



Universidade Federal de Pernambuco  
Centro de Ciências Exatas e da Natureza  
Departamento de Química Fundamental

# **Mecanismo de Ação da Enzima Protease de Serina NS2B-NS3 do Vírus da Dengue**

**Maria Carolina Pacheco Lima**

Recife  
2016



Universidade Federal de Pernambuco  
Centro de Ciências Exatas e da Natureza  
Departamento de Química Fundamental

## **Mecanismo de Ação da Enzima Protease de Serina NS2B-NS3 do Vírus da Dengue**

Maria Carolina Pacheco Lima

Tese apresentada ao Programa de  
Pós-Graduação em Química  
Fundamental da UFPE como parte  
dos requisitos para obtenção do grau  
de Doutor em Química.

Orientador: Prof. Dr. Gustavo de Miranda Seabra  
Co-Orientador: Prof. Dr. Ricardo L. Longo

Recife  
2016

Catálogo na fonte  
Bibliotecária Leonice Maria Cavalcante CRB 4-P1714

L732m    Lima, Maria Carolina Pacheco.  
            Mecanismo de Ação da Enzima Protease de Serina  
            NS2B-NS3 do Vírus da Dengue/ Maria Carolina Pacheco Lima– 2016.  
            169 f.: fig.; tab.

            Orientador: Prof. Dr. Gustavo de Miranda Seabra.  
            Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN.  
            Química. Recife, 2017.  
            Inclui referências e apêndices.

            1.    Dengue.NS2B-NS3 protease. 2. Dinâmica molecular. I. Seabra,  
            Gustavo de Miranda. (Orientador). II. Título.

616.921              CDD (22. ed.)                              UFPE-FQ2017-53

*À minha família e aos meus amigos.*



**MARIA CAROLINA PACHECO LIMA**

**Mecanismo de Ação da Enzima Protease de  
Serina NS2B-NS3 do Vírus da Dengue**

Tese apresentada ao Programa de Pós-Graduação em Química, do Departamento de Química Fundamental, Universidade Federal de Pernambuco, como requisito para a obtenção do título de Doutor em Química.

Aprovado em: 16/09/2016

**BANCA EXAMINADORA**

---

**Profº Gustavo de Miranda Seabra**  
Departamento de Química Fundamental  
Universidade Federal de Pernambuco

---

**Profº João Bosco Paraíso da Silva**  
Departamento de Química Fundamental  
Universidade Federal de Pernambuco

---

**Profº Carlos Henrique Madeiros Castelletti**  
Instituto Agrônômico de Pernambuco

---

**Profº Willian Ricardo Rocha**  
Departamento de Química  
Universidade Federal de Minas Gerais

---

**Profº Jorge Luiz Neves**  
Departamento de Química Fundamental  
Universidade Federal de Pernambuco

## AGRADECIMENTOS

---

Agradeço aos meus orientadores, professor Dr. Gustavo de Miranda Seabra, professor Dr. Ricardo L. Longo e professora Dra. Thereza Soares, pelos ensinamentos, discussões, conselhos, oportunidades e paciência que permitiram a execução e desenvolvimento deste trabalho.

Ao Gabriel Costa da Hora, pelas dicas durante a escrita deste documento e pelo suporte emocional.

Aos amigos de laboratório do LQTC (Júlio Cosme, Victor Rusu, Diego de Paula, Miguel Ângelo, Fred Pontes, Paulo McMiller, Carlos Henrique, Renaldo Júnior, Albano Neto, Eduardo Aguiar), pelas ideias trocadas e ajuda durante esse tempo.

À administração do CENAPAD-PE, *Oak Ridge Leadership Facilities* e LQTC pelo suporte técnico, pois sem a estrutura computacional não seria possível à realização do trabalho.

À CAPES e ao INAMI pelo suporte financeiro.

## RESUMO

---

A Dengue é uma doença que atinge milhões de pessoas em todo o mundo, e pode levar à morte. O Brasil é o país que apresenta o maior número de casos combinados de Dengue e febre hemorrágica da Dengue (D/FHD). No entanto, não existem medicamentos específicos para o tratamento de D/FHD e, uma vez infectado, as recomendações da OMS estão limitadas à observação e tratamento sintomático.

Esforços recentes têm revelado uma série de proteínas essenciais para o ciclo de vida do vírus do Dengue, que podem ser utilizados como alvos para novos medicamentos. O objetivo deste estudo é entender o mecanismo de ação de uma dessas enzimas, a proteína não estrutural NS3 protease complexada com o seu cofator NS2B (NS2B-NS3), responsável pela clivagem da poliproteína viral durante a etapa de replicação do vírus, e a compreensão da interação desta proteína com um inibidor e dois modelos do substrato através de estudos de dinâmica molecular. Aqui usamos DM-QM/MM e *Umbrella Sampling* para estudar o primeiro passo (acilação) da reação catalisada pela NS2B-NS3, usando o método semi-empírico PDDG/PM3 para o subsistema QM, e o campo de força *Amber* FF99SB para o subsistema MM.

É necessária a construção de um modelo estrutural que represente a NS2B-NS3 protease na forma ativa (fechada), chamada DENF2f, e o estudo do comportamento dos sistemas por dinâmica molecular da estrutura NS2B-NS3 inativa (aberta), DENV2a, assim como o modelo de estrutura ativa na presença de um inibidor, e também com dois modelos diferentes de substrato, a fim de conhecer as peculiaridades de cada sistema e validar o modelo estrutural da NS2B-NS3 protease ativo do vírus da Dengue.

Os resultados revelam que a protease NS2B-NS3 pode existir também na conformação fechada em solução, mesmo na ausência do inibidor. O DENV2f revelou-se um modelo realístico para o estudo de modelagem para triagem de inibidores, assim como para o estudo do mecanismo de reação durante a replicação viral. Com relação ao mecanismo de reação, nossos resultados indicam que o ataque nucleofílico da SER117 sobre o substrato ocorre gradualmente, em que uma transferência de próton para a HIS33 primeiro ativa a SER117, que só depois ataca o substrato. O passo determinante para velocidade de reação é a ativação da SER117, com uma barreira de 24,1 kcal/mol. A cavidade oxiânion é formada por moléculas de água que estabilizam a carga negativa formada sobre o oxigênio carboxílico do substrato. O passo final da reação é a transferência do próton da HIS33 ao nitrogênio do substrato, o que acontece com uma barreira de ativação mais baixa de 5,1 kcal/mol, e induz diretamente a quebra da ligação peptídica.

**Palavras-chave:** Dengue. NS2B-NS3 protease. Dinâmica molecular

## ABSTRACT

---

Dengue is a vector tropical disease affecting millions of people, and may lead to death. Brazil is the country with the highest number of Dengue and Dengue hemorrhagic fever (D/DHF) cases combined. However, there are no specific medicines for the treatment of D/DHF and, once infected, the WHO recommendations are limited to observation and symptomatic treatment.

Recent efforts have revealed a series of proteins essential to the Dengue virus's life cycle, which may be used as targets for new medicines. The objective of this study is to understand the mechanism of action of such an enzyme, the nonstructural protein NS3 protease complexed with the cofactor NS2B (NS2B-NS3), responsible for cleaving the viral polyprotein during the virus replication step, and understanding this protein's interaction with its substrate and inhibitor through molecular dynamics studies.

For the reaction study, a preparation is needed that includes the construction of a model for the protease structure in active (closed) form, named DENF2f, and the study of the systems' behavior by molecular dynamics of the NS2B-NS3 inactive structure (open), the model for active structure, and the model in the presence of an inhibitor or and two different models of the substrate, in order to know the peculiarities of each system and validate the structural model of the active NS2B-NS3 Dengue virus protease. Here we used QM/MM-MD and Umbrella Sampling to study the first step (acylation) of the reaction catalyzed by NS2B-NS3pro, using the PDDG/PM3 semi-empirical Hamiltonian for the QM subsystem, and Amber FF99SB for the MM subsystem.

The results show that the NS2B-NS3 protease may also exist in a closed conformation in solution, even in the absence of inhibitor. The DENV2f proved to be a realistic model for a modeling study for screening for inhibitors as well as for the study of the reaction mechanism during viral replication. Our results indicate that the nucleophilic attack on the substrate by SER117 occurs in a stepwise manner, where a proton transfer to HIS33 first activates SER117, which only later attacks the substrate. The rate-determining step is the SER117 activation, with a barrier of 24.1 kcal/mol. Water molecules completing the oxyanion hole stabilize the negative charge formed on the carbonyl oxygen of the substrate. The final step in the process is a proton transfer from HIS33 to the substrate's nitrogen, which happens with a lower barrier of 5.1 kcal/mol, and leads directly to the peptide bond breaking.

**Keywords:** Dengue. NS2B-NS3 protease. Molecular dynamics.

## SIGLAS E ABREVIações

---

|             |                                                                  |
|-------------|------------------------------------------------------------------|
| AMBER       | <i>Assisted Model Building with Energy Refinement</i>            |
| CHARMM      | <i>Chemistry at Harvard Macromolecular Mechanics</i>             |
| DENV        | Vírus da Dengue                                                  |
| DENV2       | Vírus da Dengue sorotipo 2                                       |
| DENV2a      | Estrutura aberta da NS2B-NS3 protease para a DENV2 (PDB-ID:2FOM) |
| DENV2f      | Modelo da estrutura fechada da NS2B-NS3 protease para a DENV2    |
| DENV2f-INIB | Modelo da estrutura fechada da NS2B-NS3 protease com inibidor    |
| DENV2f-SUBS | Modelo da estrutura fechada da NS2B-NS3 protease com substrato   |
| GROMOS      | <i>Groningen Molecular Simulation</i>                            |
| HID         | Histidina (HIS) protonada na posição delta                       |
| MD          | <i>Molecular Dynamics</i>                                        |
| MM          | <i>Molecular Mechanics</i>                                       |
| QM/MM       | <i>Quantum Mechanics/Molecular Mechanics</i>                     |
| OMS         | <i>Organização Mundial da Saúde</i>                              |
| PDB         | <i>Protein Data Bank</i>                                         |
| PME         | <i>Particle Mesh Ewald</i>                                       |
| PMEMD       | <i>Particle Mesh Ewald Molecular Dynamics</i>                    |
| RESP        | <i>Restrained Electrostatic Potential</i>                        |
| RMN         | Ressonância Magnética Nuclear                                    |
| SANDER      | <i>Simulated Annealing with NMR-Derived Energy Restraints</i>    |
| WHAM        | <i>Weighted Histogram Analysis Method</i>                        |
| WHO         | <i>World Health Organization</i>                                 |
| WNV         | <i>West Nile Virus</i>                                           |

## LISTA DE FIGURAS

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1: Mapa da Dengue no Mundo. As esferas em azul: extensão geográfica da Dengue de 2000 até 2007, região colorida: risco de transmissão da Dengue, a linha em preto demarca a região onde o vetor que transmite o vírus existe. Adaptado da referência (WORLD HEALTH ORGANIZATION, 2009a).....                                                                                                                                                                                                                       | 18 |
| Figura 2: Casos prováveis de Dengue, por semana epidemiológica de início de sintomas, 2014, 2015 e 2016. (VIGILÂNCIA, 2016).....                                                                                                                                                                                                                                                                                                                                                                                          | 20 |
| Figura 3: Mosquito <i>Aedes aegypti</i> , vetor transmissor do vírus da Dengue. ....                                                                                                                                                                                                                                                                                                                                                                                                                                      | 22 |
| Figura 4: Organização estrutural de uma partícula de vírus. O ácido nucleico encontra-se no meio da partícula, envolvido pelo capsídeo (proteína). Muitos vírus possuem também um envelope de membrana, que normalmente é coberto por "grampos" proteicos. Adaptado da referência (FARRELL, 2008).....                                                                                                                                                                                                                    | 24 |
| Figura 5: Superfície do vírion (partícula viral completa, ou seja, infecciosa) maduro da Dengue (sorotipo 2), na sua forma nativa, obtida por crio-microscopia eletrônica, PDB: 3J27.(ZHANG et al., 2012).....                                                                                                                                                                                                                                                                                                            | 25 |
| Figura 6: Ilustração da infecção viral. Os vírions aderem à superfície de uma célula hospedeira e, subsequentemente, entram na célula por endocitose mediada pelo receptor (descrito no texto). Adaptado (QI; ZHANG; CHI, 2008b).....                                                                                                                                                                                                                                                                                     | 26 |
| Figura 7: Esquema mostra os sítios na poliproteína do flavivírus que são clivadas por proteases do hospedeiro (setas preenchidas em baixo) e a protease do vírus (complexo NS2B.NS3) (setas abertas acima). O cofator NS2B e o domínio proteolítico da proteína NS3 (NS3pro) estão sombreados. E os *, na parte de baixo, estão indicando os resíduos do sítio ativo do complexo NS2B.NS3pro, H51, D57 e S135. Adaptado da referência (ARAKAKI et al., 2002).....                                                         | 27 |
| Figura 8: Tríade catalítica da quimotripsina (Código PDB: 4CHA). (POLGÁR, 2005) .....                                                                                                                                                                                                                                                                                                                                                                                                                                     | 28 |
| Figura 9: Representação esquemática da região da cavidade oxianion e sua funcionalidade durante a formação do intermediário tetraédrico. Tríade catalítica representada por Asp, His e Ser, onde a serina está formando a ligação acil com o substrato na formação do intermediário tetraédrico (carga formal negativa). Onde X representa átomos de nitrogênio ou de oxigênio.(WAHLGREN et al., 2011a).....                                                                                                              | 31 |
| Figura 10: Esquema para o mecanismo geral aceito das proteases de serina.(BLOW, 1976; HEDSTROM, 2002b) a) Etapa da acilação (acil-enzima) e quebra da ligação peptídica. b) Etapa da quebra da ligação acil – enzima. Na ilustração da cavidade oxianion, o X pode representar átomos de nitrogênio (N) ou de oxigênio (O). ....                                                                                                                                                                                          | 32 |
| Figura 11: Esquema do mecanismo "His Flip". Depois da formação inicial do intermediário tetraédrico, o N <sub>ε1</sub> -H estará formando uma ligação de hidrogênio com o átomo γO da SER195. A rotação do His57 supostamente coloca o N <sub>ε1</sub> -H perto do grupo de saída NH. Embora as ligações de hidrogênio entre N <sub>δ1</sub> -H e ASP102 e, entre C <sub>ε1</sub> e Ser214 sejam rompidas, novas ligações de hidrogênio irão se formar para substituí-las. Adaptada da referência (HEDSTROM, 2002c) ..... | 35 |
| Figura 12: Superposição de estruturas da região sítio ativo da tripsina para ES (em vermelho), TI1 (em verde), TI2 (em azul) e EA1 (em laranja) no modo de palito. O átomo do tipo HG da SER195 é apresentado no modo de esfera. As estruturas enzima-substrato(ES), intermediário tetraédrico 1 (TI1), intermediário tetraédrico 2 (TI2) e acil-enzima (EA1) pode ser melhor vista na Figura 13. Adaptado da referência (ZHOU; ZHANG, 2011b).....                                                                        | 41 |
| Figura 13: Ilustração das estruturas para a fase de acilação da tripsina. ES, complexo enzima-substrato; TI1 e TI2 duas configurações distintas intermediários tetraédricos; EA1, acil-enzima. Adaptado da referência (ZHOU; ZHANG, 2011b) .....                                                                                                                                                                                                                                                                          | 43 |
| Figura 14: Perfil de energia livre para a etapa da reação de acilação da tripsina determinado pela simulação de dinâmica molecular QM/MM por amostragem guarda-chuva (do inglês, umbrella sampling). Adaptado da referência (ZHOU; ZHANG, 2011b) .....                                                                                                                                                                                                                                                                    | 44 |
| Figura 15: Ilustração das estruturas de NS2B-NS3pro na ausência e na presença de um inibidor. (a) DEN NS2B-NS3pro. Cinza, NS3pro; amarelo, NS2B. Não foi observada densidade de elétrons dos resíduos NS2B 77-84. (b) WNV NS2B-NS3pro em complexo com Bz-Nle-Lys-Arg-Arg-H (laranja). Adaptado da referência (ERBEL et al., 2006) .....                                                                                                                                                                                   | 46 |
| Figura 16: Ilustração da estrutura NS2B-NS3 protease da Dengue ligada ao Bz-nKRR-H. A estrutura é mostrada na representação de cartoon, NS3 em verde e NS2B em azul, com o Bz-nKRR-H em representação stick. Adaptada da referência (NOBLE et al., 2012).....                                                                                                                                                                                                                                                             | 47 |
| Figura 17: Os potenciais sítios de interação do substrato com a enzima são designados P (mostrados em vermelho), e os correspondentes locais de ligação à enzima são designados S (subsítio). A ligação peptídica (em vermelho) é o ponto de referência. Adaptado da referência (HEDSTROM, 2002c).....                                                                                                                                                                                                                    | 48 |
| Figura 18: Detalhe das interações entre NS2B-NS3 protease da DEN-3 e a aprotinina (inibidor). Diagrama dos resíduos da aprotinina (mostrado como stick amarelo) na superfície da protease NS2B-NS3 da DEN-3                                                                                                                                                                                                                                                                                                               |    |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| (cinzento). As possíveis ligações de hidrogênio estão indicadas por linhas tracejadas. Outros resíduos que revestem os bolsões são mostrados com stick verdes, com a tríade catalítica com stick azuis. Imagem em estéreo, adaptada da referência (NOBLE et al., 2012).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 49 |
| Figura 19: Ilustração detalhada da estrutura da região catalítica para a protease (NS2B-NS3) da DENV-3 na presença de inibidores, (a) Bz-nKRR-H e (b) aprotinina. (a) representação em stick para a cavidade oxianion (verde), para a tríade catalítica (H51, D75, e S135) (azul) e do inibidor Bz-nKRR-H (rosa). Apenas resíduos na posição P1 e P2 são mostrados por motivos de clareza. A água e o sulfato são mostrados como uma esfera e stick (vermelho), respectivamente. As possíveis ligações de hidrogênio (até 3,2 Å) são indicadas pelas linhas tracejadas. (b) A representação para a tríade e para a cavidade oxianion é a mesma do (a). Os resíduos P1 a P2 da aprotinina são mostrados como stick de cor amarela, e o grupo carbonilo do resíduo P1 ocupa a cavidade oxianion. Adaptada da referência (NOBLE et al., 2012)..... | 50 |
| Figura 20: Condições de contorno periódicas. Quando uma partícula se move para fora da caixa (célula) de simulação, uma partícula se move na imagem para substituí-lo. Ao calcular as interações das partículas dentro do raio de corte, os vizinhos reais e imagem estão incluídos. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 59 |
| Figura 21: Separação da coordenada de reação entre dois estados em janelas distintas (linha entre A e B), aqui representados por dois mínimos da superfície de energia potencial. O sistema é principalmente amostrado na região perpendicular a coordenada de reação de cada janela. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 63 |
| Figura 22: Ilustração da divisão do sistema onde será tratado pelo método híbrido QM/MM. A maior parte será tratada por mecânica molecular e a menor parte calculada por química quântica.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 65 |
| Figura 23: Estrutura da 2FOM (DENV2a protease). Em azul a NS3, e em vermelho a NS2B. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 70 |
| Figura 24: Estrutura da 2FOM (DENV2a) com acréscimo dos resíduos 183-194. Em azul a NS3 (1-150), em vermelho a NS2B (151-204), e em verde os resíduos adicionados à estrutura (183-194).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 71 |
| Figura 25: Modelo para a estrutura ativa da NS2B-NS3 (DENV2f). Em azul a NS3, e em vermelho a NS2B. ...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 74 |
| Figura 26: Inibidor Bz-NLE-ARG-ARG-LYS-H.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 75 |
| Figura 27: DENV2 com inibidor (DENV2f-INIB). Em azul a NS3, em vermelho NS2B, canudo a tríade catalítica (HIS33, ASP57 e SER117), e as bolas representam o inibidor.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 75 |
| Figura 28: Substrato ALA-GLY-ARG-LYS-SER-ILE-NME. NME é um resíduo para completar a valência, -NHCH <sub>3</sub> .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 76 |
| Figura 29: DENV2 com o substrato (DENV2-SUBS1). Em azul a NS3, em vermelho NS2B, canudo a tríade catalítica (HIS33, ASP57 e SER117), e as bolas representam o substrato.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 77 |
| Figura 30: Substrato CH <sub>3</sub> NH-RRPV-COCH <sub>3</sub> , onde os grupos CH <sub>3</sub> NH e COCH <sub>3</sub> são para completar a valência da sequência de aminoácidos do modelo do substrato. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 77 |
| Figura 31: Resíduos que foram parametrizados. a) Bz (benzoil) e b) NLE (norleucina) tem um grupo -CH <sub>3</sub> a mais quando comparado com a ILE.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 78 |
| Figura 32: (a) Ilustração de algumas das distâncias usadas para esse estudo, na qual, d1(H <sub>γ1</sub> - N <sub>ε2</sub> ), d2(C <sub>γ2</sub> - O <sub>γ1</sub> ), d3(N <sub>ε2</sub> - H <sub>γ1</sub> ), d4(C <sub>γ2</sub> - N <sub>γ2</sub> ), d5(O <sub>δ1</sub> - N <sub>δ1</sub> ) e d6(N <sub>δ1</sub> - H <sub>δ1</sub> ). (b) Ilustração da região QM, os átomos de hidrogênio em vermelho são átomos para completar a valência da região QM. ....                                                                                                                                                                                                                                                                                                                                                                                 | 81 |
| Figura 33: Definição dos vizinhos a uma região da proteína. Onde R é o raio que inclui a vizinhança. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 84 |
| Figura 34: Definição para a ligação de hidrogênio. Onde r é a distância entre doador e aceitador de ligação de hidrogênio, e α é o ângulo.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 85 |
| Figura 35: Conformação da NS2B-NS3 obtida de uma simulação de dinâmica molecular num ensemble NPT. As moléculas de água da caixa de simulação foram omitidas para melhor visualização. Em Azul NS3, vermelho a NS2B, e representado em canudo estão os resíduos do sítio ativo (Asp, His e Ser). ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 87 |
| Figura 36: RMSD para DENV2a (preto), assim como, o RMSD só para a NS3 (azul) e para o cofator NS2B (vermelho).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 88 |
| Figura 37: Superposição para as 700 conformações extraídas a intervalos regulares da simulação de dinâmica molecular da DENV2a. a) superposição da NS3; b) superposição da NS2B. Em Azul tem-se a NS3, e em vermelho a NS2B. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 89 |
| Figura 38: RMSD para NS2B completo (preto), e para algumas regiões do cofator: resíduos 151 a 156 em azul; 196 a 204 em vermelho, e 182 a 196 em verde. Em todos os casos, a referência foi a estrutura inicial da simulação.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 90 |
| Figura 39: Ilustração da DENV2a mostrando os resíduos da NS2B analisados por RMSD na Figura 38. Em cinza está a NS3, já a NS2B tem seus resíduos divididos em: azul os resíduos 151 a 156, vermelho claro os resíduos 157 a 181, resíduos 182 a 196 em verde, resíduos 196 a 204 em vermelho.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 91 |
| Figura 40: RMSF para os resíduos da DENV2a. Onde os resíduos 1-152 são da NS3 e os resíduos 153-192 da NS2B.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 92 |
| Figura 41: Uma conformação da DENV2 extraída de uma simulação de dinâmica molecular num ensemble NPT. As moléculas de água da caixa de simulação não estão representadas na figura para melhor                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| visualização. Em Azul NS3, vermelho a NS2B, e representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117).....                                                                                                                                                                                                                                                                                                                                                                                                                    | 93  |
| Figura 42: RMSD para NS2B-NS3 do modelo DENV2f completa (preto), apenas NS3 (resíduos 1 a 152) em azul e apenas NS2B (resíduos 153 a 192) em vermelho.....                                                                                                                                                                                                                                                                                                                                                                                            | 94  |
| Figura 43: Superposição de 700 conformações extraídas em intervalos regulares da simulação de dinâmica molecular NPT da DENV2f. a) superposição da NS3; b) superposição da NS2B. Em Azul tem-se a NS3, e em vermelho a NS2B. ....                                                                                                                                                                                                                                                                                                                     | 94  |
| Figura 44: RMSF para os resíduos da DENV2f. Os resíduos HIS33, ASP57 e SER117, do sítio ativo, estão indicados. Onde os resíduos 1-152 são da NS3 e os resíduos 153-192 da NS2B. ....                                                                                                                                                                                                                                                                                                                                                                 | 95  |
| Figura 45: Resíduos da tríade catalítica (HIS33, ASP57 e SER117) na DENV2. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 96  |
| Figura 46: Distância entre átomos do sítio ativo, conforme ilustrado na Figura 45. Preto: r1 (SER117 HG – NE2 HIS33); azul: r2 (HIS33 HD1 – OD1 ASP57); Vermelho: r3 (HIS33 HD1 – OD2 ASP57).....                                                                                                                                                                                                                                                                                                                                                     | 96  |
| Figura 47: Resíduos encontrados dentro de um raio de 5.0 Å dos resíduos 153-156. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada. ....                                                                                                                                                                                                                                                                                                                             | 97  |
| Figura 48: Resíduos que estão num raio de 5.0 Å dos resíduos 182-192. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada. ....                                                                                                                                                                                                                                                                                                                                        | 98  |
| Figura 49: Resíduos que estão num raio de 5.0 Å dos resíduos 33 57 e 117. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada. ....                                                                                                                                                                                                                                                                                                                                    | 99  |
| Figura 50: Uma conformação do sistema DENV2f-INIB extraída da simulação de dinâmica molecular com ensemble NPT. As moléculas de água da caixa de simulação não estão representadas na figura para facilitar a visualização. Em Azul NS3, vermelho a NS2B, representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117), e a representação em bola são do inibidor sem os hidrogênios para a melhor visualização. ....                                                                                                              | 102 |
| Figura 51: RMSD para o sistema DENV2f-INIB, tendo como referência a estrutura inicial para toda as etapas de simulação.....                                                                                                                                                                                                                                                                                                                                                                                                                           | 103 |
| Figura 52: RMSF para os resíduos da DENV2f-INIB, região delimitada com seta azul são os resíduos da NS3, a seta em vermelho são os resíduos da NS2B, já a seta em verde delimita os resíduos do inibidor. Onde os resíduos 1-152 são da NS3 e os resíduos 153-192 da NS2B.....                                                                                                                                                                                                                                                                        | 104 |
| Figura 53: Superposição para as 700 conformações extraídas em intervalos regulares dos últimos 150 ns da simulação de dinâmica molecular NPT do sistema DENV2f-INIB. a) superposição da NS3; b) superposição da NS2B. A região da NS3 (azul), em vermelho a NS2B e representado com bolas o inibidor. ....                                                                                                                                                                                                                                            | 105 |
| Figura 54: Resíduos que estão num raio de 5.0 Å dos resíduos do inibidor. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.....                                                                                                                                                                                                                                                                                                                                     | 106 |
| Figura 55: Ilustração dos resíduos que estão num raio de 5.0 Å dos resíduos do inibidor. a) os resíduos estão representados por uma superfície, com o inibidor sobre essa superfície. b) os resíduos estão representados em verde na imagem, e a numeração se refere à região que está representada no gráfico da Figura 54. Representado em azul a NS3, em vermelho a NS2B, em bolas o inibidor, e em verde os resíduos que estão num raio de 5.0 Å do inibidor. No inibidor não há a representação dos hidrogênios para a melhor visualização. .... | 107 |
| Figura 56: Ilustração do inibidor com a cadeia lateral dos resíduos do sítio ativo (HIS33, ASP57 e SER117). O inibidor só está representado nessa imagem pelos resíduos 195, 196 e 197, que são os resíduos da tabela 3. Os postos tracejados se referem a ligações de hidrogênio entre resíduos do inibidor e os indicados na figura da proteína. Os átomos de hidrogênio ligados a átomos de carbono foram retirados para melhorar a visualização. ....                                                                                             | 108 |
| Figura 57: Ilustração do inibidor com os resíduos ASP111, GLY15, GLY135 e TYR143, todos estão indicados na figura. O inibidor só estar representado nessa imagem pelos resíduos 195, 196 e 197, que são os resíduos da Tabela 2. Representações em azul são os hidrogênios, em amarelo os carbonos, em rosa os nitrogênios, e em vermelho, os oxigênios. Os postos tracejados se referem a ligações de hidrogênio entre resíduos do inibidor e os indicados na figura da proteína. ....                                                               | 109 |
| Figura 58: Distâncias de ligação de hidrogênio versus os frames da simulação. As distâncias r1 (SER117-HIS33); r2 (ASP57-HIS33); r3 (197-SER117); r4 (ARG195-GLY135); r5 (ARG195-GLY135); r6 (ARG197-GLY115); r7 (ASP57-LYS196) e r8 (ARG195-TYR:143). Esses frames são referentes aos 150 ns finais da dinâmica molecular.....                                                                                                                                                                                                                       | 110 |
| Figura 59: Número de moléculas de H2O que estão a uma distância de 5.0 Å dos resíduos do sítio ativo. Em preto são as águas que estão próximas a tríade do sítio ativo, em azul ao resíduo HIS33, em vermelho ao resíduo ASP57 e verde ao resíduo SER117. Esses frames são referentes aos 150 ns finais da dinâmica molecular .....                                                                                                                                                                                                                   | 111 |
| Figura 60: Superposição da última estrutura da simulação de dinâmica molecular da DENV2 com a estrutura cristalográfica da NS2B-NS3 (PDB ID: 3U1I). Em azul, a NS3 da (PDB ID: 3U1I); azul claro a NS3 da DENV2;                                                                                                                                                                                                                                                                                                                                      |     |



|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| em vermelho a NS2B da (PDB ID: 3U11); em rosa a NS2B da DENV2; verde os resíduos do sítio ativo para as duas estruturas.....                                                                                                                                                                                                                                                                                                                                                                                                                                     | 112 |
| Figura 61: Uma conformação da DENV2f-SUBS retirada da simulação de dinâmica molecular num ensemble NPT. As moléculas de água da caixa de simulação não estão representadas na figura para melhor visualização. Em Azul NS3, vermelho a NS2B, representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117), e a representação em bola são do inibidor sem os hidrogênios para a melhor visualização.                                                                                                                                           | 113 |
| Figura 62: RMSD para DENV2f (preto), assim como, o RMSD só para a NS3 resíduos 1 a 152 em azul, NS2B resíduos 153 a 192 em vermelho e para o substrato em verde. ....                                                                                                                                                                                                                                                                                                                                                                                            | 114 |
| Figura 63: RMSF para os resíduos da DENV2f-SUBS1, região delimitada com seta azul são os resíduos da NS3, a seta em vermelho são os resíduos da NS2B, já a seta em verde delimita os resíduos do substrato. A numeração é para os resíduos que apresentam maior mobilidade. ....                                                                                                                                                                                                                                                                                 | 115 |
| Figura 64: Ilustração da superposição para as 700 conformações retiradas da simulação de dinâmica molecular NPT da DENV2f-SUBS. a) superposição da NS3; b) superposição da NS2B. A região da NS3 (azul), em vermelho a NS2B e representado com bolas o substrato.....                                                                                                                                                                                                                                                                                            | 116 |
| Figura 65: Ilustração da superposição das estruturas da DENV2f-INIB e DENV2f-SUBS, onde o inibidor estar representado em lilás, e o substrato em verde. NS3 em azul escuro parte da DENV2f-INIB; NS3 em azul claro parte da DENV2f-SUBS; NS2B em vermelho da estrutura DENV2f-INIB; NS2B em laranja da estrutura DENV2f-SUBS. ....                                                                                                                                                                                                                               | 117 |
| Figura 66: Ilustração para algumas regiões da proteína. Em amarelo os resíduos 10 a 14; azul os resíduos 112 a 116; laranja os resíduos 138 a 141; rosa resíduos 153 a 158 e rosa escuro os resíduos 182 a 192. ....                                                                                                                                                                                                                                                                                                                                             | 118 |
| Figura 67: RMSD para algumas regiões da proteína (Figura 66). RMSD: amarelo para os resíduos 10 a 14; azul os resíduos 112 a 116; laranja os resíduos 138 a 141; rosa resíduos 153 a 158 e vermelho os resíduos 182 a 192. ....                                                                                                                                                                                                                                                                                                                                  | 119 |
| Figura 68: Distância entre os resíduos ILE12 e SER140. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 120 |
| Figura 69: Ilustração de algumas estruturas da DENV2f-SUBS em que as regiões dos resíduos 112-116 e os 138-141 estão fazendo um movimento de se afastar e se aproximar uns dos outros, relativamente. Distância relativa entre as regiões amarela e laranja a) 45,3 Å; b) 35,5 Å; c) 44,5 Å e d) 34, 1 Å. Em azul a NS3; vermelho a NS2B; verde o substrato e amarelo os resíduos 112-116 e, laranja os resíduos 138-141. ....                                                                                                                                   | 121 |
| Figura 70: Resíduos que estão num raio de 5.0 Å dos resíduos do substrato. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada. ....                                                                                                                                                                                                                                                                                                                                              | 122 |
| Figura 71: Ilustração dos resíduos que estão num raio de 5.0 Å dos resíduos do substrato. a) os resíduos estão representados por uma superfície, e o substrato estar sobre essa superfície. ; b) os resíduos estão representados em verde na imagem, e a numeração se refere à região que está representada no gráfico da Figura 70. Representado em azul a NS3, em vermelho a NS2B, em bolas o substrato, e em verde os resíduos que estão num raio de 5.0 Å do substrato. No substrato não há a representação dos hidrogênios para a melhor visualização. .... | 123 |
| Figura 72: Distâncias de ligação de hidrogênio versus os frames da simulação. As distâncias r1 (SER117-HIS33); r2 (ASP57-HIS33); r3 (GLY133-LYS196); r4 (ARG195-GLY115); r5 (PHE112-ARG195); r6 (ASP11-ARG195); r7 (AILE198-TYR143). ....                                                                                                                                                                                                                                                                                                                        | 124 |
| Figura 73: Número de moléculas de H <sub>2</sub> O que estão a uma distância de 5.0 Å dos resíduos do sítio ativo. Em preto são as águas que estão próximas a tríade do sítio ativo, em azul ao resíduo HIS33, em vermelho ao resíduo ASP57 e verde ao resíduo SER117. ....                                                                                                                                                                                                                                                                                      | 126 |
| Figura 74: Uma conformação da DENV2f-SUBS2 retirada da simulação de dinâmica molecular numa ensemble NPT. As moléculas de água da caixa de simulação não estão representadas na figura para melhor visualização. Em Azul NS3, vermelho a NS2B, representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117), e a representação em bola são do inibidor sem os hidrogênios para a melhor visualização.                                                                                                                                         | 128 |
| Figura 75: RMSD para DENV2f (preto), assim como, o RMSD só para a NS3 resíduos 1 a 152 em azul, NS2B resíduos 153 a 192 em vermelho e para o substrato (2) em verde.....                                                                                                                                                                                                                                                                                                                                                                                         | 129 |
| Figura 76: RMSF para os resíduos da DENV2f-SUBS, região delimitada com seta azul são os resíduos da NS3, a seta em vermelho são os resíduos da NS2B, já a seta em verde delimita os resíduos do substrato. A numeração é para os resíduos que apresentam maior mobilidade. ....                                                                                                                                                                                                                                                                                  | 130 |
| Figura 77: Ilustração da superposição para as 750 conformações retiradas da simulação de dinâmica molecular NPT da DENV2f-SUBS2. a) superposição da NS3; b) superposição da NS2B. A região da NS3 (azul), em vermelho a NS2B. ....                                                                                                                                                                                                                                                                                                                               | 130 |
| Figura 78: Gráfico dos resíduos que estão num raio de 5.0 Å dos resíduos do substrato (2). Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.....                                                                                                                                                                                                                                                                                                                               | 131 |

|                                                                                                                                                                                                                                                                                                                                                                                                |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figura 79: Gráfico de distâncias e ilustração das mesmas na imagem ao lado. A distância $d1$ é definida entre o hidrogênio $H_{Y1}$ da serina e o nitrogênio $N_{E2}$ da histidina, e $d2$ é a distância entre o oxigênio $O_{Y1}$ da serina e o carbono carboxílico $C_{Y2}$ do substrato-2. ....                                                                                             | 132 |
| Figura 80: Perfil de energia livre bidimensional (ou PMF) QM (PM3PDDG) / MM para a formação do intermediário tetraédrico (TI) pela reação da NS2B-NS3 protease com o substrato (VPRR). O zero da energia corresponde à formação do complexo enzima-substrato (ES). A linha tracejada representa o caminho de menor energia para a formação do intermediário tetraédrico. ....                  | 133 |
| Figura 81: Ilustração da diferença de energia livre dos pontos estacionários em relação ao complexo enzima-substrato (ES) que representa o caminho de menor energia observado na Figura 80. Deste modo, apresentando o patamar de energia para cada ponto estacionário e, uma melhor visualização das barreiras de energia entre os estados. ....                                              | 134 |
| Figura 82: Cargas de Mulliken para alguns dos átomos da região de reação no processo de formação do intermediário tetraédrico. As barras coloridas são os valores médios para cada nos pontos estacionário. O desvio padrão é mostrado pela barra em preto em cada uma das cargas. Isso corresponde a 50 ps em cada um dos pontos estacionários. ....                                          | 136 |
| Figura 83: Número de moléculas de água em percentagem dos frames de cada ponto estacionário. Cada uma das cores representa um número específico de moléculas de água. Esse dado foi obtido em 50 ps de cada um dos pontos estacionários. ....                                                                                                                                                  | 138 |
| Figura 84: a) Perfil de energia livre para a reação de quebra da ligação peptídica. b) Esquema da diferença de energia livre dos pontos estacionários em relação ao produto AE. Ambos correspondendo a coordenada de reação unidimensional para a transferência do hidrogênio ( $H_{Y1}$ ) da histidina ( $N_{E2}$ ) para o nitrogênio ( $N_{Y2}$ ) do grupo de saída do substrato (ARG). .... | 140 |
| Figura 85: Cargas de Mulliken para alguns dos átomos da região de reação durante o processo de quebra da ligação peptídica. As barras coloridas são os valores médios para cada nos pontos estacionário. O desvio padrão é mostrado pela barra em preto em cada uma das cargas. Isso corresponde a 30 ps em cada um dos pontos estacionários. ....                                             | 141 |
| Figura 86: Número de moléculas de água em percentagem dos frames de cada ponto estacionário. Cada uma das cores representa um número específico de moléculas de água. Isso corresponde a 30 ps em cada um dos pontos estacionários. ....                                                                                                                                                       | 142 |

## LISTA DE TABELAS

---

|                                                                                                                                                                                                                                                                                                                                                   |            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <i>Tabela 1: Dados das ligações de hidrogênio entre a NS3 e a NS2B, onde os resíduos da NS3 estão como aceitadores de ligação de hidrogênio, e os resíduos da NS2B estão como doadores de ligação de hidrogênio.....</i>                                                                                                                          | <i>100</i> |
| <i>Tabela 2: Ligações de hidrogênio entre a DENV2 e o inibidor. Resíduos do inibidor estão destacados em negrito.....</i>                                                                                                                                                                                                                         | <i>107</i> |
| <i>Tabela 3: Dados das ligações de hidrogênio entre a DENV2f e o substrato-1. Resíduos do substrato estão destacados em negrito. ....</i>                                                                                                                                                                                                         | <i>124</i> |
| <i>Tabela 4: Dados das ligações de hidrogênio entre a DENV2f e o substrato (2). Resíduos do substrato estão destacados em negrito. ....</i>                                                                                                                                                                                                       | <i>131</i> |
| <i>Tabela 5: Valores médios e desvios padrão das distâncias em cada um dos pontos estacionários, na qual estão identificadas na Figura 80. Médias obtidas sobre 50 ps de simulação em cada um dos pontos estacionários. Valores em Å.....</i>                                                                                                     | <i>135</i> |
| <i>Tabela 6: Valores médios e desvios padrão das distâncias em cada um dos pontos estacionários, na qual estão identificadas na. Médias obtidas sobre 30 ps de simulação em cada um dos pontos estacionários. Valores em Å.....</i>                                                                                                               | <i>139</i> |
| <i>Tabela 7: Valores médios e desvios padrão dos ângulos (seis ângulos que definem uma estrutura tetraédrica) em cada um dos pontos estacionários, na qual estão identificadas na Figura 84. Médias obtidas sobre 30 ps de simulação em cada um dos pontos estacionários. O carbono alpha C<sub>γ3</sub> pertence ao resíduo da prolina. ....</i> | <i>140</i> |

## SUMÁRIO

---

|                                                     |    |
|-----------------------------------------------------|----|
| 1. INTRODUÇÃO .....                                 | 18 |
| 1.1 DENGUE .....                                    | 18 |
| 1.1.1 Epidemiologia .....                           | 18 |
| 1.1.2 Histórico .....                               | 20 |
| 1.1.3 Vetor .....                                   | 21 |
| 1.1.4 Sintomas .....                                | 22 |
| 1.1.5 Combate .....                                 | 23 |
| 1.2 VÍRUS .....                                     | 24 |
| 1.2.1 Vírus da Dengue .....                         | 25 |
| 1.2.2 Replicação Viral .....                        | 25 |
| 1.3 PROTEASES DE SERINA .....                       | 27 |
| 1.3.1 Aspectos Gerais .....                         | 27 |
| 1.3.2 Breve histórico .....                         | 28 |
| 1.3.3 Mecanismo de reação .....                     | 29 |
| 1.3.4 Aspectos Teóricos .....                       | 36 |
| 1.4 NS2B-NS3 PROTEASE DO VÍRUS DA DENGUE .....      | 44 |
| 1.4.1 Inibição da NS2B-NS3 do vírus da Dengue ..... | 47 |
| 2. OBJETIVO .....                                   | 52 |
| 2.1 OBJETIVOS GERAIS .....                          | 52 |
| 2.1.1 Objetivos Específicos .....                   | 52 |
| 3. METODOLOGIA .....                                | 53 |
| 3.1 SUPERFÍCIE DE ENERGIA POTENCIAL .....           | 54 |
| 3.2 CAMPOS DE FORÇAS .....                          | 54 |
| 3.3 DINÂMICA MOLECULAR .....                        | 56 |

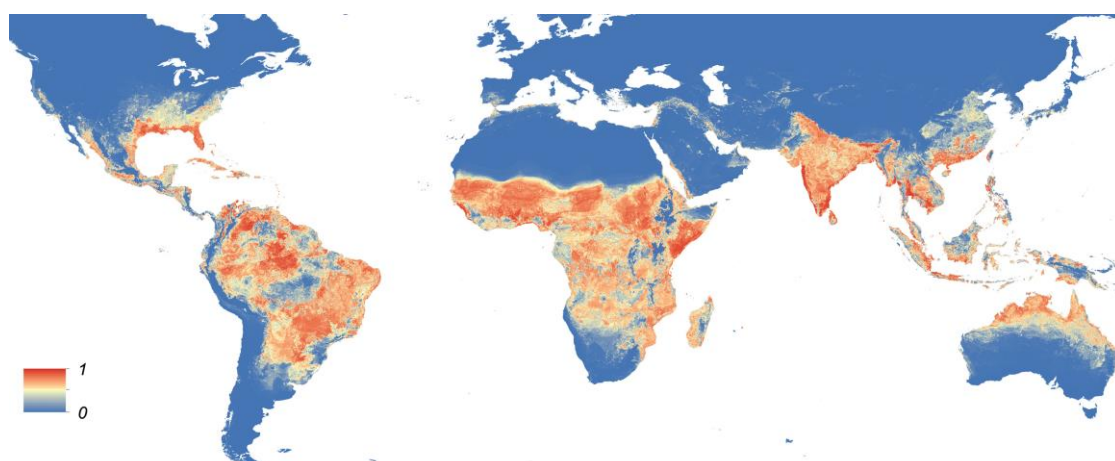
|                                                                                                 |     |
|-------------------------------------------------------------------------------------------------|-----|
| <b>3.3.1 Formulação</b>                                                                         | 56  |
| <b>3.4 MECÂNICA ESTATÍSTICA</b>                                                                 | 59  |
| <b>3.5 HOMOLOGIA</b>                                                                            | 61  |
| <b>3.6 ENERGIA LIVRE</b>                                                                        | 61  |
| <b>3.6.1 Amostragem guarda-chuva (Umbrella Sampling)</b>                                        | 62  |
| <b>3.7 MÉTODO HÍBRIDO QM/MM</b>                                                                 | 64  |
| <b>3.7.1 MÉTODO SEMI-EMPÍRICO</b>                                                               | 67  |
| <b>4. PROCEDIMENTO COMPUTACIONAL</b>                                                            | 70  |
| <b>4.1 ESCOLHA DA ESTRUTURA INICIAL</b>                                                         | 70  |
| <b>4.2 CRIAÇÃO DO MODELO ATIVO DA NS2B-NS3 (DENV2f)</b>                                         | 73  |
| <b>4.3 DENV2f-INIB</b>                                                                          | 74  |
| <b>4.4 DENV2f-SUBS</b>                                                                          | 76  |
| <b>4.5 PARAMETRIZAÇÃO DO INIBIDOR</b>                                                           | 78  |
| <b>4.6 MINIMIZAÇÃO DA ESTRUTURA</b>                                                             | 78  |
| <b>4.7 DINÂMICA MOLECULAR</b>                                                                   | 79  |
| <b>4.8 CÁLCULOS HÍBRIDOS (QM/MM)</b>                                                            | 80  |
| <b>4.9 ANÁLISE DOS RESULTADOS</b>                                                               | 83  |
| <b>5. RESULTADOS e DISCUSSÃO</b>                                                                | 86  |
| <b>5.1 SISTEMA DENV2<math>\alpha</math>: NS2B-NS3 (2FOM)</b>                                    | 86  |
| <b>5.2 MODELO DA ESTRUTURA FECHADA (ATIVA) DA NS2B-NS3 PROTEASE DO VÍRUS DA DENGUE (DENV2f)</b> | 92  |
| <b>5.3 MODELO DENV2f COM INIBIDOR (DENV2f-INIB)</b>                                             | 101 |
| <b>5.4 MODELO DENV2f COM O SUBSTRATO 1 (DENV2f-SUBS1)</b>                                       | 112 |
| <b>5.5 MODELO DENV2f COM SUBSTRATO 2 (DENV2f-SUBS2)</b>                                         | 126 |
| <b>5.6 MECANISMO DE REAÇÃO</b>                                                                  | 133 |
| <b>5.6.1 Etapa de Acilação</b>                                                                  | 133 |

|                                                         |            |
|---------------------------------------------------------|------------|
| <b>5.6.2 Etapa de Quebra da Ligação Peptídica .....</b> | <b>138</b> |
| <b>6. CONSIDERAÇÕES FINAIS .....</b>                    | <b>144</b> |
| <b>7. PERSPECTIVAS .....</b>                            | <b>146</b> |
| <b>REFERÊNCIAS .....</b>                                | <b>147</b> |
| <b>APÊNDICE A: Protocolos de Simulação .....</b>        | <b>163</b> |
| <b>APÊNDICE B: Artigos Publicados .....</b>             | <b>167</b> |
| <b>APÊNDICE C: Apresentações em Eventos .....</b>       | <b>169</b> |

## 1. INTRODUÇÃO

### 1.1 DENGUE

A Dengue é uma doença viral infecciosa sistêmica que é transmitida entre os seres humanos por mosquito *Aedes (aegypti e albopictus)* (BHATT et al., 2013), sendo uma doença endêmica na África, Ásia, Caribe e América Latina. O número de casos de Dengue tem crescido em todo o mundo nas últimas décadas. Os números reais de casos de Dengue são subnotificados e muitos casos são classificados erroneamente. Uma estimativa recente indica 390 milhões de infecções de Dengue por ano (95% intervalo de credibilidade 284-528 milhões), dos quais 96 milhões manifestam clinicamente os sintomas (com qualquer gravidade da doença). (BHATT et al., 2013) Outro estudo, da prevalência da Dengue, estima que 3,9 bilhão de pessoas, em 128 países, estão em risco de infecção por DENV (SHEPARD et al., 2016). Como visto na Figura 1, os países do hemisfério sul são os mais atingidos pelo vírus da Dengue, devido as condições climáticas da região que favorece a disseminação do mosquito *Aedes aegypti*, vetor do vírus (KRAEMER et al., 2015). Notícias recentes mostram que o vírus da Dengue atingiu regiões onde, até então, não havia sido registrado casos de pessoas infectadas, como por exemplo a região metropolitana de Paris, França (LA RUCHE et al., 2010); Flórida, Estados Unidos (FRANCO et al., 2010) e em Nova Caledônia, Oceania (KLINE et al., 2013).



