

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA



Ciências
ULisboa

**Modelação molecular/Bioinformática estrutural da
hemaglutinina do vírus influenza**

Mestrado em Bioquímica
Especialização em Bioquímica Médica

Jorge Miguel Penela de Deus Antunes

Dissertação orientada por:
Doutor Cláudio Soares
Doutor Francisco Pinto

Resumo

O vírus influenza é responsável por vários surtos de gripe a nível mundial, como o recente surto causado pelo H1N1. A infeção começa com a ligação da proteína Hemaglutinina (HA) do envelope viral ao ácido siálico presente na membrana das células epiteliais do trato respiratório. O vírus é então internalizado por endocitose. No endossoma a diminuição de pH provoca alterações conformacionais da HA, levando por último à fusão das membranas.

Neste trabalho reportamos um estudo de simulação molecular tanto da proteína HA, como também de dois péptidos específicos. O péptido de fusão (PF), que apresenta propriedades fusógenicas *in vitro*, sendo considerado o principal responsável pela fusão das membranas. E o péptido transmembranar (PT), responsável pela ligação da hemaglutinina ao envelope viral, que não possui estrutura conhecida, tanto a pH neutro como a pH de fusão. Nos estudos dos péptidos foram utilizados dois campos de força e níveis de detalhe distintos. Um campo de forças *coarse-grain* (MARTINI), mais rápido mas com menos detalhe, foi utilizado para simular a montagem do sistema e um campo de forças *united-atom* (GROMOS 54A7) para simular os resultados obtidos com um detalhe superior. Nos estudos da HA foi utilizado apenas o campo de forças *coarse-grain*, com a modificação para incluir redes elásticas (ELNEDIN).

Com este estudo apresentamos estruturas, e algumas variações, para ambos os péptidos numa bicamada composta por 128 moléculas de dimiristoilfosfatidilcolina (DMPC). No caso do PT foi a primeira estrutura obtida com o C-terminal completo (46 resíduos) e também do subtipo H3. Foram ainda calculados os ângulos de inclinação e a posição, de cada resíduo, na membrana para ambos os péptidos estudados. No caso da HA o resultado principal foi o esclarecimento dos locais de contacto preferenciais entre esta e os lípidos (DMPC). Observou-se que, para além do domínio transmembranar, também existem resíduos à volta do local de ligação ao ácido siálico com afinidade para o DMPC.

Palavras-chave: Vírus Influenza, Hemaglutinina, Dinâmica molecular, MARTINI

Abstract

The influenza virus (IV) is responsible for worldwide flu outbreaks like the recent one due to H1N1. Infection starts with the binding of the membrane-bound virus Hemagglutinin (HA) to sialic acid at the surface of membranes from epithelial cells of the respiratory tract. The virus is then internalized into an endosome, where the low pH triggers conformational changes in HA, ultimately leading to membrane fusion.

Here we report a molecular simulation study, not only of the HA, but also of two specific peptides. The fusion peptide (FP), which has fusogenic properties *in vitro*, and is considered to have a main role in the membrane fusion process. And the transmembrane peptide (TP), responsible for the connection between HA and the viral envelope, which has no known structure. We used a multi-scale approach in the studies with the peptides, by using two different force-fields with different levels of detail; a coarse-grain force field (MARTINI), which is faster but with less detail, to study the self-assembling of the system, and an united-atom force-field (GROMOS 54A7), to study with more detail the results obtained using the MARTINI force-field. In the whole HA studies, we used the coarse-grain force-field only, but with the modification to include elastic networks (ELNEDIN).

With this work we found some preferential structures and some variations, for both fusion and transmembrane peptides, in a bilayer composed by 128 molecules of dimyristoylphosphatidylcholine (DMPC). It is the first structure of the TP obtained with the complete C-terminus (46 residues) and also for the subtype H3. The tilt angles and the membrane position, of each residue, of both peptides studied were also calculated. In the case of the HA, the main result was to find that main sites of contact between this protein with lipids (DMPC). We found that apart from the transmembrane domain, the residues around the sialic acid binding site have also affinity to the DMPC lipids.

Key words: Influenza virus, Hemagglutinin, Molecular dynamics, MARTINI

Agradecimentos

Primeiro de tudo gostaria de agradecer aos meus orientadores externos, Bruno Victor e Cláudio Soares por despendarem do seu precioso tempo para me ensinar e orientar ao longo desta aprendizagem. Não esquecendo o meu orientador interno, Francisco Pinto, que se dispôs a tal sem qualquer entrave.

Além dos meus orientadores também devo parte da minha aprendizagem aos meus colegas do laboratório, João, Carla, Diana, Sofia, bem como o pessoal do laboratório de simulação molecular, Pedro, Catarina, Sara, Luís e o António, por estarem sempre predispostas a ajudar ou tirar qualquer dúvida que tivesse.

Fora o apoio/ajuda com o trabalho, tenho também a agradecer a quem me apoiou emocionalmente, os meus amigos, Daniela, Catarina (aka Naci), Ana, Luís e Sara que sempre me apoiaram e acreditaram em mim, mesmo quando eu não acreditava em mim próprio.

Por ultimo e não menos importante a minha família, que se não fosse por causa e por eles eu não teria conseguido entregar este trabalho. Fazendo menção honrosa à minha irmã que além do grau de parentesco é também a minha melhor amiga, e a minha avó, a quem tudo devo.

Seria ingrato se também não agradecesse ao Bruno, que surtiu de certo modo um abaixamento de pH ao péptido de fusão.

Índice

Resumo	iii
Abstract	v
Agradecimentos.....	vii
1. Introdução	1
1.1 Vírus influenza	1
1.2 Hemaglutinina	5
1.3 Ciclo de vida do vírus	8
1.3.1 Clivagem da hemaglutinina	8
1.3.2 Ligação ao recetor	10
1.3.3 Entrada na célula	14
1.4 Alterações conformacionais – mecanismo de <i>spring-loaded</i>	14
1.5 Péptido de fusão	24
1.6 Péptido transmembranar	27
1.6.1 Domínio transmembranar	27
1.6.2 Cauda citoplasmática	30
1.6.3 <i>Linker</i>	32
2. Modelação e simulação molecular	33
2.1 Mecânica molecular	35
2.1.1 Função de energia potencial empírica	36
2.1.2 Dinâmica molecular	41
2.1.3 Solvatação e condições periódicas de fronteira	44
2.1.4 Banhos de temperatura e pressão	46
2.2 Campos de força não atomísticos	48
2.2.1 Campos de força united-atom	49
2.2.2 Campos de força MARTINI	50
2.2.3 Campos de força ELNEDIN	56
2.3 Métodos	58
2.3.1 Preparação/Solvatação	58
2.3.2 Minimização	59
2.3.3 Inicialização	60
2.3.4 Corrida de produção	62

2.4	Transformação reversa	62
2.5	Técnica de “crescimento lento”	65
3.	Tratamento de resultados	67
3.1	Péptido de fusão	67
3.1.1	Péptido de fusão em água usando o campo de forças 54A7.....	68
3.1.2	Péptido de fusão em água usando o campo de forças MARTINI.....	76
3.1.3	Self-assembling de lípidos	78
3.1.4	Interação do péptido de fusão com a bicamada de DMPC ...	82
3.1.5	Análises	84
3.1.5.1	Partição na membrana – profundidade de inserção.....	84
3.1.5.2	Conteúdo em hélice	88
3.1.5.3	Ângulo de inclinação	92
3.1.5.4	Conformações finais	99
3.2	Péptido transmembranar	108
3.2.1	Construção do péptido transmembranar	109
3.2.2	Self-assembling do péptido transmembranar	111
3.2.3	Análises	115
3.2.3.1	Conformação	115
3.2.3.2	Conteúdo em hélice	117
3.2.3.3	Ângulo de inclinação	121
3.2.3.4	Partição na membrana	127
3.2.3.5	Parâmetros de ordem das caudas dos lípidos ...	132
3.3	Hemaglutinina	135
3.3.1	Estabilidade em água – aplicação de ligações elásticas	135
3.3.2	Alteração conformacional com o método de crescimento lento.....	139
3.3.3	Self-assembling da sequência completa da hemaglutinina ..	142
4.	Conclusões	148
5.	Bibliografia	151

1. Introdução

1.1. Vírus Influenza

O Vírus Influenza (VI), mais conhecido por vírus da gripe, é responsável por infecções que ocorrem sazonalmente, provocando normalmente febre, tosse, dor corporal e corrimento nasal [Luo 2012]. Podem, no entanto ocorrer epidemias que provocam complicações respiratórias/pulmonares graves, levando inclusivamente à morte. O VI apresenta uma mortalidade particularmente elevada em indivíduos com certas condições médicas de alto risco, incluindo imunodeficientes, disfunções renais, problemas cardiovasculares e pulmonares como a asma, ou mesmo crianças saudáveis com uma idade inferior a dois anos [Treanor 2012].

O VI faz parte da família Orthomyxoviridae que inclui 5 géneros: Influenzavirus A, influenzavirus B, influenzavirus C, Isavirus e Thogotovirus. O VI é dividido então em três tipos: A, B e C. Os VI do tipo B e C são menos comuns e causam efeitos mais suaves que o tipo A. Para além destes dois apenas infetarem mamíferos, não existem subtipos destes vírus e não causam pandemias. A maior diferença entre os VI do tipo C e o tipo A ou B é a glicoproteína na superfície viral. O VI do tipo C contém apenas uma glicoproteína em contraste com os outros tipos que possuem duas. Esta proteína possui uma estrutura semelhante à hemaglutinina, mas além da atividade da hemaglutinina também apresenta atividade da neuramidase. Na verdade a proteína hemaglutinina dos tipos A e B têm um domínio que possui semelhanças a um centro ativo com função de esterase, sugerindo que esses dois tipos tiveram como antepassado o tipo C [Luo 2012].

No decorrer desta dissertação será tratado o VI do tipo A, exceto se for referido especificamente o tipo. Ao contrário dos outros tipos, o principal organismo alvo do tipo A são as aves, mas os humanos, porcos, cavalos e mamíferos aquáticos (por exemplo focas) também podem ser infetados [Treanor 2012].

O VI é constituído por um envelope derivado da membrana celular do hospedeiro apresentando um diâmetro entre 80-120nm. Por debaixo da membrana

existe uma camada proteica formada pela proteína matriz (M1). Estas duas camadas criam uma barreira que protege os oito segmentos de RNA de cadeia simples, num total de 10-14k par de bases, que codificam para 10 proteínas: três subunidades (PA, PB1 e PB2) da RNA polimerase dependente de RNA específica de vírus; duas glicoproteínas de membrana, a hemaglutinina (HA) e a neuramidase (NA); nucleoproteína (NNP), proteína matriz (M1) que forma a estrutura do virião, o canal de protões M2 e duas proteínas não estruturais NS1 e NS2, esta última também conhecida como proteína exportadora do núcleo [Luo 2012].

A nomenclatura dos vírus influenza inclui o tipo de vírus, o local do primeiro isolamento do vírus, número da estirpe e o ano do seu isolamento. Por exemplo, o VI A/aichi/2/68 é do tipo A, da segunda estirpe isolada no Japão, o local Aichi e no ano 1968. Dentro do tipo A existe ainda outra forma de nomenclatura onde os vírus são divididos pelos subtipos das suas duas proteínas membranares, a hemaglutinina (H ou HA) e a neuramidase (N ou NA). Por exemplo, o H3N2 é um VI do tipo A com hemaglutinina do subtipo H3 e neuramidase do subtipo N2 [Treanor 2012].

A entrada do VI nas células do hospedeiro pode ser resumida em quatro passos: i) A partícula viral por intermédio da HA liga-se ao seu recetor que se encontra à superfície das células do hospedeiro; ii) A associação do vírus à membrana induz a endocitose do vírus por esta, gerando um endossoma com o vírus no interior; iii) O endossoma é translocado para o centro da célula, onde a diminuição do pH provoca alterações conformacionais na HA e iv) Fusão entre a membrana do endossoma e o envelope viral por intermédio de várias hemaglutininas. Deste modo as proteínas e os segmentos de RNA são libertados para o interior da célula do hospedeiro. Ao contrário da maioria de vírus de cadeia de RNA simples, a transcrição do RNA viral para mRNA só ocorre no núcleo, dando então início à replicação viral [Luo 2012].

A HA possui um papel direto na entrada do VI na célula, e como se encontra na superfície do envelope viral é um dos alvos do sistema imunitário do hospedeiro. Contudo, visto que a transcriptase de RNA possui uma taxa de erro alta, durante a replicação formam-se inúmeras variantes da HA, podendo no processo ser sintetizada uma hemaglutinina para a qual não se têm imunidade devido a variações antigénicas.

Este processo, que também ocorre na NA, designa-se por derivação antigénica (*antigenic drift*) e é a causa para que apareçam, todos os anos, novas estirpes circulantes de gripe sazonal [Luo 2012, Treonor 2012]. Estas alterações dão-se maioritariamente à superfície da molécula onde os anticorpos interagem, exceto no local de ligação ao recetor, o péptido de fusão e os aminoácidos da bolsa onde este está inserido, que são locais com resíduos conservados [Wilson 1981].

O fenómeno de derivação antigénica pode ser repetido sucessivamente provocando várias epidemias por parte do vírus influenza mas, no entanto, ao final de 10-30 anos a imunidade na população a certo subtipo é alta. A existência de um genoma segmentado e coinfeção torna possível a troca de segmentos entre os diferentes vírus, num processo designado por recombinação/rearranjo. A variação neste caso é muito superior, designando-se por mudança antigénica (*antigenic shift*). Este processo ocorre com uma maior facilidade em organismos que possuam tanto os recetores dos VI das aves como dos VI dos humanos como, por exemplo, os porcos. É um facto que a hemaglutinina é um passo crucial para a entrada do vírus na célula devido à sua ligação ao recetor, fator este que controla a transmissão entre espécies, mas outros genes são importantes, por exemplo, na replicação. Portanto a junção de uma nova HA provinda do VI das aves, ao qual não existe imunidade, com outros genes do VI dos humanos potencia uma maior patogenicidade, sobretudo se o vírus for capaz de propagar contágios entre humanos [Luo 2012, Treonor 2012].

Um desses casos foi o recente surto de H1N1 em 2009, também conhecida por gripe A, que possui sequências genéticas do VI dos humanos, das aves e dos porcos. Teve início nos finais de março/inícios de abril no México e EUA, propagando-se globalmente devido ao fluxo de pessoas entre esses e outros países. Estima-se que o número de mortes tenha sido aproximadamente 280 mil em todo o mundo. Mesmo parecendo um número elevado, este não é superior ao número de casos de morte que ocorrem sazonalmente pelo VI [Smith 2009, Dawood 2012]. Em contrapartida o mesmo vírus (H1N1) provocou a maior pandemia da história moderna em 1918-1920, também conhecida como gripe espanhola, matando milhões de pessoas, número superior às vítimas da primeira guerra mundial. Não se sabe ao certo o número exato de mortes,

mas uma contagem realizada em 2002 sugeriu mortes na ordem dos 50 milhões, podendo mesmo este valor estar subestimado até 200% **[Jonhson 2002]**.

A gripe asiática de 1957-58 foi provocada por um VI do subtipo H2N2 que foi criado pela recombinação entre dois vírus, contendo os genes HA, a NA e a polimerase básica I (PBI) do VI das aves H2N2 e os restantes genes do VI dos humanos H1N1. Esta epidemia provocou muito menos mortes que a sua antecessora: 1-4 milhões de mortes. Este subtipo circulou na população humana durante cerca de 10 anos até sofrer uma recombinação para o subtipo H3, recebendo os genes da HA e PBI deste, e os restantes seis genes da estirpe anterior H2N2. Este subtipo (H3N2) para o qual os humanos ainda não tinham imunidade provocou outra epidemia em 1968-1969, conhecida como gripe de Hong Kong, causando também cerca de um milhão de mortes. **[Lindstorm 2004]**.

A substituição no resíduo de aminoácido 225, uma Gly, do vírus H3N2 de 1968 para Asp em 2004 e depois para Asn em 2005, provocou uma diminuição progressiva na patogenicidade deste. Mesmo continuando a ocorrer infeções, como por exemplo na América do Norte em 2010/2011 ou na Europa em 2011/2012, estas não possuem a mesma severidade que as epidemias dos anos 1968,1989,1994 e 1999. Provavelmente este facto deve-se às interações que este resíduo possui com o recetor **[Lin 2012]**.

Estes vírus relatados anteriormente possuem hemaglutininas do subtipo H1, H2 e H3 que são os únicos subtipos de HA que existem no VI dos humanos. Até 1977, pensava-se que os humanos apenas podiam ser infetados por esses subtipos, sendo que os restantes infetavam as aves. No entanto nesse ano uma criança morreu em Hong Kong por um VI do subtipo H5 (H5N1) mostrando que podem ocorrer “saltos” de contágio entre aves e humanos **[Mandell 2010]**. Existiram também casos de infeções do VI com uma HA do subtipo H7 (H7N7), sendo que em 2003, na Holanda, ocorreu um caso de morte dentro dos vários indivíduos infetados **[Fouchier 2004]**. Em 2013 ocorreu um novo surto, neste caso do subtipo H7N9, com início na China, onde até finais de junho verificaram-se 40 casos de morte **[To 2013]**. Embora o VI do subtipo H9, nomeadamente H9N2, não ter provocado ainda nenhuma morte, já ocorreram até à data, pelo menos, infeções em três crianças em Hong Kong. Mesmo sendo um VI de aves com uma patogenicidade reduzida, possui características como uma grande circulação nos

aviários e pode mostrar, daqui a alguns anos, uma alta patogenicidade como os VI referidos anteriormente [Song 2011].

Como se pode observar o VI continua ativo e com uma grande capacidade de recombinação, originando regularmente novas estirpes muito patogénicas, capazes de causar a morte aos indivíduos infetados. Posto isto torna-se clara a necessidade de se perceber o mecanismo de fusão, que é geral em todos os vírus influenza do tipo A, para que se possa atuar a esse nível.

1.2. Hemaglutinina

A hemaglutinina do vírus influenza foi a primeira proteína ligante a glícidos (PLG) de micro-organismos a ser identificada, sendo-o nos anos 50, por Alfred Gottschalk. O nome hemaglutinina provém da sua habilidade em agregar eritrócitos, processo designado por hemaglutinação, devido à ligação da hemaglutinina aos ácidos siálicos existentes na superfície da membrana dos eritrócitos [Varki 2009].

A HA é uma glicoproteína ancorada ao envelope membranar do vírus influenza constituída por três subunidades iguais, as quais se dividem em duas subunidades diferentes (HA1 e HA2) ligadas por uma ponte persulfureto (figura 1), ou seja, é um homotrímico de heterodímeros. Cada monómero, de aproximadamente 550 resíduos, pode ser dividido em três regiões: cerca de 510 constituem o ectodomínio, 27 o domínio transmembranar e 10-11 resíduos (na zona do C-terminal) o domínio intraviral. A sua estrutura foi determinada em 1981 utilizando a protease bromelaína, que hidrolisa os últimos 46 resíduos da HA, que compreendem o domínio transmembranar, gerando-se assim uma proteína solúvel. Em contrapartida a estrutura desses 46 resíduos da HA é ainda desconhecida [Wilson 1981, Dopheide 1981].

Esta proteína é facilmente reconhecida em microscopia eletrónica devido à sua forma em pico na superfície do envelope viral com dimensões de 4 nm de largura por 14 nm de comprimento [Harris 2006].

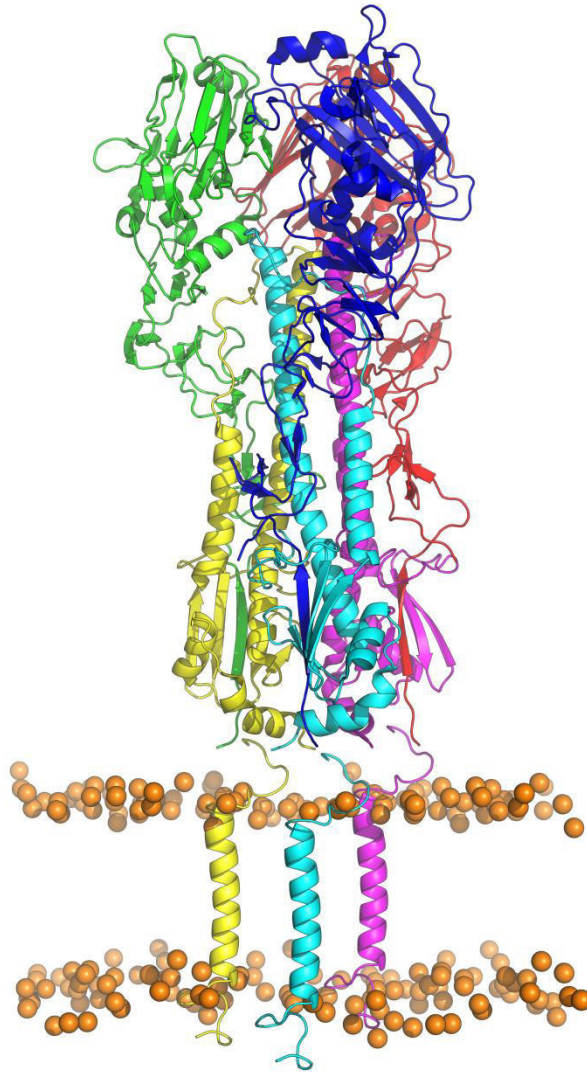


Figura 1 – Estrutura cristalográfica da hemaglutinina (PDB ID - 3VUN), representado em *cartoon*, com uma cor diferente para cada subunidade: Verde, vermelho e azul para as HA2 e amarelo, ciano e magenta para a HA1. Estrutura da parte transmembranar é apenas uma hipótese figurativa.

A HA2 forma o centro e a base da proteína, formando uma *coiled-coil* entre os três monómeros, o que lhe proporciona a sua estrutura em bastão, característica esta das proteínas de fusão do tipo I. No N-terminal da HA2 encontra-se o péptido de fusão hidrofóbico resguardado no interior de uma cavidade formada entre as diferentes subunidades e monómeros. De seguida e após os resíduos que formam uma folha beta com a HA1 temos uma estrutura hélice-*loop*-hélice que forma a já descrita *coiled-coil*. O loop entre as hélices é designado por péptido dobradiça, devido às alterações conformacionais que sofre a pH baixo [Chen 1999].

A HA1 é constituída em grande parte pela parte superior da HA, região globular marcada maioritariamente pelo barril de folhas beta e o local de ligação ao recetor. A HA1 é estabilizada pelas interações entre monómeros, principalmente por pontes salinas [Huang 2002, Xu 2011]. Os primeiros 18 resíduos da HA1 (zona inferior da HA1) formam uma folha beta tanto com resíduos da região N-terminal como também da região C-terminal da HA2, sendo um fator de estabilidade das duas subunidades, apenas ligadas por uma ponte persulfureto.

Por fim no C-terminal encontra-se o domínio transmembranar e o domínio intraviral, também referido como cauda citoplasmática, que não possuem estrutura tridimensional conhecida.

Para além da ponte persulfureto que liga as duas subunidades HA1 e HA2, existem mais 5 pontes conservadas, sugerindo um possível papel na função/estrutura, bem como no *folding* da hemaglutinina.

A hemaglutinina é glicosilada em 3-10 locais, dependendo da estirpe do vírus, aquando do seu transporte do retículo endoplasmático para a membrana celular [Gamblin 2004]. As glicosilações possuem um papel importante, tanto no *folding* através do recrutamento de *chaperones*, como também um efeito na afinidade da HA para com os diferentes recetores. Por exemplo, remover a glicosilação no resíduo Asn158 altera drasticamente a afinidade para com um determinado recetor [Wang 2007].

Existem 16 subtipos diferentes de hemaglutinina, os quais são classificados consoante o antissoro que inibe um determinado subtipo. Mesmo sendo a estrutura quaternária da hemaglutinina idêntica entre os diferentes subtipos, estes apresentam uma similaridade de 30-70% [Nobusawa 1991, Luo 2012]. As diferenças ocorrem maioritariamente nos *loops* da superfície da região globular da HA1 onde ocorre a maior variação antigénica sem afetar a entrada do vírus. Ocorrem também no local de ligação, variando a afinidade para os diferentes recetores do hospedeiro. No decorrer desta dissertação iremos apenas focarmos no subtipo H3 utilizado neste trabalho, a não ser que seja referido explicitamente o subtipo, como ocorrerá nos estudos feitos sobre o péptido de fusão.

1.3. Ciclo de vida do vírus

A hemaglutinina precursora é sintetizada ao nível do retículo endoplasmático, sendo encaminhada para a membrana do hospedeiro através do complexo de Golgi e vesículas secretórias. O vírus adquire o envelope lipídico através da gemulação, ou seja através da formação de protuberâncias, da membrana apical de células infetadas. É na superfície da membrana que as proteínas são reunidas num processo designado por “reunião” (*assembly*), onde a NA, a HA e a M2 são concentradas na membrana e incorporadas no envelope viral.

As hemaglutininas, nas células do hospedeiro, reúnem-se em jangadas lipídicas devido à zona transmembranar, principalmente devido à palmitoilação da cauda citoplasmática [Scheiffele 1997]. Mutações pontuais ou proteínas quimeras podem originar diferentes partições para esses microdomínios insolúveis em detergentes [Scheiffele 1997, Melikyan 1999]. Este fator aumenta a probabilidade das hemaglutininas se reunirem na mesma região da membrana da célula do hospedeiro, aquando do processo de gemulação.

1.3.1. Clivagem da hemaglutinina precursora (HA0)

A hemaglutinina é traduzida como uma única molécula precursora (HA0), com 550 resíduos, que se associa de seguida a outras duas moléculas iguais, formando um trímero. Essa estrutura mantém-se até a hemaglutinina ser transferida para a superfície membranar. No exterior do envelope a HA só se torna ativa após ser clivada por endoproteases de serina do tipo tripsina, entre o resíduo 328 e 329, resultando nas subunidades HA1 e HA2. No caso específico das HA do subtipo H1, H2 e H3 que infetam humanos o enzima candidato é a triptase Clara extracelular, produzida nas células Clara do epitélio bronquíolar ou na mucosa gastrointestinal. Este enzima reconhece um local específico (Q/E-X-R) normalmente constituído por um único resíduo básico (usualmente arginina) [Taubengerger 1998]. Contudo, o enzima responsável pode provir de bactérias ou por respostas imunitárias, o que pode explicar o aumento da severidade da doença aquando de coinfeções [Zambon 2001].

Utilizando a mutação R329Q foi possível impedir a clivagem da hemaglutinina mesmo após tratamento com tripsina, estando a comparação entre a estrutura clivada e não clivada representada na figura 2 [Chen 1998].

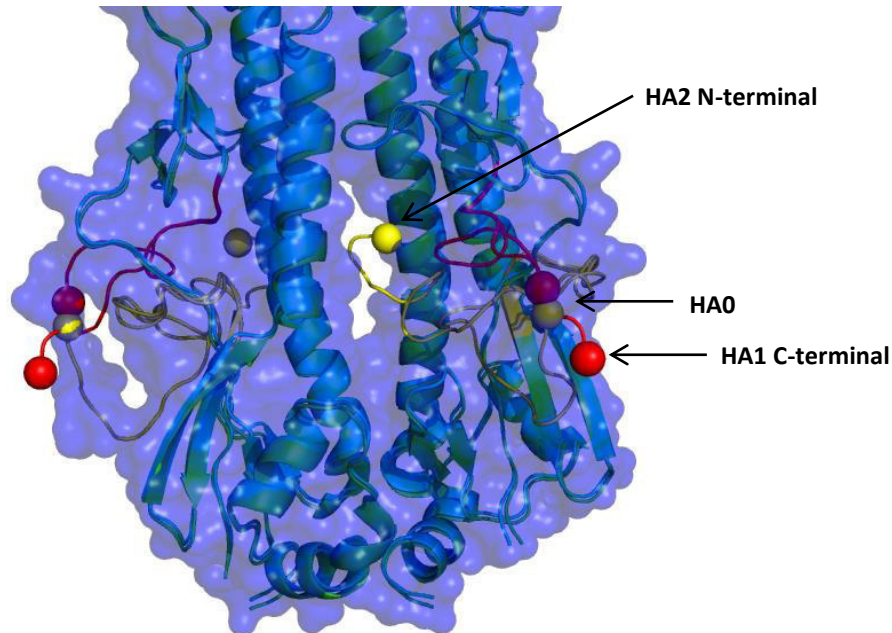


Figura 2 – Comparação entre as estruturas sobrepostas da hemaglutinina precursora (HA0) e da hemaglutinina clivada (HA1+HA2), onde a subunidade que ficaria por trás foi omissa para melhor percepção das cavidades. N-terminal e C-terminal representados como esfera amarela e vermelha, respetivamente. Foi ainda aplicada uma superfície para se perceber onde se situa a bolsa entre os monómeros, para onde migra o N-terminal da HA2 após clivagem.

Esta clivagem possui um efeito muito pouco significativo na estrutura da hemaglutinina, apenas alterando a posição de 10 resíduos de aminoácidos antes e depois do local de clivagem, como apresentado a amarelo e vermelho na figura 2. Após a clivagem o N-terminal apolar da HA2, segmento este designado por péptido de fusão, é deslocado para uma cavidade entre subunidades, diminuindo a acessibilidade ao solvente por parte dos resíduos constituintes da cavidade. Esta bolsa é maioritariamente hidrofóbica, mas existem resíduos ionizáveis, nomeadamente Asp109, Asp112, Lys117 da HA2 e a His17 da HA1 que formam pontes de hidrogénio e interações eletrostáticas com os resíduos carregados (a pH neutro) do péptido de fusão [Chen 1998]. No caso do C-terminal da HA1, os últimos resíduos não sofrem grandes alterações, sendo apenas afastados da cavidade apontando para o solvente.

Após a clivagem, a estrutura originada a pH neutro torna-se fortemente resistente à ação de outras proteínas digestivas mas não possui capacidades fusogênicas [Thoennes 2008]. A HA0 a pH baixo também não é fusogênica. A clivagem é necessária para a criação de uma estrutura metastável que serve de gatilho para as alterações conformacionais que ocorrem com a diminuição do pH [Bullough 1994].

O tropismo das infecções do VI, ou seja, a propensão que o VI tem para infectar um determinado tipo de tecido, deve-se não só às diferenças de afinidade para os diferentes recetores, mas também da protease capaz de clivar a HA. No caso humano, a existência da triptase Clara apenas no epitélio bronquiolar promove a infecção só nessa zona. Em contraste existem outros subtipos de vírus mais virulentos, nomeadamente do subtipo H5 e H7, que possuem vários resíduos básicos no local de clivagem. Para além de aumentar o tamanho do *loop* (mais acessível), os resíduos básicos permitem a clivagem por proteases do tipo furina (reconhecem sequências R-X-R/K-R), que existem numa ampla distribuição de tecidos. Tornando assim estas estirpes mais virulentas e por sua vez mais patogênicas, uma vez que são capazes de infectar um maior número de tecidos [Zambon 2001, Skehel 2000].

Note-se que não são apenas os resíduos no local de clivagem que possuem um papel no reconhecimento desse local pelo enzima e por consequência na patogenicidade do vírus. Tanto os resíduos à volta do local como outros fatores, não relacionados com o processo de clivagem, podem influenciar [Zambon 2001].

1.3.2. Ligação ao recetor

A entrada do VI nas células por via endocítica necessita, primeiro, que o vírus se ligue à membrana, como ocorre noutros vírus de animais. No caso do VI a HA é a ponte de ligação entre o envelope viral e a membrana do hospedeiro através do reconhecimento do seu recetor, o ácido siálico ou ácido N-acetilneuramínico (figura 3) [Luo 2012].

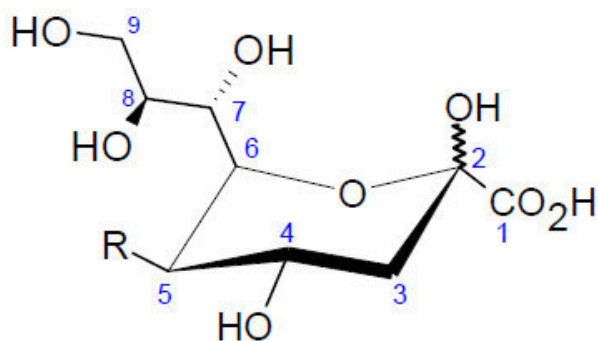


Figura 3 – Estrutura geral de um ácido siálico. Imagem criada no programa Chemscketch.

O nome de ácido siálico provém da descoberta desta família de moléculas nas mucinas/muco salivar. É uma família de moléculas ampla onde todas as moléculas derivam do ácido neuramínico, apenas variando no grupo ligado ao carbono 5 (C5), onde o Neu5Ac (Ac, grupo amida) é o mais comum nos mamíferos. O ácido siálico encontra-se normalmente no final de ramificações de N-glicanos, O-glicanos (glicoproteínas) e glicosfingolípidos (gangliósido) situados no lado exterior da membrana celular, no interior de lisossomas ou em glico-proteínas secretadas. Os ácidos siálicos possuem diferentes papéis: como estabilizador, para ligar e transportar iões, para proteger células de proteases ou glicosidasas, para aumentar a viscosidade dos mucos, para regular afinidade de recetores e para modelar processos como sinalização, crescimento e diferenciação [Varki 2009].

O local de ligação ao ácido siálico encontra-se numa depressão da zona globular de cada subunidade HA1 (figura 4), revestido por resíduos conservados nos locais designados por *loop* 130, hélice 190 e *loop* 220. Uma vez que o local de ligação se encontra numa depressão, torna-o inacessível aos anticorpos, podendo manter os resíduos conservados necessários á ligação ao ácido siálico.

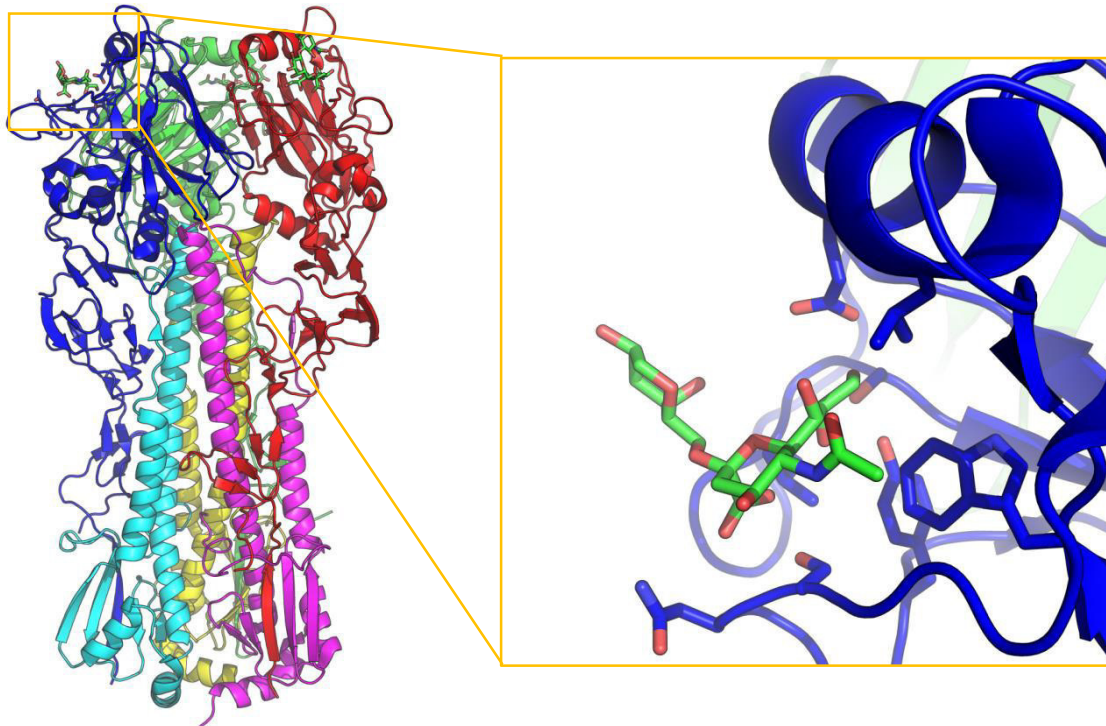


Figura 4 – Estrutura da hemaglutinina realçando o local de ligação ao ácido siálico. No pormenor a hemaglutinina está representada como *cartoon* azul ao passo que o resíduo de ácido siálico e os resíduos que interagem com ele estão representados de tubos a verde e azul, respetivamente.

Embora as ligações sejam fracas, a existência de uma ligação multivalente de três locais de ligação (um por cada HA1), resulta numa ligação forte à membrana [Takemoto 1996]. No caso da HA do subtipo H3 o anel do ácido siálico assenta por cima das cadeias aromáticas dos resíduos Trp153 e Tyr98, com os grupos funcionais ligados ao anel de piranose a interagir com resíduos da HA, exceto o grupo hidróxido do C4 (figura 3).

O grupo carboxilo do C1 forma duas pontes de hidrogénio com a Ser136 e o azoto da ligação peptídica com a Asn137. No grupo amida ligado ao C5 o azoto forma uma ponte de hidrogénio com o grupo carbonil da ligação peptídica da Gly135 e o grupo metilo forma interações de van der Waals com o Trp153. O grupo hidróxido do C7 faz uma ponte de hidrogénio com o oxigénio da amida (C5), e ambos formam contactos de van der Waals com a Leu194. Os grupos hidróxido C8 e C9 formam pontes de hidrogénio com Tyr98 e His183/Glu190, respetivamente. Note-se que a estrutura e orientação do ácido siálico nas estruturas cristalográficas, é independente do análogo a que está ligado (tipo de ligação e molécula ligada) [Skehel 2000].

Os dois tipos mais comuns de ligação entre o ácido siálico e a penúltima galactose são Neu5Ac α (2,6)-Gal e Neu5Ac α (2,3)-Gal. A especificidade da HA para os diferentes receptores varia dependendo dos resíduos do local de ligação, e por consequência varia entre os diversos subtipos de HA. Atualmente os VI dos humanos podem conter HA do subtipo H1, H2 ou H3 e estas têm uma maior afinidade para o recetor Neu5Ac α (2,6)-Gal existentes nas células epiteliais do trato respiratório superior. As mucinas do muco que revestem esse trato possuem muitos receptores Neu5Ac α (2,3)-Gal, diminuindo a possível infecção deste tipo de vírus. Os vírus influenza das aves podem conter HA de todos os subtipos e possuem uma maior afinidade para o recetor Neu5Ac α (2,3)-Gal, existentes no intestino. No caso específico do porco, ambos os tipos de receptores existem nas células do trato respiratório, o que lhe dá a característica de poder servir de veículo para a passagem do vírus das aves para os humanos, através da mudança antigénica **[Skehel 2000, Thompson 2006, Luo 2012]**.

Curiosamente, a alteração de um único resíduo é suficiente para alterar a preferência pelo recetor, nomeadamente a mutação L226Q **[Rogers 1983]**. As hemaglutininas com preferência para o recetor com ligação α (2,6) possuem uma leucina na posição 226 e uma serina na posição 228, ao passo que as hemaglutininas com preferência para o recetor com outra ligação, α (2,3), possuem uma glutamina e uma glicina nas posições equivalentes. Contudo outros resíduos têm influência na especificidade da HA, por exemplo, os resíduos nas posições 186 são prolinas nas aves e serinas nos humanos e os resíduos na posição 225 são glicinas nas aves e aspartatos nos humanos. Outra diferença, com um potencial papel na especificidade do recetor, é dos grupos ligados ao ácido siálico “saírem” em direções opostas do local de ligação, embora, a posição do ácido siálico seja igual para ambos os receptores **[Skehel 2000]**.

Foi demonstrado que VI mutantes que não expressam neuramidase (responsável pela hidrólise da ligação HA-recetor) podem crescer num meio de células epiteliais modelo se forem continuamente tratadas com neuramidase exógena retirando, portanto os ácidos siálicos à superfície da célula. Surpreendentemente, isto apenas diminui parcialmente a ligação do vírus às células, sugerindo uma via de entrada alternativa, onde o ácido siálico tem um papel secundário, por exemplo, aumentando a afinidade da HA para outro recetor **[Stray 2000]**.

1.3.3. Entrada na célula

Após a ligação da partícula viral à membrana do hospedeiro por intermédio de várias hemaglutininas ligadas ao seu recetor, a entrada do vírus na célula segue maioritariamente uma via endocítica dependente de clatrina. No entanto, pode também seguir uma via independente de clatrina [Sieczkarski 2002], que pode estar relacionado com a entrada do vírus sem ligação ao ácido siálico, discutido atrás.

O transporte da vesícula para o núcleo pode ser dividido em três passos, ocorrendo no primeiro passo o transporte dependente de actina na periferia da célula do endossoma precoce. No segundo passo ocorre um transporte rápido do endossoma precoce para a região perinuclear através dos microtúbulos de forma dependente de dineína. Nesta região ocorre um primeiro abaixamento de pH para aproximadamente 6.0, sinal do início do amadurecimento do endossoma. Por fim, no terceiro passo, o movimento unidirecional e rápido do passo anterior é alterado para um movimento bidirecional e transiente, sinal da formação de um endossoma maduro, onde ocorre o último abaixamento de pH para aproximadamente 5.0 [Lakadamyali 2003]. É nesta fase final, na região perinuclear, que ocorrem as alterações conformacionais da hemaglutinina devido ao abaixamento de pH, que serão descritas de seguida, levando no final à fusão entre o envelope viral e a membrana do endossoma tardio.

1.4. Alterações conformacionais – mecanismo *spring-loaded*

Ainda se está longe do completo conhecimento do complexo processo de fusão entre as membranas do vírus e da célula do hospedeiro. A estrutura da hemaglutinina anteriormente apresentada foi obtida a pH neutro (não fusogénico) e corresponde a um estado metastável, ou seja, é um estado estável mas não é o estado de menor energia da hemaglutinina [Carr 1993, Madhusoodan 2003]. A pH fusogénico a hemaglutinina sofre alterações conformacionais, atingindo uma estrutura de menor energia, onde a diferença de energia é utilizada para intermediar a fusão entre as membranas [Bullough 1994]. Contudo as alterações conformacionais que da hemaglutinina entre o estado nativo e o estado fusogénico não estão completamente asseguradas, nomeadamente a

ordem e reversibilidade dos estados intermédios [Fontana 2012]. A labilidade dos diferentes estados e a rapidez das alterações conformacionais não permitem estudos estruturais dos intermediários, havendo apenas estudos indiretos de proteólise, ligação a anticorpos e microscopia eletrónica a temperaturas criogénicas (cryo-EM) [Korte 1999, Bottcher 1999, White 1987]. Ainda, a mera exposição de hemaglutininas a pH baixo, sem uma membrana alvo, resulta na formação de agregados. É provável que isto se deva à exposição no estado final e/ou estados intermediários do péptido de fusão altamente hidrofóbico [Skehel 1982, Doms 1985].

O mecanismo atualmente aceite para estas alterações é o designado por *spring-loaded*, existindo uma diversidade de propostas, onde as maiores diferenças entre estas são a ordem das alterações conformacionais, que serão debatidos de seguida.

Os resíduos que regulam a estabilidade acídica encontram-se em quatro regiões: a interface entre as subunidades HA1 e HA2; a zona *coiled-coil* da HA2; o péptido de fusão (~25 resíduos no N-terminal da HA2) e a cavidade da qual o péptido de fusão se encontra inserido a pH neutro [Reed 2009]. Ambas são importantes na regulação das alterações conformacionais que ocorrem na HA a pH baixo [Weis 1990].

A subunidade HA1 tem uma carga global levemente positiva a pH neutro, sobretudo nas zonas de contacto com a subunidade HA2. Algo semelhante ocorre para a subunidade HA2, mas neste caso com uma carga negativa. Assim sendo, as subunidades têm uma leve repulsão entre subunidades iguais, mas uma forte atração entre subunidades diferentes. No entanto, com a diminuição do pH a subunidade HA1 que já possuía uma carga global positiva, fica ainda mais positiva, portanto aumentando a repulsão entre as restantes subunidades HA1 do trímero [Huang 2002]. Não obstante a HA1 mantém consideravelmente a estrutura terciária do estado nativo [Madhusoodanan 2003, Korte 1999]. É essencial que a HA sofra este efeito de destabilização para que ocorra a fusão, uma vez que em estudos onde se estudaram mutantes com pontes persulfureto entre as diferentes subunidades HA1 do trímero, impossibilitando a sua separação, o processo de fusão não ocorreu [Godley 1992, Kemble 1992]. Existem ainda outros estudos de microscopia eletrónica [Böttcher 1999], ligação a anticorpos [White 1987], proteólise [Doms 1986, Madhusoodanan 2003] ou

estudos onde a HA se encontrava ligada ao ácido siálico [Leikina2000], que apoiam a ideia de que a destabilização é necessária para que outras alterações conformacionais ocorram na HA, até ao estado fusogénico.

O pH não é o único fator externo que serve de gatilho para as alterações conformacionais. A temperatura ou agentes desnaturantes também podem surtir o mesmo efeito, sugerindo que as alterações resultam de uma estrutura metastável e de certa desestabilização [Carr 1997]. Existe uma relação praticamente linear quer tanto com o aumento de temperatura a pH fixo como com o decréscimo de pH a temperatura fixa [Ruigrok 1986] originando em ambos casos alterações conformacionais que são indistinguíveis entre elas [Huang 2002].

Na estrutura a pH neutro a subunidade HA2 está parcialmente escudada do solvente pela HA1, no entanto a diminuição do pH diminui não só essa proteção ao solvente como também a forte interação entre a HA2 e a HA1. Como exemplo, existem dois locais de fortes interações iónicas a pH neutro, sendo o primeiro um quarteto de carga envolvendo a Glu69 do péptido dobradiça e o segundo uma tríade iónica envolvendo a Lys310 da parte inferior da HA1 e os resíduos Asp86 e Asp90 da hélice central da HA2 [Xu 2011]. Ambos os fatores permitem uma certa destabilização na HA2 entre si, formando uma cavidade central ao longo de toda a estrutura do trímero [Böttcher 1999].

Para ultrapassar o problema de agregação da HA a pH baixo, Bullough et al. em 1994 utilizaram um fragmento solúvel do trímero da HA, designado por TBHA2. Esse fragmento foi obtido pelo tratamento da BHA a pH 5.0 com digestões sucessivas de tripsina e termolisina, contendo no final os resíduos 38-175 da HA2 ligados por uma ponte persulfureto aos resíduos 1-27 da HA1 [Bullough 1994]. Uma estrutura cristalográfica deste fragmento TBHA2 foi determinada a pH baixo, estando a sua estrutura representada na figura 5.

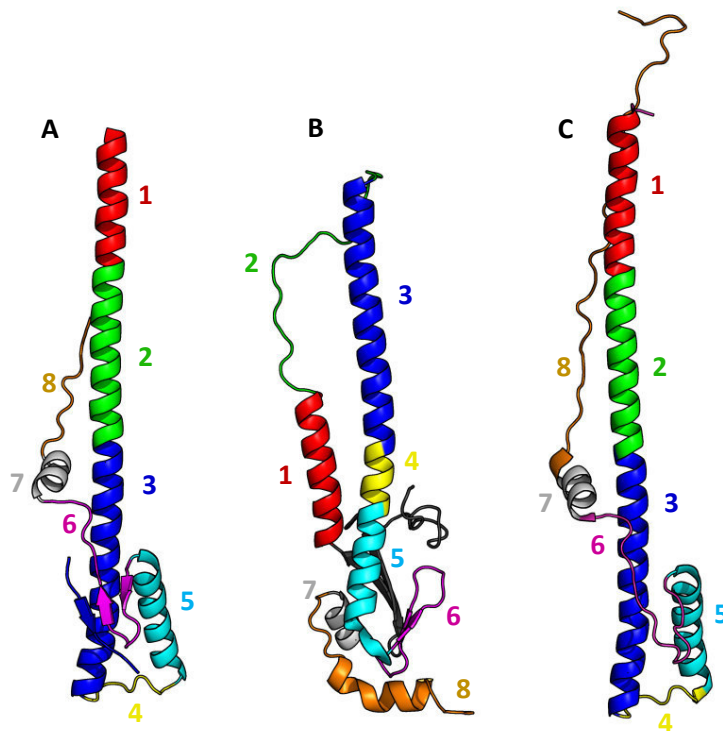


Figura 5 – Estruturas da subunidade HA2 em diferentes condições. A) Estrutura da TBHA2 a pH 5.0, B) Estrutura de apenas a HA2 a pH neutro e C) Estrutura da EHA2(23-185) a pH 5.0. Cada uma das estruturas está representada por *cartoons* e a seguinte sequência de cores exemplificada para a estrutura B (HA2 a pH neutro): 1 – a vermelho a primeira hélice da HA2; 2 – a verde o péptido dobradiça; 3 – a azul escuro a primeira metade da hélice central; 4 – a amarelo a hélice que a pH baixo se torna um *loop*; 5 – a ciano a segunda metade da hélice central; 6 – a magenta as folhas beta; 7 – a branco a hélice do C-terminal que mantém a estrutura e 8 – a laranja a hélice do C-terminal que perde estrutura. Na estrutura B, está representado a preto os resíduos do N-terminal da HA2 (incluindo o péptido de fusão) que foi hidrolisado nas restantes estruturas.

Segundo esta estrutura a pH baixo ocorrem três grandes rearranjos conformacionais na HA2 (figura 5A) em comparação com a estrutura a pH neutro (figura 5B), onde cada parte dessa estrutura que sofreu rearranjos está numerada por ordem do N-terminal ao C-terminal da HA2 (figura 5). O primeiro rearranjo é a formação de uma hélice central composta pela hélice 1 (resíduos 40-55), pelo *loop* 2, também designado por péptido dobradiça (resíduos 56-74) e metade da segunda hélice (hélice 3, resíduos 75-105). Basicamente o abaixamento de pH provocou a passagem de uma estrutura desordenada do *loop* 2 para uma estrutura secundária em hélice, o que já tinha sido sugerido através de estudos com péptidos sintéticos [Carr 1993]. Esta mudança parece estar relacionada, pelo menos inicialmente, tanto com a protonação dos resíduos Glu64 e Glu69 que formavam pontes salinas com hélices da HA2, como também com os resíduos Phe63 e Phe70 que estabeleciam contactos com o interior da

molécula e passam a ficar virados para o solvente [Xu 2011]. O segundo passo consiste na perda de estrutura no interior da segunda hélice (hélice 4, resíduos 106-112), resultando numa volta de 180° na estrutura e colocando a outra metade da segunda hélice (hélice 5, resíduos 113-129) antiparalela à primeira metade (hélice 3). A terceira alteração é a perda da estrutura compactada, principalmente a partir do resíduo 154 (hélice 8) onde a anterior estrutura em hélice adota uma estrutura estendida, acomodando-se nas cavidades entre duas hélices centrais de forma antiparalela. Esta estrutura em *hairpin* permite que tanto os péptidos de fusão no N-terminal da HA2, como os péptidos transmembranar no C-terminal da HA2 estejam virados para o mesmo lado da membrana [Bullough 1994].

Após 5 anos, foi obtida uma nova estrutura, desta vez de um fragmento sem quaisquer resíduos da HA1 e com os resíduos 23-185 da subunidade HA2, designado por EHA2(23-185) (figura 5C) [Chen 1999]. À parte dos resíduos que não se encontravam na estrutura anterior a estrutura obtida foi exatamente igual à obtida anteriormente pela TBHA2, onde se pode ver que os novos resíduos continuam a extensão antiparalela ao longo de toda a *coiled-coil* em forma de *hairpin*. Os últimos resíduos do C-terminal estendem-se até à zona de início da hélice ao qual o péptido de fusão se encontrará ligado. Os resíduos 174-178 formam diversas pontes de hidrogénio com os resíduos 34-37 (a hélice 1 começa no resíduo 38) formando uma *N-cap* trimérica, principalmente no Asp37, estabilizando as hélices centrais [Chen 1999].

Curiosamente, a expressão em *Escherichia coli* do fragmento 38-175 da HA2, designado por EBHA2 (que não possui a ligação à HA1), origina uma proteína que parece adotar, a pH neutro, a mesma estrutura que a TBHA2. Isto foi estudado através de estudos de dicroísmo circular, proteólise, ligação de anticorpos e microscopia eletrónica. A adição de um octapéptido altamente carregado no N-terminal torna a HA2 (resíduos 1-185) solúvel e a clivagem desse octapéptido provoca agregação de fragmentos ou ligação a vesículas lipídicas, mimetizando as interações da HA a pH baixo [Chen 1995]. Isto sugere fortemente que a estrutura da HA2 a pH baixo é o estado mais baixo de energia, sugerindo por sua vez que a HA1 funciona como uma pinça que prende cineticamente a HA2 numa estrutura metastável até à diminuição do pH, ou seja, até ao endossoma chegar à região perinuclear da célula do hospedeiro [Carr 1993, Chen 1995].

Segundo estas observações poderíamos considerar que a suposta necessidade de um pH baixo para que ocorra fusão deve-se apenas à destabilização da HA1. Isto é ainda suportado pela observação de que o passo limitante da fusão é independente do pH e da composição lipídica, mas dependente da temperatura [Markosyan 2001]. No entanto, estudos por Kim et al. mostraram que mesmo este fragmento da HA2 (resíduos 1-185 com o péptido hidrofílico no N-terminal) apresenta capacidades fusogénicas de uma forma dependente do pH e do tamanho do próprio fragmento. O fragmento 1-127 da HA2 apenas permitiu a mistura de lípidos entre vesículas, ao passo que um fragmento maior com os resíduos 1-185 consegue provocar a fusão entre membranas e criar pequenos poros [Epand 1999, Kim 2011]. Estes poros não são capazes de alargar, diminuindo a eficiência da fusão, o que já tinha sido sugerido ser o papel do péptido transmembranar nos últimos passos da fusão [Markosyan 2000]. Surpreendentemente 2 mutações, como por exemplo a G1E, inibem fortemente a atividade fusogénica do segmento, bem como na HA nativa [Kim 2011]. Note-se que apenas no segundo segmento (segmento 1-185) é que estão presentes os resíduos 174-178, que formam uma *N-cap* com os resíduos imediatamente antes da hélice central, estabilizando-as.

Uma vez que não ocorreram alterações significativas da estrutura na HA2 com o abaixamento do pH [Kim 2011] isto sugere que a destabilização da HA1 não é o único fator para que a fusão, a nível fisiológico, ocorra a pH baixo, mas também a interação da HA2 entre monómeros e com a membrana.

Para além das alterações conformacionais mencionadas anteriormente que ocorrem na HA2 com o abaixamento de pH, existe um fenómeno paralelo: a saída do péptido de fusão da cavidade onde se encontra inserido, a qual é dependente do pH. Como já foi mencionado, a HA0, mesmo a pH baixo, não é fusogénica. Após a clivagem ocorre a deslocação do péptido de fusão para a cavidade entre monómeros, enterrando certos resíduos ionizáveis conservados da HA, assim formando potenciais gatilhos das alterações conformacionais (figura 6) [Wilson 1981, Gamblin 2004]. Existem, no entanto, outros resíduos específicos do subtipo de hemaglutinina, que regulam a estabilidade a diferentes pH's [Reed 2009, Thoennes 2008].

