **Mol. Phys.**, 1984. 52: p.255–268.
61. HOOVER, W. G. Canonical dynamics: equilibrium phase-space distributions. **Phys. Rev.**, 1985. 31:p. 1695–1697.
62. BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **J. Chem. Phys.**, 2007. 126 (1): 014101.
63. FRISHMAN, D.; ARGO, P. Knowledge Based Protein Secondary Structure Assignment. **Proteins: Structure, Function, and Genetics**, 1995. 23: P.566-579.
64. BAKER, N. A. et al. Electrostatics of nanosystems: application to microtubules and the ribosome. **Proc. Natl. Acad. Sci.**, 2001, 98(18): p.10037-10041.
65. VAN DER SPOEL, D.; HESS, B. GROMACS - the road ahead. **Computational Molecular Science**, 2011. 1(5): p.710-715.
66. HESS, B. et al. GROMACS 4: Algorithms for highly efficient, loadbalanced, and scalable molecular simulation. **Journal of Chemical Theory and Computation**, 2008. 4(3): p. 435447.
67. VAN DER SPOEL, D. et al. GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, 2005. 26(16): p. 17011718.
68. LINDAHL, E.; HESS, B.; VAN DER SPOEL, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. **Journal of Molecular Modeling**, 2001. 7(8): p. 306-317.
69. BEKKER, H. et al. Gromacs a Parallel Computer for MolecularDynamics Simulations. **Physics Computing**, 1993. 517: p. 252-256 .
70. MARTIN, M. G. Comparison of the AMBER, CHARMM, COMPASS, GROMOS, OPLS, TraPPE and UFF force fields for prediction of vapor-liquid coexistence curves and liquid densities. **Fluid Phase Equilibria**, 2006. 248(1): p.50-55.
71. DELANO, W.L.; LAM, J. W. PyMOL: A communications tool for computational models. **Abstracts of Papers of the American Chemical Society**, 2005. 230: p.U1371-U1372.

72. HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, 1996. 14(1): p. 33-&.
73. TURNER, P. J. Center for Coastal and Land-Margin Research Oregon Graduate Institute of Science and Technology Beaverton, Oregon. <http://plasma-gate.weizmann.ac.il/Grace>.
74. VASILEF, I. QtiPlot: data analysis and scientific visualization, 2009. <http://soft.proindependent.com/qtiplot.html>.
75. TOVCHIGRECHKO A.; VAKSER, I. A. GRAMM-X public web server for protein-protein docking. **Nucleic Acids Res**, 2006. 34: p.310-314.
76. PARK, H.; BOYINGTON, J. C. The 1.5 Å Crystal Structure of Human Receptor for Advanced Glycation Endproducts (RAGE) Ectodomains Reveals Unique Features Determining Ligand Binding . **The journal of biological chemistry**, 2010. 285 (52): p. 40762–40770.
77. CUENDET, M. A.; MICHIELIN, O. Protein-Protein Interaction Investigated by Steered Molecular Dynamics: The TCR-pMHC Complex. **Biophysical Journal**, 2008. 95: p.3575–3590.
78. JARZYNSKI, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. **Phys. Rev.**, 1995. 56(5): 5018-5035
79. ECHEVERRIA, I.; AMZEL, L. M. Helix propensities calculations for amino acids in alanine based peptides using Jarzynski's equality. **Proteins**, 2010. (78): p.1302–1310.
80. HUNG, S. W.; HSIAO, P. Y.; CHIENGA, C. C. Dynamic information for cardiotoxin protein desorption from a methyl-terminated self-assembled monolayer using steered molecular dynamics simulation. **The journal of chemical physics**, 2011. 134: p.194705.
81. PARK, S.; SCHULTEN, K. Calculating potentials of mean force from steered molecular dynamics simulations. **J. Chem. Phys.**, 2004. 120 (13): p. 5946-5961.