**Figura 1:** Mapa da distribuição global previsto para o mosquito *Ae. aegypti*. O mapa mostra a probabilidade de ocorrência (a partir de 0 (azul) para 1 (vermelho)) com uma resolução espacial de 5 km × 5 km (KRAEMER et al., 2015). ..

#### 1.1.1 Epidemiologia

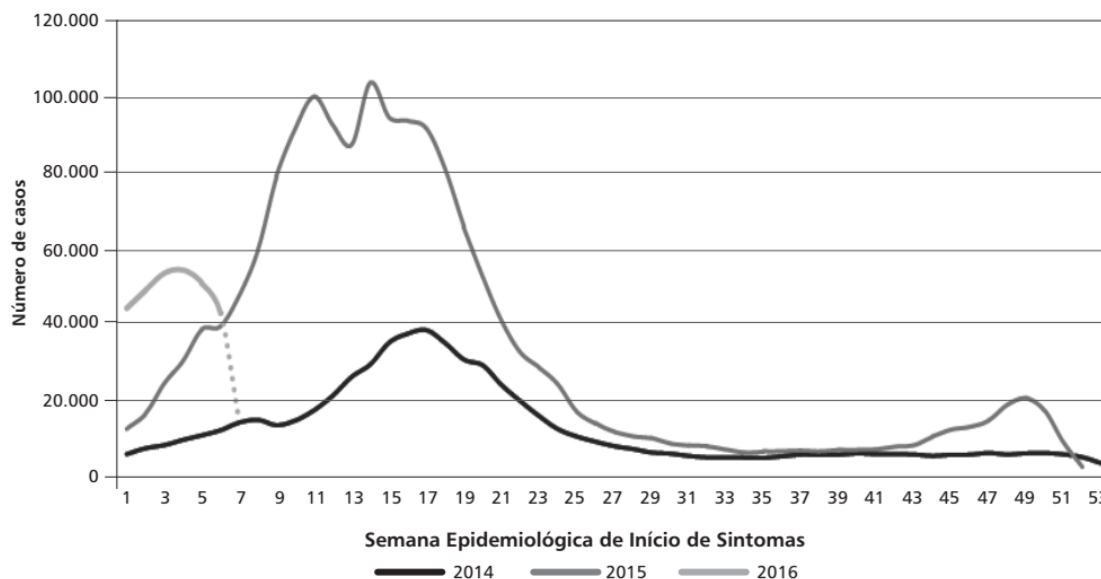
O continente asiático contribuiu com 70% dos casos da doença, só a Índia com 34%, já a América com 14 % dos infectados do mundo, onde mais da metade dos casos ocorreram

em Brasil e México. Os países da África apresentam uma taxa de 16 % dos infectados. Os países da Oceania contribuem com menos de 0,2% das infecções aparentes(BHATT et al., 2013). A taxa de mortalidade é de aproximadamente 5% dos infectados em todo o mundo (WORLD HEALTH ORGANIZATION, 2009a). O número de casos nas Américas, Sudeste Asiático e na região Ocidental do Pacífico excedeu 1,2 milhões em 2008 e mais de 3 milhões em 2013 (com base em dados oficiais apresentados pelos Estados-Membros). (“WHO | Dengue and severe dengue”, 2016) Recentemente, o número de casos relatados tem continuado a aumentar. Em 2015, 2,35 milhões de casos de dengue foram notificados somente no continente Americano, dos quais 10.200 casos foram diagnosticados como dengue grave causando 1.181 mortes. O ano de 2015 foi caracterizado por grandes surtos de dengue em todo o mundo, com o relato das Filipinas mais de 169.000 casos e na Malásia superior a 111.000 casos suspeitos de dengue, representando um aumento de 59,5% e de 16% no número de casos ao ano anterior, respectivamente (“WHO | Dengue and severe dengue”, 2016). Só o Brasil relatou mais de 1,5 milhões de casos em 2015, cerca de três vezes maior do que em 2014. Também em 2015, Délhi, Índia, registrou sua pior epidemia desde 2006, com mais de 15.000 casos (“WHO | Dengue and severe dengue”, 2016).

No século 21, o Brasil se tornou o país do mundo com mais casos notificados de Dengue, ocupando o primeiro lugar no *ranking* internacional para o total de casos da doença, com mais de três milhões de casos relatados de 2000 a 2005 (COSTA, 2009). Isso representou 78% de todos os casos notificados nas Américas e 61% de todos os casos relatados à Organização Mundial da Saúde (OMS).

O Ministério da Saúde já registrou até a Semana Epidemiológica (SE), distribuição temporal dos casos que permite a comparabilidade dos dados, (3/1/2015 a 6/2/2016) (Figura 2), em 2016, mais de 170.103 casos prováveis de Dengue no país (Figura 2). A região Sudeste registrou o maior número de casos (96.664), seguida das regiões Nordeste (25.636), Centro-Oeste (25.246), Sul (13.522) e Norte (9.035)(VIGILÂNCIA, 2016). Neste período, foram descartados 19.249 casos suspeitos de dengue (VIGILÂNCIA, 2016), pois a Dengue, Zika e Chikungunha apresentam sintomas similares.





**Figura 2:** Casos prováveis de Dengue, por semana epidemiológica de início de sintomas, 2014, 2015 e 2016. (VIGILÂNCIA, 2016) Semana epidemiológica é a distribuição temporal dos casos que permite a comparabilidade dos dados, padronização internacional.

No estado de Pernambuco, quando comparamos os casos prováveis de Dengue até a Semana Epidemiológica 7 entre os anos de 2015 (4.798 casos) e 2016 (17.703), ocorreu um aumento de 28,12 % no número de casos.(VIGILÂNCIA, 2016)

### 1.1.2 Histórico

Embora a Dengue só tenha atingindo níveis dramáticos de incidência que trouxeram maiores consequências após a cessação da Segunda Guerra Mundial, a doença tem uma longa história de interação com o homem. É conhecida a partir do registro histórico na China uma doença com descrição clínica semelhante à da Dengue, já no século III durante a dinastia Chin. Relatos semelhantes foram descritos durante o século VII e X do calendário cristão (Dinastia Tang e Dinastia Sung, respectivamente). Estes relatórios descreveram uma doença chamada "veneno da água", devido à sua associação com a água e insetos voadores, e cuja, descrição clínica incluía febre, erupção cutânea, mialgia e manifestações hemorrágicas. Depois de uma longa ausência de registro histórico, o relato de uma doença semelhante apareceu quase sete séculos mais tarde, descrevendo uma doença aguda com convalescença prolongada na região das Antilhas Francesas e no Panamá durante o período de 1635 a 1699. Entretanto, só no século XX a Organização Mundial de Saúde a reconheceu como doença (SCHEINER; LIPSCOMB-T, 1976)(WEAVER; VASILAKIS, 2009).

No Brasil a Dengue tem seu primeiro registro em 1685 na cidade de Recife-PE,(COSTA, 2009) entretanto, foi em 1692 que houve o primeiro surto de Dengue levando a óbito duas mil pessoas na cidade de Salvador-BA.(COSTA, 2009) O vírus possivelmente chegou ao Brasil na época do tráfico de escravos, nos navios negreiros que chegavam da África. Neste período a África já havia registrado casos da doença. O precursor no combate à Dengue (combate ao mosquito transmissor) foi Oswaldo Cruz, Diretor Geral de Saúde Pública em 1903 e, decorrente dessa luta contra a doença, no ano de 1957 a Dengue foi considerada erradicada no Brasil.

Entretanto, reaparecimento do *Aedes aegypti* (mosquito transmissor do vírus) na segunda metade da década de 1970, em um país como o Brasil, com um clima tropical onde as condições ambientais são propícias ao vetor, e durante um período em que a vigilância entomológica e a estrutura para combate ao vetor eram inexistentes, permitiu não só a sua reintrodução, mas também a sua expansão territorial.(COSTA, 2009) Embora a primeira epidemia de Dengue com isolamento viral ocorreu em 1981, em Boa Vista-Roraima, apenas em 1986, quando o sorotipo 1 (DENV-1) foi introduzido no Rio de Janeiro, que a doença se tornou um importante problema de saúde pública. Dias depois, algumas capitais do Nordeste já tinham sido afetadas e o número total de casos acumulados notificados em dois anos chegou a 134.721. A força de transmissão do vírus da Dengue foi tão intensa que, de acordo com as estimativas com base em um inquérito sorológico em 1986 e 1987, no Rio de Janeiro mais de um milhão de pessoas tinham sido infectadas com o vírus DENV-1. Uma vez que nos anos seguintes as condições ambientais permaneceram receptivas a circulação do vírus da Dengue, em 1990 um novo sorotipo, DENV-2, foi detectado também no Rio de Janeiro, com o recrudescimento da Dengue e os primeiros casos de febre hemorrágica da Dengue. Com a introdução do DENV-3 em 2002, houve um aumento dramático em casos notificados de Dengue (794.219) e febre hemorrágica da Dengue (2.702). E em 2010 tivemos a confirmação da circulação do sorotipo 4 do vírus da Dengue no Brasil.(COSTA, 2009)

Portanto, nos dias atuais, no Brasil a Dengue é considerada uma doença endêmica, ou seja, um problema de saúde pública.

### **1.1.3 Vetor**

O vetor transmissor do vírus da Dengue é o *Aedes aegypti* (Figura 3), um mosquito que vive de preferência em regiões tropicais e subtropicais (Figura 1), que apresentam temperaturas elevadas e, chuvas frequentes no período mais quente do ano, como também em

regiões de baixa altitude. Esse mosquito é originário do Egito, na África, e se espalhou pelas regiões tropicais e subtropicais durante o século XVI no período das Grandes Navegações. (HENCHAL; PUTNAK, 1990)



**Figura 3:** Mosquito *Aedes aegypti*, vetor transmissor do vírus da Dengue.

Os mosquitos machos do *Aedes aegypti* não transmitem o vírus da dengue, devido a sua forma de alimentação, pois os mesmos não se alimentam de sangue. Entretanto, os mosquitos fêmeas do *Aedes aegypti* se alimentam de sangue animal, principalmente o sangue humano, assim tornando-se capaz de transmitir o vírus aos humanos. No momento em que se alimentam, a fêmea infectada pelo vírus da Dengue transmite o mesmo para o ser humano, devido ao contato com o sangue, que por sua vez poderá infectar outro mosquito fêmea, e assim, propagando a doença. (WORLD HEALTH ORGANIZATION, 2009a)

#### **1.1.4 Sintomas**

Depois de infectado, o indivíduo pode manifestar alguns sintomas como febre com duração de aproximadamente sete dias, acompanhada por pelo menos dois dos seguintes sintomas: dor de cabeça, dores musculares, dor na região posterior dos olhos, dores nas articulações, vermelhidão no corpo e prostração. Os sintomas da Dengue hemorrágica se assemelham aos da Dengue clássica, entretanto há um agravamento do quadro infeccioso. Existem alguns sinais que alertam para esse agravamento, como dores abdominais fortes e contínuas, sangramento no nariz, boca e gengiva, vômitos persistentes, pele pálida, fria e úmida, sonolência, agitação e confusão mental, dificuldade respiratória e perda de consciência. Esse quadro clínico da Dengue hemorrágica se agrava rapidamente, apresentando sinais de insuficiência circulatória e choque, podendo levar a pessoa à morte em apenas 24 horas. (WORLD HEALTH ORGANIZATION, 2009a)

### 1.1.5 Combate

Atualmente a prevenção primária da Dengue se concentra em combater o vetor transmissor do vírus, o *Aedes aegypti*, através de pulverização de pesticidas e aplicação de larvicidas. Entretanto, esse tratamento também pode atingir outros insetos e animais, como também o mosquito pode adquirir resistência aos compostos utilizados.(WORLD HEALTH ORGANIZATION, 2009a)

A primeira vacina para a Dengue foi desenvolvida em 2015 (DengVax, Sanofi-Pasteur), e já foi licenciada no Brasil, México e Filipinas, apesar de apresentar eficácia de apenas 65,6% na prevenção da doença.(HADINEGORO et al., 2015) Outros esforços existem para desenvolver vacinas mais eficientes, como a que está em desenvolvimento em uma parceria da Fundação Oswaldo Cruz (FIOCRUZ) e Instituto Butantan no Brasil.(KONISHI, 2011) Outra linha de ação em testes é o controle dos vetores através da liberação de espécimes geneticamente modificados para gerar crias que não chegam à fase adulta,(FRANZ et al., 2006) ou a esterilização dos mosquitos por irradiação.(BENEDICT, 2003) Esses procedimentos ainda estão em fase de testes e não tem eficiência emergencial na rápida disseminação global do vírus.(BHATT et al., 2013) No entanto, uma vez contraída a doença, não há um tratamento específico para a Dengue, e as orientações da OMS se limitam ao repouso e tratamento dos sintomas. (WORLD HEALTH ORGANIZATION, 2009b)

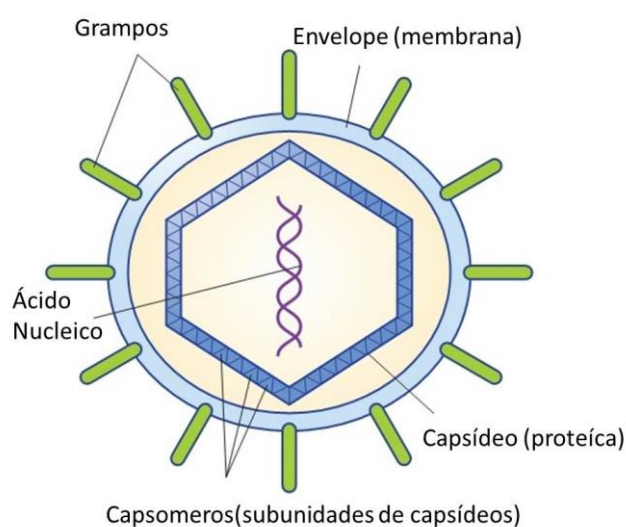
Combater o mosquito transmissor do vírus é de fundamental importância, assim como investir em pesquisa que vise o aprimoramento das técnicas de diagnósticos, tendo em vista a dificuldade na identificação da doença, como também que tipo de sorotipo da Dengue o indivíduo fora infectado. Outro ponto fundamental é a grande dificuldade de conseguir uma vacina eficiente aos quatro sorotipos da Dengue, simultaneamente.

Atualmente, não existe nenhum medicamento no mercado que possa deter a replicação viral, e o indivíduo, depois de infectado pelo vírus, tem um tratamento apenas sintomático: repouso, ingestão de líquidos e paracetamol para febre e dores. Durante esse período, o paciente pode ainda transmitir a Dengue, caso seja picado novamente por um mosquito. Portanto, com o aumento do número de casos de Dengue em todo o mundo, se fazem necessárias pesquisas na área de substâncias que possam deter a replicação viral depois que o indivíduo for infectado. Embora a Dengue atinja milhares de pessoas em todo o mundo, existe muito a ser explorado nessa área, haja vista, muitas questões ainda se encontram sem solução.

## 1.2 VÍRUS

Os vírus não podem se reproduzir de forma independente, pois não produzem proteínas ou energia de forma independente. Pode-se defini-los como uma quantidade relativamente pequena de material genético envolto por um envelope de proteína. Uma vez que o vírus infecta uma célula, este utiliza-se dos mecanismos da própria célula infectada para reproduzir-se. Os vírus têm como material genético apenas um tipo de ácido nucleico, DNA ou RNA. Dependendo do vírus, este ácido nucléico pode ser de cadeia simples ou de cadeia dupla. O RNA ou DNA podem codificar desde apenas quatro proteínas no vírus mais simples (*Brome mosaic virus*, BMV), até de 100 a 200 proteínas em vírus mais complexos, como os Baculovírus, que infectam artrópodes, assim são usados como agentes de controle biológico de insetos na agricultura.(FARRELL, 2008)

O vírus tem uma estrutura de partícula muito pequena e simples, em que seu material genético se encontra no centro da partícula, envolto pelo o capsídeo, que é um revestimento formado por proteínas estruturais; esse arranjo é chamado de nucleocapsídeo (Figura 4). Alguns vírus também apresentam uma membrana envolvendo esse nucleocapsídeo, podendo também possuir algumas proteínas externas como “grampos” para ajudar na ligação com a célula hospedeira.



**Figura 4:** Organização estrutural de uma partícula de vírus. O ácido nucleico encontra-se no meio da partícula, envolvido pelo capsídeo (proteína). Muitos vírus possuem também um envelope de membrana, que normalmente é coberto por “grampos” proteicos. Adaptado da referência (FARRELL, 2008)

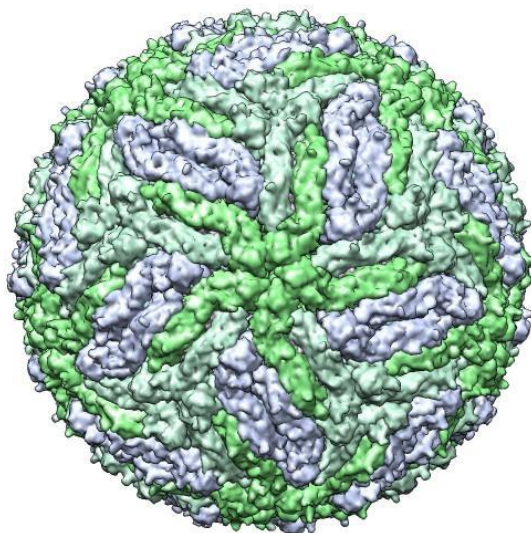
Os vírus não se mantêm vivos por longos períodos fora da célula e, portanto, procuram mecanismos para invadir as células. A indústria farmacêutica trabalha muito em buscar

maneiras de bloquear o acesso dos vírus à célula como uma das maneiras de se evitar a infecção do hospedeiro pelo vírus.(FARRELL, 2008)

### 1.2.1 Vírus da Dengue

Os vírus da Dengue, febre amarela, hepatite C e febre do Nilo Ocidental fazem parte da família *Flaviviridae*, do gênero *Flavivirus*, caracterizados por serem vírus RNA fita simples senso positivo ((+)ssRNA), deste modo, o RNA genômico serve de RNA mensageiro e é traduzido pelos ribossomos da célula hospedeira.(AHLQUIST, 2006) Existem quatro tipos virais, chamados sorotipos, para o vírus da Dengue (Figura 5), e todos são transmitidos pelo mesmo mosquito.(QI; ZHANG; CHI, 2008a) Atualmente, no Brasil há coexistência dos quatro sorotipos. (BARRETO; TEIXEIRA, 2008)

O genoma do vírus da Dengue é traduzido em três proteínas estruturais e sete não estruturais. O primeiro grupo é composto pelo capsídeo (C), a proteína precursora de membrana (prM) e a glicoproteína de envelope (E), essencial para a entrada do vírus na célula. No segundo grupo estão as proteínas envolvidas na replicação viral no interior da célula, denominadas de NS1, NS2A, NS2B, NS3, NS4A, NS4B, e NS5. (MURRELL; WU; BUTLER, 2011)

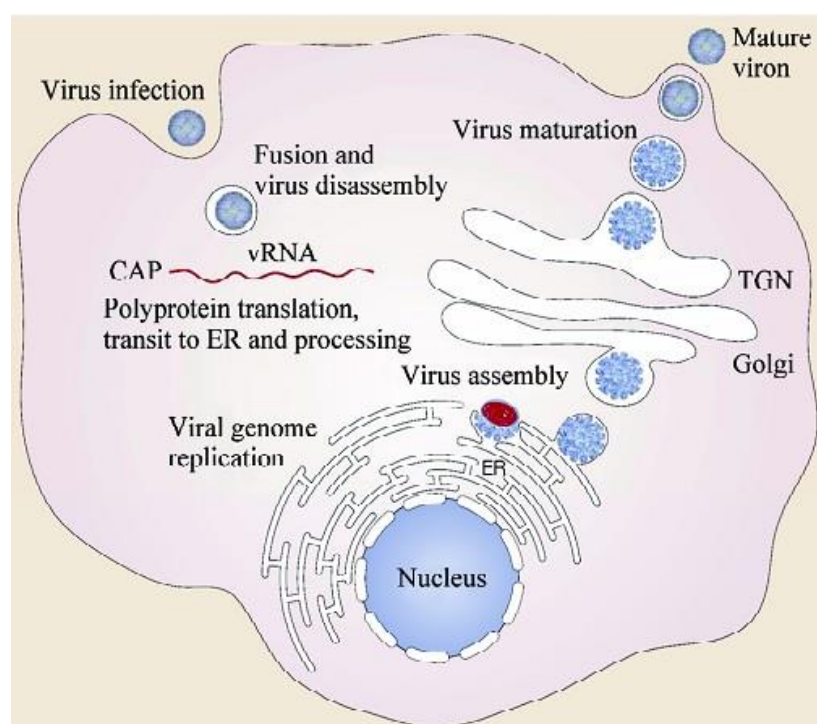


**Figura 5:** Superfície do vírion (partícula viral completa, ou seja, infecciosa) maduro da Dengue (sorotipo 2), na sua forma nativa, obtida por crio-microscopia eletrônica, PDB: 3J27.(ZHANG et al., 2012)

### 1.2.2 Replicação Viral

Os vírions (partícula viral completa, ou seja, infecciosa) aderem à superfície de uma célula hospedeira e, subsequentemente, entram na célula por endocitose mediada pelo

receptor. Após ter ocorrido a fusão da membrana viral e celular, o nucleocapsídeo (NC) é libertado para o citoplasma, que conduz à dissociação da proteína C e do RNA. Uma vez que o material genético é liberado no citoplasma, o RNA de senso positivo ((+)ssRNA) é traduzido em uma única poliproteína que é então clivada por proteases virais e do hospedeiro para a formação de um novo vírion. A replicação do material genético ocorre nas membranas intracelulares, e a montagem e formação das partículas virais imaturas ocorrem na superfície do retículo endoplasmático (RE). As partículas de vírions imaturos são então clivadas pela protease do hospedeiro, resultando em partículas maduras, infecciosas. Os vírions maduros e de partículas subvirais são posteriormente liberados da célula hospedeira por exocitose (Figura 6). (QI; ZHANG; CHI, 2008a)

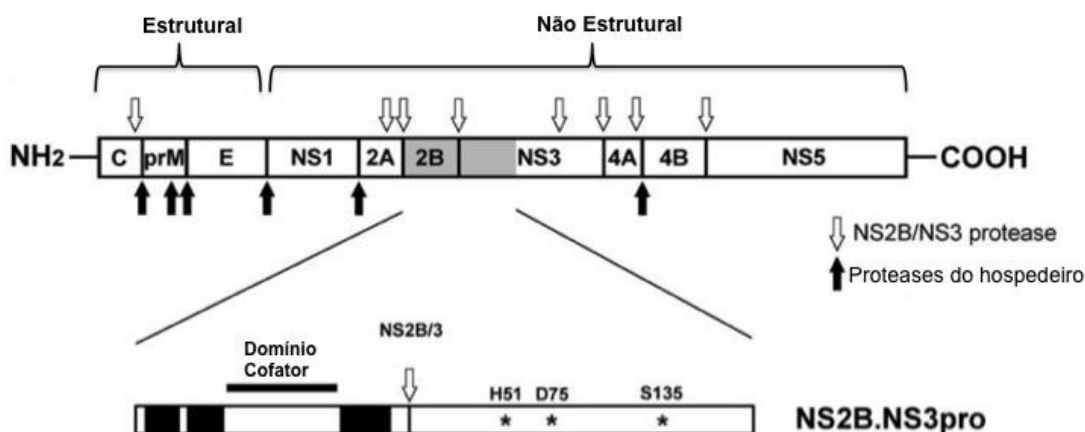


**Figura 6:** Ilustração da infecção viral. Os vírions aderem à superfície de uma célula hospedeira e, subsequentemente, entram na célula por endocitose mediada pelo receptor (descrito no texto). Adaptado (QI; ZHANG; CHI, 2008b)

A peptidase, enzimas que quebram ligações peptídicas entre aminoácidos, da célula do hospedeiro cliva entre as junções C-PrM, PrM-E, E-NS1 e NS4A-NS4B. Já a protease de



serina que é codificada pelo próprio vírus cliva entre as junções NS2A-NS2B, NS2B-NS3, NS3-NS4A, NS4B-NS5 (Figura 7). (NATARAJAN, 2010)(FALGOUT; MARKOFF, 1995)



**Figura 7:** Esquema mostra os sítios na poliproteína do flavivírus que são clivadas por proteases do hospedeiro (setas preenchidas em baixo) e a protease do vírus (complexo NS2B.NS3) (setas abertas acima). O cofator NS2B e o domínio proteolítico da proteína NS3 (NS3pro) estão sombreados. E os \*, na parte de baixo, estão indicando os resíduos do sitio ativo do complexo NS2B.NS3pro, H51, D57 e S135. Adaptado da referência (ARAKAKI et al., 2002)

### 1.3 PROTEASES DE SERINA

#### 1.3.1 Aspectos Gerais

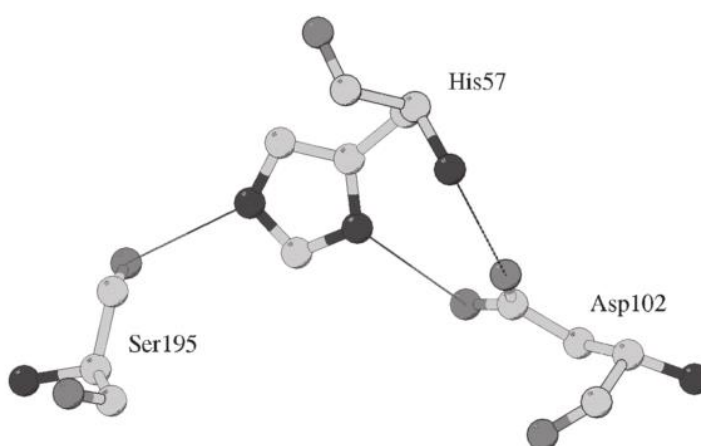
As proteases são enzimas que catalisam a degradação de proteínas por hidrólise de ligações peptídicas e desempenham funções indispensáveis em todas as células vivas, que estão relacionadas a digestão, a coagulação do sangue, a fibrinólise, o desenvolvimento, a fertilização, a apoptose, e a imunidade celular.(TURK, 2006) Elas são associadas com uma série de doenças do organismo humano, por exemplo, doenças cardiovasculares, doença de Alzheimer, alguns tipos de câncer, doenças autoimunes, inflamações, e hipertensão.(COUGHLIN, 2000; HEDSTROM, 2002a; NEURATH, 1984) A compreensão do mecanismo catalítico e da estrutura dessas proteases pode facilitar o design de drogas. (EKICI; PAETZEL; DALBEY, 2008)

Um terço das proteases é classificada como proteases de serina, as proteases mais estudadas dessa classe sendo a quimotripsina, a tripsina, a elastase e a subtilisina. Algumas proteases de serina estão associadas a doenças como febre amarela, vírus da influenza A, vírus da hepatite C, vírus da Dengue e febre do Nilo. As proteases de serina são divididas em duas grandes famílias: a família da quimotripsina (que inclui a tripsina) e a família da



subtilisina, embora existam outras famílias distintas de protease de serina. O termo "família" denota um grupo dessas enzimas que são homólogas por sua sequência de aminoácidos e muito similares em sua geometria tridimensional.

A característica fundamental das proteases de serina é a presença do resíduo serina (SER), muito reativo, em seu sítio ativo. Em geral, as proteases de serina, no caso das famílias da quimotripsina e subtilisina, apresentam uma tríade catalítica, ASP-HIS-SER (Figura 8). Entretanto, há também pequenas diferenças na tríade catalítica em outras famílias de proteases de serina, como SER-HIS-GLU, SER-HIS-HIS e SER-GLU-ASP, assim como díades, SER-HIS ou SER-LYS. (EKICI; PAETZEL; DALBEY, 2008)



**Figura 8:** Tríade catalítica da quimotripsina (Código PDB: 4CHA). (POLGÁR, 2005)

### 1.3.2 Breve histórico

O modo de ação das proteases de serina tem sido objeto de estudo por mais de trinta anos. Os detalhes estruturais dessa classe de proteínas foram observados por diversas técnicas experimentais como difração de Raio-X, difração de nêutrons e RMN de proteínas.

Em 1964, pesquisadores determinaram a sequência de aminoácidos do tripsinogênio (WALSH; NEURATH, 1964) e da quimotripsinogênio (HARTLEY, 1964), precursoras da tripsina e da quimotripsina, respectivamente, revelando que essas proteases eram homólogas entre si. A estrutura tridimensional da quimotripsina foi elucidada em 1967 (MATTHEWS et al., 1967a), já a estrutura da tripsina foi resolvida em 1974. (STROUD; KAY; DICKERSON, 1974)

Em 1969, David Blow et al. realizou um estudo detalhado da estrutura tridimensional da quimotripsina pelo método de difração de Raio-X, mostrando algumas características dessa enzima. (BLOW; BIRKTOFT; HARTLEY, 1969a) Essa análise estrutural revelou alguns

aspectos químicos importantes sobre os resíduos do sítio ativo, em que o ácido aspártico (ASP102) forma ligação de hidrogênio com a histidina (HIS57), que por sua vez forma ligação de hidrogênio com a serina (SER195), todos do sítio ativo. Além disso, Blow propôs um possível mecanismo para esse sítio ativo, propondo um mecanismo de redistribuição de carga, onde duas trocas de prótons ocorreriam (a partir de SER195 para HIS57 e de HIS57 para ASP102). (A numeração dos resíduos do sítio ativo – serina, aspartato e histidina mencionada nesse parágrafo se refere à sequência de aminoácidos da quimotripsina estudada por Blow). Uma de suas conclusões foi de que a polarização do sistema devido à carga negativa presente no ácido aspártico tornaria o oxigênio da serina mais nucleofílico, assim explicaria sua reatividade maior para amidas e ésteres. (BLOW; BIRKTOFT; HARTLEY, 1969a)

Com o tempo alguns estudos mostraram a função do ASP102 no sítio ativo. Um desses estudos mostrou que, ao substituir o resíduo ASP02 (que possui carga negativa) por um grupo ASN102 (com carga neutra), a atividade catalítica dessa enzima diminuiu na ordem de  $10^4$  e a ação dos inibidores em até oito vezes. (SPRANG et al., 1987) Isso evidencia que a carga negativa do ácido aspártico (ASP102) é fundamental para a distribuição de carga do sítio ativo, favorecendo a maior nucleofilicidade da serina (SER195).

Outro estudo interessante foi o de RMN-H em quimotripsina A, que confirmou a formação de ligação de hidrogênio entre os resíduos do sítio ativo. (ROBILLARD; SHULMAN, 1972)

Mais tarde o trabalho de Anthony Kossiakoff et al., realizado com difração de nêutrons em tripsina bovina, mostrou que não há transferência de próton entre os resíduos ASP102 e HIS57. A HIS57 age como base no sítio ativo da proteína, sugerindo que só há uma transferência de próton entre SER195 e HIS57. (KOSSIAKOFF; SPENCER, 1981)

Com essas informações foi possível entender e propor melhores modelos para o mecanismo de ação das proteases de serina.

### **1.3.3 Mecanismo de reação**

#### **1.3.3.1 Aspectos Gerais**

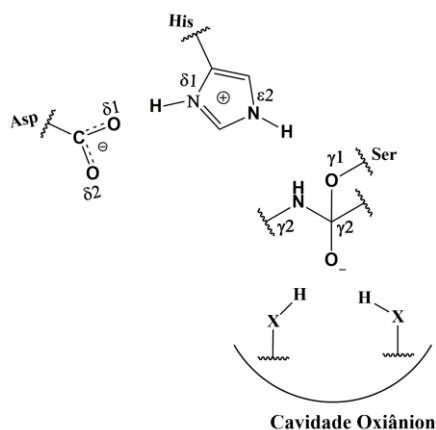
A hidrólise de ligações peptídicas é termodinamicamente favorável. Entretanto, essas reações de hidrólise são extremamente lentas. Tipicamente, na ausência de um catalisador e em pH neutro, a reação tem um tempo de vida médio entre 10 e 1.000 anos. Em processos biológicos, com a presença de proteases, essa reação ocorre em milissegundos. Essa

dificuldade com relação à parte cinética da quebra da ligação peptídica é devida à natureza química dessa ligação, em que a ligação C–N tem um caráter parcial de dupla ligação, não permitindo que a molécula rotacione nesse eixo da ligação e favorecendo uma estrutura plana. Os demais átomos que compõem a região da ligação peptídica, como C, N, O e H formam então uma estrutura plana, resultando na estabilização da ligação peptídica por ressonância. Outra característica importante é que o átomo de carbono do grupo carboxila é menos eletrofílico e menos susceptível ao ataque nucleofílico quando comparado ao mesmo grupo em compostos de ésteres carboxilato, por exemplo. Por conseguinte, para a quebra da ligação peptídica, a enzima funciona como facilitador nesse processo reacional (FARRELL, 2008).

As proteases executam a reação de maneira muito eficiente. As taxas de hidrólise peptídica por proteases de serina são na ordem de  $10^{10}$  vezes maiores do que as reações não catalisadas. Este mecanismo catalítico não é exclusivo da hidrólise de peptídeos: proteases de serina também catalisam a hidrólise de outros compostos como acilas, amidas, ésteres e tioésteres.

#### 1.3.3.2 Cavity Oxyanion

Durante esse período de estudo sobre protease de serina, uma região de interação importante da proteína foi proposta devido a sua recorrência em diversas proteínas da classe, chamada “cavidade oxiânion” (do inglês, *oxyanion hole*). Normalmente, a cavidade oxiânion é formada por grupos amida da cadeia principal (*backbone*) da enzima, por exemplo, GLY193 e SER195, bem como moléculas de água, no caso da quimotripsina. A cavidade oxiânion é formada pelos grupamentos NH da cadeia principal dos resíduos GLY193 e SER195. Estes átomos formam uma região de carga positiva que ativa o grupo carbonila da ligação peptídica clivada e estabiliza o íon oxiânion carregado negativamente do intermediário tetraédrico, (próxima seção), Figura 9. (HEDSTROM, 2002b; MATTHEWS et al., 1967b; WARSHEL; RUSSELL, 1986)

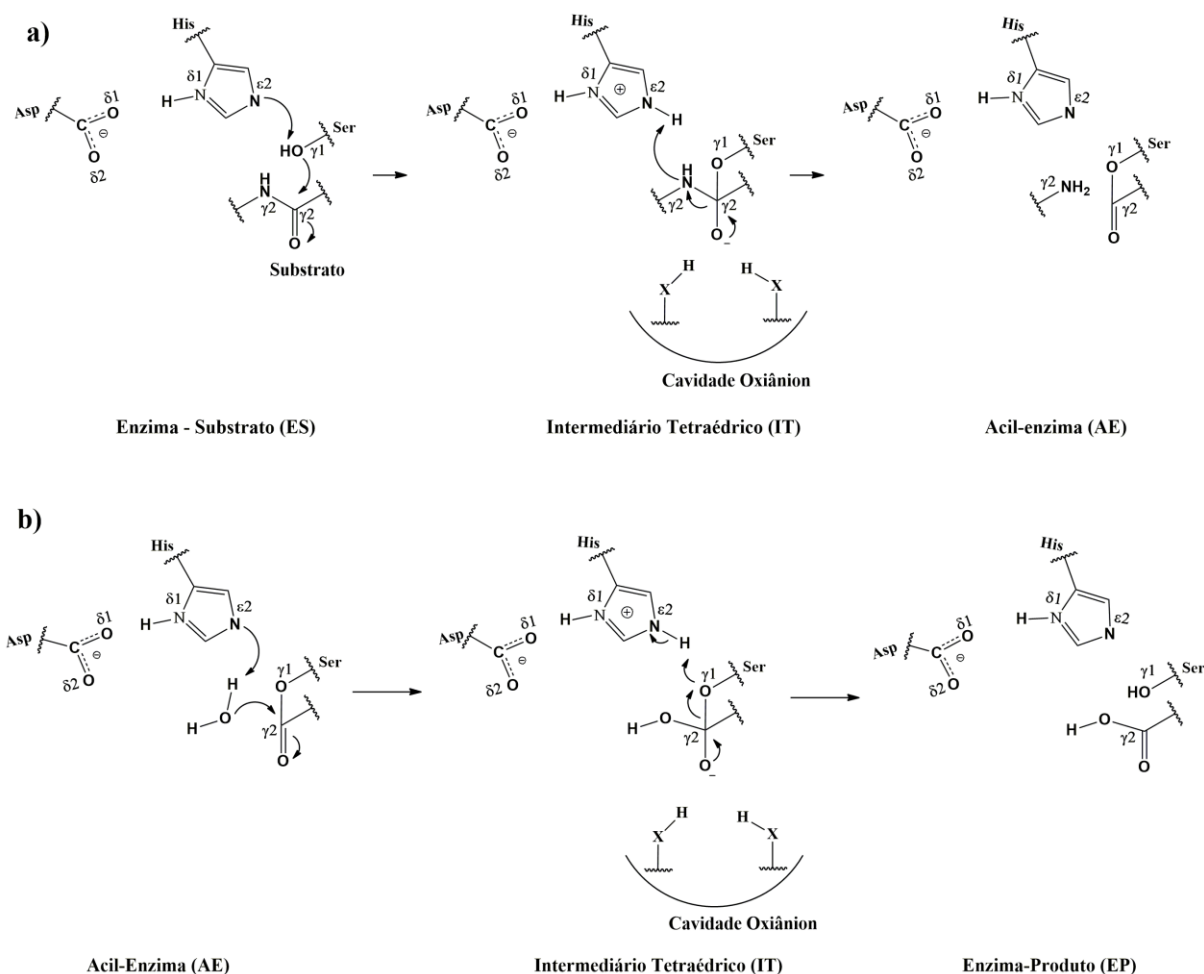


Intermediário Tetraédrico (IT)

**Figura 9:** Representação esquemática da região da cavidade oxianion e sua funcionalidade durante a formação do intermediário tetraédrico. Triade catalítica representada por Asp, His e Ser, onde a serina está formando a ligação acil com o substrato na formação do intermediário tetraédrico (carga formal negativa). Onde X representa átomos de nitrogênio ou de oxigênio. (WAHLGREN et al., 2011a)

### 1.3.3.3 Mecanismo

As principais características do mecanismo de ação das proteases de serina foram primeiramente resumidos por Bender e Zerner (1964) (ZERNER; BENDER, 1964) e Steitz (1970) (ERLANGER et al., 1970). Diversos outros trabalhos contribuíram para que hoje tenhamos o mecanismo das proteases de serina descrito na literatura.