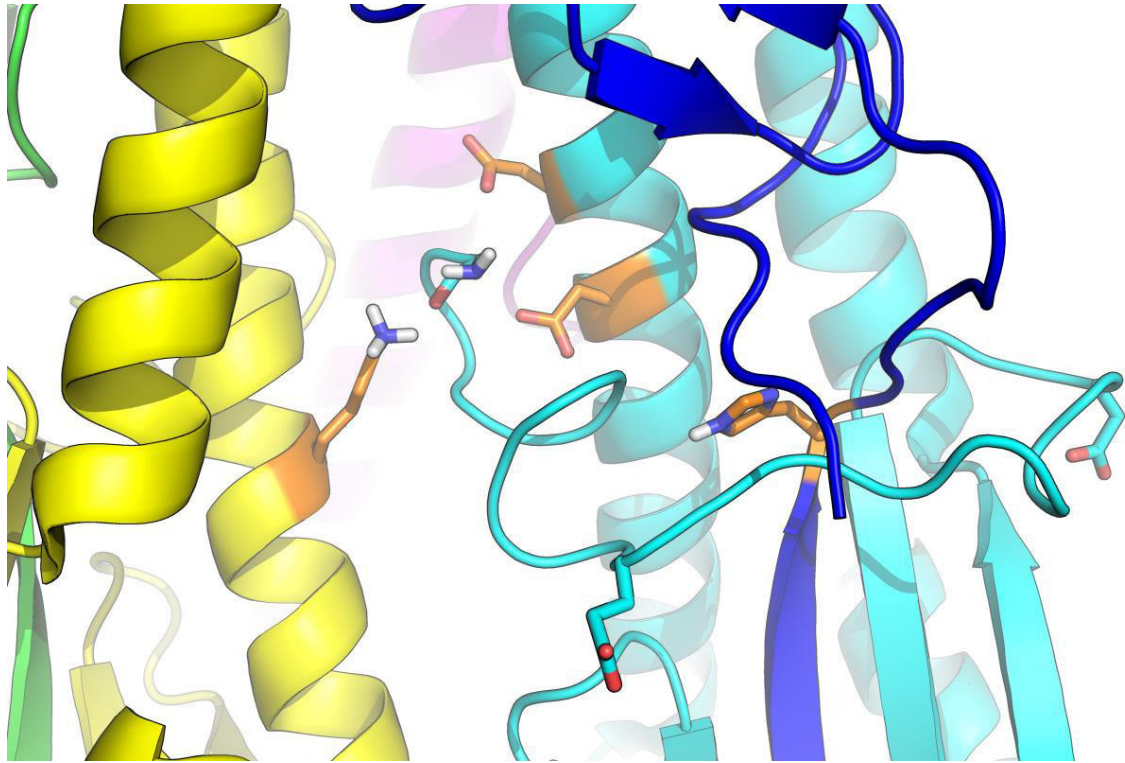


Figura 6 – Pormenor da estrutura envolvente da cavidade onde o péptido de fusão se insere após clivagem, representada em *cartoon*, sendo utilizados *sticks* para salientar os aminoácidos mais importantes (Gly1, Glu11, Asp19, Asp109, Asp112 e Lys114 da HA2 e His17 da HA1).

Os aspartatos conservados na posição 109 e 112, na mesma face da hélice e na mesma subunidade que o péptido de fusão, formam pontes de hidrogénio com os primeiros 2-6 resíduos deste último. Principalmente o Asp112, que está mais próximo do péptido de fusão e forma mais pontes de hidrogénio com este. Quando o Asp112 é mutado para alanina provoca uma diminuição do pH de fusão em 0.5 unidades. Para além disso mutações duplas nesses dois aspartatos impossibilitam o *fold*ing da HA e por consequência o processo de fusão [Thoennes 2008].

Outro exemplo de um resíduo ionizável, que é enterrado pela translocação do péptido de fusão, é a histidina na posição 17 da subunidade HA1. Este resíduo forma duas pontes de hidrogénio com os resíduos 6 e 10 do péptido de fusão, por intermédio de moléculas de água [Daniels 1986]. A mutação H17 da HA1 para alanina, arginina e glutamina aumenta o pH de ativação por 0.4, 0.7 e 0.9 unidades, respetivamente, sendo estes dois últimos das hemaglutininas menos estáveis observadas em vírus. Por outro lado a mutação para tirosina (resíduo nativo do subtipo H1) diminui o pH de fusão por 0.3 unidades, uma vez que além de não ser um resíduo protonável também forma

pontes de hidrogénio sem o intermédio de moléculas de água. Ao passo que a histidina (nativa do subtipo H3) só forma pontes de hidrogénio com o intermédio da água [Daniels 1986, Thoennes 2008].

Isto sugere que o abaixamento de pH pode provocar a protonação desses resíduos, conferindo carga ao resíduo e/ou perda de pontes de hidrogénio entre o péptido de fusão e os resíduos da cavidade.

As histidinas livres em água tem um pK_a de 6, que é aproximadamente o valor ao qual começam a ocorrer as alterações conformacionais da hemaglutinina. A existência de muitas histidinas conservadas nos diferentes subtipos de hemaglutinina, sugere que a protonação das histidinas pode ser o gatilho para as alterações conformacionais. Contudo, dependendo de diversos fatores como a formação de pontes de hidrogénio ou baixa acessibilidade ao solvente, a protonação das histidinas podem ocorrer a valores de pH mais baixos. Isto no entanto não impossibilita o papel das histidinas, podendo apenas apresentar uma protonação contínua que segue as alterações conformacionais.

Existem ainda resíduos ionizáveis, não tão conservados, que mesmo não possuindo interações diretas com o péptido de fusão interagem com outros resíduos intermediários, como é o caso do Q105 e E114 para a H3. As mutações E114K e Q105K (ou Q105R) aumentam respetivamente o pH de fusão em 0.5 e 0.3 unidades, ou seja, ambas as mutações têm um efeito desestabilizador [Thoennes 2008].

Os resíduos 106-111 da HA2 que interagem com o péptido de fusão (por exemplo o resíduo Asp109) são os mesmos onde ocorre a transformação de hélice para *turn* com o abaixamento de pH, podendo portanto haver uma relação entre essa transformação e a saída do péptido de fusão. De facto um estudo de simulação computacional sugeriu que a saída do péptido de fusão ocorre após esta transformação de hélice em *turn* [Madhussodanan 2003].

Segundo as interações exemplificadas nos parágrafos acima, é notório que o péptido de fusão tem um papel importante na estrutura metastável da HA a pH neutro, existindo mesmo estudos que apontam a saída do péptido de fusão como o primeiro passo das alterações conformacionais [White 1987, Korte 1999, Fontana 2012]. Contudo, por exemplo, o estudo realizado por Korte et al. não descarta a possibilidade

de nos primeiros passos haver uma pequena destabilização da HA2, o suficiente para deixar o péptido de fusão subir e o péptido dobradiça assumir estrutura em hélice [Korte 1999].

No entanto todos estes estudos que sugerem a existência e sequência de intermediários são indiretos, não existindo nenhuma estrutura desses intermediários obtida por cristalografia ou NMR. Recentemente Xu et al. conseguiram obter a estrutura cristalográfica de uma hemaglutinina do subtipo H2, com uma mutação da arginina na posição 106 (primeiro resíduo da hélice para *loop*) por uma histidina, a qual passa a apresentar um ponto de equivalência a pH 4,2 (figura 7). Por volta do pH 5 consegue-se, portanto, obter uma estrutura anterior às alterações conformacionais irreversíveis. Assim sendo, essa estrutura pode providenciar informação para os eventos iniciais que a HA sofre com o abaixamento de pH [Xu 2011].

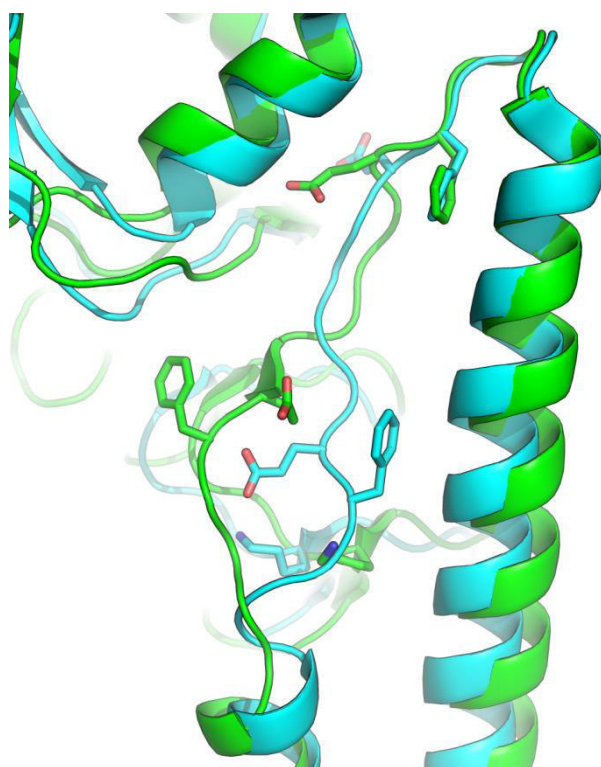


Figura 7 – Estrutura da HA com a mutação R106H. Está representada em *cartoon*, sendo utilizado *sticks* para salientar as cadeias laterais dos resíduos de Phe e Glu. A verde está representada a proteína a pH ácido (PDB ID: 3QQO) e a ciano a proteína a pH neutro (PDB ID: 3QQB).

Resumidamente, a hipótese aceita é que o abaixamento de pH causa a protonação dos resíduos de glutamato do péptido dobradiça, eliminando as pontes salinas formadas por estes. Por consequência o péptido sai do seu estado compactado

junto à hélice central e afasta-se, deixando o Phe64, altamente conservado, exposto ao solvente. Por outro lado a HA1 sofre destabilização terciária na parte inferior, onde se observa uma curvatura do C-terminal da HA1 para compensar a distanciação do péptido dobradiça. É portanto sugerido que os primeiros passos intermédios das alterações conformacionais ocorrem no péptido dobradiça e na destabilização da HA1 [Xu 2011].

Neste caso em específico foi testado ainda a reversibilidade das alterações conformacionais, e de facto estas são completamente revertidos após incubar a hemaglutinina em pH neutro [Xu 2011]. Mesmo no caso da BHA já tinha sido observado que as alterações conformacionais após incubação a pH baixo durante alguns minutos poderiam ser revertidas, sem incapacitar definitivamente a capacidade fusogénica da proteína. Em contraste com a HA nativa qua após longos períodos de incubação a pH ácido, na ausência de membrana alvo, incapacitam a sua capacidade fusogénica [Korte 1999]. Isto sugere que em algum ponto das alterações conformacionais, devido ao abaixamento do pH, ocorre uma alteração (ou acumulação de várias) que torna impossível a reversão para o estado nativo.

No entanto parece que não é apenas a pH ácido que estas alterações ocorrem, uma vez que, segundo um estudo de Markosyan et al., algo semelhante também ocorre a pH neutro [Markosyan 2001]. Neste estudo observou-se que a eficiência da fusão diminui progressivamente com o tempo em que os intermediários da HA são mantidos a pH neutro, provavelmente por perda ao longo do tempo de possíveis locais de fusão [Markosyan 2001].

No final das alterações conformacionais temos, portanto, a HA em *hairpin* com ambos, o C-terminal e o N-terminal da HA2, cada um ligado à membrana viral e à membrana do hospedeiro, respetivamente, virados para o mesmo lado da proteína. É de se esperar que ambos tenham um papel importante na destabilização das células e formação de poros. Ambos os terminais da HA2, ou seja, o péptido transmembranar e o péptido de fusão, serão apresentados mais detalhadamente de seguida.

1.5. Péptido de fusão

Designa-se por péptido de fusão um segmento das proteínas de fusão virais com alta afinidade para membranas, tendo como papel a ancoragem à membrana do hospedeiro. Esses segmentos devem seguir quatro critérios gerais: serem ricos em glicinas, maioritariamente hidrofóbicos, localizados no N-terminal da proteína e mutações na sua sequência levar a perda de atividade (mesmo possuindo resíduos que mutados não levam a uma perda completa de função) [Epanand 2003]. O segmento constituído pelos primeiros 20/23 resíduos do N-terminal da HA2, onde todos, exceto 5, são conservados nos 16 subtipos, segue estes quatro critérios.

Estes domínios sozinhos induzem troca de lípidos entre lipossomas, hemólise de eritrócitos e promovem fusão de vesículas uni-lamelares [Lear 1987] de uma forma dependente do pH e da sequência, à semelhança da respetiva proteína completa.

Mesmo existindo estruturas atómicas dos ectodomínios solúveis das proteínas de fusão, a obtenção da estrutura de apenas o domínio de fusão é um desafio, devido à sua hidrofobicidade e propensão para agregar. Han et al. desenvolveram um sistema de “hospedeiro-hóspede” onde se liga um heptapéptido com resíduos polares (hospedeiro) ao péptido de fusão (hóspede), aumentando a solubilidade e diminuindo a agregação [Han 2000]. Assim conseguiram obter a primeira estrutura de um domínio de fusão [Han 2001]. A estrutura dos primeiros 20 resíduos do N-terminal da HA2 do subtipo H3 em micelas foi obtida por NMR, tanto a pH neutro (7.4) como a pH de fusão (4.0), estando representadas na figura 8.

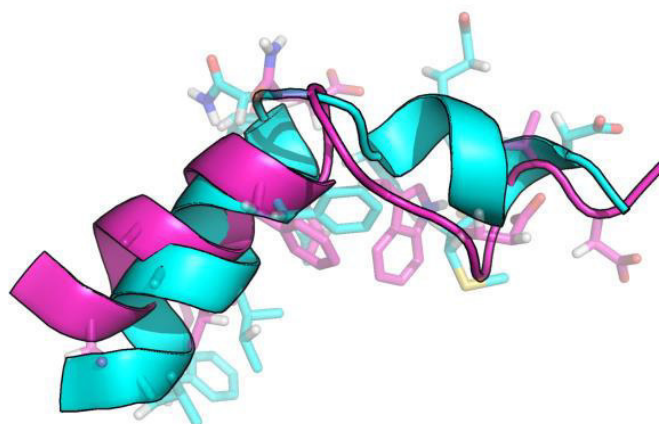


Figura 8 – Péptido de fusão constituído por 20 resíduos de aminoácidos representados em ciano a pH 4.0 (PDB ID. 1IBN) e em magenta a pH7.4 (PDB ID. 1IBO). Cadeia principal está representada como *cartoon* e as cadeias laterais por tubos transparentes.

Este péptido apresenta, segundo análise de dicroísmo circular, uma estrutura completamente desordenada em água, mas em contacto com micelas ou bicamadas lipídicas apresenta estrutura parcial em hélice, sendo que em bicamadas é apenas levemente superior que em micelas [Han 2001]. Ambas as estruturas, segundo análises de EPR, apresentam uma forma de V inverso ao nível da interface membrana-água. Onde ambos os terminais estão direccionados para o interior hidrofóbico da membrana e os resíduos da curva (Glu11 e Asn12) são os mais exteriores. O péptido é anfipático possuindo os resíduos hidrofóbicos virados para a membrana e os resíduos polares e glicinas virados para o solvente. A pH neutro o C-terminal apresenta uma estrutura estendida, mas a pH ácido forma-se uma curta hélice [Han 2001]. Esta alteração conformacional puxa os resíduos Glu15 e Asp19 para a parte superior (figura 8 a ciano), aumentando ainda mais o carácter anfipático do péptido e favorecendo a sua maior inserção na membrana [Chernomordik 2008].

Contudo este péptido não continha os resíduos Trp21, Tyr22 e Gly23, os quais além de serem conservados apresentam características importantes. Normalmente os resíduos aromáticos influenciam a ancoragem de péptidos ao nível da interface [Strandberg 2002, Vostrikov 2010] e a glicina faz parte dos motivos GXXXG e GXXG (onde X é qualquer aminoácido) importantes em interações entre hélices [Russ 2000]. Por esse motivo Lorieau et al. utilizou o mesmo sistema de “hospedeiro-hóspede” para

obter a estrutura por NMR do péptido de fusão, desta vez com 23 resíduos de aminoácidos e do subtipo H1 a pH neutro (figura 9) [Lorieau 2010].

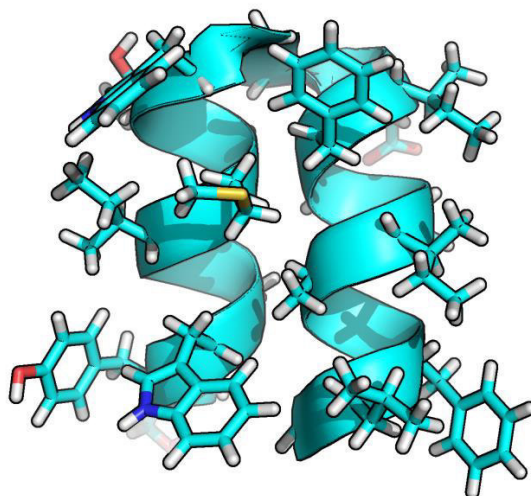


Figura 9 – Péptido de fusão de 23 resíduos de aminoácidos do subtipo H1 (PDB ID. 2KXA). A cadeia principal está representada a *cartoon* e as cadeias laterais representadas por tubos.

Com a adição destes três resíduos a interação entre os terminais é maior, principalmente devido às interações entre a Gly1 e os resíduos Trp21 e Gly23, nomeadamente por uma interação carga-dipolo [Lorieau 2011]. Isto confere uma estrutura de um *hairpin* em hélice, com uma volta nos resíduos Gly12 e Gly13. Esta estrutura compacta é possível graças às glicinas que cobrem a superfície de contacto entre as duas hélices, com motivos GXXXG, que se encontram também noutras proteínas. Esta estrutura é também anfipática com os resíduos hidrofóbicos virados para a membrana e os resíduos polares e as glicinas viradas para o solvente [Lorieau 2010]. Note-se que no péptido com 20 resíduos de aminoácidos também existiam NOEs fracos de longo alcance, os quais sugerem que uma estrutura em *hairpin* também possa ser assumida, embora transitoriamente [Han 2001].

Ao contrário do que acontecia para o péptido de 20 resíduos de aminoácidos, neste não existem diferenças significativas na estrutura tanto a pH neutro como a pH ácido [Han 2001], o qual se pode dever à superior estabilidade da estrutura com 23 resíduos de aminoácidos.

A dimensão do péptido de fusão em *hairpin* não é suficiente para atravessar a membrana, contudo a análise por NMR sugeriu uma alteração conformacional em

pequena escala sentida a pH baixo que não ocorre a pH neutro. Isto pode sugerir por sua vez a existência da passagem do estado em *hairpin* para um outro estado, por exemplo de uma hélice contígua aquando da formação do poro [Lorieau 2012].

1.6. Péptido transmembranar

O péptido de ancoragem ou transmembranar (PT) encontra-se no C-terminal, mais concretamente o segmento que não se encontra na estrutura cristalográfica da hemaglutinina (após tratamento com a protease bromelaína). Esse segmento contém: o domínio intraviral (comumente designado como cauda citoplasmática), ou seja os resíduos que se encontram no lado interior do envelope viral; o domínio transmembranar constituído pelos resíduos que estão inseridos no envelope viral; e o *linker*, resíduos da junção entre os resíduos transmembranares e os resíduos da parte solúvel, presente na estrutura cristalográfica, representados na figura 10.

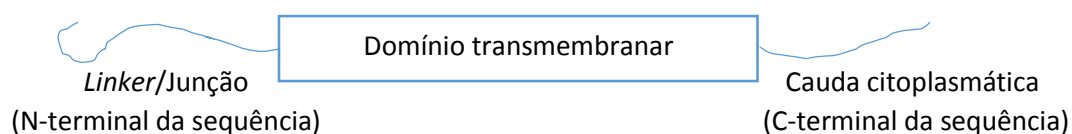


Figura 10 - Esquema dos diferentes segmentos da sequência C-terminal da Hemaglutinina, a cauda citoplasmática, o péptido transmembranar e a junção.

1.6.1. Domínio transmembranar

O domínio transmembranar é maioritariamente hidrofóbico, sendo constituído, por exemplo, pelos resíduos conservados Ile186, Leu187, Ile189, Leu197 e Trp208. No entanto possui também grupos polares conservados, principalmente serinas. Este segmento possui três locais conservados e para o segundo grupo filogenético (contendo as HA do subtipo H3, H4, H7, H10, H14 e H15) consistem em: ILolx no N-terminal onde “I” é isoleucina, “L” é leucina, “x” é maioritariamente Ser ou raramente Ala ou Tyr e “o” é um aromático grande; hSCFL onde “h” é um resíduo alifático ramificado; por último, no C-terminal a Gly204, o Trp208 e a Cys210 [Tatulian 2000].

O domínio transmembranar não é apenas importante para ancorar a hemaglutinina ao envelope viral, mas parece também ser necessário para os últimos

passos do processo de fusão. Tal facto foi suportado pelo estudo de Kemble et al. **[Kemble 1994]** no qual, células expressando ectodomínios da hemaglutinina sem o domínio transmembranar e cauda citoplasmática, mas ligados apenas a uma âncora de glicosil fosfatidil inositol, formam hemifusão entre membranas, sem a ocorrência de poros **[Kemble 1994, Melikyan 1995]**. A hemifusão define-se como um provável intermediário da fusão, onde existe uma continuidade das camadas exteriores de duas membranas, sem que haja a fusão das camadas interiores. Não havendo, portanto, permuta de conteúdo entre células/lisossomas.

Não é apenas a ausência do PT que incapacita a fusão, mas também alterações neste. É o caso observado no estudo de células que expressavam hemaglutininas quiméricas com péptidos transmembranares de diferentes tamanhos, ou seja, diferentes números de resíduos na zona do domínio transmembranar. Estes estudos mostraram que: deleções superiores a 14 aminoácidos não são expressas na superfície celular, deleções de 11 resíduos de aminoácidos mostraram uma grande debilidade a provocar fusão e deleções de 12 resíduos de aminoácidos incapacitaram totalmente a passagem do estado de hemifusão para fusão. Isto sugere um comprimento mínimo de 17 resíduos no domínio transmembranar para ocorrer uma eficiente fusão completa **[Armstrong 2000]**.

A adição de uma arginina, resíduo hidrofílico, ao domínio transmembranar de 15 resíduos conseguiu reabilitar completamente a fusão, possivelmente por favorecer uma localização na interface e assim “obrigar” o PT a assumir uma estrutura transmembranar, mesmo sendo mais curto, com apenas 16 resíduos. Se assim for, existe uma necessidade do péptido transmembranar exibir um comprimento mínimo para poder atravessar toda a membrana e ativar o processo final de fusão, através de um mecanismo ainda não totalmente conhecido **[Armstrong 2000]**.

Mesmo sendo o domínio transmembranar essencial para a fusão, a sua sequência não é restrita, uma vez que HA quimeras com domínios transmembranais de outras proteínas de fusão viral não surtem efeito na criação de poros de fusão **[Schroth-Diez 2000, Melikyan 1999]**. Algo semelhante ocorreu também para outras proteínas de fusão viral, como a proteína G do vírus da estomatite vesicular ou a gp160 do HIV, em

que os domínios transmembranares foram alterados por outros domínios transmembranares de outras proteínas [Wilk 1996, Odell 1997]. Ambos os estudos sugerem que apenas é necessário um domínio transmembranar (com as características comuns a todos) que penetre em ambos os folhetos da membrana, para que ocorra fusão.

No entanto um estudo realizado por Melikyan et al. mostrou que a mutação G520L provoca apenas hemifusão restrita (sem transferência de lípidos), ao contrário da mutação G520S, sendo que a serina é o resíduo da maioria das estirpes. Isto mostra que mutações pontuais num local crítico podem alterar completamente o processo [Melikyan 1999], mesmo existindo outras mutações como W185A, W188A, S194A e G204A que não têm qualquer efeito [Armstrong 2000].

Em contraste, foi realizado anos mais tardes por Markosyan et al. um estudo semelhante com HA acoplada a glicosil fosfatidil inositol onde, em condições ótimas de fusão, estas hemaglutininas foram capazes de formar pequenos poros dependendo do pH, temperatura e constituição lipídica [Markosyan 2000]. Sugerindo que o poro inicial é fundamentalmente constituído por lípidos. Este mesmo resultado foi obtido por Kim et al., onde fragmentos da HA2 sem o péptido transmembranar formaram pequenos poros em membranas, contudo sem a capacidade de os alargar [Kim 2011]. Perante estes testes, o péptido transmembranar além de aumentar a eficiência da formação dos poros, também é necessário para os últimos passos de fusão aquando do alargamento dos poros [Markosyan 2000, Kim 2011]. Este péptido faz provavelmente parte da estrutura do poro, revestindo-o em conjunto com o péptido de fusão, através de interações entre os péptidos [Chang 2008].

Existem dois possíveis mecanismos para o papel do PT na passagem do estado de hemifusão para a formação final de um poro. Num, o domínio transmembranar insere-se na zona formada pela junção das camadas exteriores das duas membranas designado por diafragma, induzindo esses lípidos a restabelecer a orientação transmembranar. Noutro, o domínio transmembranar pode sair do diafragma e gerar uma tensão membranar que rompe o mesmo [Melikyan 1999].

A via de fusão com um estado de hemifusão como intermediário é favorecida não só pela já descrita situação de hemifusões restritas, mas também pelo largo espectro de sequências que o domínio transmembranar pode assumir. Uma via direta de fusão, necessitava da criação de poros aquosos revestidos por proteína, mas essas proteínas têm sequências bastante restritas **[Melikyan 1999]**.

Eficiência, velocidade, alargamento dos poros e por conseguinte a fusão em si são dependentes da combinação do PT e da cauda citoplasmática. Péptidos da mesma proteína são mais eficientes que de proteínas diferentes, mostrando que ambas as sequências devem ser compatíveis pois o efeito do domínio transmembranar na fusão depende da sequência da cauda **[Melikyan 1999]**. Foi sugerido que a cauda citoplasmática ou mutações podem afetar a hipotética estrutura necessária para o processo de fusão, sendo que essa estrutura deverá ser comum entre a maioria dos domínios transmembranares **[Melikyan 1999]**.

Uma vez que a maioria dos estudos acerca do domínio intraviral foram realizados em células que expressavam a hemaglutinina (glicoproteína do envelope viral) na membrana plasmática chamaremos daqui para a frente como cauda citoplasmática.

1.6.2. Cauda citoplasmática

A cauda citoplasmática (CC) consiste em 10-11 resíduos de aminoácidos (assumindo que o resíduo carregado delimita o domínio transmembranar) com alta conservação entre todos os subtipos de HA, sobretudo em cinco dos subtipos **[Nobusawa 1991]**. Dois desses resíduos são cisteínas e se considerarmos a cisteína na parte final do domínio transmembranar (ocupando no subtipo H3 da HA as posições 555,562 e 565) geram três locais onde são adicionados, pós-traducionalmente, grupos palmitoil.

A função da cauda citoplasmática não é ainda totalmente conhecida, contudo vários estudos foram realizados para tentar elucidar o papel desta, das cisteínas conservadas e da palmitoilação.

Estudos de células a expressar HA quiméricas sem cauda citoplasmática indicam que esta não é necessária para a fusão **[Jin 1994, Melikyan 1997, Melikyan 1999]**, nem

para a “reunião” (*assembly*) das hemaglutininas [Jin 1994] e nem para a estrutura quaternária [Simpson 1992]. Outros estudos indicam ainda que a substituição da CC pela CC de outras proteínas não abole a capacidade fusogénica da HA [Schroth-Diez 2000]. Curiosamente existem estudos onde mutações na CC podem abolir por completo a fusão, o que não é necessariamente contraditório com os estudos anteriores, uma vez que, certas mutações podem provocar efeitos distintos [Ohuchi 1998].

No entanto, a existência da cauda deve facultar alguma vantagem ao vírus, já que em culturas de vírus, onde foram introduzidos codões stop imediatamente antes dos codões que codificam para a CC, a HA reverteu ao estado nativo (sem codão stop) [Jin 1994]. Algumas dessas vantagens podem ser, por exemplo, uma certa vantagem no crescimento [Jin 1994] ou aumentando a rapidez e eficiência do transporte da HA entre o RE e o complexo de Golgi [Simpson 1992].

O papel das palmitoilações ainda não é claro, principalmente porque diferentes subtipos de HA, nomeadamente H1, H2, H3 e H7 apresentam diferentes respostas, podendo variar de fusão completa até inatividade. No caso da H3 foi amplamente estudado que a HA não palmitoilada não influencia a cinética da fusão [Melikyan 1997, Chen 2005], contudo parece favorecer o fenómeno de “oscilação”, ou seja, abertura e encurtamento do poro. [Melikyan 1997].

As duas últimas palmitoilações não são necessárias para a infecciosidade do vírus [Jin 1994]. Contudo modificações dessas duas cisteínas inibem severamente a atividade fusogénica [Naevé 1990]. Para o subtipo H1 mutações nas cisteínas também incapacitaram a recuperação de partículas virais [Zurcher 1994].

Pelo menos para a HA do subtipo H3 as palmitoilações das cisteínas são importantes para a sua reunião ao nível das membranas, muito provavelmente devido a uma maior capacidade de associação às jangadas lipídicas [Chen 2005]. Jangadas lipídicas são pequenas zonas na membrana, ricas em colesterol e esfingomiélin e assume-se que funcionam como um instrumento para a concentração seletiva de proteínas [Varma 1998]. Outra hipótese é a existência de interações específicas entre a CC e/ou palmitoilações e as proteínas da matriz M1. Por exemplo, um vírus do subtipo H3 com mutação nas cisteínas da CC possui uma inferior associação às jangadas lipídicas

e reunião ao nível da membrana, ao passo que o gene HA do subtipo H3 num vírus H1 não possui qualquer efeito na reunião [Jin 1996, Chen 2005].

A partição da HA nas jangadas lipídicas não é relevante para a fusão [Scheiffele 1997], podendo mesmo ocorrer numa maior extensão que na HA nativa quando não está em jangadas [Melikyan 1999]. Contudo a afinidade da HA para frações insolúveis em detergente (o que se crê estar relacionado com as jangadas lipídicas) é influenciada pelo PT [Scheiffele 1997]. Por exemplo, a palmitoilação das cisteínas do PT é necessária para recrutar a HA para jangadas lipídicas [Scolari 2009].

O domínio transmembranar e a cauda citoplasmática não são os únicos segmentos com implicações no mecanismo de fusão. O próprio *linker* entre o ectodomínio da estrutura cristalográfica e o domínio transmembranar também influencia.

1.6.3. *Linker*

Inserções entre o domínio transmembranar e o *linker* de até 8 resíduos são capazes de originar hemaglutininas capazes de formar poros de fusão, ao passo que inserção de 10 resíduos apenas permite mistura de lípidos, ou seja, fica preso no estado intermédio de hemifusão. Deleções com mais de 8 resíduos ou inserções superiores a 10 resíduos abolem completamente o processo de fusão, possivelmente devido ao aumento da distância entre os terminais da HA2. Dois fatores que apoiam esta sugestão são: quanto maior foi o tamanho do fragmento inserido menor foi a eficiência da fusão, e a destabilização das interações na *N-cap* também a reduz [Li 2008].

Curiosamente a distância entre a arginina na posição 170 e a leucina na posição 187, que são completamente conservadas (e as únicas nesse segmento), difere entre todos os subtipos de HA em apenas um resíduo de aminoácido, mostrando uma alta conservação no tamanho do *linker* entre o ectodomínio e o domínio transmembranar [Li 2008].

2. Modelação e simulação molecular

Em ciência e sobretudo em ciências biológicas ou ciências da vida a forma mais comum para procurar a resposta a um problema é através da experimentação. Contudo apenas algumas propriedades de um sistema biomolecular estão acessíveis a medições experimentais, quer por limitações metodológicas, quer pelo custo da experiência ou mesmo pelas dimensões e escalas de tempo a que uma determinada propriedade ocorre. Este último argumento é a razão chave para a utilidade da modelação/simulação molecular, uma vez que usando esta técnica tanto podemos obter médias, distribuições e evolução temporal de uma propriedade à escala microscópica e a curtas escalas de tempo [Van Gunsteren 2006].

As observações experimentais ocorrem sobre sistemas macroscópicos, sendo a medição uma média de uma certa propriedade durante o tempo da medição e de todas as moléculas do sistema. Idealmente dever-se-ia simular sistemas com $\sim 10^{23}$ moléculas, mas devido a limitações computacionais apenas é permitido simular um número restrito de moléculas ou mesmo uma molécula apenas. Então como podemos inferir propriedades macroscópicas se a simulação molecular possui uma escala microscópica? A resposta é através da mecânica estatística, uma “ferramenta” matemática que consegue interligar as propriedades de uma molécula individual (microscópica) às propriedades termodinâmicas de sistemas macroscópicos [Chandler 1987, Bem-Naim 1992].

O pressuposto básico da mecânica estatística é a hipótese ergódica, a qual estabelece que se seguirmos uma única molécula, durante um período de tempo suficientemente longo, as propriedades termodinâmicas médias serão iguais às propriedades médias de todas as moléculas de um *ensemble*, a um determinado instante [Van Holde 1998, Frenkel 2002]. *Ensemble* é um conjunto de todos os microestados possíveis que uma molécula pode assumir, dependendo das restrições impostas: por exemplo para um *ensemble* isotérmico-isobárico (*ensemble* NPT), à temperatura e pressão constantes [Chandler 1987].

A técnica, utilizada neste trabalho, para seguir uma proteína ao longo do tempo e que será descrita mais à frente, foi a simulação da dinâmica molecular, primeiramente introduzida nos finais dos anos 50 por Alder e Wainwright utilizando esferas rígidas **[Alder 1957]**. A primeira simulação utilizando uma função de energia potencial molecular em argon líquido foi realizada em 1964 por Rahman et al. **[Rahman1964]**. Apenas 10 anos mais tarde é que foi realizada a primeira simulação de um sistema realista em água **[Stillinger 1974]** e três anos mais tarde a simulação da primeira proteína, o inibidor da tripsina pancreática bovina **[McCammon 1977]**. Nessa altura o poder computacional era uma barreira para as dimensões do sistema e para o tempo que se podia simular. Contudo, o contínuo aumento do poder computacional, diminuição do seu custo e melhoramento dos modelos, tornou possível a simulação de sistemas mais complexos e durante períodos de tempo superiores. Por exemplo, utilizando o supercomputador Anton desenhado e construído pelo Instituto de bioquímica computacional de David Shaw, Lindorff-Larsen et al. simulou 12 proteínas (com simulações representando todos os átomos) ao longo de centenas de microssegundos, chegando algumas mesmo a milissegundos **[Lindorff-Larsen 2011]**.

Mesmo com a evolução computacional a simulação de moléculas biológicas englobando todos os seus eletrões seria impraticável para tempos de simulação de interesse. Na mecânica molecular podem-se desprezar os eletrões devido à aproximação de Born-Oppenheimer. Esta aproximação assume que sendo a massa dos eletrões muito inferior ao núcleo, este último vai ter uma velocidade muito inferior aos eletrões. Portanto na escala de tempo do movimento do núcleo, os eletrões vão rapidamente relaxar e assumir a distribuição do estado fundamental, ou seja, podem-se calcular separadamente os movimentos do núcleo e dos eletrões. Ao considerar apenas o movimento dos núcleos perde-se no entanto a capacidade de analisar propriedades que provém da distribuição eletrónica, como a formação ou destruição de ligações químicas **[Leach 2001, Ramachandran 2008]**.

A simulação computacional apresenta claras limitações, mas pode ser vista como um complemento da experimentação, podendo ajudar a interpretar e perceber certas observações experimentais, correlacionar detalhe microscópico com propriedades macroscópicas e fornecer informações que não se podem obter experimentalmente.

2.1. Mecânica molecular

O termo mecânica molecular (MM) apareceu nos anos 70 e é atualmente o termo mais usado, mesmo sendo também conhecido como método de campo de forças [Hinchliffe 2008]. O modelo considerado pela mecânica molecular assume tanto um modelo estrutural em que os átomos são considerados bolas esféricas, como também necessita de um conjunto de diferentes funções potenciais empíricas. As funções potenciais são também conhecidas por campos de força, porque descrevem as interações/forças entre os átomos. Cada uma destas funções é um equilíbrio entre o custo computacional a calcular e a capacidade de reproduzir os resultados físicos experimentais. Logo, as funções potenciais utilizadas provém de princípios físicos, mas a sua parametrização é empírica, reproduzindo aproximadamente o comportamento do sistema físico, através da comparação com dados experimentais [Saxena 2009]. Os valores dos parâmetros são obtidos através de parametrizações/otimizações, ou seja, a tentativa de correlacionar os dados obtidos pelas simulações com os dados experimentais. As parametrizações podem ter âmbitos mais específicos, ou seja, os parâmetros utilizados são ajustados apenas para reproduzir as propriedades de um pequeno número/tipo de moléculas, ou então um âmbito mais geral. Neste último a transmissibilidade torna-se importante num campo de forças, onde os parâmetros ajustados para algumas moléculas esperam-se que reproduzam os efeitos de um grande número de moléculas. Posto isto é necessário saber o que se quer estudar e que análises queremos obter antes de escolher o campo de forças a utilizar [Becker 2001].

Em suma, um campo de forças é definido pelas equações potenciais e os respetivos parâmetros. As funções potenciais para cada tipo de interação são englobadas numa única função de energia potencial, que serão debatidas com mais detalhe de seguida.

2.1.1. Função de energia potencial empírica

A função de energia potencial possui uma componente ligante, ou seja, a contribuição dada pela vibração das ligações, dos ângulos e dos diedros e uma componente não ligante, isto é, resultante das interações eletrostáticas e de van der Waals:

$$\begin{aligned} V_{\text{total}} &= V_{\text{ligante}} + V_{\text{não ligante}} \\ &= V_{\text{ligações}} + V_{\text{ângulos}} + V_{\text{diedros}} + V_{\text{diedros impróprios}} + V_{\text{electrostática}} + V_{\text{LJ}} \end{aligned}$$

Interações ligantes

A primeira parcela da função de energia potencial relaciona-se com a variação de energia causada com a deformação da ligação entre dois átomos. Visto que, a temperaturas normais, a distância não varia significativamente do valor de equilíbrio, podemos assumir que uma função simples como a de Hooke é adequada:

$$V_{\text{ligações}} = \frac{1}{2} K_l (d_{ij} - d_l)^2$$

Onde K_l é a constante de força, d_{ij} é a distância instantânea entre os átomos e d_l é a distância em equilíbrio dos dois átomos, ou seja, é o valor adotado pela estrutura de menor energia. Note-se que quanto maior for a diferença entre a distância de equilíbrio e a distância num determinado momento, maior será a energia. O mesmo ocorre com a constante de força, onde um valor elevado produzirá uma maior rigidez nas variações.

A segunda parcela das interações ligantes é a flexão dos ângulos entre três átomos (i, j e k), sendo também descrita por um potencial harmônico:

$$V_{\text{ângulos}} = \frac{1}{2} K_a (\theta_{ijk} - \theta_a)^2$$

Onde K_a é a constante de força e θ_{ijk} , θ_a são respectivamente o ângulo instantâneo entre os átomos e o ângulo de equilíbrio.

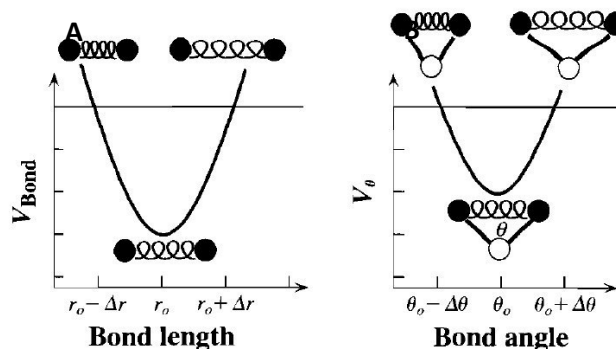


Figura 11 – Variação da energia potencial com a variação da distância entre átomos (A) ou da variação do ângulo entre três átomos (B). Adaptado de [van Holde 1998].

Variações na distância das ligações e no ângulo entre duas ligações necessitam de muita energia em comparação com outras deformações, como as torções que ocorrem entre quatro átomos ligados em sequência (i , j , k e l). As torções são caracterizadas por ângulos diedros, ou seja, o ângulo entre o plano constituído pelos três átomos $i - i+2$ e o plano constituído pelos átomos $i+1 - i+3$, cujas energias são contabilizadas na função de energia potencial através de uma equação sinusoidal:

$$V_{\text{diedros}} = K_d [1 + \cos(n\psi_{ijkl} - \psi_d)]$$

Onde K_d é a constante de força que define a amplitude, n é a multiplicidade, ou seja, é o número de mínimos que a função assume ao longo de uma rotação completa de 360° , ψ_{ijkl} é o ângulo diedro instantâneo e ψ_d é a fase que define os máximos e pode assumir valores de 0° ou 180° para torções entre ligações simples ou duplas, respetivamente. Note-se que a energia é mínima quando os átomos estão em conformação trans (para configurações sp^2) ou conformação estrelada (para configurações sp^3) e máxima quando em conformação cis (sp^2) ou eclipsada (sp^3), devido às interações antiligantes entre os átomos opostos.

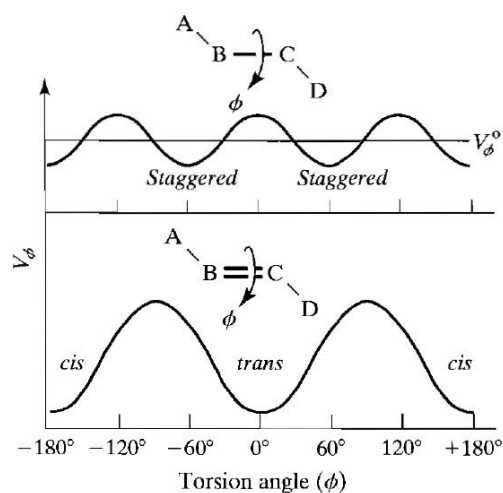


Figura 12 – Variação da energia potencial com a variação do ângulo diedro entre quatro átomos. Adaptado de [van Holde 1998].

O último termo corresponde a um caso específico dos diedros, que são os chamados de diedros impróprios, descritos por uma equação harmônica com vista a manter os anéis aromáticos planares e impossibilitando a alteração de enantiômeros (mantendo a quiralidade).

$$V_{\text{diedros impróprios}} = \sum_{\text{diedros impróprios}} K_{di} (\psi_{ijkl} - \psi_{di})^2$$

Onde K_{di} é a constante de força, ψ_{ijkl} é o ângulo diedro instantâneo e ψ_{di} é o ângulo diedro de equilíbrio.

Interações não ligantes

A contribuição das interações não ligantes é bastante importante, ocorrendo entre todos os pares de partículas do sistema. Traduzindo-se não só num maior número de interações em comparação com as interações ligantes, como também na existência de interações com as moléculas do ambiente envolvente. As interações não ligantes dividem-se entre dois tipos: um potencial de Lennard-Jones (LJ) para todas as partículas e uma função de energia de Coulomb para todos os átomos com carga parcial.

O potencial de Lennard-Jones apresenta uma variação do tipo apresentado na figura 13, onde se mostra a energia potencial em função da distância entre o centro de massa de duas partículas. As partículas a distâncias muito longas interagem pouco, mas

quando se aproximam sofrem interações do tipo dipolo instantâneo – dipolo induzido (forças dispersivas de London) fazendo as partículas aproximarem-se até a uma distância de equilíbrio (mínimo da função). Mas quando dois átomos se aproximam demasiado e as suas nuvens eletrônicas sobrepõem-se, segundo o princípio de exclusão de Pauli, os elétrões do mesmo spin não podem partilhar o mesmo estado, acabando por provocar uma força repulsiva entre os átomos [Hinchliffe 2008].

As interações entre átomos covalentemente ligados (i+1) ou unidos por um (i+2) átomo, são designados normalmente por primeiro e segundo vizinho, respetivamente. Estas interações são do campo da mecânica quântica, logo não podem ser descritas por um potencial de Lennard-Jones. No entanto considera-se que estas interações são descritas aceitavelmente pelas interações ligantes. No caso dos terceiros vizinhos (i+3), a repulsão entre alguns átomos pode ser demasiado forte, como por exemplo, a interação entre carbonos no cis-butano, portanto para esses a força de repulsão é diminuída [Leach 2001, Van der Spoel 2005].

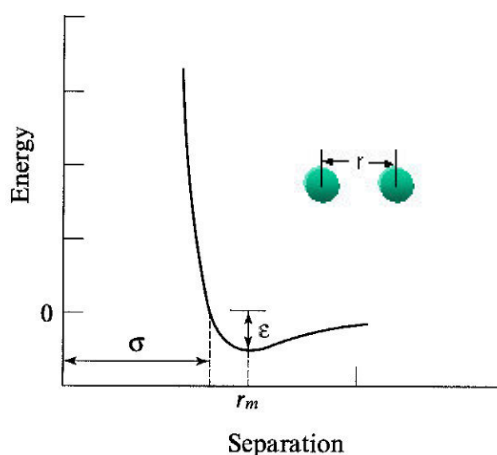


Figura 13 – Potencial de Lennard-Jones. Adaptado de [Leach 2001].

A equação que descreve a função é:

$$V_{\text{Lennard-Jones}}(r_{ij}) = \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

Onde ϵ_{ij} representa a energia mínima da função, σ_{ij} representa o tamanho efetivo, ou seja, a distância mínima efetiva da aproximação de duas partículas ($E_{\perp}=0$), C_{12} e C_6 refletem, respetivamente, a componente repulsiva e a componente atrativa.

A segunda parcela das interações não ligantes são as interações eletrostáticas que são descritas pela lei de Coulomb:

$$V_{el} = \sum_{\text{pares não ligantes}} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_{rel} r_{ij}}$$

Onde q_i e q_j são as cargas parciais dos átomos i e j , respetivamente, ϵ_{rel} a constante dieléctrica do meio e ϵ_0 a permitividade em vácuo.

Na figura 14 estão esquematizados os diferentes tipos de contribuições, na mecânica molecular, utilizados na função de energia potencial.

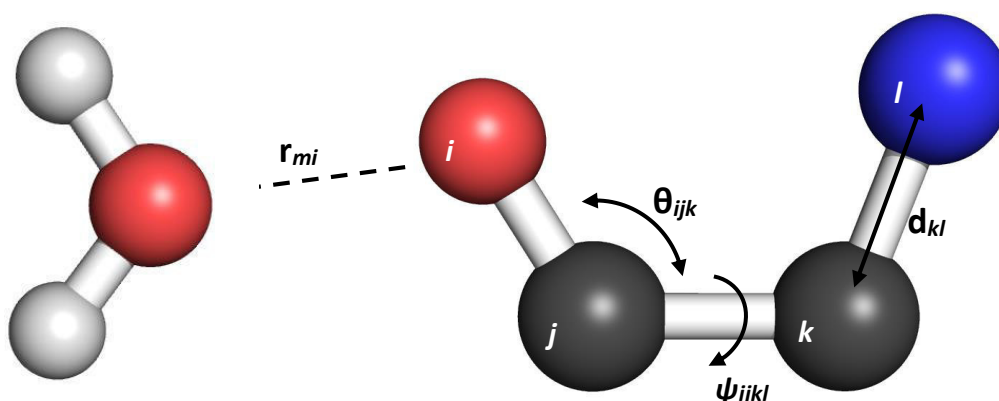


Figura 14 – Representação das diferentes interações que ocorrem entre os átomos pesados da glicina, a vibração das ligações (d_{kl}), a dobra dos ângulos (θ_{ijk}), torção dos diedros (ψ_{ijkl}) e as interações não ligantes (r_{mi}). Representação em *balls and sticks*, a azul o átomo de azoto, a vermelho o átomo de oxigénio, a preto os átomos de carbono e a branco os átomos de hidrogénio.

No âmbito deste trabalho utilizou-se o campo de forças GROMOS 54A7 [Schmid 2011], visto ser, o mais recente da família, incluir parâmetros de lípidos e não ter o problema do seu antecessor (53A6), de destabilizar demasiado as hélices [Schmid 2011]. Em parte serve também como comparação e continuação de um trabalho realizado no grupo de modelação de proteínas do ITQB onde se simulou o péptido de fusão e o péptido transmembranar usando o campo de forças 54A7 [Victor 2012, Victor 2015]. Utilizaram-se ainda outros campos de força, o MARTINI [Marrink 2007, Monticelli 2008] e o ELNEDIN [Periole 2009], que são campos de força *coarse-grain* onde as suas vantagens e desvantagens serão debatidas noutra capítulo desta tese (capítulo 2.2).

A mecânica molecular, só por si, dá-nos apenas informação sobre a energia de um sistema em determinadas posições relativas dos átomos, por exemplo, a energia associada a uma conformação de uma proteína, sendo portanto estática. Para se poder

criar a superfície de energia do sistema é necessário que a proteína apresente o máximo de conformações (idealmente, todas as conformações possíveis), e assim aplicar a mecânica estatística.

2.1.2. Dinâmica molecular

A simulação de dinâmica molecular é um método de amostragem do espaço conformacional, onde são geradas conformações relacionadas temporalmente entre si, utilizando sempre a conformação anterior para gerar uma nova conformação, após um determinado tempo, através da mecânica clássica. Sendo assim a simulação de dinâmica molecular é uma tentativa de simular o comportamento das moléculas ao longo do tempo.

Um exemplo comum na mecânica clássica é um sistema de uma esfera e uma mola: ao empurrar a esfera contra a mola, esta exerce um trabalho negativo sobre a esfera, pois, a força da mola transforma a energia cinética na esfera em energia potencial elástica. Quando a energia cinética se anula, a esfera para, e começa a mover-se no sentido contrário, na mesma direção da força da mola, passando a exercer trabalho positivo, transformando por sua vez, energia potencial em energia cinética [Halliday 2001]. Posto tal, temos que a diferença da energia potencial (ΔV) é igual ao trabalho (W) realizado pela força, esquematizado nas seguintes equações:

$$\Delta V = -W = -F_i \Delta r_i$$

$$F_i = -\frac{\delta V}{\delta r_i} = -\nabla_i V$$

Onde F_i é a força aplicada no átomo i e r_i é o vetor com as posições cartesianas do átomo i . Obtemos portanto que a força aplicada nos átomos é o gradiente da energia potencial, sendo calculada apenas pela derivação da equação de energia potencial em ordem a todas as coordenadas. A superfície de energia potencial é um campo escalar, portanto o seu gradiente vai originar um campo vetorial de forças, daí o nome dado como “campo de forças” ao conjunto da energia potencial e os correspondentes parâmetros.

Após o cálculo da energia potencial, e da força correspondente a partir do gradiente, utiliza-se as leis de movimento de Newton. A primeira lei estabelece que um corpo continua a mover-se numa linha reta com velocidade constante até que uma força atue sobre o corpo. A segunda lei estabelece que a força é igual à taxa de variação do momento linear ($F = dp/dt$). Por último a terceira lei estabelece que para qualquer ação ocorre uma igual e oposta reação, ou seja, se um corpo exerce uma força F_1 sobre um outro corpo, este último vai exercer uma força $F_2 = -F_1$ sobre o primeiro corpo [Leach 2001].

Visto que, nos sistemas simulados a massa é mantida constante e $p = m \times v$, a trajetória dos átomos perante uma certa força é resolvida pela segunda lei de Newton, ou equação de movimento de Newton:

$$F_i = m_i \frac{dv_i}{dt} \quad \text{e} \quad v_i = \frac{dr_i}{dt} \quad \text{então:} \quad F_i = m_i \frac{d^2 r_i}{dt^2} = m_i a_i$$

Onde F_i, m_i, v_i, r_i e a_i são respetivamente a força exercida sobre o átomo i , a massa do átomo i , a velocidade do átomo i após a força atuar, o vetor de posições cartesianas do átomo i após a força atuar e a aceleração do átomo i após a força atuar.

A partir da integração da equação da segunda lei de Newton podemos determinar as novas posições e velocidades de todos os átomos do sistema, sendo no entanto necessário métodos numéricos, devido à complexidade dos sistemas em causa. O método de integração usado foi o “salto de rã” (*leapfrog*) [Hockney 1974], que é um método simples, possui propriedades de conservação de energia e é amplamente utilizado [Rapaport 2004]. Este método baseia-se em três passos: primeiro utiliza as posições r_i a um determinado tempo t , para calcular a força F_i ; segundo usa essa força, anteriormente calculada e as velocidades do meio passo anterior ($t - \Delta t/2$) para calcular as velocidades do meio passo posterior ($t + \Delta t/2$); por último utilizam-se as velocidades, anteriormente calculadas no segundo passo, para calcular as posições dos átomos no novo passo de integração ($t + \Delta t$) [Becker 2001, Van der Spoel 2005]. As equações utilizadas na integração são as seguintes:

$$v_i \left(t + \frac{\Delta t}{2} \right) = v_i \left(t - \frac{\Delta t}{2} \right) + \frac{\Delta t}{m_i} F_i(t)$$

$$r_i(t - \Delta t) = r_i(t) + \Delta t v_i\left(t + \frac{\Delta t}{2}\right)$$

O tempo entre cada passo de integração não é algo estrito, dependendo de várias variáveis. Vejamos a figura 15, que mostra o resultado para diferentes tempos. Ao utilizar um espaço de tempo curto para cada passo de integração a amostragem do espaço conformacional ocorre lentamente, aumentando o tempo de computação. Por outro lado o uso de espaços de tempo longos para os passos de integração provoca instabilidades, choques entre átomos ou mesmo átomos a afastarem-se demasiado. O passo de integração deve ser, portanto, um equilíbrio entre uma amostragem eficiente do espaço conformacional e uma simulação o mais próxima do real possível [Leach 2001].

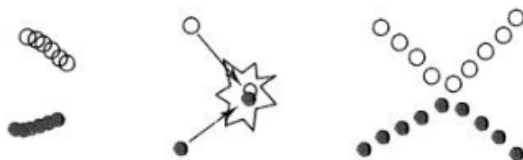


Figura 15 – Variações dos átomos a diferentes tempos de passos de integração. Adaptado de [Leach 2001].

A escolha do valor do passo de integração depende portanto do sistema, e o espaço de tempo onde ocorrem os movimentos mais rápidos desse sistema. No caso deste trabalho, não se simulam átomos, mas sim moléculas flexíveis, portanto as vibrações mais rápidas são as vibrações das ligações entre átomos ligados covalentemente, principalmente com átomos de hidrogénio. Os períodos de vibração de ligações entre um carbono e um hidrogénio ocorrem a cada 10fs [Leach 2001, Becker 2001]. Dependendo da fonte bibliográfica, o tempo de integração a utilizar nestes sistemas deveria ser aproximadamente um décimo [Leach 2001] ou um vigésimo [Becker 2001] da vibração mais rápida, ou seja, um valor de 0,5-1fs para o passo de integração é aceitável.

Visto que o tempo de integração está diretamente relacionado com a vibração mais frequente, e esta ocorre nas ligações covalentes entre os átomos, uma forma de aumentar esse tempo é constranger as ligações [Van Gunsteren 1990]. A aplicação das constrições é feita por algoritmos como o SHAKE [Ryckaert 1977], o SETTLE [Miyamoto 1992] ou o LINCS [Hess 1997], sendo este último o mais comum, devido à sua maior

estabilidade e rapidez. O algoritmo utilizado para os péptidos/proteínas neste trabalho foi o LINCS que atua em dois passos, na recolocação dos átomos após um passo de integração, de modo a que as distâncias de equilíbrio entre os átomos se mantenham. Num primeiro passo o algoritmo projeta a nova ligação na ligação antiga e num segundo passo é aplicado um acerto na distância para se ter em conta a rotação da ligação (figura 16) [Hess 1997].

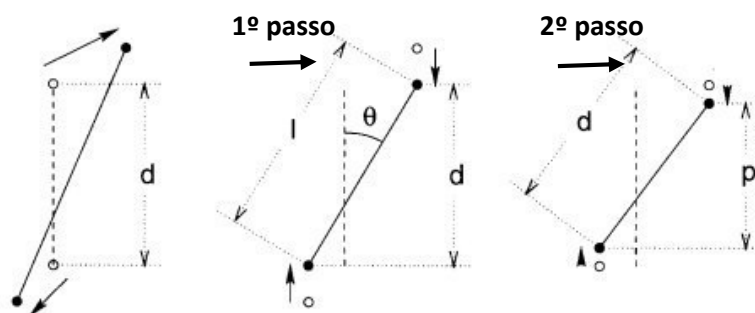


Figura 16 – Esquema da atualização das posições a cada passo de integração. A linha a tracejado simboliza a ligação anterior com comprimento d e a linha a preto é a nova ligação. $I = d \cos \theta$ e $p = (2d^2I^2)^{1/2}$. Adaptado de [Hess 1997].

Contudo, o LINCS apenas permite restrições nas ligações entre os átomos. No caso da água é necessário também restringer os ângulos. Para tal é necessário utilizar o algoritmo SETTLE, que é uma solução analítica do algoritmo SHAKE para moléculas pequenas.

Ao restringer as ligações, fonte de maior vibração, tempos de integração de 2fs podem ser utilizados.

2.1.3. Solvatação e condições periódicas de fronteira

Em dinâmica molecular podem-se realizar simulações com a(s) molécula(s) de interesse em vácuo, mas o estudo de sistemas biológicos implica a existência de um solvente, normalmente a água. Para tal utiliza-se uma caixa de simulação, por exemplo cúbica, onde são colocadas as moléculas de interesse (soluto) e preenche-se o resto com água. Contudo tal procedimento apresenta dois problemas: num sistema microscópico, normalmente simulado com um número de moléculas entre 10^5 e 10^6 átomos, as superfícies da fronteira têm um grande impacto no sistema [Van Gunsteren 2006], onde moléculas perto do limite vão sofrer forças diferentes das moléculas no centro da caixa;

o segundo problema ocorre quando após um passo de integração a molécula sai da caixa, o que altera a densidade do sistema.

A solução para ultrapassar este problema é utilizar as condições periódicas de fronteira, que são uma abstração, onde consideramos que a caixa inicial é rodeada por cópias idênticas da mesma (figura 17). Nesta condição qualquer movimento de um átomo é reproduzido nas suas imagens periódicas, solucionando o problema das fronteiras, pois a força a atuar num átomo depende dos átomos nas restantes caixas. O problema de a densidade variar é também solucionado, visto que, quando um átomo sai da caixa ele volta a entrar através de uma imagem periódica pelo lado contrário da caixa, com igual velocidade.

No entanto ao utilizar esta metodologia originam-se outros problemas, como por exemplo, o número de interações não ligantes aumenta infinitamente. A convenção mínima da imagem estipula também que nenhum átomo deve interagir com o soluto (por exemplo uma macromolécula) e a imagem periódica desta.

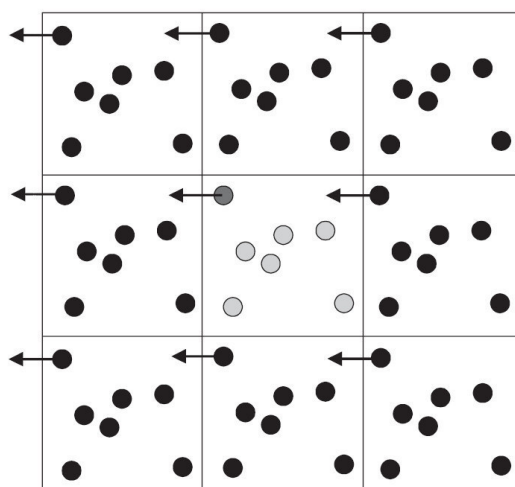


Figura 17 – Esquema do efeito das condições periódicas de fronteira, com apenas algumas caixas representadas. Os átomos reais estão representados a cizento e as imagens periódicas a preto. Adaptado de [Hinchliffe 2008]

Uma solução para ultrapassar esses problemas é considerar uma esfera imaginária com centro em cada átomo e com um raio pré-determinado (designado por *cut-off*), onde as energias das interações não ligantes, com átomos fora desse *cut-off*, são consideradas 0. Mas só por si, ainda teríamos de calcular a cada passo de integração as distâncias entre os átomos para saber que átomos se encontravam dentro do *cut-off*.

Visto que, os átomos vizinhos não variam significativamente entre 5-10 passos de integração, cria-se uma lista com os vizinhos a uma dada distância e a um dado tempo, atualizando a lista após n passos de integração, ganhando assim tempo de cálculo. A taxa de atualização da lista de vizinhos deve ter em atenção, não só a eficiência da simulação (quanto maior for o n mais rápida a simulação), mas também uma maior probabilidade dos átomos saírem ou entrarem dentro do *cut-off*, antes da atualização, provocando cálculos de forças errados.

Existem dois métodos para limitar as interações não ligantes utilizando *cut-offs*, o *twin-range* [Berendsen 1985] e o *shift* [Van der Spoel 2005]. No caso do método *twin-range* são definidos dois raios, onde as posições dos átomos e correspondentes forças de interação dentro do primeiro *cut-off* (mais curto) são calculadas em todos os passos de integração. Para os átomos entre o primeiro e o segundo *cut-off* as forças da interação são apenas calculadas quando se atualiza da lista de vizinhos, mantendo-se constantes entre as atualizações.

No caso do método *shift* utilizam-se também dois raios, mas usam a mesma lista de vizinhos, onde a partir do primeiro *cut-off* (r_{shift}) as interações começam gradualmente a ser zero à medida que se aproxima do segundo *cut-off* (r_{cut}).

No caso da interação não ligante de Lennard-Jones a aproximação ao utilizar o *cut-off* é mínima, uma vez que a uma distância de 2.5σ (figura 13) o potencial de Lennard-Jones possui 1% do valor da distância de σ e os valores de *cut-off* usados são superiores a 2.5σ . Já no caso das interações eletrostáticas isto já não acontece, pelo que é importante ter em consideração interações de longo alcance. Utiliza-se portanto o método do *Reaction-field* [Tironi 1995] que trata as moléculas, que se encontram a uma distância superior ao *cut-off*, de uma forma média, sendo consideradas como um dielétrico contínuo.

2.1.4. Banhos de temperatura e pressão

Enquanto a utilização direta da simulação de dinâmica molecular seria simular condições NVE (número de partículas, volume e energia constantes), se quisermos estudar sistemas biológicos devemos simular condições NPT (número de partículas,

pressão e temperatura constantes). No entanto dentro dos sistemas simulados pode ocorrer interconversão entre energia potencial e energia cinética, ou seja, a energia cinética do sistema muda, alterando por consequente a temperatura do sistema devido à seguinte equação:

$$E_{\text{cinética}}(t) = \sum_{i=1}^N \frac{1}{2} m_i v_i^2(t) = \frac{1}{2} N_{gl} k_B T(t)$$

Onde m_i é a massa do átomo i , v_i é a velocidade do átomo i , N_{gl} é o número de graus de liberdade e k_B é a constante de Boltzmann.

Como ocorrem variações na pressão e na temperatura ao longo das simulações é necessário utilizar algoritmos que assegurem a conservação da temperatura e da pressão. Por exemplo, o método de acoplamento fraco de Berendsen para a temperatura [Berendsen 1984] utilizado neste trabalho, acopla, a cada passo de integração, o nosso sistema a um banho externo à temperatura desejada (T_{banho}), que vai retirando ou introduzindo calor consoante a diferença entre a temperatura do banho e a temperatura do sistema:

$$\frac{dT}{dt} = \frac{(T_{\text{banho}} - T)}{\tau_t}$$

$$\Delta T = \frac{\Delta t}{\tau_t} (T_{\text{banho}} - T)$$

Onde Δt é o tempo do passo de integração e τ_t é a constante de acoplamento, que pode variar consoante se deseja um acoplamento mais forte (τ_t pequeno), permitindo menos variação na temperatura, ou um acoplamento mais fraco (τ_t grande).

Uma vez que a temperatura está relacionada com a velocidade dos átomos, a forma como o banho introduz calor no nosso sistema é através do escalamento das velocidades. Através das equações acima, podemos inferir a equação:

$$\lambda^2 = 1 + \frac{\Delta t}{\tau_t} \left(\frac{T_{\text{banho}}}{T(t)} - 1 \right)$$

Onde λ é o fator de escalamento das velocidades, o qual é multiplicado a todas as velocidades dos átomos em cada passo de integração.

No caso do acoplamento fraco de Berendsen para a pressão [Berendsen 1984], também se usa um “banho” de pressão e um fator de escalamento, que neste caso se aplica ao volume da caixa de simulação, segundo as seguintes equações:

$$\frac{dP}{dt} = \frac{(P_{\text{banho}} - P)}{\tau_p}$$
$$\lambda = 1 - \kappa \frac{\Delta t}{\tau_p} \left(\frac{P_{\text{banho}}}{P(t)} - 1 \right)$$

Onde κ é a compressibilidade isotérmica, que mede a variação relativa do volume em resposta a uma variação da pressão.

Já que o volume da caixa se altera ao multiplicar-se por λ , a posição relativa dos átomos também necessita de ser alterada, multiplicando-se por $\lambda^{1/3}$, sendo as novas posições dadas por:

$$r'_i = \lambda^{1/3} r_i$$

Existem três tipos de acoplamento para a pressão. Isotrópico, onde as constantes de acoplamento dos três eixos da caixa (x,y e z) são iguais, normalmente utilizado para simulações de proteína em água. Semi-isotrópico, onde duas constantes são iguais e outra é diferente, normalmente utilizado em simulações com membranas, sendo os eixos xy acoplados com a mesma constante e o eixo z com uma constante diferente. Por último o acoplamento anisotrópico, onde todas as constantes de acoplamento de pressão, nos três eixos, são diferentes.

2.2. Campos de força não atomísticos

A vantagem dos campos de força com alta resolução é a capacidade de aceder às interações átomo-átomo, mas esta resolução reduz, na prática, o tamanho dos sistemas a dezenas de nanómetros, e o tempo de simulação a centenas de nanossegundos. No entanto a maioria dos processos biomoleculares ocorrem a escalas de tempo superiores, como o *folding* de proteínas, formação de vesículas lipídicas e o agrupamento de sistemas complexos em bicamadas [Marrink 2007, Bennun 2009].

Num sistema molecular as interações não ligantes, por ocorrerem entre todos os átomos (exceto os primeiros vizinhos ligados covalentemente), são em número superior às interações ligantes. Uma vez que os cálculos da energia potencial, especialmente os termos de interações não ligantes, são o que mais contribui para o consumo de tempo no cálculo das forças, então quanto menor for o número de átomos no sistema a simular, menor será o número de interações e por sua vez mais rápidos serão os cálculos. Assim sendo, uma das formas de aumentar a velocidade das simulações é reduzindo a resolução do sistema, ou seja, o número de partículas que o compõem [Van Gunsteren 1990, Bennun 2009].

A questão principal para saber que resolução utilizar é: que propriedades se querem estudar? Se queremos estudar algo mais estrutural como interações entre moléculas deve-se utilizar uma resolução alta. Contudo se quisermos estudar comportamentos mais globais a grandes escalas de tempo, onde as interações de alta resolução são dispensáveis, modelos de baixa resolução podem ser vantajosos [Bennun 2009].

O *multi-scale*, ou seja, uso de diferentes resoluções, surge como um método que pode juntar as vantagens de ambas as resoluções. Neste trabalho utiliza-se primeiro a baixa resolução para aceder a eventos que ocorrem a largas escalas de tempo (eficiência) e depois utiliza-se uma resolução superior para se estudar os detalhes atomísticos (precisão) [Marrink 2004, Rzepiela 2010]. Podem também ser realizadas simulações onde diferentes componentes do sistema apresentam diferentes resoluções, no entanto para o campo de forças MARTINI, este método ainda necessita de melhoramentos.

Os campos de força utilizados neste trabalho foram o 54A7 para a alta resolução e o MARTINI para a resolução mais baixa, ambos explicados de seguida.

2.2.1. Campos de força *united-atom*

Campos de força *united-atom* tratam os hidrogénios apolares como parte integrante do átomo pesado a que estão covalentemente ligados. A esse átomo pesado

é lhe atribuído um novo tipo de átomo com parâmetros de eletrostática e de Lennard-Jones diferentes, para que simule a presença dos hidrogénios. Usualmente este novo átomo é colocado no centro de massa do átomo pesado, como o campo de forças GROMOS utilizado neste trabalho, mas também existem campos de força onde o novo átomo é colocado no centro geométrico. Os hidrogénios polares, dado que são necessários para a formação de pontes de hidrogénio e outras interações polares, são tratados explicitamente [Leach 2001, Hinchliffe 2008].

Um ponto importante nos campos de força *united-atom* é a transição entre diferentes enantiômeros devido à perda de resolução. No caso de proteínas à exceção da glicina e da prolina todos os resíduos de aminoácidos possuem no carbono alfa um centro quiral, mas ao perder o hidrogénio altera a sua estereoquímica. O uso de um termo de torção imprópria para os átomos N-C- α -R mantém a posição relativa da cadeia lateral correta [Leach 2001].

2.2.2. Campo de forças MARTINI

Na criação do campo de forças MARTINI, houve a preocupação de criar um equilíbrio entre quatro aspetos: a velocidade de cálculo, através do uso de interações de curto alcance e grandes passos de integração (discutidos mais à frente); versatilidade, por poder ser aplicado a qualquer molécula mantendo algum do seu detalhe estrutural; aplicabilidade, ao ser um campo de forças simples e fácil de usar, com poucos parâmetros; e precisão semi-quantitativa através da reprodução de resultados dinâmicos, estruturais e termodinâmicos de simulações atomísticas. [Marrink 2004, Marrink 2007].

As moléculas em MARTINI têm uma resolução inferior do que a encontrada nos tratamentos atomísticos ou mesmo nos *united-atom*, sendo também designada por resolução *coarse-grain*. Nesta resolução é usado um mapeamento de quatro para um, isto é, em média, quatro átomos pesados (hidrogénios não são contabilizados) são transformados em apenas uma partícula [Marrink 2003]. Para moléculas com anéis, como por exemplo, o benzeno ou a cadeia lateral do triptofano, usa-se uma resolução superior de dois para um [Marrink 2007].

Cada partícula utilizada na resolução *coarse-grain* pode representar diferentes tipos de átomos, logo é necessário criar diferentes partículas que definam a origem química da molécula. O campo de forças MARTINI possui quatro tipos diferentes de partículas. Partículas polares (P) que representam grupos neutros de átomos que facilmente se dissolvem em água, como por exemplo, o etanol ou a acetamida. Partículas apolares (C) que representam grupos de átomos hidrofóbicos, como o butano ou o clorofórmio. Partículas não polares (N) que representam grupos de átomos mistos, com porções polares e porções apolares, por exemplo, o propanal ou o 1-butanol. Por último, partículas carregadas (Q) que representam grupos ionizados, como é o caso dos grupos colina e fosfato.

Dentro dos quatro diferentes tipos, existem ainda subtipos que ajudam a definir melhor a natureza química da partícula. Para os tipos Q e N existem quatro subtipos (*da*, *d*, *a* e *0*), consoante a sua capacidade para formar pontes de hidrogénio, onde *0* significa que não tem capacidade, *a* significa que é aceitador, *d* que é dador e *da* significa que tanto pode ser dador como aceitador de pontes de hidrogénio. No caso dos tipos P e C são divididos em cinco subtipos (1, 2, 3, 4 e 5) consoante a sua polaridade, onde 1 significa baixa polaridade e 5 alta polaridade [Marrink 2007]. Um exemplo da correspondência entre a representação de moléculas em campos de força *united-atom* e no campo de forças MARTINI, encontra-se na figura 18.

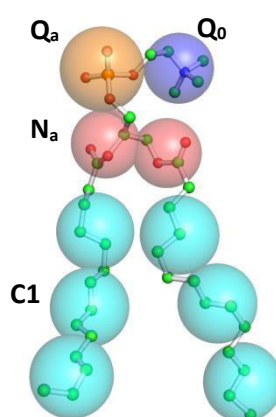


Figura 18 – Representação de um lípido DMPC em representação atomística (sem hidrogénios) e *coarse-grain*. Esferas pequenas a verde são carbonos, a vermelho são oxigénios, a azul azotos e laranja fósforos. As esferas grandes representam a cadeia apolar do tipo C1 a azul claro, a parte do glicerol N_a a vermelho, o grupo fosfato Q_a a laranja e o grupo colina Q₀ a azul escuro.

Esses diferentes tipos de partículas vão interagir entre si com diferentes forças, visto que uma partícula polar sofre muito mais atração por uma partícula semelhante que por uma partícula apolar. A matriz com os diferentes níveis de interação encontra-se na tabela 1.

Tabela 1 - Matriz dos níveis de interação entre as diferentes partículas, onde *O* ocorre entre partículas com maior atração ($\epsilon=4.6\text{kJ/mol}$) e IX nas partículas com maior repulsão ($\epsilon=2.0\text{kJ/mol}$).

		Q					P					N					C											
sub		da	d	a	0	5	4	3	2	1	da	d	a	0	5	4	3	2	1	da	d	a	0	5	4	3	2	1
Q	da	O	O	O	II	O	O	O	I	I	I	I	I	IV	V	VI	VII	IX	IX	I	I	I	IV	V	VI	VII	IX	IX
	d	O	I	O	II	O	O	O	I	I	I	III	I	IV	V	VI	VII	IX	IX	I	III	I	IV	V	VI	VII	IX	IX
	a	O	O	I	II	O	O	O	I	I	I	I	III	IV	V	VI	VII	IX	IX	I	I	III	IV	V	VI	VII	IX	IX
	0	II	II	II	IV	I	O	I	II	III	III	III	III	IV	V	VI	VII	IX	IX	III	III	III	IV	V	VI	VII	IX	IX
P	5	O	O	O	I	O	O	O	O	O	I	I	I	IV	V	VI	VI	VII	VIII	I	I	I	IV	V	VI	VI	VII	VIII
	4	O	O	O	O	O	I	I	II	II	III	III	III	IV	V	VI	VI	VII	VIII	III	III	III	IV	V	VI	VI	VII	VIII
	3	O	O	O	I	O	I	I	II	II	II	II	II	IV	IV	V	V	VI	VII	II	II	II	IV	IV	V	V	VI	VII
	2	I	I	I	II	O	II	II	II	II	II	II	II	III	IV	IV	V	VI	VII	II	II	II	III	IV	IV	V	VI	VII
	1	I	I	I	III	O	II	II	II	II	II	II	II	III	IV	IV	IV	V	VI	II	II	II	III	IV	IV	IV	V	VI
N	da	I	I	I	III	I	III	II	II	II	II	II	II	IV	IV	V	VI	VI	VI	II	II	II	IV	IV	V	VI	VI	VI
	d	I	III	I	III	I	III	II	II	II	II	III	II	IV	IV	V	VI	VI	VI	II	III	II	IV	IV	V	VI	VI	VI
	a	I	I	III	III	I	III	II	II	II	II	II	III	IV	IV	V	VI	VI	VI	II	II	III	IV	IV	V	VI	VI	VI
	0	IV	IV	IV	IV	IV	IV	IV	III	III	IV	IV	IV	IV	IV	IV	IV	V	VI	IV	IV	IV	IV	IV	IV	IV	V	VI
C	5	V	V	V	V	V	V	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	V	V	IV	IV	IV	IV	IV	IV	IV	V	V
	4	VI	VI	VI	VI	VI	VI	V	IV	IV	V	V	V	IV	IV	IV	IV	V	V	V	V	V	IV	IV	IV	IV	V	V
	3	VII	VII	VII	VII	VI	VI	V	V	IV	VI	VI	VI	IV	IV	IV	IV	IV	IV	VI	VI	VI	IV	IV	IV	IV	IV	IV
	2	IX	IX	IX	IX	VII	VII	VI	VI	V	VI	VI	VI	V	V	V	IV	IV	IV	VI	VI	VI	V	V	V	IV	IV	IV
	1	IX	IX	IX	IX	VIII	VIII	VII	VII	VI	VI	VI	VI	VI	V	V	IV	IV	IV	VI	VI	VI	VI	V	V	IV	IV	IV

Por exemplo, no caso da molécula de DMPC, representada na figura 18, os grupos colina e fosfato são representados, respetivamente, por uma partícula carregada positivamente do tipo Q_0 e uma partícula carregada negativamente do tipo Q_a . Os dois grupos glicerol são representados por duas partículas N_a e os grupos alceno por

partículas C1, onde duas correspondem a cinco carbonos e uma (a ponta da cauda) a quatro carbonos.

Todas as partículas têm um tamanho efetivo (σ) de 0.47 nm, exceto para as partículas que compõem anéis, onde σ é igual a 0.43 nm, sendo compensados com um ϵ (equação do termo LJ da função potencial) reduzido em 25%. A distância de equilíbrio entre duas partículas dos lípidos é igual ao tamanho efetivo, exceto para ligações entre duas partículas N_a onde é 0.37 nm, com constantes de força de $1250 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. O ângulo de equilíbrio entre as partículas é de 180° e constantes de força de $25 \text{ kJ mol}^{-1} \text{ nm}^{-2}$, exceto para o ângulo entre a partícula do grupo fosfato e do glicerol, que é definido como 120° , e para as ligações duplas que têm constante de força 45 kJ mol^{-1} e ângulos de 120° para conformações *cis*, e 180° para conformações *trans* [Marrink 2007].

Em 2008 o modelo MARTINI 2.0 foi estendido para proteínas passando a ser designado por MARTINI versão 2.1 [Monticelli 2008]. Para avaliar quais eram os melhores tipos e subtipos de partículas para as diferentes cadeias laterais, foi feita uma comparação entre o coeficiente de partição em meios água/óleo, calculado experimentalmente, e a razão das densidades nos diferentes meios, calculadas utilizando simulação computacional. O resultado da escolha das diferentes partículas está exemplificado na figura 19.

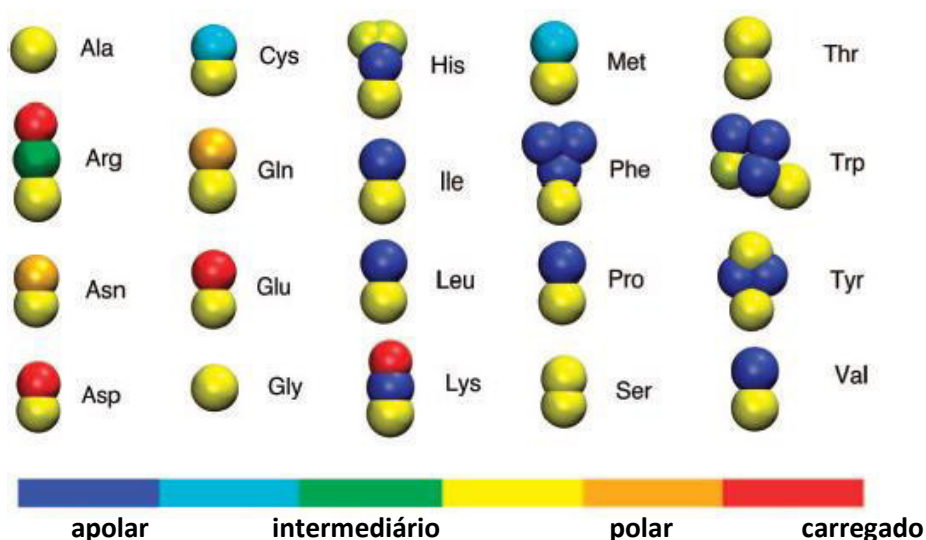


Figura 19 – Representação dos diferentes aminoácidos no MARTINI, em que a cor define o tipo de partícula. Imagem adaptada de [Monticelli2008].

As cadeias laterais dos aminoácidos apolares (Leu, Pro, Ile, Val, Cys e Met), são representadas por partículas do tipo C e as cadeias laterais dos aminoácidos polares (Thr, Ser, Asn e Gln) por partículas do tipo P. Para os aminoácidos carregados (Glu, Asp, Arg e Lys), as cadeias laterais são representadas por partículas do tipo Q, em particular as cadeias laterais das argininas e lisinas são ainda representadas por uma partícula não carregada. No caso dos aminoácidos com cadeias laterais em anel, são representadas por três (His, Phe e Tyr) ou quatro (Trp) partículas especiais para moléculas em anel, SC ou SP (1, 2, 3, e 4 consoante a polaridade). Os aminoácidos mais pequenos (Gly e Ala) não possuem cadeia lateral, sendo apenas representados pela partícula da cadeia principal **[Monticelli 2008]**.

A partícula da cadeia principal depende da estrutura secundária, uma vez que quando em *coil* ou em dobra é fortemente polar (partícula do tipo P5) mas quando em hélice ou folha beta a existência de pontes de hidrogénio diminuem a polaridade, passando a ser representada por uma partícula do tipo N0 (hélice) ou Nda (folha beta). Exceto no início e no fim da hélice, onde se representa por partículas Nd e Na, respetivamente. São casos especiais a glicina, a alanina e a prolina em que, por exemplo, esta última não possui a capacidade de ser dador de hidrogénios para pontes **[Monticelli 2008]**.

No caso especial das moléculas de água, que são o solvente mais habitual das simulações de sistemas biológicos, o conjunto de quatro moléculas atomísticas são modeladas como uma única partícula do tipo P4, designando-se por água MARTINI. Contudo desta forma a temperatura da água MARTINI apresenta uma temperatura de solidificação superior ao normal, necessitando em certos sistemas de ser adicionado cerca de 10% de uma nova partícula, denominada BP4 (*big* P4). A única diferença desta partícula é a sua interação com as partículas P4 da água, onde se aumenta o tamanho efetivo de 0.47nm para 0.57nm para destabilizar o agrupamento uniforme das moléculas de água e a força de interação para o nível O ($\epsilon = 4.6$ kJ/mol), para não haver separação de fases entre as duas partículas de solvente **[Marrink 2007]**. Esta água MARTINI anticongelante será utilizada nos estudos com a hemaglutinina completa.

Os parâmetros de ligação, ou seja, as ligações, os ângulos e os diedros entre as partículas foram parametrizados através da comparação entre aproximadamente 2000 proteínas do “*protein data bank*” (PDB). Através da análise das distribuições dos parâmetros para todas as estruturas secundárias e para todos os tipos de aminoácidos foi estipulado que:

- Os parâmetros de ligação entre a cadeia principal são dependentes da estrutura secundária e independentes do tipo de aminoácidos, exceto para o valor da distância entre partículas da cadeia principal que é sempre 0.35 nm;

- No caso anterior, se as ligações ocorrem entre partículas com estruturas secundárias diferentes, o valor dos parâmetros é igual ao menor valor das duas estruturas secundárias. Por exemplo, o ângulo entre duas partículas em hélice e uma partícula em *coil* é escolhido o valor de ângulo e constante de força das estruturas em *coil*;

- Os diedros são apenas definidos quando quatro partículas seguidas têm estrutura em hélice ou estendida (folha beta);

- Os parâmetros entre a cadeia principal e a cadeia lateral, ou entre cadeias laterais, são independentes da estrutura secundária mas variam consoante o aminoácido;

- Os ângulos entre duas partículas da cadeia principal e uma da cadeia lateral ou uma partícula da cadeia principal e duas da cadeia lateral são independentes tanto da estrutura secundária como do aminoácido [Monticelli 2008].

Escala de tempo no MARTINI

Para uma maior eficiência computacional (um dos pontos principais no desenvolvimento do MARTINI) a massa de todas as partículas deste campo de forças é igual à massa de quatro águas ($m = 72$ amu), exceto para as partículas de estruturas em anel, que é mais reduzida ($m = 45$ amu). Como a dinâmica dos sistemas é afetada pelas massas das partículas, passos de integração com tempos de 30-40fs podem ser utilizados, mas para prevenir que não ocorrem flutuações da energia total é mais seguro usar passos entre 20-30fs, principalmente em simulações de moléculas de anéis, como por exemplo o triptofano [Marrink 2010].

Devido a um menor número de graus de liberdade, comparativamente a simulações atomísticas, e ao tamanho das partículas, a interação entre elas é mais suave, causando uma menor fricção efetiva. Tal provoca uma superfície de energia muito menos rugosa, tornando a amostragem do espaço conformacional 3 a 8 vezes mais rápida. Dessa forma, na interpretação da escala de tempo nas simulações usando o modelo MARTINI deve-se multiplicar o tempo por um fator de 4. A multiplicação por esse fator torna a velocidade de eventos como a difusão da água, a formação de vesículas ou bicamadas semelhante à velocidade das simulações atomísticas. Contudo, o fator de aceleração depende dos sistemas a estudar, sendo sempre preciso interpretar cuidadosamente o tempo das simulações usando o campo de forças MARTINI [Marrink 2003, Monticelli 2008, Voth 2008].

Limitações

Para além da óbvia aproximação que resulta do uso deste campo de forças, devido à resolução reduzida dos sistemas, existem também outros pontos, que poderão ser melhorados em desenvolvimentos futuros do modelo.

- O campo de forças foi parametrizado para a fase condensada/líquida, portanto simulações com sistemas em estado gasoso ou em estado sólido não serão recomendadas.

- O campo de forças MARTINI foi parametrizado considerando as energias livres de partição, mas devido à resolução reduzida a entropia é maior, sendo por sua vez, a entalpia menor para compensar. Isso altera a dependência pela temperatura das simulações, devendo-se usar temperaturas próximas da temperatura de parametrização (270-330K).

- O cálculo de interações eletrostáticas apenas entre átomos a uma distância de 1.2 nm entre si (equivalente a entre duas a três partículas) não refletindo com precisão as interações eletrostáticas de longo alcance. Isto, aliado ao *screening* implícito torna as interações de moléculas polares ou carregadas pouco corretas, principalmente em solventes não polarizáveis, como a água MARTINI normal (P4).

- A estrutura secundária em péptidos ou proteínas é distinguida pela adição e fixação dos diedros entre as partículas, logo não se podem esperar alterações conformacionais usando o campo de forças MARTINI [Marrink 2007, Voth 2008].

2.2.3. Campo de forças ELNEDIN

Nas simulações de proteínas usando o campo de forças MARTINI ocorrem na maioria das vezes alterações conformacionais da estrutura terciária que pouco provavelmente ocorreriam experimentalmente. Para ultrapassar tal problema, Marrink e seus colaboradores desenvolveram um novo campo de forças, designado por ELNEDIN [Periole 2009]. Mais especificamente, este campo de forças consiste na junção do MARTINI 2.1, usado para descrever as interações inter e intramoleculares, com a aplicação de uma rede de ligações elásticas que restringe a estrutura da molécula à sua conformação inicial [Periole 2009].

A rede de ligações elásticas é criada ao aplicar ligações extra entre as partículas da cadeia principal que se encontram na estrutura inicial a uma distância inferior a um determinado *cut-off*, e separados pelo menos por duas ligações. A distância entre partículas e a constante de força das ligações é estipulado pelo utilizador, sendo que, os valores 0.8 a 1.0nm para a distância (R_C , que define a extensão da rede) e 500-1000 kJ mol⁻¹ nm⁻² para as constantes de força (K_{spring} , que define a rigidez da rede), reproduzem melhor os valores das simulações atomísticas. Contudo o valor ideal varia de proteína para proteína [Periole 2009], sendo necessário investigar que conjunto de valores é o mais indicado para a Hemaglutinina simulada neste trabalho.

A aplicação deste campo de forças numa proteína fará com que a sua estrutura inicial seja mantida ao longo do tempo, tornando impossível estudos sobre alterações conformacionais como, por exemplo, estudos de *folding*. Contudo, o ELNEDIN é capaz de reproduzir deformações globais (alterações da estrutura terciária) de uma proteína, a flutuação dos resíduos e movimentos coletivos. Um exemplo de aplicações para este campo de forças são as interações proteína-proteína, que foram testadas no artigo de Periole et al. [Periole 2009], e conclui-se que o reconhecimento molecular neste tipo de

interações é dependente da flexibilidade (nestes casos não se deve usar redes muito rígidas), sendo consistente com outros artigos [Yamniuk 2004, Basdevant 2006].

Note-se que este campo de forças não é meramente o campo de forças MARTINI com a adição de simples redes elásticas. Com a adição das redes elásticas foi preciso efetuar alterações no campo de forças. A partícula da cadeia principal de um resíduo no ELNEDIN é colocada na posição do carbono alfa, de modo à distância ser sempre de 0.38 nm, removendo assim, a dependência da distância entre partículas da cadeia principal e a estrutura secundária, que existia no MARTINI. Assim sendo, os parâmetros de ligação foram também re-parametrizados e o valor de referência da distância de equilíbrio e a constante de força foram alterados para 0.38 nm e $150000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$, respetivamente para as ligações entre duas partículas da cadeia principal e 120° e $40 \text{ kJ mol}^{-1} \text{ nm}^{-2}$, respetivamente para o ângulo entre três partículas da cadeia principal. A estrutura dos anéis das cadeias laterais também foi alterada, tanto a nível das ligações na fenilalanina e na tirosina, como a implementação de assimetria nas cadeias laterais da histidina e do triptofano [Periole 2009].

2.3. Métodos

Neste capítulo serão explicadas passo a passo as metodologias e protocolos da simulação computacional dos diferentes sistemas em estudo.

2.3.1. Preparação / Solvatação

Num capítulo anterior sobre solvatação e condições periódicas de fronteira foi referido que o sistema em estudo era colocado no centro da caixa de simulação, mas nem sempre isso ocorre. Neste trabalho foram utilizados diferentes configurações iniciais, variando consoante o sistema simulado e o que se pretende estudar.

Foram utilizados dois tipos diferentes de caixas de simulação: quadradas/retangulares ou octaédricas truncadas. As primeiras foram utilizadas nas simulações com lípidos, onde a bicamada é formada nos eixos x e y que possuem tanto o mesmo comprimento como o mesmo valor de acoplamento de pressão (τ_p). A bicamada fica portanto perpendicular ao eixo dos z, que possui um τ_p diferente dos

anteriores, uma vez que se utiliza um acoplamento de pressão semi-isotrópico. O segundo tipo de caixas é utilizado para proteínas/péptidos flexíveis ou proteínas globulares em água, onde a distância da proteína aos cantos de uma caixa quadrada é superior à distância nas direções verticais ou horizontais da proteína à caixa, aumentando o número de águas. Ao utilizar uma caixa mais parecida com a forma da proteína diminuimos o número de águas sem alterar a distância mínima às imagens periódicas.

Normalmente as estruturas iniciais provêm de estruturas obtidas por cristalografia de raios-x, como a estrutura da hemaglutinina H3N2 (PDB code: 2VIU) utilizada neste trabalho, ou obtidos por NMR, como a estrutura de 24 aminoácidos do péptido de fusão (PDB code: 2KXA). Mas se não existir nenhuma estrutura pelos métodos descritos atrás, a solução é gerar uma estrutura usando técnicas de modelação, como por exemplo o péptido transmembranar que foi criado usando o programa PyMOL [PyMOL].

Após a colocação das moléculas, que se querem estudar, na caixa de simulação o passo seguinte é a sua solvatação com moléculas de água, colocadas ao acaso até encher a caixa de simulação. Nas simulações com lípidos foi utilizado uma caixa onde o número de águas necessárias para encher fossem aproximadamente 2000 águas MARTINI (8000 águas reais) e adicionaram-se rigorosamente 2000 águas em todas as caixas. O controlo do número de águas adicionado foi realizado para manter a razão entre lípidos e águas, ou seja, a hidratação da membrana ser constante em todas as simulações.

O modelo da molécula de água utilizado neste trabalho foi o *simple point charge* (SPC) [Berendsen 1981] para o campo de forças 54A7 e, consoante a necessidade, águas MARTINI “normais” (P4) ou águas MARTINI anticongelantes (BP4) para o campo de forças MARTINI e ELNEDIN.

A seguir à solvatação do sistema, e antes de se iniciar as simulações de dinâmica molecular é necessário fazer uma minimização de energia, ou seja, encontrar uma estrutura com menor energia comparativamente à estrutura inicial, através de uma leve alteração nas posições dos átomos. Tal é necessário uma vez que a estrutura inicial sem

minimizações pode conter forças excessivamente grandes e a simulação falhar [Saxena 2009, Van der Spoel 2005].

2.3.2. Minimização

A superfície de energia de sistemas biológicos relaciona a posição dos átomos (conformação de uma molécula) com a energia potencial correspondente. Devido à existência de inúmeros graus de liberdade e por consequência inúmeras conformações, a superfície de energia é bastante complexa, estando repleta de mínimos locais. Tornando difícil, desta forma, encontrar o mínimo global, mas para o efeito da minimização basta um mínimo local.

Uma forma de identificar esses mínimos é através das derivadas de primeira e segunda ordem dessa mesma superfície de energia. A primeira derivada indica se nos encontramos num vale (mínimo) ou num pico (máximo) da superfície, sem os diferenciar. Por isso usa-se a segunda derivada que distingue mínimos da função, onde se obtém valores positivos, em oposição dos máximos da função onde se obtém valores negativos [Leach 2001].

O método de “descida íngreme” (*Steepest-descent*) é um método derivativo de primeira ordem que num primeiro passo calcula o gradiente da energia potencial, o qual é de seguida utilizado para calcular as forças (igual ao gradiente negativo). O próximo passo será calcular as novas posições dos átomos (r_{n+1}), segundo a seguinte equação:

$$r_{n+1} = r_n + \frac{F_n}{\max(|F_n|)} h_n$$

Onde h_n é o deslocamento máximo, F_n é a força (ou o negativo do gradiente da energia potencial V) e $\max(|F_n|)$ é o valor absoluto máximo entre todos os componentes da força.

A energia potencial é então calculada para estas novas posições, em que:

Se $V_{n+1} < V_n$ as novas posições são aceites e o $h_{n+1} = 1.2h_n$

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

Estes passos são repetidos em ciclos até que um número de ciclos definido pelo utilizador seja alcançado ou então o valor $\max(|F_n|)$ seja menor do que um valor estipulado pelo utilizador. Em suma este método encontra mínimos sucessivos da função de energia potencial, seguindo a direção do gradiente negativo dessa mesma função. Neste trabalho utilizaram-se como limite os 2000 ciclos ou força de $10 \text{ kJ mol}^{-1} \text{ nm}^{-2}$.

Este método não é o mais eficiente na procura de mínimos globais, podendo ficar preso num mínimo local, não podendo “pular” para outros mínimos locais que tenham uma menor energia. Contudo, é um método bastante robusto, fácil de implementar e converge rapidamente para o mínimo local mais próximo [Van der Spoel 2005].

2.3.3. Inicialização

Na minimização a temperatura não é contabilizada, mas as simulações de sistemas biológicos devem ser realizadas por volta dos 300-310 K, dependendo do sistema em causa. Neste trabalho as simulações foram todas realizadas a 303 K. Este valor foi utilizado, devido ao uso de membranas de DMPC, o qual tem uma temperatura de transição de fase de 298 K, pelo que se deve usar uma temperatura superior a este último valor. Algo semelhante ocorre para a pressão, onde se usou sempre 1 bar, equivalente à pressão atmosférica.

De modo a ajustar o sistema a estes valores de temperatura e de pressão, os “banhos” são ligados e as constantes de acoplamento são lentamente ajustadas ao valor final que se pretende. De uma maneira geral é realizado um primeiro passo onde apenas é ligado o banho de temperatura, onde são geradas aleatoriamente as velocidades iniciais dos átomos para uma determinada temperatura. No segundo passo o banho de pressão é ligado e os restantes passos servem para ajustar o sistema ao acoplamento de temperatura e pressão corretos, que serão utilizados nas corridas de produção.

Tanto na minimização, como nos passos de inicialização é utilizado um termo externo ao potencial, designado por restrição das posições, e é aplicado aos átomos que se deseja imobilizar, através da adição da seguinte equação à função de energia potencial:

$$V_{rp}(r_i) = \frac{1}{2} K_{rp} \|r_i - R_i\|^2$$

Onde, R_i é a posição de referência do átomo i e K_{rp} é a constante de força definida pelo utilizador, que no caso deste trabalho foi sempre de $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$, à exceção dos dois últimos passos de inicialização com lípidos, durante os quais o valor de $100 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ foi usado. Este método é utilizado para evitar a destruição da proteína, deixando primeiro o solvente relaxar à volta de uma proteína imóvel. Inicializações suaves têm impacto na estabilidade das simulações, portanto nos primeiros passos restringe-se todos os átomos da proteína ou dos lípidos, mas nos últimos passos apenas são restringidos os carbonos alfa da cadeia principal (no caso do MARTINI, as partículas da cadeia principal) e os lípidos, mas estes últimos com menor energia.

Ao realizar apenas uma simulação, por exemplo de uma proteína, teríamos o problema desta tomar uma conformação correspondente a um mínimo local e não conseguir sair desse mínimo durante o resto da simulação. Assim sendo, deve-se realizar várias simulações replicadas para se obter uma maior amostragem do espaço conformacional e assim ter uma maior certeza estatística das propriedades após análises das trajetórias.

Visto que a velocidade inicial para cada átomo é atribuída aleatoriamente, obedecendo a uma distribuição de Boltzmann, é necessário apenas mudar a *seed* do gerador de números aleatórios para se obter um conjunto diferente de velocidades atómicas, originando uma simulação diferente.

2.3.4. Corrida de produção

Após a inicialização estar terminada a simulação é mantida igual ao último passo de inicialização mas sem restrições na posição dos átomos. Durante esta fase do processo, existe ainda uma subfase que é a equilibração. Uma proteína diz-se em

equilíbrio quando alcança estabilidade e não apresenta flutuações irregulares, podendo ser averiguado através de análises estruturais, como a “root mean square deviation” (RMSD).

Nesta fase, a simulação continua até alcançar um tempo definido pelo utilizador (podendo-se sempre simular por mais tempo caso o utilizador encontre necessidade). As análises são realizadas na parte equilibrada das trajetórias.

2.4. Transformação reversa

O método de transformação reversa será utilizado para transformar a estrutura de um sistema em resolução MARTINI para uma estrutura *united-atom*. Este método é apenas utilizado neste sentido, uma vez que a estrutura em MARTINI possui menos informação que uma estrutura atomística, esta última pode ser facilmente transformada numa estrutura MARTINI. Este processo é realizado por um algoritmo de reconstrução com três passos:

1º Passo – Colocação inicial das partículas atómicas

Todos os átomos que são representados por cada partícula MARTINI são colocados, ao acaso, numa esfera (0.3 nm de raio) com o seu centro no centro de massa da partícula correspondente. Mesmo sendo necessário um mapeamento que indique que conjunto de átomos correspondem a cada partícula MARTINI, este passo acelera o processo, ao colocar os átomos relativamente próximos da sua posição final.

Existe, no entanto, um problema com este passo. A colocação, ao acaso, dos vários átomos pode criar, no início da simulação, forças elevadas que provocam instabilidades. De modo a ultrapassar este problema, todas as forças superiores a 15 000 kJ mol⁻¹ nm⁻² ($F_{cap,0}$) são reduzidas para este último valor. Ao longo da simulação ocorre um aumento linear da força limite por um fator At , onde A é por norma 100kJ nmol⁻¹ nm⁻² ps⁻¹ e t o tempo em picossegundos.

2º Passo – Simulated annealing restrito

O *simulated annealing* (SA) é um método que utiliza cálculos de dinâmica molecular, no qual ocorre um aumento da temperatura do sistema até um valor

elevado. As moléculas a elevadas temperaturas tem uma maior facilidade em passar as barreiras energéticas, permitindo assim uma maior facilidade na exploração do espaço conformacional. De seguida a temperatura do sistema é reduzida gradualmente até ao valor normal, ocorrendo variações mais seletivas à medida que se chega a um mínimo de energia [Kirkpatrick 1983 & Cernv 1985].

De forma a obter uma estrutura final de baixa energia mas sem se desviar muito da estrutura CG inicial, foi utilizado um SA com um potencial de restrição. A introdução do potencial U^{tot} mantém o centro de massa dos átomos perto do centro de massa da partícula MARTINI correspondente:

$$U^{tot} = U^{UA} + U^{restr}$$

Onde U^{UA} representa o campo de forças UA e U^{restr} é o potencial harmónico calculado a cada passo de integração,

$$U^{restr} = \sum_{i=1}^n \frac{k}{2} (r_i^{CG} - r_i^{UA,com})^2$$

Onde r_i^{CG} é a posição de referência da partícula MARTINI i , $r_i^{UA,com}$ é o centro de massa dos átomos que foram mapeados pela partícula i , n é o número total de partículas e k a constante de força de restrição.

3º Passo – Libertação gradual das restrições

No final, quando a temperatura se encontra no valor pretendido, a nova estrutura encontra-se demasiado ligada à estrutura CG. No entanto, as estruturas de energia mínima para o UA e para o CG são ligeiramente diferentes, devido a diferenças entre os campos de forças. Uma forma de ultrapassar isto, é aliviar a restrição anterior à nova estrutura. Portanto o U^{restr} é removido, suavemente, ao longo dos últimos 5000 passos do SA, de forma a obter uma configuração de menor energia.

No caso especial da água MARTINI cada partícula representa quatro moléculas de água. De forma a impedir que se dispersem da partícula respetiva, outro potencial é adicionado a U^{tot} :

$$U_j^{restr,W} = \begin{cases} 0 & r_{ij} \leq r_{CGW} \\ \frac{k_w}{2}(r_{ij} - r_{CGW})^2 & r_{ij} > r_{CGW} \end{cases}$$

Onde r_{ij} é a distância entre um átomo de oxigénio e o centro da partícula MARTINI i (que corresponde a quatro moléculas de água e por sua vez quatro oxigénios), r_{CGW} é o raio limite (*cut-off*) e k_w é a constante de força de restrição.

Desta forma se uma das moléculas de água ultrapassar o raio limite é aplicada uma força que puxa a molécula de volta. Finalmente para assegurar que a força externa, aplicada a cada molécula de água, não contribui para o movimento do centro de massa, uma contra força é distribuída por todas as moléculas de água restantes, para que $\sum_{j=1}^4 F_j = 0$ [Rzepiela 2010].

Outro fator a ter em conta é os diedros, uma vez que se existir uma barreira energética muito elevada entre dois estados com diferentes diedros, o resultado da SA pode ser um mínimo local não desejado. Uma solução é a implementação de uma restrição do ângulo diedro, definido pelo próprio ângulo e uma constante de força. Por exemplo, para as ligações peptídicas de péptidos e proteínas, usadas neste trabalho, foi imposto uma restrição à conformação *trans*, ou seja um diedro com um ângulo de 180°.

Rzepiela et al. recomendam um SA de 60-80 ps seguido de 10-20 ps de equilíbrio à temperatura final, ainda sobre o efeito do potencial de restrição. Por fim depois de remover esse potencial deve-se simular por um período de 10 ps, equivalente aos já referidos 5000 passos. Portanto, para todas as transformações reversas foi simulado 100 ps de SA com temperaturas a variar de 1300 K para 303 K para proteínas e lípidos, e 400 K para 303 K para as águas. Foram ainda simulados 20 ps à temperatura final, onde os 10 ps finais tiveram o potencial de restrição gradualmente diminuído.

2.5. Técnica de “crescimento lento”

Como foi mencionado na introdução, o péptido dobradiça, com a descida de pH, passa de uma estrutura desordenada para uma estrutura em hélice. Para averiguar as possíveis alterações conformacionais, que daí poderiam surgir na proteína, um possível estudo é obrigar a parte desordenada do péptido a adotar uma estrutura em hélice. A

técnica utilizada para tal foi o método de “crescimento lento” (*slow-growth*) que consiste, resumidamente, na conversão de um estado A de um sistema noutro estado B do mesmo sistema, de uma forma lenta e contínua.

Note-se que o método de crescimento lento foi primeiro desenvolvido para o cálculo de energia livre, por exemplo, energia livre de hidratação, através do desaparecimento lento de uma molécula solvatada [Kollman 1993]. No entanto, para este trabalho o interesse não é calcular energias livres, mas sim a evolução lenta da estrutura desordenada (estado A) do péptido dobradiça para uma estrutura em hélice (estado B). Isto é realizado sobretudo a nível dos diedros, onde a constante de força K_d é 0 no estado A, pois estruturas desordenadas não são definidos por diedros no campo de forças MARTINI, e é lentamente aumentado para $1250 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ no estado B. Sendo o valor da constante de força ao longo da simulação igual ao resultado da seguinte equação:

$$K_d(\lambda) = (1 - \lambda)K_d(\text{estado A}) + \lambda K_d(\text{estado B})$$

Onde λ é o parâmetro de acoplamento que vai desde o valor 0, descrevendo o estado A, até 1, descrevendo o estado B.

As simulações, para o péptido dobradiça em água, foram corridas durante 200 ns com um passo de integração de 0.01 ps, onde λ aumentou, por cada passo de integração, $0,5 \times 10^{-7}$ (1/20.000.000 passos) de 0 até 1. Para a hemaglutinina (estrutura cristalográfica) em água foram realizadas corridas de 50 ns com passos de integração de 0,005 ps em que λ aumentou $0,1 \times 10^{-6}$ a cada passo. No estudo do péptido dobradiça e da hemaglutinina foram utilizados, respetivamente, os campos de força MARTINI e ELNEDIN, com os respetivos protocolos normais de simulação em água.

3. Tratamento dos Resultados

3.1. Péptido de fusão

A estrutura do péptido de fusão do subtipo H3 com 20 resíduos de aminoácidos obtido por Han et al. [Han 2001] já foi amplamente estudada através de simulação computacional, tanto com *united-atom* [Huang 2004, Légaré 2012] como com *coarse-grain* [Fuhrmans 2009, Fuhrmans 2012]. Em contraste a estrutura do péptido de fusão do subtipo H1 com 24 aminoácidos, obtido recentemente por Lorieau et al. [Lorieau 2010], ainda não foi praticamente estudada, o que torna o seu estudo apelativo.

Uma vez que a nova estrutura foi obtida por NMR, não existe apenas uma conformação, mas sim 10 conformações, representadas na figura 20.

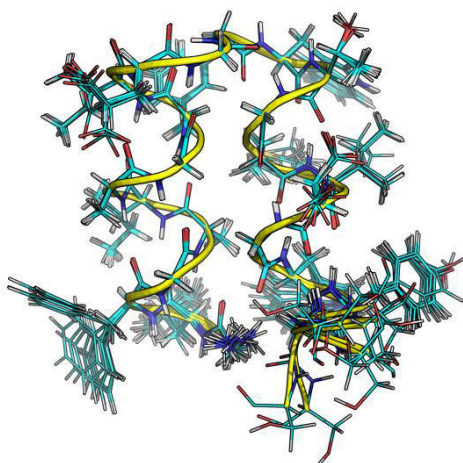


Figura 20 – Representação das 10 conformações obtidas por NMR, onde a amarelo está representado o estilo *ribbon* para salientar a estrutura em hélice e as variações no C-terminal.

Através da sobreposição das diferentes conformações é fácil constatar que com exceção de algumas cadeias laterais, que não são tão relevantes no decorrer da simulação, apenas o último resíduo (Ser24) sofre flutuações. É intuitivo que estando num terminal do péptido, a estabilidade deste resíduo fosse menor que os restantes. Principalmente por não fazer parte da sequência da hemaglutinina, mas sim da cauda hidrofílica adicionada ao péptido (altamente hidrofóbico), impedindo a sua agregação [Lorieau 2010]. Posto isto decidiu-se utilizar a primeira conformação (figura 9, página

26) para a simulação computacional, visto que sendo as estruturas praticamente idênticas, a escolha não influencia significativamente.

3.1.1. Péptido de fusão em água usando o campo de forças 54A7

Com a finalidade de realizar uns estudos preliminares do péptido de fusão em água criaram-se quatro sistemas (tabela 2): i) péptido completo com 24 resíduos, em que os resíduos tituláveis, glutamato na posição 15 (Glu15) e aspartato na posição 19 (Asp19), encontram-se protonados (P_24); ii) péptido completo com 24 resíduos mas com os resíduos tituláveis desprotonados (D_24); iii) péptido com 23 resíduos, onde a Ser24 foi eliminada da estrutura e os resíduos tituláveis estão protonados (P_23); iv) péptido com 20 resíduos, onde os três últimos (Trp21, Tyr22, Gly23 e Ser24) foram eliminados e os resíduos tituláveis estão protonados (P_20). Todos os péptidos têm os terminais protonados, estando o N-terminal carregado e o C-terminal neutro.

Tabela 2 – Sumário das simulações para o péptido de fusão em água, usando o 54A7

Sistema	Número de aminoácidos	Estrutura	Estado de protonação de Glu15 e Asp19	Estado de protonação dos terminais
P_24	24	1º estrutura do NMR	Protonado	Protonado
D_24			Desprotonado	
P_23	23	1º estrutura sem Ser24	Protonado	
P_20	20	1º estrutura sem Trp21, Tyr22, Gly23 e Ser24		

Como referido acima, neste péptido existem dois resíduos tituláveis, Glu15 e Asp19, que possuem um pK_a em água de 4.07 e 3.86, respetivamente. A pH de fusão, ou seja a pH 4-5, estes resíduos estariam maioritariamente desprotonados. Contudo neste sistema os aminoácidos encontram-se inseridos num péptido e diferentes ambientes, alterando, de forma não previsível, o pK_a destes resíduos. Assim sendo assumimos para esta dissertação que o sistema P-24 reproduz o péptido de fusão a pH baixo, idealmente a pH de fusão e o sistema D_24 reproduz o péptido a pH neutro.

O pK_a do N-terminal em péptidos é aproximadamente 8, portanto tanto a pH neutro como a pH de fusão vai estar protonado (no caso do primeiro, possivelmente não na totalidade). Tanto para os sistemas a pH baixo como a pH alto, o C-terminal ficou protonado para não introduzir uma carga ao sistema que não existiria no sistema completo (hemaglutinina), mesmo que o pK_a do C-terminal de péptidos ronde os 3.