**Figura 10:** Esquema para o mecanismo geral aceito das proteases de serina.(BLOW, 1976; HEDSTROM, 2002b) a) Etapa da acilação (acil-enzima) e quebra da ligação peptídica. b) Etapa da quebra da ligação acil – enzima. Na ilustração da cavidade oxianion, o X pode representar átomos de nitrogênio (N) ou de oxigênio (O).

Depois do substrato se aproximar do sítio ativo, o primeiro passo da reação é o ataque nucleofílico do oxigênio da hidroxila da serina (SER195 no caso da quimotripsina) ao carbono da carbonila do substrato. Já o hidrogênio da hidroxila da serina é atacado pelo par de elétrons do nitrogênio da histidina (sítio ativo), havendo uma transferência de próton entre serina e histidina. No segundo passo a ligação carbono-nitrogênio da ligação peptídica é quebrada e o intermediário denominado de acila-enzima é formado, composto pelo grupo acila (acila mais a enzima), formada pela enzima e a parte do substrato ligado ao carbono da carbonila, que foi atacada pela serina. O resto do substrato fica livre no meio. No terceiro passo, uma molécula de água presente irá hidrolisar o complexo acila-enzima, através da transferência de um próton da água ao grupo éster do fragmento acila. Assim, esse fragmento é eliminado e a serina volta à sua estrutura original, recompondo as características do sítio ativo. Com relação às mudanças estruturais e químicas durante a reação pode-se observar que,

depois do ataque nucleofílico do primeiro passo, ocorre uma alteração na geometria do carbono da carbonila que, antes trigonal plana, passa a ser tetraédrica. Esse intermediário tetraédrico é instável, devido à formação de uma carga formal negativa no oxigênio derivado da carbonila. A carga negativa é estabilizada por grupos NH de resíduos da proteína que ficam nessa região, e também estabiliza o estado de transição que precede a formação da estrutura tetraédrica. Para avaliar de maneira ilustrada o mecanismo geral ver a Figura 10. (HEDSTROM, 2002c)

Mais recentemente, um trabalho com estruturas de alta resolução mostrou outros aspectos interessantes sobre a catálise. Em 2006, Fodor et al. mostraram a relevância das ligações de hidrogênio durante o mecanismo de reação da tripsina. Estruturas intermediárias de proteases de serina revelaram que as forças das ligações de hidrogênio entre a enzima e o substrato são alteradas durante a catálise. As ligações de hidrogênio são conservadas entre a estrutura folha- $\beta$  da enzima e o substrato; entretanto, tornaram-se significativamente mais curtas no complexo de transição, de forma síncrona com o ataque nucleofílico ao átomo de carbono carboxílico da ligação peptídica que será quebrada. Assim o mecanismo utiliza a energia liberada a partir das ligações de hidrogênio fortes para superar a barreira energética do ataque nucleofílico pelo grupo hidroxila da serina catalítica. As ligações de hidrogênio da cavidade oxiânion proporcionam estabilização adicional para o complexo acil-enzima que terá sua ligação peptídica quebrada. No complexo de acil-enzima a distribuição de carga da ligação carbonílica é distorcida para o átomo de oxigênio em comparação com outras ligações peptídicas da estrutura, indicando um efeito de polarização da cavidade oxiânion. (FODOR et al., 2006)

Outro ponto importante, é o papel do aspartato (ASP) no processo catalítico. Quando a tripsina é mutada de forma que o aspartato é substituído por um resíduo neutro, como a asparagina, a constante de velocidade da reação é reduzida na ordem de  $10^4$ , suportando a importância da carga negativa próxima a histidina (HIS) (KOSSIAKOFF; SPENCER, 1981). Trabalhos experimentais e teóricos indicam que o aspartato (ASP) desprotona a histidina na posição delta ( $N_{\delta 1}$ , Figura 10), aumentando a basicidade da posição épsilon ( $N_{\epsilon 2}$ , Figura 10) para a transferência de próton da serina para histidina, ativando a serina para a formação do intermediário tetraédrico (KRAUT, 1977; WAHLGREN et al., 2011b).

#### 1.3.3.4 Evidência do intermediário acil-enzima

O experimento de Hartley e Kilbey, em 1954, evidenciou pela primeira vez o intermediário acil-enzima. Através do estudo da cinética da reação de hidrólise do acetato de *p*-nitrofenil na presença da enzima quimiotripsina. A hidrólise do acetato de *p*-nitrofenil pela quimotripsina indica a formação de um intermediário enzima-acetato e a decomposição deste intermediário foi o limitante da velocidade. Esse intermediário foi isolado em várias reações de ésteres com a quimotripsina, e a cinética química destes intermediários foi demonstrada. (HARTLEY; KILBY, 1954) Os intermediários acil-enzima foram posteriormente isolados em várias reações de ésteres com quimotripsina, assim como a cinética destes intermediários foram determinadas. (HESS et al., 1970a)

#### 1.3.3.5 Evidência do intermediário tetraédrico

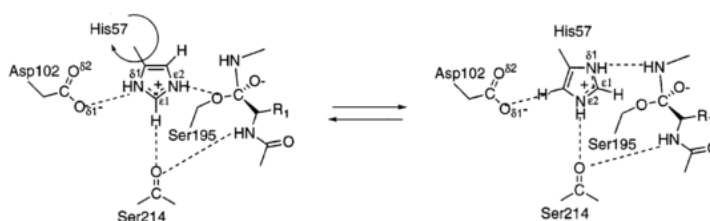
Entretanto, a existência do intermediário tetraédrico ainda não foi rigorosamente demonstrada nessas reações, mas é inferida apenas a partir de uma solução química. Intermediários tetraédricos são instáveis e têm tempo de vida curto. O tempo de vida do estado intermediário tetraédrico está associado ao grupo de saída quando se forma o complexo acil-enzima: quanto melhor o grupo de saída, menor o tempo de vida do estado intermediário tetraédrico. Isso pode ser evidenciado no caso do grupo de saída do *p*-nitrofenol, onde o tempo de vida do estado intermediário tetraédrico é menor que a menor vibração da molécula. A reação torna-se concertada, e o estado tetraédrico é considerado um estado de transição. Em solução, a hidrólise do *p*-nitrofenil acetato pela quimotripsina é uma reação concertada, e é formalmente possível que a hidrólise de peptídeos catalisada por esta enzima também seja concertada. (HEDSTROM, 2002c)

#### 1.3.3.6 Mecanismo de movimento da histidina

No mecanismo geralmente aceito para a catálise de protease de serina (Figura 10), a HIS57 retira o próton da SER195 e transfere-o para o grupo de saída. Esta transferência de próton intrigou alguns pesquisadores da época, que argumentavam que, se a HIS57 retira um próton da SER195, formando assim uma histidina com carga positiva  $HIS57-H^+$  e se posicionando perto da SER195 no intermediário tetraédrico, assim favorecendo a reprotonação da SER195, regenerando o estado inicial dos dois resíduos, HIS57 e SER195. Para reação proceder, a  $HIS57-H^+$  deveria estar posicionada próxima ao grupo de saída. O mecanismo chamado *ring flip* (inglês), um giro de  $180^\circ$  no anel da  $HIS57-H^+$ , foi proposto

para resolver esse problema: após a formação do intermediário tetraédrico, a HIS57-H<sup>+</sup> gira, o que deve colocar o próton do átomo N<sub>δ1</sub> perto do grupo de saída (Figura 11). A conformação da HIS57 invertida foi observada em subtilisina em 50 % de dimetilformamida. Além disso, a presença da ligação hidrogênio entre os resíduos SER214-HIS57 e, a dinâmica incomum de HIS57 foi uma evidência para o mecanismo "*His Flip*". (ASH et al., 2000)

Existem vários argumentos contrários ao mecanismo "*His Flip*". Primeiro, a quebra e formação de tantas ligações de hidrogênio num curto tempo de vida em que o intermediário tetraédrico é formado, parece improvável. O seu mecanismo de inversão parece violar o princípio de menor movimento, que sugere que as reações enzimáticas ocorrem com economia de movimento. Além disso, HIS57 está interagindo com resíduos do substrato, o que cria uma barreira estérica para giro. A dinâmica não usual do resíduo HIS57 pode ser explicada pelo movimento de entrada e saída do substrato e do solvente na região do sítio ativo, ou mesmo por uma transformação química, não necessariamente uma inversão de posição. (HEDSTROM, 2002c)



**Figura 11:** Esquema do mecanismo "*His Flip*". Depois da formação inicial do intermediário tetraédrico, o N<sub>ε1</sub>-H estará formando uma ligação de hidrogênio com o átomo γO da SER195. A rotação do His57 supostamente coloca o N<sub>ε1</sub>-H perto do grupo de saída NH. Embora as ligações de hidrogênio entre N<sub>δ1</sub>-H e ASP102 e, entre C<sub>ε1</sub> e Ser214 sejam rompidas, novas ligações de hidrogênio irão se formar para substituí-las (HEDSTROM, 2002c).

Alguns movimentos são inevitáveis, como do substrato e da protease para que a reação de hidrólise prossiga. Para a formação do intermediário tetraédrico, o resíduo SER195 da protease deve se movimentar de pelo menos 1 Å, assim como pequenos movimentos do substrato são requeridos para que o carbono da carbonila mude de uma hibridização sp<sup>2</sup> para uma sp<sup>3</sup>, permitindo a formação do intermediário. Portanto, a HIS57 pode não ser orientada para a direção da SER195 no intermediário tetraédrico como assumida, anteriormente. Mesmo se HIS57 estiver posicionada de modo que SER195 seja protonada no intermediário tetraédrico, a reação não está destinada a entrar em colapso, e voltar aos reagentes, pois a protonação da amina é favorecida em geral, e o próton necessita apenas se transferir para a amina antes da quebra do intermediário tetraédrico. Além disso, as várias estruturas dos



complexos análogos do estado de transição indicam que HIS57-H<sup>+</sup> pode interagir com a posição do grupo de saída, mantendo a ligação de hidrogênio ASP102-HIS57.(HEDSTROM, 2002b)

### 1.3.4 Aspectos Teóricos

Compreender o mecanismo detalhado de catálise enzimática é um dos problemas fundamentais em bioquímica. Uma das famílias de enzimas mais estudada experimentalmente e teoricamente são as proteases de serina que catalisam a hidrólise de ligações peptídicas. As características gerais do mecanismo de reação são amplamente aceitas, no entanto, alguns detalhes importantes ainda são controversos, como a transferência de próton entre o aspartato e a histidina do sítio ativo, assim como se o mecanismo de formação do intermediário acil-enzima ocorre em um mecanismo sincronizado ou em etapas, o movimento *flip* da histidina durante a reação, e estão sujeitos a pesquisa extensiva (HEDSTROM, 2002b).

Embora a função da tríade catalítica e da cavidade de oxianion possa ser pensada em termos de estabilização eletrostática das cargas no estado de transição, a catálise e a especificidade não são simplesmente controladas por alguns resíduos, mas são uma propriedade de toda a estrutura da proteína, controlada através da distribuição de carga por uma rede de ligações de hidrogênio e, talvez, também pelo acoplamento do movimento do domínio da protease para que haja a transformação química. Proteases de serina apresentam uma oportunidade única para o estudo teórico desses sistemas.

Compreender o funcionamento de enzimas em nível atomístico é uma questão central que vem sendo abordada por diversos estudos teóricos sobre estrutura e função de enzima (WANG; HU; ZHANG, 2007; WARSHEL; LEVITT, 1976a; WARSHEL; PAPAZYAN; KOLLMAN, 1995; WARSHEL; RUSSELL, 1986; ZHOU; ZHANG, 2011a).

Grande parte dos trabalhos teóricos se concentra em modelos em fase gás (DAGGETT; SCHROEDER; KOLLMAN, 1991; SCHROEDER; DAGGETT; KOLLMAN, 1991; TOPF; RICHARDS, 2004; TOPF; VÁRNAI; RICHARDS, 2002a; UMEYAMA; HIRONO; NAKAGAWA, 1984; UMEYAMA; IMAMURA; NAGATA, 1973; WARSHEL; LEVITT, 1976a; ZHOU; ZHANG, 2011a) do sítio ativo enzimático, isso devido ao menor poder computacional que até pouco tempo atrás era disponível. Estes modelos podem dar uma ideia qualitativa da atividade da enzima. Entretanto, devido à ausência do meio solvente, não são capazes de quantificar o efeito do ambiente enzimático sobre os processos. A utilização

de um modelo de solvente implícito (FEIG; BROOKS, 2004; ROUX; SIMONSON, 1999) pode superar essa limitação, mas também será deficiente em tratar o microambiente presente na região do sítio ativo da enzima, tornando necessário muitas vezes o tratamento explícito do solvente e outros resíduos do ambiente reacional. No entanto, tudo isso demanda maior poder computacional.

O estudo teórico de reações enzimáticas teve uma grande mudança depois da publicação do trabalho de Warshel e Levitt, em 1976, que tiveram uma abordagem pioneira para simular reações enzimáticas (WARSHEL; LEVITT, 1976a). Este foi o primeiro estudo que combinou fatores internos da reação com questões do ambiente reacional, por meio de um modelo que combinava mecânica molecular e mecânica quântica (semi-empíricos) para avaliar energeticamente a quebra de ligação química durante a reação enzimática. Os resultados mostraram que os efeitos eletrostáticos são importantes para a estabilização do íon carbânion que é formado no intermediário dessa reação. A partir desse trabalho tornou-se um momento propício para o desenvolvimento de uma nova abordagem no estudo de reações enzimáticas (WARSHEL; LEVITT, 1976b). Alguns estudos teóricos usaram essa metodologia para melhor entender o mecanismo reacional das proteases de serina.

Como já mencionado anteriormente por trabalhos experimentais (BLOW, 1976; BLOW; BIRKTOFT; HARTLEY, 1969b; HESS et al., 1970b; KRAUT, 1977; SPENCER; STURTEVANT, 1959), existia uma questão que levantava muita controvérsia no caso das proteases de serina: a de que existia uma dupla transferência de próton no sítio ativo durante o mecanismo reacional. Isso também foi estudado na época por trabalhos teóricos (DAGGETT; SCHROEDER; KOLLMAN, 1991; SCHROEDER; DAGGETT; KOLLMAN, 1991; TOPF; RICHARDS, 2004; TOPF; VÁRNAI; RICHARDS, 2002a; UMEYAMA; HIRONO; NAKAGAWA, 1984; UMEYAMA; IMAMURA; NAGATA, 1973; WARSHEL; LEVITT, 1976a; ZHOU; ZHANG, 2011a). O primeiro estudo questionando essa conclusão foi o de Hayes e Kollman apontando que, ao serem realizadas correções para os erros nos cálculos de afinidade de próton, e quando são incluídas ligações de hidrogênio no ácido aspártico, os cálculos *ab initio* mostram que só deve haver uma transferência de próton, que se dá entre serina e histidina (KOLLMAN; HAYES, 1981). Esta análise foi apoiada depois por Umeyama et al. e Kollman e Hayes, com um modelo de química quântica (QM) de nível superior. (KOLLMAN; HAYES, 1981; UMEYAMA; IMAMURA; NAGATA, 1973) A primeira resposta experimental definitiva à esta questão foi dada por Kossiakoff e Spencer. (KOSSIAKOFF; SPENCER, 1981) Por estudo de difração de nêutrons da tripsina na

presença de um inibidor análogo ao substrato, observou-se a protonação dos resíduos do sítio ativo, indicando que o hidrogênio  $H_{\delta 1}$  (Figura 10) está posicionado no  $N_{\delta 1}$  da histidina. Concluíram assim, portanto, que a eficácia catalítica da tríade ASP-HIS-SER não é devida a algum efeito químico perpetuado através do sistema de ligação de hidrogênio tal como proposto na troca de carga (*charge-relay*), (BLOW; BIRKTOFT; HARTLEY, 1969b) e nem pelo mecanismo de dupla transferência de prótons, (HUNKAPILLER et al., 1973) mas em função das características químicas destes grupos (ASP, HIS, SER) que maximizam a eficiência e eficácia pelo posicionamento estereoquímico no complexo enzima-substrato. Estudos espectroscópicos por RMN- $H^1$  mostram que o  $H_{\delta 1}$  estaria fazendo uma ligação de hidrogênio entre o  $O_{\delta 2}$  e o  $N_{\delta 1}$ , como também, indicam que essa ligação de hidrogênio apresenta uma baixa barreira de energia para a transferência do próton ( $H_{\delta 1}$ ) entre  $N_{\delta 1}$  (histidina) e o  $O_{\delta 2}$  (aspartato), deste modo facilitando a formação do intermediário tetraédrico. (BACHOVCHIN, 2001; CLELAND; KREEVOY, 1994; FREY; WHITT; TOBIN, 1994)

Warshel et al. também mostraram que a reação em fase gás é caracterizada por uma barreira grande de ativação (associada com a formação do nucleófilo carregado a partir de fragmentos neutros), embora o ataque nucleofílico seja essencialmente por barreiras menores nas enzimas. Por outro lado, as reações enzimáticas e em solução apresentam um perfil energético reacional semelhante, com uma barreira mais baixa para a reação enzimática. Eles concluem que os sítios ativos das enzimas proporcionam ambientes polares específicos que não se assemelham a fase gás, mas que são projetados para a estabilização eletrostática de estados de transição iônicos, e que solvatam esses estados mais do que só o solvente. (WARSHEL; AQVIST; CREIGHTON, 1989)

Schroeder et al. examinaram uma outra questão sobre a catálise em protease de serina: qual é o passo limitante para a velocidade global da hidrólise de amida e de éster? Descreveram cálculos semi-empíricos de orbitais moleculares da hidrólise de amida e ésteres catalisados pela tripsina, observando que o ataque da serina ao substrato, na formação do intermediário tetraédrico, foi a etapa limitante com ambos os substratos. Assim, o caminho reacional de menor energia para formação do intermediário tetraédrico se realizou pela aproximação da serina ao substrato, em seguida pelo movimento acoplado de átomos pesados e, da transferência de próton para completar a reação. Esse trabalho foi um dos poucos onde todo o caminho da catálise foi simulado e o passo limitante da velocidade foi sugerido. O efeito do meio sobre a catálise também foi estimado por cálculos de mecânica molecular,

usando um modelo não covalente entre o peptídeo e a tripsina, mostrando a importância do ambiente na estabilização do intermediário tetraédrico. Uma simulação de dinâmica molecular do complexo não covalente também foi realizada, observando a importância do movimento local do sítio ativo nas fases iniciais da catálise. Ficou evidente a necessidade de cálculos precisos de química quântica, assim como uma representação do ambiente químico para definir o caminho da catálise enzimática.(SCHROEDER; DAGGETT; KOLLMAN, 1991)

Em 2002, um trabalho de simulação de dinâmica QM/MM foi utilizado para examinar a estabilidade do intermediário tetraédrico durante a etapa de desacilação da hidrólise de um peptídeo simples, que foi catalisada pela enzima elastase.(TOPF; VÁRNAI; RICHARDS, 2002b) A região quântica incluía a tríade catalítica, a estrutura tetraédrica, e a cavidade oxianion. Os cálculos indicaram que o intermediário tetraédrico é uma espécie estável na escala de tempo de picossegundos, além de outras observações como a mobilidade do anel imidazol entre dois dos átomos de oxigênio da estrutura tetraédrica, que mostra como o intermediário pode proceder para o estado dos produtos sem um “mecanismo de inversão de anel”. Além da proposta de ligação de hidrogênio entre o anel de imidazol e da cadeia principal da SER214, observa-se uma ligação de hidrogênio alternativa com a cadeia principal do resíduo THR213, que pode estabilizar o intermediário durante o movimento do anel imidazol. O próton se desloca entre ASP102 e HIS57 durante a simulação. Entretanto, o próton fica grande parte do tempo localizado no átomo de nitrogênio da histidina. (TOPF; VÁRNAI; RICHARDS, 2002b)

Ishida e Kato, em 2003, estudaram o mecanismo de reação da tripsina por dinâmica molecular QM/MM e cálculos de energia livre. Calculando o perfil de energia livre para o primeiro passo da reação, acilação, os cálculos mostraram que o passo determinante para a velocidade de acilação é a formação do intermediário tetraédrico, e a quebra desse intermediário tem uma pequena barreira energética. (ISHIDA; KATO, 2003a) A energia livre de ativação para a formação desse intermediário foi de 17,8 kcal/mol usando QM/MM (MP2/(aug)-cc-pVDZ//HF/6-31+G\*\*/AMBER) combinado com *Molecular Dynamics-Free Energy Perturbation: MD-FEP*.(ISHIDA; KATO, 2003b) Já simulações de dinâmica molecular para o complexo enzima-substrato (ES), enzima livre, e em fase gás mostram que a ligação do substrato induz alterações conformacionais em torno do sítio ativo, as quais favorecem o alinhamento dos fragmentos reativos (HIS57, ASP102 e SER195) que ficam numa orientação reativa. Observou-se também que a etapa de transferência do próton entre a

serina e a histidina, e o ataque nucleofílico da serina ocorrem de maneira concertada. Como relação ao ambiente, os resultados mostraram que ele desempenha maior estabilização no intermediário tetraédrico quando comparado ao complexo enzima-substrato. O fator mais importante de estabilização do intermediário tetraédrico é a interação eletrostática entre o sítio ativo e algumas regiões específicas da proteína: os grupos NH da cadeia principal dos resíduos GLY193 e SER195 (cavidade oxiânion) estabilizam a carga negativa gerada no oxigênio da carbonila quando forma o intermediário, e os grupos carbonila da cadeia principal dos resíduos ILE212 e SER214 estabilizam uma carga positiva gerada no anel imidazol da histidina (His57).(ISHIDA; KATO, 2003a)

Ishida e Kato, em 2004, fizeram um estudo teórico sobre o papel do ASP102 no sistema catalítico de protease de serina, quimotripsina, calculando o perfil da energia livre de reação da transferência de um próton para esse resíduo, usando cálculos de estrutura eletrônica QM/MM (HF/6-31+G\*\*/AMBER) combinadas com dinâmica molecular e o método da perturbação para energia livre. (ISHIDA; KATO, 2004) O ASP102 desempenha dois papéis importantes no processo catalítico: um é para estabilizar a HIS57 protonada, e o outro é o de manter a configuração em torno do sítio ativo, pois favorece o processo catalítico. Esses dois fatores estão intimamente relacionados entre si e são indispensáveis para uma catálise eficiente. Além disso, os cálculos apresentados sugerem a importância da interação entre a HIS57 e os resíduos SER214-VAL213 no estado de transição da etapa de acilação (numeração dos resíduos referente à quimotripsina).(ISHIDA; KATO, 2004)

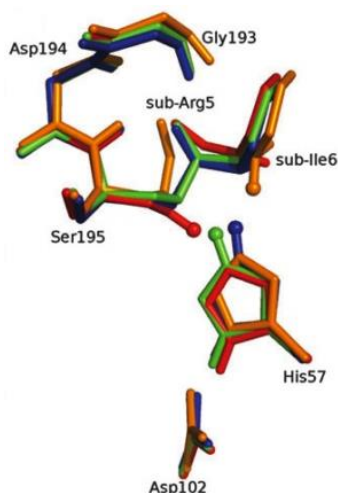
Em 2006, Ishida publicou um outro trabalho que analisava de maneira teórica o RMN para a tríade catalítica da tripsina, observando a correlação entre a variação dos deslocamentos químicos e o rearranjo estrutural do sítio ativo no processo de acilação do mecanismo. (ISHIDA, 2006) Para isso fez cálculos de dinâmica molecular QM/MM (HF/6-31+G\*\*/AMBER). Os deslocamentos químicos obtidos foram consistentes com os experimentos, mostrando que o próton entre a ASP102 e HIS57 fica localizado próximo a HIS57, existindo uma forte interação eletrostática entre os dois resíduos ASP102 e HIS57. Quando o intermediário se formou se observou uma interação eletrostática entre a HIS57 e a cadeia principal dos resíduos VAL213 e SER214. (ISHIDA, 2006)

Em 2011, Zhang e Zhou determinaram o mecanismo de acilação de um protótipo de serina protease (tripsina) e seu perfil de energia livre por simulação de dinâmica molecular QM/MM e amostragem de guarda-chuva (do inglês, *Umbrella Sampling*). É importante mencionar alguns aspectos desse artigo, pois faz diversas observações sobre pontos já

mencionados anteriormente e polêmicos sobre a reação, além de utilizar uma metodologia mais recente para o tratamento de reações enzimáticas.(ZHOU; ZHANG, 2011b)

A estrutura inicial do complexo de substrato-enzima foi modelada com base na estrutura de raios-X do complexo tripsina-inibidor (PDB ID: 1MCT). A sequência de aminoácidos para o substrato foi alinhada a partir da posição do inibidor no sítio ativo da enzima, e a ligação a ser quebrada entre ARG5 e ILE6. O subsistema QM (nível B3LYP/6-31+G\*) incluiu as cadeias laterais da tríade catalítica (HIS57, ASP102 e SER195) e a porção do substrato que teve a ligação peptídica quebrada. Deste modo, usando 41 janelas, cada janela com 30 ps, de *Umbrella Sampling* para percorrer a coordenada de reação. Eles obtiveram uma barreira de ativação total de 16,9 kcal/mol, em boa concordância com o valor experimental de 15-20 kcal/mol, dependendo do substrato e das condições experimentais.(ZHOU; ZHANG, 2011b)

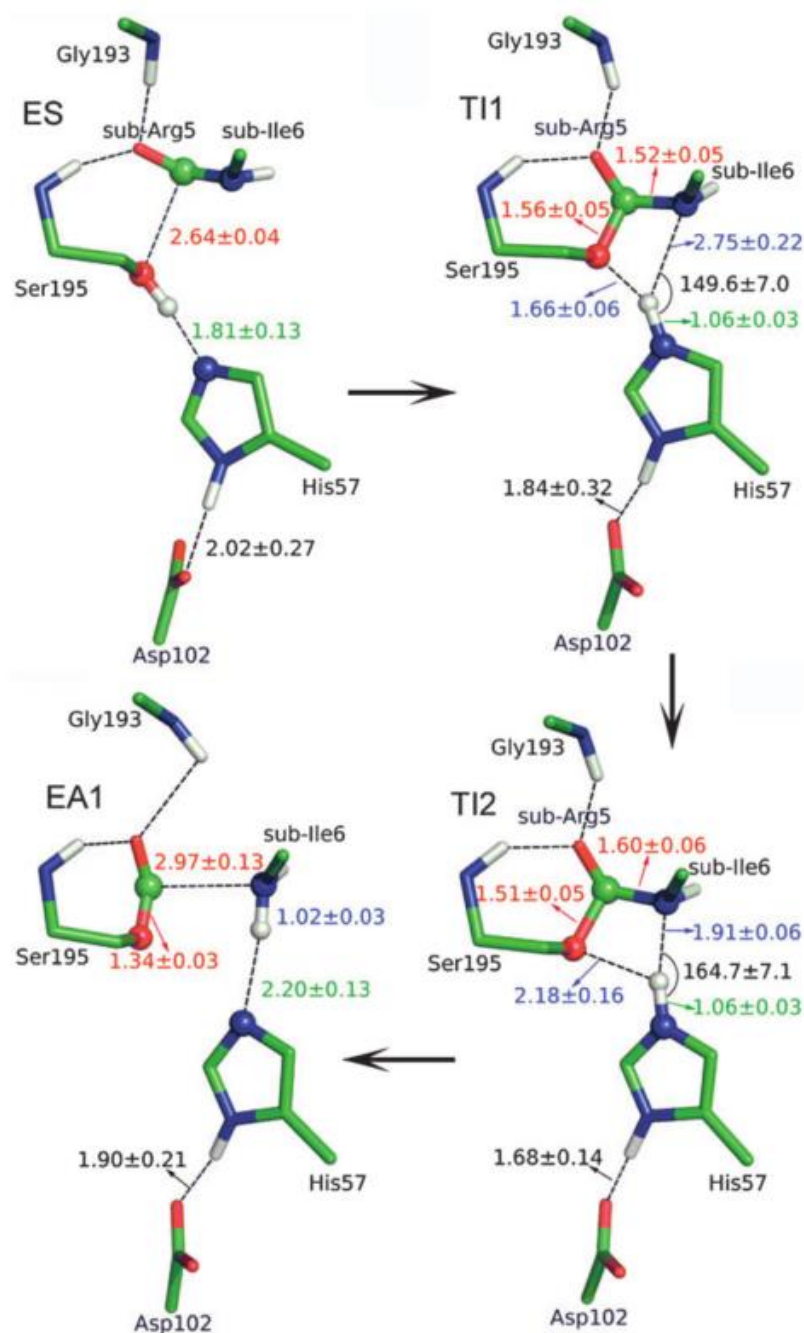
Os resultados desse estudo indicam que a posição do anel de HIS57 se altera pouco durante a reação (Figura 12).



**Figura 12:** Superposição de estruturas da região sítio ativo da tripsina para ES (em vermelho), TI1 (em verde), TI2 (em azul) e EA1 (em laranja) no modo de palito. O átomo do tipo HG da SER195 é apresentado no modo de esfera. As estruturas enzima-substrato(ES), intermediário tetraédrico 1 (TI1), intermediário tetraédrico 2 (TI2) e acil-enzima (EA1) pode ser melhor vista na **Figura 13**. Adaptado da referência (ZHOU; ZHANG, 2011b)

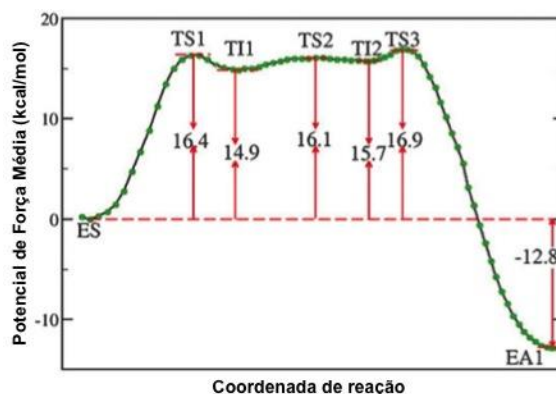
A sutil reorientação do seu anel no intermediário tetraédrico serve como uma ponte que liga duas etapas químicas: primeiro HIS57 atua como uma base, aceitando o próton da SER195, e depois age como um ácido protonando o grupo amina de saída (substrato). Esta troca de prótons pela histidina prossegue com uma economia de movimento atômico, contrariando a teoria do '*His-Flip*' (1.3.3.6 Mecanismo de movimento da histidina ), ou

mesmo a inversão do nitrogênio do grupo amina de saída (substrato). Para a ligação de hidrogênio entre HIS57 e ASP102, ela se torna um pouco mais curta no estado de transição e nos intermediários do que nos estados de menor energia (ES e EA1), indicando uma interação mais forte devido à carga positiva no anel (HIS57), na Figura 13. Já a Figura 14 mostra o perfil energético para todas as estruturas, como suas barreiras de ativação. Durante a simulação nenhuma transferência de próton espontânea entre ASP102 e HIS57 foi observada. Os resultados mostram que o principal papel do resíduo ASP102 na catalise é estabilizar a carga positiva do anel imidazol (HIS57) nos estados TS e TI através de interações eletrostáticas. Isso é mais uma evidência contrária ao mecanismo de redistribuição de carga (do inglês, *charge-relay mechanism*) proposto por Blow et al.(BLOW; BIRKTOFT; HARTLEY, 1969a)



**Figura 13:** Ilustração das estruturas para a fase de acilação da tripsina. ES, complexo enzima-substrato; TI1 e TI2 duas configurações distintas intermediários tetraédricos; EA1, acil-enzima. Adaptado da referência (ZHOU; ZHANG, 2011b)





**Figura 14:** Perfil de energia livre para a etapa da reação de acilação da tripsina determinado pela simulação de dinâmica molecular QM/MM por amostragem guarda-chuva (do inglês, *umbrella sampling*). Adaptado da referência (ZHOU; ZHANG, 2011b)

#### 1.4 NS2B-NS3 PROTEASE DO VÍRUS DA DENGUE

A protease viral NS3 tem uma tríade catalítica (HIS51, ASP75, SER135) que é essencial para a replicação viral (como visto para as proteases de serina, caso da NS3) e é responsável pela clivagem das junções NS2A/NS2B, NS2B/NS3, NS3/NS4A e NS4B/NS5 na poliproteína.(CHAMBERS et al., 1990)

Li et al., em 1999, revelaram que o domínio responsável pela atividade da protease do vírus da Dengue é constituído por 167 aminoácidos da parte N-terminal da NS3.(LI et al., 1999) O resíduo His51, que pertence a NS3 protease e também à tríade catalítica, sendo fundamental para a função da protease. (FALGOUT; MILLER; LAI, 1993)(CLUM, 1997)(YUSOF et al., 2000)

Experimentos *in vitro* (YUSOF et al., 2000) mostraram que a NS3 protease por si só apresenta uma atividade limitada para vários substratos peptídicos sintéticos enquanto, para processar os substratos naturais, o domínio da protease da NS3 é co-expressa com um cofator, a proteína NS2B.(YUSOF et al., 2000)

Falgout et al., em 1993, mostrou que preservar um fragmento central de 40 aminoácidos (resíduos 53-92 da NS2B) no precursor NS2B é suficiente para produzir um produto ativo junto com a NS3 (NS2B-NS3). A atividade desse complexo NS2B-NS3 foi demonstrada em sistemas de expressão tanto eucarióticas quanto procarióticas.(FALGOUT; MILLER; LAI, 1993; YUSOF et al., 2000)

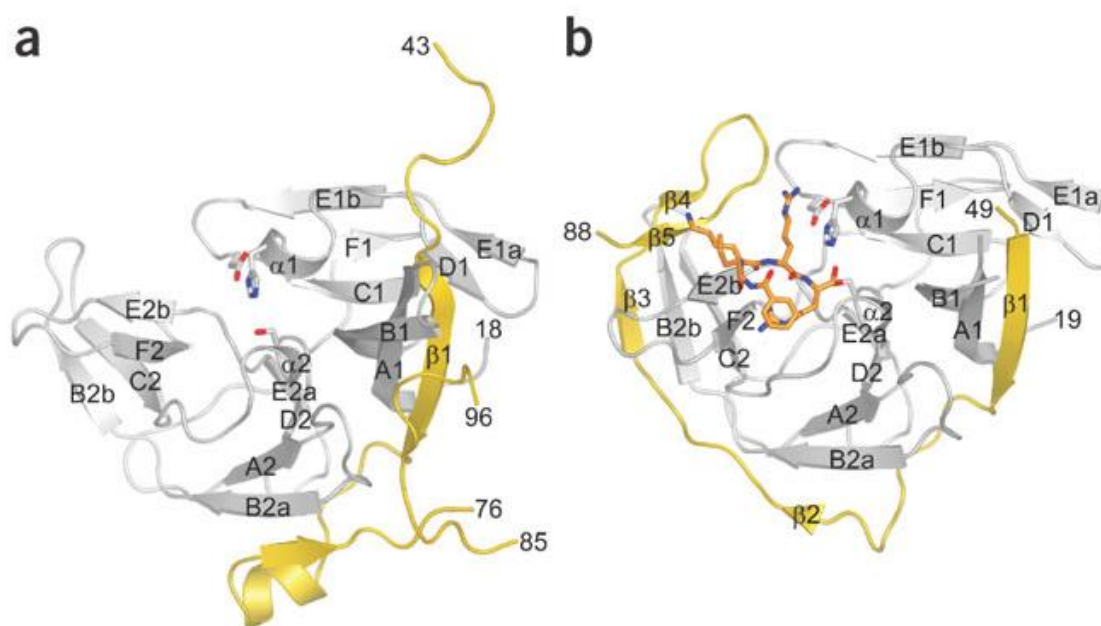
Erbel et al., em 2006, relatou as estruturas cristalinas de um complexo NS2B-NS3pro do vírus da Dengue (DENV) e, um complexo NS2B-NS3pro com um inibidor do vírus do Nilo Ocidental (WNV). Cristais de alta qualidade pra NS2B-NS3pro da Dengue (1,5 Å) e para o Nilo Ocidental (1,68 Å), e o complexo WNV NS2b-NS3pro com inibidor baseado em

modelo de substrato benzoil-norleucina (P4) - lisina (P3) - arginina (P2) - arginina (P1) - aldeído (Bz-NLE-LYS-ARG-ARG-H) foram obtidos e as estruturas foram resolvidas. Com essas estruturas foi possível identificar resíduos importantes para o reconhecimento do substrato na NS3pro e esclarecer a importância do cofator NS2B para estrutura ativa da protease no caso do WNV. (ERBEL et al., 2006)

Os domínios NS3pro em ambas as estruturas adotaram a conformação semelhante à da quimotripsina, ou seja, dois barris- $\beta$ , cada um formado por seis folhas- $\beta$ , com a tríade catalítica (HIS51-ASP75-SER135) localizada na fenda entre barris (Figura 15). Os dois vírus apresentam 50 % de identificação das sequências dos aminoácidos para a NS3pro, e a superposição das estruturas apresenta um r.m.s de apenas 1.37 Å. (ERBEL et al., 2006)

A conformação adotada pela parte C-terminal da NS2B revelou diferenças marcantes entre os complexos cofator-protease com inibidor ligado e não ligado. Na estrutura da NS2B-NS3pro da Dengue (Figura 15), a densidade de elétrons para resíduos além do resíduo 76 é descontinua, o que sugere que a parte C-terminal do cofator para essa estrutura adota diversas conformações quando em solução. NS2B não interage com o local de ligação do inibidor, e a atividade da NS2B-NS3pro não pode ser explicada por esta estrutura. Em contraste, para o caso da estrutura do Nilo Ocidental (Figura 15), o complexo NS2B-NS3pro com inibidor revela interações diretas da NS2B (ASP82-PHE85) com o sítio ativo da NS3pro. (ERBEL et al., 2006)

A NS2B não só estabiliza ambos os barris de N-terminal e C-terminal, fornecendo folhas- $\beta$  adicionais, como também completa o local de ligação ao inibidor com a sua região C-terminal. As estruturas apresentadas proporcionaram, pela primeira vez, uma explicação para o aumento na atividade da NS3pro na presença de NS2B. (ERBEL et al., 2006) As coordenadas para essas estruturas estão no PDB (*Protein Data Bank*) com os códigos: 2FOM (DENV NS2B-NS3pro) e 2FP7 (WNV NS2B-NS3pro-Bz-NLE-LYS-ARG-ARG-H). (ERBEL et al., 2006)

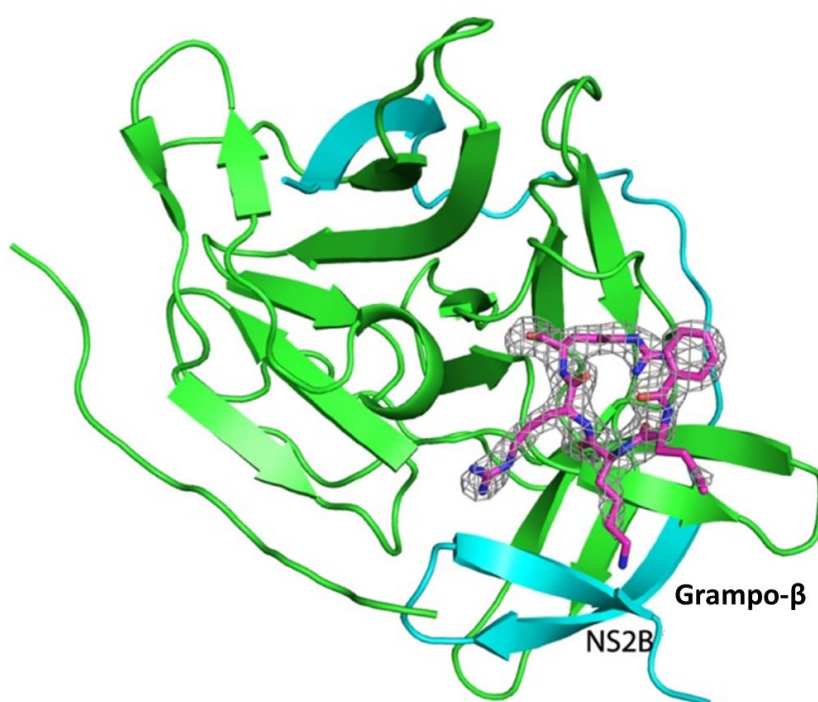


**Figura 15:** Ilustração das estruturas de NS2B-NS3pro na ausência e na presença de um inibidor. (a) DEN NS2B-NS3pro. Cinza, NS3pro; amarelo, NS2B. Não foi observada densidade de elétrons dos resíduos NS2B 77-84. (b) WNV NS2B-NS3pro em complexo com Bz-Nle-Lys-Arg-Arg-H (laranja). Adaptado da referência (ERBEL et al., 2006)

Wichapong et al., em 2010, geraram e validaram um modelo de homologia da forma ligada do inibidor com protease NS2B-NS3pro para o vírus da Dengue (DV). O modelo DV-2 por homologia foi construído com base unicamente na estrutura de Raio-X do homólogo do complexo NS2B-NS3pro com inibidor, e apresentou resultados que se encontram em bom acordo com os dados experimentais. Dessa maneira, sugeriram que o modelo DV-2 representava a protease NS2B-NS3 para Dengue na forma ativa. O modelo obtido e as simulações dinâmica molecular (DM) revelaram elevada similaridade com os resultados obtidos para o cristal de protease do caso do vírus do Nilo Ocidental (WNV). As simulações DM revelaram a existência de fortes interações entre o domínio C-terminal da NS2B e a NS3, ajudando na estabilidade das regiões de loop da NS3 protease. Os resultados indicam que o domínio C-terminal da NS2B não é apenas importante para a ligação com a NS3 protease, mas também desempenha um papel importante para a interação com o resíduo da posição P2 do inibidor, além de mostrar outras interações importantes entre a NS2B-NS3 com os resíduos do inibidor e outros resíduos da proteína. Assim, este trabalho descreve pela primeira vez a interação do inibidor peptídico com NS2B-NS3 protease da Dengue utilizando uma abordagem de simulação DM. (WICHAPONG et al., 2010b)

Foi apenas em 2012 que uma estrutura cristalina para o complexo formado por um peptídeo ligado covalentemente a NS2B-NS3pro do vírus da Dengue (sorotipo 3, DEN-3),

bem como o inibidor de serina-protease aprotinina ligado à mesma enzima, foram obtidas. (Noble et al., em 2012) Estas estruturas revelaram, pela primeira vez, uma conformação fechada que é cataliticamente ativa para a protease do DENV. Na presença do peptídeo, a protease do DENV-3 constitui a conformação fechada, em que a região hidrófila do grampo- $\beta$  da NS2B fica em torno da NS3 protease, de uma maneira análoga à da estrutura da protease do vírus do Nilo Ocidental (WNV). Os resultados confirmaram que a protease (NS2B-NS3) para o vírus da Dengue forma uma conformação fechada durante proteólise, como anteriormente proposto para a WNV. A Figura 16, mostra a estrutura para o complexo NS2B-NS3 protease da Dengue tipo 3 com inibidor Bz-nKRR-H (benzoila-norleucina-LYS-ARG-ARG-aldeído). (NOBLE et al., 2012)



**Figura 16:** Ilustração da estrutura NS2B-NS3 protease da Dengue ligada ao Bz-nKRR-H. A estrutura é mostrada na representação de *cartoon*, NS3 em verde e NS2B em azul, com o Bz-nKRR-H em representação stick. Adaptada da referência (NOBLE et al., 2012).