Outra questão que se colocou foi a influência da Ser24, que não faz parte da sequência de aminoácidos da hemaglutinina do subtipo H1, na estrutura do péptido, obtendo-se o sistema P-23. Um raciocínio semelhante foi usado na construção do sistema P-20, com apenas os primeiros 20 resíduos, de forma a ter o mesmo número de resíduos que a anterior estrutura obtida por Han et al. [Han 2001] e assim averiguar se a diferença entre as estruturas se deve ao diferente número de aminoácidos. A estrutura dos diferentes sistemas está representada na figura 21.

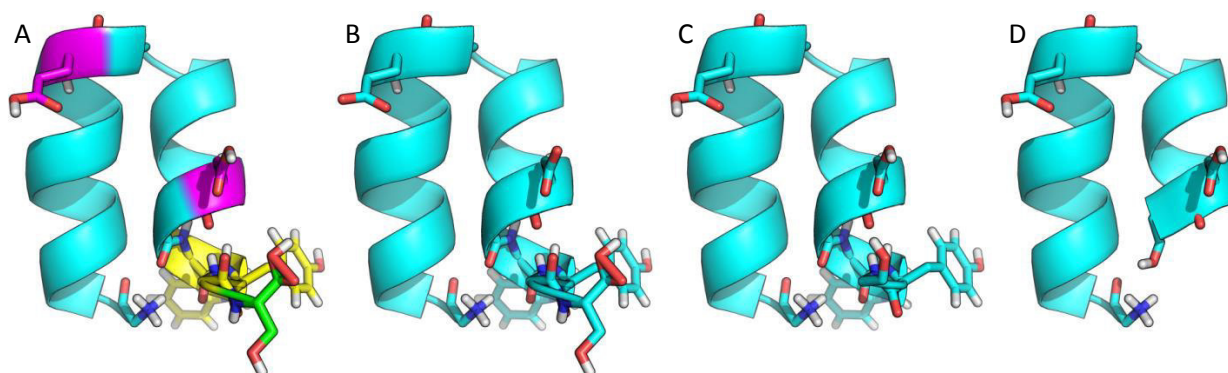


Figura 21 – Estrutura inicial dos quatro diferentes sistemas (A) P-24, (B) D-24, (C) P-23 e (D) P-20. Estão representados como *cartoon* em ciano a estrutura secundária e as cadeias laterais dos resíduos mais importantes (os tituláveis e terminais) em *sticks* com os oxigénios representados a vermelho e os azotos a azul-escuro. No caso A estão ainda representados a magenta os resíduos tituláveis, amarelo os resíduos 21-23, e a verde a serina terminal.

Todos os sistemas foram simulados, em triplicado, em água, numa caixa de simulação octaédrica durante 500ns. Para seguir as alterações conformacionais, foi analisada a estrutura secundária do péptido, ao longo do tempo de simulação. Para tal utilizou-se o algoritmo “Define Secondary Structure of Proteins” (DSSP) que resumidamente utiliza as coordenadas da proteína para identificar as pontes de hidrogénio da cadeia principal e os ângulos entre os carbonos alfa. As variações da estrutura secundária para todas as simulações encontram-se nas figuras 22-25.

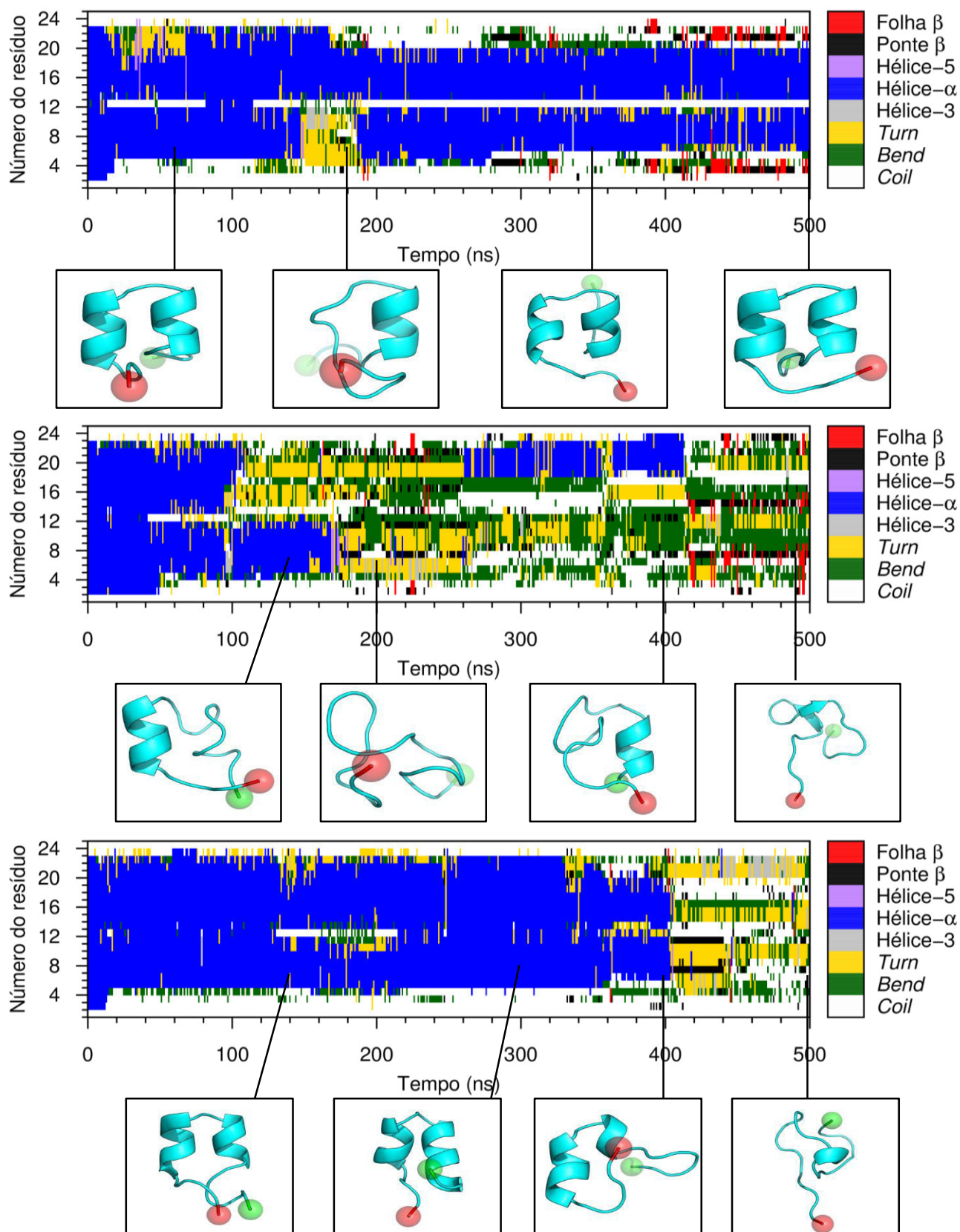


Figura 22 – Resultado da análise de DSSP para o sistema P_24, onde foi atribuído a cada nanossegundo e a cada resíduo uma cor correspondente à sua estrutura secundária. Em pormenor encontram-se as conformações a um dado tempo, onde o N-terminal e o C-terminal estão representados por esferas vermelhas e verdes, respetivamente.

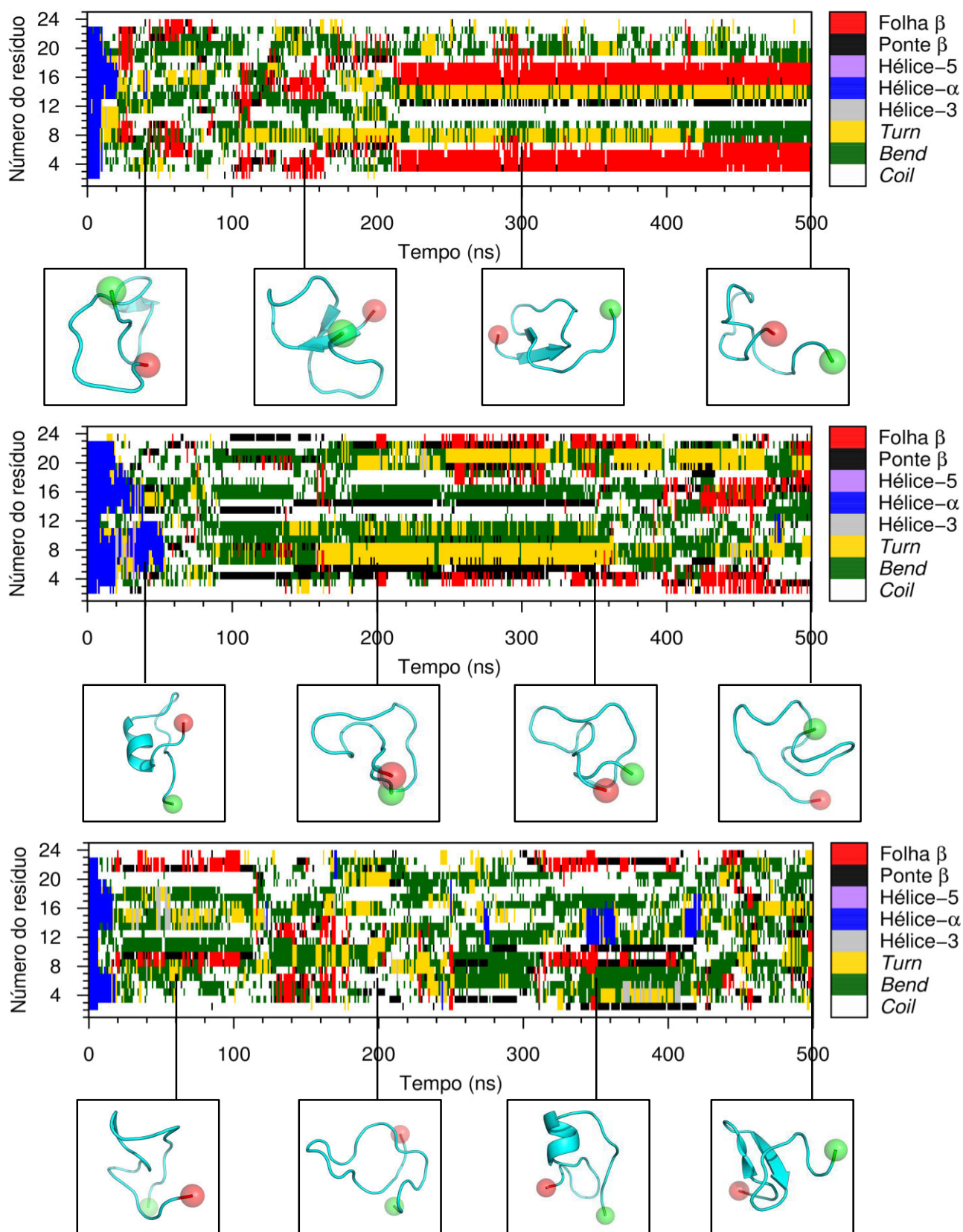


Figura 23 – Resultado da análise de DSSP para o sistema D_24, onde foi atribuído a cada nanossegundo e a cada resíduo uma cor correspondente à sua estrutura secundária. Em pormenor encontram-se as conformações a um dado tempo, onde o N-terminal e o C-terminal estão representados por esferas vermelhas e verdes, respetivamente.

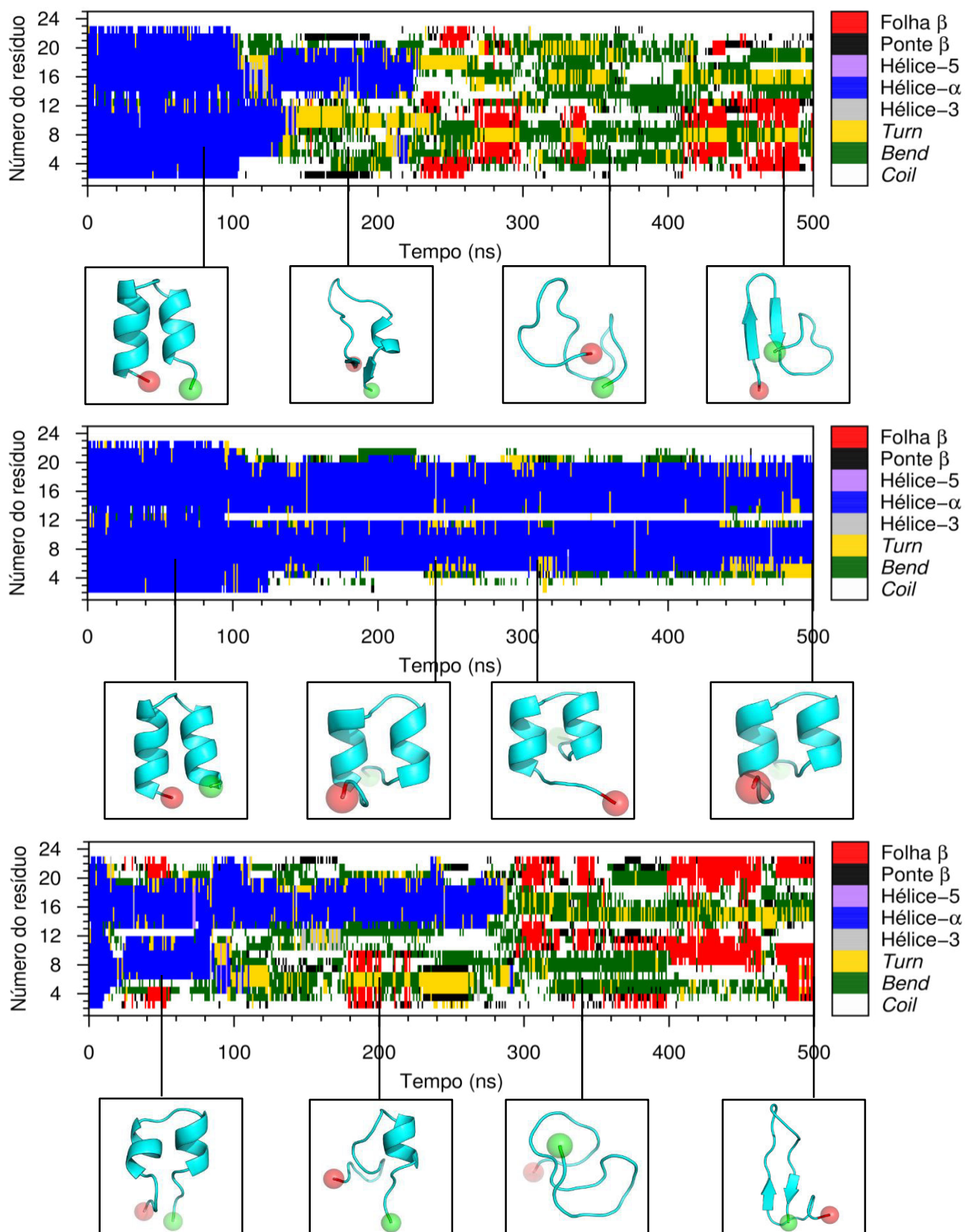


Figura 24 – Resultado da análise de DSSP para o sistema P_23, onde foi atribuído a cada nanossegundo e a cada resíduo uma cor correspondente à sua estrutura secundária. Em pormenor encontram-se as conformações a um dado tempo, onde o N-terminal e o C-terminal estão representados por esferas vermelhas e verdes, respetivamente.

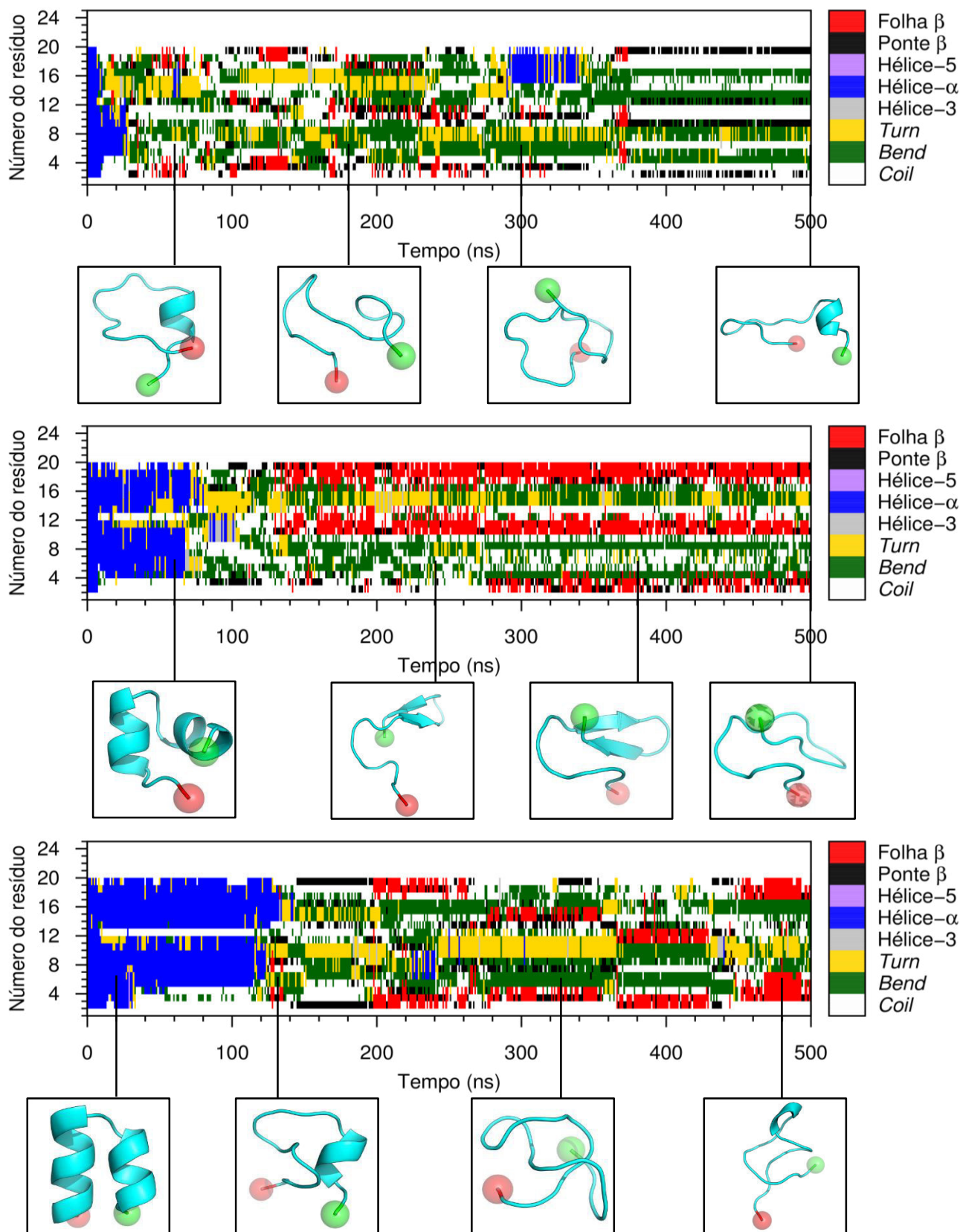


Figura 25 – Resultado da análise de DSSP para o sistema P_20, onde foi atribuído a cada nanossegundo e a cada resíduo uma cor correspondente à sua estrutura secundária. Em pormenor encontram-se as conformações a um dado tempo, onde o N-terminal e o C-terminal estão representados por esferas vermelhas e verdes, respetivamente.

Observa-se que em água a estrutura em *hairpin* é perdida, no intervalo de tempo simulado, em todas as simulações exceto numa réplica do sistema P_24 e numa réplica do sistema P_23. Para poder comparar a estabilidade da estrutura em hélice entre os diferentes sistemas, calculou-se o número de resíduos que se encontravam em hélice, ao longo do tempo de simulação (figura 26).

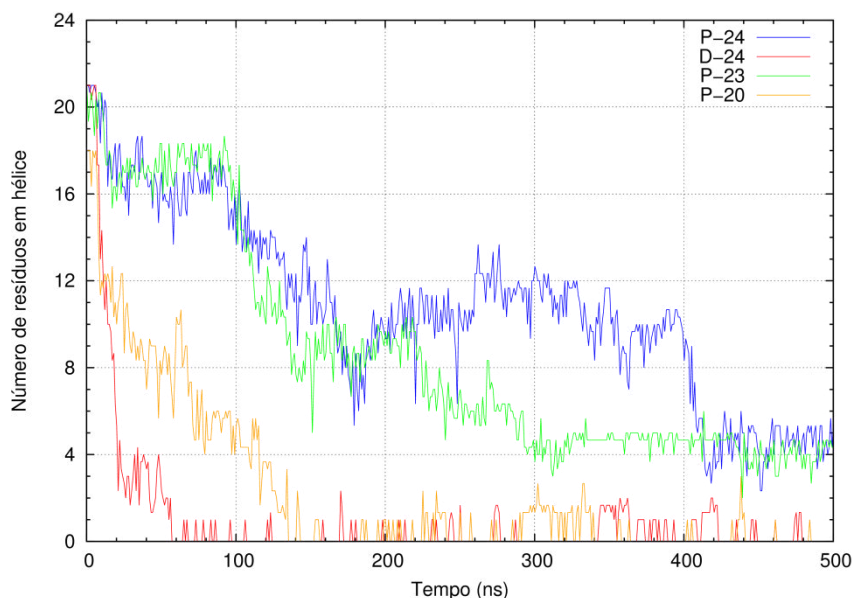


Figura 26 – Variação do número médio (das 3 réplicas) de resíduos em hélice do péptido de fusão, em cada um dos sistemas estudados, ao longo do tempo de simulação.

Como esperado, observa-se uma diminuição do número de resíduos em hélice, principalmente nos sistemas D_24 e P_20 que rapidamente perdem o conteúdo em hélice até praticamente não apresentar qualquer resíduo nessa estrutura secundária. Embora o estado de protonação das cadeias laterais dos resíduos tituláveis não alterem a estrutura do péptido de fusão em membrana [Lorieau 2010], parece que em água existe uma maior destabilização quando desprotonado. Isto sugere, hipoteticamente, que as alterações conformacionais aquando do abaixamento de pH podem também estar relacionadas com o aumento da preferência para a estrutura em hélice pelo péptido. O péptido com apenas 20 resíduos é logicamente mais instável que o péptido com 24 resíduos, uma vez que os resíduos do N-terminal do *hairpin* não vão interagir com os resíduos do C-terminal do *hairpin*, o qual estabilizava a estrutura através da formação de pontes de hidrogénio.

Estes resultados vão de encontro com os resultados de dicroísmo circular obtidos por Han et al., os quais mostram que o PF de 20 resíduos não tem estrutura em solução aquosa [Han 2001]. Os resultados e conclusões retirados das análises dos sistemas P_24 e D_24 encontram-se publicadas num artigo de Victor et al. [Victor 2015].

No caso dos sistemas P_24 e P_23 a perda de resíduos em hélice é muito semelhante entre si e além de ser mais lenta que as anteriores, estes sistemas parecem estabilizar num patamar de 4-5 resíduos em hélice. Isto deve-se à existência de uma réplica em cada um desses dois sistemas onde o péptido apresenta uma estrutura de *hairpin* estável, dentro do tempo simulado (figura 27).

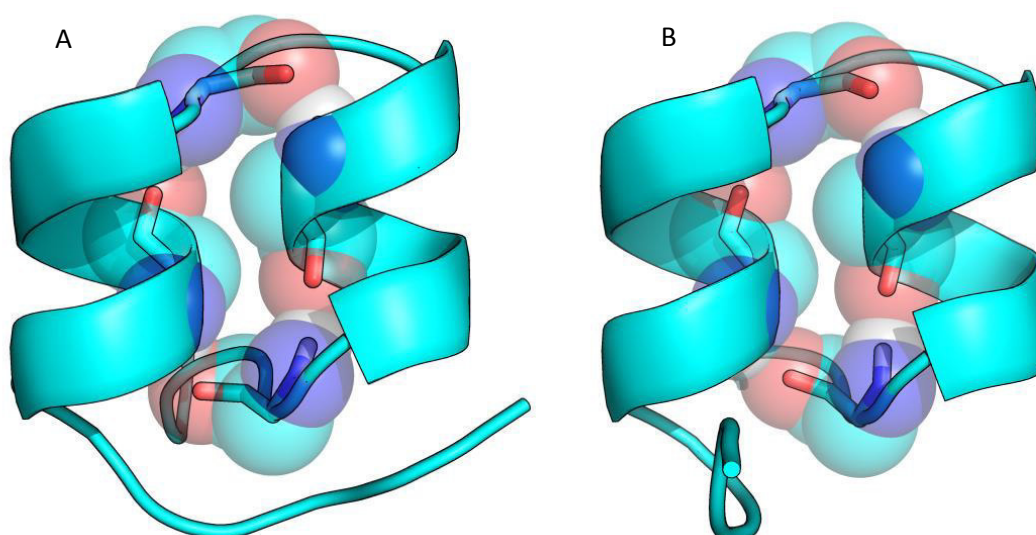


Figura 27 – Representação em *cartoon* da primeira réplica do sistema P_24, à esquerda, e da segunda réplica do sistema P_23, ambas aos 500ns, onde as glicinas Gly8, Gly12, Gly16 e Gly20 estão representadas como *sticks* e esferas.

Uma das características deste péptido é a elevada percentagem em glicinas, distanciadas por três resíduos, formando motivos GXXXG, comuns em interações hélice-hélice devido à possibilidade de um melhor empacotamento. Nas duas réplicas que mantiveram uma estrutura em *hairpin* ao longo da simulação (figura 27), pode-se visualizar as interações GXXXG que delimitam um conjunto de hélices mais curtas que no início da simulação.

O nosso objetivo principal é realizar o *self-assembly* do péptido de fusão numa membrana, mas utilizando o campo de forças 54A7 seria um processo com um custo

computacional demasiado elevado. O campo de forças MARTINI surge então como uma solução, pois já mostrou anteriormente bons resultados em *self-assembly* de lípidos [Marrink 2007] e mesmo *self-assembly* com uma proteína no sistema [Fuhrmans 2009, Fuhrmans 2012].

O próximo passo será portanto transformar a resolução atômica do péptido de fusão (como é obtido no NMR) na resolução *coarse-grain* do MARTINI e fazer simulações preliminares do péptido em água para observar o seu comportamento.

3.1.2. Péptido de fusão em água usando MARTINI

Para estudar o comportamento do péptido de fusão (sistema P-24) em água utilizando o campo de forças MARTINI primeiro é preciso transformar as coordenadas atômicas do péptido numa representação *coarse-grain* (figura 28). A transformação é realizada através de um programa em python disponibilizado na página do MARTINI (<http://md.chem.rug.nl/cgmartini/>).

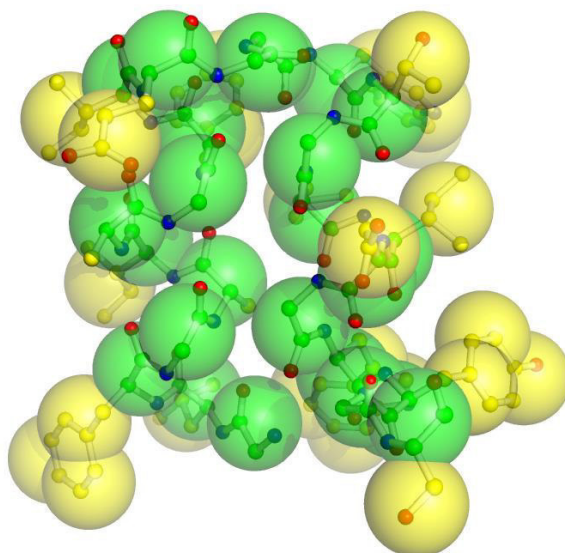


Figura 28 – Representação da estrutura de NMR do péptido em *balls and sticks*. A verde estão representados os carbonos da cadeia principal, a amarelo os carbonos da cadeia lateral, a azul os azotos e a vermelho os oxigénios. Na mesma figura também podemos ver a representação no MARTINI. As esferas verdes representam a cadeia principal e as esferas amarelas representam a cadeia lateral.

Utilizando os parâmetros padrão do campo de forças MARTINI ocorreram, ao longo do tempo de simulação, alterações anormais no péptido. Mesmo impondo a estrutura secundária das hélices, as glicinas do *turn* central são bastante flexíveis,

provocando aberturas do *hairpin* e afastamento das hélices em direções opostas. Adicionalmente, o campo de forças MARTINI não foi parametrizado para modelar alterações conformacionais de péptidos ou proteínas. Optou-se, portanto, por fixar o péptido à estrutura de NMR permitindo, no entanto, alguma flexibilidade.

Foi necessário definir diedros para os resíduos da *turn*, uma vez que normalmente, apenas as hélices e folhas beta são definidas por diedros. Após várias combinações de constantes de força para diedros, para ligações e para ângulos de ligação, chegou-se a um conjunto de parâmetros que mantinha a estrutura fixa, mas ao mesmo tempo minimamente flexível. Aplicaram-se os correspondentes diedros da estrutura de NMR a todos os resíduos, com exceção dos terminais desordenados (Gly1, Gly23, Ser24), com as constantes de força iguais às usadas normalmente para hélices (400 kJ/mol/nm^2). Os ângulos foram também alterados para os ângulos da estrutura de NMR, com constante de forças iguais às das hélices (700 kJ/mol/nm^2), exceto para os 3 resíduos terminais desordenados, aos quais usou-se o valor padrão para resíduos desordenados (25 kJ/mol/nm^2). As constantes de força das ligações dos resíduos em dobra foram alteradas para o mesmo valor das hélices alfa (1250 kJ/mol/nm^2). Ao mesmo tempo que chegamos a esta solução, uma topologia semelhante foi utilizada por Fuhrmans et al. para a primeira estrutura de NMR do péptido de 20 aminoácidos **[Fuhrmans 2012]**.

Com esta nova topologia não ocorreram distanciamentos incorretos das hélices, mas ela impede alterações conformacionais da estrutura secundária que eventualmente poderiam ocorrer. De qualquer das formas o que se pretende estudar de seguida é sobretudo a partição e as interações do péptido de fusão numa bicamada, o que segundo estudos anteriores **[Han 2001]**, não altera (em comparação com micelas) a estrutura secundária do péptido em hélice. Sendo assim a imposição da estrutura secundária não trará grandes desvantagens, visto que o péptido mesmo com a estrutura em *hairpin* fixa, apresenta ao mesmo tempo alguma flexibilidade, mesmo que mínima.

Para criar uma bicamada lipídica foi utilizado um método de *self-assembly*, que consiste num processo onde um sistema de componentes desorganizados, neste caso lípidos e o péptido de fusão, formam, ao longo do tempo de simulação, um sistema organizado através de interações entre eles. Optou-se por usar este processo para se ter uma simulação imparcial, não dependendo da bicamada inicial nem da posição do péptido de fusão nesta.

3.1.3. Self-assembly dos lípidos

Ao contrário das proteínas, não existe nenhuma ferramenta que transforme um lípido com resolução atomística num lípido com resolução *coarse-grain* do campo de forças MARTINI, sendo portanto necessário usar uma molécula pré-criada. A fosfatidilcolina (PC) mais utilizada no MARTINI é o dipalmitoil fosfatidilcolina (DPPC), cujas propriedades (área por lípido, difusão lateral, etc.) foram já estudadas. Os resultados reproduziram os dados experimentais na maioria dos casos, sendo que na minoria dos casos reproduziram aproximadamente [Marrink 2003, Marrink 2004, Marrink 2007]. O DPPC é um glicerofosfolípido com duas cadeias derivadas do ácido palmítico (16:0), e cada uma é simulada em MARTINI com quatro partículas C1, ou seja, três partículas reproduzem 4 carbonos e uma partícula reproduz 3 carbonos, como descrito na figura 29.

Os estudos anteriores realizados por Victor et al. tanto com o péptido de fusão [Victor 2015] como com o péptido transmembranar [Victor 2012] foram realizados usando DMPC, um glicerofosfolípido com duas cadeias derivadas do ácido mirístico (14:0), o qual iremos usar também neste estudo. O modelo de MARTINI propõe o uso de três partículas C1 em cada cauda para fosfatidilcolinas com 12 a 15 carbonos, uma vez que o mapeamento em MARTINI são em média de 4 átomos (sem contar os hidrogénios) para uma partícula.

Já foram simulados anteriormente DMPC ou mesmo DPPC com três partículas C1 em cada cauda [Thøgersen 2008, Rzeplia 2009], mas nunca foram estudadas as propriedades deste tipo de membranas. Há, portanto, necessidade de se estudar as propriedades das bicamadas formadas por *self-assembly*, como a área por lípido, a espessura e a difusão lateral. Uma vez que, o nível de hidratação (número de moléculas

de água por cada lípido) das bicamadas tem efeito nas propriedades, nós optamos por um sistema hidratado de 128 DMPCs e 2000 águas MARTINI (equivalente a 4 águas atomísticas), dando uma razão de 62.5 águas atomísticas por lípido.

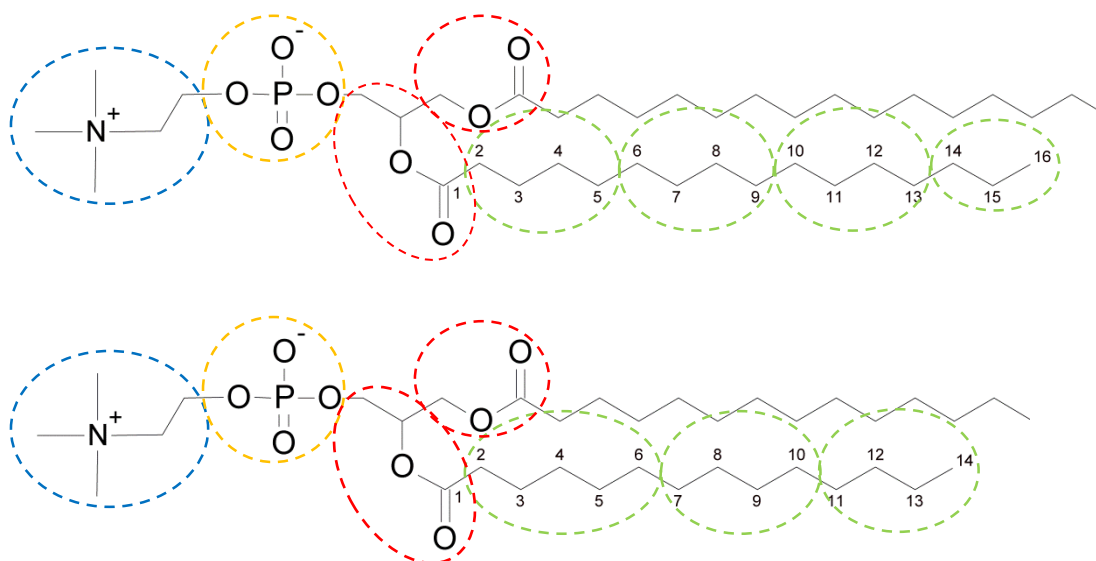


Figura 29 – Representação da fórmula estrutural do DPPC em cima e DMPC em baixo, com círculos a tracejado a definir as partículas na resolução *coarse-grain* de MARTINI. Imagem realizada no programa Chemskech.

Foram iniciadas inúmeras simulações de *self-assembly* sendo apenas selecionadas as que seguiam três critérios após a bicamada estar formada: primeiro, a bicamada deve-se encontrar ao longo dos planos xy, ou seja, com normal no eixo dos z; segundo, a bicamada devia ser constituída por todos os 128 lípidos; por último, foram considerados para estas análises apenas as bicamadas com o mesmo número de lípidos em cada folheto desta, ou seja 64 lípidos. No total obtivemos cinco simulações que respeitam estes critérios.

Área por lípido

A área por lípido de cada uma das cinco simulações foi estudada após 25 ns, tempo suficiente para se formar a bicamada em todas as simulações. Realizou-se uma média de todas as simulações estando o resultado apresentado na figura 30.

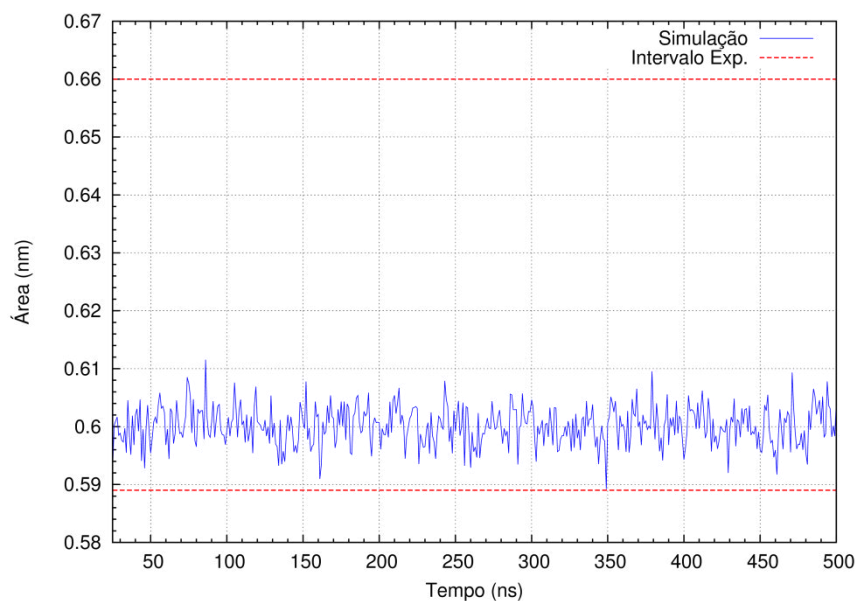


Figura 30 – Variação da área por lípido, a azul e o intervalo experimental à temperatura 303 K a vermelho e tracejado [Kucerka 2004, Poger 2010].

Observa-se pelo gráfico que a área por cada lípido de resolução *coarse-grain* (e com apenas 3 partículas em cada cauda) encontra-se dentro do intervalo experimental. Foi calculada a média nos últimos 200 ns, resultando num valor de $0.599 \pm 0.003 \text{ nm}^2$, intervalo este que está de acordo com outros valores experimentais para DMPC à mesma temperatura (0.597 nm^2 , 0.596 nm^2 e 0.6 nm^2) [Petrache 1998, Nagle 2000, Petrache 2000].

Espessura da membrana

A espessura é maioritariamente importante em estudos com o péptido transmembranar, uma vez que esta tem uma grande influência no ângulo adotado pelos péptidos inseridos na membrana, sendo portanto um fator a ter em conta. Consideramos a espessura de uma membrana como a distância entre os fósforos de cada folheto da bicamada. Para tal calculou-se a densidade do grupo PO₄, que simboliza o grupo fosfato ao longo da caixa, utilizando os últimos 400 ns. Um exemplo do resultado final desta análise encontra-se na figura 31.

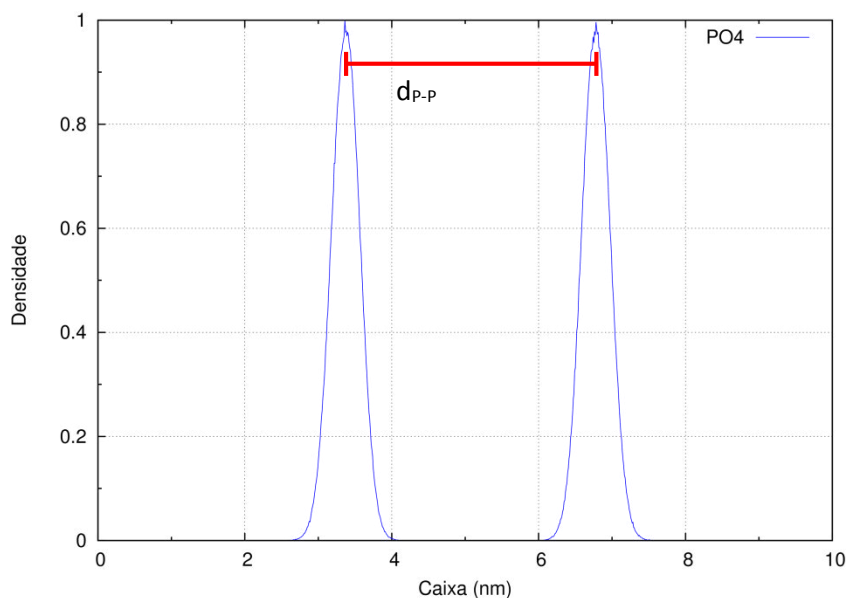


Figura 31 – Densidade normalizada do grupo fosfato ao longo do eixo do z da caixa de simulação. Sendo d_{p-p} , a distância entre os dois picos de densidade de grupos fosfato.

Através do cálculo da distância entre os máximos de densidade do grupo fosfato para as cinco simulações obteve-se um valor de 3.432 ± 0.068 nm, reproduzindo o valor experimental de 3.45 nm [Kucерka 2011]. Dado que existia alguma discrepância (embora pouco significativa) nos valores obtidos nas diferentes réplicas, usou-se outra forma de calcular a espessura. A ferramenta GridMAT divide ambos os folhetos da bicamada numa grelha, determina os pares de fósforos que estão numa posição semelhante das duas grelhas e calcula a distância para cada par de fósforos [Allen 2009]. Ao longo da bicamada apenas ocorria uma variação da espessura de aproximadamente 0.04 nm, e a média das espessuras ronda o valor experimental.

Difusão lateral

A difusão lateral dos lípidos é também uma propriedade analisada normalmente, e ao contrário das duas anteriores é uma propriedade dinâmica. A difusão lateral em simulação computacional é calculada através do declive da função de “mean square displacement” (msd) ao longo do tempo. O valor obtido após a análise foi de $(2.34 \pm 0.15) \times 10^{-7}$ cm²/s, que se encontra na mesma ordem de grandeza dos valores experimentais de aproximadamente 1×10^{-7} e 1.43×10^{-7} cm²/s [Sheats 1978, Orädd 2004]. Encontra-se também na mesma ordem de grandeza do valor calculado por

Marrink et al. para o DPPC de $(3 \pm 1) \times 10^{-7} \text{ cm}^2/\text{s}$ [Marrink 2004], o qual corresponde também a um valor experimental de aproximadamente $1 \times 10^{-7} \text{ cm}^2/\text{s}$ [Sheats 1978].

Visto que as análises da bicamada de DMPC utilizando um modelo do MARTINI para moléculas PC com três partículas em cada cauda do lípido reproduziram os dados experimentais, podemos iniciar os estudos da interação do péptido de fusão com a bicamada lipídica.

3.1.4. Interação do péptido de fusão com bicamada de DMPC

Para estudar a interação do péptido de fusão com uma bicamada de DMPC foram utilizados dois métodos. Um primeiro foi o *self-assembly* (sistema SA) dos lípidos e do péptido de fusão, todos colocados em posições ao acaso (figura 32 A), para que o péptido se insira na bicamada ao mesmo tempo em que esta é formada. Assim sendo, o péptido agrega-se na membrana sem a necessidade de penetrar a parte polar da bicamada. Outro método é colocar o péptido de fusão a uma certa distância, neste caso específico 2.5 nm (figura 32 B), de uma bicamada já pré-equilibrada (sistema PB) e deixar que ele penetre a bicamada vindo do solvente, como aconteceria *in vivo*.

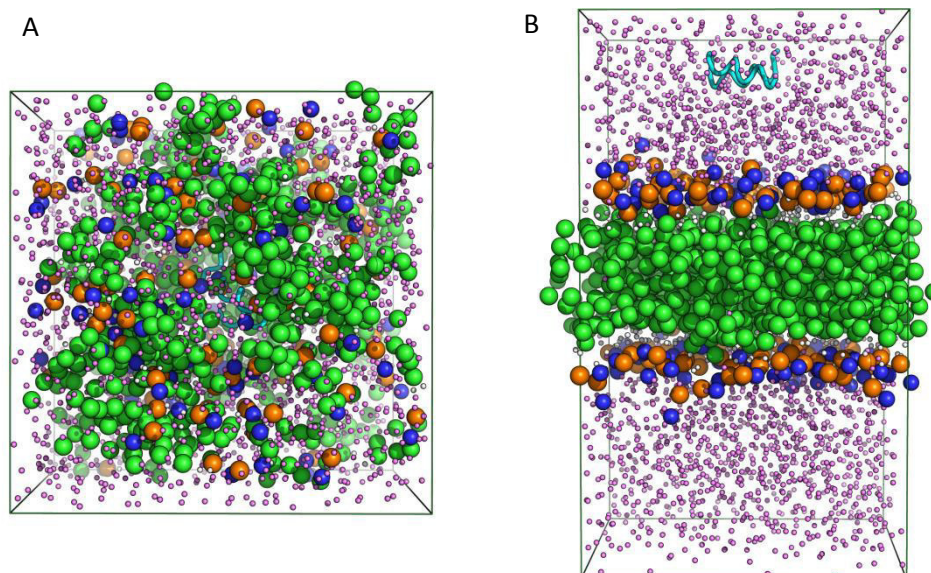


Figura 32 – Arranjos iniciais dos sistemas com membrana formada por *self-assembly* (A) e com a membrana formada e equilibrada previamente (B). O péptido está representado por tubos a ciano, as águas representadas por pequenas esferas violetas e os lípidos representados também por esferas, onde a laranja são representados o grupo fosfato, a azul o grupo ciano, a branco os grupos de glicerol e verde os grupos da cauda.

Foi ainda testado o péptido de fusão com as cadeias laterais dos resíduos ionizáveis tanto no estado protonado, como também no estado desprotonado, representando, respetivamente o péptido a pH de fusão e a pH neutro. Isto para que se possa averiguar a existência de diferenças na interação do péptido com a membrana a diferentes estados de protonação. Os sistemas estudados estão esquematizados na tabela 3.

Tabela 3 – Sumário dos diferentes sistemas das simulações com o péptido de fusão e DMPC.

Sistema	Estado de protonação das cadeias laterais	Membrana
SA_P	Protonadas	<i>Self-assembling</i>
SA_D	Desprotonadas	<i>Self-assembling</i>
PB_P	Protonadas	Pré-equilibrada
PB_D	Desprotonadas	Pré-equilibrada

Nas simulações usando o campo de forças MARTINI, todas as simulações dos sistemas de *self-assembling* (sistemas SA) foram corridas durante 1 μ s, ao passo que para as simulações dos sistemas da pré-bicamada (sistemas PB) o tempo de 1 μ s foi apenas contabilizado após a inserção do péptido na membrana. No fim das simulações CG todas as estruturas finais de todas as réplicas tiveram a sua resolução *coarse-grain* transformada em resolução *united-atom*, utilizando o método de transformação reversa (descrito no segundo capítulo). Após a transformação, cada uma dessas réplicas foi simulada por 500 ns utilizando o campo de forças *united-atom* 54A7. Cada sistema *united-atom* (sistemas UA) vai, portanto, comportar 9 (3x3) simulações.

Devido ao elevado número de simulações é preciso utilizar uma nomenclatura simples que consiga identificar cada uma das simulações. Dentro de um determinado sistema (por exemplo o SA-P), usaremos o nome rXY, para cada uma das réplicas específicas, onde X corresponde à réplica da simulação CG e Y à réplica da simulação UA de uma dada simulação CG. Por exemplo, a réplica r12 corresponde à segunda réplica das simulações *united-atom*, que se realizaram a partir da transformação reversa do final da primeira réplica da simulação *coarse-grain*. De modo a simplificar a escrita e os gráficos. De modo a simplificar também os gráficos e as análises quando analisarmos as

simulações *united-atom*, a referência às réplicas r1 estaremos a referir-nos às três réplicas *united-atom* que advieram da primeira réplica das simulações *coarse-grain*, ou seja às réplicas r11, r12 e r13. Quando nos quisermos referir à primeira réplica das simulações *coarse-grain*, indicaremos especificamente.

Fizeram-se então várias análises, como a partição do péptido ao longo da membrana, o conteúdo em hélice, o ângulo de inserção do péptido em relação ao plano da membrana, as diferentes conformações que adotou depois das simulações UA.

3.1.5. Análises

3.1.5.1. Partição na membrana – Profundidade de inserção

A partição dos péptidos na membrana foi calculada a partir da distância média, no eixo dos z 's, de cada resíduo (centro de massa de todos os átomos do resíduo) a uma das camadas de fósforos (centro de massa de todos os fósforos de uma camada). Para tal foram utilizados os últimos 800 ns das simulações *coarse-grain*, visto que a partir dos 200 ns a distância dos resíduos não variava significativamente. Essas distâncias, assim como a distância da camada de fósforos inferior e a média dos fósforos referência estão representadas na figura 33.

Comparando as profundidades de inserção dos sistemas PB e SA podemos constatar que a utilização de uma membrana formada por *self-assembling* ou pré-equilibrada não tem qualquer influência. Um resultado semelhante já tinha sido obtido num estudo onde se utilizou uma metodologia semelhante [Wee 2008].

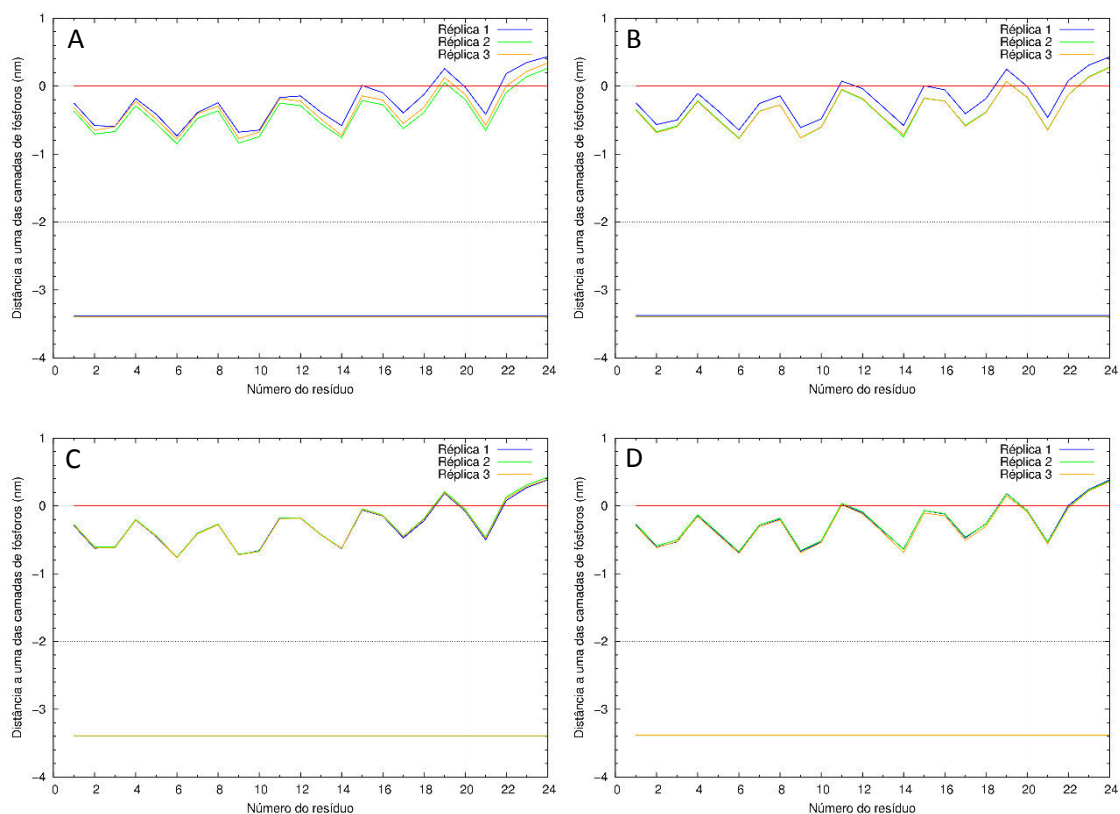


Figura 33 – Distância média, no eixo dos z, a uma das camadas de fósforos para todos os resíduos de aminoácidos das simulações *coarse-grain* (A) SA_P, (B) SA_D (C) PB_P e (D) PB_D.

Analisando as partições de cada resíduo na membrana, observa-se que todas as réplicas possuem uma variação semelhante, exceto para o resíduo Glu11, uma vez que este possui diferentes estados de protonação consoante o sistema simulado. Nos sistemas SA_D e PB_D o resíduo Glu11 encontra-se desprotonado, conferindo-lhe uma carga negativa. Consequentemente aumenta ainda mais a repulsão com o interior apolar da membrana e torna as interações com o grupo colina e a água mais favorável. O mesmo não ocorre para o resíduo Asp19 que também possui diferentes estados de protonação, provavelmente porque, mesmo protonado, já se encontra no lado exterior da interface.

Nos sistemas PB, a variação da penetração do PF na membrana nas diferentes réplicas, são sobreponíveis entre si, o que não ocorre para os sistemas SA (figura 33). Os péptidos simulados com uma membrana pré-equilibrada partem do mesmo local e interagem com uma membrana já formada, logo a única variável nestas simulações seria a forma de inserção do péptido. Analisando a variação da distância para cada resíduo ao longo do 1 μ s de simulação, observa-se que rapidamente assumem e mantêm uma

posição estável, para um péptido de fusão a penetrar a membrana, vindo do solvente, num modelo MARTINI. Os péptidos simulados num processo de *self-assembly* partem também da mesma posição, contudo os lípidos espalhados ao acaso pela caixa de simulação vão formar, de forma estocástica, uma membrana à volta do péptido. Assim sendo a posição inicial do péptido, aquando da formação da membrana, é variável podendo, portanto ficar preso numa certa posição da membrana, causando as leves diferenças apresentadas nos sistemas SA, que não ocorre nos sistemas PB.

No caso das simulações usando o campo de forças *united-atom* 54A7 foram utilizados nos cálculos apenas as distâncias de cada resíduo após 300 ns, tempo necessário para ocorrer equilíbrio do sistema e os valores não variarem significativamente. Além disso, para cada resíduo foi calculada ainda uma média entre cada conjunto de três réplicas UA que provieram da réplica CG correspondente (após a transformação reversa). Na figura 34 estão representadas as médias de cada conjunto de simulações *united-atom* (r1, r2 e r3) incluindo a média de todos os péptidos de um determinado sistema, ou seja, a média global do sistema (SA_P, SA_D, PB_P e PB_D).

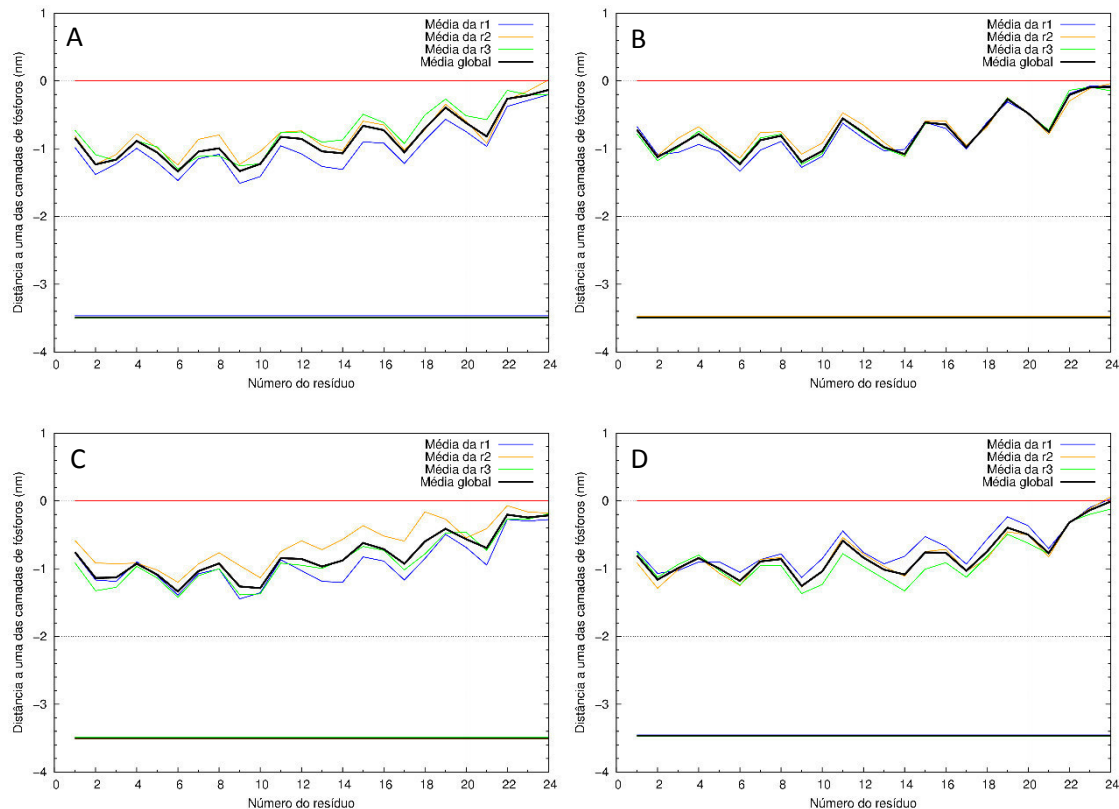


Figura 34 – Distância média, no eixo dos z, a uma das camadas de fósforos para todos os resíduos de aminoácidos das simulações *united-atom* (A) SA_P, (B) SA_D, (C) PB_P e (D) PB_D.

Em comparação com a partição dos mesmos sistemas antes da transformação reversa pode-se notar que existem algumas variações diferentes entre réplicas, o qual se deve provavelmente há existência de deformações das hélices e da estrutura em *hairpin*. No entanto, pelas médias podemos ver que apresentam a mesma distribuição ao longo da membrana como ocorria nos sistemas CG. Também aqui a maior diferença entre os sistemas protonados e desprotonados é resíduo Glu11, onde se nota um pico proeminente quando este resíduo está desprotonado, devido à sua menor afinidade para o interior hidrofóbico da membrana.

Nota-se ainda que o efeito da primeira metade do péptido se apresentar paralelo aos fósforos e a segunda metade apresentar certo declive em direção à interface, é mais proeminente nos péptidos simulados com o 54A7. A comparação entre as médias dos diferentes sistemas pode ser visualizada na figura 35.

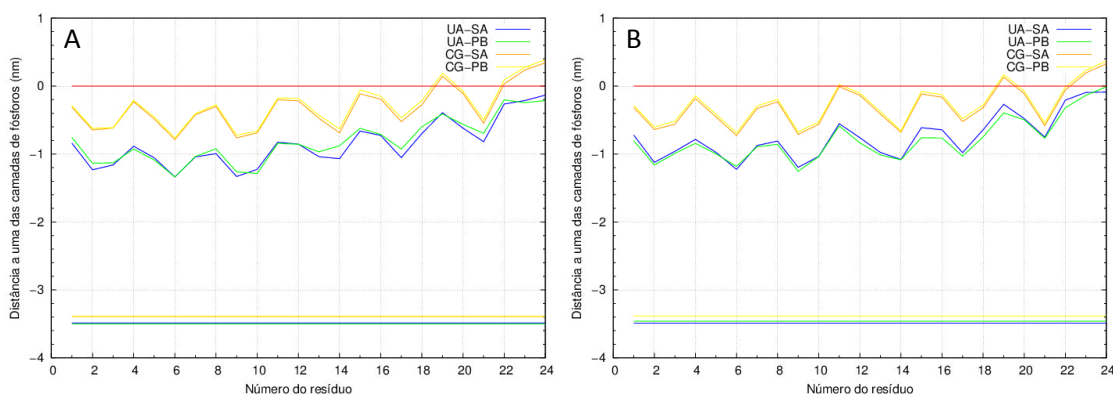


Figura 35 – Distância média, no eixo dos z, a uma das camadas de fósforos média de cada uma das réplicas de cada sistema para todos os resíduos de aminoácidos para o péptido de fusão com os resíduos ácidos (A) protonados ou (B) desprotonados.

Mesmo com a diferença de espessuras das membranas das duas diferentes representações (cerca de 0,1 nm) é notório que as simulações utilizando o campo de forças com mais detalhe permitiram o péptido de fusão penetrar mais profundamente na membrana lipídica. Um dos problemas conhecidos no campo de forças MARTINI é a menor penetração da água na bicamada, em comparação com campos de força GROMOS, provavelmente devido à representação de quatro moléculas de água numa partícula de água MARTINI [Marrink 2004, Monticelli 2008]. Portanto não é surpreendente que algo parecido também ocorra com o péptido de fusão, onde as

partículas tanto da bicamada como do próprio péptido possuem raios de van der Waals significativamente grandes.

Note-se ainda que em todas as simulações os resíduos hidrofóbicos encontram-se mais inseridos que os restantes, como ocorre experimentalmente [Lorieau 2010], nomeadamente os resíduos Iso6, Phe9, Iso10, Trp14, Met17 e Trp21.

A preferência do C-terminal para se localizar na interface membrana-água e consequentemente menos inserido deve-se provavelmente à concentração de triptofanos e tirosinas nessa região. Estes resíduos em proteínas transmembranares encontram-se preferencialmente em hélices na interface membrana-água do que em *loops* ou mesmo hélices transmembranares [Granseth 2005].

A comparação desta, ou de outras análises, com resultados experimentais do péptido de fusão precisa de ser feita com cuidado, uma vez que, excetuando os estudos de NMR por Lorieau et al. [Lorieau 2010], não existem outros estudos experimentais ou atomísticos do péptido de fusão com 23 resíduos de aminoácidos e do subtipo H1. A par disto, existem estudos tanto experimentares como computacionais com o péptido de fusão de 20 aminoácidos e do subtipo H3 que mostram resultados próximos dos nossos [Han 2001, Vaccaro 2005, Vaidya 2010, Légaré 2012]. Em todos estes estudos o péptido de fusão apresentou uma inserção superior no N-terminal do que no C-terminal do péptido, e, por vezes, uma conformação semelhante, com o N-terminal plano e o C-terminal inclinado para a superfície. No entanto, o valor de inserção do péptido aqui apresentado é superior ao de outros estudos, com valores de inserção máxima de 0.7 nm [Han 2001] e 1 nm [Vaidya 2010, Légaré 2012].

3.1.5.2. Conteúdo em hélice

O conteúdo em hélice do péptido de fusão em membrana é calculado de forma semelhante ao realizado anteriormente para o péptido de fusão em água, utilizando a ferramenta DSSP. A evolução do número de resíduos em hélice ao longo do tempo de simulação, para cada um dos sistemas com resolução *united-atom* encontra-se na figura 36.

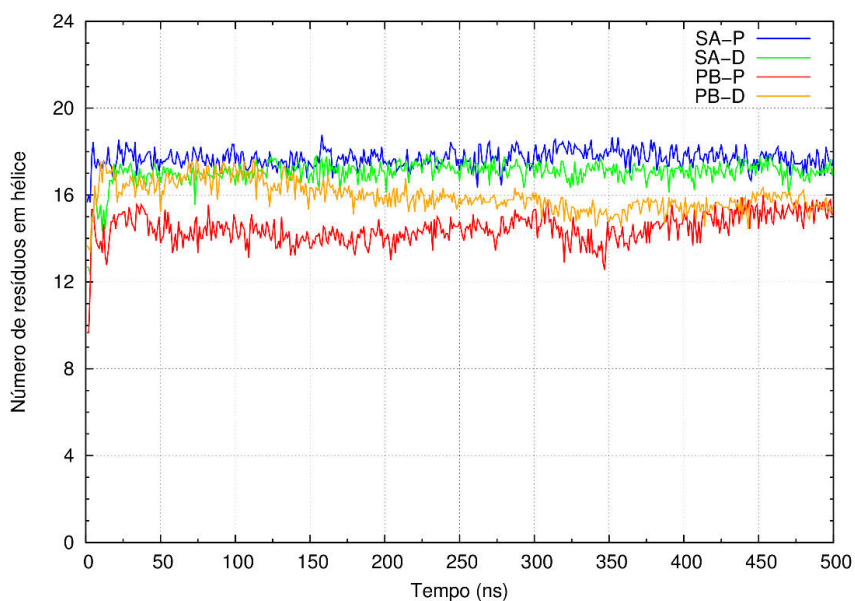


Figura 36 – Número de resíduos em estrutura secundária de hélice ao longo do tempo para os quatro diferentes sistemas das simulações *united-atom*.

Após análise do gráfico a observação mais clara é o número inferior de resíduos em hélice do péptido de fusão nos sistemas com uma membrana pré-equilibrada, em comparação aos sistemas com uma membrana formada por *self-assembly*. A média das réplicas com ambas as protonações nos sistemas SA têm uma variação semelhante, e se considerarmos os últimos nanossegundos (quando o sistema está mais equilibrado), também os sistemas PB apresentam uma variação semelhante. Os resultados do número de resíduos em hélice médio ao longo dos últimos 200 ns foram: 17.81 ± 0.37 para o sistema SA_P, 17.10 ± 0.33 para o SA_D, 14.72 ± 0.64 para o PB_P e 15.48 ± 0.38 para o PB_D. A (tabela 4). Na verdade tanto se pode observar pelo gráfico como pela variância que o sistema PB_P apresenta um constante aumento do número de resíduos até perto dos 450 ns. Os mesmos resultados para os últimos 50 ns foram: 17.69 ± 0.33 para o sistema SA_P, 17.04 ± 0.35 para o SA_D, 15.29 ± 0.34 para o PB_P e 15.69 ± 0.37 para o PB_D (tabela 4).

Tabela 4 - Número médio de resíduos em hélice analisado durante diferentes tempos de simulação

Sistema	Últimos 200 ns (nº de resíduos \pm variância)	Últimos 50 ns (nº de resíduos \pm variância)
SA_P	17.81 \pm 0.37	17.69 \pm 0.33
SA_D	17.10 \pm 0.33	17.04 \pm 0.35
PB_P	14.72 \pm 0.64	15.29 \pm 0.34
PB_D	15.48 \pm 0.38	15.69 \pm 0.37

Os sistemas com a protonação das cadeias laterais dos grupos acídicos do péptido possuem em média um maior número de resíduos em hélice para o caso do *self-assembly*. No entanto, o oposto ocorreu nas simulações feitas com bicamadas pré-equilibradas. As diferenças são no entanto muito pouco significativas e dentro do intervalo da variância, sugerindo que a protonação dos resíduos não possui relevância na estrutura do péptido na membrana, como já tinha sido referido na literatura [Lorieau 2010].

Por outro lado parece que a forma como a membrana foi originada é um fator para a estabilidade das hélices. A maior diferença entre as duas membranas é o número de lípidos em cada uma das faces da membrana. Para as membranas pré-equilibradas sem o péptido, que foi adicionado posteriormente, o número de lípidos em cada camada é de 64. Em contraste, no caso das membranas que foram formadas com o péptido presente (*self-assembly*), foi permitido haver um número diferente de lípidos nos dois folhetos da bicamada. Uma vez que se esperava os lípidos organizarem-se entre os dois folhetos de modo a acomodar livremente o péptido no folheto em que estava inserido. Como seria expectável, a maioria (5 em 6 réplicas) destas simulações CG apresentam um menor número de lípidos na camada onde o péptido se encontra inserido. Nas três simulações UA que partiram dessa réplica CG com um maior número de lípidos no folheto onde o péptido estava inserido, a média de resíduos em hélice é inferior à média das outras simulações, reforçando o que foi sugerido no início do parágrafo.

Seguindo a linha do que tinha sido referido na análise anterior, a única forma de comparação do conteúdo em hélice com resultados experimentais é através da estrutura obtida por NMR [Lorieau 2010]. Esta estrutura apresenta um resíduo no N-

terminal e dois no C-terminal com estrutura desordenada, mas as duas glicinas do *loop* são por vezes consideradas como hélice pelo DSSP, o que daria 19-21 resíduos em hélice. Pela análise anterior vimos que o sistema SA-P foi o que apresentou um maior conteúdo de hélice, mas mesmo assim inferior à estrutura inicial.

Ao analisar as estruturas resultantes da transformação reversa de *coarse-grain* para *united-atom*, observam-se resíduos que se encontram numa conformação próxima de hélice, mas que não é contabilizada pelo DSSP. É possível que esta estrutura inicial mais instável, proveniente da transformação reversa, possa surtir um efeito na evolução conformacional do péptido. Embora o número de resíduos em hélice aumente inicialmente (figura 36) sugerindo um aumento de estabilidade, pode ser que apenas ocorra localmente e a nível global a estrutura inicial continue a ter um impacto. Os resultados instáveis da transformação reversa em proteínas é um problema conhecido, o qual tem sido alvo de melhorias por parte do grupo do Marrink.

De forma a ter uma visão sobre que resíduos possuem uma maior propensão para se encontrar em hélice realizou-se uma análise onde foi contabilizado, para cada resíduo, mas a percentagem de tempo (após os 300 ns de simulação) que apresentavam uma estrutura em hélice (figura 37).

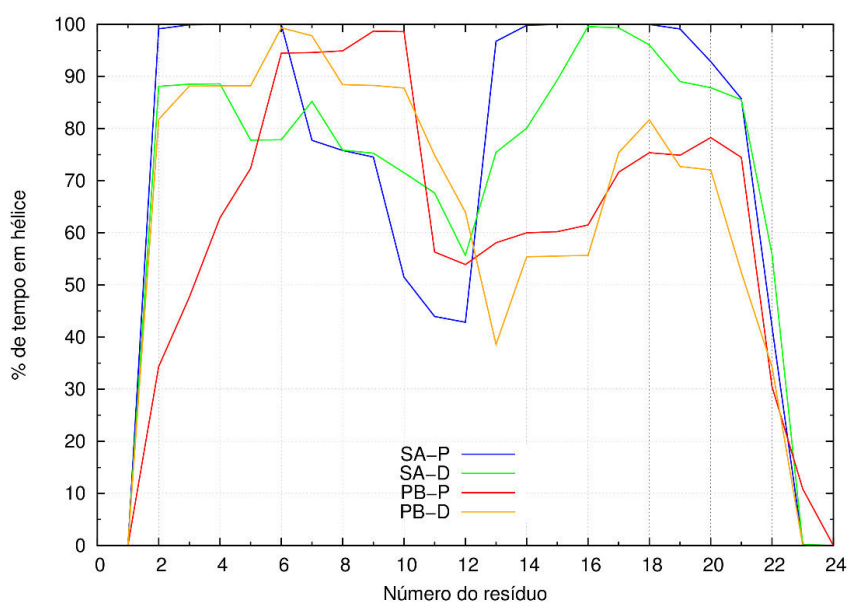


Figura 37 – Percentagem de tempo que cada resíduo do péptido de fusão apresenta uma estrutura em hélice, nos últimos 200 ns de simulação, para os quatro diferentes sistemas das simulações *united-atom*.

Como já tinha sido referido, as glicinas na posição 12 e 13, que correspondem ao loop entre as duas hélices do hairpin na estrutura de NMR, são por vezes contabilizadas como estrutura em hélice pelo DSSP. No entanto é notório que nessa zona existe um vale com uma menor percentagem de estrutura em hélice, mesmo que não tanto como os resíduos terminais, principalmente Gly1, Gly23 e Ser24 que não possuem estrutura secundária na estrutura de NMR. De novo observam-se semelhanças entre os sistemas PB entre si e SA entre si, principalmente na segunda hélice.

Comparando os resultados da partição na membrana e o conteúdo em hélice do péptido observa-se uma relação entre uma maior profundidade do péptido na membrana e um maior número de resíduos em hélice. Uma vez que um péptido mais inserido numa membrana encontra um meio mais hidrofóbico e menor contacto com o solvente, o qual favorece a estrutura em hélice. Contudo, analisando em paralelo a partição do péptido na membrana e o seu conteúdo em hélice para cada uma das réplicas essa mesma relação não se encontra presente em todos os casos. Mesmo sendo um fator importante podem existir outros fatores a contribuir, como por exemplo, a estrutura inicial do péptido após a transformação reversa, a qual não é o processo ideal e apresenta problemas, como já foi debatido mais acima.

3.1.5.3. Ângulo de inclinação

O ângulo de inclinação do péptido transmembranar ao plano da membrana foi calculado usando a ferramenta `g_bundle` do Gromacs. Esta ferramenta calcula o ângulo entre um vetor, definido por dois centros de massa de dois respetivos conjuntos de resíduos, e um plano. Neste caso específico foram escolhidos 4 resíduos para cada centro de massa, número de resíduos ao qual corresponde uma volta de uma hélice alfa. Foram escolhidos então os resíduos 3-6 e 7-10 para a primeira hélice e 14-17 e 18-21 para a segunda hélice. Estes resíduos foram escolhidos tendo em atenção não só a estrutura inicial, mas também o conteúdo em hélice de cada resíduo. Uma vez que uma volta de uma hélice alfa é composta por 4 resíduos, o centro de massa da cadeia central desses 4 resíduos será o centro da hélice, diminuindo assim algum ruído produzido pelo movimento de um resíduo em particular. É portanto imperativo que se utilizem os resíduos que adotem um maior tempo de simulação em hélice

Na primeira hélice, os resíduos após o terceiro resíduo e até aos resíduos próximos do *turn*, encontram-se a maioria do tempo de simulação em hélice. Tanto os resíduos na posição 2 como na posição 11 em alguns sistemas apresentam um conteúdo em hélice abaixo dos 50% (figura 37). Optou-se, portanto, pelos resíduos 3 a 6 para o primeiro ponto (centro de massa) e os resíduos 7 a 10 para o segundo ponto do vetor.

Na segunda hélice, os resíduos 12 e 13 correspondem às glicinas do *turn* do péptido e ao mesmo tempo aos mínimos de conteúdo em hélice, portanto os resíduos selecionados começariam no resíduo 14 ou 15. A partir dos resíduos 20-21 ocorre um decréscimo abrupto de conteúdo em hélice, onde mesmo o pior sistema (sistema PB-D) o resíduo 21 ainda se encontra mais de metade do tempo em hélice. Optou-se portanto pelos resíduos 14 a 17 para o primeiro ponto e os resíduos 18-21 para o segundo ponto do vetor.

1ª hélice

Foram calculados os ângulos de inclinação entre o vetor criado pela primeira hélice do péptido (resíduos 3-10) e o plano xy da membrana para todas as simulações, estando representadas na figura 38 as médias globais de cada um dos sistemas.

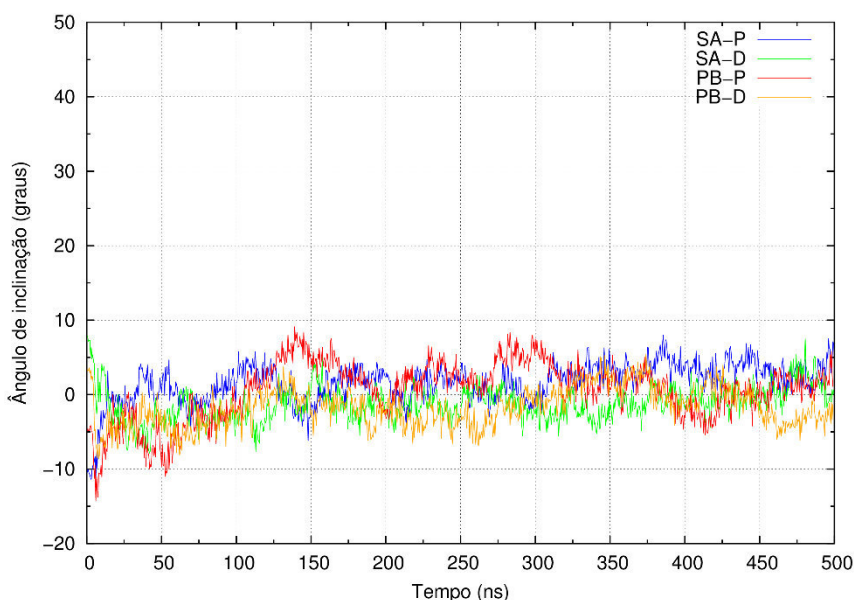


Figura 38 – Variação do ângulo de inclinação para a primeira hélice (resíduos 3-10) para todos os sistemas estudados usando o campo de forças 54A7.

Não existe uma grande variação dos ângulos ao longo do tempo de simulação, exceto nos instantes iniciais que se devem provavelmente à destruição parcial após a transformação reversa. O ângulo da primeira hélice flutua aproximadamente à volta dos zero graus, ou seja de uma orientação paralela ao plano da membrana.

Este valor pode ser utilizado como comparação com os dados experimentais, embora a sua informação seja limitada. Para avaliar melhor a variação do ângulo de inclinação, efetuamos histogramas de frequência de cada ângulo, estando estes a média de cada réplica e a média global para cada sistema representados na figura 39.

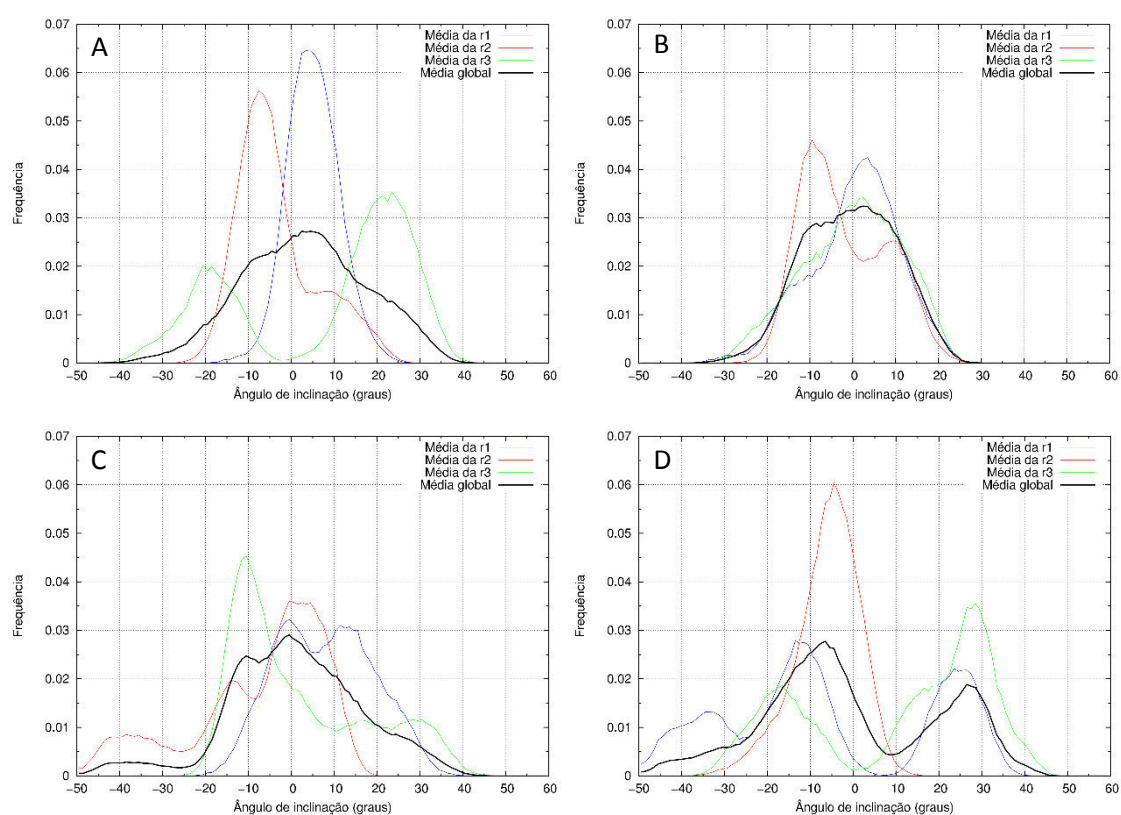


Figura 39 – Histograma da frequência de cada ângulo de inclinação da primeira hélice para to sistema (A) SA_P, (B) SA_D, (C) PB_P e (D) PB_D usando o campo de forças 54A7, para cada uma das réplicas e uma média global de todos os valores.

Ao analisar os histogramas reparamos que embora haja alguma variação, por norma os ângulos à volta dos zero graus são os mais frequentes. Isto mostra alguma liberdade para a hélice variar a sua inclinação, tanto em direção ao centro da bicamada como em direção à superfície desta. Note-se que os resultados com variações mais longe da norma são das simulações com uma bicamada pré-equilibrada a qual apresenta uma

menor inserção na membrana e menor conteúdo em hélice, sugerindo que as variações superiores são causadas por um péptido mais destruído.

A comparação com as simulações utilizando o campo de forças CG podem ser realizadas através do histograma de variação (figura 40) e dos valores médios com respectivos desvios padrão (tabela 5).

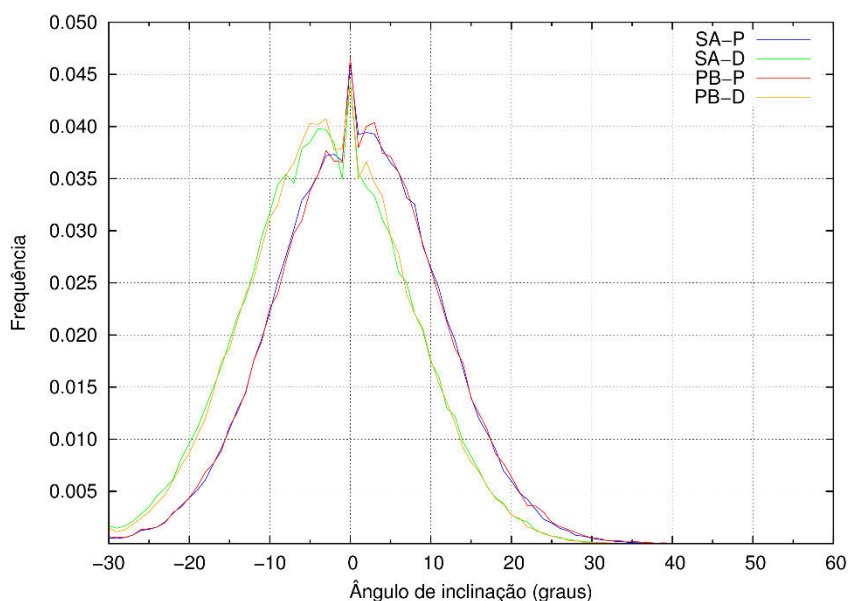


Figura 40 – Histograma da frequência de cada ângulo de inclinação da primeira hélice para todos os sistemas estudados usando o campo de forças MARTINI.

Tabela 5 – Valores médios e respectivos desvios padrão da segunda hélice para cada um dos sistemas e campos de força.

Sistema	UA	CG
SA_P	3.24 ± 1.74	1.32 ± 5.78
SA_D	-0.59 ± 2.13	-2.42 ± 5.75
PB_P	0.82 ± 2.40	1.38 ± 5.7
PB_D	-0.37 ± 2.38	-2.25 ± 5.64

Aparentemente, segundo os gráficos e os valores, os péptidos protonados apresentam um maior ângulo de inclinação tanto para o campo de forças UA como CG. No entanto a variação está dentro do desvio, portanto não podes extrapolar tal resultado.

2ª Hélice

De forma semelhante ao realizado para a primeira hélice do péptido de fusão, também se estudou a evolução do ângulo de inclinação da segunda hélice (resíduos 14-21) ao longo do tempo de simulação para os diferentes sistemas, estando a média de cada sistema representado na figura 41.

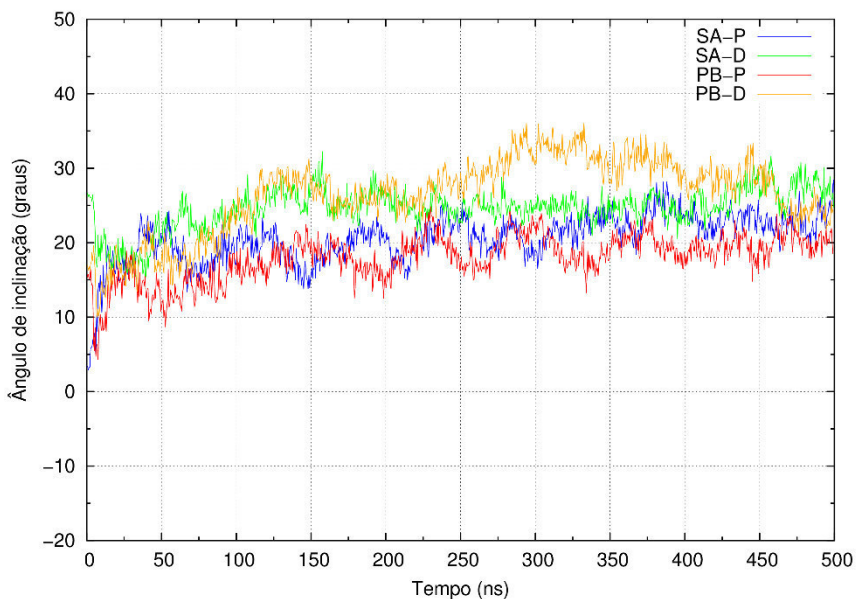


Figura 41 – Variação do ângulo da segunda hélice do péptido de fusão com o plano da membrana ao longo do tempo, para os diferentes sistemas estudados.

Assim como na primeira hélice não houve grande variação após os momentos iniciais (embora o sistema PB_D apresente uma variação superior). No entanto, a segunda hélice apresenta ângulos de inclinação superiores, a variar aproximadamente entre os 20 e os 30, estando a hélice orientada para o exterior. Tal como na primeira hélice também se realizou o calculo de histogramas de frequência (figura 42 e 43).

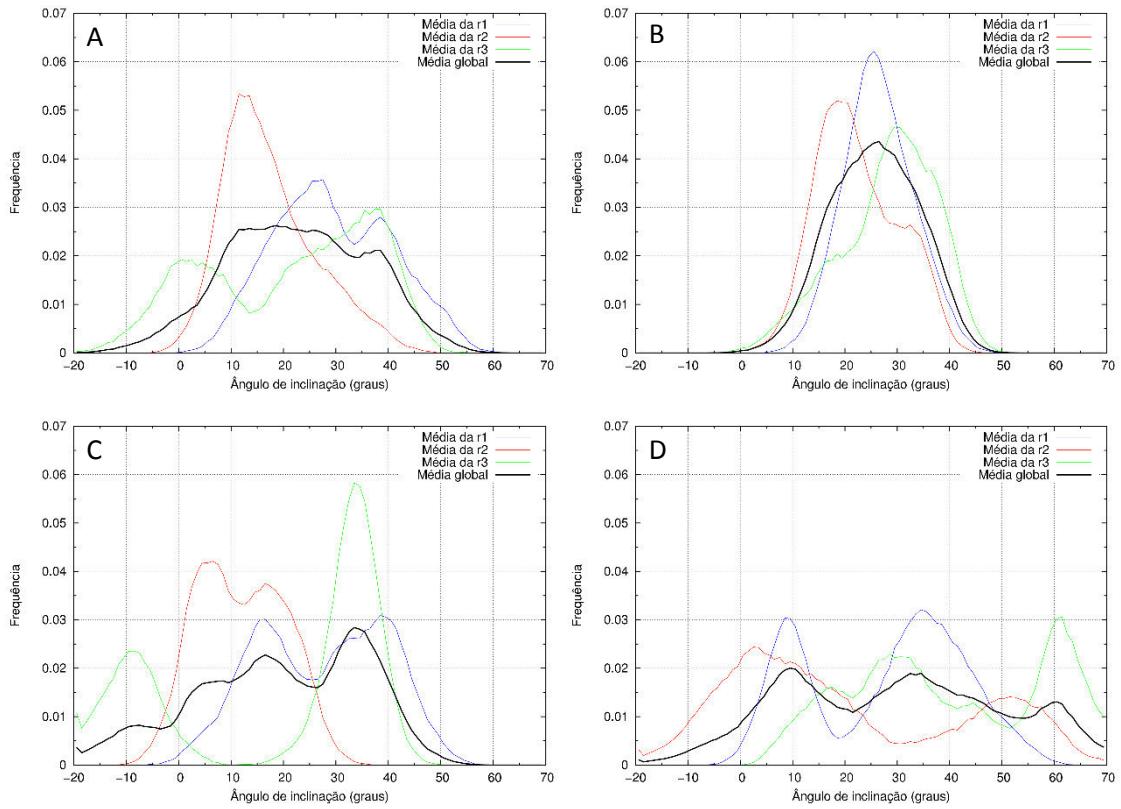


Figura 42 – Histograma da frequência de cada ângulo de inclinação da segunda hélice para to sistema (A) SA_P, (B) SA_D, (C) PB_P e (D) PB_D usando o campo de forças 54A7, para cada uma das réplicas e uma média global de todos os valores.

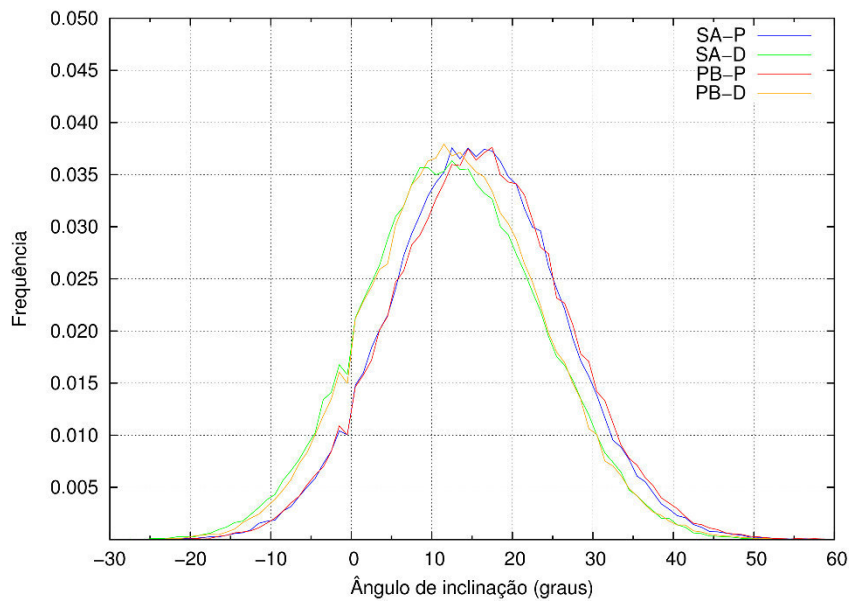


Figura 43 - Histograma da frequência de cada ângulo de inclinação da segunda hélice para todos os sistemas estudados usando o campo de forças MARTINI.

Estando os ângulos médios da segunda hélice e respetivos desvios padrões para cada um dos sistemas e campos de forças estudados na tabela 6.

Tabela 6 – Valores médios e respetivos desvios padrão da segunda hélice para cada um dos sistemas e campos de força.

Sistema	UA	CG
SA_P	22.74 ± 1.83	15.56 ± 6.11
SA_D	25.67 ± 2.01	12.60 ± 6.29
PB_P	19.52 ± 1.83	15.89 ± 6.18
PB_D	29.24 ± 3.23	12.43 ± 6.15

Neste caso podem se notar dois resultados opostos. Ao contrário do que acontecia na primeira hélice, nos sistemas UA os sistemas protonados apresentam um ângulo inferior em média que os sistemas desprotonados. Na primeira hélice as diferenças eram pouco significativas e dentro do desvio padrão. Se tivermos em atenção que os dois resíduos tituláveis estão na segunda hélice, não é estranho que apenas observemos uma grande variação na segunda hélice. O segundo resultado oposto é entre os sistemas UA e CG da segunda hélice, onde neste último os sistemas protonados tem um ângulo superior aos sistemas desprotonados, seguindo uma variação semelhante à primeira hélice. O campo de forças MARTINI apresenta dificuldades em simular corretamente resíduos tituláveis quando carregados, o que pode explicar a diferença com as simulações *united-atom*.

Podemos concluir então que a primeira hélice em média apresenta-se praticamente paralela à membrana, ao passo que a segunda hélice apresenta um ângulo de inserção de 20 a 30 graus, indicando que se encontra direcionado para o exterior da membrana. Tal tinha sido verificado no estudo da partição do péptido de fusão na membrana, onde se observava que a segunda hélice diminuía gradualmente a sua inserção. Tal des de facto a hemaglutinina encontra-se no C-terminal e exterior da bicamada.

Os resultados aqui apresentados reproduzem a estrutura de NMR, onde existe uma hélice planar e outra inclinada [Lorieau2010]. Outras comparações com estudos de péptidos de fusão de 20 resíduos de aminoácidos e de outros subtipos (H3 ou H5), pois

este péptido apresenta uma estrutura na membrana completamente diferente, dando ângulos de 38°, 25° e 33° [Macosko1997, Han2001, Vaccaro2009].

3.1.5.4. Conformações finais

Após os 500ns de simulação, tempo que se considera os sistemas já equilibrados, foram realizadas imagens mostrando a conformação final de todas as simulações utilizando o campo de forças 54A7, estando representadas das figuras 44-47.

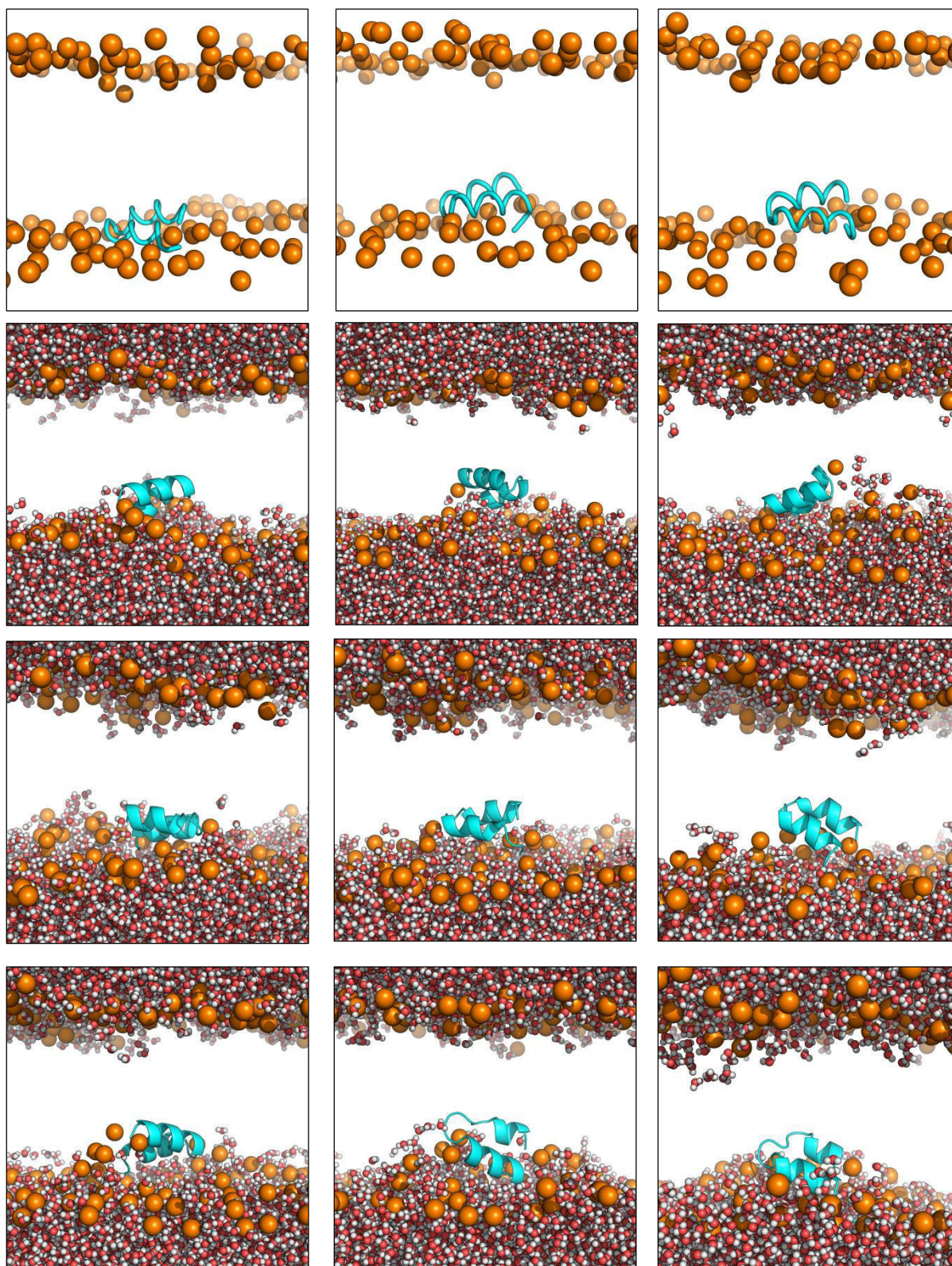


Figura 44 – Conformações finais para cada um dos replicados do péptido de fusão para o sistema SA.P. As três primeiras imagens representam as réplicas coarse-grain, e por baixo delas encontram-se as três respetivas réplicas *united-atom* que advieram da réplica coarse-grain correspondente. Os lípidos são representados apenas pelo grupo fosfato como uma esfera laranja, no caso das conformações UA as águas são representadas por esferas vermelhas para o oxigénio e brancas para os hidrogénios. O péptido está representado por tubos nas conformações CG e por *cartoon* nas conformações UA, ambas a ciano e com o C-terminal virado para o ponto de observação.

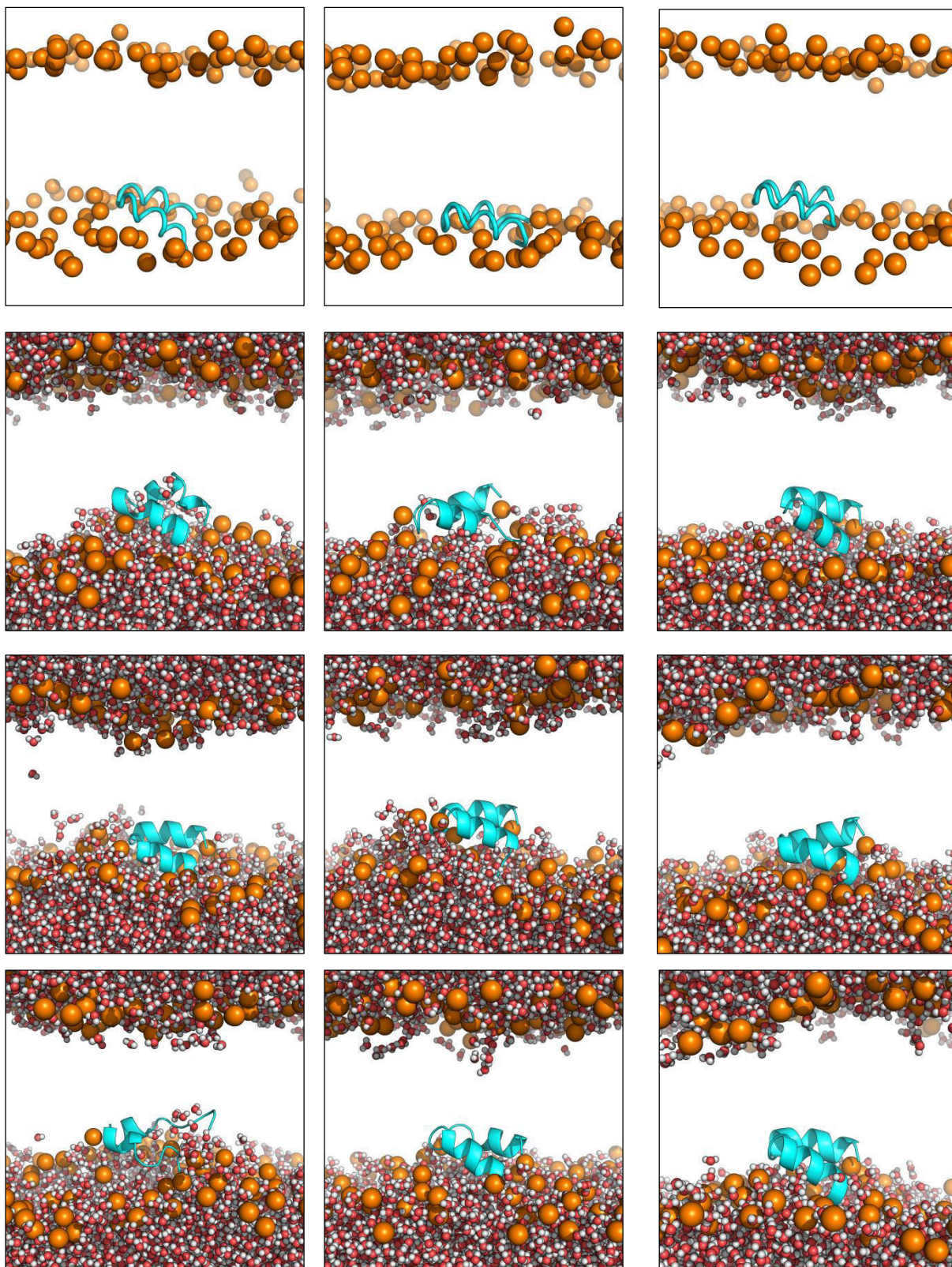


Figura 45 – Conformações finais para cada um dos replicados do péptido de fusão para o sistema SA-D. As três primeiras imagens representam as réplicas *coarse-grain*, e por baixo delas encontram-se as três respetivas réplicas *united-atom* que advieram da réplica *coarse-grain* correspondente. Os lípidos são representados apenas pelo grupo fosfato como uma esfera laranja, no caso das conformações UA as águas são representadas por esferas vermelhas para o oxigénio e brancas para os hidrogénios. O péptido está representado por tubos nas conformações CG e por *cartoon* nas conformações UA, ambas a ciano e com o C-terminal virado para o ponto de observação.

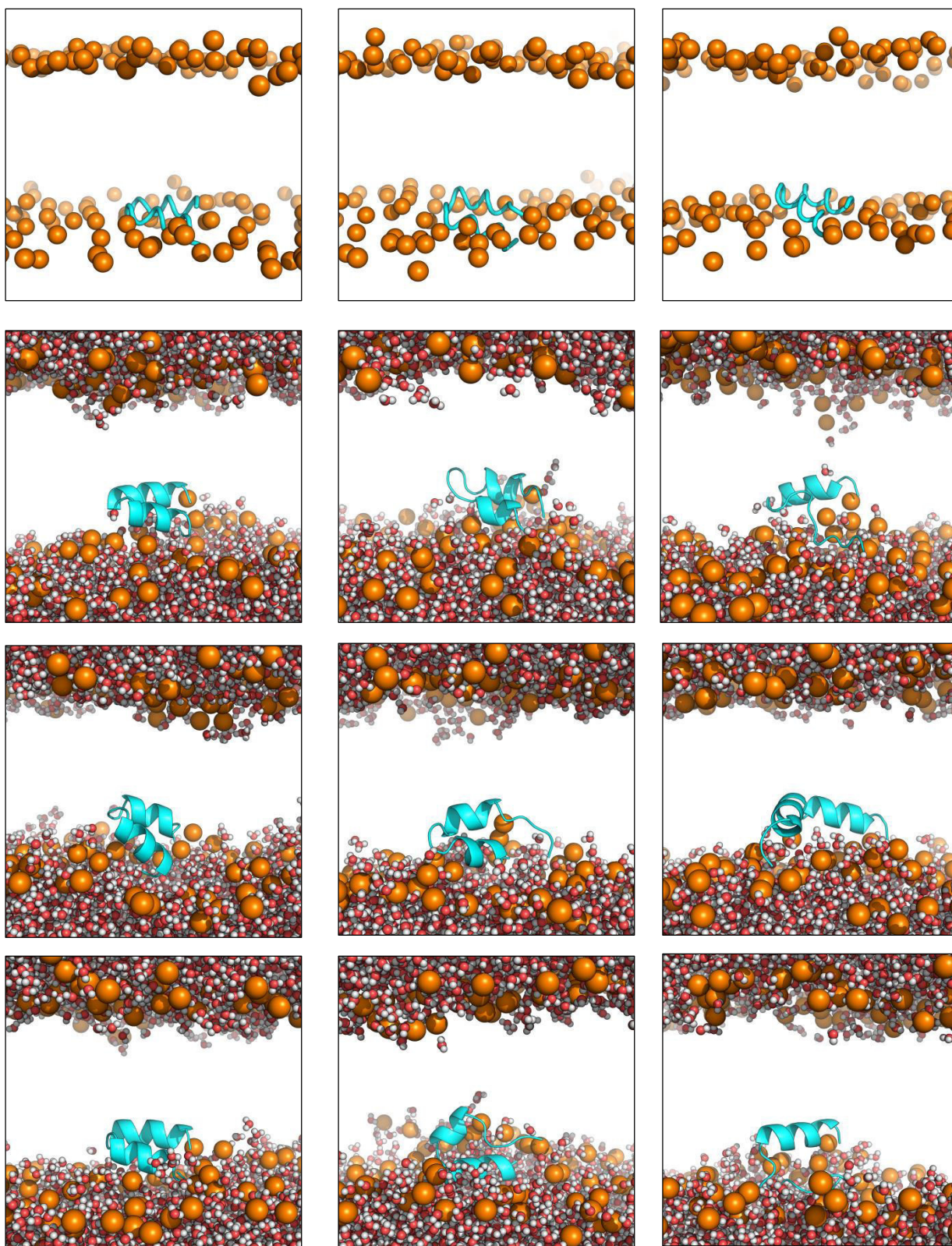


Figura 46 – Conformações finais para cada um dos replicados do péptido de fusão para o sistema PB-P. As três primeiras imagens representam as réplicas *coarse-grain*, e por baixo delas encontram-se as três respetivas réplicas *united-atom* que advieram da réplica *coarse-grain* correspondente. Os lípidos são representados apenas pelo grupo fosfato como uma esfera laranja, no caso das conformações UA as águas são representadas por esferas vermelhas para o oxigénio e brancas para os hidrogénios. O péptido está representado por tubos nas conformações CG e por *cartoon* nas conformações UA, ambas a ciano e com o C-terminal virado para o ponto de observação.

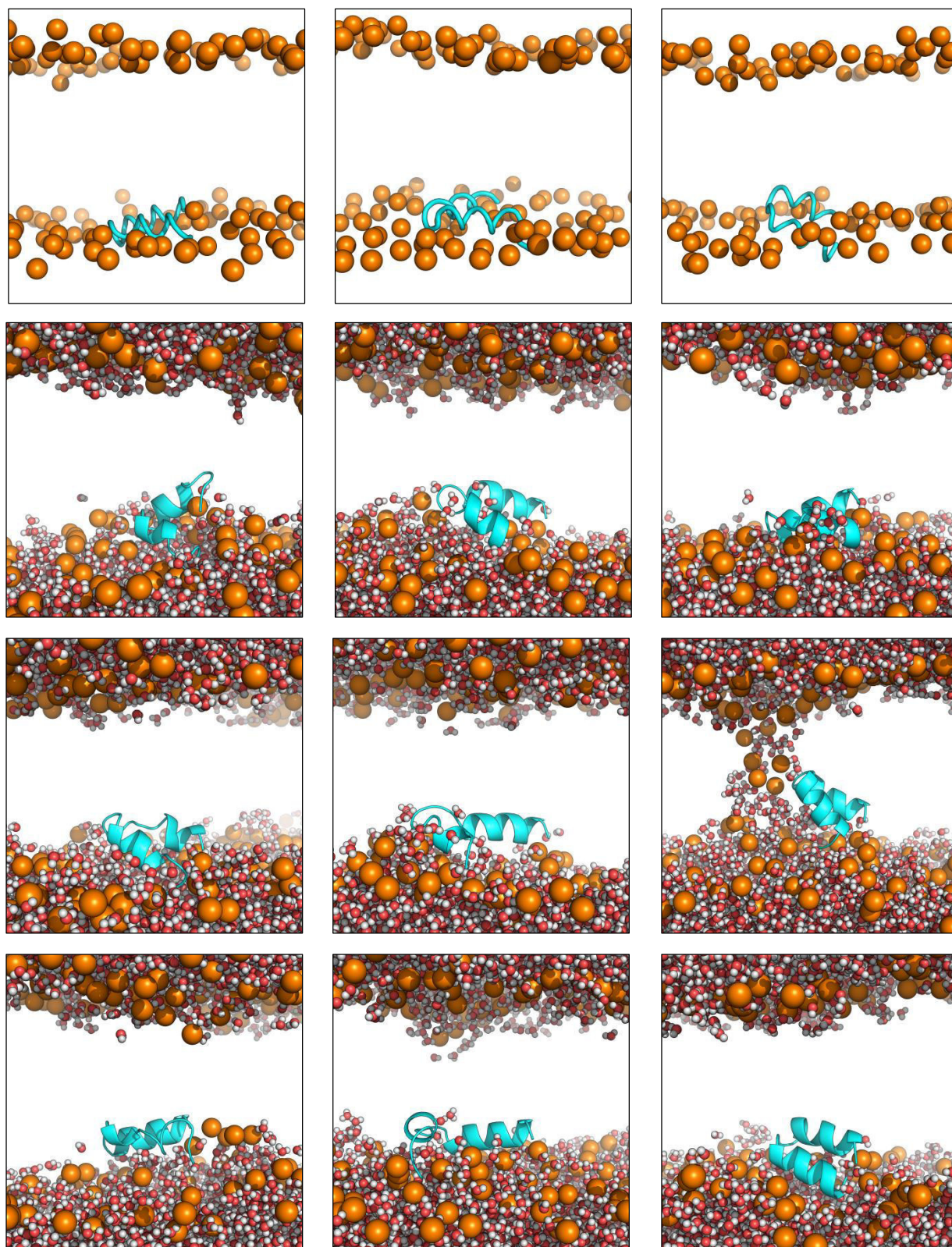


Figura 47 – Conformações finais para cada um dos replicados do péptido de fusão para o sistema PB-D. As três primeiras imagens representam as réplicas *coarse-grain*, e por baixo delas encontram-se as três respetivas réplicas *united-atom* que advieram da réplica *coarse-grain* correspondente. Os lípidos são representados apenas pelo grupo fosfato como uma esfera laranja, no caso das conformações UA as águas são representadas por esferas vermelhas para o oxigénio e brancas para os hidrogénios. O péptido está representado por tubos nas conformações CG e por *cartoon* nas conformações UA, ambas a ciano e com o C-terminal virado para o ponto de observação.

Aparte de algumas deformações, o péptido apresenta uma orientação paralela à membrana, onde a segunda hélice, por norma, tem um maior declive que a primeira, como já se tinha observado pelas análises de partição e dos ângulos. Esta estrutura e disposição das hélices reproduz a estrutura obtida por NMR, onde existe uma hélice planar e uma outra mais inclinada [Lorieau 2010].

Como também já tinha sido referido em análises anteriores, o péptido encontra-se no interior do plano dos fósforos, de uma forma mais inserida nas simulações UA que nas CG. Ainda, os grupos aromáticos encontram-se virados para o interior hidrofóbico da membrana.