#### 1.4.1 Inibição da NS2B-NS3 do vírus da Dengue

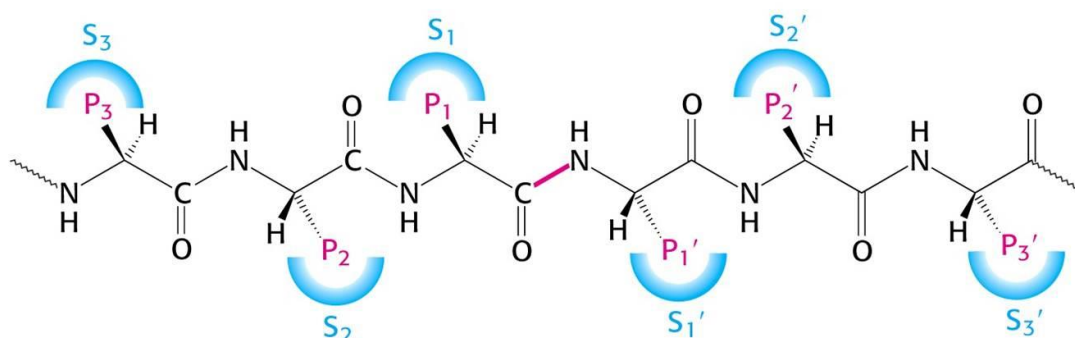
##### 1.4.1.1 Generalidades

Compostos que agem inibindo ou modulando a atividade das proteases tem efeitos nos processos biológicos em que essas proteases atuam. Muitos desses inibidores naturais se assemelham aos substratos com relação à estrutura. (FERSHT; REQUENA, 1971; HIROARA;

BENDER; STARK, 1974) O fato de a enzima estabilizar o estado de transição vem sendo demonstrado muitas vezes através da utilização de análogos do estado de transição (SCHRAMM, 2011), que são moléculas que mimetizam o estado de transição de substrato. No caso das proteases de serina é comum construir inibidores que se pareçam com a estrutura tetraédrica observada no mecanismo da reação. (HEDSTROM, 2002c)

Estudos mostram um padrão de interação dos resíduos que compõem o inibidor ou substrato com regiões específicas da enzima. Os locais de reconhecimento do substrato incluem o próprio sítio ativo da enzima, como também, regiões de bolsões de interação que acomodam as cadeias laterais dos peptídeos do substrato. (HEDSTROM, 2002c)

Os estudos normalmente focam em uma interação específica da protease com o substrato. Com base na nomenclatura proposta por Harper e Berger, (HARPER; BERGER, 1972) a interação mais focada é a P1-S1, onde P1 e P1' (Figura 17) indicam resíduos peptídicos no grupo acila, em que um ficará ligado a serina e outro sairá como grupo de saída pela clivagem da ligação peptídica. S1 é a região da proteína com a qual P1 interage. A numeração é dada segundo a Figura 17, e os sítios de interação  $S_n$  interagem com  $P_n$  correspondente. Cada bolsão tem características químicas diferentes para interagir com resíduos distintos: uns acomodam melhor resíduos com cargas negativas, outros com carga positiva, outros ainda interagem melhor com resíduos neutros, hidrofílicos, ou também com hidrofóbicos. Isso deve ser levado em consideração quando se cria modelos para se estudar a interação do substrato com a proteína (HEDSTROM, 2002c)

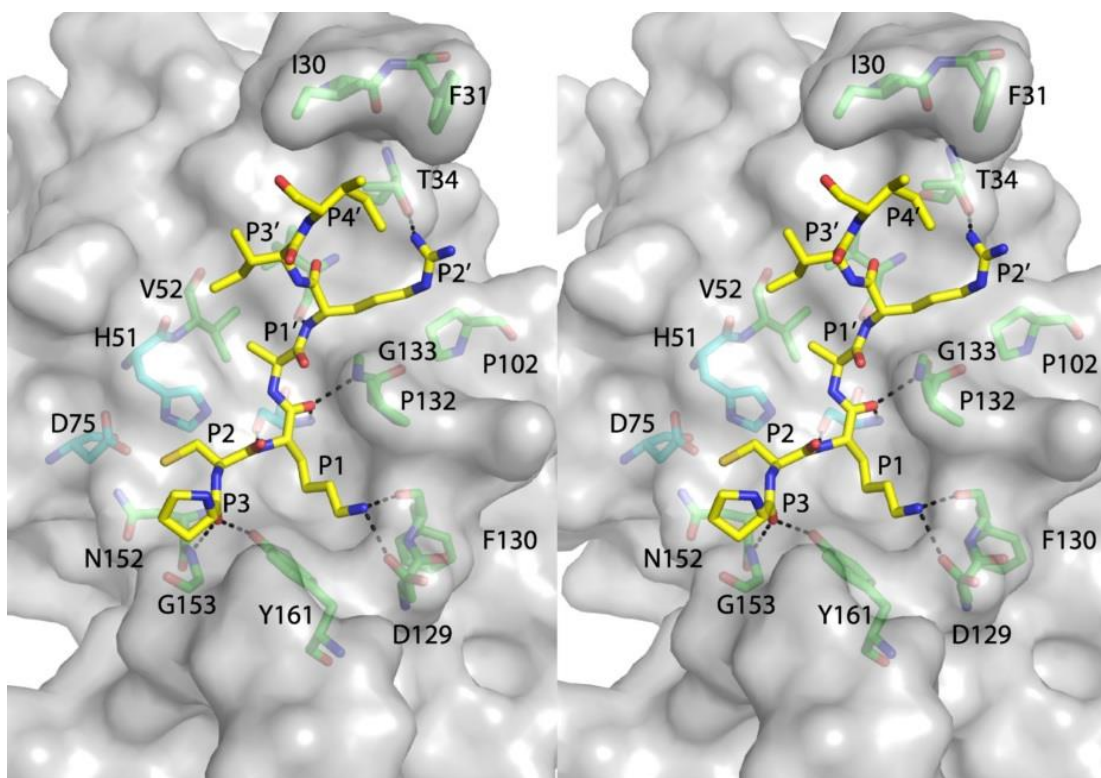


**Figura 17:** Os potenciais sítios de interação do substrato com a enzima são designados P (mostrados em vermelho), e os correspondentes locais de ligação à enzima são designados S (subsítio). A ligação peptídica (em vermelho) é o ponto de referência. Adaptado da referência (HEDSTROM, 2002c)

Para as posições P1' e P2', a protease do WNV mostra uma preferência por GLY (P1')-GLY(P2') (SHIRYAEV et al., 2007), ao passo que no caso do DENV, há preferência pela SER na posição P1', como também mostra uma fraca preferência para os resíduos ácidos na posição P2' (LI et al., 2005). No caso da DENV, a posição P3' do substrato mostra uma

forte preferência para SER (LI et al., 2005). A DENV também exibe preferência para GLY na posição P4' do substrato.(LI et al., 2005)

Com relação ao estudo da preferência da clivagem da ligação peptídica do substrato, estudos mostram que a NS2B-NS3pro do vírus da dengue apresenta uma preferência para clivar ligações peptídicas com uma sequência de aminoácidos característica. (SHIRYAEV et al., 2007) Representando a sequência de um substrato peptídico por -P4-P3-P2-P1-P1'-P2'-P3'-P4'-, na qual a ligação peptídica clivada ocorre entre P1 e P1', a díade (LYS/ARG)-ARG é comumente encontrada nas posições P2 e P1 em várias sequências na poliproteína precursora. Em estudos com modelos de substrato de oito aminoácidos de comprimento, a sequência que apresentou maior eficiência catalítica foi GKKRR↓PVR (a seta para baixo indica a região da quebra da ligação entre os resíduos).(SHIRYAEV et al., 2011)



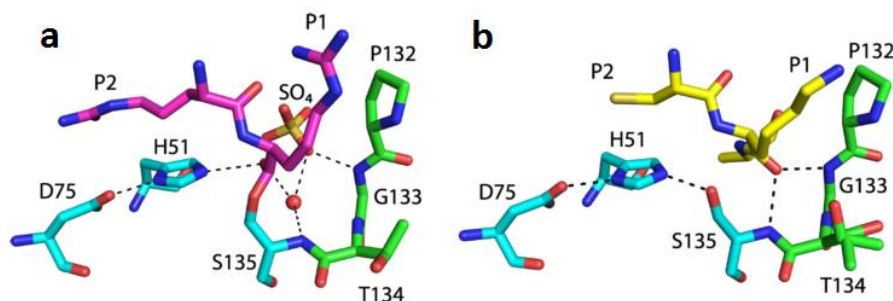
**Figura 18:** Detalhe das interações entre NS2B-NS3 protease da DEN-3 e a aprotinina (inibidor). Diagrama dos resíduos da aprotinina (mostrado como *stick* amarelo) na superfície da protease NS2B-NS3 da DEN-3 (cinzento). As possíveis ligações de hidrogênio estão indicadas por linhas tracejadas. Outros resíduos que revestem os bolsões são mostrados com *stick* verdes, com a tríade catalítica com *stick* azuis. Imagem em estéreo, adaptada da referência (NOBLE et al., 2012)

Em ambas as estruturas da protease da DENV-3 com Bz-nKRR-H ou com aprotinina (Figura 18), a cavidade oxiânion é formada pelos resíduos G133, T134, e S135. A cavidade oxiânion é uma região para acomodar o intermediário tetraédrico carregado negativamente, e



que é formada durante a reação do substrato com a protease. No caso da protease da DENV-3 com a Bz-nKRR-H, a região da cavidade oxianion foi ocupada por uma molécula de água e um sulfato. Quando a ligação peptídica GLY133 é invertida, leva a estrutura da cavidade oxianion “fechada” (uma conformação inativa). Foi proposto anteriormente que o fechamento da cavidade oxianion é devido à ausência de um resíduo na posição P1. (NOBLE; SHI, 2012) Entretanto, no caso da Figura 19(a) vemos uma conformação ativa da cavidade oxianion, mesmo sem a presença do resíduo P1 do Bz-nKRR-H. Este resultado indica que um resíduo em P1’ não é necessário para esta conformação ativa.(NOBLE et al., 2012)

Na estrutura da protease (DENV-3) com aprotinina, existem fortes interações do tipo de ligação de hidrogênio entre os resíduos da tríade catalítica (HIS51, ASP75, e SER135); todas as ligações de hidrogênio estão dentro de uma distância de 2,7 Å (Figura 19a). O grupo carboxílico do resíduo LYS (P1) na aprotinina está posicionado na cavidade oxianion (Figura 19b), indicando que a cavidade oxianion está na conformação ativa. Uma conformação semelhante foi observada num complexo protease-aprotinina para o WNV.(ALESHIN et al., 2007) Estes resultados indicam fortemente que as estruturas representam conformações cataliticamente competentes.(NOBLE et al., 2012)



**Figura 19:** Ilustração detalhada da estrutura da região catalítica para a protease (NS2B-NS3) da DENV-3 na presença de inibidores, (a) Bz-nKRR-H e (b) aprotinina. (a) representação em stick para a cavidade oxianion (verde), para a tríade catalítica (H51, D75, e S135) (azul) e do inibidor Bz-nKRR-H (rosa). Apenas resíduos na posição P1 e P2 são mostrados por motivos de clareza. A água e o sulfato são mostrados como uma esfera e stick (vermelho), respectivamente. As possíveis ligações de hidrogênio (até 3,2 Å) são indicadas pelas linhas tracejadas. (b) A representação para a tríade e para a cavidade oxianion é a mesma do (a). Os resíduos P1 a P2 da aprotinina são mostrados como stick de cor amarela, e o grupo carbonilo do resíduo P1 ocupa a cavidade oxianion. Adaptada da referência (NOBLE et al., 2012)

Inibidores de NS2B/NS3pro contêm uma sequência de peptídeos altamente carregados, com elemento dibásico P1-P2 acoplado a uma variedade de grupos eletrofílicos (aldeídos, trifluorometil cetonas, e ácidos borônicos).(ERBEL et al., 2006; YIN et al., 2006a, 2006b) Inibidores não peptídicos como  $\alpha$ -acetamidas,(STEUER et al., 2011), derivados à base de ftalazina,(BODENREIDER et al., 2009) acil-ciano-acrilamidas,(NITSCHKE; STEUER;

KLEIN, 2011) peptídeos híbridos,(NITSCHKE et al., 2012) derivados de benzo [d] isotiazol-3 (2H)-ona(TIEW et al., 2012), e outros tem sido relatados.(PARKINSON; PRYDE, 2010) (ARAVAPALLI et al., 2012)

Muito se tem discutido sobre as proteases de serina. Entretanto, existem algumas proteínas dessa classe para as quais temos poucas informações, mesmo sendo importantes devido ao impacto na saúde de milhões de pessoas, como é o caso da protease NS2B-NS3 do vírus da Dengue.

Logo, além das questões dinâmicas relacionadas a esse sistema, há muitos desafios a respeito da protease NS2B-NS3 do vírus da Dengue, a começar pela necessidade de um modelo estrutural ativo, e a elucidação do mecanismo da reação de quebra da ligação peptídica do substrato durante replicação viral.

Este trabalho pretende contribuir para cobrir esse espaço, estudando a protease do NS2B-NS3 do vírus da Dengue por métodos de química teórica e computacional.



## 2. OBJETIVO

---

### 2.1 Objetivos Gerais

O objetivo deste trabalho é o estudo teórico do mecanismo reacional da replicação viral da NS2B-NS3, elucidando cada etapa da reação para a quebra da ligação peptídica do substrato. Com a elucidação do mecanismo será possível usar a estrutura do estado de transição como molde na varredura de potenciais inibidores em um banco de substâncias.

#### 2.1.1 Objetivos Específicos

Para o estudo reacional é necessário um preparo que inclui a construção de um modelo para estrutura da NS2B-NS3 protease na forma ativa, e o estudo do comportamento dos sistemas através de dinâmica molecular da NS2B-NS3 na estrutura inativa (fechada), do modelo para estrutura ativa, e o modelo na presença de inibidor ou substrato, a fim de conhecer as peculiaridades de cada sistema e validar o modelo estrutural ativo da NS2B-NS3 protease do vírus da Dengue. As etapas estão divididas deste modo:

1. Dinâmica Molecular: DENV2a (2FOM) – estrutura inativa da NS2B-NS3 (aberta)
2. Criação do modelo estrutural: DENV2f (molde) – estrutura ativa da NS2B-NS3 (fechada)
3. Dinâmica Molecular: DENV2f
4. Dinâmica Molecular: DENV2f com inibidor (BZO-NLE-LYS-ARG-ARG-H).
5. Dinâmica Molecular: DENV2f com substrato (ALA-GLY-ARG-LYS-SER-ILE) e (CH<sub>3</sub>OC-VAL-PRO-ARG-ARG-NHCH<sub>3</sub>).
6. Validação do modelo estrutural DENV2f para estudo reacional.
7. Estudo reacional da NS2B-NS3 protease ativa (DENV2f) na quebra da ligação peptídica do substrato (NHCH<sub>3</sub>-ARG-ARG-PRO-VAL-CH<sub>3</sub>OC).
8. Busca por novos inibidores para a replicação viral da NS2B-NS3 protease do vírus da Dengue

A metodologia usada para alcançar esses objetivos será apresentada a seguir.

### 3. METODOLOGIA

---

Essencial para o sucesso de qualquer abordagem computacional no estudo de sistemas químicos é a qualidade do modelo matemático utilizado para calcular a energia do sistema como uma função da sua estrutura. Para sistemas químicos menores estudados na fase gás, abordagens da mecânica quântica (QM) são apropriados. Estes métodos, no entanto, são tipicamente limitados a sistemas de aproximadamente centenas de átomos ou menos.(JENSEN, 2007; YOUNG, 2004) Abordagens para o tratamento sistemas grandes têm sido desenvolvidas (LEACH, 2001). Sistemas de interesse em áreas como a bioquímica ou biofísica tipicamente envolvem macromoléculas que contêm 1000-5000 (KOLLMAN, 1987) ou mais átomos, além do seu ambiente de fase condensada. Isto pode levar a sistemas que chegam a conter dezenas de milhares de átomos. Além disso, a natureza dinâmica inerente desses sistemas e do ambiente químico exige a amostragem de um grande número de conformações, que podem ser geradas através de diferentes métodos de amostragem, envolvendo cálculos de energia para cada conformação. Assim, há necessidade de uma função de energia que permita a realização de um grande número de cálculos de energia em sistemas que contêm na ordem de  $>10^5$  átomos.

Funções de energia empíricas podem atender a tais demandas exigidas para estudos computacionais de sistemas bioquímicos e biofísicos. As equações matemáticas em funções de energia empíricas incluem termos relativamente simples para descrever as interações físicas que determinam a estrutura e propriedades dinâmicas de moléculas biológicas. Além disso, os campos de forças empíricas, em que a energia eletrônica é descrita como uma função paramétrica das coordenadas nucleares, e ajustam os parâmetros para dados experimentais ou dados computacionais de nível mais elevado, (JENSEN, 2007) usam um modelo atomístico no qual os átomos são as menores partículas do sistema, em vez de os elétrons e núcleos usados na mecânica quântica. Estas duas simplificações permitem a velocidade computacional necessária para executar o grande número de cálculos de energia em biomoléculas nos seus ambientes exigido e, mais importante, através da utilização de parâmetros devidamente otimizados nos modelos matemáticos, a precisão necessária pode ser alcançada (MACKERELL, 2004). O uso de funções de energia empíricas foi inicialmente aplicado a pequenas moléculas orgânicas (ALLINGER, 1977; LII; ALLINGER, 1989) e, nas últimas décadas, para os sistemas biológicos (CORNELL et al., 1995; WEINER et al., 1984).

### 3.1 SUPERFÍCIE DE ENERGIA POTENCIAL

Uma superfície de energia potencial (PES, do inglês *Potential Energy Surface*) descreve a energia potencial do sistema, especialmente um conjunto de átomos, em função de determinados parâmetros, normalmente, as posições dos átomos. Se existe um conjunto de átomos, a PES é uma superfície definida pela energia potencial de todos os arranjos atômicos possíveis desse conjunto de átomos. Assim, a PES tem dimensões de  $3N - 6$ , onde  $N$  é o número de átomos. A energia potencial do sistema será uma função das posições dos  $N$  átomos interagentes. Essa energia potencial pode ser descrita tanto por métodos clássicos, quanto por métodos quânticos de química teórica.

A PES é uma ferramenta conceitual para auxiliar a análise da geometria molecular e dinâmica de reação química. Os pontos estacionários têm significado físico, podem ser mínimos de energia que corresponde a espécie fisicamente estável, como também os pontos de sela que correspondem estados de transição.

### 3.2 CAMPOS DE FORÇAS

Quando se tem o interesse de estudar estrutura e interação em sistemas moleculares muito grandes, compreendendo centenas ou mesmo milhares de átomos, as técnicas de modelagem molecular mais sofisticadas exigem também, maior recurso computacional. Portanto, cálculos de mecânica molecular (campo de força), podem ser de grande utilidade na descrição desses sistemas. Neste tipo de modelagem, o aspecto estrutural é mais relevante e não se tem informação sobre as propriedades eletrônicas e/ou espectroscópicas, pois baseia-se num modelo matemático, onde o sistema molecular é descrito por um conjunto de esferas (o que corresponde aos átomos) que são mantidas juntas por meio de “molas” (correspondentes as ligações e ângulos), assim como interações de carga e van der Waals para as interações intermoleculares.

Deste modo, modelos atômicos de campos de forças são baseados em potenciais empíricos com uma forma funcional específica, representando as propriedades físicas e químicas dos sistemas de interesse. Parâmetros ajustáveis são escolhidos de tal modo que o potencial empírico represente um bom ajuste para as regiões relevantes da superfície de energia potencial, ou podem basear-se em dados experimentais. Um campo de forças típico, utilizado nas simulações de biosistemas, tem a forma mostrada na equação 1.

$$\begin{aligned}
U(r_1, \dots, r_N) = & \sum_{\text{ligação}} \frac{a_i}{2} (l_i - l_{i0})^2 + \sum_{\text{ângulo}} \frac{b_i}{2} (\theta_i - \theta_{i0})^2 \\
& + \sum_{\text{torção}} \frac{c_i}{2} [1 + (n\omega_i - \gamma_i)] \\
& + \sum_{\text{pares átomos}} 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{\text{pares átomos}} k \frac{q_i q_j}{r_{ij}}
\end{aligned} \tag{1}$$

Os três primeiros termos da soma estão relacionados a todas as ligações, ângulos e ângulos de torção definidas pela estrutura covalente do sistema, enquanto que os dois últimos termos somam os pares de átomos (ou locais ocupados por ponto de carga  $q_i$ ), separados por distâncias  $r_{ij} = |r_i - r_j|$  e não ligados quimicamente.

Fisicamente, os dois primeiros termos descrevem energias de deformações do comprimento de ligação  $l_i$  e ângulos  $\theta_i$  em relação aos seus respectivos valores de equilíbrio  $l_{i0}$  e  $\theta_{i0}$ . A forma harmônica desses termos (com constantes de força  $a_i$  e  $b_i$ ) garante a estrutura química correta, mas impede que aconteçam alterações químicas na modelagem, do tipo quebra e formação de ligações. O terceiro termo descreve rotações em torno da ligação química, que são caracterizadas por termos energéticos periódicos (com periodicidade determinada por  $n$  e alturas de barreiras de rotação definidos pelas  $c_i$ ). O quarto termo descreve forças interatômicas repulsivas e atrativas de van der Waals (dispersão) na forma do potencial de Lennard-Jones 12-6. Finalmente, o último termo é o potencial eletrostático de Coulomb.

O potencial deve ter uma forma simples e funcional para facilitar a avaliação das forças. Idealmente, os potenciais empíricos também devem ser transferíveis e aplicáveis a muitos sistemas e em diferentes condições. A abordagem padrão combina dados experimentais e os resultados de cálculos *ab initio* em sistemas modelo que podem ser utilizados como blocos de construção para uma macromolécula. Por exemplo, parâmetros estruturais, tais como o comprimento de ligação de equilíbrio pode ser feita a partir de estudos de cristalografia e a estimativa inicial das cargas parciais atômicas pode basear-se em densidades eletrônicas de cálculos *ab initio*. Os campos de força disponíveis, como o AMBER (Cornell et al., 1995), CHARMM (Brooks et al., 1983) ou GROMOS (van Gunsteren et al., 1996), têm-se revelado suficientemente precisos em termos de propriedades cinéticas e termodinâmicas obtidas por meio de simulações de dinâmica molecular para proteínas ou ácidos nucleicos.

### 3.3 DINÂMICA MOLECULAR

A simulação de dinâmica molecular é uma ferramenta importante para a compreensão da base física da estrutura e função de macromoléculas biológicas. A visão inicial de proteínas como estruturas relativamente rígidas tem sido substituída por um modelo dinâmico em que os movimentos internos e mudanças conformacionais resultantes desempenham um papel essencial na sua função.(KARPLUS; KURIYAN, 2005; KARPLUS; MCCAMMON, 2002)

Simulações dinâmicas monitoram processos dependentes do tempo em sistemas moleculares, bem como suas propriedades estruturais, dinâmicas e termodinâmicas, através da resolução numérica de uma equação de movimento, que é a formulação das regras que governam o movimento executado pela molécula. Ou seja, a amostragem através de dinâmica molecular (MD) fornece informações sobre a dependência do tempo e magnitude das flutuações em ambas as posições e velocidades (JENSEN, 2007; LEACH, 2001)

Simulações por dinâmica molecular de macromoléculas biológicas têm fornecido muitas informações sobre os movimentos internos desses sistemas, desde a primeira proteína estudada a mais de 25 anos atrás.(MCCAMMON; GELIN; KARPLUS, 1977) Com os avanços contínuos na metodologia e a velocidade dos computadores, que ainda está dobrando a cada 18 meses ou menos, os estudos de dinâmica molecular estão sendo estendidos para sistemas maiores, com maiores mudanças conformacionais e as escalas de tempo cada vez mais longas. Isso torna possível a investigação de movimentos que têm particulares implicações funcionais e a obtenção de informações que não são acessíveis a partir de experimento. Os resultados disponíveis hoje deixam claro que as aplicações de dinâmica molecular irão desempenhar um papel ainda mais importante para a nossa compreensão da biologia no futuro. (FRENKEL; SMIT, 2001)

#### 3.3.1 Formulação

Nas simulações de dinâmica molecular a evolução temporal de um conjunto de partículas que interagem é acompanhada através da solução das equações do movimento de Newton, equação 2, onde  $r_i(t) = (x_i(t), y_i(t), z_i(t))$  é o vetor posição da partícula  $i$  e  $F_i$  é a força que age sobre a partícula  $i$  no tempo  $t$  e,  $m_i$  é a massa da partícula.

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \quad [2]$$

"Partículas" geralmente correspondem a átomos, embora possam representar quaisquer entidades distintas (por exemplo, grupos químicos específicos ou cargas pontuais) que podem ser convenientemente descritos em termos de uma determinada lei de interação (JENSEN, 2007; LEACH, 2001). Para integrar as equações diferenciais das forças instantâneas que atuam sobre as partículas, as suas posições iniciais e velocidades precisam ser especificadas, pois a equação de movimento de Newton é de segunda ordem em relação a  $r$  (posição). Devido à natureza de muitos corpos do problema as equações de movimento são discretizadas e resolvidas numericamente (JENSEN, 2007; LEACH, 2001). As trajetórias da dinâmica molecular são definidas por ambos vetores das posições e das velocidades, e assim descrevem a evolução temporal do sistema no espaço de fase. Deste modo, as posições e velocidades são propagadas com um intervalo de tempo finito utilizando integradores numéricos, por exemplo, o algoritmo de Verlet (VERLET, 1967). A posição instantânea de cada partícula no espaço é definida por  $r_i(t)$ , enquanto que as velocidades instantâneas  $v_i(t)$  determinam a energia cinética e a temperatura do sistema em um determinado momento. Como as partículas se movem, suas trajetórias podem ser exibidas e analisadas, proporcionando propriedades médias. Os eventos dinâmicos que podem influenciar as propriedades funcionais do sistema podem ser diretamente relacionados ao nível atômico, tornando dinâmica molecular especialmente valiosa em biologia molecular. (FIELD, 2007)

### 3.3.1.1 Integração numérica das equações de movimento

O objetivo da integração numérica das equações do movimento de Newton é encontrar uma expressão que defina as posições  $r_i(t + \Delta t)$  no momento  $t + \Delta t$  em termos das posições já conhecidas no tempo  $t$ . As coordenadas e velocidades moleculares no tempo  $t + \Delta t$  são obtidos a partir das coordenadas moleculares em um momento anterior  $t$ . A posição no tempo  $t + \Delta t$  pode ser representada como uma expansão de Taylor para as posições  $r_i(t)$ , que está escrita na equação [3]:

$$r_i(t + \Delta t) \cong 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i} \Delta t^2 \quad [3]$$

Na equação [3] pode se incluir termos até quarta ordem em  $\Delta t$ .

As trajetórias exatas correspondem ao limite de um passo de integração ( $\Delta t$ ) infinitamente pequeno. É, no entanto, desejável utilizar passos de tempo maiores para amostrar trajetórias mais longas. Na prática  $\Delta t$  é determinado por rápidos movimentos no

sistema. Ligações envolvendo átomos leves (por exemplo, a ligação O-H) vibram com períodos de alguns femtosegundos, implicando que  $\Delta t$  deve estar numa escala de sub-femtosegundos para garantir a estabilidade da integração. Embora as vibrações mais rápidas e não cruciais podem ser eliminadas mediante a imposição de restrições sobre o comprimento de ligação no algoritmo de integração, um passo de tempo de mais que 2,0 fs raramente pode ser conseguida em simulações de biomoléculas.(FIELD, 2007)

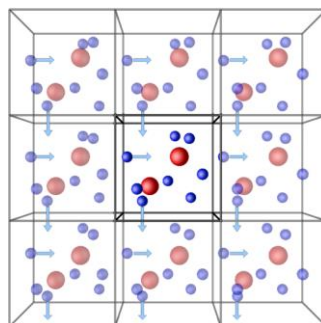
### 3.3.1.2 Cálculos das Forças e das Interações de Longo Alcance

A atualização passo a passo das posições e velocidades no processo de integração numérica requer que as forças que atuam sobre os átomos (que mudam suas posições relativas a cada período de tempo) sejam recalculadas a cada passo. Campos de força de biomoléculas incluem interações de longo alcance eletrostáticas e de dispersão. Assim, um somatório de ordem  $N^2$  precisa ser realizado em uma implementação simples para abranger todos os pares não ligados. Portanto, o cálculo repetido das forças define a complexidade geral do algoritmo de MD pois, com o aumento no tamanho do sistema, e pela inclusão de moléculas de água explícita, o cálculo de todas as interações possíveis, levaria mais de 90% do tempo computacional(LEACH, 2001), muitas técnicas têm sido desenvolvidas para lidar com o problema das forças de longo alcance.

Uma solução aquosa é o ambiente típico para macromolécula biológica que deve ser considerada numa simulação realista, de preferência usando uma representação atomística do solvente. Devido à memória computacional limitada e, também, para acelerar os cálculos, apenas uma amostra finita do sistema estendido (infinito) pode ser representada de forma explícita em um modelo no computador. O tratamento das forças de longo alcance está relacionado com a escolha das condições de fronteira impostas a um sistema para lidar com os efeitos de tamanho finito e superfície. Duas abordagens comuns são baseadas em condições periódicas e o método de Ewald ou em condições de contorno esféricas e no método de campo de reação. Algoritmos recentes, por exemplo, o método chamado de *particle meshed Ewald*, permitem o cálculo eficiente das interações de longo alcance, sem recorrer à aproximação de corte abrupto, em que as contribuições de partículas separadas por uma distância maior do que certo limite são negligenciadas.(FRENKEL; SMIT, 2001)

### 3.3.1.3 Breve introdução sobre condições periódicas de contorno

Num sistema realístico teríamos que atingir o limite termodinâmico de  $10^{23}$  moléculas. Entretanto, isso não ainda não é viável, e o número  $N$  de moléculas que constituem o sistema não pode ser maior que algumas centenas. Porém esse problema é contornado, usando-se as condições de contorno periódicas (Figura 20), em que se define uma célula de referência com  $N$  moléculas e, esta célula, cujas paredes são fictícias, é então envolvida por um número infinito de réplicas nas três dimensões. Nessas condições de contorno, caso uma molécula, sob o efeito das forças que atuam sobre os átomos, escape da célula de referência, a saída será compensada pelo movimento de outra molécula proveniente da célula imagem oposta. Utilizando esse método, as moléculas do sistema (célula principal) sentem um ambiente equivalente ao de dentro do volume e não mais da superfície. Assim o número de partículas dentro da caixa central se mantém constante. Algo que pode parecer estranho é a interação da molécula com ela mesma em outra célula, entretanto, isso não acontece se for introduzido um “raio de corte” menor que o comprimento da célula, garantindo a não repetição das interações. Importante dizer que só a coordenada e parâmetros dos átomos da célula principal são considerados na configuração do sistema.(FRENKEL; SMIT, 2001)



**Figura 20:** Condições de contorno periódicas. Quando uma partícula se move para fora da caixa (célula) de simulação, uma partícula se move na imagem para substituí-lo. Ao calcular as interações das partículas dentro do raio de corte, os vizinhos reais e imagem estão incluídos.

## 3.4 MECÂNICA ESTATÍSTICA

De acordo com a mecânica estatística, quantidades físicas são representadas por médias sobre os estados microscópicos (configurações) do sistema, distribuídos de acordo com um determinado conjunto estatístico(JENSEN, 2007; LEACH, 2001). Um exemplo importante é o *ensemble* micro canônico, em que apenas os diferentes estados que correspondem a uma energia  $E$  específica tem probabilidade não nula de ocorrência. Outro



exemplo é o *ensemble* canônico, em que a temperatura  $T$  é constante, enquanto que a energia pode ser trocada com o ambiente (banho térmico) e a distribuição de estados é dado pela função de Boltzmann.

A Dinâmica newtoniana implica na conservação da energia(LEACH, 2001), e trajetórias de MD fornecem um conjunto de configurações distribuídas de acordo com o *ensemble* micro canônico. Generalizações de MD podem ser usadas para provar outros ensembles estatísticos. Por exemplo, através da introdução de um acoplamento com o banho térmico, podemos obter um conjunto de trajetórias que representam o *ensemble* canônico. Verificou-se que a energia não é estritamente conservada(JENSEN, 2007), mesmo na simulação micro canônica, por causa dos erros que se acumulam na integração numérica (estimar a variação de energia é um bom teste para a qualidade das simulações de DM). Entretanto, o reescalonamento das velocidades pode ser usado para ajustar a energia.

Observada sob o ponto de vista da mecânica estatística, MD é apenas um método de amostragem conformacional que produz propriedades estruturais e termodinâmicas médias. A hipótese da ergodicidade garante que todos os microestados acessíveis são igualmente prováveis, ao longo de um período de tempo prolongado. Ou seja, a média de uma propriedade sobre o tempo é igual a média sobre o conjunto (*ensemble*)(JENSEN, 2007). Na prática, uma dada propriedade termodinâmica será obtida computando-se médias sobre as  $M$  configurações do sistema definidas pelas trajetórias de dinâmica molecular, independentemente do tipo de *ensemble* utilizada:

$$\langle A \rangle = \frac{1}{M} \sum_{t=1}^M A_t(r^N, p^N) \quad [4]$$

Portanto, uma quantidade física pode ser medida por meio de simulação MD, tendo uma média aritmética de valores instantâneos ao longo das trajetórias. No limite de tempo de simulação infinito tais médias convergem para o valor real das propriedades termodinâmicas:

$$\langle A \rangle = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{t=1}^M A_t(r^N, p^N) \quad [5]$$

### 3.5 HOMOLOGIA

O objetivo da modelagem comparativa estrutural de proteínas ou homologia (JOHN; SALI, 2003; SALI; BLUNDELL, 1993) é de construir um modelo tridimensional (3D) de uma proteína de estrutura desconhecida (o alvo), com base em uma ou mais proteínas relacionadas de estrutura conhecida (os moldes). As condições necessárias para a obtenção de um modelo útil são que a semelhança entre a sequência do alvo e das estruturas de modelo seja grande o suficiente, e que o alinhamento correto entre elas possa ser construído (JOHN; SALI, 2003). Esta abordagem para estruturar a predição é possível porque uma pequena alteração na sequência de proteína resulta normalmente numa pequena mudança na sua estrutura 3D (SALI; BLUNDELL, 1993). Modelagem comparativa da estrutura de proteínas continua a ser o método de previsão mais preciso (ERAMIAN et al., 2008; JOHN; SALI, 2003).

Todos os métodos atuais de modelagem comparativa consistem em quatro passos sequenciais. O primeiro passo é identificar as proteínas com estruturas 3D conhecidas que estão relacionados com a sequência alvo. A segunda etapa é alinhá-los com a sequência alvo e escolher aquelas estruturas conhecidas que serão usados como modelos. O terceiro passo é construir o modelo para a sequência alvo dado o seu alinhamento com as estruturas de modelo. No quarto passo, o modelo é avaliado utilizando uma variedade de critérios. Se necessário, a seleção do modelo, alinhamento, e modelo de construção são repetidas até que um modelo satisfatório seja obtido. A principal diferença entre os métodos de modelagem comparativa está na forma como o modelo 3D é calculado a partir de um determinado alinhamento (terceira etapa). Para cada uma das etapas do processo de modelagem, existem programas e servidores disponíveis, por exemplo, Modeller (JOHN; SALI, 2003), Swiss-model (ARNOLD et al., 2006) e JACKAL (PETREY et al., 2003).

### 3.6 ENERGIA LIVRE

A energia livre é frequentemente considerada uma das quantidades termodinâmicas mais importantes, usualmente expressa como uma função de Helmholtz,  $A$ , ou pela função de Gibbs,  $G$ . A energia livre de Helmholtz é apropriadamente usada para sistemas que o número de partículas, volume e temperatura são constantes ( $NVT$ ), enquanto que a função de Gibbs é adequadamente usada quando o sistema apresenta número de partículas, pressão e temperatura constantes ( $NPT$ ). A maioria dos experimentos são realizados sob as condições

de pressão e temperatura constante, onde a função de Gibbs é função adequada para quantificação da energia livre.

A função de partição canônica  $Q$  de um sistema pode ser calculada por uma integral que inclui todo o espaço de fase, isto é, o espaço de configuração e de momento. Se a energia potencial,  $E$ , é independente do momento, a integral do mesmo é uma constante multiplicativa de  $Q$ , que pode ser ignorada. Então,  $Q$  é obtida por:

$$Q = \int \exp[-\beta E(r)] dr^N \quad [6]$$

Com  $\beta = 1/k_B T$ ,  $k_B$  é a constante de Boltzmann,  $T$  a temperatura absoluta, e  $N$  o número de graus de liberdade do sistema.

A energia livre de Helmholtz  $A$  é relacionada com  $Q$  por  $A = -1/\beta \ln Q$ . Se a pressão for mantida constante, então a energia livre de Gibbs,  $G$ , é obtida. Em fase condensada, que é relevante para a maioria das aplicações, os sistemas são dificilmente compressíveis, logo  $\Delta A$  e  $\Delta G$  são numericamente similares.

Entretanto, existe uma dificuldade para obter a energia livre em sistemas líquidos ou macromoléculas flexíveis que apresentam muitas configurações que são separadas por pequenas barreiras de energia. Quantidades associadas como entropia e potencial químico são também dificilmente calculados. A energia livre não pode ser determinada com acurácia por simulações padrão de Monte Carlo e dinâmica molecular, pois essas simulações não fazem amostragem adequada das regiões de fase que tem contribuições importantes para a energia livre. Uma maneira de contornar esse problema é usando técnicas de amostragem do tipo amostragem guarda-chuva (*Umbrella Sampling*).

### 3.6.1 Amostragem guarda-chuva (*Umbrella Sampling*)

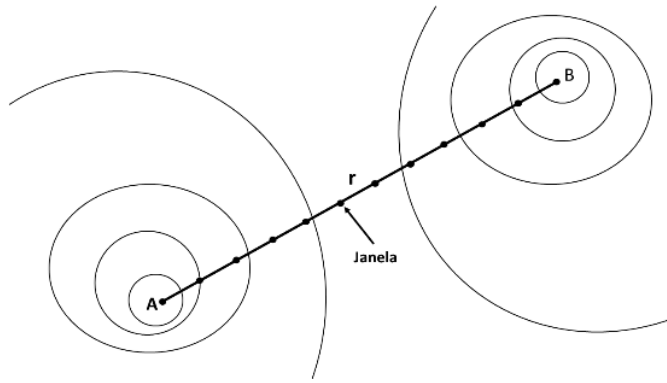
Essa técnica de amostragem modifica a função potencial para que os estados não favoráveis sejam amostrados suficientemente. O método pode ser usado com ambos os métodos de simulação Monte Carlo e dinâmica molecular. (TORRIE; VALLEAU, 1974, 1977) A modificação da função potencial pode ser escrita como uma perturbação:

$$\mathcal{V}'(r^N) = \mathcal{V}(r^N) + \mathcal{W}(r^N) \quad [7]$$

Onde  $\mathcal{W}(r^N)$  é uma função peso, que frequentemente tem uma forma quadrática (harmônica):

$$\mathcal{W}(r^N) = k_w(r^N - r_0^N)^2 \quad [8]$$

Onde  $k_w$  é a constante de força da função; quanto maior essa constante, maior o valor do potencial. Para configurações que estão distantes do estado de equilíbrio  $r_0^N$  a função terá um peso maior, então uma simulação que usa essa função de energia modificada ( $\mathcal{V}'(r^N)$ ) levará a um caminho relevante na direção de configurações longe da distância de equilíbrio,  $r_0^N$ . Por exemplo, se aplicada a um sistema para garantir uma boa amostragem ao longo de uma coordenada de reação (Figura 21). Isso pode ser feito em uma única simulação ou em diferentes simulações, chamadas de janelas (Figura 21), em que as distribuições em cada janela se sobrepõem. O potencial de restrição,  $\mathcal{W}_i(r^N)$ , da janela  $i$  é um termo de energia adicional, que só depende da coordenada de reação.



**Figura 21:** Separação da coordenada de reação entre dois estados em janelas distintas (linha entre A e B), aqui representados por dois mínimos da superfície de energia potencial. O sistema é principalmente amostrado na região perpendicular a coordenada de reação de cada janela.

O resultado dessa distribuição de configuração será não Boltzmann. Entretanto, as médias correspondentes a distribuição do tipo Boltzmann podem ser extraídas da distribuição não Boltzmann, pelo método desenvolvido por Torrie e Valleau em 1977.(TORRIE; VALLEAU, 1977) O resultado é:

$$\langle A \rangle = \frac{\langle A(r^N) \exp \left[ + \frac{W(r^N)}{k_B T} \right] \rangle_w}{\langle \exp \left[ + \frac{W(r^N)}{k_B T} \right] \rangle_w} \quad [9]$$

Onde,  $A$  é a energia livre de Helmholtz, o  $W$  subscrito na equação (9) é para indicar que as médias são baseadas na probabilidade,  $P_w(r^N)$ , que por sua vez é determinado pela função de energia modificada,  $\mathcal{V}'(r^N)$ .(KÄSTNER, 2011)

Quando o sistema está sobre restrição por uma função peso (*bias*),  $\mathcal{W}_i(r^N)$ , a distribuição de probabilidade do sistema com restrição (*biased*),  $P_i(r^N)$ , para cada janela é (KÄSTNER, 2011):

$$P_i(r^N) = e^{-\mathcal{W}_i(r^N)/k_B T} P_i^u(r^N) \langle e^{-\mathcal{W}_i(r^N)/k_B T} \rangle^{-1} \quad [10]$$

Onde  $P_i^u(r^N)$  é a distribuição de probabilidade sem restrição (*unbiased*), e  $\mathcal{W}_i(r^N)$  é a função peso que se refere a restrição (*bias*)(KÄSTNER, 2011).

Assim a energia livre de Helmholtz da janela  $i$  sem a restrição (*unbiased*) dada por:

$$A_i^u(r^N) = -\frac{1}{\beta} \ln P_i(r^N) - \mathcal{W}_i(r^N) - F_i \quad [11]$$

Onde  $A_i^u$  é o potencial de força média referente a simulação sem a perturbação, assim como  $P_i$  é a probabilidade dessas configurações com a perturbação (*bias*), e  $\mathcal{W}_i$  é a função peso que se refere a perturbação do sistema. E os valores de  $F_i$  correspondem a energia livre introduzida devido a introdução da função peso. Apesar de ser independente da variável  $r^N$ , o  $F_i$  depende do perfil global de energia livre (equação 12)(KÄSTNER, 2011):

$$F_i = \frac{1}{\beta} \int \exp\{-\beta[A^u(r^N) + \mathcal{W}_i(r^N)]\} dr^N \quad [12]$$

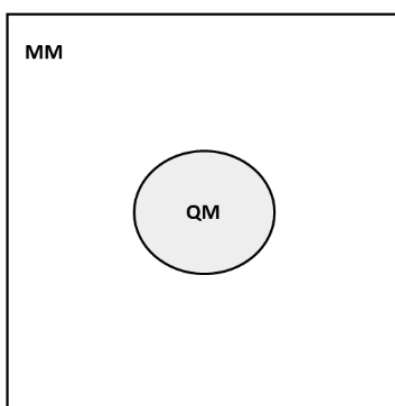
O método é exato, a aproximação é para o tempo finito. Para combinar as janelas, pode-se calcular o  $F_i$ , usando o método do histograma ponderado (WHAM).

As equações do WHAM apresentam duas quantidades não conhecidas,  $F_i$  e  $P_i^u$ . O programa WHAM consiste em ajustar as curvas  $A_i^u$  pela equação 11, calculando a média para as diferentes janelas, até a convergência ser atingida(HUB; DE GROOT; VAN DER SPOEL, 2010).

### 3.7 MÉTODO HÍBRIDO QM/MM

Os métodos híbridos que combinam cálculos de química quântica com mecânica molecular (QM/MM, Figura 22) tem sido usados para modelagem de processos reacionais em sistemas biomoleculares (WARSHEL; LEVITT, 1976a; WEINER; SINGH; KOLLMAN, 1985). Os métodos de química quântica são necessários para os processos de reações químicas, transferência de carga e excitação eletrônica. Entretanto, os métodos QM estão limitados a sistemas de até algumas centenas de átomos. No entanto, o tamanho e a complexidade conformacional de biopolímeros exige métodos capazes de tratar um número de átomos na ordem de  $10^6$ , assim como permitir simulações em escalas de tempo de dezenas

de nanosegundos. Isto é possível com alta eficiência pelo método de mecânica molecular baseado em campos de força. A abordagem lógica para tratar sistemas biomoleculares é combinar as duas técnicas, ou seja, usando o método QM para a região quimicamente ativa (por exemplo, substrato e cofator numa reação enzimática), e um tratamento por método de mecânica molecular (MM) para tratar a vizinhança (proteína e solvente). Eles permitem a modelagem de sistemas biomoleculares reativos com esforço computacional plausível, e proporcionando a precisão necessária.



**Figura 22:** Ilustração da divisão do sistema onde será tratado pelo método híbrido QM/MM. A maior parte será tratada por mecânica molecular e a menor parte calculada por química quântica.

Assim, a partir dessa perspectiva de modelagem, pode-se considerar o tratamento do sistema como descrito resumidamente na Figura 22. Dentro de uma região limitada, faz-se uso das ferramentas da mecânica quântica para modelar com precisão um problema de estrutura eletrônica, enquanto que no entorno dessa região a representação explícita do resto do sistema é importante, mas o nível de modelo aplicado pode ser reduzido em sua complexidade, pois essa região é mais externa ao sistema onde ocorre o processo, aplicando a essa vizinhança o método MM. O Hamiltoniano completo para o sistema deve ser um híbrido que inclua a parte QM e MM, assim como a região de fronteira QM/MM. De maneira simples:

$$\hat{H} = \hat{H}_{MM} + \hat{H}_{QM} + \hat{H}_{QM/MM} \quad [13]$$

Onde  $\hat{H}_{QM}$  representa o operador da energia total de as interações entre as partículas quânticas,  $\hat{H}_{MM}$  é o operador da energia total de interação de todas as partículas clássica, e  $\hat{H}_{QM/MM}$  representa a energia de todas as interações entre as partículas quânticas e as partículas clássicas.

Uma das principais diferenças entre os vários métodos QM/MM é o tipo de tratamento para a região de fronteira QM/MM, ou seja, na forma como são tratadas as interações (se existirem) entre os átomos QM e MM. Um método mais simples de tratar a ligação entre a parte QM e MM, envolve uma interação direta da parte QM e o ambiente MM. Em que as interações entre as regiões QM e MM são tratadas puramente classicamente, isto é, o sistema QM é representado por cargas pontuais (MM) na sua interação com o ambiente MM. Para o caso em que não existem ligações covalentes entre as regiões QM e MM, este termo é o somatório da contribuição eletrostática e do termo Lenard-Jones.

A partição do sistema é caracterizada por um Hamiltoniano efetivo, que opera sobre a função de onda do sistema, que é dependente das posições dos núcleos QM e MM, para se obter a energia efetiva do sistema,  $E_{eff}$  (WALKER; CROWLEY; CASE, 2008):

$$E_{eff} = \langle \Psi | \hat{H}_{QM} + \hat{H}_{QM/MM} | \Psi \rangle + E_{MM} \quad [14]$$

Na implementação QM/MM no *sander* (WALKER; CROWLEY; CASE, 2008), programa do pacote de programas Amber, a  $E_{MM}$  (energia da parte MM) é calculada classicamente a partir das posições atômicas da região MM, utilizando a equação e os parâmetros campo de força Amber, enquanto a região QM é avaliada através do método QM escolhido (WALKER; CROWLEY; CASE, 2008).

O termo de interação  $\hat{H}_{QM/MM}$  é mais complicado. Por padrão, o *sander* usa um esquema de incorporação eletrostática (também referida como esquema aditivo), em que a interação das cargas pontuais MM com os elétrons da região QM, bem como, a interação entre as cargas pontuais MM e os núcleos da região QM (núcleos atômicos para os métodos semiempíricos) é explicitamente calculado. Deste modo, a região MM polariza a densidade de eletrônica da região QM. Para o caso onde não existem ligações covalentes entre os átomos da região QM e região MM, o Hamiltoniano de interação é a soma de um termo eletrostático e um termo Lennard-Jones (VDW). (WALKER; CROWLEY; CASE, 2008)

Há uma variedade de abordagens para lidar com o problema dos limites da região QM/MM nas ligações covalentes. (AMARA et al., 2000; GAO et al., 1998; GARCIA-VILOCA; GAO, 2004) O mais comum é o uso da metodologia átomo de ligação (*link atom*) (WALKER; CROWLEY; CASE, 2008) para satisfazer a valência da região QM, deste modo, a ligação covalente do limite da região QM é saturado por um átomo adicional QM, mais comumente um átomo de hidrogênio, colocados ao longo da ligação entre os limites da

fronteira QM/MM. Compartilhando a mesma lista de pares para as interações QM/MM como átomos reais QM. Os átomos de ligação estão presentes apenas no cálculo QM, e são invisíveis para os átomos MM. Em princípio, cada *link atom* apresenta três graus de liberdade adicionais para o sistema. No entanto, na prática, o *link atom* é colocado numa posição fixa ao longo da ligação que separa o átomo QM do MM na ligação covalente, em cada passo da simulação, de modo que estes graus de liberdade adicionais são removidos de novo. Em cada passo, a força que atua sobre o *link atom* está distribuída ao longo dos átomos QM e MM da ligação.

O método QM que foi empregado no estudo será brevemente comentado na próxima secção.

### 3.7.1 MÉTODO SEMI-EMPÍRICO

Os métodos semi-empíricos são baseados em aproximações das equações de Hartree-Fock-Roothaan (HFR)(ROOTHAAN, 1951), introduzindo nestas equações funções com parâmetros empíricos. Esses parâmetros são ajustados a certos conjuntos de dados experimentais para melhorar a qualidade dos resultados. O custo real de cálculo HF deve-se ao cálculo das integrais de dois-elétrons que são simplificadas e/ou desprezadas nos métodos semi-empíricos. A maioria dos métodos semi-empíricos possui três aproximações básicas.

#### 3.7.1.1 Eliminação dos elétrons internos

Elétrons internos ou de caroço não contribuem significativamente para atividade química, o que torna possível remover as funções que descrevem estes elétrons da matriz de Fock. Normalmente, o caroço inteiro (o núcleo e os elétrons de caroço) de um átomo é substituído por uma função parametrizada. Isto reduz a complexidade do cálculo drasticamente sem muito impacto na precisão do cálculo. No caso dos métodos semi-empíricos, em geral, a simplificação é ainda maior, pois simplesmente subtrai-se o número de elétrons de caroço da carga nuclear.(THIEL, 2005)

#### 3.7.1.2 O uso do conjunto de funções de base mínimo

Assim como no método HFR, são introduzidas funções de base para descrever os elétrons valência, entretanto apenas um número mínimo de funções base é utilizado. Em geral, todos os métodos semi-empíricos utilizam funções do tipo Slater (STO – *Slater Type*



*Orbitals*)(SLATER, 1968) em seu conjunto de base mínima, pois as integrais de dois-elétrons e dois-centros têm soluções analíticas. Para integrais com mais de dois-centros não soluções analíticas, mas os métodos semi-empíricos desprezam estas integrais, e, portanto, não causam problemas computacionais. A utilização de funções STO é relevante, pois elas apresentam as mesmas condições assintóticas e de cúspide que as funções de onda exatas.(THIEL, 2005)

#### 3.7.1.3 A redução de integrais de dois elétrons

Dada a grande demanda computacional do cálculo de integrais,  $M^4$ , uma das principais aproximações dos métodos semi-empíricos está na redução do número de integrais de dois elétrons calculadas. A aproximação de maior importância com relação à redução do custo computacional, para o cálculo das integrais de dois-elétrons, é conhecida como aproximação ZDO (*Zero Differential Overlap*)(THIEL, 2005) que anula todas as integrais que contém produtos de funções de base diferentes no seu integrando. O nível de aplicação da aproximação ZDO leva aos três métodos básicos semi-empíricos em ordem crescente de sofisticação e demanda computacional: CNDO, INDO e NDDO.(THIEL, 2005) Destas aproximações, a CNDO praticamente não é mais utilizada e a INDO está presente em alguns poucos métodos semi-empíricos como ZINDO (ou INDO/S-CIS)(ZERNER, 1991) e SINDO(NANDA; JUG, 1980). A maioria dos métodos semi-empíricos é baseada nas equações e aproximações NDDO e a família de métodos desenvolvida por Dewar e colaboradores(DEWAR; THIEL, 1977), a saber, MNDO(DEWAR; THIEL, 1977), AM1(DEWAR et al., 1985), PM3(STEWART, 1989), PDDG-PM3(REPASKY; CHANDRASEKHAR; JORGENSEN, 2002) e, mais recentemente, RM1(ROCHA et al., 2006).

#### 3.7.1.4 Interações intermoleculares

Sabe-se que os métodos semi-empíricos não descrevem tão bem as interações intermoleculares como ligações de hidrogênio. Os modelos mais populares têm sido derivados da aproximação MNDO, que não consegue descrever complexos formados por ligações de hidrogênio, devido à sua tendência a superestimar a repulsão entre os átomos que estão separados por distâncias próximas à soma dos seus raios de van der Waals. Muito tem se feito para corrigir a deficiência da aproximação MNDO, por exemplo, AM1, PM3, PDDG-PM3. Todos esses propõem uma modificação no termo de repulsão núcleo-núcleo. Nesses casos, tem aproximações e equações semelhantes ao método MNDO com relação a integrais de dois elétrons, mas usa uma expressão modificada para repulsão caroço-caroço com a inclusão de

funções Gaussianas para aprimorar as energias de interação intermoleculares, assim, proporcionando uma contribuição extra para estabilizar a distância de equilíbrio dessa interação. Essas funções são parametrizadas para reproduzir valores experimentais de calores de formação para pequenos complexos formados por ligações de hidrogênio.

#### 3.7.1.5 Estudo com proteases

Em estudo de Schroder *et al.* (SCHROEDER; DAGGETT; KOLLMAN, 1991) afim de determinar o melhor método semi-empírico (AM1 e PM3) para representam compostos modelo pertinentes à catalise da protease de serina pela hidrólise de amidas e ésteres. Tendo como resultado que o PM3 obteve melhores resultados para as geometrias de ligação de hidrogênio do que o AM1, obtendo-se geometrias que foram muito semelhantes aos encontrados na estrutura cristalina da tripsina.(SCHROEDER; DAGGETT; KOLLMAN, 1991) O PM3 foi o melhor semi-empírico para sistemas de ligação de hidrogênio, onde é claramente importante e onde o processo de interesse está a ocorrer no local ativo da enzima. A comparação com um maior nível de cálculo *ab initio* mostrara que as energias relativas a catalise da reação de transferência de prótons estão em uma escala comparável.(SCHROEDER; DAGGETT; KOLLMAN, 1991)

Além disso, valores de barreiras de energia para transferência de prótons em sistemas contendo ligações de hidrogênio, em geral, são sobrestimados por PM3.(MORPURGO; BOSSA; MORPURGO, 1998)

A modificação *Pairwise Distance Directed Gaussian* (PDDG) fornece uma boa descrição da van der Waals atração entre átomos, deste modo, e o modelo PDDG-PM3 mostra ser apropriado para cálculos de complexos intermoleculares. PDDG-PM3 produz as energias ligação de hidrogênio mais confiáveis para um conjunto de geometrias de dímeros de água e melhor energia para dímero de benzeno.(REPASKY; CHANDRASEKHAR; JORGENSEN, 2002; TUBERT-BROHMAN et al., 2004).Uma combinação desses métodos foi utilizada para estudar os sistemas de interesse. No próximo capítulo veremos os métodos sendo aplicados aos sistemas.

#### 4. PROCEDIMENTO COMPUTACIONAL

---

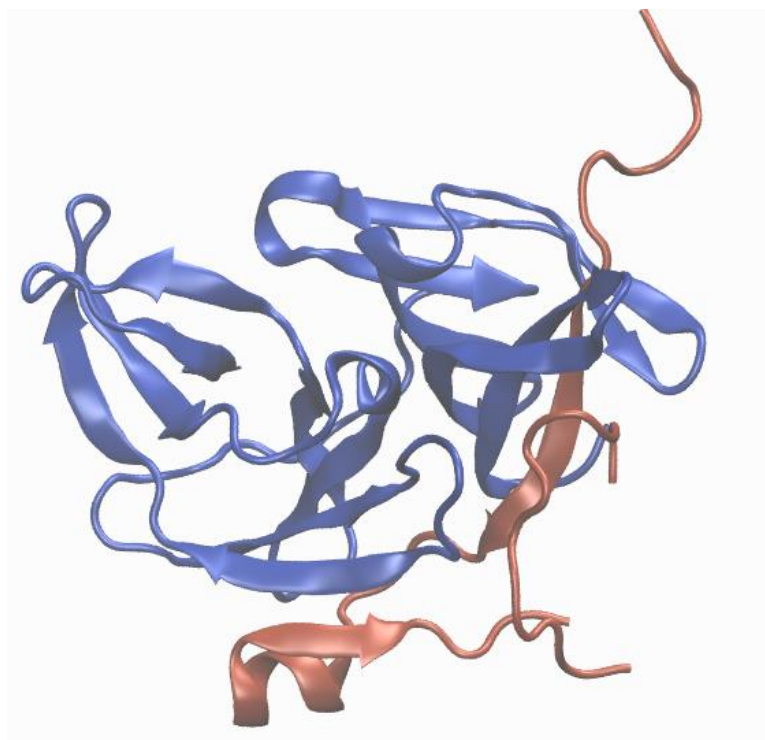
Para simular um sistema de partículas através de dinâmica molecular são necessárias algumas etapas, de preparação do sistema, execução dos cálculos, controle e análise dos resultados. Nesta seção será apresentada a metodologia utilizada em cada etapa do processo.

Em resumo:

- Gerar a configuração inicial das moléculas
- Ensemble e controles de simulação
- Análise das configurações (trajetórias)

##### 4.1 ESCOLHA DA ESTRUTURA INICIAL

A estrutura da proteína NS2B-NS3 protease na sua forma inativa foi obtida através do banco de dados de proteínas, *Protein Data Bank* – PDB. Na busca procuramos a NS2B-NS3 do vírus da Dengue para o sorotipo tipo 2, sorotipo de grande circulação no Brasil. Na época (2010), a estrutura com a melhor resolução (1,50 Å) foi 2FOM (ID, identificador), obtido por técnica de difração de Raio-X (Figura 23).(ERBEL et al., 2006)

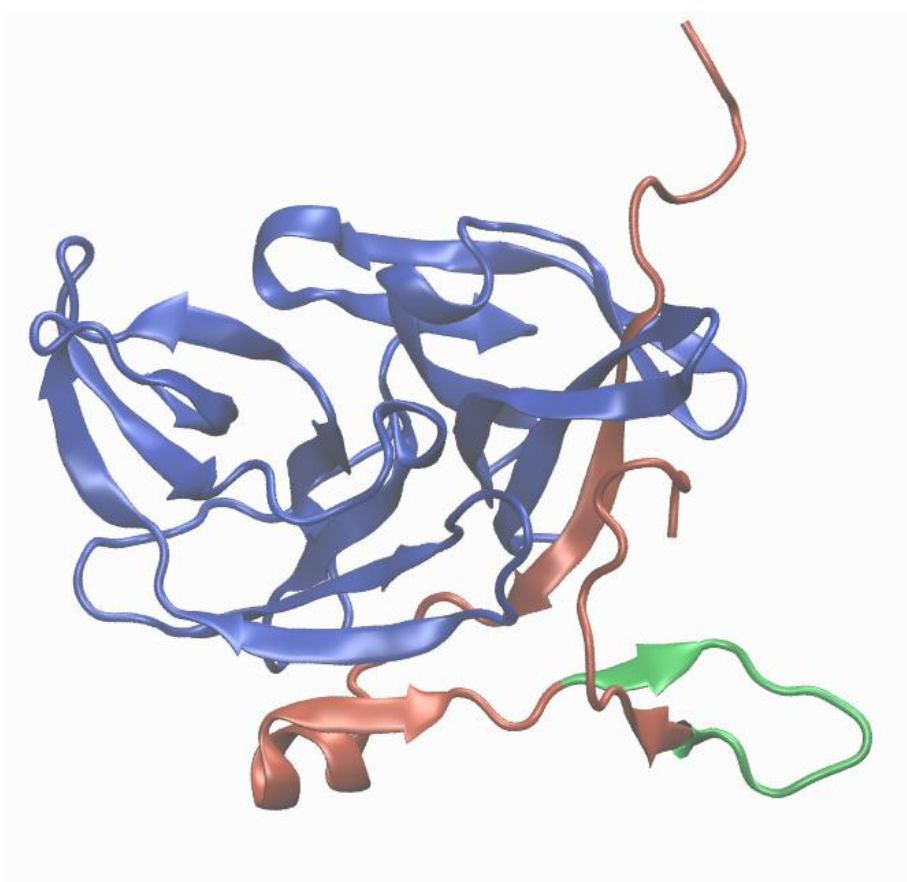


**Figura 23:** Estrutura da 2FOM (DENV2a protease). Em azul a NS3, e em vermelho a NS2B.

#### 4.1.1 Complementação de regiões da proteína

Dependendo da técnica de elucidação de estrutura, nem todas as posições atômicas são solucionadas, podendo haver regiões sem definição estrutural, até mesmo por características próprias da proteína. Entretanto, é possível utilizar-se programas que comparam estrutura de proteínas e, por similaridade, completem as regiões não elucidadas por técnicas experimentais.

A região da protease que precisou ser adicionada compreende os resíduos 183-194 da NS2B. Esses resíduos fazem parte de uma região altamente móvel da NS2B e, para reconstruir essa região, utilizamos um programa de comparação estrutural. Neste caso, utilizamos o programa *Swiss-Model* (ARNOLD et al., 2006), onde se submete online a sequência para a qual se deseja a estrutura, e o PDB ID (2FOM)(ERBEL et al., 2006) da sequência na qual se deseja completar regiões não solucionadas. O programa, por um método de comparação (similaridade), completa as regiões e fornece um novo arquivo de coordenadas (Figura 24).



**Figura 24:** Estrutura da 2FOM (DENV2a) com acréscimo dos resíduos 183-194. Em azul a NS3 (1-150), em vermelho a NS2B (151-204), e em verde os resíduos adicionados à estrutura (183-194).

#### 4.1.2 Estado de Protonação de Resíduos Ionizáveis

Para verificar os estados de protonação dos resíduos ionizáveis da proteína usamos o programa *MolProbity*. (CHEN et al., 2010)

Ao entrar com o *ID:PDB* da proteína no servidor do *MolProbity* (<http://molprobity.biochem.duke.edu>), o programa oferece uma opção de adicionar hidrogênios, que é feito a partir do cálculo de *pKa* dos resíduos. Depois verificamos o estado de protonação das histidinas, aspartato e glutamatos calculadas pelo programa, foi possível verificar quais mudanças seriam necessárias no arquivo PDB original. A única alteração que efetuamos foi a mudança da histidina do sitio ativo, que apresentou uma protonação delta; as demais histidinas têm sua protonação épsilon. No PDB mudamos a chamada HIS do sitio ativo para HID, para posteriormente os hidrogênios serem adicionados corretamente por um outro programa. Assim, não salvamos as coordenadas dos hidrogênios fornecida pelo *MolProbity*, essa etapa foi realizada pelo utilitário *LEaP* do *AmberTools11* (CASE et al., 2010a), que adiciona os hidrogênios a partir de moldes, ou seja, ao ler no PDB o molde HID, automaticamente a histidina é protonada na posição delta, o padrão é a protonação na posição épsilon

#### 4.1.3 Arquivos de entrada

A etapa seguinte é adicionar os átomos de hidrogênios dos resíduos, moléculas do solvente e contra íons. Se o sistema a ser simulado for um sistema isotrópico homogêneo (fluido), o procedimento mais usual é a criação de uma “caixa de simulação” ortorrômica onde se encerram as *N* moléculas do sistema.

Para começar os cálculos no programa *pmemd* do pacote de programas *Amber* (CASE et al., 2010a) é necessário que se tenha as coordenadas da estrutura, assim como um arquivo com os dados de parâmetros do campo de força utilizado, e outro arquivo contendo os comandos de controle da simulação.

O programa utilizado para gerar a estrutura inicial e o arquivo com os parâmetros foi o utilitário *LEaP*, parte do pacote *AmberTools* (CASE et al., 2010a). Ao final se tem o arquivo de coordenadas que inclui a proteína, íons e solvente (H<sub>2</sub>O) e, o arquivo de parâmetros.

O campo de força usado para a proteína foi o Amber ff99SB, (HORNAK et al., 2006) e o modelo para as moléculas de H<sub>2</sub>O o TIP3P.(JORGENSEN et al., 1983) Em todos os casos, os tamanhos das caixas de simulação foram tais que a distância entre qualquer átomo da proteína e qualquer parede da caixa seja de, no mínimo, 10 Å. Condições periódicas de contorno foram utilizadas para simular o sistema.

#### 4.2 CRIAÇÃO DO MODELO ATIVO DA NS2B-NS3 (DENV2f)

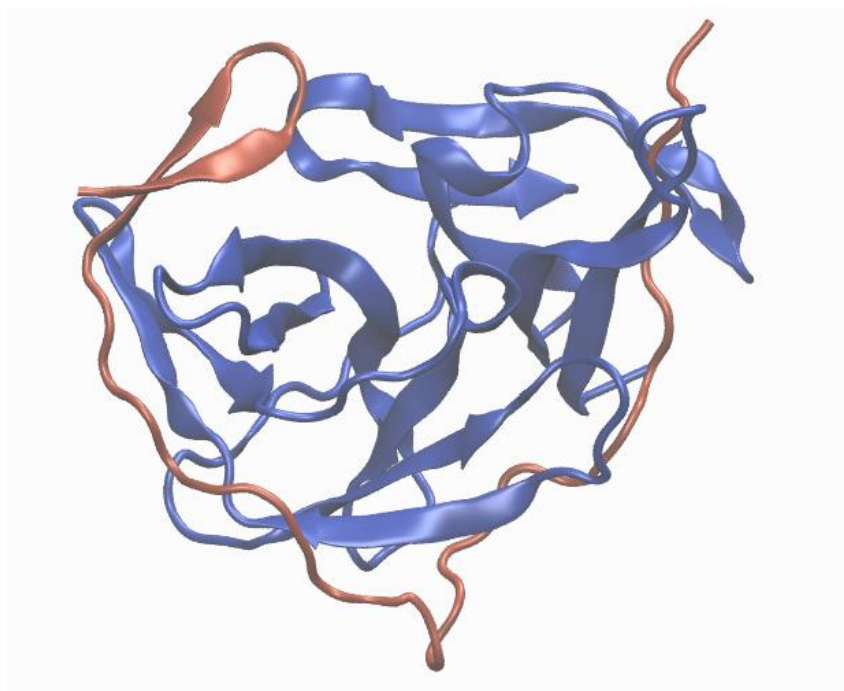
Como no início desses trabalhos ainda não havia sido determinada a estrutura cristalina para a forma ativa da NS2B-NS3 protease, criamos um modelo para a estrutura ativa a partir de um molde da mesma proteína de outro vírus, o Vírus do Nilo Ocidental, (WNV, PDB ID:2FP7(ERBEL et al., 2006)). O modelo foi criado utilizando o programa *Swiss-Model* (ARNOLD et al., 2006), em que se forneceu a sequência FASTA\* (LIPMAN; PEARSON, 1985) do vírus da Dengue (2FOM), e o molde da estrutura da protease do vírus do Nilo Ocidental (2FP7).

O objetivo de modelar a estrutura de proteínas por comparação ou homologia é a construção de um modelo tridimensional (3D) para proteínas de estrutura desconhecida (o alvo) com base na similaridade de sequência de proteínas de estrutura conhecida (o molde) (MARTÍ-RENO et al., 2000). Usualmente, uma maior similaridade de sequência entre o alvo e o molde produz um modelo melhor. Portanto, a maior precisão estrutural do modelo depende de quão próximos seja a sequência de identidade entre o alvo e molde, estudos indicam que uma sequência de identidade > 30% gera modelo estrutural de alta similaridade entre as estruturas alvo e molde (BORDOLI et al., 2009; MARTÍ-RENO et al., 2000).

A alta sequência de identidade entre WNV e DENV, sendo de 57% para o domínio NS3 e 40% para o domínio NS2B (WICHAPONG et al., 2010a), indica uma alta similaridade entre estas duas estruturas, ou seja, um modelo estrutural confiável para o estudo.

---

\* O formato FASTA, em bioinformática, é um formato baseado em texto para identificar, através de códigos de uma única letra, sequências de nucleotídeos ou sequências de aminoácidos.



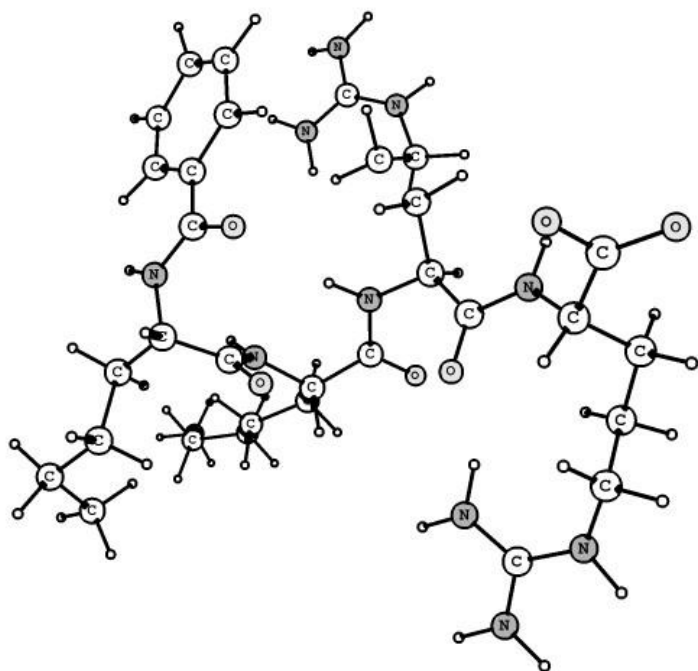
**Figura 25:** Modelo para a estrutura ativa da NS2B-NS3 (DENV2f). Em azul a NS3, e em vermelho a NS2B.

A partir desse modelo (Figura 25) foram criados também mais dois sistemas: um na presença de um inibidor e outro na presença de um modelo de substrato, como descrito a seguir.

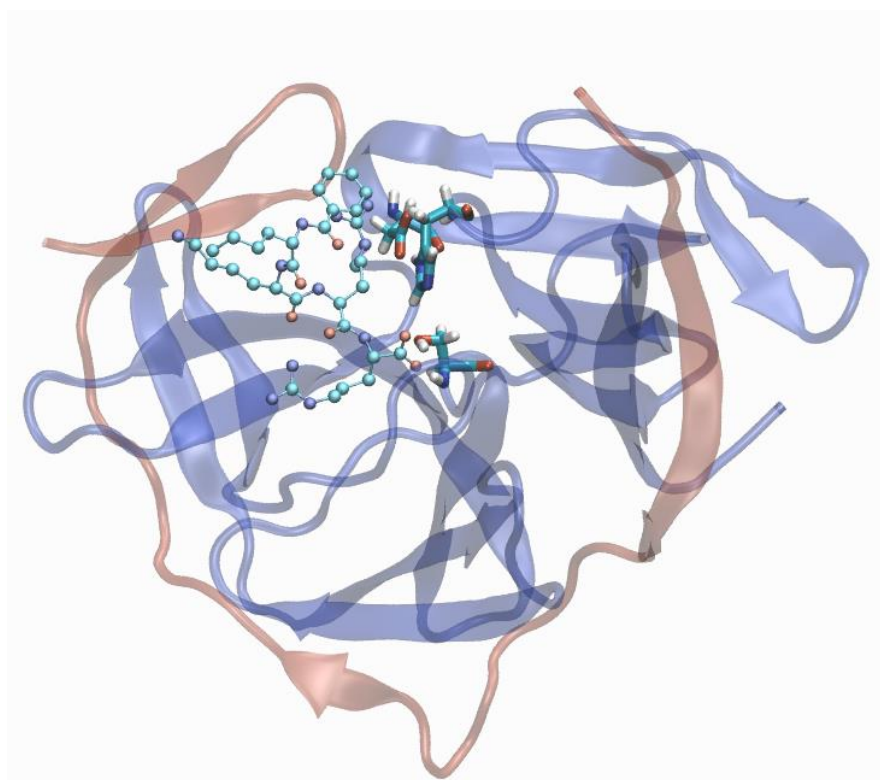
#### 4.3 DENV2f-INIB

Usamos o modelo DENV2f para protease, enquanto a posição do inibidor (Bz-NLE-ARG-ARG-LYS-H, Figura 26) foi a mesma posição do caso do WNV (2FP7).(ERBEL et al., 2006) Utilizamos o programa VMD (*Visual Molecular Dynamics*)(HUMPHREY; DALKE; SCHULTEN, 1996) para alinhar as estruturas da DENV2f e WNV, assim adicionar as coordenadas do inibidor na estrutura da DENV2f, tornando o arquivo de coordenadas DENV2f-INIB. A estrutura da DENV2f com inibidor pode ser vista na Figura 27.





**Figura 26:** Inibidor Bz-NLE-ARG-ARG-LYS-H.



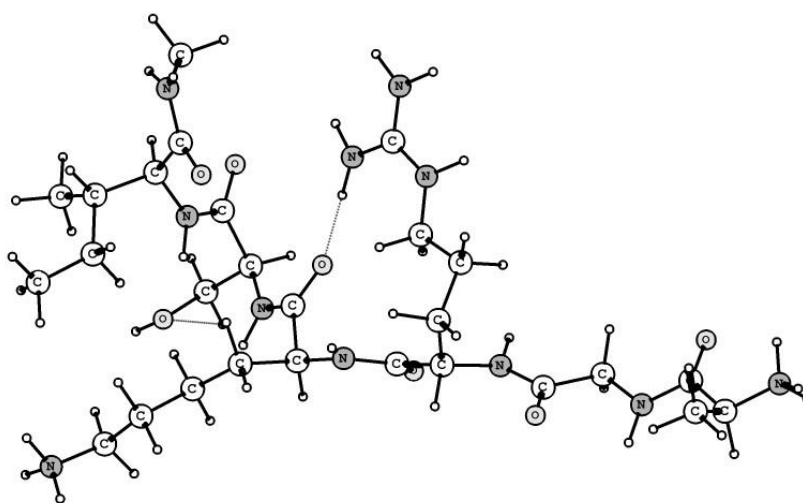
**Figura 27:** DENV2com inibidor (DENV2f-INIB). Em azul a NS3, em vermelho NS2B, canudo a tríade catalítica (HIS33, ASP57 e SER117), e as bolas representam o inibidor.



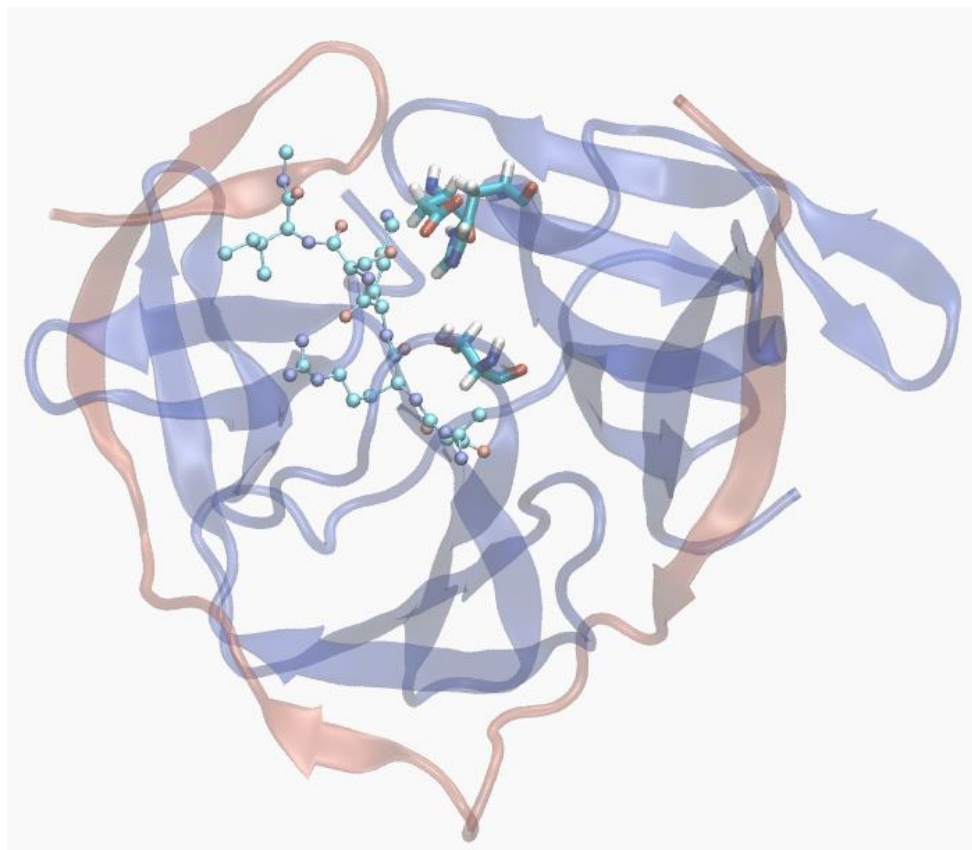
#### 4.4 DENV2f-SUBS

##### 4.4.1 Substrato ALA-GLY-ARG-LYS-SER-ILE-NME (DENV2f-SUBS1)

Neste sistema o modelo DENV2f também foi usado, e a posição do substrato (Figura 28) foi alinhada à posição do inibidor, alinhando os átomos da espinha dorsal (*backbone*) do substrato aos átomos do *backbone* do inibidor, com o programa VMD (HUMPHREY; DALKE; SCHULTEN, 1996). (Figura 29).



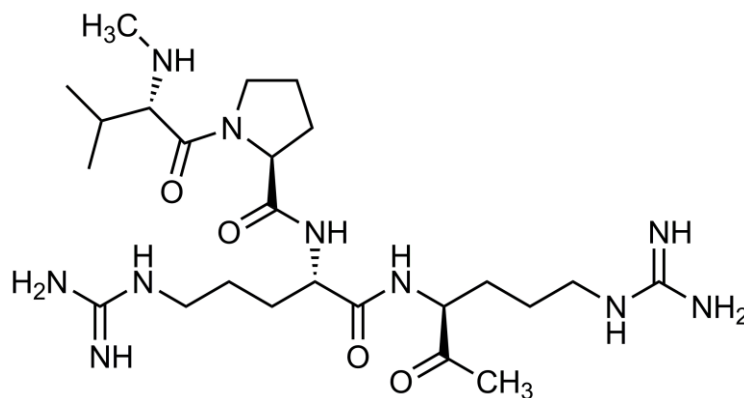
**Figura 28:** Substrato ALA-GLY-ARG-LYS-SER-ILE-NME. NME é um resíduo para completar a valência, -NHCH<sub>3</sub>.



**Figura 29:** DENV2 com o substrato (DENV2-SUBS1). Em azul a NS3, em vermelho NS2B, canudo a tríade catalítica (HIS33, ASP57 e SER117), e as bolas representam o substrato.

#### 4.4.2 Substrato $\text{CH}_3\text{NH-ARG-ARG-PRO-VAL-COCH}_3$ (DENV2f-SUBS2)

A escolha desse substrato foi baseada em estudos experimentais que analisaram a eficiência da catálise da protease NS2B-NS3 para várias pequenas sequências de aminoácidos. (SHIRYAEV et al., 2007) A sequência de resíduos com melhor eficiência catalítica foi  $\text{GKKRR}\downarrow\text{PVR}$  (SHIRYAEV et al., 2007). Com base nesse substrato escolhemos a sequência de quatro aminoácidos  $\text{CH}_3\text{NH-RRPV-COCH}_3$  (Figura 30).

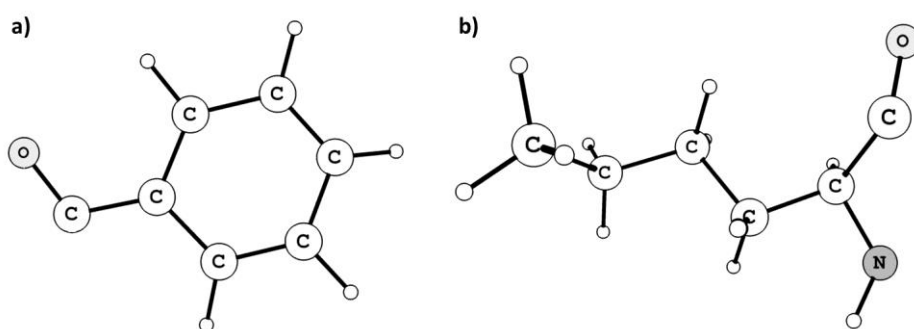


**Figura 30:** Substrato  $\text{CH}_3\text{NH-RRPV-COCH}_3$ , onde os grupos  $\text{CH}_3\text{NH}$  e  $\text{COCH}_3$  são para completar a valência da sequência de aminoácidos do modelo do substrato.

Para posicionar esse substrato no sítio ativo da enzima, usamos o programa VMD (HUMPHREY; DALKE; SCHULTEN, 1996), alinhando os átomos do *backbone* do substrato aos átomos do *backbone* dos resíduos do inibidor Bz-Nle-Lys-Arg-Arg-H, presente na estrutura obtida por raios-X,(ERBEL et al., 2006) o que deixou a região -Pro-Arg- (ligação peptídica a ser quebrada) próxima ao sítio ativo.

#### 4.5 PARAMETRIZAÇÃO DO INIBIDOR

Houve a necessidade de parametrizar dois resíduos do inibidor (Bz e NLE, Figura 31). Para que o programa possa gerar os parâmetros de campo de força para o inibidor são necessárias a geometria e as cargas, que foram calculadas utilizando-se o programa *Gaussian09* (M. J. FRISCH, G. W. TRUCKS, H. B. SCHLEGEL, G. E. SCUSERIA, M. A. ROBB, J. R. CHEESEMAN, G. SCALMANI, V. BARONE, B. MENNUCCI, G. A. PETERSSON, H. NAKATSUJI, M. CARICATO, X. LI, H. P. HRATCHIAN, A. F. IZMAYLOV, J. BLOINO, G. ZHENG, J. L. SONNENBERG, M. HAD, 2009) e, depois, ajustados no programa *Antechamber* (CASE et al., 2010b) utilizando o método RESP (BAYLY et al., 1993). Os parâmetros utilizados pelo *Antechamber* são do campo de força GAFF (WANG et al., 2004), que já inclui parâmetros para moléculas orgânicas ( C, N, O, S, P, H, F, Cl, Br e I).



**Figura 31:** Resíduos que foram parametrizados. a) Bz (benzoil) e b) NLE (norleucina) tem um grupo  $-CH_3$  a mais quando comparado com a ILE.

#### 4.6 MINIMIZAÇÃO DA ESTRUTURA

O objetivo desta fase é ajustar a estrutura em função dos campos de forças, em particular a distribuição das moléculas do solvente, e relaxar possíveis impedimentos estéricos

criados durante a preparação da estrutura inicial, em especial a adição de moléculas de solvente.

Em primeiro lugar, é feita uma minimização de energia com 100 passos com o algoritmo *steepest descente* (JENSEN, 2007, p. 383–385), seguidos de 900 passos de gradiente conjugado (*conjugate gradient*) (JENSEN, 2007, p. 383–385) mantendo a restrição em todo sistema (proteína e/ou inibidor e/ou substrato) constante de força de  $100 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ , exceto nas moléculas de água, assim relaxando a posição dos seus átomos. Um raio de corte (*cutoff*) para as interações não ligadas de  $10 \text{ \AA}$ . Essa etapa é importante para deixar as moléculas de água mais bem distribuídas.

Posteriormente, minimizamos todo o sistema, deste modo, retiramos a restrição de coordenada sobre o sistema (proteína e/ou inibidor e/ou substrato), com 100 passos com algoritmo *steepest descente*, seguidos de 900 passos de *conjugate gradient*. E um raio de corte (*cutoff*) para as interações não ligadas de  $10 \text{ \AA}$ .