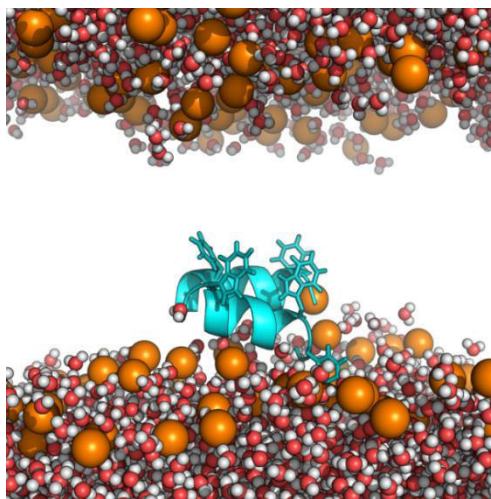


Figura 48 – Estrutura representativa, de uma das réplicas do péptido de fusão, com uma orientação paralela à membrana. Onde a ciano e *cartoon* está representado o péptido de fusão, com os resíduos aromáticos representados em tubos. As águas estão representadas por esferas vermelhas e brancas para os oxigénios e os hidrogénios, respetivamente. Os lípidos são representados pelo fósforo em esferas laranja.

Na variedade de estruturas de todas as simulações aqui estudadas vamos referir dois casos especiais: a abertura do *hairpin*, adotando a forma de um “L” e a formação de um poro na membrana.

Abertura do *hairpin*

Embora a estrutura de NMR do péptido de fusão a pH baixo e pH neutro sejam iguais, a diminuição do pH aumentou a taxa de relaxação do ^{15}N da cadeia principal de alguns resíduos [Lorieau 2010]. Lorieau et al. assumiram tal como alterações conformacionais transientes, as quais possuem desvios químicos de ^{15}N semelhantes ao

mutante G8A, que encontra-se maioritariamente numa estrutura de *hairpin* aberta. Utilizando esse mutante como modelo, obtiveram uma estrutura, a pH baixo, em forma de “L” e uma estendida, sugerindo que a primeira era apenas um estado intermediário para a estrutura estendida. Segundo Lorieau et al. esta estrutura estendida já possui capacidade de atravessar a membrana e assim potenciar a formação de poros na membrana [Lorieau 2012].

As transições acima descritas são muito lentas para os intervalos de tempo simulados normalmente em dinâmica molecular, mas a corrida de diversas simulações diferentes, facilitou a diversidade da amostragem. Durante a simulação usando o campo de forças MARTINI, um lípido colocou-se entre as duas hélices do *hairpin* na zona dos terminais, tendo-se aí mantido até a transformação reversa. Uma vez que já existia a destabilização inicial do *hairpin*, devido ao afastamento dos terminais, as duas hélices acabaram por se afastar, formando duas hélices perfeitas num estado aberto em forma de “L” (figura 49). Note-se que mesmo que esta estrutura tenha sido formada devido à destabilização inicial que ocorreu na simulação CG, a estrutura afastou-se e manteve-se estável na simulação UA, mostrando que o péptido de fusão pode manter a estrutura em hélice mesmo sem o contacto entre hélices. Curiosamente e de acordo com o relatado acima, este evento ocorreu numa simulação do péptido protonado (pH baixo).

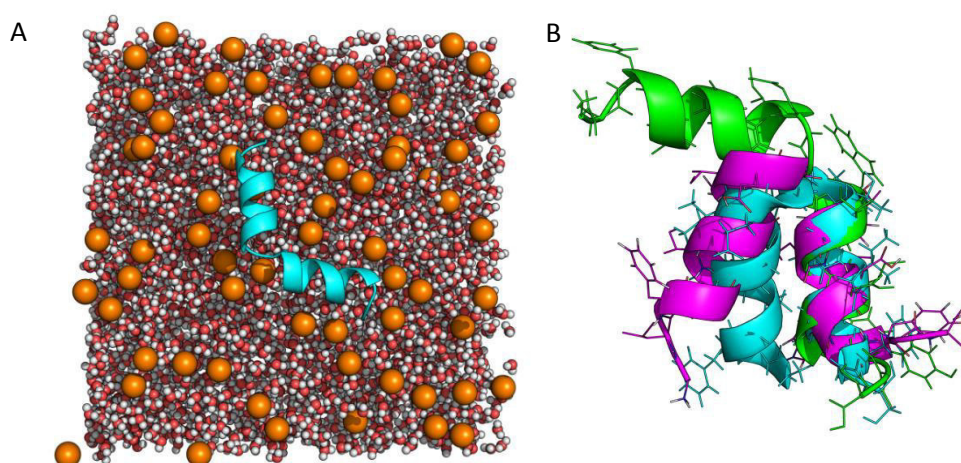


Figura 49 – Estrutura do péptido de fusão em forma de “L” (A) inserido na membrana, visualizado com o ponto de vista do interior da membrana. (B) comparação entre a estrutura de NMR a ciano, a estrutura inicial da simulação *united-atom* a magenta, e estrutura final da mesma simulação a verde. Todos os péptidos estão representados por cartoon e as cadeias laterais por linhas, as águas representadas em esferas onde os oxigénios estão em vermelho e os hidrogénios em branco, os lípidos apenas estão representados pelo fósforo em esferas laranjas.

Formação de poro

Num dos nove replicados do sistema PB-D, sistema com a membrana pré-equilibrada e as cadeias laterais desprotonadas, formou-se um poro, ou seja, existem águas a preencher um canal na membrana, sendo possível para as águas passarem de um lado para o outro lado da membrana. A estrutura do sistema aos 500 ns encontra-se na figura 50.

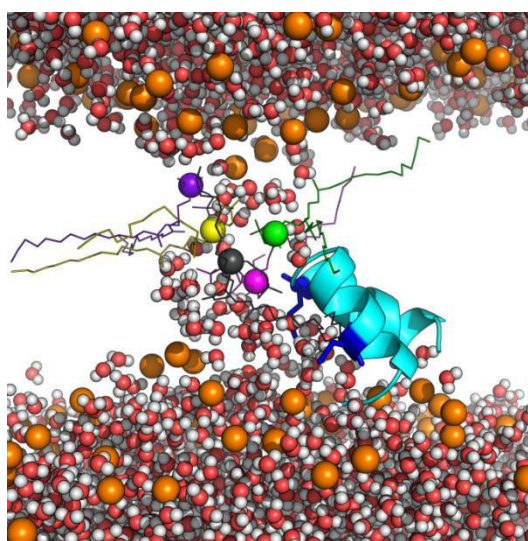


Figura 50 – Estrutura final da simulação onde se formo um poro na membrana. O péptido de fusão está representado em *cartoon* a ciano, com os resíduos ácidos a azul-escuro. As águas representadas como esferas vermelhas e brancas para os oxigénios e os hidrogénios, respetivamente. Os lípidos cuja cabeça se encontra no interior da membrana estão representados por linhas e esferas de diferentes cores, para os restantes apenas é representado o fósforo com esferas laranja.

Através da observação da estrutura (figura 50) ao fim da simulação nota-se que o péptido se encontra inclinado em direção ao interior da membrana e do poro, com os resíduos ionizáveis direcionados para o interior deste. Dentro do poro apenas se encontram cinco fósforos, sendo que as cadeias laterais desses resíduos interagem maioritariamente com água. Uma análise pormenorizada da simulação completa mostra uma rápida inclinação do péptido e a entrada das moléculas de água, que praticamente acompanham a penetração dos resíduos ionizáveis. Passadas algumas dezenas de nanossegundos, existem algumas águas a entrar no folheto oposto onde o péptido está inserido e interagem com o Trp14 (e em menor extensão a Phe9) que se encontra virado

para o interior da membrana. A entrada das primeiras águas, muito provavelmente, foi o catalisador da entrada de 1-3 cabeças de lípidos no interior da camada superior. Causando, por consequente, a entrada de um maior número de águas até se formar, por fim, o poro. A formação de um poro estável demorou cerca de 150 ns desde a entrada dos primeiros lípidos na membrana.

A forma como a membrana foi formada antes do início da simulação é provavelmente irrelevante. O mais surpreendente é o péptido que mimetiza o pH neutro (não fusogénico) resultar num poro e nenhuma das simulações do sistema protonado formar um. Contudo existem outras simulações onde existe água dentro do plano da membrana, incluindo uma simulação do sistema SA-P (figura 51) onde o péptido se encontra com a mesma inclinação e águas dentro da camada inferior como no primeiro passo de formação do poro explicado anteriormente. Será interessante continuar esta simulação por mais tempo, para averiguar se será possível a formação de um poro.

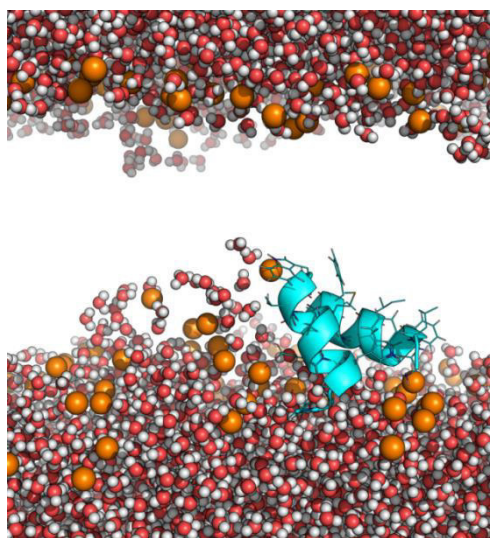


Figura 51 – Representação de um péptido de fusão na membrana com algumas águas e lípidos quase a meio da membrana, sugerindo a formação de um possível poro. onde a azul e tubos está representado o péptido de fusão As águas estão representadas por esferas vermelhas e brancas para os oxigénios e os hidrogénios, respetivamente. Os lípidos são representados pelo fósforo em esferas laranja.

A formação do poro, provavelmente, deve-se a vários fatores, portanto a capacidade do péptido desprotonado ter formado um poro e o péptido protonado pode apenas se dever à falta de amostragem. Por outro lado aqui apenas foi simulado um péptido, mas a hemaglutinina por si possui três segmentos do péptido fusogénico. Muito

provavelmente, um maior número de péptidos, dispendo-se ao redor do poro formado, facilitaria a sua formação.

A obtenção de uma simulação com a formação de um poro, sem qualquer fator externo para que tal ocorra, vai de certo modo contra o estipulado por Lorieau et. al [Lorieau 2012]. Aí é sugerido que o estado ativado do péptido de fusão é uma estrutura estendida, uma vez que a estrutura em *hairpin* não é suficiente para atravessar por completo a membrana. Contudo os nossos resultados sugerem que não é necessário os péptidos de fusão revestirem o poro, mas precisam apenas de destabilizar a membrana.

3.2. Péptido transmembranar

No âmbito do trabalho foram estudados dois tipos de péptidos transmembranares (PT) diferentes: O subtipo H5 para comparar com estudos realizados por outro membro do grupo de modelação de proteínas do ITQB, mas utilizando uma metodologia diferente; e o subtipo H3, que será também utilizado, posteriormente, em junção com a estrutura cristalográfica da hemaglutinina.

O C-terminal da HA não estava presente quando se obteve a estrutura cristalográfica da H3N2, sendo assim desconhecida a sua estrutura. Essa sequência que será aqui designada como péptido transmembranar contém a cauda citoplasmática (CC) no C-terminal do péptido, o domínio transmembranar (resíduos inseridos na membrana) e a junção entre este último e o ectodomínio da hemaglutinina no N-terminal do péptido. Uma figura ilustrativa dos diferentes segmentos foi esquematizada na figura 10.

A informação de que resíduos constituem o domínio transmembranar é importante, tanto pela sua importância fundamental, como também para se simular apenas esse domínio. Outro fator importante é a estrutura secundária do segmento, visto que, ao usar o campo de forças MARTINI é necessário definir previamente a estrutura secundária. De forma a abordar estas duas questões, foram utilizadas várias ferramentas bioinformáticas diferentes, para além dos resíduos do domínio transmembranar sugeridos na bibliografia.

3.2.1. Construção do péptido transmembranar

Péptido do subtipo H5

Este péptido de 27 resíduos (sequência: ILSIYSTVASSLALAIMVAGLSLWMCS) foi utilizado num trabalho realizado por um membro do grupo de modelação molecular [Victor 2012], sendo a estrutura gerada através do programa PyMOL, uma vez que não existe estrutura 3D. Neste trabalho a única diferença foi a alteração da resolução para *course-grained* com vista a simular esse sistema em MARTINI. A estrutura secundária em MARTINI foi definida como em hélice, exceto os resíduos das pontas, que foram definidos como estrutura desordenada.

Péptido do subtipo H3

Para a H3N2 foram estudados dois péptidos, um com a sequência completa do C-terminal da HA não contabilizada na estrutura 3D da proteína, ou seja, os resíduos 521-566. Um outro péptido será apenas constituído pelos resíduos que deverão estar inseridos na membrana (domínio transmembranar). Para construir ambos utilizou-se a sequência: VELKSGYKDWILWISFAISCFLLCVLLGFIMWACQQRGNIRCNICI, que foi analisada para inferir que resíduos tinham uma maior probabilidade de fazer parte do domínio transmembranar, sendo o resto a cauda citoplasmática e a junção entre o ectodomínio e o PT. Para isso usaram-se diferentes métodos, descritos adiante. A numeração para o péptido transmembranar inicia-se na valina na posição 1 (Val1) e termina na isoleucina da posição 46 (Ile46).

1) **RHYTHM** é um servidor que utiliza inicialmente a ferramenta HMMTOP, onde a topologia das proteínas membranares é determinada pela composição e distribuição de aminoácidos das várias partes estruturais [Tusnady 1998], para prever a estrutura secundária, ou seja, que resíduos compõem a hélice transmembranar. Seguidamente, utiliza bases de dados com informações estruturais e informações de sequências para prever, na estrutura terciária, os contactos entre hélices ou com lípidos [Rose 2009]. O método previu hélice para o segmento Ile11-Cys35.

2) **Clustal omega** é um complexo programa de sequenciamento múltiplo, que utiliza “árvores guia” para construir vetores para cada sequência que serão posteriormente alinhados utilizando perfis de *hidden markov models* [Sievers 2011].

```

H5N1      -----ILSIYSTVASSLALAIMVAGLSLWMC----- 27
H3N2      VELKSGYKDWILWISFAISCFLLCVLLGFIMWACQQRGNIRCNICI 46
          : * . * . : * . : : * : : * * .

```

Figura 52 – Resultado da ferramenta de bioinformática online Clustal ómega.

3) **Phyre** (*protein homology/analogy recognition engine*) é um servidor que utiliza algoritmos de correspondências perfil-perfil. De uma forma geral utiliza uma biblioteca de proteínas com estrutura conhecida, retirada da base de dados, das quais serão utilizadas as sequências para serem comparadas com outra base de dados de sequências não redundante. Esse perfil é analisado posteriormente por três programas independentes, prevendo para cada um dos resíduos da sequência, uma das três estruturas secundárias: hélice (h), folha beta (e) e desordenado (c) [Kelley 2011].

Índice10.....2.....3.....4.....
Sequência	VELKSGYKDWI LWI SFAI SCFLLCVLLGFIMWACQQRNI RCNICI
psipred	ccccch hhhHhhh hhhhhhhhhhhhhhhhhhh h hh cccceeee
jnet	ccccce eeE ee eehhhhhhhhhhhhhhhhh h hh cccceeee
sspro	ccccch hhhHhhh hhhhhhhhhhhhhhhhh h hh cccceeee
Consensus	ccccch hhhHhhh hhhhhhhhhhhhhhhhh h hh cccceeee
Cons_prob	85677466666 666 6788789999999999999 8758996677766

Figura 53 – Resultado da ferramenta de bioinformática online Phyre.

A sequência do Trp10 até ao Gln36 é sugerida como o domínio transmembranar [Armstrong 2000] e de facto obtemos os mesmos resultados com a ferramenta RYTHM. Por outro lado o resultado da ferramenta clustal omega mostra um alinhamento com grandes semelhanças entre o domínio simulado por Victor et al. [Victor 2012]. Portanto o segmento simulado como domínio transmembranar será o segmento Trp10-Gln36.

A nível da estrutura secundária obteve-se um grande conteúdo em hélice através do servidor Phyre, sendo o maior consenso desde a Gly6 até à Gln36. Contudo o valor de confiança obtido para os resíduos Gly6-Ile14 e para a Gln36 são, respetivamente 6 e 5, dando uma menor confiança na predição da estrutura em hélice desses resíduos. Visto que consideramos o domínio transmembranar como Trp10-Gln36, nas simulações

do campo de forças MARTINI, iremos definir esses resíduos como em hélice, exceto os terminais Trp10 e Gln36 (de forma semelhante ao utilizado por Victor et al. [Victor 2012]).

Para as simulações com o péptido completo, a estrutura para o segmento de junção (Val1-Asp9) será simulado em MARTINI como desordenado, uma vez que apresenta estrutura estendida flexível, algo que também ocorre em outras proteínas de fusão [Li 2008 & Liao 2006]. A cauda citoplasmática (Arg37-Ile46) foi também simulada como estrutura desordenada.

A orientação do péptido já era conhecida, ou seja, N-terminal no lado extracelular onde o ectodomínio da hemaglutinina se encontra, e o C-terminal no lado do citoplasma. Isto foi corretamente reproduzida pela ferramenta RHYTHM. Contudo nas nossas simulações não existe diferença entre os dois lados da membrana, visto que, as caixas de simulação apenas são constituídas pelos péptidos, os lípidos e água.

3.2.2. *Self-assembly* do péptido transmembranar

Para se obter a estrutura dos péptidos inseridos numa bicamada sem qualquer manipulação optou-se pelo protocolo de *self-assembly*, onde três péptidos são colocados numa caixa de simulação juntamente com 128 lípidos e 2000 águas MARTINI.

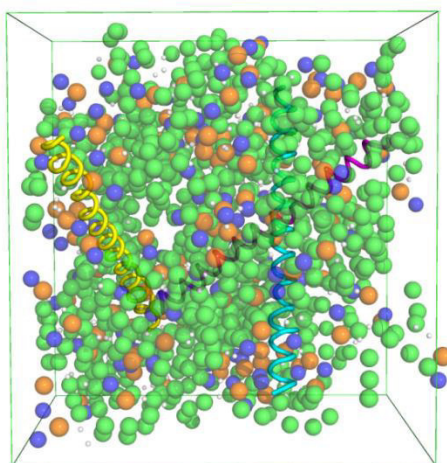


Figura 54 – Arranjo inicial de um dos sistemas. Os péptidos estão representados por tubos a ciano, amarelo e magenta, os lípidos estão representados também por esferas, onde a laranja são representados o grupo fosfato, a azul o grupo ciano, a branco os grupos de glicerol e verde os grupos da cauda. As águas não foram apresentadas para melhor visualização.

Após 25 ns, foram selecionadas as simulações que formaram uma bicamada nos eixos dos x e y, com os três péptidos inseridos na membrana e com a mesma orientação. Cada um dos diferentes péptidos, isto é, o domínio transmembranar do subtipo H5, o domínio transmembranar do subtipo H3 e o segmento C-terminal completo da H3, foram simulados como triplicados por 10 μ s. Uma vez que não há evidências de o pH ter algum efeito direto no péptido transmembranar, as simulações foram feitas com os resíduos ácidos e básicos (Glu2, Lys4, Lys8, Asp9, Arg37 e Arg41) carregados, mimetizando um pH neutro. Tanto o C-terminal como o N-terminal nos péptidos truncados (apenas domínio transmembranar) foram simulados sem carga para mimetizar o facto de estarem ligados ao resto da proteína. O mesmo raciocínio foi utilizado para o N-terminal do segmento completo, embora o C-terminal deste, ou seja o último resíduo da proteína, encontra-se carregado, como ocorreria *in vivo* a pH neutro. O resumo dos sistemas simulados para o self-assembly do péptido transmembranar encontra-se na tabela 7.

Tabela 7 – Especificações dos diferentes sistemas em *coarse-grain* utilizados no processo de *self-assembly*.

Sistema	Número de resíduos de aminoácidos	Setup	Carga do N-terminal	Carga do C-terminal	Tempo de simulação (μ s)
H5_27	27	Sequência do segmento transmembranar	Neutra	Neutra	10
H3_27				Negativa	
H3_46	46	Sequência completa			

A formação da bicamada lipídica ocorre muito rapidamente, por volta dos 25 ns, mas nestes tempos os péptidos ainda se encontram distantes, necessitando portanto, de um maior tempo de simulação. No primeiro microssegundo os péptidos já se encontram empacotados e estáveis, mas apresentam distribuições dos péptidos diferentes. Observaram-se dois tipos de arranjo entre os péptidos: num deles os péptidos assumiam uma distribuição triangular semelhante à distribuição da *coiled-coil* central do ectodomínio da hemaglutinina, noutro os terminais dos péptidos encontram-se alinhados linearmente. Embora este último arranjo seja minoritário apareceu em

todos os sistemas diferentes, sugerindo ser uma conformação estável mas não maioritária. Por essa razão decidiu-se aumentar o tempo de simulação até aos 10 μ s. Embora uma das três simulações do sistema H3_46, (representados na figura 55) não tenha evoluído para o arranjo triangular, decidiu-se iniciar a transformação reversa na mesma. Desta forma podemos também ter uma maior variação conformacional, nas simulações utilizando o campo de forças 54A7.

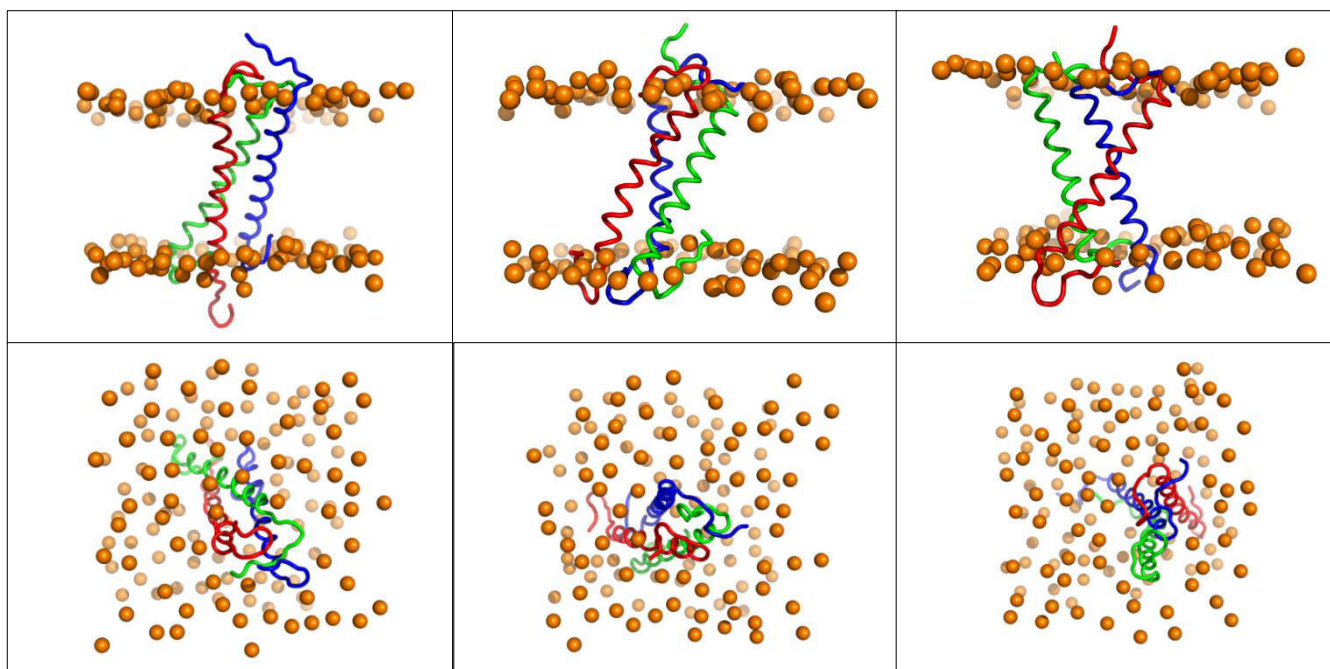


Figura 55 – Conformações finais para cada replicado do sistema H3_46 em MARTINI, onde os péptidos estão representados como tubos e os fósforos da bicamada como esferas laranjas. As águas e as restantes partículas da bicamada foram ocultadas para melhor visualização. As imagens de cima e as de baixo representam, respetivamente, o ponto de vista por dentro da membrana e o ponto de vista por cima da membrana.

A conformação final de cada um dos três replicados do sistema H3_46 (réplicas CG), ou seja, aos 10 μ s, encontra numa resolução *coarse-grain* a qual foi reconstruída para uma resolução *united-atom*, através do método de transformação reversa (figura 56).

Não se continuou o estudo do sistema H5_27 para *united-atom*, porque este último campo de forças apresenta um maior custo computacional e ambos os sistemas parecem ter, aproximadamente, os mesmos comportamentos estruturais. O péptido transmembranar da estirpe H5 foi também estudada anteriormente por Victor et al., embora sem utilizar o *self-assembly* [Victor 2012]. O estudo do sistema H3_27 também não será continuado, uma vez que este foi simulado em *coarse-grain*

basicamente como um intermédio de comparação entre o sistema H3_46 e H5_27, por ter o mesmo comprimento que o sistema H5_27 e ser da estirpe que o sistema H3_46.

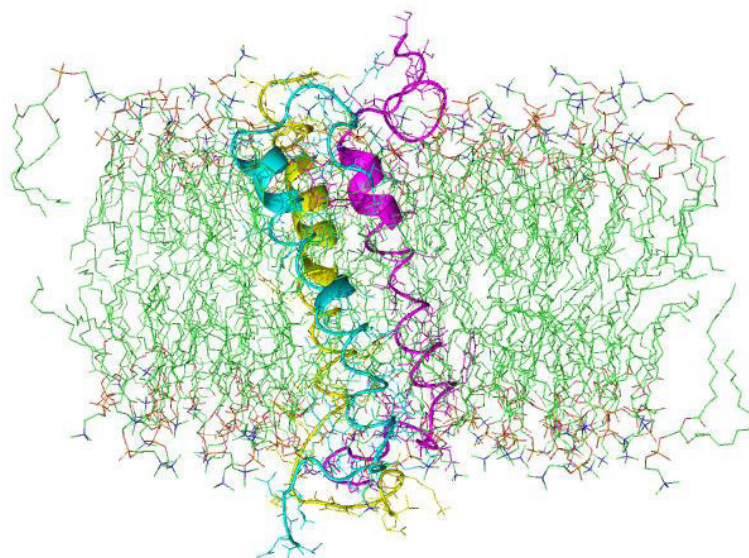


Figura 56 – Estrutura de uma das réplicas após transformação reversa. Lípidos definidos como linhas e os péptidos a ciano, magenta e amarelo com representação *cartoon*.

Como se pode verificar na figura 51 e também pela análise da ferramenta DSSP, existem algumas zonas, maioritariamente a zona N-terminal do segmento helical, que não se encontram em conformação de hélice. O que seria de prever devido aos problemas conhecidos da transformação reversa e do mesmo ter ocorrido no estudo do péptido de fusão. No entanto, pode-se também observar na figura, que mesmo não sendo uma hélice perfeita apresenta uma estrutura helical, e após alguns nanossegundos de simulação, a estrutura em hélice foi restabelecida, pelo menos parcialmente.

Cada uma das três conformações obtidas pela transformação reversa foi simulada em triplicado durante 500ns usando o campo de forças 54A7. Com isso fizeram-se 9 réplicas em *united-atom* (réplicas UA) e 4.5 μ s de tempo de simulação total.

3.2.3 Análises

3.2.3.1 - Conformação

Após 500ns de simulação, tempo em que se considera os sistemas já equilibrados, foram retiradas imagens mostrando a conformação final de todas as simulações UA realizadas, estando representadas na figura 57.

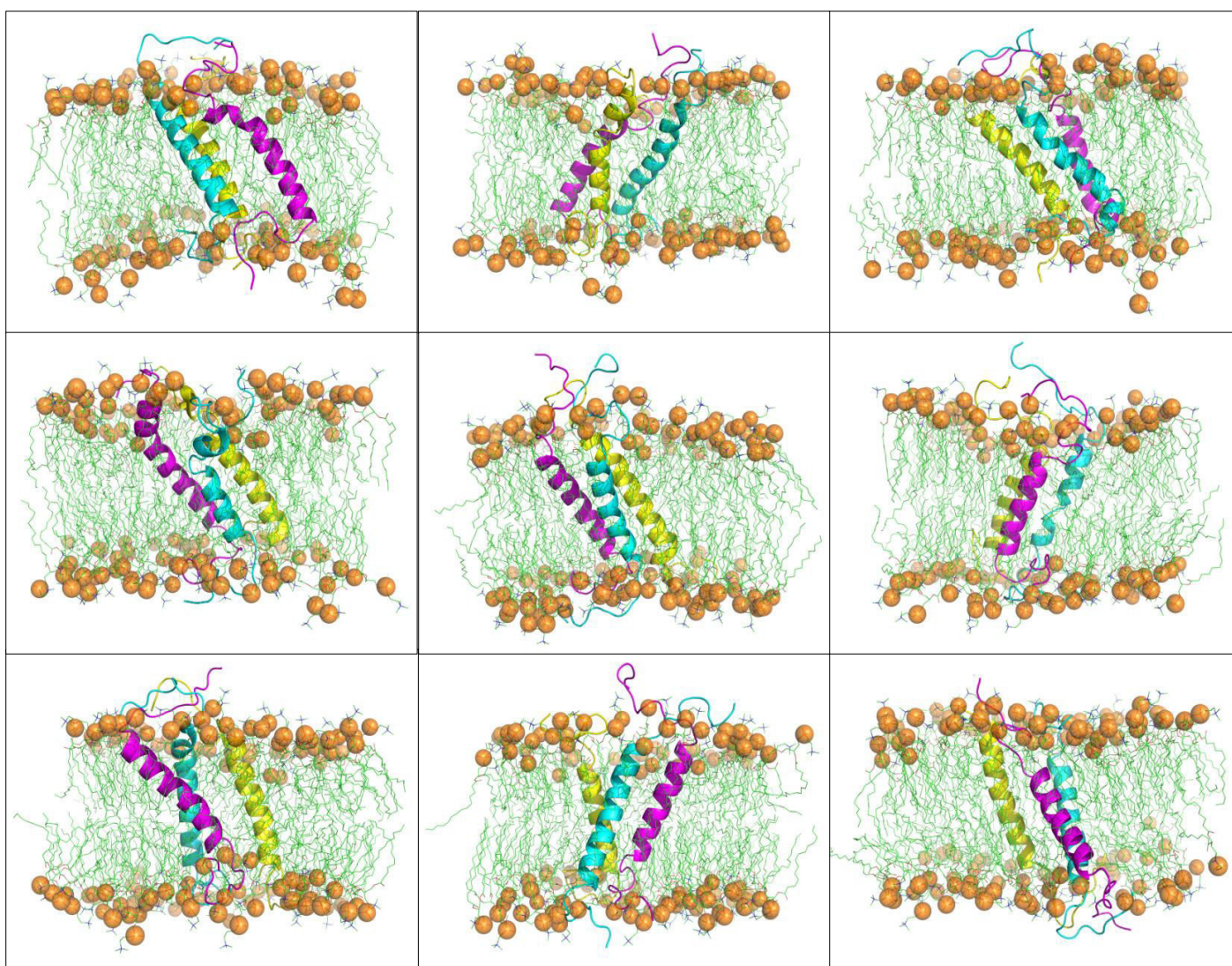


Figura 57 – Conformações finais de cada uma das simulações usando o campo de forças 54A7. Cada um dos peptídeos é representado a ciano, amarelo e violeta ambos definidos como *cartoon*, os lípidos como linhas verdes e os fósforos representados como esferas laranjas.

Ao analisarem-se todas as conformações finais dos peptídeos nas diferentes réplicas observa-se que estas são bastante semelhantes, apresentando sempre uma

interação direta entre péptidos, maioritariamente numa estrutura de *coiled coil*. Os péptidos parecem também ter um conteúdo em hélice e ângulo de inclinação semelhantes. Nota-se ainda uma preferência acrescida por parte da cauda citoplasmática para interagir com a membrana, em comparação com o segmento de junção ao ectodomínio. Tal corrobora estudos de dicroísmo circular e FTIR com um péptido sintético de 30 aminoácidos da zona transmembranar, que mostravam uma tendência para esse péptido estar maioritariamente em hélice e formar *coiled coils* numa membrana de DMPC [Tatulian 2000].

O empacotamento dos péptidos por nós obtido é superior ao obtido por Victor et al. [Victor 2012] em que os péptidos apenas se aproximavam. Isto pode-se dever ao uso de diferentes subtipos de hemaglutinina, à diferença dos números de resíduos e/ou ao protocolo de montagem dos péptidos, uma vez que não usou o campo de forças MARTINI. Victor et al. inseriu manualmente os péptidos na membrana paralelos à normal desta e afastados entre si. Em contraste, no nosso estudo, o *self-assembling* permite uma maior amostragem do espaço e uma difusão das partículas também superior. No nosso estudo, comparando o empacotamento obtido ao utilizarem-se os diferentes campos de força, foi notório um relaxamento e afastamento entre os péptidos ao utilizar o campo de forças 54A7 comparativamente ao MARTINI, o qual se deve às diferentes propriedades destes dois campos de força.

Curiosamente o rearranjo entre os péptidos apresenta uma disposição triangular semelhante à estrutura trimérica da hemaglutinina, como se pode observar pela, figura 58. Tal sugere que, os péptidos por si só, são capazes de se organizar, mesmo sem a presença do ectodomínio da hemaglutinina. Embora haja, no entanto, a possibilidade da presença do ectodomínio poder alterar a estrutura por nós obtida.

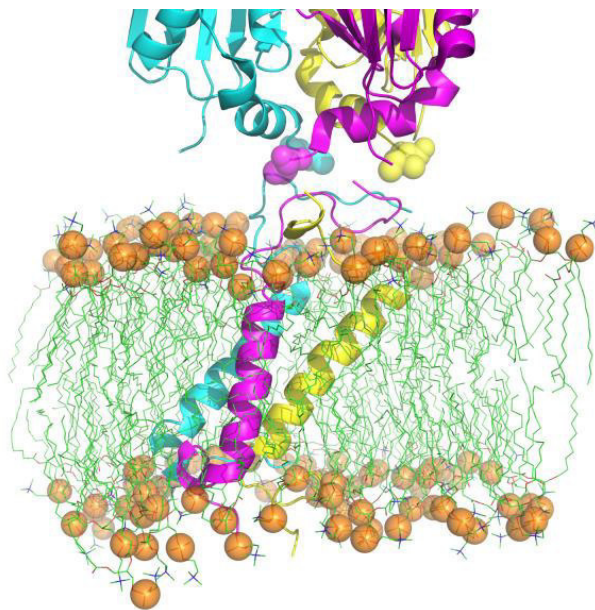


Figura 58 – Parte inferior da estrutura cristalográfica da hemaglutinina sobreposta à conformação final de uma das simulações dos péptidos transmembranares. As esferas transparentes na estrutura da hemaglutinina representam o último resíduo de aminoácido que se liga ao péptido transmembranar. As estruturas foram sobrepostas manualmente utilizando o programa PyMOL.

3.2.3.2. Conteúdo em hélice

De forma análoga ao feito para o péptido de fusão, também aqui, para uma melhor percepção dos resultados obtidos pelo DSSP, foi analisada a percentagem de tempo que cada resíduo passou em hélice (figura 59). Foram utilizados os últimos 200ns de simulação, por se considerar o sistema já equilibrado aos 300ns, segundo algumas análises preliminares.

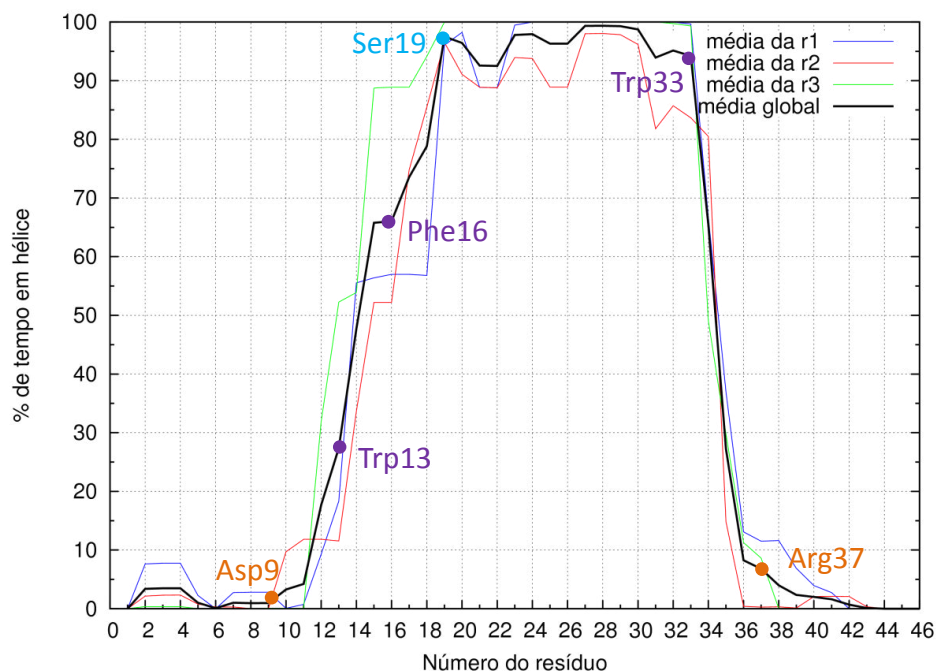


Figura 59 – Percentagem de tempo médio que cada resíduo apresentou uma conformação de estrutura em hélice ao longo dos últimos 200ns da simulação.

Os resultados apresentam uma variabilidade significativa, não só entre os 3 péptidos dentro de uma única simulação, como também dos péptidos entre as outras (oito) simulações. Mas uma vez que estamos a estudar o mesmo sistema e cada péptido encontra-se num ambiente semelhante aos restantes, podemos calcular a média da percentagem em hélice de cada réplica inicial que vieram das simulações MARTINI (três réplicas r1, r2 e r3 que por sua vez foram triplicadas em 54A7).

Mais uma vez, como foi explicado na análise dos resultados do péptido de fusão, iremos tratar as simulações com a mesma denominação. Por exemplo, na análise das simulações *united-atom*, quando referirmos a réplica r1 ou a média da réplica r1 estaremos a referir todas as três simulações *united-atom* que provieram da primeira réplica simulada em *coarse-grain*, e a média de todos os péptidos transmembranares correspondentes a essas simulações, respetivamente.

A conformação em hélice na parte N-terminal do segmento transmembranar é menos estável do que a parte C-terminal, pois apresenta nos últimos resíduos um conteúdo em hélice intermédio. Em contraste, na outra ponta do domínio transmembranar, a percentagem de tempo em hélice diminui de forma rápida, contínua e reproduzível ao longo dos últimos resíduos. Uma conclusão semelhante foi obtida no

trabalho realizado por Victor et al. [Victor 2012] e também anteriormente sugerido num estudo experimental [Tatulian 2000]. Isto pode-se dever à maior concentração de resíduos polares na metade do N-terminal do domínio transmembranar. Os quais, por sua vez, são bastante conservados entre os diferentes subtipos da hemaglutinina. Portanto, para além do papel que as suas cargas possam ter na estrutura ou no processo de fusão, possivelmente, uma maior desorganização estrutural no N-terminal pode ter também algum papel na fusão. Talvez nos últimos estágios na fusão, aquando da interação do péptido de fusão com o péptido transmembranar.

Se considerarmos em hélice os resíduos do péptido transmembranar que possuem uma percentagem superior a 50% de tempo nesta estrutura secundária, podemos verificar que o péptido transmembranar possui cerca de 19 resíduos em hélice. O segmento helical é composto pelos resíduos compreendidos entre a Ser15 e a Ala34, com os resíduos aromáticos Phe16 e Trp33 a flanquear cada uma das pontas do segmento.

Ao analisar o número médio de resíduos em hélice ao longo do tempo, para cada um dos replicados CG ou a média global/total, obtém-se um número semelhante ao sugerido pela análise acima descrita. Observa-se uma flutuação entre os 19 e 20 resíduos em hélice, sendo a média dos últimos 200ns de 19.64 ± 0.23 resíduos em hélice.

Se considerarmos apenas o domínio transmembranar (proposto anteriormente), ou seja, apenas 30 resíduos, obtemos uma percentagem aproximada dos resíduos em hélice de 65.5%. O valor estará ligeiramente inflacionado mas tanto os resíduos da cauda citoplasmática como da junção péptido-ectodomínio possuem <5% em hélice. Este valor é próximo do obtido experimentalmente por Tatulian et al. [Tatulian 2000] de 62% usando um péptido de 30 resíduos e bicamada composta só por DMPC. Principalmente porque o nosso valor está sobrestimado, devido à consideração da formação de hélice fora desses 30 resíduos (embora inferior a 5%). Curiosamente o aumento de cargas negativas com a adição de DMPG aumenta a percentagem de resíduos em hélice, tendo sido obtido valores de 79% para 40% de DMPG, 88% para 60%DMPG [Tatulian 2000] e $64\% \pm 1$ para 50% de DMPG e péptido de 33 resíduos [Chang 2008]. Tal está

provavelmente relacionado com a interação dos grupos carregados positivamente, que flanqueiam o domínio transmembranar, e as cargas negativas desses lípidos.

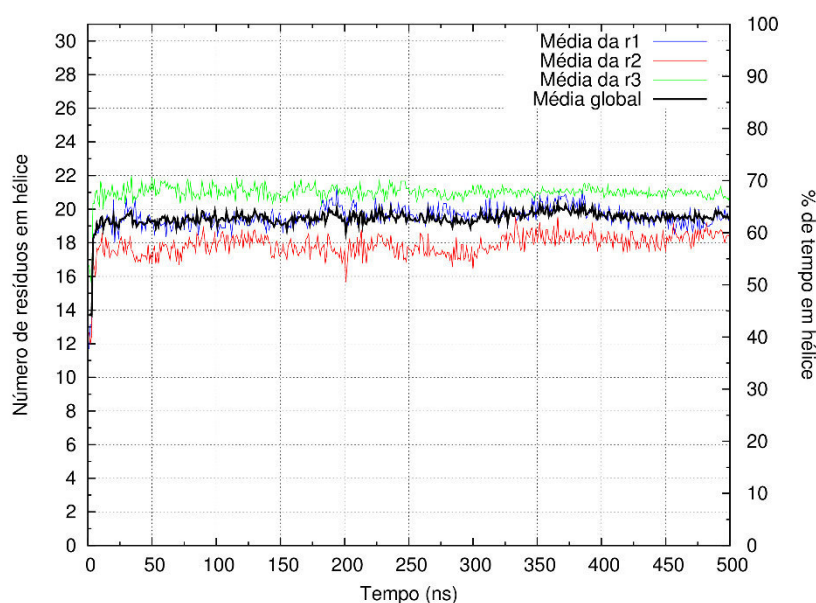


Figura 60 – Número médio e porcentagem de resíduos em hélice ao longo de todo o tempo de simulação, para a média de cada replicado CG (r1, r2 e r3) e para a média de todos os péptidos, média global.

Para além do número e porcentagem de resíduos em hélice foi ainda realizada uma análise suplementar para averiguar a frequência do número de resíduos em hélice (figura 61). Curiosamente o número de resíduos em hélice com maior frequência é por volta dos 22 resíduos, o que é reproduzido na média das 6 simulações oriundas de 2 dos 3 sistemas iniciais após transformação reversa (réplicas r1 e r3). Na média das outras 3 simulações (réplicas r2) os 22 resíduos é apenas um máximo local. Contudo o máximo da média dessas mesmas simulações da réplica r2, por volta dos 17-18 resíduos, é igual ao segundo máximo de frequência da média de todas as réplicas (média global).

O valor do segundo pico, ou seja, aproximadamente 17 resíduos em hélice corrobora com o valor médio de 17 resíduos em hélice obtido por Victor et al. ao usar um péptido transmembranar de 27 resíduos do subtipo H5 [Victor 2012]. Neste caso a diferença nas médias não se deve ao protocolo utilizado, uma vez que o conteúdo em hélice é semelhante tanto para os sistemas com três péptidos como com apenas um péptido, mas provavelmente deve-se à diferença no número de resíduos. O valor de 17 resíduos já tinha sido mostrado ser o número mínimo de resíduos do domínio transmembranar necessário para que ocorra uma fusão eficiente. Foi sugerido que este

valor é o comprimento mínimo para que o domínio transmembrantar atravesse de um lado ao outro da membrana [Armstrong 2000].

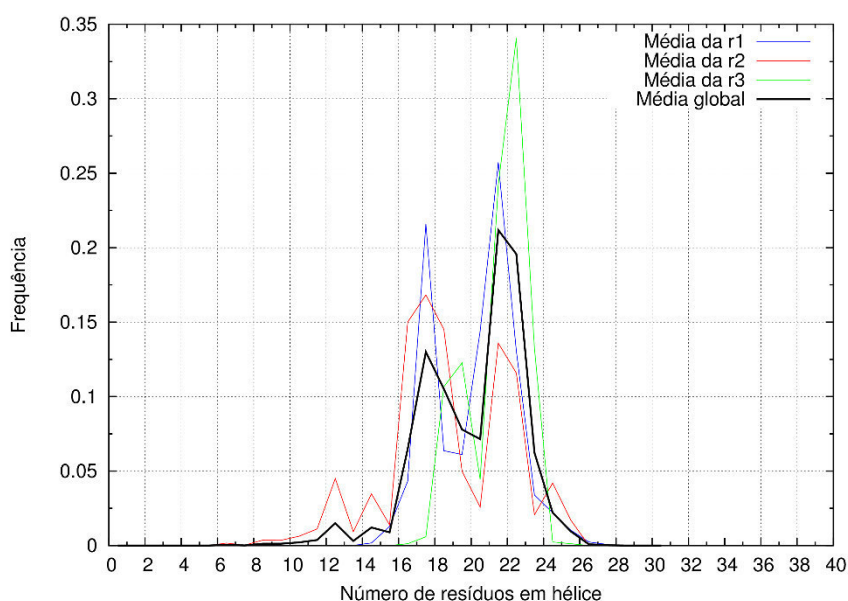


Figura 61 – Histograma de frequência do número médio de resíduos em hélice calculado para os últimos 200ns de cada simulação, para a média de cada replicado CG (r1, r2 e r3) e para a média de todos os péptidos, média global.

3.2.3.3 Ângulo de inclinação

O ângulo de inclinação do péptido transmembrantar em relação ao plano da membrana foi calculado de forma semelhante ao calculado para o péptido de fusão. Contudo, neste caso foram selecionados os resíduos 19-22 e 24-26 por apresentarem uma estrutura em hélice superior a 90% do tempo analisado (figura 59) e encontram-se no meio do péptido.

Na imagem 62 estão representados os ângulos, ao longo do tempo, de cada péptido de todas as simulações e na imagem A-B estão representados os ângulos médios dos três péptidos e de todos os replicados ou apenas dos replicados UA provindos de cada um dos replicados CG.

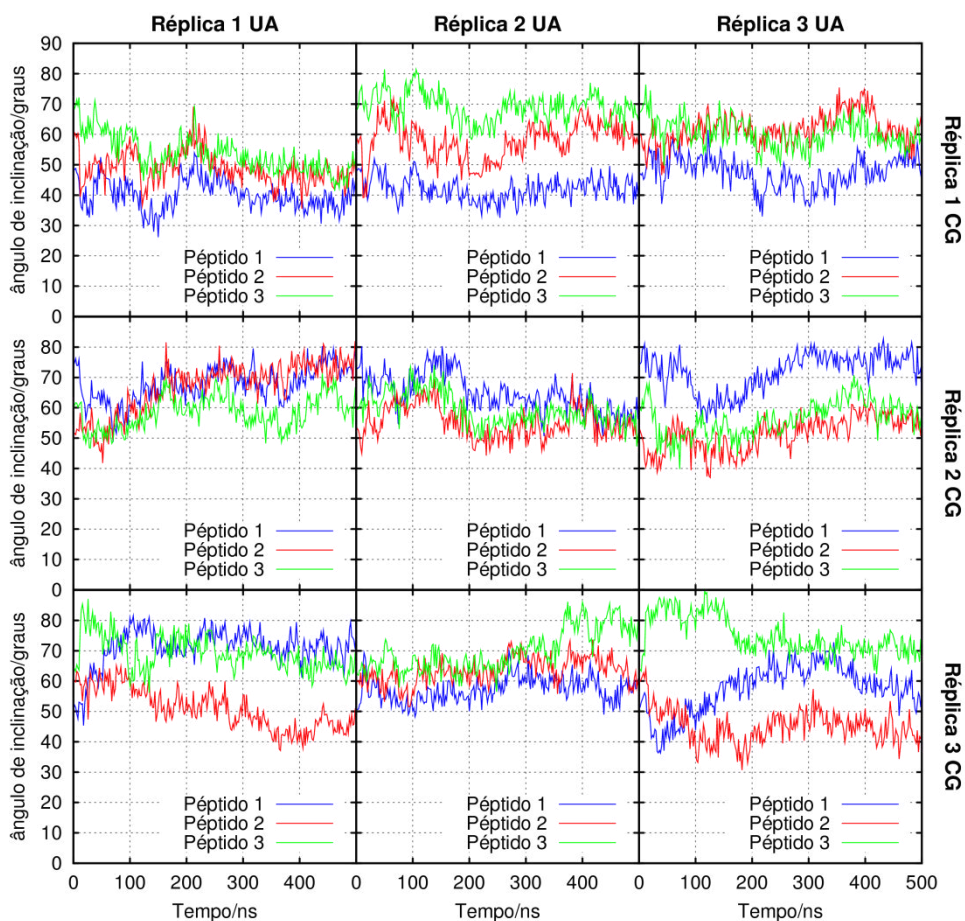


Figura 62 – Representação do ângulo de inclinação, ao longo do tempo, dos péptidos transmembranares para todas as simulações.

Através da análise da figura 62 observa-se que o péptido transmembranar apresenta um ângulo de inclinação bastante variável, podendo ir de um ângulo de 30 até 90 graus (perpendicular à membrana). Tal variação observa-se mesmo dentro de cada simulação em particular, mostrando uma grande dinâmica, não obstante os péptidos se encontrarem numa forma compacta entre os péptidos adjacentes e os lípidos. Contudo, nos últimos 200 ns da simulação, a maioria dos péptidos parecem estar estabilizados, mas com ângulos diferentes entre si. Tal variação já tinha sido antes reportada, observando-se que o ângulo de inclinação variava com a concentração de péptido e metodologia, podendo abranger um intervalo entre 50° e 90° [Tatulian 2000].

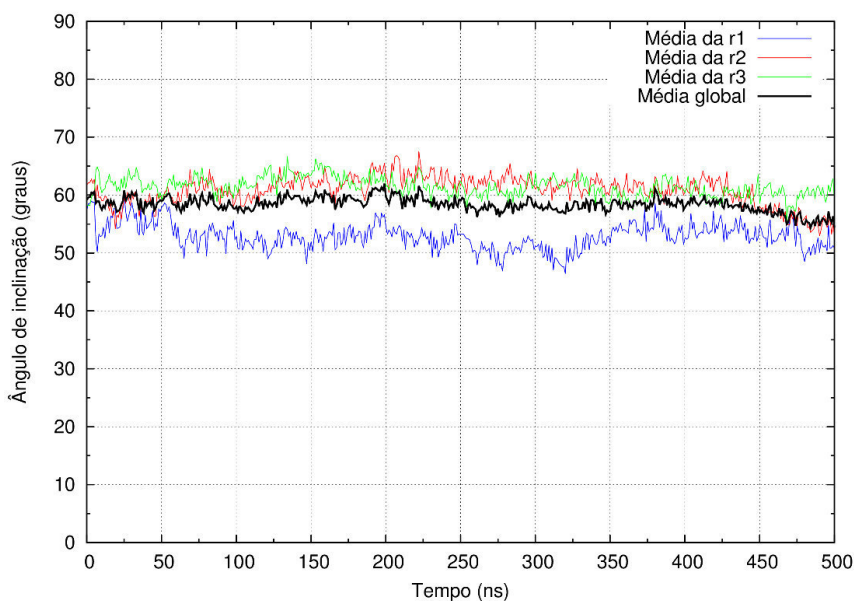


Figura 63 – Representação do ângulo de inclinação médio, ao longo do tempo, para a média de cada replicado CG (r1, r2 e r3) e para a média de todos os péptidos, média global.

A média do ângulo de inclinação calculada para todos os péptidos das nove réplicas varia, ao longo do tempo, em redor dos 60 graus, apresentando uma ligeira diminuição nos últimos nanossegundos devido às simulações da réplica r2. O valor do ângulo de inclinação médio dos últimos 200 ns é de $58.4^{\circ} \pm 1.25$, semelhante ao valor obtido por Chang et al. ($56^{\circ} \pm 1$) considerando o desvio padrão e dentro do intervalo obtido por Tatulian et al. ($53-62^{\circ}$) [Chang 2008, Tatulian 2000]. Mesmo sendo o valor obtido por nós um pouco superior ao obtido por Chang et al. observa-se, nos últimos 50 ns, uma descida do ângulo médio devido à réplica r2 que diminui até por volta dos 55 ns, onde parece ter estabilizado, sugerindo que houve uma alteração no péptido e/ou na membrana, encontrando um ângulo mais favorável.

Victor et al. obteve um ângulo de inclinação para a simulação de um único péptido em redor dos 60 graus e um intervalo de 70 a 80 graus nas simulações com três péptidos [Victor 2012]. O valor de 60 graus está em concordância com o obtido por nós, contudo este valor foi para um péptido e segundo estudo de Tatulian et al. a concentração do péptido influencia o ângulo do péptido [Tatulian 2000]. O valor obtido por Victor et al. com concentrações iguais às que estamos a estudar é superior em pelo menos 10 graus. Esta diferença pode-se dever tanto ao uso de péptidos de diferentes estirpes, ao diferente tamanho dos péptidos (27 e 46 resíduos) e o facto de o self-

assembling ter sido feito com *coarse-grain*, o que permite uma proximidade superior entre os péptidos transmembranares.

É de notar que as nossas simulações e os dados experimentais foram obtidos sem a presença do ectodomínio da hemaglutinina, o que pode influenciar tanto a dinâmica do péptido como o valor do ângulo em si. Não existem dados experimentais sobre o péptido transmembranar com a presença do ectodomínio da hemaglutinina.

Utilizando os ângulos de inclinação dos últimos 200ns de cada simulação foi construído um gráfico da frequência de cada ângulo entre 0° (paralelo) a 90° (perpendicular) em relação ao plano da bicamada, representado na figura 64.