#### 4.7 DINÂMICA MOLECULAR

Depois que o sistema passou pela etapa de minimização, o aquecimento foi feito com uma simulação a volume constante e temperatura de  $100 \text{ K}$  por  $1 \text{ ns}$ , seguido de  $1 \text{ ns}$  a  $300 \text{ K}$ , a volume e temperatura constantes (*ensemble* NVT). Uma vez que o sistema é aquecido (NVT), então continuamos a dinâmica em pressão constante (NPT) para permitir que a densidade atinja o equilíbrio, por  $20 \text{ ns}$  à pressão e temperatura constantes de  $1 \text{ bar}$  e  $300 \text{ K}$ , verificando a convergência por meio do monitoramento das propriedades energéticas do sistema, densidade e RMSD. Por último, depois que o sistema entrou em equilíbrio, a etapa de produção foi realizada num *ensemble* NPT, importante para garantir que o sistema líquido esteja com a densidade correta durante a simulação. O sistema foi finalmente submetido a  $200 \text{ ns}$  de simulação a uma temperatura de  $300 \text{ K}$  e uma pressão constante de  $1 \text{ bar}$  para estabilização. Todas as simulações usaram condições periódicas de contorno.

No *ensemble* NVT, usamos termostato de *Langevin* (UBERUAGA; ANGHEL; VOTER, 2004) com frequência de colisões  $2,0 \text{ ps}^{-1}$ , raio corte (*cutoff*) de  $10 \text{ \AA}$  para interações não ligadas, e o método de *Particle Mesh Ewald* (PME)(DARDEN; YORK; PEDERSEN, 1993) para as interações eletrostáticas e de longa distância. O algoritmo *SHAKE* (RYCKAERT; CICCOTTI; BERENDSEN, 1977) foi utilizado para restringir as ligações envolvendo átomos de hidrogênio, permitindo um passo de simulação de  $2 \text{ fs}$ .

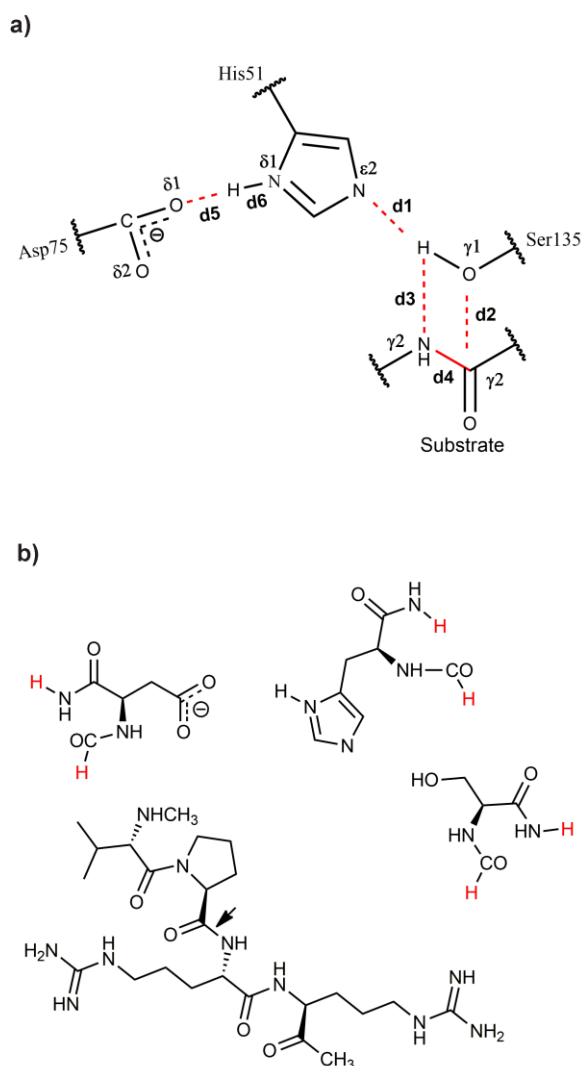
Em um *ensemble* NPT utilizamos os mesmos parâmetros anteriores, somados ao barostato isotrópico de *Berendsen* (BERENDSEN et al., 1984) com constante de acoplamento 1,0 ps para manter a pressão em 1 atm. Novamente, algoritmo *SHAKE* (RYCKAERT; CICCOTTI; BERENDSEN, 1977) foi utilizado para restringir as ligações envolvendo átomos de hidrogênios. Detalhes dos protocolos de simulação, com os arquivos utilizados, estão reproduzidos no Apêndice A.

Portanto, todos os sistemas DENV2a (2FOM) (ERBEL et al., 2006), DENV2f (modelo), DENV2f-INIB e DENV2f-SUBS1/2 (substrato 1 e substrato 2) passaram pelas mesmas etapas de minimização e dinâmica molecular (*ensembles* NVT e NPT). O programa *pmemd* que está no pacote de programas Amber11 foi utilizado para os cálculos.(CASE et al., 2010a)

#### 4.8 Cálculos Híbridos (QM/MM)

Para as simulações que envolvem quebra ou formação de ligações, utilizamos o hamiltoniano híbrido QM/MM (WARSHEL; LEVITT, 1976a; WEINER; SINGH; KOLLMAN, 1985). Ou seja, O sistema foi dividido em duas regiões: a parte QM definida pelo substrato (CH<sub>3</sub>OC-V-P-R-R-NHCH<sub>3</sub>) e pelos resíduos catalíticos (HIS51, ARG75 e SER135), e mais doze átomos de resíduos vizinhos (N e CO da espinha dorsal) que fazem ligação peptídica com os resíduos do sitio ativo, evitando substituir uma ligação peptídica por uma ligação com hidrogênio, totalizando 148 átomos QM, descrito pelo método PM3PDDG (REPASKY; CHANDRASEKHAR; JORGENSEN, 2002) (Figura 32). *Link Atoms* do tipo hidrogênio foram utilizados para completar as valências das ligações que atravessam a fronteira entre as regiões QM e MM. O algoritmo *SHAKE* não foi utilizado para a região QM. A região MM incluía todo o resto do sistema (proteína, moléculas de água e íons) sendo descritos pelos parâmetros de campos de força mencionados anteriormente.

Escolhemos uma geometria inicial para o estudo reacional, vinda da simulação de dinâmica molecular (MM) do sistema DENV2f-SUBS2 descrita acima, uma estrutura com a geometria apropriada, ou seja, em que os resíduos que teriam sua ligação peptídica quebrada ficassem próximo ao sitio ativo da protease. Assim, optamos por selecionar uma configuração em que d1 e d2 tivessem os menores valores, d1 = 1,85 Å e d2 = 2,51 Å, respectivamente (Figura 32, Página 81). Assim, tendo a melhor condição para o estudo da reação.



**Figura 32:** (a) Ilustração de algumas das distâncias usadas para esse estudo, na qual,  $d1(H_{\gamma1} - N_{\epsilon2})$ ,  $d2(C_{\gamma2} - O_{\gamma1})$ ,  $d3(N_{\epsilon2} - H_{\gamma1})$ ,  $d4(C_{\gamma2} - N_{\gamma2})$ ,  $d5(O_{\delta1} - N_{\delta1})$  e  $d6(N_{\delta1} - H_{\delta1})$ . (b) Ilustração da região QM, os átomos de hidrogênio em vermelho são átomos para completar a valência da região QM.

#### 4.8.1 Coordenada de Reação

Para explorar a coordenada de reação utilizamos o método de amostragem de guarda-chuva (*Umbrella Sampling*). (KUMAR et al., 1992) A reação foi dividida em dois passos: o primeiro passo, caracterizado pela formação da ligação acil-enzima (intermediário tetraédrico, Figura 10, Página 32), e o segundo passo a quebra da ligação peptídica a partir do intermediário tetraédrico.

Para o primeiro passo, fizemos uma varredura bidimensional. Um potencial de restrição foi utilizado para controlar a transferência de próton ( $d1$  na Figura 32, Página 81), com distância variando de 1.90 Å a 0.90 Å, e outro para o ataque nucleofílico ( $d2$  na Figura 32), variando de 2.40 Å a 1.35 Å.

Na segunda etapa, as restrições utilizadas na etapa anterior foram removidas. A transferência do próton da histidina ( $H_{\gamma1}$ ) para o nitrogênio da amina ( $N_{\gamma2}$ ) no substrato, foi controlada através da restrição da coordenada  $d3$  na Figura 32, com valores variando de 2.60 Å a 0.85 Å. Essa aproximação leva naturalmente à quebra da ligação peptídica ( $d4$  na Figura 32), de forma que o controle de uma única coordenada de reação foi suficiente.

Um potencial quadrático foi utilizado em todas as restrições. As distâncias foram varridas em intervalos de 0.1 Å, e as constantes de força entre 200 e 1500 kcal mol<sup>-1</sup> Å<sup>-2</sup>. Assim, foram utilizadas 120 janelas para a primeira etapa, e 16 para a segunda etapa da reação. Necessitou-se de 136 janelas de simulação para a obter os perfis completos de energia.

A simulação de cada janela foi realizada em três etapas. Em primeiro lugar, é feita uma minimização de energia com 100 passos com o algoritmo *steepest descent*, seguidos de 900 passos de gradiente conjugado (*conjugate gradient*) mantendo as restrições e relaxando a posição dos demais átomos. O sistema é então aquecido à 300 K por 5 ps a volume constante, seguidos de 35 ps à pressão e temperatura constantes de 1 bar e 300 K, sempre mantendo as restrições. O termostato de *Langevin* foi usado para controle da temperatura, com uma frequência de colisão de 2.0 ps<sup>-1</sup>, (UBERUAGA; ANGHEL; VOTER, 2004) e o controle da pressão com o barostato de *Berendsen*, (BERENDSEN et al., 1984) com tempo de relaxação 1 ps. Assim, usou-se 10 Å de raio de corte (*cutoff*) para as interações não ligadas. As análises foram realizadas utilizando-se apenas os últimos 30 ps de simulação. Os pontos críticos da reação foram ainda submetidos a mais 50 ps de simulação para os pontos da etapa de acilação, e 30 ps para os pontos críticos da etapa da quebra da ligação peptídica, para melhor caracterizar os estados. Todas as simulações QM/MM foram realizadas utilizando o programa *sander* (CASE et al., 2010b) e um passo de simulação de 1 fs.

As energias livres foram obtidas com o *Weighted Histogram Analysis Method* (WHAM), (KUMAR et al., 1992) utilizando uma temperatura de 300 K, 110 e 85 bins para as etapas de acilação e quebra da ligação peptídica, respectivamente, e tolerância de convergência de 0.01.

#### 4.9 ANÁLISE DOS RESULTADOS

Uma parte crucial dos estudos de simulação é a análise das trajetórias obtidas. No nível mais básico, a análise de uma trajetória dinâmica consiste em calcular uma propriedade para cada quadro da trajetória e ver como ela muda em função do tempo. Como a simulação de dinâmica molecular gera um número grande de dados é necessário o uso de programas, e por vezes, criar pequenos programas (*scripts*) para a análise desejada.

##### 4.9.1 RMSD e RMSF

O deslocamento quadrático médio ou RMSD (*root mean square deviation*) é uma medida da mobilidade relativa do sistema.(CRAMER, 2004)

A evolução temporal do RMSD é usada em dinâmica molecular para avaliar o quanto a estrutura molecular de interesse se modifica durante a simulação, em relação à uma estrutura de referência.(CRAMER, 2004) . Assim se faz o RMSD para um conjunto de átomos específicos a cada passo de simulação. Desta maneira, esse cálculo pode ser analisado fazendo um gráfico RMSD vs. tempo, para avaliar quais mudanças estruturais ocorrem ao longo da simulação. Variações pequenas em torno da posição de equilíbrio indica baixa mobilidade, enquanto, mudanças bruscas, durante a dinâmica molecular pode indicar uma mudança na conformação da proteína, equação 6.

$$RMSD(t_i) = \sqrt{\frac{\sum_{j=1}^N [r_j(t_i) - r_0]^2}{N}} \quad [6]$$

Onde N é número total de átomos considerados. Normalmente, usa-se os átomos da cadeia principal dos resíduos. E  $r_j(t_i)$  é a posição do átomo  $j$  num tempo  $t_i$ , e  $r_0$  é a coordenada de referência.

A flutuação quadrática media ou RMSF é uma medida do desvio entre a posição da partícula  $i$  e a posição da referência com relação ao tempo ( $t$ ), equação 7.(CRAMER, 2004)

$$RMSF_i = \sqrt{\frac{\sum_{j=1}^t [r_i(t_j) - r_0]^2}{t}} \quad [7]$$

RMSD é uma medida refletindo desvio de todas as coordenadas de referência para as coordenadas em cada quadro, enquanto RMSF reflete o movimento global de átomos (ou

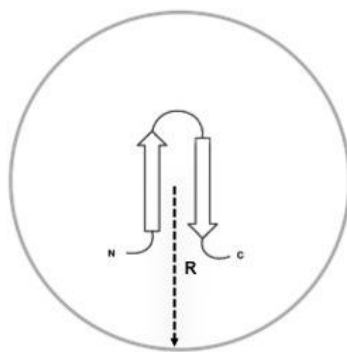


resíduos depender do seu tipo de saída) ao longo de todos os quadros. (VOTANO; PARHAM; HALL, 2004)

O programa *cpptraj* (parte do pacote *AmberTools11*) (CASE et al., 2010b) foi utilizado para os cálculos de RMSD e RMSF.

#### 4.9.2 Análise dos vizinhos

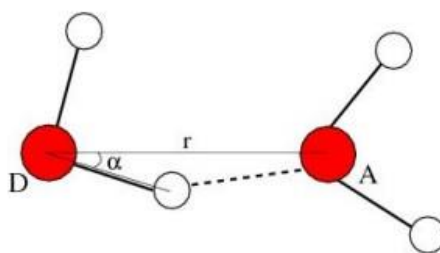
Usando o programa VMD (*Visual Molecular Dynamics*)(HUMPHREY; DALKE; SCHULTEN, 1996) selecionamos os resíduos os quais desejávamos avaliar os vizinhos próximos, dentro de um raio de 5,0 Å (Figura 33), contando isso a cada configuração da trajetória da simulação. Deste modo, gerando um histograma a partir dessa informação.



**Figura 33:** Definição dos vizinhos a uma região da proteína. Onde R é o raio que inclui a vizinhança.

#### 4.9.3 Ligações de Hidrogênio

Para determinar se existe uma ligação de hidrogênio entre um átomo doador D e um aceitador A, utilizamos um critério geométrico (Figura 34):  $r \leq 3,5 \text{ Å}$  e  $\alpha \leq 30^\circ$ . Este é o critério padrão, entretanto, o programa *cpptraj* (parte do pacote *AmberTools*) permite que os critérios sejam alterados.



**Figura 34:** Definição para a ligação de hidrogênio. Onde  $r$  é a distância entre doador e aceitador de ligação de hidrogênio, e  $\alpha$  é o ângulo.

#### 4.9.4 Número de moléculas de H<sub>2</sub>O

Para avaliar quantas moléculas de água estavam próximas a um grupo de resíduos, usamos uma opção do programa *cpptraj* (incluso do pacote *Ambertools*) que realiza a contagem do número médio de moléculas de H<sub>2</sub>O que se encontram dentro de um raio de 5,0 Å do centro de massa de um grupo de resíduos, para cada quadro de simulação.

#### 4.10 Performance dos Cálculos

Existe um custo computacional na execução do estudo, tanto nos cálculos de dinâmica molecular clássica (MM), quanto para a dinâmica molecular usando o método híbrido QM/MM.

Ao usarmos 64 processadores Xeon 2,67 GHz, CENAPAD-PE, o desempenho da dinâmica molecular clássica (MM) foi de 2,2 ns/dia. Mas ao usar 128 processadores AMD Opteron 2.2GHz (supercomputador TITAN), Oak Ridge National Laboratory (EUA), a performance passou para 8,5 ns/dia. Entretanto, ao usarmos uma GPU Geforce GTX 780, a performance passou para 34 ns/dia.

Já os cálculos de dinâmica molecular QM/MM em uma máquina com 8 processadores Intel Xeon 1,80 GHz, a performance foi 0,02 ns/dia. Pois o cálculo QM exige maior demanda computacional para sua execução.

## 5. RESULTADOS e DISCUSSÃO

---

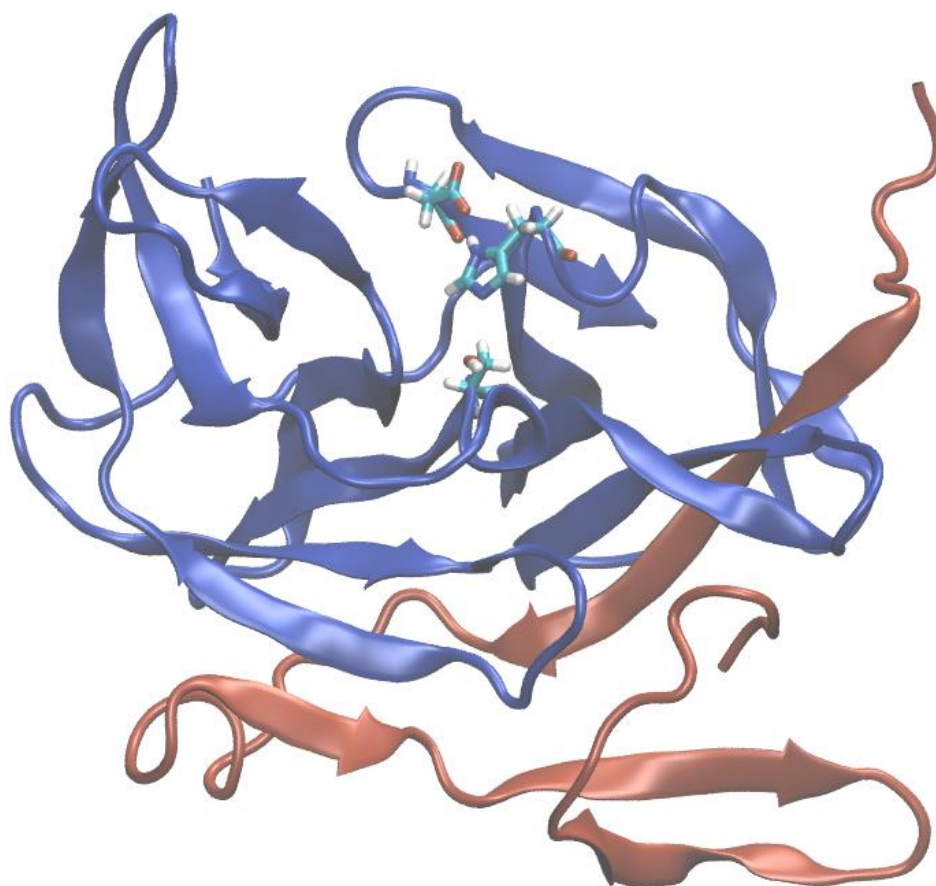
Estudos estruturais sugeriram que as proteases tanto da DENV (Dengue) quanto WNV (Nilo Ocidental) existam em equilíbrio entre as formas aberta e fechada, assim como outras conformações, e que este equilíbrio é deslocado para a forma fechada, na presença de um inibidor ou substrato.(KIM et al., 2013) Em solução, a forma fechada da protease do WNV é a conformação predominante, independentemente da presença ou ausência de inibidores. Entretanto, o mesmo estudo indica que a conformação aberta é a forma predominante na ausência de inibidores para a protease do DENV. (NOBLE et al., 2012)(SU et al., 2009) Para entender o funcionamento da NS2B-NS3 protease, realizamos cálculos de dinâmica molecular tanto para estrutura fechada quanto a estrutura aberta, na presença do inibidor e do substrato.

Realizamos ainda o estudo da reação enzimática de quebra da ligação peptídica do substrato-2 pela NS2B-NS3 protease do vírus da Dengue (DENV2f). Os resultados indicam que a formação do intermediário tetraédrico deve ocorrer em duas etapas claramente distintas, sendo a primeira a ativação da Ser135 pela transferência do próton entre a serina e a His51 e, posteriormente, o ataque nucleofílico da Ser135 ao carbono da carbonila no substrato. Esse intermediário tetraédrico tem seu ânion estabilizado por moléculas de água formando a cavidade oxiânion.

### 5.1 SISTEMA DENV2a: NS2B-NS3 (2FOM)

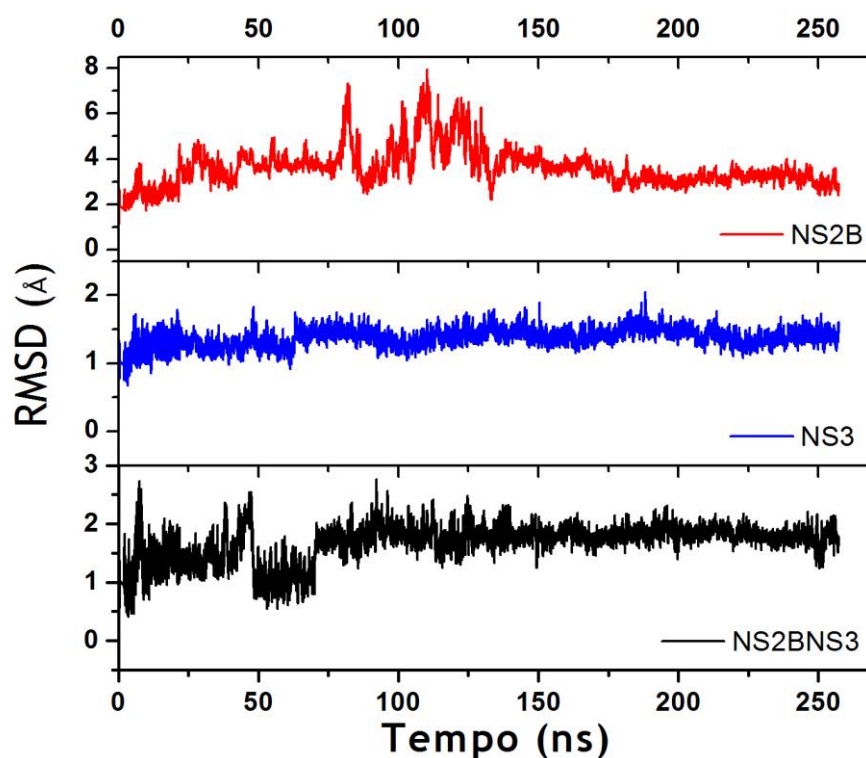
Nessa secção serão apresentados os resultados da simulação para a estrutura aberta da NS2B-NS3 protease (DENV2a), que foi modelada baseada na estrutura cristalográfica retirada do banco de dados de proteína (PDB ID: 2FOM) (NOBLE et al., 2012).

A Figura 35 representa uma conformação da NS2B-NS3 durante a simulação num ensemble NPT. Essa estrutura mostra que o sítio ativo da protease se encontra numa região externa, e que a NS2B (cofator) se localiza próxima a uma pequena região da NS3, e a outra parte do cofator fica imersa na região do solvente.



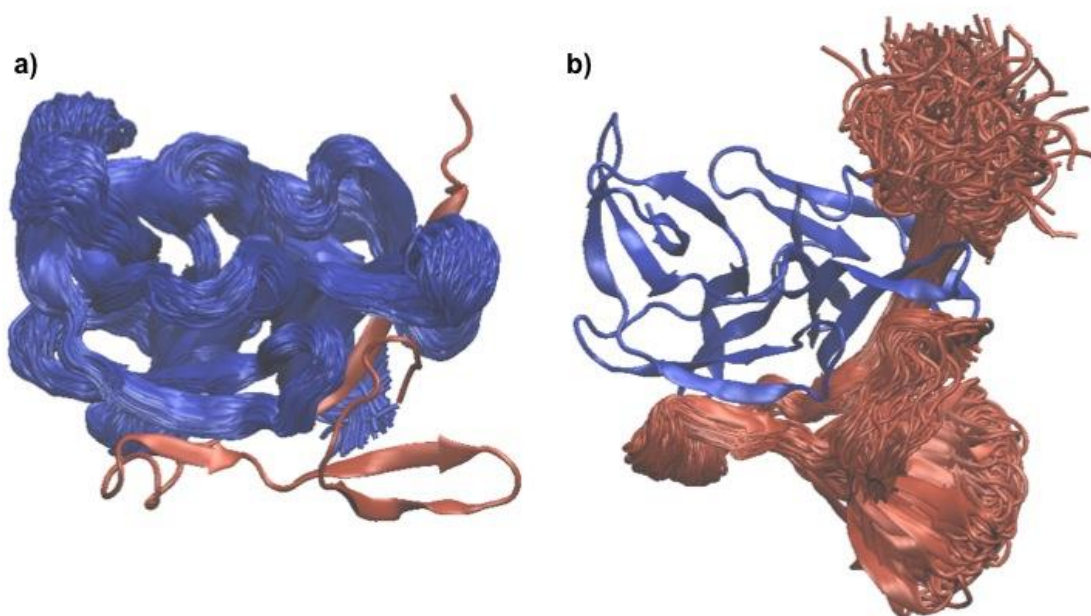
**Figura 35:** Conformação da NS2B-NS3 obtida de uma simulação de dinâmica molecular num ensemble NPT. As moléculas de água da caixa de simulação foram omitidas para melhor visualização. Em Azul NS3, vermelho a NS2B, e representado em canudo estão os resíduos do sítio ativo (Asp, His e Ser).

Os dados de RMSD (Figura 36), podemos ver mostram que a simulação atinge o equilíbrio após os 75-100 ns iniciais. Entretanto, quando separamos os RMSDs das regiões da proteína, observamos uma flutuação maior do cofator (NS2B – em vermelho na Figura 36), já um comportamento de pouca flutuação para a NS3 protease (gráfico azul da Figura 36). Essa maior flutuação observada no cofator já era esperada, uma vez que o mesmo se encontra mais livre no meio solvente, além de interagir pouco com a NS3. Assim, a NS2B passa maior parte do tempo procurando outras conformações.



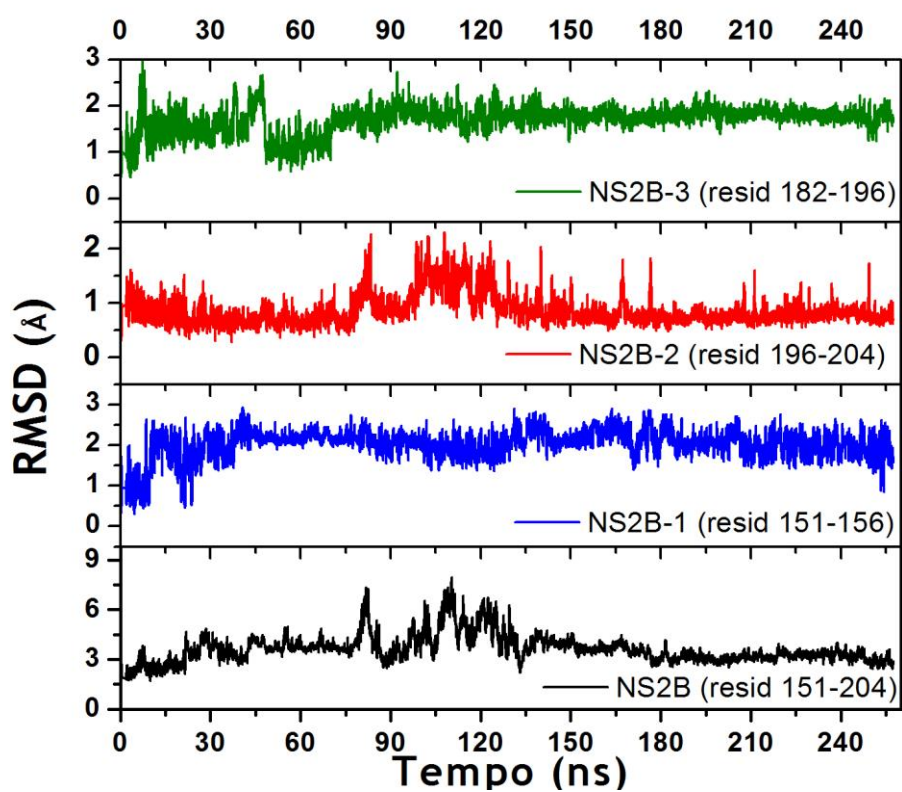
**Figura 36:** RMSD para DENV2a (preto), assim como, o RMSD só para a NS3 (azul) e para o cofator NS2B (vermelho).

Essa maior mobilidade do cofator (NS2B) pode ser vista na Figura 37, em que há superposição de 700 conformações da NS2B da simulação de dinâmica molecular NPT. Essa superposição evidencia as regiões do cofator que possuem maior mobilidade que são, como seria esperado, as regiões terminais e a região que fica mais exposta ao solvente e tem menos contato com a NS3. Já as regiões com maior contato com a NS3, apresentam menor mobilidade. Por outro lado, na superposição das 700 conformações da NS3 (Figura 37), fica evidente a estabilidade conformacional da proteína.



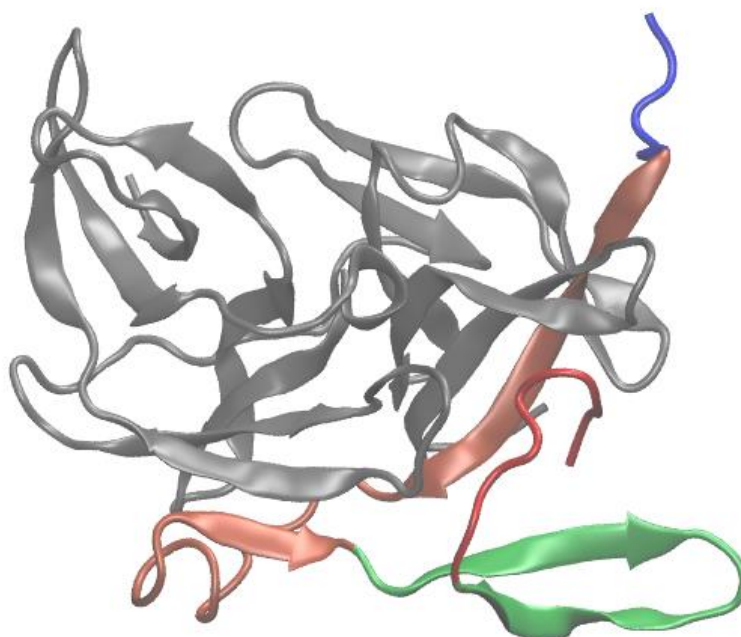
**Figura 37:** Superposição para as 700 conformações extraídas a intervalos regulares da simulação de dinâmica molecular da DENV2a. a) superposição da NS3; b) superposição da NS2B. Em Azul tem-se a NS3, e em vermelho a NS2B.

Um exame mais detalhado das regiões de maior mobilidade da NS2B pode ser feito analisando-se separadamente algumas regiões, como mostrado na Figura 38. Em preto, o RMSD do NS2B inteiro mostra, como já se viu anteriormente (vermelho na Figura 36), uma maior amplitude entre os tempos de 70 a 130 ns, que também é vista para a região dos resíduos 196 a 204 (em vermelho), ou seja, a maior mobilidade é devida a esses resíduos. A Figura 39 mostra que esses resíduos correspondem à região terminal da NS2B. Segundo a literatura, esta é a região que precisa se ajustar para a ativação da proteína, formação de uma conformação fechada. (NOBLE et al., 2012)



**Figura 38:** RMSD para NS2B completo (preto), e para algumas regiões do cofator: resíduos 151 a 156 em azul; 196 a 204 em vermelho, e 182 a 196 em verde. Em todos os casos, a referência foi a estrutura inicial da simulação.

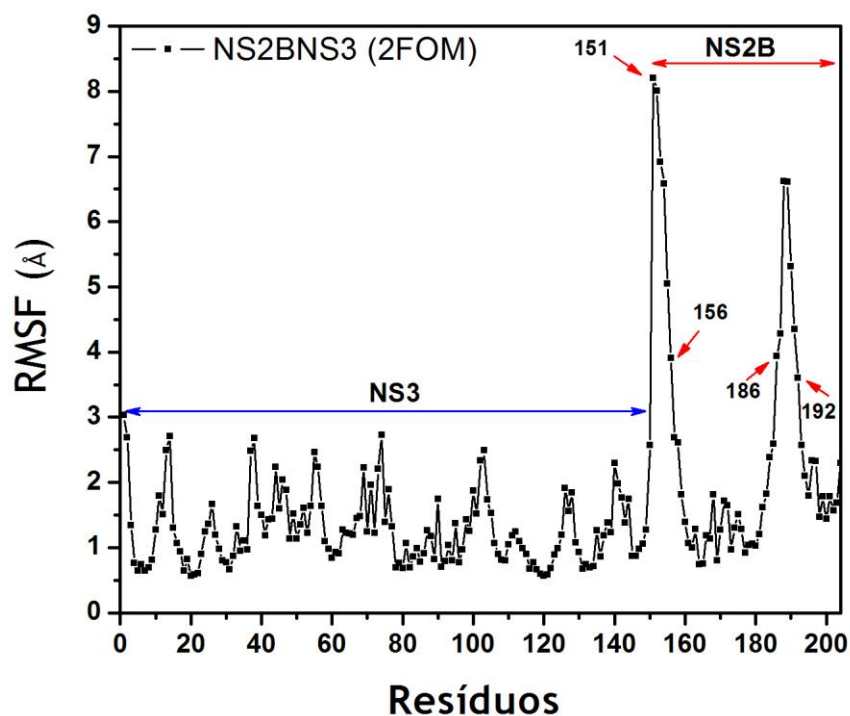




**Figura 39:** Ilustração da DENV2a mostrando os resíduos da NS2B analisados por RMSD na **Figura 38**. Em cinza está a NS3, já a NS2B tem seus resíduos divididos em: azul os resíduos 151 a 156, vermelho claro os resíduos 157 a 181, resíduos 182 a 196 em verde, resíduos 196 a 204 em vermelho.

A mesma informação pode ser obtida a partir das flutuações relativas dos resíduos (RMSF na Figura 40). Como esperado, a maior mobilidade ocorre com os resíduos terminais da NS2B 151 a 156 e 186 a 192. Isso está de acordo com experimento de difração de Raios-X, (ERBEL et al., 2006) que encontraram uma descontinuidade na densidade eletrônica nessa região, indicando que essa última região assume múltiplas conformações em solução.





**Figura 40:** RMSF para os resíduos da DENV2a. Onde os resíduos 1-152 são da NS3 e os resíduos 153-192 da NS2B.

O estudo da estrutura aberta (DENV2a) foi realizado para comparar com o modelo criado para a estrutura ativa da NS2B-NS3 protease (fechada), uma vez que até meados de 2013 não havia sido publicada a elucidação da estrutura cristalográfica dessa protease ativa para o vírus da Dengue.

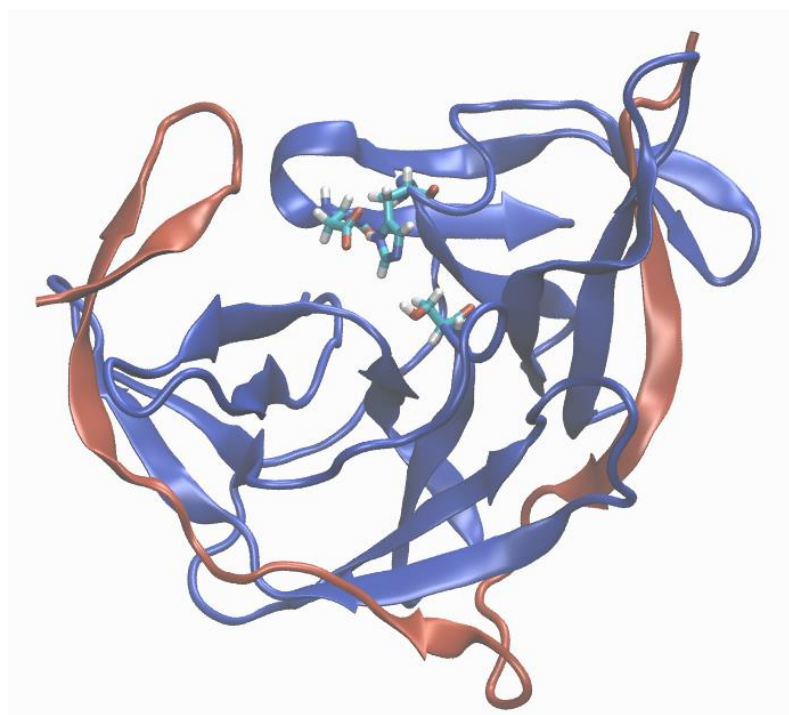
A simulação mostrou que a estrutura aberta é estável em solução. Entretanto, existe uma grande mobilidade do cofator por causa da exposição ao solvente e sua baixa interação com a NS3.

## 5.2 MODELO DA ESTRUTURA FECHADA (ATIVA) DA NS2B-NS3 PROTEASE DO VÍRUS DA DENGUE (DENV2f)

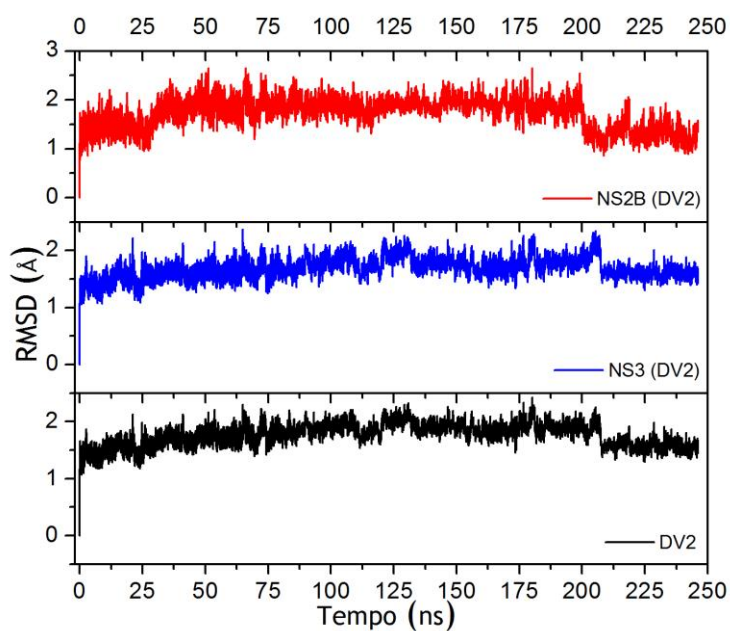
Havia a necessidade da estrutura na sua forma ativa. Para o WNV era conhecida a estrutura com inibidor e (ERBEL et al., 2006) desta mesma protease, que apresenta uma interação maior do cofator com a NS3. Como essas proteínas são homólogas, criamos um modelo por homologia para a estrutura ativa da NS2B-NS3 protease do vírus da Dengue, chamado aqui de DENV2f.

O modelo DENV2f (Figura 41) se mostrou estável com relação a sua estrutura durante a simulação de dinâmica molecular (246 ns). Não se observa alterações significativas no RMSD (Figura 42), indicando uma estabilidade estrutural mesmo a estrutura ativa na ausência de inibidor ou substrato.

A NS3 e a NS2B apresentam também estabilidade quando se observa o RMSD separadamente (Figura 42), indicando que tanto NS3, quanto NS2B se encontram numa região amostral de mínimo local.

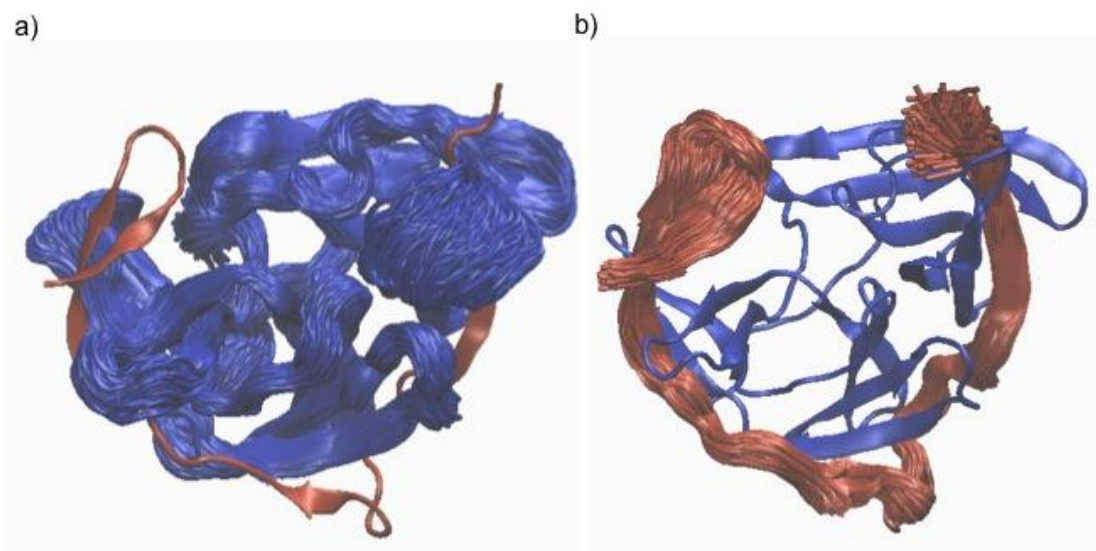


**Figura 41:** Uma conformação da DENV2 extraída de uma simulação de dinâmica molecular num *ensemble* NPT. As moléculas de água da caixa de simulação não estão representadas na figura para melhor visualização. Em Azul NS3, vermelho a NS2B, e representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117).



**Figura 42:** RMSD para NS2B-NS3 do modelo DENV2f completa (preto), apenas NS3 (resíduos 1 a 152) em azul e apenas NS2B (resíduos 153 a 192) em vermelho.

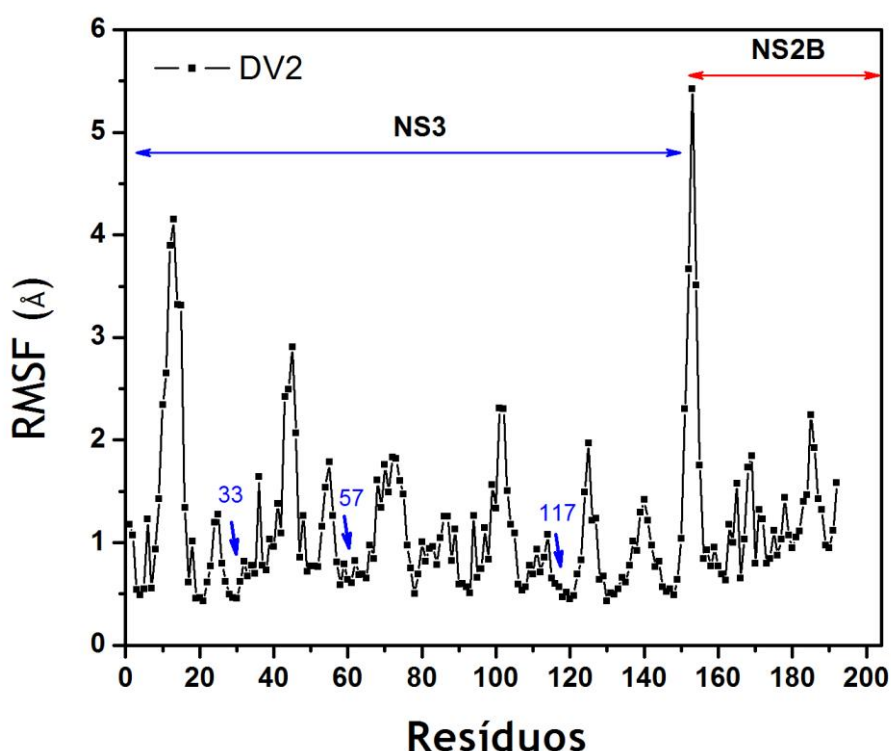
A Figura 43 mostra a superposição de 700 conformações tanto da NS3 quanto da NS2B. É possível notar que as duas estruturas variam pouco de posição, confirmando a estabilidade da estrutura.