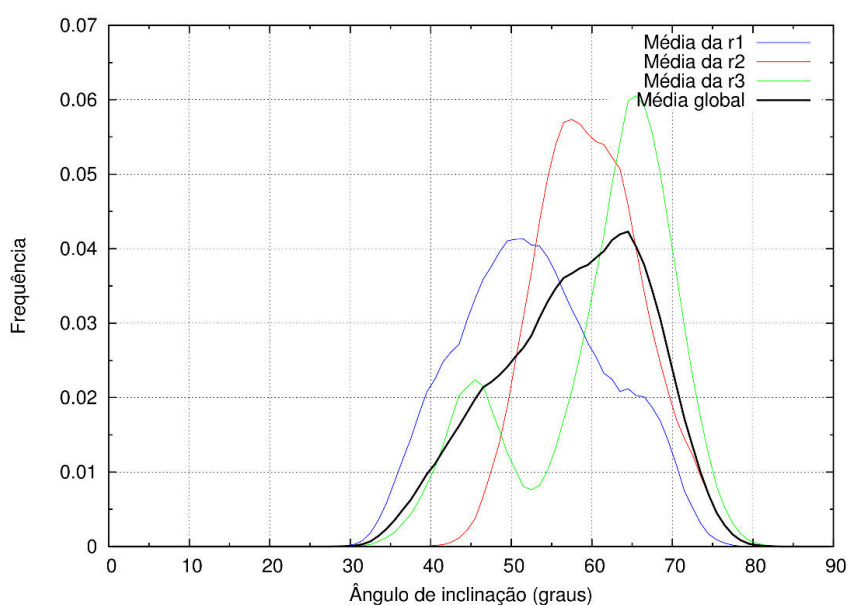


Figura 64 – Histograma de frequência do número médio de resíduos em hélice calculado para os últimos 200ns, para a média de cada replicado CG (r1, r2 e r3) e para a média de todos os péptidos, média global.

O péptido transmembranar apresenta um ângulo de inclinação muito variável. Contudo, analisando os últimos 200 ns, apenas uma réplica (r32) ultrapassou os 80 graus, apresentando aqui uma variação entre os 30 e os 80 graus.

A média das três simulações da primeira réplica CG (r1) é a que apresenta uma maior variedade de ângulos, como se pode ver pela largura e altura do pico, apresentando um máximo por volta dos 50 graus, que é um pouco inferior ao valor do pico obtido pela média global. Contudo esta mesma média r1 apresenta uma protuberância na zona dos 65 graus, sugerindo uma certa estabilidade nesses ângulos.

A réplica r2 apresenta um pico por volta dos 57 graus, valor este também inferior ao pico da média global. Por outro lado este valor do máximo da média destas três simulações é concordante com o valor médio do ângulo (calculado acima) obtido por nós e o valor obtido por Chang et al. [Chang 2008]. A réplica r2 para valores superiores ao máximo apresenta uma protuberância que só começa a diminuir depois dos 63 graus, concordante com o intervalo obtido por Tatulian et al. de 53-62° [Tatulian 2000]. Para a réplica r3 existem dois picos, um por volta dos 45 graus, que deve-se exclusivamente a dois péptidos (de 9 péptidos, 3 péptidos por cada simulação), cujos ângulos variam entre os 40 e os 50 graus. O segundo pico nos 65 graus é superior ao pico da média global, valores estes mais próximos dos valores obtidos por Victor et al. [Victor2012], mostrando também alguma variabilidade e dinâmica neste péptido.

A média das frequências de todos os péptidos apresenta uma distribuição maioritária entre os 50 e os 70 graus, ocorrendo o máximo aproximadamente aos 64 graus. Este valor é semelhante ao valor máximo obtido na média das simulações oriundas da terceira réplica CG (r3), e mesmo nas outras réplicas CG (r1 e r2) a distribuição sofre algumas flutuações por volta desse ângulo ($\pm 1^\circ$), sugerindo que esse ângulo é também especialmente favorável em todas as simulações. Este valor é no entanto diferente da média de todos os ângulos de todos os péptidos obtido por nós (58 graus), e dos intervalos experimentais 53-62 e 75-90 graus [Tatulian2000]. Note-se que os resultados experimentais são médias o que não quer dizer que a sua distribuição seja normal. Possivelmente a maioria dos péptidos encontra-se com um ângulo de inclinação superior, mas existe uma maior distribuição com ângulo inferior.

De seguida foram feitas as mesmas análises aos outros sistemas que apenas foram simulados utilizando o campo de forças MARTINI. Podendo desta forma comparar tanto as diferenças entre o tamanho do péptido, o subtipo da HA e o campo de forças utilizado. Na figura 65 estão representados a frequência com que cada ângulo é populado para todos os sistemas ao longo do tempo.

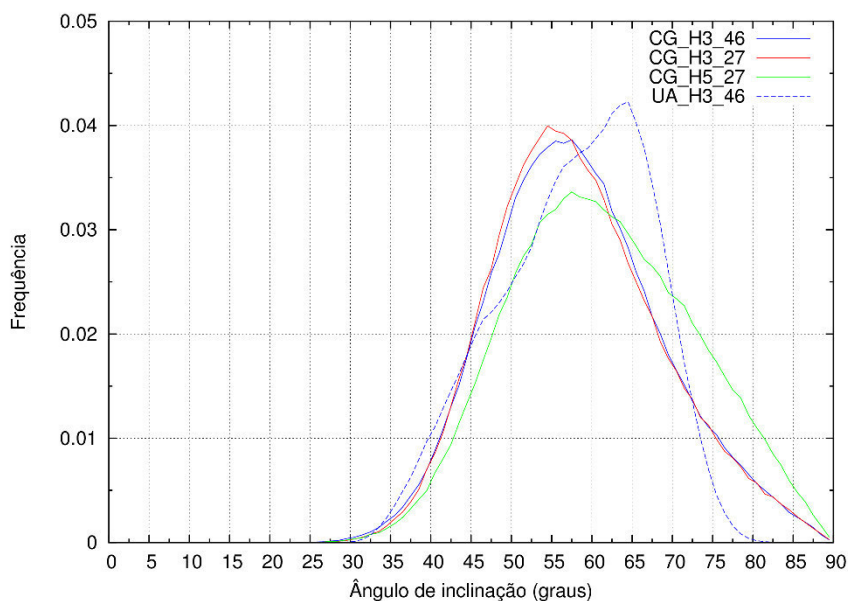


Figura 65 – Histograma de frequência do número médio de resíduos em hélice calculado para todos os sistemas estudados.

A distribuição das médias das três réplicas dos sistemas CG_H3_46 e CG_H3_27 são iguais, havendo apenas uma ínfima diferença ao nível do pico da distribuição. A única diferença entre estes dois sistemas é a existência dos resíduos do *linker* e da cauda no sistema CG_H3_46, que não são contabilizados no cálculo dos ângulos. Isto sugere que esses resíduos não afetam a conformação do domínio transmembranar na membrana. Por outro lado o sistema CG_H5_27, de uma estirpe diferente, possui um pico a um ângulo superior que o das anteriores simulações CG, e uma distribuição mais ampla nos ângulos superiores a 65°. Se a estirpe H5 apresenta ângulos superiores em *coarse-grain* é de se esperar que também apresente ângulos superiores em *united-atom*. Assim sendo, é plausível que os ângulos obtidos por Victor et al. [Victor 2012] sejam superiores aos obtidos neste trabalho, uma vez que nesse trabalho foi estudado o péptido transmembranar da estirpe H5.

É notório que o ângulo médio do sistema H3_46 é diferente nos diferentes campos de força, o que se deve provavelmente, a diferenças tanto na resolução como nos parâmetros dos campos de força. O grupo indol dos triptofanos da interface membranar aparenta ter um papel importante no ângulo das hélices transmembranares ao acomodar as caudas dos lípidos do espaço criado pelo grupo indol e outras cadeias laterais pequenas [Beswick 2011]. Isto mostra assim que diferenças estruturais podem ser relevantes para esta diferença de ângulos.

Como se irá observar na análise seguinte também na partição do péptido na membrana existem diferenças entre campos de força, e que o ângulo de um determinado péptido transmembranar é bastante influenciado pela existência de âncoras aromáticas, ou seja, de resíduos aromáticos nos limiares dos domínios transmembranares [Vostrikov 2010].

Os ângulos e conteúdo em hélice do péptido transmembranar aqui simulado reproduzem os valores experimentais, o que indica que este modelo e este protocolo foram capazes de produzir uma estrutura realista.

3.2.3.4. Partição na membrana

A partição do péptido transmembranar na membrana foi calculada de forma idêntica para o péptido de fusão, ou seja, pela distância de cada resíduo à camada de fósforos (média de todos os fósforos no eixo transversal à membrana) ao longo do tempo. No entanto, o péptido transmembranar apresentou uma variação mais robusta, observando-se que após os 100 ns a distância não variava significativamente, e portanto foram utilizados os últimos 400 ns para o cálculo da média. Essas distâncias estão representadas na 61 para todos os péptidos de todas as simulações UA e na 62 tanto a média de cada replicado proveniente das simulações CG, bem como da média global.

A partir desta análise pode-se determinar que resíduos se encontram na interface da membrana, ou seja, flanqueiam o domínio transmembranar, delimitando a penetração do péptido na membrana e por consequência o seu ângulo. Começando pelo N-terminal, o Trp13 é o primeiro resíduo do domínio transmembranar a estar debaixo da linha de fósforos em média, embora em alguns péptidos (maioritariamente na primeira réplica CG) seja o Trp10. Os resíduos Trp10, Ile11 e Leu12 encontram-se ao mesmo nível, em média, e imediatamente acima do centro dos fósforos, o que significa que interagem com estes. A existência de dois resíduos carregados, Lys8 (carga positiva) e Asp9 (carga negativa) imediatamente antes do Trp10 explica, tanto a diminuição acentuada da distância aos fósforos do resíduo 9 para o resíduo 10, como a distância do Trp 10 aos fósforos. Isto deve-se ao número superior de resíduos carregados como a carga negativa que interage repulsivamente com os fósforos.

No C-terminal o Trp33 não é o responsável pela ancoragem à interface, mas sim o grupo carregado positivamente Arg37. Mesmo não sendo o esperado, é facilmente explicado pela favorável interação entre grupos com diferentes cargas, neste caso a arginina carregada positivamente com o grupo fosfato carregado negativamente.

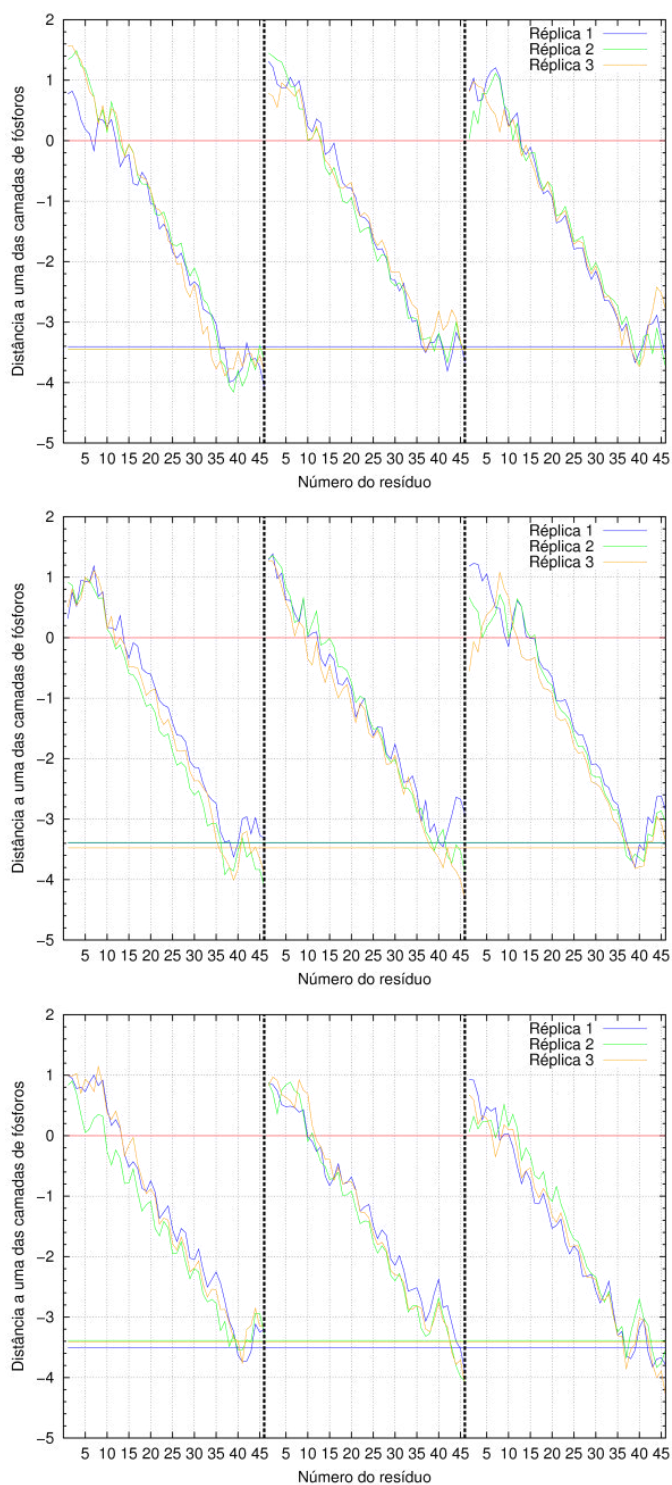


Figura 66 – Representação da partição de cada um dos péptidos em separado na membrana para cada uma das réplicas da réplica CG1 (A), réplica CG2 (B) e réplica CG3 (C).

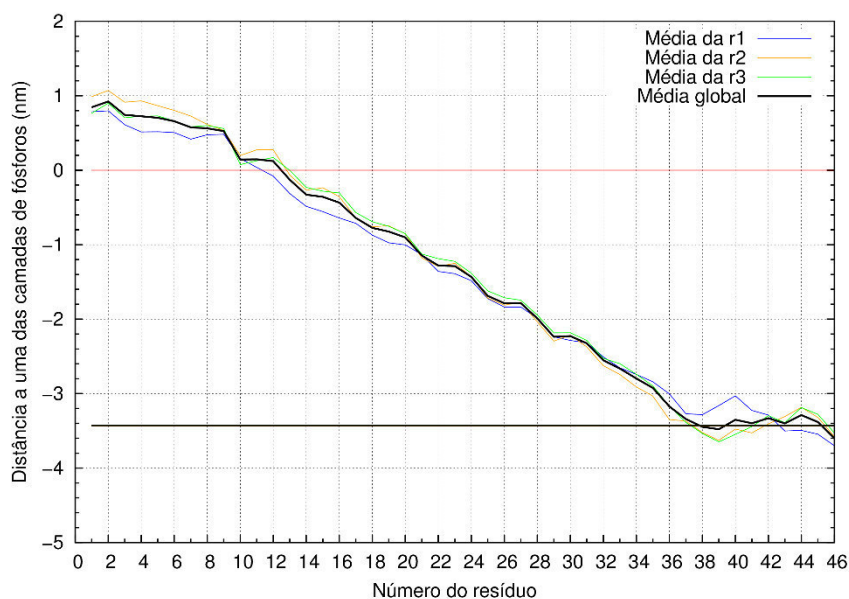


Figura 67 – Representação das distâncias médias de cada resíduo ao plano dos fósforos para as diferentes réplicas.

Na maioria das simulações, observa-se no C-terminal, após a Arg37, a existência de uns resíduos mais exteriores à membrana, mas os resíduos seguintes voltam a penetrar na membrana. Ao nível fisiológico as cisteínas da cauda citoplasmática estão ligadas covalentemente a grupos acilo, adicionados pós-traducionalmente, os quais vão interagir com a membrana. É curioso portanto que mesmo sem esses grupos acilo a cauda em si têm afinidade para a membrana, especialmente a interface, uma vez que a cauda citoplasmática também possui um papel na interação da hemaglutinina com as proteínas do citoesqueleto [Chen 2005]. Uma possível explicação seria que, mesmo sendo as cisteínas conservadas, a evolução selecionou resíduos com afinidade para a interface, visto serem possíveis mutações nos resíduos de cisteína, impedindo a adição de grupos acilo à cauda citoplasmática.

Observa-se ainda que as extremidades que correspondem ao *linker* e à cauda citoplasmática se encontram na interface ou no exterior da membrana, o qual aumenta a sua flexibilidade. Isto reproduz assim a observação de que ambos os segmentos apresentam uma estrutura desordenada.

O resultado aqui apresentado, onde os triptofanos e os resíduos carregados funcionam como âncoras, é bastante reportado para este sistema e outros sistemas, inclusive um estudo sugeriu que o Trp33 ou mesmo os triptofanos no N-terminal são

importantes para o posicionamento correto do péptido transmembranar na membrana [Wimley 1996].

Os resíduos carregados, arginina e lisina, são bifuncionais, possuindo uma parte hidrofóbica e conseguem formar pontes de hidrogénio com grupos polares nos lípidos, o que leva a que sejam encontrados predominantemente em *loops* citoplasmáticos de proteínas transmembranares. No caso dos triptofanos, estes conseguem formar pontes de hidrogénio com a cabeça dos lípidos e a parte aromática promove interações hidrofóbicas com a zona mais interior da membrana, resultando em que estes resíduos se localizem preferencialmente na interface membrana/água [Schiffer 1992, Nugent 2012]. Na imagem seguinte podemos observar a distribuição, tanto dos triptofanos como dos resíduos ácidos na membrana (figura 68).

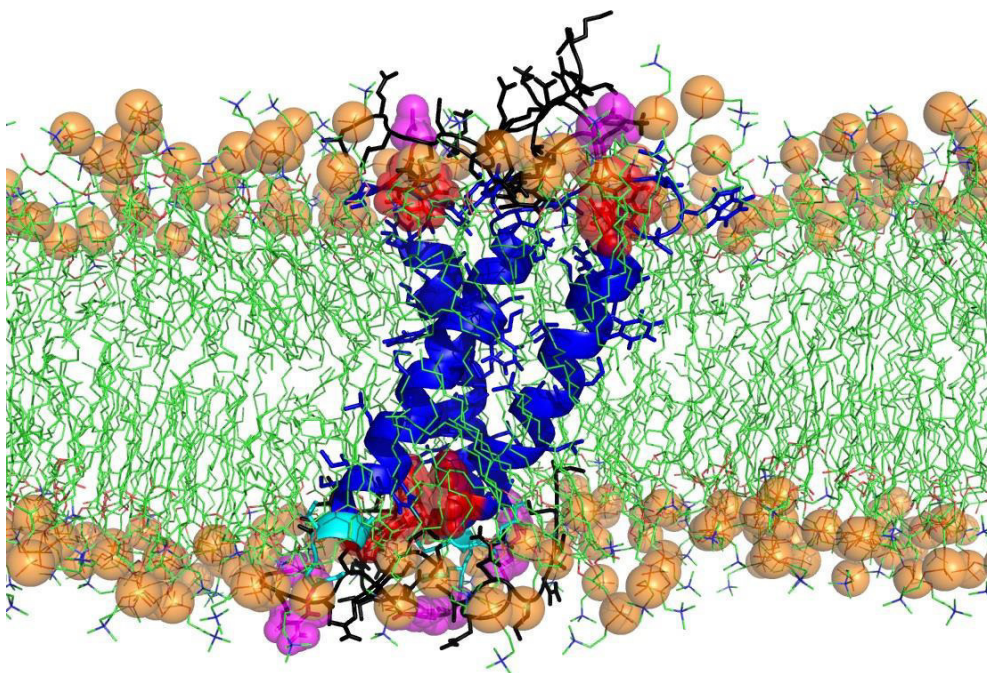


Figura 68 – Estrutura do trímero dos péptidos transmembranares realçando os triptofanos e os resíduos ácidos. O péptido está representado por *cartoon* com a zona definida por hélice a azul e a restante a preto, os triptofanos e os resíduos ácidos estão representados por esferas vermelhas e magentas, respetivamente. Os lípidos estão representados por linhas, exceto os fósforos que estão representados por esferas.

De modo a se poder comparar as diferenças na partição devido ao comprimento, subtipo da HA e campo de forças utilizado, apresentamos de seguida a partição do péptido transmembranar para todos os sistemas estudados, representado na figura 69.

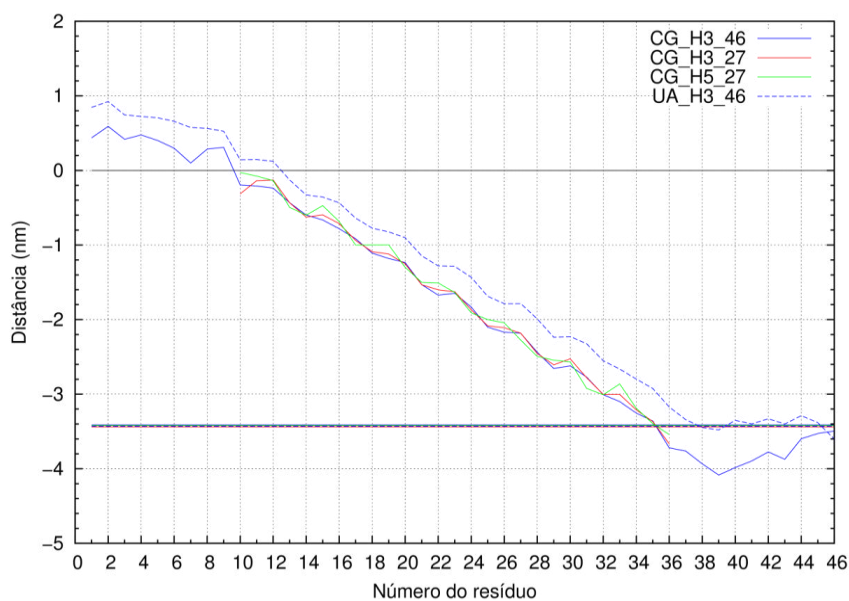


Figura 69 – Representação das distâncias médias de cada resíduo para os diferentes sistemas estudados.

É facilmente observado que a partição dos três sistemas utilizando o campo de forças MARTINI é semelhante, o que era expectável, uma vez que mesmo sendo a sequência de aminoácidos dos subtipos H3 e H5 diferente, ambos fazem parte do domínio transmembranar. Contudo como já tinha sido observado pela análise dos ângulos, existem diferenças entre os diferentes campos de força, neste caso nas partições dos péptidos na membrana.

No caso do péptido usando o campo de forças MARTINI o N-terminal é ancorado na interface pelo Trp10, em contraste ao Trp13 que é preferencial nas simulações UA. No C-terminal o resíduo na interface é a Cys35 (uma das cisteínas conservadas, à qual é adicionado um grupo acilo pós-traducionalmente) que se encontra entre o Trp33 e a Arg37, sugerindo que ambos tenham um papel na ancoragem.

Aparte das diferenças intrínsecas dos campos de força MARTINI e 54A7, nota-se que nas simulações com MARTINI os resíduos 11-35, definidos como estrutura secundária em hélice, são os resíduos que se encontram inseridos na membrana. Isto poderia sugerir que a estrutura em hélice favorece a presença dos resíduos na membrana. Contudo, mesmo sendo uma explicação para a partição do N-terminal, o C-terminal perde a helicidade a partir do resíduo 35 das simulações UA, coincidindo com os resíduos em hélice das simulações CG, mas apresentando partições diferentes.

Embora tal possa ser explicado pela necessidade de existirem um certo número de resíduos em hélice para atravessar a membrana.

Tanto o ângulo como a partição do péptido transmembrantar ao longo da membrana são em parte controlados pela incompatibilidade hidrofóbica (“*hydrophobic mismatch*”), ou seja, a diferença entre o comprimento hidrofóbico da hélice e a espessura hidrofóbica da membrana [Kim 2010]. Todavia, pela observação da figura 69 não existem diferenças significativas entre a espessura da bicamada de DMPCs para ambos os campos de força.

3.2.3.5. Parâmetros de ordem das caudas dos lípidos

Um dos efeitos do péptido transmembrantar já demonstrado é a ordem que provoca nas membranas onde se encontra inserido, ou seja, aumenta o parâmetro de ordem dos lípidos. De forma a podermos ter uma maior confiança no nosso modelo do péptido transmembrantar e ter uma visão mais pormenorizada analisou-se o parâmetro de ordem e cada um dos carbonos da cauda dos lípidos (figura 70), de modo a comprovar o efeito sugerido no início do parágrafo.

A forma mais comum de se analisar a ordem de bicamadas lipídicas é através do parâmetro de ordem S_{CD} (CD, carbono-deutério), que mede a orientação relativa das ligações C-D relativamente à normal da membrana.

$$S_{CD} = \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle$$

Onde θ é o ângulo entre a ligação C-D e um eixo de referência, neste caso será a normal à membrana.

No caso de simulações *united-atom* uma vez que os hidrogénios não são contabilizados, é preciso fazer uma reconstrução do vetor utilizando três carbonos sucessivos da cauda. Razão essa por não existir parâmetro de ordem para o último resíduo da cauda de DMPC (figura 70).

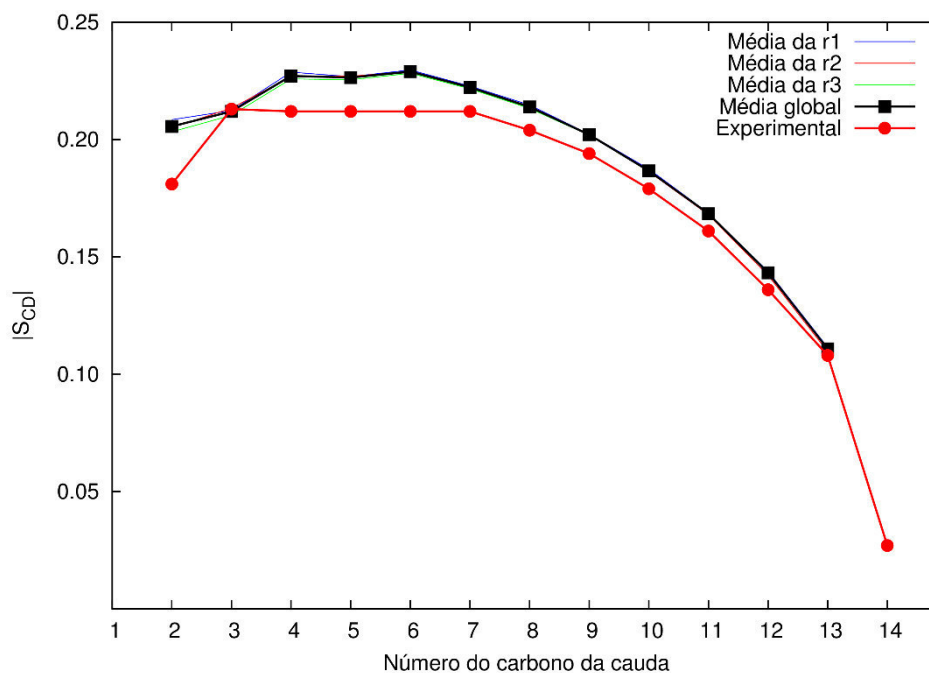


Figura 70 – Representação do parâmetro de ordem S_{CD} médio para cada um dos carbonos da cauda dos lípidos. A vermelho estão representados os resultados experimentais por Petrache et al. [Petrache 2000].

Como se pode observar pelo gráfico a ordem por lípido nestes sistemas é bastante reprodutível, dado que as médias das réplicas de cada uma das três estruturas iniciais são muito próximas. Com exceção do terceiro carbono da cauda os valores obtidos na simulação são superiores aos valores de referência experimental, sendo a diferença cada vez menor à medida que nos aproximamos da extremidade da cauda. No caso excepcional do terceiro carbono, nota-se que, além de ter valor semelhante ao experimental, este é inferior ao quarto carbono, o oposto do que ocorre nos valores experimentais. Contudo a curva obtida por nós tem uma forma semelhante a outra obtida por Machuqueiro et al., com os valores a subir até ao quarto carbono, e uma subida ténue do segundo para o terceiro, em contraste com o que ocorre experimentalmente. Podemos então constatar que o péptido de fusão aumenta de facto a ordem dos lípidos na membrana onde se encontra inserido, apresentando características dinâmicas próximas do experimental. Note-se que as simulações só da membrana sem os péptidos transmembranares dariam aproximadamente o mesmo resultado que o experimental.

O péptido transmembranar aumenta a ordem dos lípidos, o que provoca uma desidratação da superfície da membrana, diminuindo a repulsão/pressão de hidratação

entre membranas. Tal é fundamental na desestabilização da membrana e na curvatura aquando da fusão [Tatulian 2000]. Este efeito não ocorre com a presença de outros péptidos na membrana. Por exemplo, péptidos de polileucinas não têm qualquer efeito na ordem dos lípidos aquando da sua inserção na membrana [Nezil 1992]. Isto demonstra, que o efeito é causado pelas propriedades do péptido transmembranar e não apenas por estar inserido na membrana.

3.3 Hemaglutinina

A hemaglutinina é uma proteína com elevado número de resíduos (cerca de 1620), que só por si apresenta um considerável custo computacional na sua simulação. Aliado a isso tem a uma forma não globular, sendo necessária uma caixa de simulação maior e por consequência um maior número de águas. Existem no entanto simulações do ectodomínio da hemaglutinina em água a valores de pH e a temperaturas diferentes, realizadas por um membro do grupo (Victor et al., resultados não publicados). Resumidamente nota-se diferenças na estabilidade dos diferentes sistemas, mas é necessário muito maior tempo de simulação para se observarem alterações conformacionais, podendo mesmo ser impossível com os atuais recursos computacionais. Como tal, apresentamos aqui uma abordagem utilizando o campo de forças MARTINI, que ao reduzir o pormenor da proteína consegue simular maiores tempos de simulação.

3.3.1 Estabilidade em água – aplicação de ligações elásticas

O campo de forças MARTINI pode ser uma solução para reduzir o custo computacional, mas este campo de forças não foi desenvolvido nem parametrizado para modelar alterações conformacionais. Em vez disso as estruturas secundárias são definidas inicialmente e são imutáveis ao longo da simulação.

Simulações realizadas com o ectodomínio da hemaglutinina em água, utilizando o MARTINI, provocaram variações aparentemente não físicas na estrutura terciária, nomeadamente um deslocamento ascendente da parte inferior (não globular). A existência de alterações não físicas da estrutura terciária já tinha sido reportada no *website* do MARTINI para outras proteínas (<http://md.chem.rug.nl/cgmartini/index.php/tutorial/elastic-network>). Note-se que as simulações foram feitas sem a parte transmembranar e sem a presença de uma membrana, o que pode ter influenciado as alterações verificadas. No entanto decidiu seguir-se o protocolo para proteínas grandes com estrutura terciária bem definida, ou seja, a utilização da representação ELNEDIN (Elastic Network in Dynamics – Rede elástica na dinâmica).

A representação ELNEDIN aparece então como solução tanto para o custo computacional, como para a manutenção da conformação terciária graças à adição de ligações elásticas entre os átomos da cadeia principal. O aumento da rigidez da estrutura tem um lado negativo, pois torna a amostragem do espaço conformacional bastante restrita, a qual já por si usando o campo de forças MARTINI apresenta as suas lacunas. Contudo, dependendo das análises que se pretendem realizar, a diminuição da amostragem e a incapacidade de alterações conformacionais podem ser menos valorizadas.

Como já foi discutido no capítulo 2, o primeiro passo é saber que conjunto de valores R_c e K_{spring} são mais eficazes para simular a proteína em estudo. Foram feitas simulações do ectodomínio da hemaglutinina em água usando 6 conjuntos diferentes de valores, $R_c = 0.8, 0.9$ e 1.0 nm (figura 71) e $K_{spring} = 500$ e 1000 $\text{kJ mol}^{-1} \text{nm}^{-2}$, para três diferentes sistemas: i) campo de forças MARTINI com ligações elásticas; ii) campo de forças ELNEDIN com ligações intra e inter monómeros; iii) campo de forças ELNEDIN com ligações apenas intra monómeros.

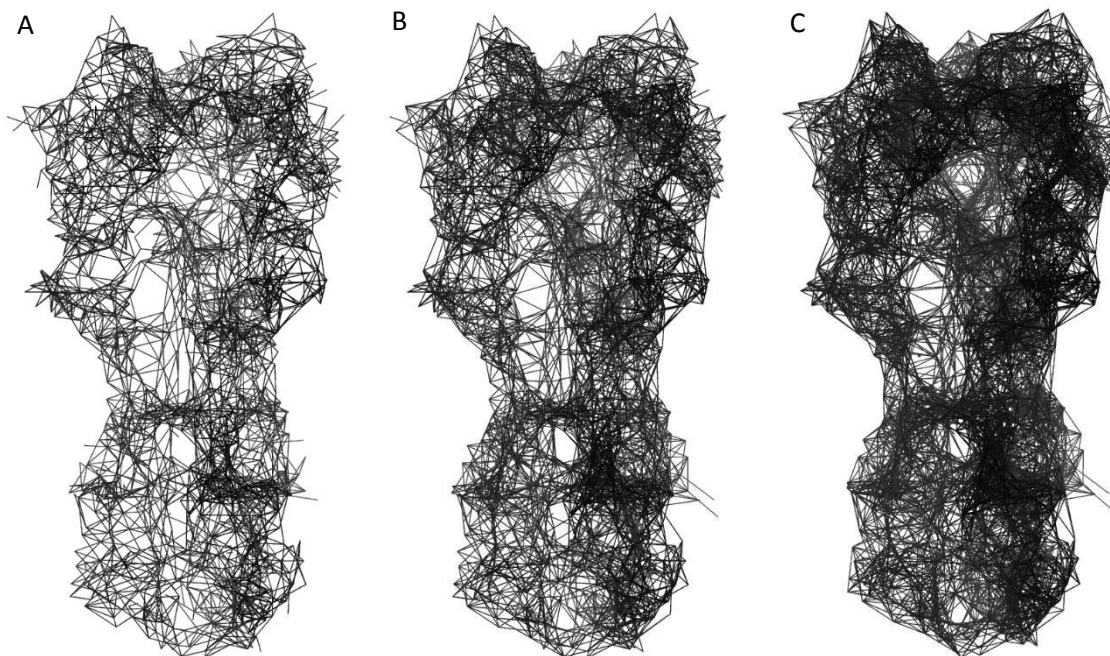


Figura 71 – Representação das ligações entre partículas da cadeia principal da hemaglutinina, com diferentes valores de R_c : 0.8 nm (A), 0.9 nm (B) e 1.0 nm (C).

O raciocínio por detrás destes 3 sistemas é verificar as diferenças entre o uso de ligações elásticas em MARTINI e o uso do campo de forças ELNEDIN, já estudadas para

outras proteínas por Periole et al. [Periole2009], e as diferenças de se adicionar ligações entre os 3 monómeros ou não. Cada uma das 18 simulações foi corrida durante 1 μ s, e as suas trajetórias foram analisadas segundo o seu RMSD e RMSF da hemaglutinina. O RMSF foi também comparado com o resultado de simulações atomísticas. As análises para o sistema ELNEDIN com ligações entre monómeros estão representados nas figuras 72 e 73.

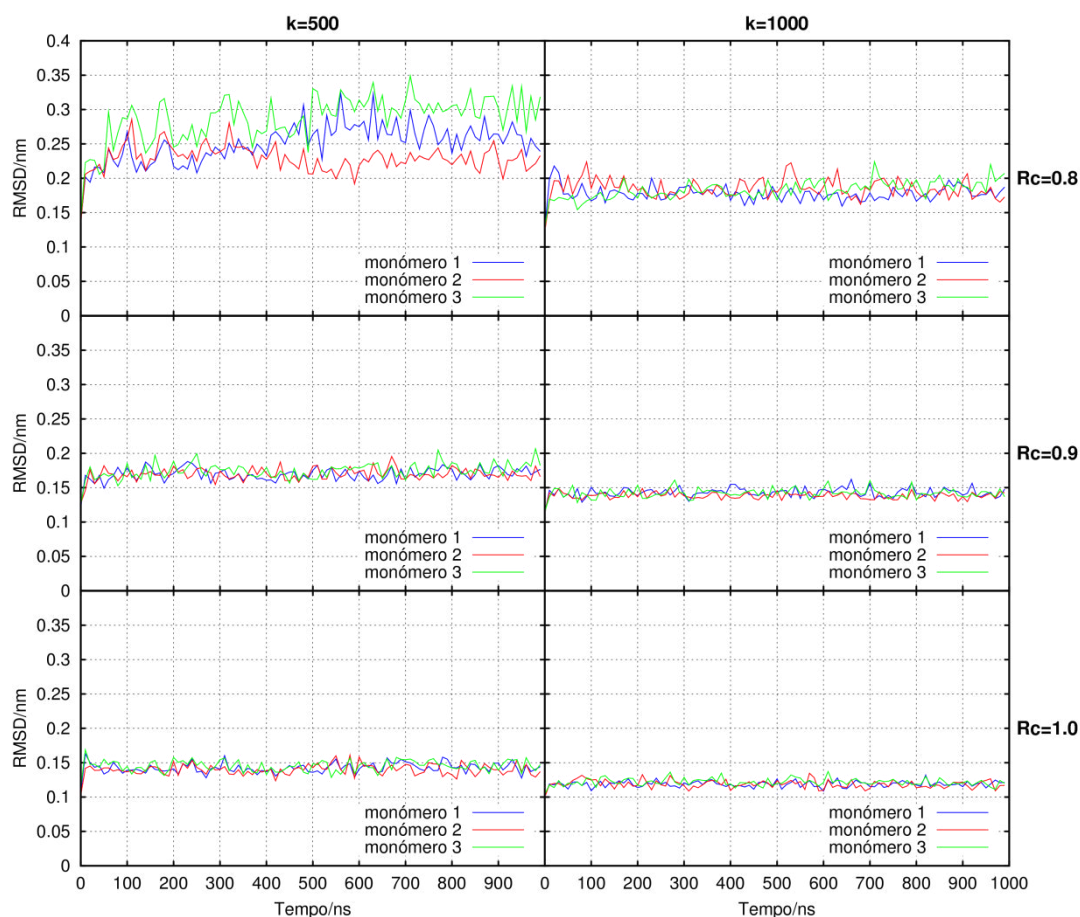


Figura 72 – Representação da evolução do RMSD ao longo do tempo para cada um dos conjuntos de 2 valores R_c (0.8, 0.9 e 1.0 nm) e K_{spring} (500 e 1000 $\text{kJ mol}^{-1} \text{nm}^{-2}$) para o sistema ELNEDIN com ligações intra e inter-monómeros.

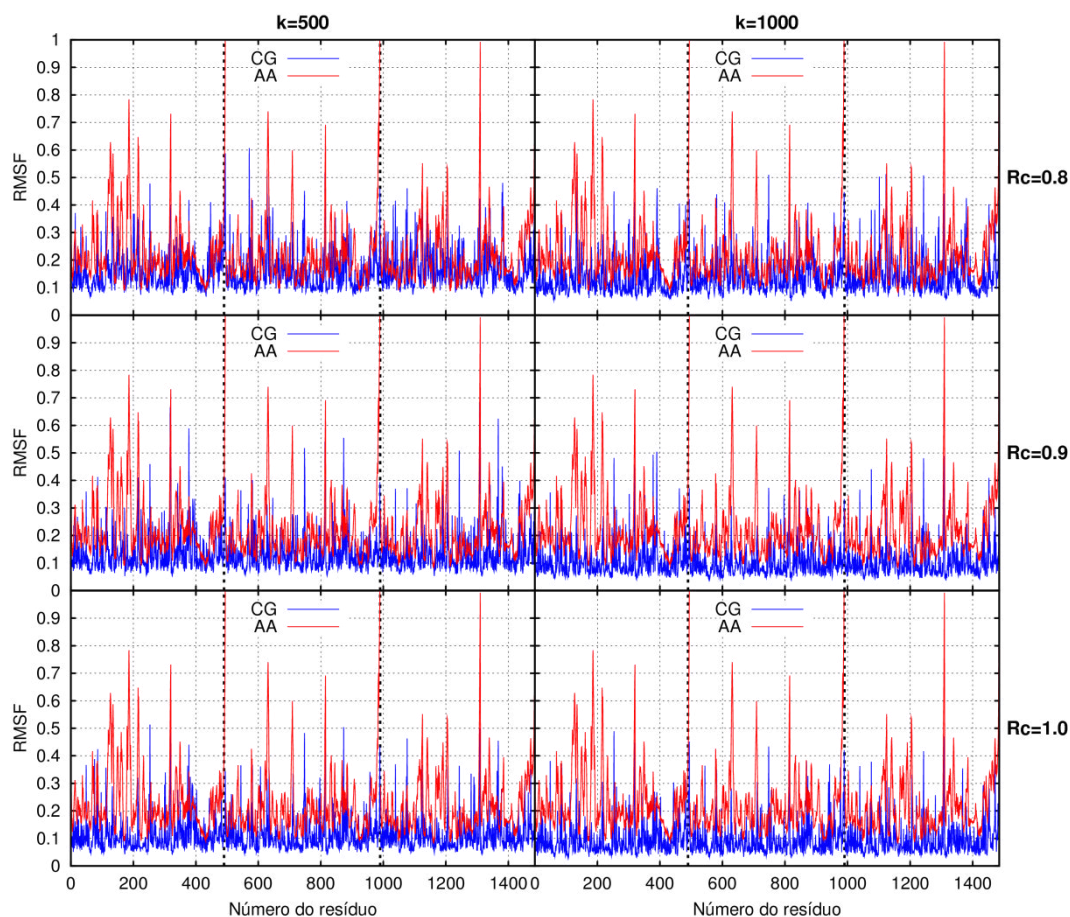


Figura 73 – Representação da evolução do RMSF ao longo do tempo para cada um dos conjuntos de 2 valores R_c (0.8, 0.9 e 1.0 nm) e K_{spring} (500 e 1000 $\text{kJ mol}^{-1} \text{nm}^{-2}$) para o sistema ELNEDIN com ligações intra e inter-monómeros. A vermelho (AA) encontra-se o RMSF calculado por Victor et al. utilizando um campo de forças *united-atom*.

O sistema sem ligações entre os monómeros, sistema iii, (ANEXO RMSF) apresenta flutuações ao mesmo nível das simulações atomísticas, contudo apresenta um desvio grande em relação à estrutura inicial, como observado no RMSD (ANEXO RMSD), concluindo-se ser necessária a ligação entre os três monómeros entre si. O sistema de ligações elásticas (sistema i) e o sistema ELNEDIN com ligações entre os monómeros (sistema ii) não têm grandes diferenças entre si (figuras 72 e 73). Apresentam apenas diferenças no RMSD usando restringimentos intermédios ou fracos, bem como no RMSF de restringimentos intermédios. Contudo optou-se pela utilização do campo de forças ELNEDIN nos estudos seguintes com a hemaglutinina, uma vez que este campo de forças possui melhorias nos parâmetros do campo de forças em comparação com o MARTINI. Ao analisar todas as simulações do sistema ii e tentando encontrar um equilíbrio entre uma estrutura estável e flutuações semelhantes às simulações atomísticas, optou-se por um restringimento intermédio utilizando os

valores 0.9 nm para o R_c e $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. Este valor foi utilizado para todas as restantes simulações da hemaglutinina neste trabalho. Uma vez que já temos os parâmetros do campo de forças ELNEDIN configurado, podemos prosseguir com as simulações.

Acreditasse que a alteração conformacional do péptido dobradiça com o abaixamento de pH é um passo essencial para a fusão do envelope viral com a membrana da célula do hospedeiro. A primeira questão abordada foi se seria possível induzir a mudança conformacional do péptido dobradiça de um estado desordenado para uma estrutura em hélice, e qual a sua repercussão no resto da proteína. De forma a averiguar isso, utilizou-se a técnica do crescimento lento (*slow-growth*), explicado ao detalhe de seguida.

3.3.2 Alterações conformacionais com o método do crescimento lento

A técnica de crescimento lento foi aqui aplicada para transformar o segmento desordenado do péptido dobradiça (estado A) numa estrutura em hélice (estado B). Para ajustar os parâmetros da simulação, utilizaram-se primeiro apenas os resíduos 40-91 da HA2 na estrutura cristalográfica da hemaglutinina do subtipo H3 (figura 74 A), visto que um reduzido número de resíduos significa simulações mais rápidas. As alterações do estado A para o estado B foram feitas alterando as constantes de força das ligações, os ângulos e os ângulos diedros da cadeia principal. Alterando os seus valores de uma estrutura desordenada (estado A) para os valores respetivos de uma estrutura em hélice (estado B). Estas alterações foram aplicadas progressivamente ao longo de 200 ns, para que as alterações fossem feitas de uma forma suave. Obteve-se no final desses 200 ns uma estrutura esticada em forma de hélice, tal como esperado, mostrando assim que o método é eficaz (figura 74 B).

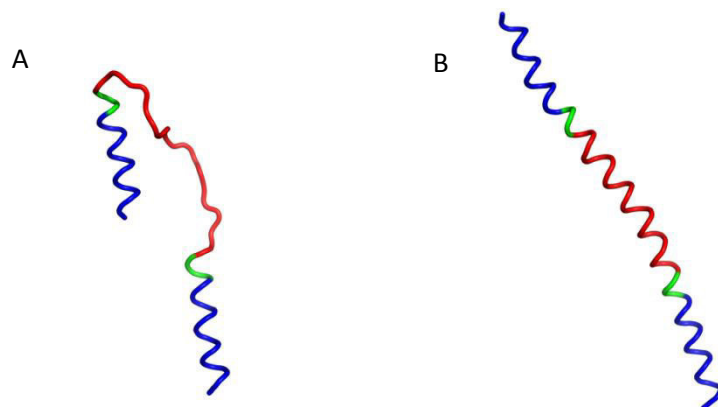


Figura 74 – Péptido dobradiça representado por tubos que ligam as partículas da cadeia principal (A) no início da simulação e (B) no fim da simulação do crescimento lento. Os resíduos que tinham estrutura desordenada estão representados a vermelho, os resíduos com estrutura em hélice estão representados a azul e os resíduos com estrutura em hélice, mas que tiveram os parâmetros de ligações e ângulos alterados, estão representados a verde.

Testado o método, o próximo passo foi aplicar essa mesma transformação, nos mesmos resíduos, mas desta vez na estrutura do ectodomínio completo. O sistema da hemaglutinina em ELNEDIN possui ligações elásticas em toda a sua estrutura, incluindo tanto o péptido dobradiça como os primeiros 40 resíduos do N-terminal (incluindo o péptido de fusão) da HA2. As ligações elásticas, ao resto da proteína, do péptido dobradiça impossibilitam este de conseguir formar uma hélice, e impedem a deslocação do N-terminal. Optou-se portanto por retirar as ligações desde o C-terminal do péptido dobradiça até ao N-terminal da HA2 com a restante proteína, mantendo apenas as ligações entre resíduos desse mesmo segmento (figura 75 A).

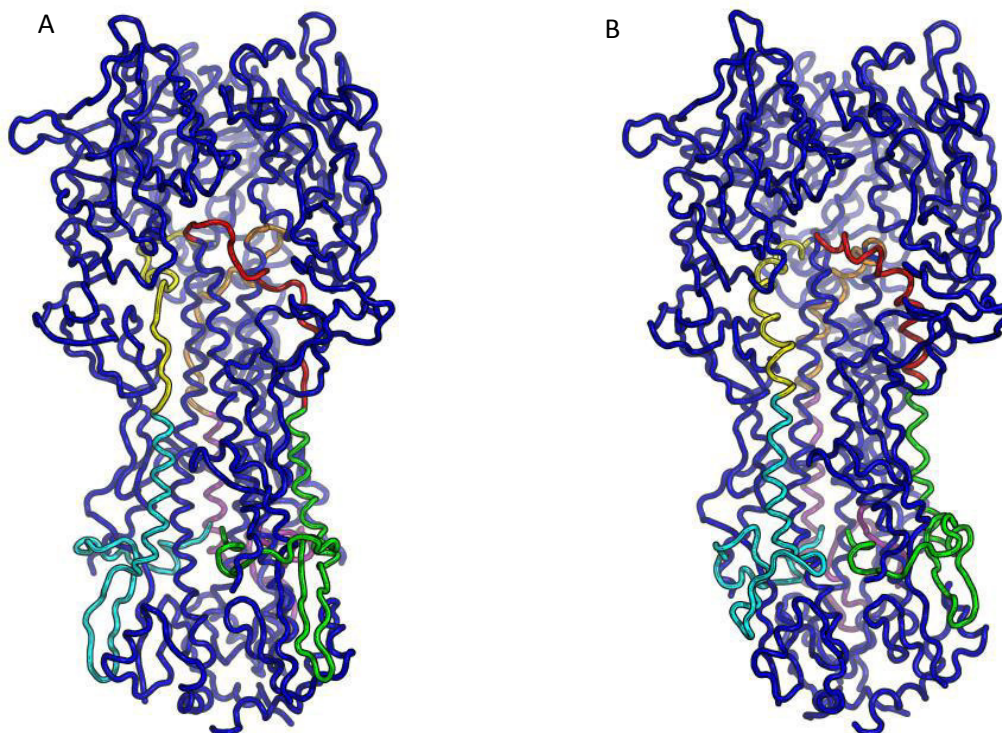


Figura 75 – Representação da conformação da hemaglutinina antes (A) e após (B) o processo de crescimento lento. Proteína representada como tubos, péptido dobradiça a vermelho, amarelo ou laranja (para cada monómero) e o N-terminal de cada HA2 como verde, ciano e roxo.

Neste caso, devido à maior complexidade do ambiente à volta do péptido dobradiça 200 ns não são suficientes para a suavidade das alterações, sendo necessário simular durante 400 ns. Ao contrário do segmento linear formando uma hélice perfeita obtido anteriormente com o péptido livre, neste caso obtivemos uma hélice curvada (figura 75 B). Isto deve-se provavelmente à parte globular da HA1 imediatamente por cima do péptido dobradiça, já que, as subunidades HA1 sofrem uma destabilização com o abaixamento do pH, ocorrendo um afastamento das mesmas, segundo alguns estudos [Bottcher 1999, Xu 2011]. No nosso modelo isto não é permitido devido há existência de interações entre subunidades.

Ainda assim, se o péptido dobradiça em hélice apresenta um comprimento inferior comparativamente ao péptido dobradiça estendido, então este encurtamento poderia obrigar a saída do péptido de fusão da bolsa onde está inserido na proteína. Porém apenas puxou a zona das folhas beta para cima, deixando o péptido de fusão inalterado.

Uma possível explicação para o ocorrido é a elevada estabilidade do péptido de fusão na bolsa a pH neutro. O que por sua vez sugere que não deverá ser apenas o movimento do péptido dobradiça, mas o próprio péptido de fusão deve ser translocado para o exterior da bolsa após variações na protonação dos resíduos chave. Esta possibilidade já tinha sido sugerida, visto que estudos anteriores observaram uma saída do péptido de fusão nos passos iniciais das alterações conformacionais [White 1987, Korte 1999].

Uma vez que não é possível estudar alterações conformacionais com um sistema tão rígido, decidiu-se estudar o *self-assembling* da hemaglutinina, processo que também foi usado para o péptido de fusão e o péptido transmembranar, ambos estudados anteriormente.

3.3.3 *Self-assembling* da sequência completa da hemaglutinina

Para se estudar o *self-assembling* da hemaglutinina é necessário adicionar o segmento C-terminal que não se encontra na estrutura cristalográfica, uma vez que é onde se encontra o domínio transmembranar. É à volta desse domínio que os lípidos se vão orientar e formar uma membrana. Foi então adicionado a estrutura inicial do péptido transmembranar (igual à usado nos estudos do péptido transmembranar) ao C-terminal da estrutura cristalográfica da hemaglutinina. Foi também necessário definir os parâmetros ligantes (ligações e ângulos) entre o N-terminal do domínio transmembranar e o C-terminal da hemaglutinina na topologia. Estas simulações são, portanto, um híbrido de campos de força, onde se usou o ELNEDIN para o ectodomínio da HA, mas o MARTINI sozinho para o domínio transmembranar.

A hemaglutinina completa foi colocada no centro de uma caixa de simulação (de dimensões 19 nm no eixo x, 19 nm no eixo y, 25 nm no eixo z) e de seguida foram adicionados 1200 lípidos, quantidade suficiente para formar uma bicamada ao longo dos eixos x e y. Partículas de água martini foram de seguida adicionadas para encher a caixa, num total de 59523, onde 10% são anticongelantes. Foram simuladas 10 réplicas para cada um dos dois tipos de acoplamento de pressão utilizados, isotrópico e semi-isotrópico.

As simulações com o acoplamento isotrópico foram corridas durante 1 μ s e em nenhuma das 10 réplicas se gerou uma bicamada lipídica entre os eixos x e y. Além de ser uma questão de probabilidade, a dimensão elevada da caixa torna os lípidos demasiado dispersos, torna difícil a formação de uma única bicamada lipídica. Mesmo em caixas mais pequenas utilizadas nos estudos dos péptidos foram necessárias eliminar certas simulações que além de uma bicamada criaram também micelas. No entanto formaram-se discos lipídicos que se comportam como uma bicamada excepto nas extremidades, onde os lípidos se arranjam como numa micela. Estes discos lipídicos são designados por bicelas. Para além da formação das bicelas ocorre ainda formação de micelas de diferentes tamanhos, as quais parecem interagir com zonas específicas da hemaglutinina, principalmente na parte superior da HA1. Um exemplo da hemaglutinina inserida numa bicela e micelas a interagir com esta, obtido no final de uma das réplicas, pode ser visualizado na figura 76.

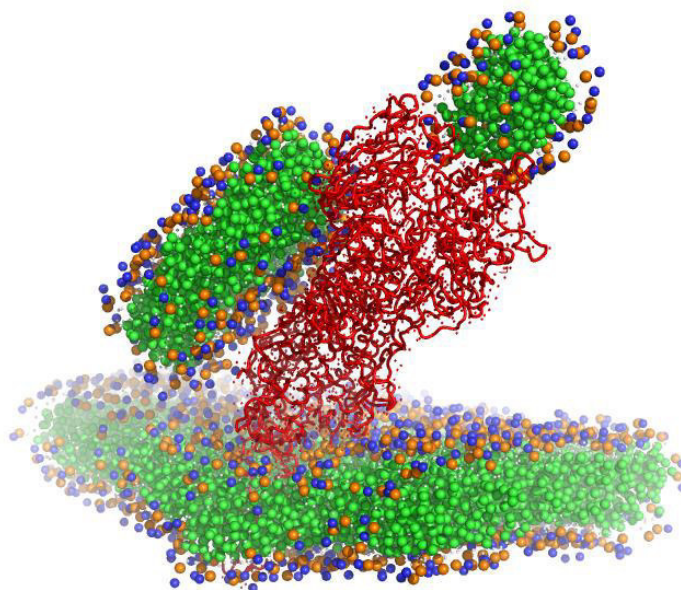


Figura 76 – Estrutura final de uma das réplicas, onde está representada a cadeia principal e as cadeias laterais da proteína como tubos vermelhos e esferas amarelas, respetivamente. Os lípidos estão representados como esferas sendo a azul o grupo de colinas, a laranja o grupo fosfato e a verde as cadeias de ácidos gordos.