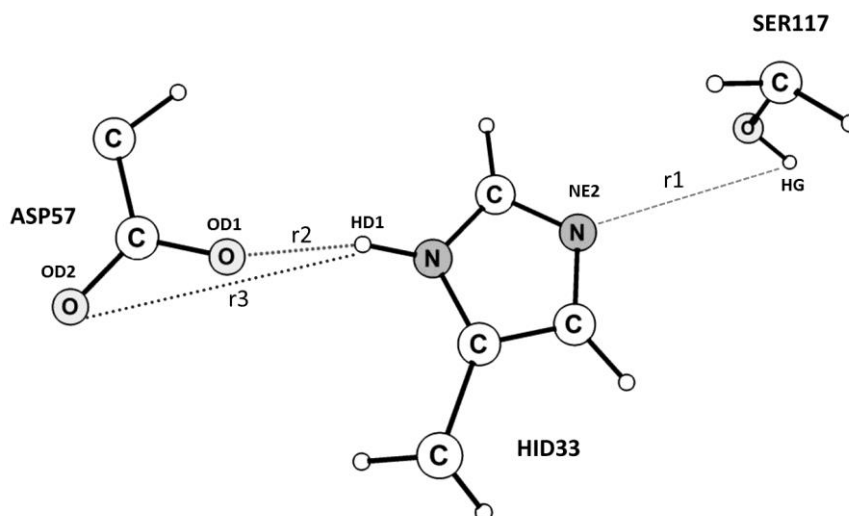
**Figura 43:** Superposição de 700 conformações extraídas em intervalos regulares da simulação de dinâmica molecular NPT da DENV2f. a) superposição da NS3; b) superposição da NS2B. Em Azul tem-se a NS3, e em vermelho a NS2B.

A análise do RMSF (Figura 44) concorda com os dados experimentais em solução (análise de RMN-H) (KIM et al., 2013), em que, os resíduos 1 a 13 (N-terminal), e os resíduos 141 a 152 (C-terminal) da NS3 são flexíveis, assim como os três primeiros resíduos da NS2B (153-156). É possível também observar que os resíduos do sítio ativo (HIS33, ASP57 e SER117) apresentam mobilidade limitada.

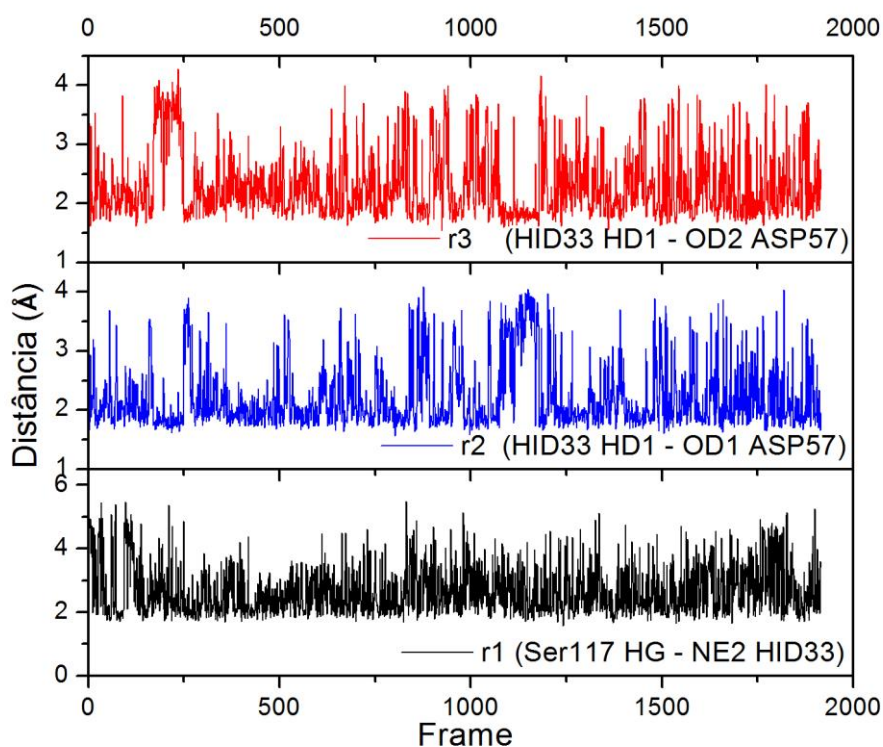
A Figura 45 mostra os resíduos da tríade catalítica, HIS33, ASP57 e SER117, enquanto a Figura 46 mostra o comportamento das ligações de hidrogênio entre estes resíduos (r1, r2, r3 na Figura 45). As interações estão presentes em grande parte da simulação, com distância menor que 2,5 Å em mais que 50% do tempo, fazendo com que os resíduos tenham sua mobilidade diminuída. As distâncias médias durante a simulação foram 2,089 Å, 1,965 Å e 2,017 Å para r1, r2 e r3, respectivamente. A ligação de hidrogênio entre HIS33 e ASP57 se alterna entre os oxigênios da cadeia lateral da ASP57 (OD1 e OD2 na Figura 46), de forma que as médias das distâncias r2 e r3 são basicamente equivalentes.



**Figura 44:** RMSF para os resíduos da DENV2f. Os resíduos HIS33, ASP57 e SER117, do sítio ativo, estão indicados. Onde os resíduos 1-152 são da NS3 e os resíduos 153-192 da NS2B.



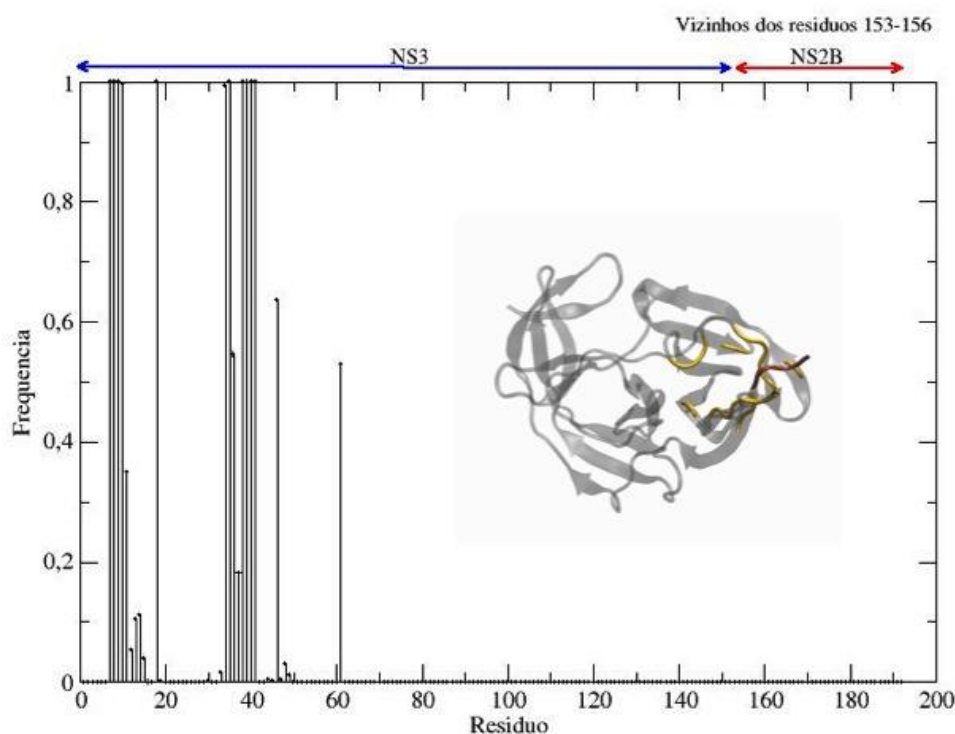
**Figura 45:** Resíduos da tríade catalítica (HIS33, ASP57 e SER117) na DENV2.



**Figura 46:** Distância entre átomos do sítio ativo, conforme ilustrado na **Figura 45**. Preto: r1 (SER117 HG - NE2 HID33); azul: r2 (HIS33 HD1 - OD1 ASP57); Vermelho: r3 (HIS33 HD1 - OD2 ASP57)

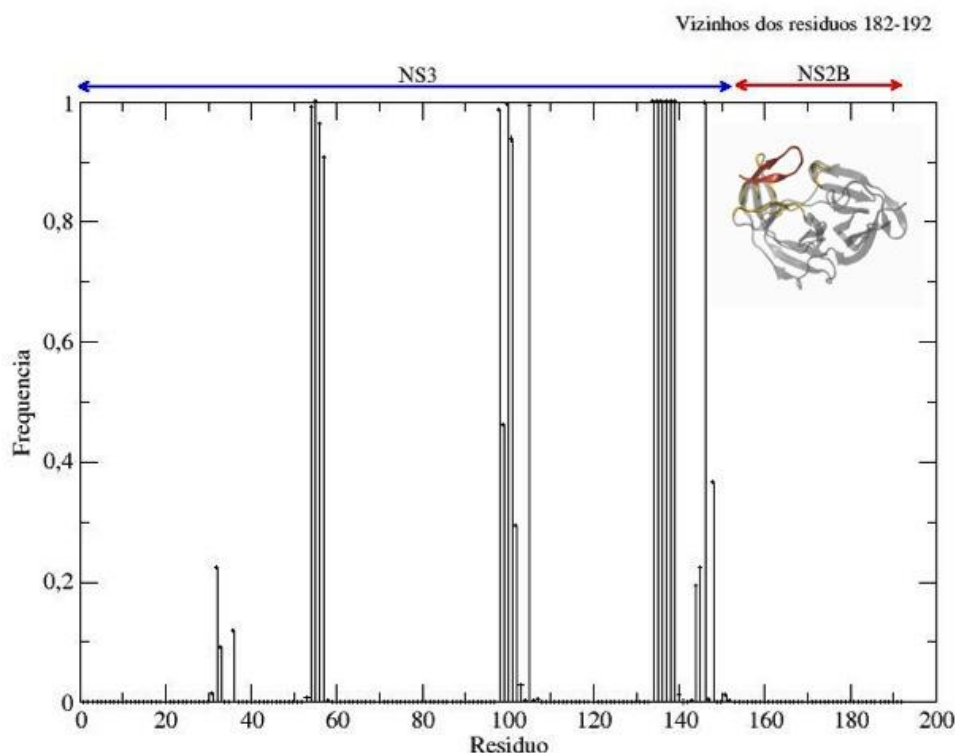
Com relação às possíveis interações dos que os resíduos 153 a 156 (resíduos da NS2B) possam fazer com a NS3, o histograma na Figura 47 mostra os resíduos da NS3 encontrados dentro de um raio de 5.0 Å em torno desses primeiros resíduos, que podem ser visualizados na mesma figura. Os resíduos vizinhos ao N-terminal da NS2B (153-158) que

tiveram contagem maior que 90% neste histograma foram ILE7, LYS8, GLN9, LYS10, ILE18, HIS33 (sítio ativo), ALA34, THR35 E MET41. Algo interessante nesse gráfico é a presença da porção N-terminal da NS2B (153-158) próxima à região do sítio ativo (HIS33) na maior parte do tempo, podendo também interagir com inibidor ou substrato.



**Figura 47:** Resíduos encontrados dentro de um raio de 5.0 Å dos resíduos 153-156. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.

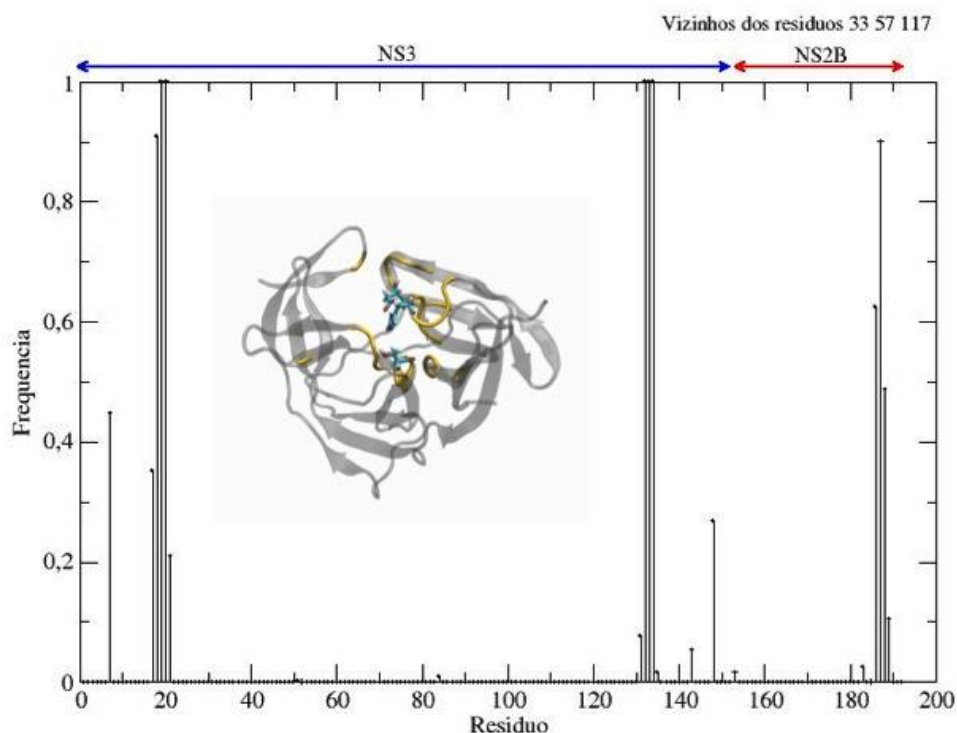
O histograma na Figura 48 mostra os resíduos da NS3 encontrados dentro de um raio de 5.0 Å em torno dos resíduos 182 a 192 da NS2B, que podem ser visualizados na mesma figura. Os resíduos vizinhos ao C-terminal da NS2B (182-192) que tiveram contagem maior que 90% no histograma foram VAL54, LYS55, LYS56, ASP57, LEU97, THR100, ASN101, THR104, ASN134, GLY135, VAL136, VAL137, THR138, ARG139 E ALA146. O ambiente químico dos resíduos 182 a 192 da NS2B estão em conformidade com os dados experimentais para a mesma região no caso do vírus da Dengue tipo 3. (KIM et al., 2013)



**Figura 48:** Resíduos que estão num raio de 5.0 Å dos resíduos 182-192. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.

O histograma na Figura 49 mostra os resíduos encontrados dentro de um raio de 5.0 Å dos resíduos do sítio ativo (HIS33, ASP57 e SER117), que podem ser visualizados na mesma figura. Os resíduos vizinhos ao sítio ativo (HIS33, ASP57 e SER117) que tiveram contagem maior que 90% no histograma foram ILE8, GLY19, ALA20, TYR132, GLY133, ASN134 e GLY187. O resíduo ASN134 forma ligação de hidrogênio com o ASP57 do sítio ativo, e essa ligação de hidrogênio é observada em 40% dos frames da simulação de dinâmica molecular. Os demais resíduos dessa região são importantes, pois essa região irá receber o inibidor, e alguns desses resíduos poderão interagir com os resíduos do inibidor ou substrato.





**Figura 49:** Resíduos que estão num raio de 5.0 Å dos resíduos 33 57 e 117. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.

A Tabela 1 mostra as principais ligações de hidrogênio entre a proteína (NS3) e o cofator (NS2B). Na análise de ligações de hidrogênio se pode escolher o conjunto de resíduos que serão doadores ou aceitadores de ligação de hidrogênio, para melhor observar as interações entre NS3 e NS2B no modelo DENV2f. A primeira parte da Tabela 1 tem os resíduos da NS3 como aceitadores de ligação de hidrogênio e os resíduos da NS2B como doadores, já a segunda parte da Tabela 1 os papéis se invertem.



**Tabela 1:** Dados das ligações de hidrogênio entre a NS3 e a NS2B, onde os resíduos da NS3 estão como aceitadores de ligação de hidrogênio, e os resíduos da NS2B estão como doadores de ligação de hidrogênio.

| <b>Aceitador</b> | <b>Doador</b> | <b>% Tempo</b> | <b>Média da distância (Å)</b> |
|------------------|---------------|----------------|-------------------------------|
| <b>NS3</b>       | <b>NS2B</b>   |                |                               |
| ALA4             | ALA161        | 73             | 2,87                          |
| LYS99            | ILE181        | 65             | 2,87                          |
| ASP2             | ARG160        | 64             | 2,85                          |
| GLN78            | GLU167        | 63             | 2,87                          |
| MET41            | LEU158        | 62             | 2,87                          |
| LYS8             | GLU157        | 57             | 2,88                          |
| GLN92            | SER173        | 51             | 2,88                          |
| PRO95            | SER176        | 49             | 2,72                          |
| ASP2             | ARG160        | 44             | 2,81                          |
| <b>NS2B</b>      | <b>NS3</b>    |                |                               |
| GLU159           | ARG6          | 94             | 2,81                          |
| SER173           | THR93         | 92             | 2,73                          |
| GLU171           | GLN92         | 89             | 2,82                          |
| ALA162           | ALA4          | 85             | 2,84                          |
| GLU157           | LYS8          | 84             | 2,84                          |
| GLU167           | GLN78         | 78             | 2,84                          |
| LEU179           | LYS99         | 64             | 2,87                          |
| LEU156           | MET41         | 62             | 2,87                          |
| SER175           | SER109        | 61             | 2,73                          |
| LYS192           | ARG139        | 46             | 2,80                          |

Os resíduos ALA4, LYS8 e GLU157, MET41 e LEU158 aparecem fazendo ligação de hidrogênio nas duas partes da tabela, mostrando que atuam tanto como doadores de ligação de hidrogênio quanto aceitadores. As ligações de hidrogênio presentes na maior porcentagem de tempo são as que estão na segunda parte da tabela, ou seja, quando os resíduos da NS3 são definidos como doadores de ligação. Algumas ligações de hidrogênio são entre regiões terminais das proteínas como: GLU159-ARG6, ALA162-ALA4, GLU157-LYS8, ALA4-ALA161. Entretanto, não se observa ligações de hidrogênio envolvendo os resíduos do sítio ativo. Todas essas interações apresentadas na Tabela 1 são importantes para manter a

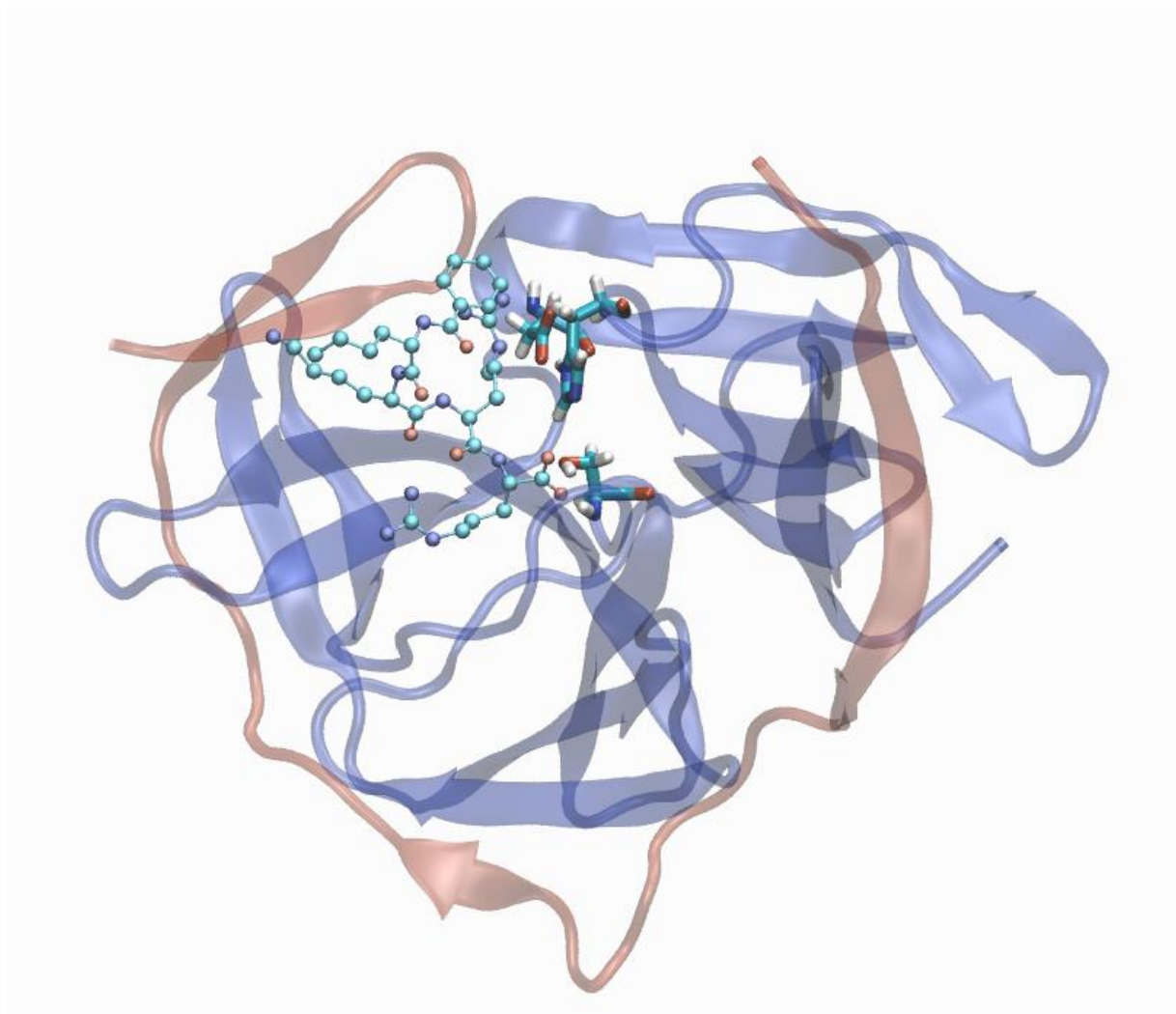
estrutura fechada (ativa), no modelo DENV2f, que se mostrou estável mesmo na ausência do inibidor.

Os resultados revelam que a protease NS2B-NS3 pode existir também na conformação fechada em solução, mesmo na ausência do inibidor. O DENV2f revelou-se um modelo realístico para o estudo de modelagem para triagem de inibidores, assim como para o estudo do mecanismo de reação durante a replicação viral. Portanto, servirá como uma ferramenta valiosa para o desenvolvimento de inibidores para a protease do vírus da Dengue (DENV).

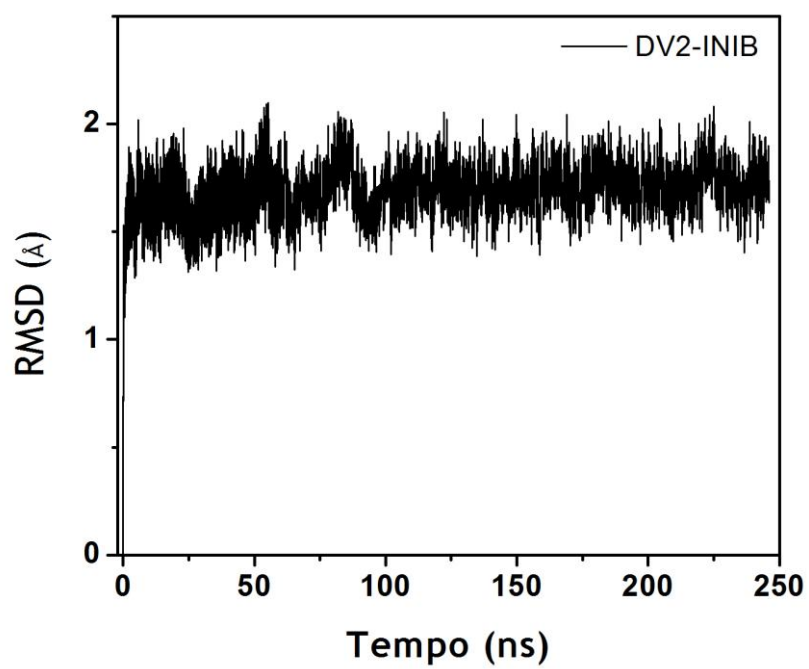
### 5.3 MODELO DENV2f COM INIBIDOR (DENV2f-INIB)

A observação do sistema DENV2f na presença do inibidor (Figura 50) dará uma melhor compreensão das interações que mantêm o inibidor na região do sítio ativo da DENV2f, assim como o entendimento de questões dinâmicas da protease, fundamentais para validação também do modelo DENV2f, já que essa estrutura “fechada” se dá na presença do inibidor (Bz-nKKR-H).

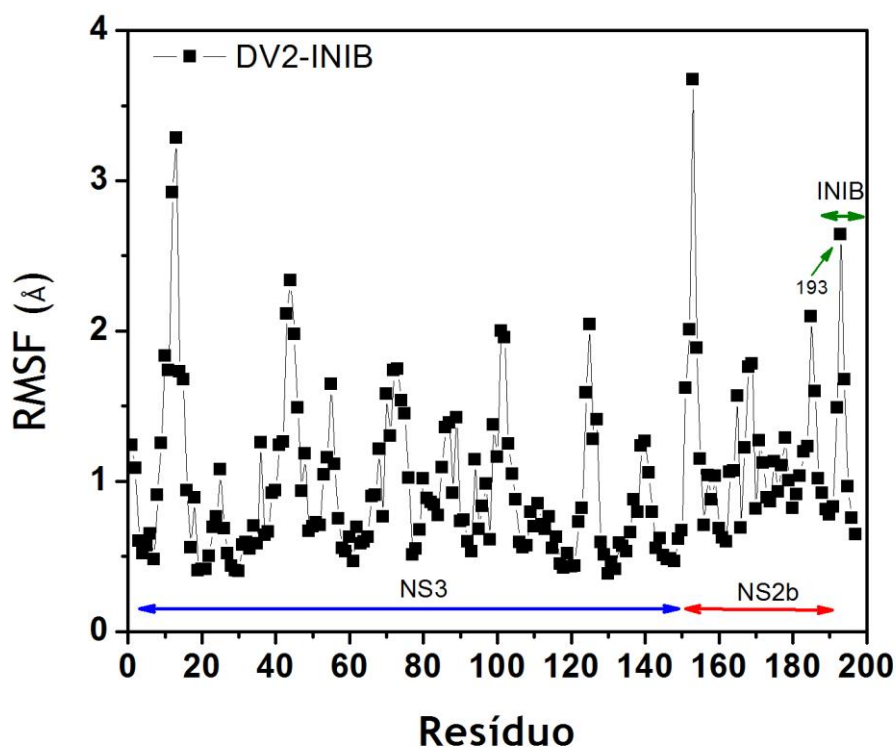
A análise do RMSD (Figura 51) obtido a partir de uma simulação de 246 ns de dinâmica molecular para estrutura da DENV2f na presença do inibidor (Figura 50) revela a estabilidade da estrutura, com RMSD médio de 1,65 Å, em relação a estrutura inicial do modelo, e sem grandes variações do valor em torno dessa média, indicando baixa mobilidade da estrutura durante esse tempo de simulação.



**Figura 50:** Uma conformação do sistema DENV2f-INIB extraída da simulação de dinâmica molecular com ensemble NPT. As moléculas de água da caixa de simulação não estão representadas na figura para facilitar a visualização. Em Azul NS3, vermelho a NS2B, representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117), e a representação em bola são do inibidor sem os hidrogênios para a melhor visualização.



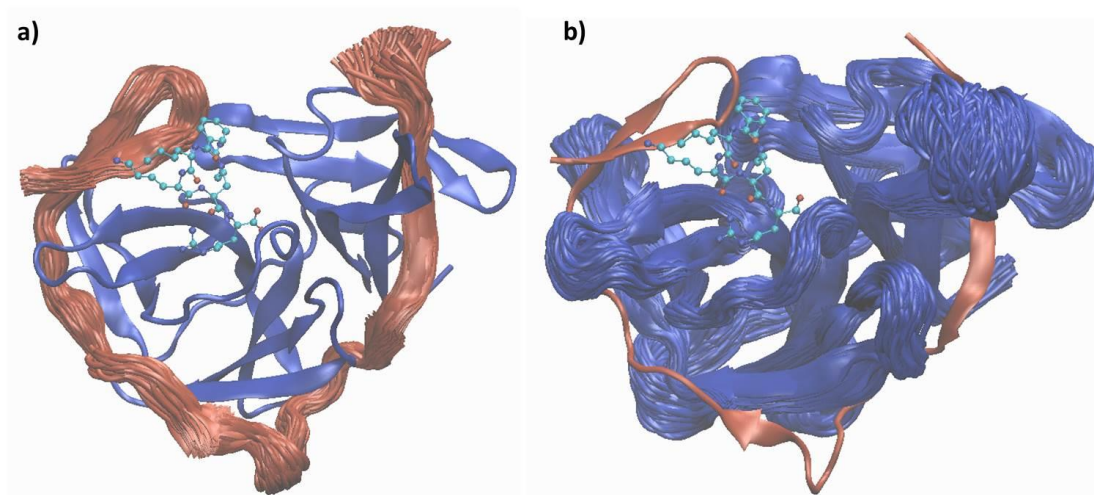
**Figura 51:** RMSD para o sistema DENV2f-INIB, tendo como referência a estrutura inicial para toda as etapas de simulação.



**Figura 52:** RMSF para os resíduos da DENV2f-INIB, região delimitada com seta azul são os resíduos da NS3, a seta em vermelho são os resíduos da NS2B, já a seta em verde delimita os resíduos do inibidor. Onde os resíduos 1-152 são da NS3 e os resíduos 153-192 da NS2B.

A Figura 52 mostra o RMSF para o sistema. Os resíduos que apresentam maior flutuação na proteína são os mesmos observados no modelo DENV2f sem inibidor. Já os resíduos do C-terminal e N-terminal da NS2B apresentam menor amplitude do RMSF na presença do inibidor quando comparados com DENV2f sem inibidor, assim, menor mobilidade. Esta mudança é devida a interações do inibidor com essas extremidades da NS2B. Na região do inibidor na Figura 52 (resíduos 193 a 197), o resíduo 193 que é terminal apresenta maior mobilidade, e interage menos com a proteína.

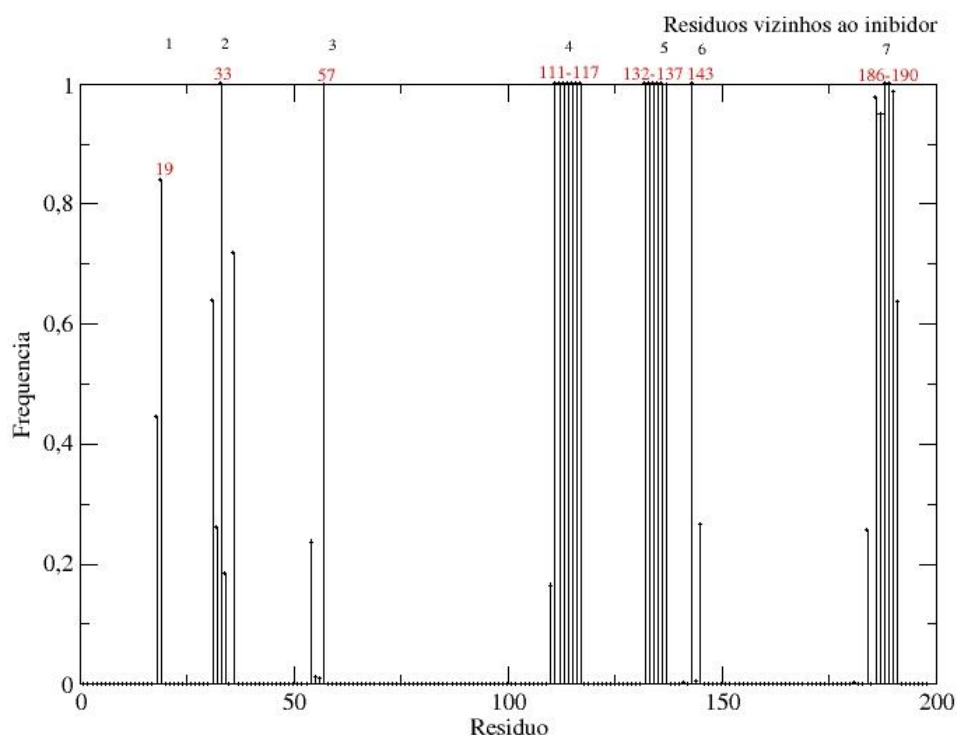
A Figura 53 mostra a superposição de 700 frames da simulação, permitindo analisar de maneira visual a mobilidade da cadeia principal de cada porção da proteína, NS3 e NS2B. Na presença do inibidor a estrutura se apresenta com pouca variação da posição dos átomos da cadeia principal tanto para NS3 quanto para a NS2B.



**Figura 53:** Superposição para as 700 conformações extraídas em intervalos regulares dos últimos 150 ns da simulação de dinâmica molecular NPT do sistema DENV2f-INIB. a) superposição da NS3; b) superposição da NS2B. A região da NS3 (azul), em vermelho a NS2B e representado com bolas o inibidor.

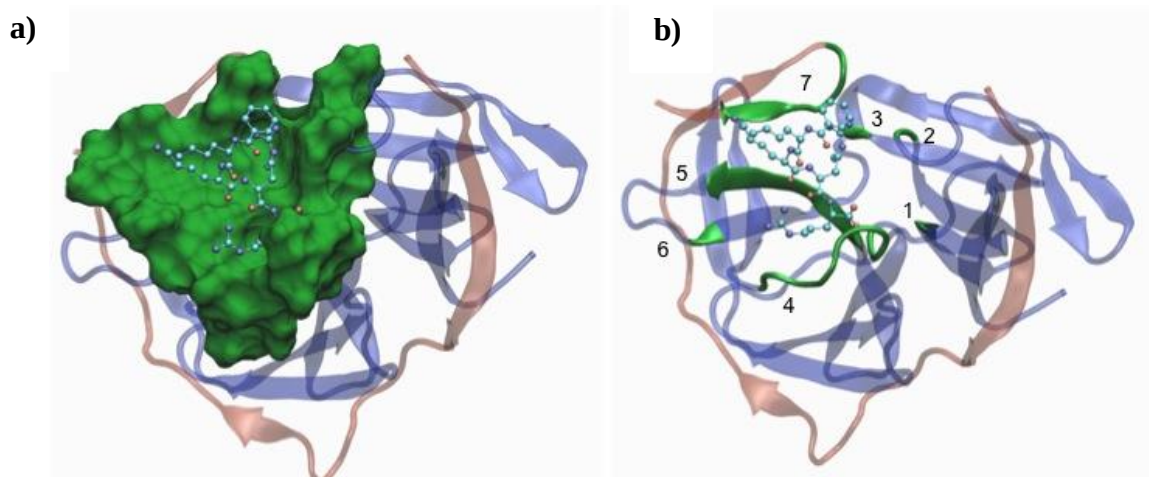
Com relação às possíveis interações que inibidor possa fazer com a protease, o histograma na Figura 54 mostra os resíduos encontrados dentro de um raio de 5.0 Å em torno do inibidor. Os resíduos da tríade catalítica (HIS33, ASP57 e SER117) aparecem dentro desse raio durante todo o tempo de simulação, assim como resíduos terminais da NS2B (186-190). Essas interações mantêm o inibidor aprisionado ao sítio ativo da DENV2f.

Na Figura 55 se observa a região dos resíduos 186 a 190 (região 7 na Figura 55b) da NS2B próxima ao inibidor, que está interagindo com alguns resíduos do inibidor. Nenhuma ligação de hidrogênio estável foi localizada entre esses resíduos da NS2B e o inibidor (Tabela 4), porém eles podem interagir através de interações do tipo eletrostáticas ou de van der Waals. Essa interação é confirmada através de dados experimentais mostrando também esse tipo de interação do inibidor com os resíduos da NS2B.(NOBLE et al., 2012)



**Figura 54:** Resíduos que estão num raio de 5.0 Å dos resíduos do inibidor. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.

Observando-se a região 4 na Figura 55 (resíduos 111 a 117) e comparando com os dados da Tabela 2, identifica-se que essa região é a que mais faz ligações de hidrogênio com o inibidor. Como já mencionado anteriormente, o resíduo SER117 faz parte do sítio ativo, e o inibidor além de estar próximo a esse resíduo, também forma ligação de hidrogênio com mesmo durante a simulação. Esse resíduo é fundamental para a quebra da ligação peptídica durante a replicação viral, atuando no ataque nucleofílico ao substrato.(HEDSTROM, 2002b)



**Figura 55:** Ilustração dos resíduos que estão num raio de 5.0 Å dos resíduos do inibidor. a) os resíduos estão representados por uma superfície, com o inibidor sobre essa superfície. b) os resíduos estão representados em verde na imagem, e a numeração se refere à região que está representada no gráfico da **Figura 54**. Representado em azul a NS3, em vermelho a NS2B, em bolas o inibidor, e em verde os resíduos que estão num raio de 5.0 Å do inibidor. No inibidor não há a representação dos hidrogênios para a melhor visualização.

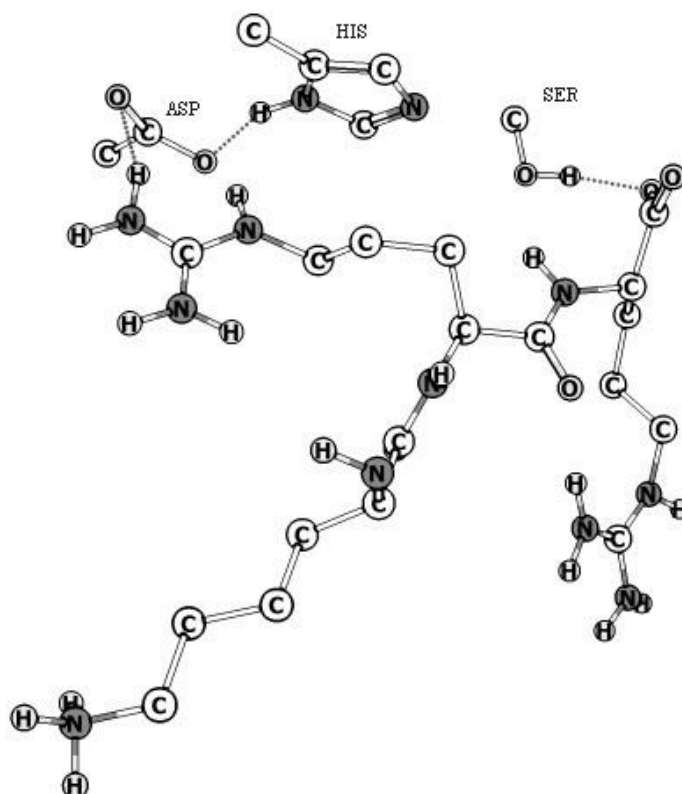
Os resíduos da DENV2f que formam ligações de hidrogênio com inibidor são ASP57, ASP111, GLY115, SER117, GLY135 e TYR143, (Tabela 2). A Figura 56 mostra esses resíduos que fazem ligação de hidrogênio com os resíduos 195, 196 e 197 do inibidor. Na Figura 56 tem-se o sítio ativo próximo ao inibidor, mostrando que o ASP57 a SER117 formam ligação de hidrogênio com inibidor, enquanto, a HIS33 forma ligação de hidrogênio com o ASP57.