As simulações utilizando acoplamento de pressão semi-isotrópico tiveram as suas caixas de simulação bastante deformadas, provocando a aproximação dos limites da caixa à proteína, o que origina interações da proteína com ela própria através das imagens periódicas. Estas simulações foram, portanto, desconsideradas. Foi no entanto observada a formação de discos e vesículas lipídicas.

Para estudar com mais detalhe as interações das micelas com a Hemaglutinina, a qual parece ter locais com maior propensão para interagir com os lípidos, realizamos mais um set de simulações para se calcular os contactos entre a HA e os lípidos.

Contactos entre hemaglutinina e lípidos

De modo a calcular os contactos entre a HA e os lípidos, nomeadamente às bicelas e às micelas, foram corridas 20 novas simulações. As partículas da cadeia central da proteína foram fixadas com restrições de posição fracas ($100 \text{ kJ mol}^{-1} \text{ nm}^{-2}$), de modo a que não ocorresse translação da proteína e assim não interagisse com as hemaglutininas e lípidos das imagens periódicas, o que ocorria nas simulações anteriores. O próprio ectodomínio da HA já continha ligações elásticas entre a sua cadeia principal, portanto a adição de uma reduzida força à posição não irá influenciar significativamente o sistema. Todas as simulações foram corridas durante 100 ns, mas para o cálculo de contactos os primeiro 25 ns foram descartados.

Mais uma vez, apenas pela visualização da simulação consegue-se observar locais preferenciais de interação, nomeadamente com o domínio transmembranar e a parte globular da HA1. Algumas imagens do final das simulações estão representadas na figura 77.

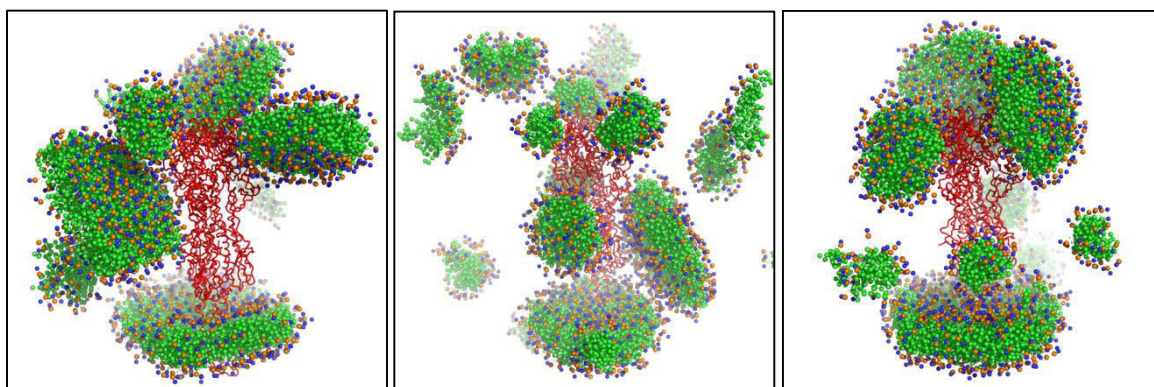


Figura 77 – Estrutura final da réplica 1 (esquerda), réplica 7 (meio) e da réplica 13 (direita). A proteína encontra-se no centro a vermelho representada como tubos e os lípidos representados como esferas: a azul as colinas, a laranja os fosfatos e verde as cadeias.

Para calcular o número de contactos, que cada um dos resíduos da HA faz com os lípidos, mede-se a distância mínima desse resíduo a todas as partículas dos lípidos, sendo contabilizadas apenas as partículas a uma distância igual ou inferior a 0.8 nm.

Este valor já tinha sido previamente utilizado para calcular o número de contactos entre uma proteína e lípidos em simulações de MARTINI [Wee 2008] e define-se pela distância plausível para ocorrerem interações de van der Waals. Este valor é consistente com o raio de van der Waals de uma partícula C1 (da cauda lipídica). Foram ainda testados outros valores de distância para o cálculo de contactos mas a única diferença significativa é a intensidade dos picos, problema resolvido aquando da normalização.

Fez-se uma média do número de contactos para cada resíduo ao longo do tempo e nas diferentes réplicas. Na figura 78 o número de contactos para cada resíduo foi mapeado na estrutura da HA com um gradiente de cores do azul-escuro (menor número de contactos) a vermelho (maior número de contactos). Na figura 79 apresenta-se essa variação ao longo do número do resíduo.

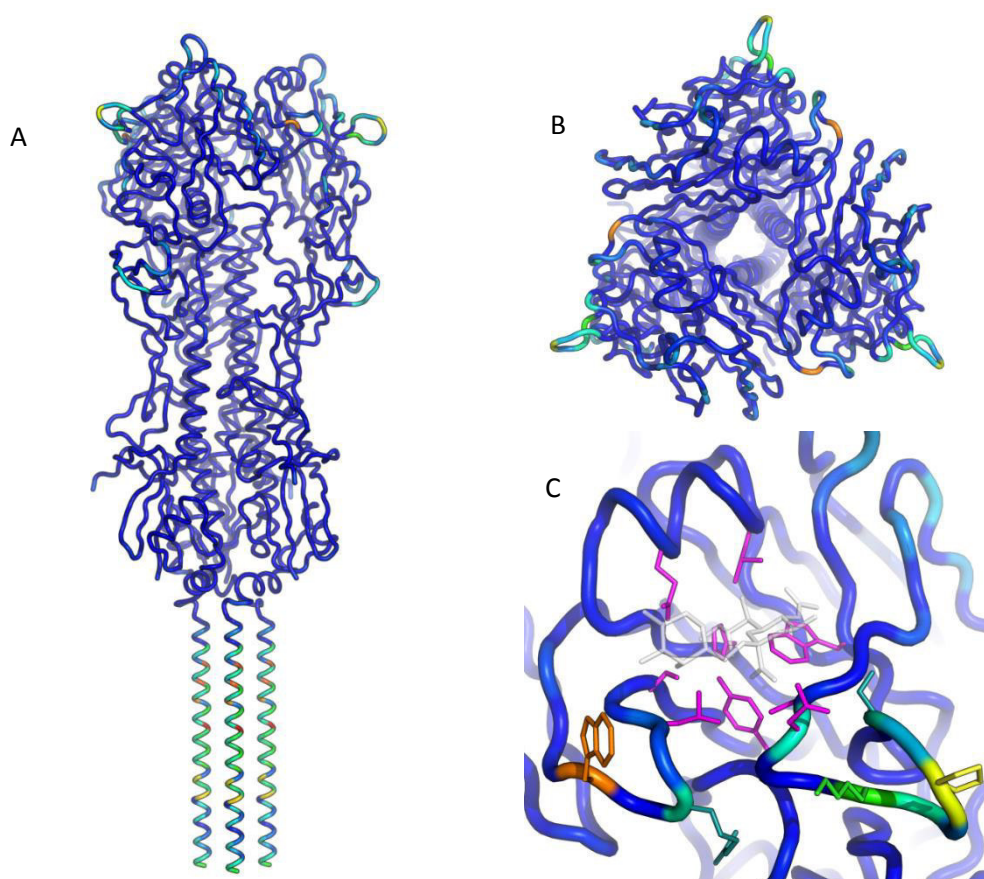


Figura 78 – Mapeamento dos contactos entre cada resíduo com os lípidos através de um gradiente de cores, onde o azul representa poucos ou nenhuns contactos e vermelho representa o número máximo de contactos. A HA encontra-se representada por tubos, ora de uma vista lateral (A) ou de uma vista superior (B). No (C) encontra-se em pormenor o local de ligação ao ácido siálico, aqui representado a branco, e a magenta os resíduos que interagem com este.

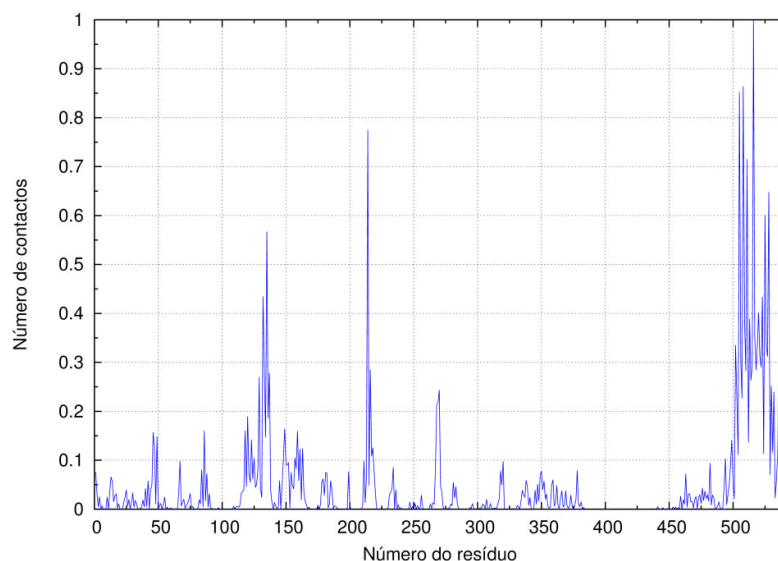


Figura 79 – Número médio normalizado de contactos para cada um dos resíduos de uma subunidade da hemaglutinina.

A zona com maior número de contactos com os lípidos é o segmento transmembranar, como já era esperado, visto ser nessa zona onde a hemaglutinina se encontra ancorada ao envelope viral. Nestas experiências não se formaram membranas, mas sim bicamadas, mas que do ponto de vista do péptido transmembranar não deve representar diferenças significativas. Consegue-se ainda distinguir as diferentes zonas do péptido. A zona inferior do C-terminal possui menos contactos que a zona interior, correspondendo ao segmento intraviral que pode, no entanto, dobrar e inserir os últimos resíduos de novo na membrana. Algo semelhante ocorre para a zona superior dos resíduos da junção entre o ectodomínio da hemaglutinina e o segmento transmembranar.

Como já tinha sido possível observar pelas imagens FH7, a zona superior da HA1 possui uma zona com afinidade para membranas, constatando através desta análise que é constituída maioritariamente pelos *loops* 130 (resíduos 135-138) e 220 (resíduos 225-228). Os resíduos à volta destes dois *loops* e a hélice 190 constituem os locais de ligação do ácido siálico à HA. Mesmo existindo um aminoácido em comum, a ASN137, os resíduos que interagem com o ácido siálico encontram-se no interior da bolsa que funciona como receptor e os resíduos que apresentam contactos com os lípidos encontram-se mais exteriores. Esta observação sugere que talvez a afinidade da HA para

a membrana do hospedeiro não se deva apenas à ligação do ácido siálico ao recetor, mas também a interações dos *loops* 130 e 220 com a membrana.

Curiosamente existe ainda outro local com um nível de contacto moderado, nos resíduos 266-278 e 54/57, que se encontram a aproximadamente 50 resíduos tanto do C-terminal como do N-terminal da HA1. A razão para a ligação de lípidos neste local não é totalmente clara. Contudo é nesta zona central da HA1 onde estes resíduos se encontram que ocorre um encurvamento da parte inferior da HA1 [Xu 2011]. Isto aliado há provável destabilização e separação da HA1, pode indicar um possível papel da HA1 no processo intermedio ou final da fusão, através de interações com a membrana.

Estudos de fotoafinidade mostraram que os péptidos de fusão e o transmembranar são os únicos domínios que interagem diretamente com a bicamada lipídica do hospedeiro ou de membranas alvo [Durrer 1996]. Contudo estes contactos aqui apresentados não necessitam de ser contactos fortes onde, por exemplo, os resíduos se inserem na membrana. Podem ser apenas contactos leves, que devido ao número de subunidades e de hemaglutininas existentes no local de fusão, se multiplicam podendo ganhar alguma relevância.

4. Conclusões

O carácter inovador deste trabalho resulta principalmente da abordagem *multi-scale*, ou seja, no uso de mais de uma resolução. Utilizou-se o campo de forças *coarse-grain* MARTINI para simular a *assembling* dos sistemas e no fim, após a transformação para *united-atom*, utilizando-se o campo de forças 54A7 para simular o sistema obtido com um maior detalhe das moléculas e das suas interações. Assim, pode-se aproveitar o melhor de ambos os campos de forças, uma vez que a assemblagem de péptidos utilizando unicamente o campo de forças 54A7 seria difícil e computacionalmente dispendiosa. Analisando os resultados, verificamos que o procedimento foi efetuado com sucesso, o que era espectável, uma vez que o campo de forças MARTINI para proteínas foi parametrizado utilizando coeficientes de partição água/óleo, tornando-o uma escolha robusta para técnicas de *self-assembling* de péptidos numa membrana. Além disso, já existiam artigos de *self-assembling* utilizando o campo de forças MARTINI com o péptido de fusão [Fuhrmans 2009, Fuhrmans 2012].

O primeiro resultado obtido para o péptido de fusão foi que a estrutura obtida por NMR em micelas [Lorieau 2010] não é estável em água, independentemente do estado de protonação do péptido. Isto vai ao encontro do estudo de dicroísmo circular realizado por Han et al. onde se observava que o péptido (composto por 20 resíduos de aminoácidos) em solução aquosa não tinha estrutura [Han 2000].

Não existem diferenças muito significativas entre os diferentes sistemas estudados (diferentes estados de protonação e formação da bicamada). Embora se observe que o péptido de fusão nos sistemas estudados com uma bicamada pré-equilibrada não encontre tanto na bicamada como nos sistemas de *self-assembling*.

O péptido de fusão encontra-se na interface da membrana com a água, onde a hélice do N-terminal se encontra paralela à membrana e a hélice do C-terminal penetra ligeiramente o interior da bicamada. Porém, obtivemos outras estruturas de interesse, nomeadamente uma estrutura mais inserida na membrana, numa simulação onde se

formou um poro. Numa outra simulação o PF adotou uma estrutura em forma de “L” paralelo à membrana, que se pensa ser um intermediário da fusão [Lorieau 2012].

Os valores obtidos nas nossas análises para este sistema, nomeadamente o conteúdo em hélice alfa do péptido, ângulo de inclinação do péptido na membrana e a posição dos resíduos do péptido na membrana, reproduzirem de forma razoável os resultados da literatura.

Relativamente ao péptido transmembranar, a conformação obtida no final do processo de *multi-scale* foi semelhante à do estudo realizado por Victor et al. [Victor 2012] com o péptido de 27 aminoácidos e do subtipo H5. Contudo o self-assembling utilizando o MARTINI permitiu que os péptidos se encontrassem mais próximos, em contraste com os resultados do estudo de Victor et al., onde apenas se usou o campo de forças 54A7. Observamos ainda que, além da conformação preferencial num arranjo triangular, também existe uma conformação em que os terminais se encontram em linha e não em triângulo.

Ao analisar a posição de cada resíduo do péptido transmembranar na membrana podemos ter uma visão da sua posição na membrana. Os resíduos aromáticos Trp10 e Trp13 delimitam o N-terminal do domínio transmembranar, ambos interagindo com os grupos fosfato da bicamada. Contudo, parece que no C-terminal não é um resíduo aromático a delimitar o domínio transmembranar, mas sim um resíduo carregado positivamente, a Arg37, que por sua vez interage com os grupos carregados negativamente dos grupos fosfato. Esta estrutura é corroborada com as análises de percentagem em hélice (~65.5%) e ângulo de inclinação (~58.4°), que estão dentro do intervalo experimental, como também pelo aumento dos parâmetros de ordem da bicamada.

Relativamente à hemaglutinina utilizou-se o campo de forças ELNEDIN, uma alteração do MARTINI para incluir ligações elásticas. Chegou-se à conclusão que os melhores valores para se utilizar, no estudo da HA, são $R_c = 0.9$ nm e $K_{spring} = 500$ kJ mol⁻¹ nm⁻², valores estes também sugeridos por Periole et al. [Periole 2009] para a maioria das proteínas.

A utilização do método de crescimento lento para transformar o péptido dobradiça de uma estrutura desordenada para uma estrutura em hélice foi apenas eficaz quando o péptido estava isolado em solução. Em contraste, este procedimento não foi eficaz com o péptido dentro da proteínas, dado que, devido ao uso do ELNEDIN a estrutura da HA fica demasiado restringida, sendo impossível ver grandes alterações conformacionais.

Nos estudos de *self-assembly* da HA com moléculas de DMPC, não se conseguiu formar bicamadas entre os eixos x e y, devido ao tamanho da caixa, formando-se, em vez disso, bicelas localizadas principalmente na zona do domínio transmembranar. Além disso, também se formaram várias micelas de DMPC que pareciam interagiam preferencialmente com certos locais da HA. Com novas simulações observamos que de facto existem locais preferenciais de contacto, especialmente nos resíduos à volta do local de ligação com o ácido siálico. Isto sugere que estes resíduos também possam ter um papel na ligação da hemaglutinina à membrana, tendo, possivelmente, um papel na entrada da hemaglutinina pela via alternativa independente do ácido siálico [Stray 2000].

5. Bibliografia

[Alder 1957]

Alder, J., Wainwright, E. (1957) Phase transition for a hard sphere system. *J. Chem. Phys.* 27:1208.

[Allen 2009]

Allen, W.J., Lemkul, J.A., Bevan, D.R. (2009) GridMAT-MD: a grid-based membrane analysis tool for use with molecular dynamics. *J. Comput. Chem.* 30:1952-1958.

[Arinaminpathy 2010]

Arinaminpathy, N., Grenfell, B. (2010) Dynamics of Glycoprotein Charge in the Evolutionary History of Human Influenza. *Plos One.* 5:1-7.

[Armstrong 2000]

Armstrong, R., Kushnir, A., White, J. (2000) The Transmembrane Domain of Influenza Hemagglutinin Exhibits a Stringent Length Requirement to Support the Hemifusion to Fusion Transition. *The Journal of Cell Biology.* 151:425-437.

[Aythya 2009]

Aythya, G., Voth, G. (2009) Systematic multiscale simulation of membrane protein systems. *Current Opinion in Structural Biology.* 19:138-144.

[Basdevant 2006]

Basdevant, N.; Weinstein, H.; Ceruso, M. (2006) Thermodynamic Basis for Promiscuity and Selectivity in Protein-Protein Interactions: PDZ Domains, a Case Study *J. Am. Chem. Soc.* 128:12766–12777.

[Becker 2001]

Becker, U.M., MacKerell Jr., A.D., Roux, B., Watanabe, M. (2001) *Computational Biochemistry and Biophysics.* Marcel Dekker Inc, New York.

[Ben-Naim 1992]

Ben-Naim, A. (1992) "Statistical thermodynamics for chemists and biochemists", Plenum press, New York.

[Bennun 2009]

Bennun, S. V., Hoopes, M. I., Xing, C., Faller, R., (2009) Coarse-grained modeling of lipids. *Chemistry and Physics of Lipids.* 159:59-66.

[Berendsen 1981]

Berendsen, C., et al (1981) "In Intermolecular Forces". Pullmann, B., Ed.; Reidel: Dordrecht. p 331.

[Berendsen 1984]

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81, 8:3684-3690.

[Berendsen 1985]

Berendsen, H.J.C., van der Spoel, D., van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43-56.

[Böttcher 1999]

Böttcher, C., et al. (1999) Structure of influenza haemagglutinin at neutral and at fusogenic pH by electron cryo-microscopy. *FEBS Lett.* 463:255-259.

[Bullough 1994]

Bullough, P., et al (1994) Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature.* 371:37-42.

[Carpenter 2008]

Carpenter, T., et al (2008) Self-Assembly of a Simple Membrane Protein: Coarse-Grained Molecular Dynamics Simulations of the Influenza M2 Channel. *Biophysical Journal.* 85:3790-3801.

[Carr 1993]

Carr, C., Kim, P. (1993) A Spring-Loaded Mechanism for the Conformational Change of Influenza Hemagglutinin. *Cell.* 73:823-832.

[Carr 1997]

Carr, C. M., Chaudhry, C., Kim, P. S. (1997) Influenza hemagglutinin is spring-loaded by a metastable native conformation. *Proc. Natl. Acad. Sci. USA* 94:14306-14313.c

[Cernv 1985]

Cernv, V. (1985) Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of optimization theory and applications.* 45:41-51.

[Chandler 1987]

Chandler, D. (1987) "Introduction to modern statistical mechanics". Oxford university press, New York.

[Chang 2008]

Chang, D., et al (2008) Membrane interaction and structure of the transmembrane domain of influenza hemagglutinin and its fusion peptide complex. *BMC Biology.* 6:2.

[Chen 1995]

Chen, J., et al (1995) A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in *Escherichia coli* into the low-pH-induced conformation. *Proc. Natl. Acad. Sci. USA* 92:12205-12209.

[Chen 1998]

Chen, J., et al (1998) Structure of the Hemagglutinin Precursor Cleavage Site, a Determinant of Influenza Pathogenicity and the Origin of the Labile Conformation. *Cell* 95:409-417.

[Chen 1999]

Chen, J., Skehel, J., Wiley, D. (1999) N-and C-terminal residues combine in the fusion-pH influenza hemagglutinin HAz submit to form an N cap that terminates the triple-stranded coiled coil. *PNAS.* 96:8967-8972.

[Chen 2005]

Chen, B., Takeda, M., Lamb, R. (1998) Influenza Virus Hemagglutinin (H3 Subtype) Requires Palmitoylation of Its Cytoplasmic Tail for Assembly: M1 Proteins of Two Subtypes Differ in Their Ability To Support Assembly. *Journal of Virology.* 79:13673-13684.

[Chernomordik 2008]

Chernomordik, L.V., Kozlov, M.M. (2008) Mechanics of membrane fusion. *Nat. Struct. Mol. Biol.* 15:675-683.

[Choi 2006]

Choi, H. S., Huh, J., Jo, W. H. (2006) Electrostatic Energy Calculation on the pH-Induced Conformational Change of Influenza Virus Hemagglutinin. *Biophys. J.* 91:55-60.

[Dawood 2012]

Dawood, F.S., et al. (2012) Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *The Lancet Infect Dis.* 12:687-695.

[Danieli 1996]

Danieli, T., et al (1996) Membrane Fusion Mediated by the Influenza Virus Hemagglutinin Requires the Concerted Action of at Least Three Hemagglutinin Trimers. *The Journal of Cell Biology.* 133:559-569.

[Dubovskii 2012]

Dubovskii, P. V. (2012) Unusual titration of the membrane-bound artificial hemagglutinin fusion peptide. *Eur. Biophys. J.* 41:1077-1084.

[Durrer 1996]

Durrer, P., et al (1996) H⁺-induced membrane insertion of influenza virus hemagglutinin involves the HA2 aminoterminal fusion peptide but not the coiled coil region. *J. Biol. Chem.* 271:13417-13421.

[Doms 1986]

Doms, R.W., Helenius, A. (1986) Quaternary structure of influenza virus hemagglutinin after acid treatment. *J. Virol.* 60:833-839.

[Dopheide 1981]

Dopheide, T.A., Ward, C.W. (1981) The location of the bromelain cleavage site in a Hong Kong influenza virus Haemagglutinin. *J. Gen. Virol.* 52:367-370.

[Downard 2013]

Downard, K. (2013) Proteotyping for the rapid identification of influenza virus and other biopathogens. *Chem Soc Rev.* 42:8584-8495.

[Epanand 1999]

Epanand, R., et al (1999) The Ectodomain of HA2 of Influenza Virus Promotes Rapid pH Dependent Membrane Fusion. *J. Mol. Biol.* 286:489-503.

[Epanand 2003]

Epanand, R. M. (2003) Fusion peptides and the mechanism of viral fusion. *Biomembranes* 1614:116-121.

[Fleishman 2011]

Fleishman, S., et al. (2011) Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. *Science.* 332:816-820.

[Fontana 2012]

Fontana, J., et al. (2012) Structural Changes in Influenza Virus at Low pH Characterized by Cryo-Electron Tomography. *Journal of Virology.* 86:2919-2929.

[Fouchier 2004]

Fouchier, R.A., et al. (2004) Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc Natl Acad Sci USA*. 101:1356-1361.

[Frenkel 2002]

Frenkel, D., Smit, B. (2002) "Understanding molecular simulation from algorithms to applications". Academic press a division of Harcourt, inc.

[Frolov 2000]

Frolov, V., et al (2000) Multiple Local Contact Sites are Induced by GPI-Linked Influenza Hemagglutinin During Hemifusion and Flickering Pore Formation. *Traffic*. 1:622-630.

[Fuhrmans 2009]

Fuhrmans, M., Knecht, V., Marrink, S. J. (2009) A Single Bicontinuous Cubic Phase Induced by Fusion Peptides. *J. Am. Chem. Soc.* 131:9166-9167.

[Fuhrmans 2012]

Fuhrmans, M., Marrink, S. J. (2012) Molecular View of the Role of Fusion Peptides in Promoting Positive Membrane Curvature. *J. Am. Chem. Soc.* 134:1543-1552.

[Gamblin 2004]

Gamblin, S. J., et al (2004) The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin. *Science* 303:1838-1842.

[Godley 1992]

Godley, L. et al. (1992) Introduction of intersubunit disulfide bonds in the membrane-distal region of the influenza hemagglutinin abolishes membrane fusion activity. *Cell* 68:635-645.

[Gofman 2012]

Gofman, Y., Haliloglu, T., Ben-Tal, N. (2012) The Transmembrane Helix Tilt May Be Determined by the Balance between Precession Entropy and Lipid Perturbation. *Journal of Chemical Theory and Computation*. 8:2896-2904.

[Granseth 2005]

Granseth, E., von Heijine, G., Elofsson, A. (2005) a study of the membrane-water interface region of membrane proteins. *J. Mol. Biol.* 346:377-385.

[Halliday 2001]

Halliday, D., Walker, J., Resnick, R. (2001) "Fundamentals of physics". Wiley, 9th edition.

[Han 2000]

Hann, X., Tamm, L. (2000) A host-guest system to study structure-function relationships of membrane fusion peptides. *PNAS*. 97: 13097-13102.

[Han 2001]

Han, X., Bushweller, J.H., Cafiso, D.S., Tamm, L.K. (2001) Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat. Struct. Biol.* 8:715-720.

[Harris 2006]

Harris, A., et al. (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci USA*. 103:19123-19127.

[Hess 1997]

Hess, B., et al (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* 18:1463–147.

[Hinchliffe 2008]

Hinchliffe, A. (2008) "Molecular Modelling for beginners". Wiley, second edition.

[Hockney 1974]

Hockney, R. W., Goel, S. P., Eastwood, J. (1974) Quiet highresolution computer models of a plasma. *J. Comp. Phys.* 14:148–158.

[Holt 2010]

Holt, A., Killian, A. (2010) Orientation and dynamics of transmembrane peptides: the power of simple models. *Eur Biophys J.* 39:609-621.

[Huang 2002]

Huang, Q., et al (2002) Protonation and stability of the Globular Domain of Influenza Virus Hemagglutinin. *Biophysical Journal.* 82:1050-1058.

[Imai 2006]

Imai, M., Mizuno, T., Kawasaki, K. (2006) Membrane Fusion by Single Influenza Hemagglutinin Trimers. *The Journal of Biological Chemistry.* 281:12729-12735.

[Jin 1994]

Jin, H., Leser, G., Lamb, R. (1994) The influenza virus hemagglutinin cytoplasmic tail is not essential for virus assembly or infectivity. *The EMBO Journal.* 13:5504-5515.

[Jin 1996]

Jin, H., et al (1996) Palmitoylation of the Influenza Virus Hemagglutinin (H3) Is Not Essential for Virus Assembly or Infectivity. *Journal of Virology.* 70:1406-1414.

[Jonhson 2002]

Jonhson, N.P, Mueller, J. (2002) Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bull Hist Med.* 76:105-115.

[Kelley 2011]

Kelley, A., Sternberg, E., (2011) Protein structure prediction on the web: a case study using the Phyre server. *Nature protocols.* 4:363-371.

[Kemble 1992]

Kemble, G.W., et al. (1992) Intermonomer disulfide bonds impair the fusion activity of influenza virus hemagglutinin. *J. Virol.* 66:4940-4950.

[Kemble 1994]

Kemble, G., Danieli, T., White, J. (1994) Lipid-Anchored Influenza Hemagglutinin Promotes Hemifusion, Not Complete Fusion. *Cell.* 76:383-391.

[Killian 1996]

Killian, J., et al (1996) Induction of Nonbilayer Structures in Diacylphosphatidylcholine Model Membranes by Transmembrane α -Helical Peptides: Importance of Hydrophobic Mismatch and Proposed Role of Tryptophans. *Biochemistry.* 35:1037-1045.

[Kim 2011]

Kim, C. et al (2011) The Final Conformation of the Complete Ectodomain of the HA2 Subunit of Influenza Hemagglutinin Can by Itself Drive Low pH-dependent Fusion. *The Journal of Biological Chemistry*. 286:13226-13234.

[Kirkpatrick 1983]

Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P. (1983) Optimization by simulated annealing. *Science*. 220:671-680.

[Kollman 1993]

Kollman, P. (1993) Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* 93:2395-2417.

[Korte 1999]

Korte, T., et al (1999) Conformational Intermediates and Fusion Activity of Influenza Virus Hemagglutinin. *Journal of Virology*. 73:4567-4574.

[Kucerka 2011]

Kucerka, N. Nieh, M.P., Katsaras, J. (2011) Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta*. 1808:2761-2771.

[Kucerka 2004]

Kucerka, N., Kiselev, M.A., Balgavý, P. (2004) Determination of bilayer thickness and lipid surface area in unilamellar dimyristoylphosphatidylcholine vesicles from small-angle neutron scattering curves: a comparison of evaluation methods. *Eur. Biophys. J.* 33:328-334.

[Lagüe 2005]

Lagüe, P., Roux, B., Pastor, R. W. (2005) Molecular Dynamics Simulations of the Influenza Hemagglutinin Fusion Peptide in Micelles and Bilayers: Conformational Analysis of Peptide and Lipids. *J. Mol. Biol.* 354:1129-1141.

[Lai 2007]

Lai, A. L., Tamm, L. K. (2007) Locking the Kink in the Influenza Hemagglutinin Fusion Domain Structure. *J. Biol. Chem.* 33:23946-23956.

[Lai 2010]

Lai, A. L., Tamm, L. K. (2010) Shallow Boomerang-shaped Influenza Hemagglutinin G13A Mutant Structure Promotes Leaky Membrane Fusion. *J. Biol. Chem.* 48:37467-37475.

[Lakadamyali 2003]

Lakadamyali, M., et al. (2003) Visualizing infection of individual influenza viruses. *Proc Natl Acad Sci USA*. 100:9280-9285.

[Langosch 2007]

Langosch, D. Hofmann, M., Ungermann, C. (2007) The role of transmembrane domains in membrane fusion. *Cell*. 64:850-864.

[Leach 2001]

Leach, A.R. (2001) "Molecular Modelling principles and applications". Pearson Education, second edition.

[Lear 1987]

Lear, J.D., De Grado, W.F. (1987) Membrane binding and conformational properties of peptide representing the amino terminus of influenza virus HA2. *J. Biol. Chem.* 262:6500-6505.

[Légaré 2012]

Légaré, S., Lague, P. (2012) The Influenza Fusion Peptide Adopts a Flexible Flat V Conformation in Membranes. *Biophys. J.* 102:2270-2278.

[Leikina 2000]

Leikina, E., Chernomordik, L.V. (2000) Reversible Merger of Membranes at the Early Stage of Influenza Hemagglutinin-mediated Fusion. *Mol. Biol. Cell.* 11:2359-2371.

[Leikina 2002]

Leikina, E., et al (2002) Reversible stages of the low-pH-triggered conformational change in influenza virus hemagglutinin. *The EMBO Journal.* 21:5701-5710.

[Li 2008]

Li, Z., et al (2008) Length Requirements for Membrane Fusion of Influenza Virus Hemagglutinin Peptide Linkers to Transmembrane or Fusion Peptide Domains. *Journal of Virology.* 82:6337-6348.

[Lin 1998]

Lin, S., et al (1998) Mutations in the Middle of the Transmembrane Domain Reverse the Polarity of transport of the Influenza Virus Hemagglutinin in MDCK Epithelial Cells. *The Journal of Cell Biology.* 42:51-57.

[Lin 2012]

Lin, Y.P., et al. (2012) Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. *Proc Natl Acad Sci USA.* 109:21474-21479.

[Lindorff-Lasen 2011]

Lindorff-Lasen, K., Piana, S., Dror, R.O., Shaw D.E. (2011) How fast-folding proteins fold. *Science.* 334:517-520.

[Lindstorm 2004]

Lindstorm, S.E, Cox, N.J, Klimov, A. (2004) Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957-1972: evidence for genetic divergence and multiple reassortment events. *Virology.* 328:101-119.

[Lorieau 2010]

Lorieau, J., Louis, J., Bax, A. (2011) The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. *PNAS.* 107:11341-11346.

[Lorieau 2011]

Lorieau, J., Louis, J. Bax, A. (2011) Helical Hairpin Structure of Influenza Hemagglutinin Fusion Peptide Stabilized by Charge-Dipole Interactions between the N-Terminal Amino Group and the Second Helix. *J. Am. Chem. Soc.* 133: 2824-2827.

[Lorieau 2012]

Lorieau, J. et al (2012) pH-tiggered, activated-state conformations of the influenza hemagglutinin fusion peptide revealed by NMR. *PNAS.* 109:19994-19999.

[Luo 2012]

Luo, M. (2012) Influenza virus entry. Em Rossman, M. ed. "Viral molecular machines".

[Macosko 1997]

Macosko, J.C., Kim, C.H., Shin, Y.K. (1997) The membrane topology of the fusion peptide region of influenza hemagglutinin determined by spin-labeling EPR. *J. Mol. Biol.* 267:1139-1148.

[Madhusoodanan 2003]

Madhusoodanan, M., Lazaridis, T. (2003) Investigation of Pathways for the Low-pH Conformational Transition in Influenza Hemagglutinin. *Biophys. J.* 84:1926-1939.

[Mandell 2010]

Mandell, G.L, Bennett, J.E, Dolin, R. (2010) "Infectious diseases".

[Markosyan 2000]

Markosyan, R., Cohen, F., Melikyan, G. (2000) The Lipid-anchored Ectodomain of Influenza Virus Hemagglutinin (GPI-HA) Is Capable of Inducing Nonenlarging Fusion Pores. *Molecular Biology of the Cell.* 11:1143-1152.

[Markosyan 2001]

Markosyan, R., Melikyan, G., Cohen, F. (2001) Evolution of Intermediates of Influenza Virus Hemagglutinin-Mediated Fusion Revealed by Kinetic Measurements of Pore Formation. *Bophysical Journal.* 80:812-821.

[Marrink 2003]

Marrink, S., J., Mark, A., E. (2003) Molecular dynamics simulation of the formation, structure, and dynamics of small phospholipid vesicles, *J. Am. Chem. Soc.* 125:15233-15242.

[Marrink 2004]

Marrink, S., Vries, A., Mark, A. (2004) Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B.* 108:750-760.

[Marrink 2007]

Marrink, S., et al (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem.* 111: 7812-7824.

[Marrink 2008]

Marrink, S., et al (2008) The MARTINI Forcefield. Em G. Voth ed. "Coarse graining of condensed phase and biomolecular systems".

[Marrink 2010]

Marrink, S.J., Periole, X., Tieleman, D.P., Vries, A.H. (2010) Comment on "On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models" by M. Winger, D. Trzesniak, R. Baron and W. F. Gunsteren. *Phys. Chem. Chem. Phys.* 11:1934. *Phys Chem Chem Phys.* 12:2254-2256.

[Martin 1999]

Martin, I. Ruyschaert, J., Epand, R. (1999) Role of the N-terminal peptides of viral envelope proteins in membrane fusion. *Advances Drug Delivery Reviews.* 38:233-255.

[McCammon 1977]

McCammon, A., Gelin, R., Karplus, M. (1997) Dynamics of folded proteins. *Nature.* 267(5612):585-590.

[Melikyan 1995]

Melikyan, G., White, J., Cohen, F. (1995) GPI-anchored Influenza Hemagglutinin Induces Hemifusion to both Red Blood Cell and Planar Bilayer Membranes. *The Journal of Cell Biology*. 131:679-691.

[Melikyan 1997]

Melikyan, G., et al (1997) The Role of the Cytoplasmic Tail Region of Influenza Virus Hemagglutinin in Formation and Growth of Fusion Pores. *Virology*. 235:118-128.

[Melikyan 1999]

Melikyan, G., Lin, S., Roth, M., Cohen, F. (1999) Amino Acid Sequence Requirements of the Transmembrane and Cytoplasmic Domains of Influenza Virus Hemagglutinin for Viable Membrane Fusion. *Molecular Biology of the Cell*. 10:1821-1836.

[Miyamoto 1992]

Miyamoto, S., Kollman, P. A. (1992) SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comp. Chem*. 13:952–962.

[Mohanram 2012]

Mohanram, H., et al (2012) NMR Structure, Localization, and Vesicle Fusion of Chikungunya Virus Fusion Peptide. 51:7863-7872.

[Mokrab 2011]

Mokrab, Y., Sansom, M. S. P. (2011) Interaction of Diverse Voltage Sensor Homologs with Lipid Bilayers Revealed by Self-Assembly Simulations. *Biophys. J*. 100:875-884.

[Monticelli 2008]

Monticelli, L., et al (2008) The MARTINI coarse-grained force field: Extension to proteins, *J. Chem. Theory and Comput*. 4:819-834.

[Mueller 2008]

Mueller, D. (2008) Histidine protonation and the activation of viral fusion proteins. *Biochemical Soc Trans*. 36:43-45.

[Naeve 1990]

Naeve, C.W., Williams, D. (1990) Fatty acids on the A/Japan/305/57 influenza virus hemagglutinin have a role in membrane fusion. *EMBO J*. 9:3857-3866.

[Nagle 2000]

Nagle, J.F., Tristram-Nagle, S. (2000) Structure of lipid bilayers. *Bioch. Biophys. Acta*. 1469:159-195.

[Neumann 2011]

Neumann, S., Langosch, D. (2011) Conserved conformational dynamics of membrane fusion protein transmembrane domains and flanking regions indicated by sequence statistics. *Proteins*. p.2418-2425.

[Nezil1992]

Nezil, A., Bloom, M. (1992) Combined influence of cholesterol and synthetic amphiphilic peptides upon bilayer thickness in model membranes. *Biophys J*. 61(5):1176-1183.

[Nishizawa 2010]

Nishizawa, M., Nishizawa, K. (2010) Molecular dynamics simulation analyses of viral fusion peptides in membranes prone to phase transition: effects on membrane curvature, phase behavior and lipid-water interface destabilization. *Journal of Biophysical Chemistry*. 1:19-32.

[Nobusawa 1991]

Nobusawa, E., et al. (1991) Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology*. 182:475-485.

[Nugent 2011]

Nugent, T., Jones, D. (2011) Membrane protein structural bioinformatics. *Journal of Structural Biology*. 179:327-337.

[Odell 1997]

Odell, D. et al. (1997) Influence of Membrane Anchoring and Cytoplasmic Domains on the Fusogenic Activity of Vesicular Stomatitis Virus Glycoprotein G. *J. Virol.* 71:7996-8000.

[Ohuchi 1998]

Ohuchi, M. et al. (1998) Elongation of the Cytoplasmic Tail Interferes with the Fusion Activity of Influenza Virus Hemagglutinin. *J. Virol.* 72:3554-3559.

[Oräd 2004]

Oräd, G., Lindblom G. (2004) Lateral diffusion studied by pulsed field gradient NMR on oriented lipid membranes. *Magn. Reson. Chem.* 42:123-131.

[Panahi 2010]

Panahi, A., Feig, M. (2010) Conformational Sampling of Influenza Fusion Peptide in Membrane Bilayers as a Function of Termini and Protonation States. *J. Phys. Chem. B.* 114:1407-1416.

[Parton 2013]

Parton, D. L., Tek, A., Baaden, M., Sansom, M. S. P. (2013) Formation of Raft-Like Assemblies within Clusters of Influenza Hemagglutinin Observed by MD Simulations. *PLoS Comput. Biol.* 9:1-10.

[Periole 2007]

Periole, X., et al (2007) Protein-Coupled Receptors Self-Assemble in Dynamics Simulations of Model Bilayers. *J. Am. Chem. Soc.* 129:10126–10132.

[Periole 2009]

Periole, X., Carvalli, M., Marrink, S., Ceruso, M. A., Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* 5:2531-2543.

[Periole 2012]

Periole, X., et al (2012) Structural Determinants of the Supramolecular Organization of G Protein-Coupled Receptors in Bilayers. *J. Am. Chem. Soc.* 134:10959-10965.

[Petrache 1998]

Petrache, H.I. et al. (1998) Interbilayer interactions from high-resolution x-ray scattering. *Phys. Rev. E.* 57:7014-7018.

[Petrache 2000]

Petrache, I. (2000) Dodd SW, Brown MF. *Biophysics journal.* 79:3172.

[Poger 2010]

Poger, D., Mark, A. (2010) On the Validation of Molecular Dynamics Simulations of Saturated and *cis*-Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *Journal of Chemical Theory and Computation.* 6:325-336.

[PyMOL]

The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

[Rahman1964]

Rahman, A. (1964) Correlations in the motion of atoms in liquid argon. *Physical review*. 136:A405-A411.

[Ramachandran 2008]

Ramachandran, K.I., et al. (2008) Chapter 4 Hückel molecular Orbital Theory. In Ramachandran, K.I., Deepa, G., Namboori K. "Computational chemistry and molecular modeling: Principles and applications. Springer 2008

[Rapaport 2004]

Rapaport, D.C. (2004) *The Art of Molecular Dynamics Simulations*. Cambridge University Press, New York.

[Reed 2009]

Reed, M., et al (2009) Amino Acid Residues in the Fusion Peptide Pocket Regulate the pH of Activation of the H5N1 Influenza Virus Hemagglutinin Protein. *Journal of Virology*. 83:3568-3580.

[Rogers 1983]

Rogers, G.N., Paulson, J.C. (1983) Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology*. 127:361-373.

[Rose 2009]

Rose, A., et al (2009) RHYTHM-a server to predict the orientation of transmembrane helices in channels and membrane-coils. *Nucleic Acids Research*. 37:575-580.

[Ruigrok 1986]

Ruigrok, R. W. (1986) Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. *Virology* 155:484-497.

[Russ 2000]

Russ, W.P., Engelman, D.M. (2000) The GxxxG Motif: A Framework for Transmembrane Helix-Helix Association. *J. Mol. Biol*. 296:911-919.

[Ryckaert 1977]

Ryckaert, J. P., Ciccotti, G., Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comp. Phys*. 23:327-341.

[Rzepiela 2009]

Rzepiela, A., et al (2009) Membrane poration by antimicrobial peptides combining atomistic and coarse-grained descriptions. *Faraday Discuss*. 144:431-443.

[Rzepiela 2010]

Rzepiela, A., et al (2010) Software News and Update Reconstruction of Atomistic Details from Coarse-Grained Structures. *J. Comput. Chem*. 31:1333-1343.

[Rzepiela 2011]

Rzepiela, A., Louhivuori, M., Peter, C., Marrink, S. (2011) Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. *Phys. Chem. Chem. Phys*. 13:10437-10448.

[Scheiffele 1997]

Scheiffele, P., Roth, M., Simons, K. (1997) Interaction of influenza virus hemagglutinin with sphingolipid-cholesterol membrane domains via its transmembrane domain. *The EMBO Journal*. 16:5501-5508.

[Schiffer 1992]

Schiffer, M., Chang, C., Stevens, F. (1992) The functions of the tryptophan residues in membrane proteins. *Protein Engineering*. 5:213-214.-Be

[Schmid 2011]

Schmid, N., et al (2011) Definition and testing of the GROMOS force-field versions 54A7 e 54B7, *Eur. Biophys. J.* 40:843-856.

[Schroth-Diez 2000]

Schoroth-Diez, B., et al (2000) The Role of the Transmembrane and of the Intraviral Domain of Glycoproteins in Membrane Fusion of Enveloped Viruses. *Bioscience Reports*. 20:571-592.

[Scolari 2009]

Scolari, S., et al (2009) Lateral Distribution of the Transmembrane Domain of Influenza Virus Hemagglutinin Revealed by Time-resolved Fluorescence Imaging. *The Journal of Biological Chemistry*. 284:15708-15716.

[Scott 2008]

Scott, K. A., et al (2008) Coarse-Grained MD Simulations of Membrane Protein-Bilayer Self-Assembly. *Structure* 16:621-630.

[Sengupta 2008]

Sengupta, D., et al (2008) Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochimica et Biophysica Acta*. 1778:2308-2317.

[Seo 2012]

Seo, M., Rauscher, S. Pomès, Tieleman, D. (2012) Improving Internal Peptide Dynamics in the Coarse-Grained MARTINI Model: Toward Large-Scale Simulations of Amyloid –and Elastin- like Peptides. *J. Chem. Theory Comput*. 8:1774-1785.

[Sheats 1978]

Sheats, J.R., McConnell, H.M. (1978) A photochemical technique for measuring lateral diffusion of spin-labeled phospholipids in membranes. *Proc. Natl. Acad. Sci. USA*. 75:4661-4663.

[Skehel 1982]

Skehel, J.J., et al. (1982) Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc Natl Acad Sci USA*. 79:968-972.

[Skehel 2000]

Skehel, J.J., Wiley, D. (2000) Receptor Binding and Membrane Fusion in Virus Entry: The Influenza Hemagglutinin. *Annu. Rev. Biochem.* 69:531-69.

[Sieczkarski 2002]

Sieczkarski, S.B., Whittaker, G.R. (2002) Dissecting virus entry via endocytosis. *J. Gen. Virol.* 83:1535-1545.

[Sievers 2011]

Sievers, F., et al (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Molecular systems biology*. 7:539.

[Simpson 1992]

Simpson, D.A., Lamb R.A. (1992) Alterations to influenza virus hemagglutinin cytoplasmic tail modulate virus infectivity. *J. Virol.* 66:790-803

[Smith 2009]

Smith, G., et al. (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*. 459:1122-1125

[Söllner 2004]

Söllner, T. (2004) Intracellular and viral membrane fusion: a uniting mechanism. *Current Opinion in Cell Biology*. 16:429-435.

[Stain Tsai 2002]

C. Stain Tsai (2002) "An Introduction to Computational Biochemistry". Wiley-Liss, Inc.

[Stansfeld 2011]

Stansfeld, P. J., Sansom, M. S. P. (2011) From Coarse Grained to Atomistic: A Serial Multiscale Approach to Membrane Protein Simulations. *J. Chem. Theory Comput.* 7:1157-1166.

[Stark 2013]

Stark, A., Andrews, C., Elcock, A. (2013) Toward Optimized Potential Functions for Protein-Protein Interactions in Aqueous Solutions: Osmotic Second Virial Coefficient Calculations Using the MARTINI Coarse-Grained Force Field. *Journal of Chemical Theory and Computation*. 9:4176-4185.

[Strandberg 2002]

Strandberg, E., et al. (2002) Lipid dependence of membrane anchoring properties and snorkeling behavior of aromatic and charged residues in transmembrane peptides. *Biochemistry*. 41:7190-7198.

[Stray 2000]

Stray, S.J., Cummings, R.D., Air, G.M (2000) Influenza virus infection of desialylated cells. *Glycobiology*. 10:649-658.

[Stillinger 1974]

Stillinger, H., Rahman, A. (1974) Improved simulation of liquid water by molecular dynamics. *J. Chem. Phys.* 60:1545-1557.

[Takemoto 1996]

Takemoto, D.K, Skehel, J.J., Wiley, D.C. (1996) A surface plasmon resonance assay for the binding of influenza virus hemagglutinin to its sialic acid receptor. *Virology*. 217:452-458.

[Tamm 2003]

Tamm, L., Jonathan, C., Kiessling, V. (2003) Membrane fusion: a structural perspective on the interplay of lipids and proteins. *Current Opinion in Structural Biology*. 13:453-466.

[Tatolian2000]

Tatolian, S., Tamm, L. (2000) Secondary structure, orientation, oligomerization and lipid interactions of the transmembrane domain of influenza hemagglutinin. *Biochemistry*. 39:496-507.

[Taubenberger 1998]

Taubenberger, J.K. (1998) Commentary, Influenza virus hemagglutinin cleavage into HA1, HA2: No laughing matter. *Proc Natl Acad Sci USA*. 95:9713-9715.

[Thoennes 2008]

Thoennes, S., et al (2008) Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion. *Virology*. 370: 403-414.

[Thøgersen 2008]

Thøgersen, L., et al (2008) Peptide Aggregation and Pore Formation in a Lipid Bilayer: A Combined Coarse-Grained and All Atom Molecular Dynamics Study. *Biophys. J*. 95:4337-3347.

[Thompson 2006]

Thompaon, C.I., Barclay, W.S., Zambon, M.C., Pickles, R.J. (2006) Infection of human airway epithelium by human and avian strains of influenza a virus. *J. Virol*. 80:8060-8068.

[Tironi 1995]

Tironi, G. et al (1995) A Generalized Reaction Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* 102:5451–5459.

[Treanor 2010]

Treanor, J.J. (2010) Influenza Viruses, Including Avian Influenza and Swine Influenza. Em Churchill Livingstone Elsevier (7th ed.) “Principles and Practice of Infectious Diseases”.

[Treanor 2012]

Treanor, J.J. et al. (2012) Effectiveness of seasonal influenza vaccines in the United States during a season with circulation of all three vaccine strains. *Clin. Infect Dis*. 55:951-959

[Tusnády 1998]

Tusnády, E., Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Bio*. 283:489-506.

[Vaidya 2010]

Vaidya, N., Huang, H., Tagaki, S. (2010) Coarse Grained Molecular Dynamics Simulation of Interaction between Hemagglutinin Fusion Peptides and Lipid Bilayer Membranes. *Adv. Appl. Math. Mech*. 2:430-450.

[Vaccaro 2005]

Vaccaro, L. et al. (2005) Plasticity of Influenza Haemagglutinin Fusion Peptides and Their Interaction with Lipid Bilayers. *Biophys. J*. 88:25-36.

[Van der Spoel 2005]

D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen, *Gromacs User Manual version 4.0*, www.gromacs.org (2005).

[Van Gunsteren 1987]

Van Gunsteren, W.F., Berendsen, H.J.C. (1987) “Groningen molecular simulation (gromos) library manual”. Technical report, Biomos, Groningen.

[Van Gunsteren 1990]

Van Gunsteren, W.F., Berendsen, H.J.C., (1990) “Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry”. *Chem. Inr. Ed. Engl*, 29, 992.

[Van Gunsteren 1996]

Van Gunsteren, F., et al (1997) "Biomolecular Simulation: The GROMOS96 Manual and User Guide"; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v.: Zürich, Groningen.

[Van Gunsteren 2006]

Van Gunsteren, W.F. et al., (2006) "Biomolecular modeling: Goals, problems, perspectives". *Angew Chem. Int. Ed. Engl.* 45, 4064.

[Van Holde 1998]

Van Holde, K.E., Johnson, W.C., Ho, P.S. (1998) "Principles of physical biochemistry". Prentice-hall, inc., New jersey.

[Varki 2009]

Varki, A., Schauer R. (2009) Chapter 14 Sialic acids. Em Varki, A. "Essentials of Glycobiology".

[Varma 1998]

Varma, R., Mayor, S. (1998) GPI- anchored proteins are organized in submicron domains at the cell surface. *Nature.* 394:798-801.

[Victor 2012]

Victor, B., Baptista, A., Soares, C. (2012) Structural Determinants for the Membrane Insertion of the Transmembrane Peptide of Hemagglutinin from Influenza Virus. *Journal of Chemical Information and Modeling.* 52:3001-3012.

[Victor 2015]

Victor, B., Lousa, D., Antunes, J., Soares, C. (2015) Self-Assembly Molecular Dynamics Simulations Shed Light into the Interaction of the Influenza Fusion Peptide with a Membrane Bilayer. *Journal of Chemical Information and Modeling.* 55:795-805.

[Vostrikov 2010]

Vostrikov, V.V, Daily, A.E., Greathouse, D.V., Koeppe, R.E. (2010) Charged or aromatic anchor residue dependence of transmembrane peptide tilt. *J. Biol. Chem.* 285:31723-31730.

[Voth 2008]

Voth, G.A. (2008) "Coarse-graining of condensed phase and biomolecular systems", chapter 2. CRC, Boca Raton USA.

[Wan 2012]

Wan, C., Han, W., Wu, Y. (2012) Parameterization of PACE Force Field for Membrane Environment and Simulation of Helical Peptides and Helix-Helix Association. *J. Chem. Theory Comput.* 8:300-313.

[Wang 2007]

Wang, Q., Tian, X., Chen, X., Ma, J., (2007) Structural basis for receptor specificity of influenza B virus hemagglutinin. *Proc Natl Acad Sci USA.* 104:16874-16879.

[Wassenaar 2013]

Wassenaar, T., et al (2013) Mixing MARTINI: Electrostatic Coupling in Hybrid Atomistic – Coarse-Grained Biomolecular Simulations. *The Journal of Physical Chemistry.* 117:3516-3530.

[Weber 1994]

Weber, T., et al (1994) Evidence for H-induced Insertion of Influenza Hemagglutinin HA2 N-terminal Segment into Viral Membrane. *The Journal of Biological Chemistry*. 269:18353-18358.

[Wee 2008]

Wee, L., et al (2008) The interaction of phospholipase A2 with a phospholipid bilayer: Coarse grained molecular dynamics simulations. *Biophys. J.* 95:1649-1657.

[Weis 1990]

Weis, W.I., Beünger, A.T., Skehel, J.J., Wiley, D.C. (1990) Refinement of the influenza virus hemagglutinin by simulated annealing. *J. Mol. Biol.* 212:737-761.

[Wel 2000]

Wel, P., et al (2000) Tryptophan-Anchored Transmembrane Peptides Promote Formation of Nonlamellar Phases in Phosphatidylethanolamine Model Membranes in a Mismatch-Dependent Manner. *Biochemistry*. 39:3124-3133.

[White1987]

White, J.M., Wilson, I.A. (1987) Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin. *J. Cell Biol.* 105:2887-2896.

[Wilk 1996]

Wilk, T. et al. (1996) Glycoprotein Incorporation and HIV-1 Infectivity Despite Exchange of the gp160 Membrane-Spanning Domain. *Virology*. 218:269-274.

[Wilson 1981]

Wilson, I.A., Skehel, J.J., Wiley, D.C (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*. 289:366-373.

[Xu 2011]

Xu, R., Wilsom, I. (2011) Structural Characterization of an Early Fusion Intermediate of Influenza Virus Hemagglutinin. *Journal of Virology*. 85:5172-5182.

[Yamniuk 2004]

Yamniuk, A. P.; Vogel, H. J. (2004) Calmodulin's Flexibility Allows for Promiscuity in Its Interactions with Target Proteins and Peptides *Mol. Biotechnol.* 27:33–57.

[Yau 1998]

Yau, W., et al (1998) The Preference of Tryptophan for Membrane Interfaces. *Biochemistry*. 37:14713-14718

[Zambon 2001]

Zambon, M. C. (2001) The pathogenesis of influenza in humans. *Rev. Med. Virol.* 11:227-241.

[Zurcher 1994]

Zurcher, T., Luo, G., Palese, P. (1994) Mutations at palmitoylation sites of the influenza virus hemagglutinin affect virus formation. *J. Virol.* 68:5748-5754.