As simulações também mostram a formação da cavidade oxianion pelos resíduos GLY135, TYR143 e SER117 (átomos de N e H desses resíduos) interagindo com inibidor, em concordância com dados experimentais (NOBLE et al., 2012). Essa interação é importante para a estabilização do intermediário tetraédrico formado durante a primeira etapa da reação de replicação viral.(HEDSTROM, 2002b)

**Tabela 2:** Ligações de hidrogênio entre a DENV2 e o inibidor. Resíduos do inibidor estão destacados em negrito.

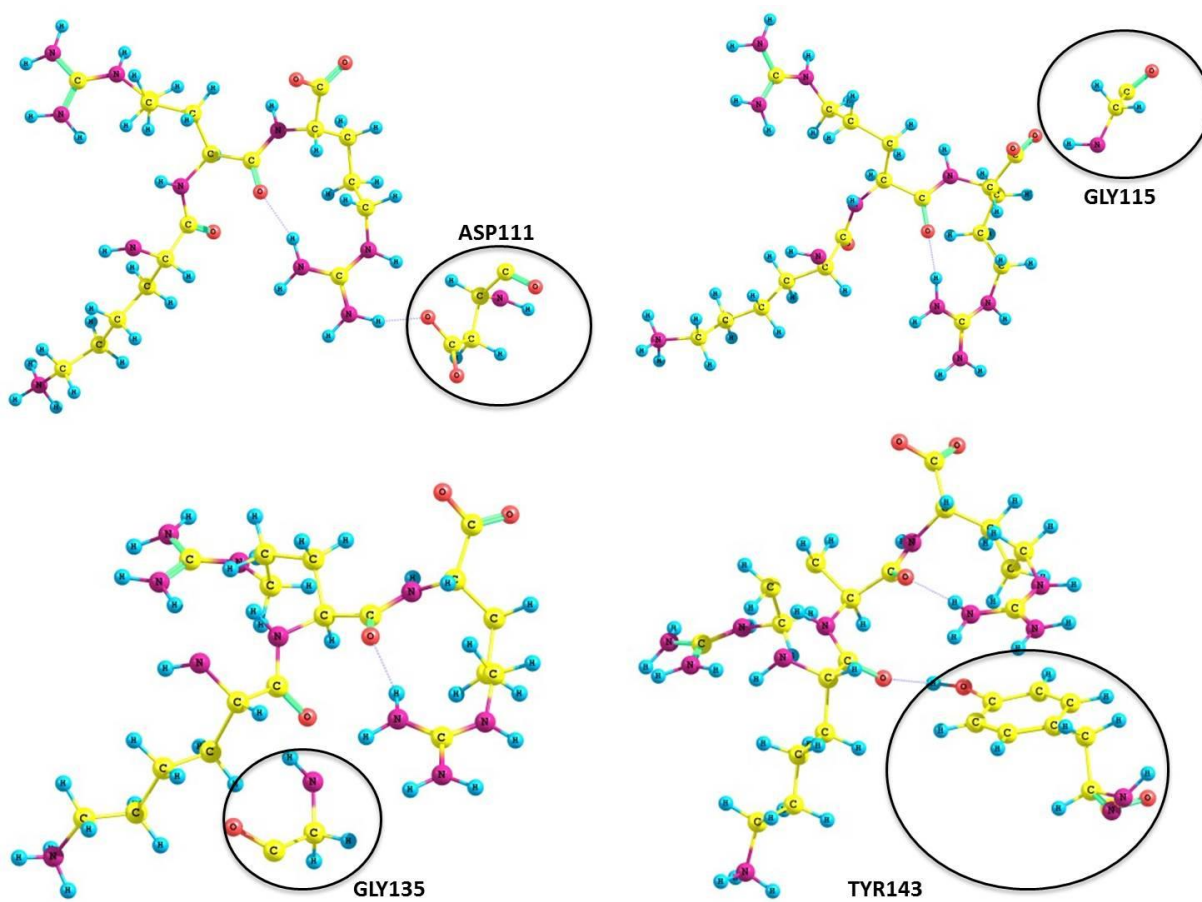
| Doador        | Aceitador     | % Tempo | Média da distância (Å) |
|---------------|---------------|---------|------------------------|
| SER117        | <b>ARG197</b> | 64      | 2,62                   |
| <b>ARG197</b> | ASP111        | 53      | 2,79                   |
| SER117        | <b>ARG197</b> | 44      | 2,85                   |
| GLY135        | <b>LYS195</b> | 43      | 2,88                   |
| GLY115        | <b>ARG197</b> | 41      | 2,86                   |
| <b>ARG196</b> | ASP57         | 36,5    | 2,80                   |
| TYR143        | <b>LYS195</b> | 36,2    | 2,79                   |



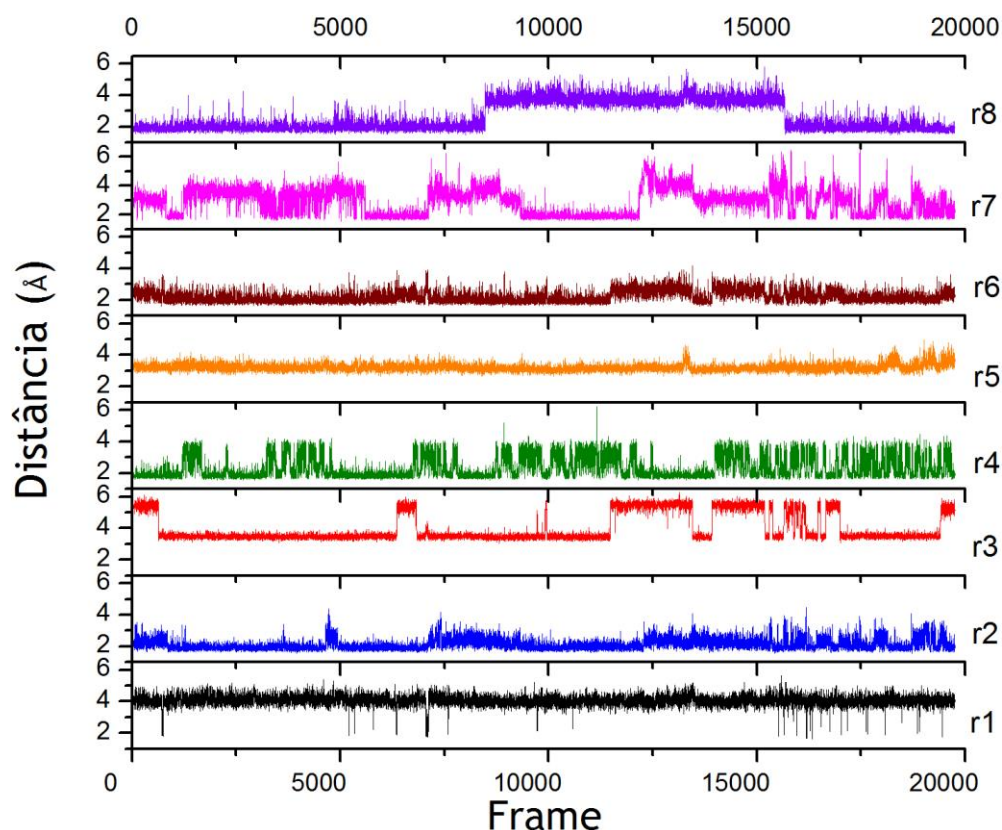


**Figura 56:** Ilustração do inibidor com a cadeia lateral dos resíduos do sítio ativo (HIS33, ASP57 e SER117). O inibidor só está representado nessa imagem pelos resíduos 195, 196 e 197, que são os resíduos da tabela 3. Os postos tracejados se referem a ligações de hidrogênio entre resíduos do inibidor e os indicados na figura da proteína. Os átomos de hidrogênio ligados a átomos de carbono foram retirados para melhorar a visualização.

Ao monitorar as distâncias das ligações de hidrogênio (Tabela 2) durante a simulação (Figura 58), observa-se que as ligações de hidrogênio que apresentam maior contagem durante a simulação para valores de distância menores que 2,5 Å foram as ASP111–ARG197 (r4), ARG197–GLY115 (r7), ASP57–LYS196 (r6) e ARG195–TYR143 (r8). Monitorou-se também as ligações de hidrogênio entre os resíduos do sítio ativo, neste caso SER117–HIS33 (r1) e ASP57–HIS33 (r2). A ligação de hidrogênio ASP57–LYS196 (r6) e a ASP57–HIS33 (r2) são as que apresentam maior estabilidade dessa ligação durante a dinâmica molecular, ou seja, o resíduo ASP57 (carga negativa) tem sua carga estabilizada por essas interações, além de manter o inibidor mais próximo ao sítio ativo. Tudo isso mostra a importância dos resíduos GLY115, TYR143, ASP57 e ASP111 (Figura 57) para a estabilização do inibidor na região do sítio de reação.

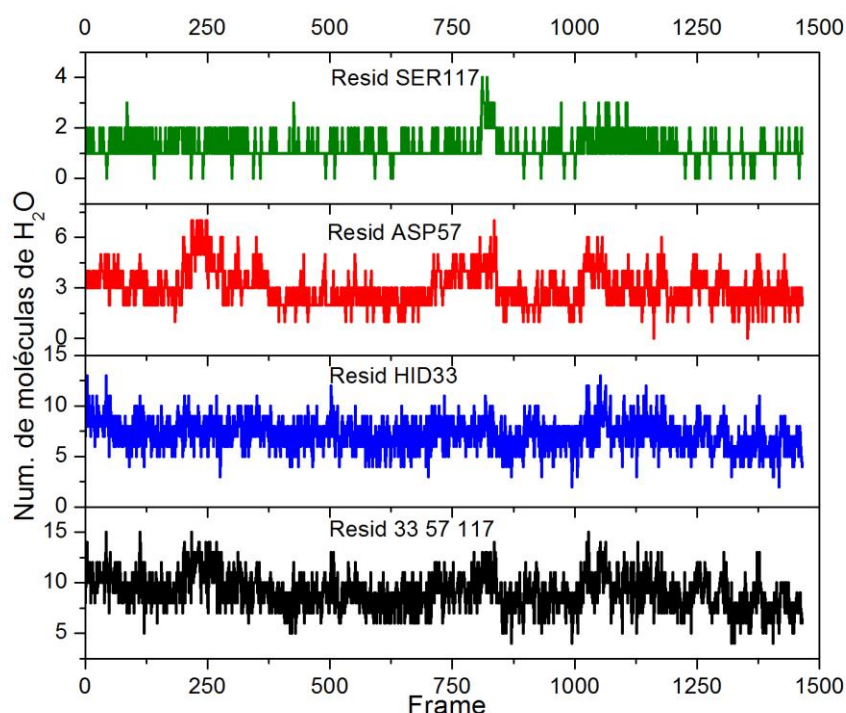


**Figura 57:** Ilustração do inibidor com os resíduos ASP111, GLY115, GLY135 e TYR143, todos estão indicados na figura. O inibidor só está representado nessa imagem pelos resíduos 195, 196 e 197, que são os resíduos da **Tabela 2**. Representações em azul são os hidrogênios, em amarelo os carbonos, em rosa os nitrogênios, e em vermelho, os oxigênios. Os pontos tracejados se referem a ligações de hidrogênio entre resíduos do inibidor e os indicados na figura da proteína.



**Figura 58:** Distâncias de ligação de hidrogênio versus os frames da simulação. As distâncias r1 (SER117-HIS33); r2 (ASP57-HIS33); r3 (197-SER117); r4 (ARG195-GLY135); r5 (ARG195-GLY135); r6 (ARG197-GLY115); r7 (ASP57-LYS196) e r8 (ARG195-TYR:143). Esses frames são referentes aos 150 ns finais da dinâmica molecular.

Algo que deve ser avaliado durante a simulação é o número médio de moléculas de água próximas ao sítio ativo (HIS33, ASP57 e SER117), pois a água pode participar de uma das fases de reação da quebra da ligação peptídica do substrato.(HEDSTROM, 2002b) Analisou-se também o número de moléculas de água próximas a cada resíduo do sítio ativo separadamente, avaliando assim qual dos resíduos interagem mais com moléculas de água. A Figura 59 mostra o número de moléculas de água encontradas dentro de um raio de 5,0 Å em torno dos resíduos da tríade catalítica durante a simulação. Pode-se ver que existe um grande número de moléculas de água próximas ao sítio ativo (média de 9 moléculas de H<sub>2</sub>O), sendo o resíduo HIS33 (média de 7 moléculas de H<sub>2</sub>O) que apresenta maior quantidade quando comparada com os demais resíduos do sítio ativo, ASP57 e SER117, com em média 4 e 1 moléculas de água, respectivamente. A presença de água nas proximidades, que pode participar em ligações de hidrogênio, explica a ausência desse tipo de interação entre o HIS33 e o inibidor, apesar da proximidade.



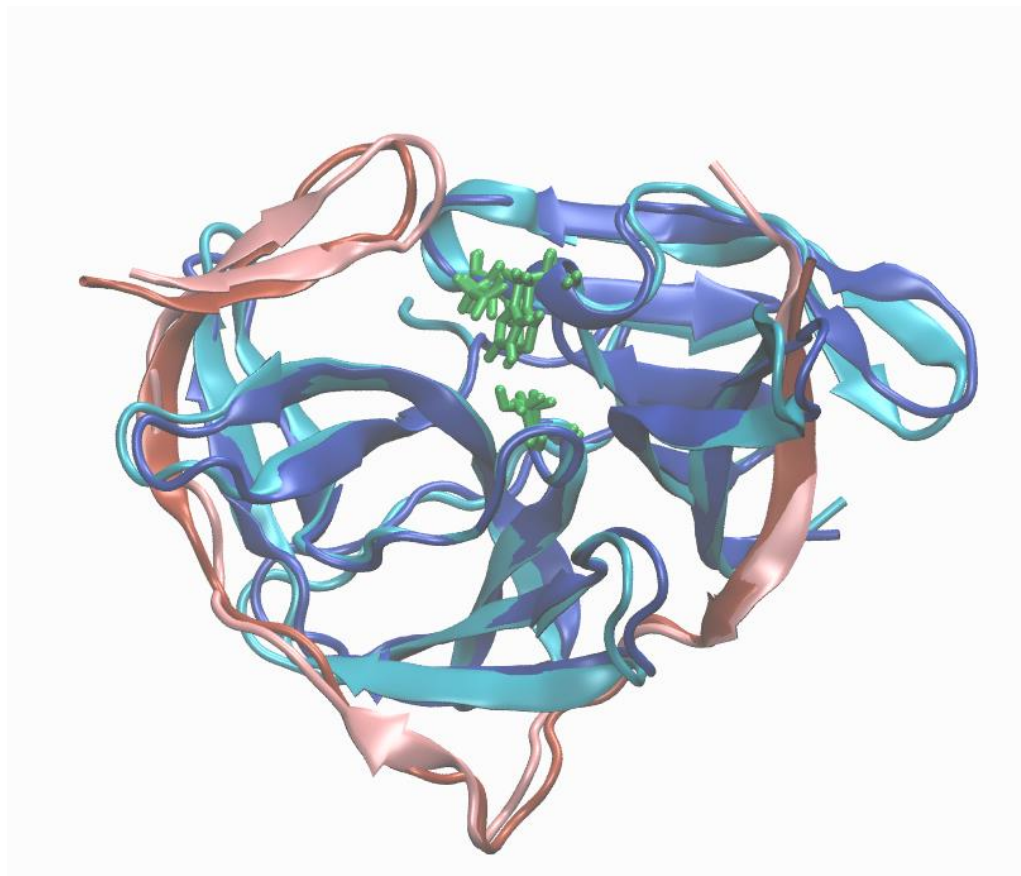
**Figura 59:** Número de moléculas de H<sub>2</sub>O que estão a uma distância de 5.0 Å dos resíduos do sítio ativo. Em preto são as águas que estão próximas a tríade do sítio ativo, em azul ao resíduo HIS33, em vermelho ao resíduo ASP57 e verde ao resíduo SER117. Esses frames são referentes aos 150 ns finais da dinâmica molecular

A cadeia lateral do resíduo ARG197 (inibidor) forma uma ponte salina com os resíduos ASP57 da NS2B e ASP111 da NS3. Essas duas interações colaboram para manter o NS2B-NS3 na conformação fechada, com NS2B próximo ao sítio ativo, em acordo com os dados experimentais. (NOBLE et al., 2012)

Os resultados dessa simulação confirmam a importância da região de grampo- $\beta$  da NS2B para a atividade, devido à interação com inibidor. (NOBLE et al., 2012) Essa região tinha sido proposta como um potencial local de alvo para inibidores. (CHAPPELL et al., 2008) Além disso, as simulações evidenciam a importância de vários resíduos da NS3 para a inibição, e mostram as interações entre os resíduos do sítio ativo, e as moléculas de água presentes na região reacional da protease.

Um dado experimental de extrema importância para a validação do modelo DENV2f foi publicado em meados de 2012, apresentado a estrutura cristalina da NS2B-NS3 do vírus da Dengue (sorotipo 3) com inibidor (Bz-NLE-ARG-ARG-LYS-H), (PDB ID: 3U1I). (NOBLE et al., 2012) A Figura 60 mostra a superposição da última estrutura da simulação de dinâmica molecular do modelo DENV2f com a estrutura cristalográfica da NS2B-NS3 com inibidor para o vírus da Dengue (PDB ID: 3U1I). (NOBLE et al., 2012) Fica

claro que o modelo DENV2f está em concordância com a estrutura ativa da NS2B-NS3 protease do vírus da Dengue.



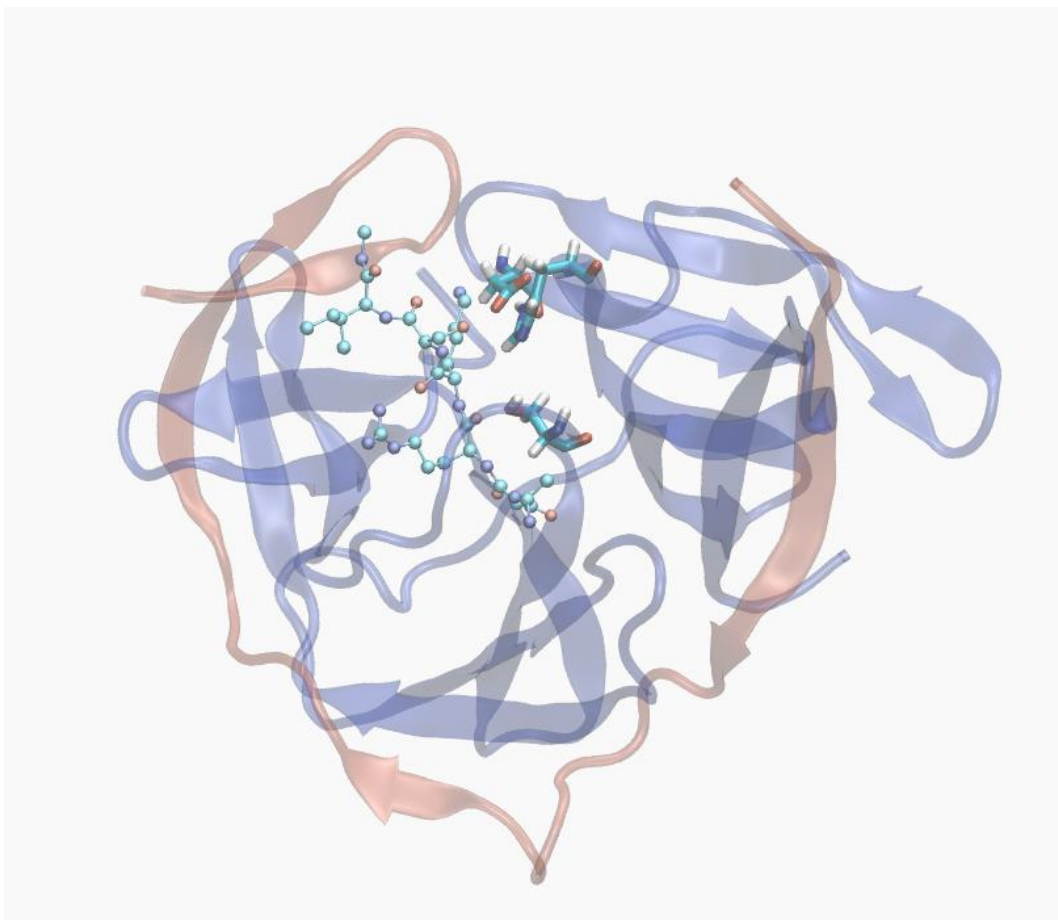
**Figura 60:** Superposição da última estrutura da simulação de dinâmica molecular da DENV2 com a estrutura cristalográfica da NS2B-NS3 (PDB ID: 3U1I). Em azul, a NS3 da (PDB ID: 3U1I); azul claro a NS3 da DENV2; em vermelho a NS2B da (PDB ID: 3U1I); em rosa a NS2B da DENV2; verde os resíduos do sítio ativo para as duas estruturas.

Quando fizemos o RMSD da simulação do modelo DENV2f usando como referência a estrutura cristalográfica do PDB ID: 3U1I, o valor foi de 1,3 Å, evidenciando que o modelo utilizado durante o estudo é apropriado para estudar a replicação viral (Figura 60), bem como, varredura de inibidores para a protease.

#### 5.4 MODELO DENV2f COM O SUBSTRATO 1 (DENV2f-SUBS1)

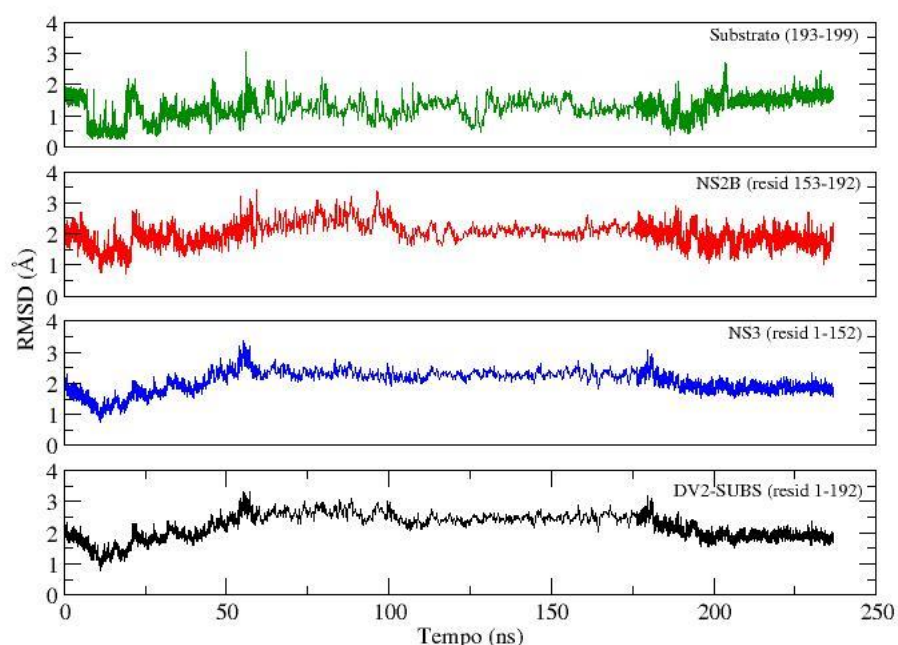
A observação do sistema DENV2f na presença do substrato (ACE-ALA-GLY-ARG-LYS-SER-ILE-NME) (Figura 61) dará uma melhor compreensão das interações que mantém o substrato na região do sítio ativo da DENV2f, assim como a análise dos resíduos que são importantes para a reação de quebra do substrato, já que na próxima etapa do trabalho será necessário saber o ambiente reacional para a construção do modelo reacional QM/MM.





**Figura 61:** Uma conformação da DENV2f-SUBS retirada da simulação de dinâmica molecular num *ensemble* NPT. As moléculas de água da caixa de simulação não estão representadas na figura para melhor visualização. Em Azul NS3, vermelho a NS2B, representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117), e a representação em bola são do inibidor sem os hidrogênios para a melhor visualização.

A Figura 62 mostra o RMSD obtido para a simulação de 237 ns de dinâmica molecular para estrutura da DENV2f-SUBS1. Nos primeiros 8 ns de simulação, em que o substrato estava preso por uma restrição com constante de força de  $200 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ , o seu RMSD (gráfico verde da Figura 62) se mantém numa região com pouca oscilação, como esperado. Retirada a restrição, o substrato se estabiliza oscilando uma região próxima a  $2,0 \text{ \AA}$ . Já o comportamento do DENV2f, assim como das porções NS3 e NS2B é de um RMSD maior, mas estabilizado depois dos 50 ns de dinâmica molecular. Na Figura 62, os gráficos em preto, azul e vermelho tem o mesmo comportamento, evidenciando que as duas porções da protease contribuem da mesma maneira para um RMSD médio de  $1,95 \text{ \AA}$ , mostrando que houve uma variação maior com relação a estrutura de partida.

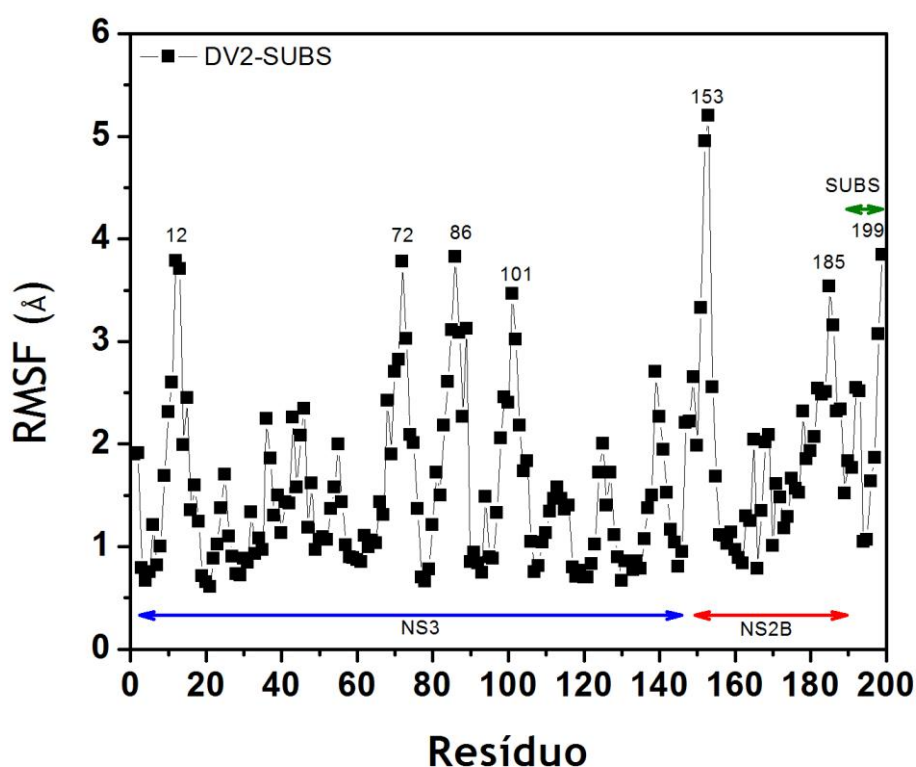


**Figura 62:** RMSD para DENV2f (preto), assim como, o RMSD só para a NS3 resíduos 1 a 152 em azul, NS2B resíduos 153 a 192 em vermelho e para o substrato em verde.

Observa-se que também existe uma maior variação em torno da média do RMSD quando comparamos com o caso DENV2f-INIB (Figura 51, Página 103), mostrando uma maior mobilidade da protease, assim como também do substrato (gráfico verde da Figura 62). O tamanho do substrato pode ser responsável por essa maior mobilidade da DENV2f, já que esta precisa se ajustar a uma molécula muito maior (7 resíduos) do que no caso anterior (DENV2f-INIB, apenas 5 resíduos).

Analisando o gráfico de RMSF, na Figura 63, observa-se que os resíduos 12, 72, 86, 101, 153 e 185 da protease apresentam maior mobilidade, além dos resíduos 198 e 199 do substrato. Os resíduos 72, 86 e 101 não apresentavam essa mobilidade no caso da DENV2f-INIB (Figura 52, Página 104). Entretanto, esses resíduos estão em regiões da proteína que não apresentam estrutura, assim é natural que resíduos dessa região sejam mais móveis. Essa maior mobilidade já tinha sido observada pelo RMSD da proteína. Os resíduos do sítio ativo apresentam os mesmos valores para RMSF (Figura 63) quando comparados com os obtidos para o DENV2f-INIB (Figura 52). Os resíduos 72, 86, 101 e 185 ganharam mobilidade, mostrando que o substrato não interage com esses resíduos como o inibidor. No sistema livre, DENV2f (Figura 44, Página 95), esses resíduos também apresentam menor mobilidade,

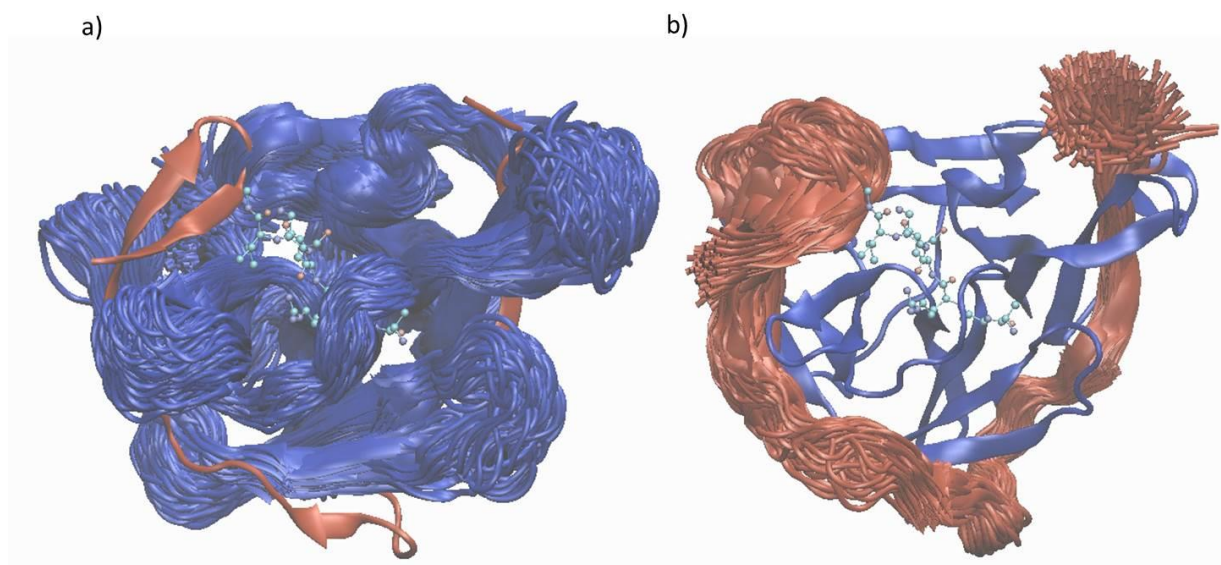
indicando que a presença do substrato gera uma perturbação na região, fazendo com que eles fiquem mais flexíveis. Já os resíduos das regiões terminais como 1-12, 150-152, 153-156 e 190-192 (Figura 63) se comportam da mesma maneira como no sistema DENV2f (Figura 44, Página 95). Entretanto, o caso DENV2f-INIB (Figura 52) essa mobilidade das extremidades é diminuída com relação ao DENV2f (Figura 44, Página 95) livre.



**Figura 63:** RMSF para os resíduos da DENV2f-SUBS1, região delimitada com seta azul são os resíduos da NS3, a seta em vermelho são os resíduos da NS2B, já a seta em verde delimita os resíduos do substrato. A numeração é para os resíduos que apresentam maior mobilidade.

A Figura 64 mostra que tanto a NS3 quando a NS2B oscilou mais em suas posições do que no caso DENV2f-INIB (Figura 53, Página 105), como dito anteriormente, o número maior de resíduos no substrato faz com que a proteína tenha que se ajustar mais, provocando uma maior mobilidade dos seus átomos. Algo interessante é a região grampo- $\beta$  da NS2B que oscila bastante quando comparamos com o caso DENV2f-INIB, e sabe-se que no caso do inibidor com poucos resíduos, essa região da proteína interage tanto com inibidor quanto com a região do sítio ativo.

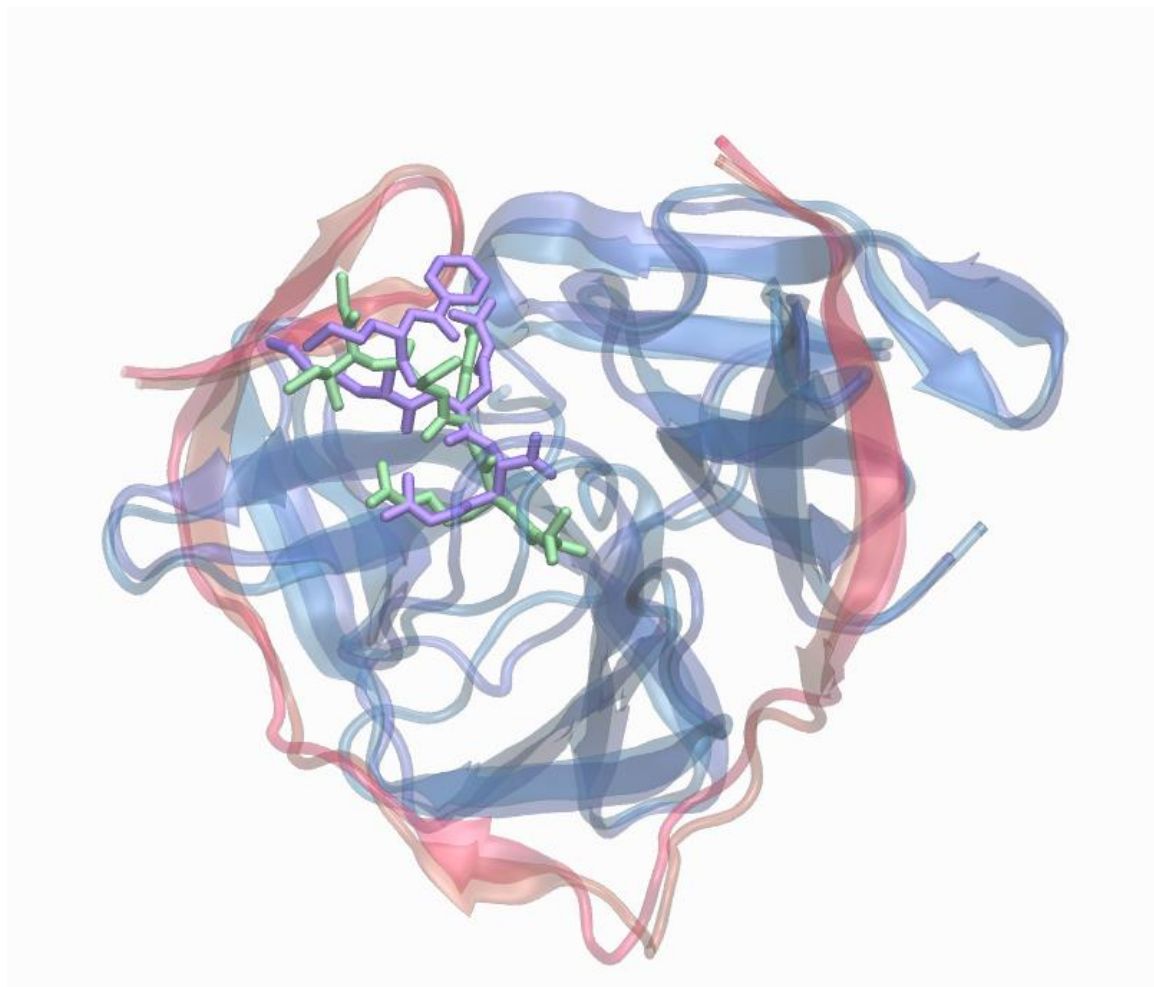




**Figura 64:** Ilustração da superposição para as 700 conformações retiradas da simulação de dinâmica molecular NPT da DENV2f-SUBS. a) superposição da NS3; b) superposição da NS2B. A região da NS3 (azul), em vermelho a NS2B e representado com bolas o substrato.

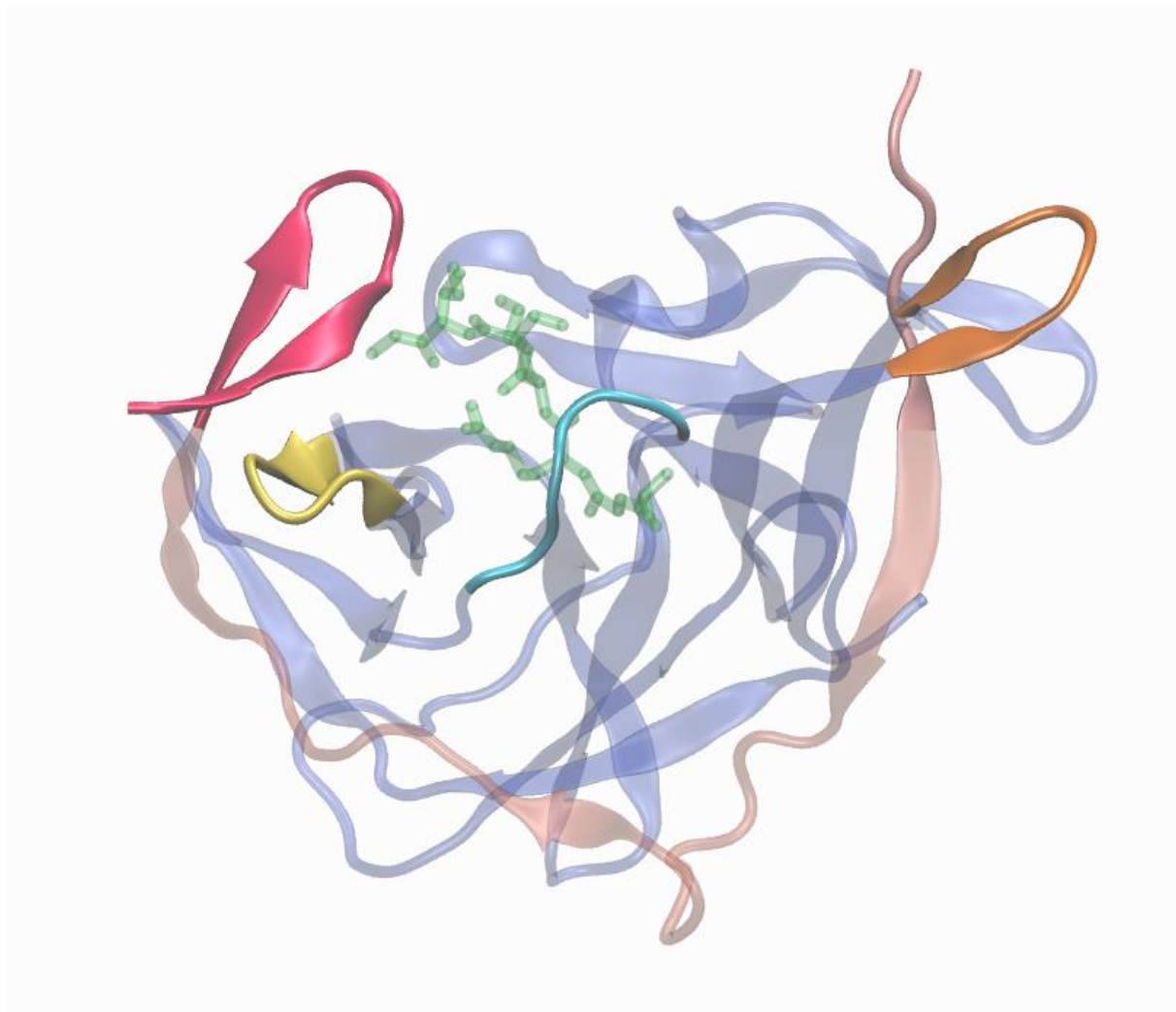
Na Figura 65 em que as imagens da estrutura DENV2f-INIB e DENV2f-SUBS1 estão superpostas é possível ver a diferença de tamanho entre os peptídeos do inibidor e substrato. Parte do substrato fica sob uma região flexível da proteína (resíduos 112-116), prendendo o substrato nessa região por efeitos estéricos.

Para tirar algumas dúvidas quanto à mobilidade da proteína, analisamos separadamente o RMSD de regiões diferentes, para que ficasse claro se havia ou não uma região da proteína com mobilidade muito maior, ou era um fenômeno da proteína com todo. Escolhemos, a partir da análise dos resíduos que apresentaram maior mobilidade pelo gráfico do RMSF (Figura 63) as regiões que são mostradas na Figura 66, incluindo regiões da NS3 e da NS2B. Essas regiões são os resíduos terminais da NS2B, e resíduos da NS3 que interagem com os resíduos terminais da NS2B (Figura 66).

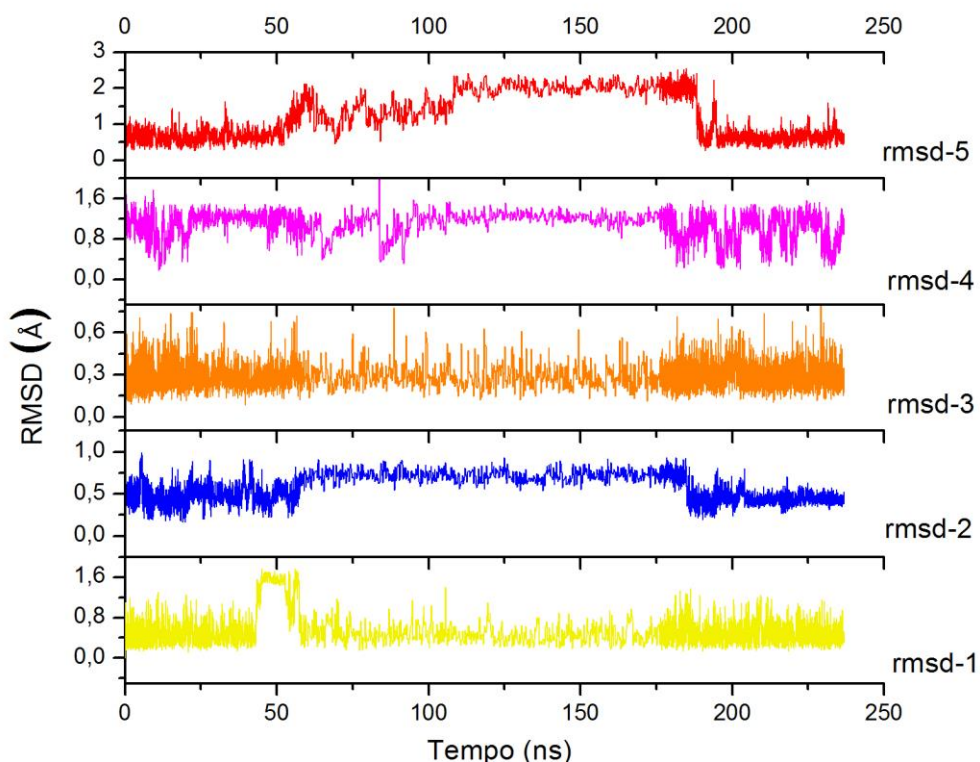


**Figura 65:** Ilustração da superposição das estruturas da DENV2f-INIB e DENV2f-SUBS, onde o inibidor estar representado em lilás, e o substrato em verde. NS3 em azul escuro parte da DENV2f-INIB; NS3 em azul claro parte da DENV2f-SUBS; NS2B em vermelho da estrutura DENV2f-INIB; NS2B em laranja da estrutura DENV2f-SUBS.

O RMSD (Figura 67) das regiões amarela e laranja da Figura 66 apresentaram maior variação em torno da média, variação de 1,3 Å e 1.1Å, respectivamente. Portanto, foram que mais oscilaram, o que a princípio não se esperaria, pois no caso do DENV2f-INIB isso não é observado, e também pela região ser bem estruturada. Como visto na Figura 66, essas duas porções ficam próximas ao sítio de reação. A região em vermelho apresentou uma variação em seu RMSD de 2,0 Å (Figura 67), essa região fica próxima à região em laranja (Figura 66). Isso tudo mostra que a presença de uma molécula maior provoca um comportamento diferente quando comparado a uma molécula menor.

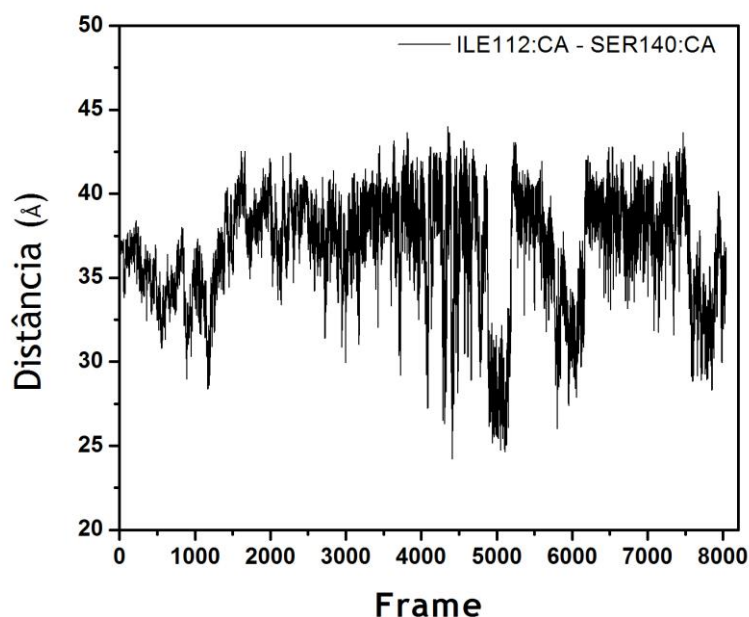


**Figura 66:** Ilustração para algumas regiões da proteína. Em amarelo os resíduos 10 a 14; azul os resíduos 112 a 116; laranja os resíduos 138 a 141; rosa resíduos 153 a 158 e rosa escuro os resíduos 182 a 192.



**Figura 67:** RMSD para algumas regiões da proteína (**Figura 66**). RMSD: amarelo para os resíduos 10 a 14; azul os resíduos 112 a 116; laranja os resíduos 138 a 141; rosa resíduos 153 a 158 e vermelho os resíduos 182 a 192.

Ao analisar a Figura 66, nota-se que os resíduos 10 a 14 (amarelo) e os resíduos 138 a 141 (laranja) se posicionam de maneira oposta um ao outro. Assim, avaliamos a distância entre dois resíduos, sendo um o resíduo ILE12 e outro o resíduo SER140, na Figura 68.

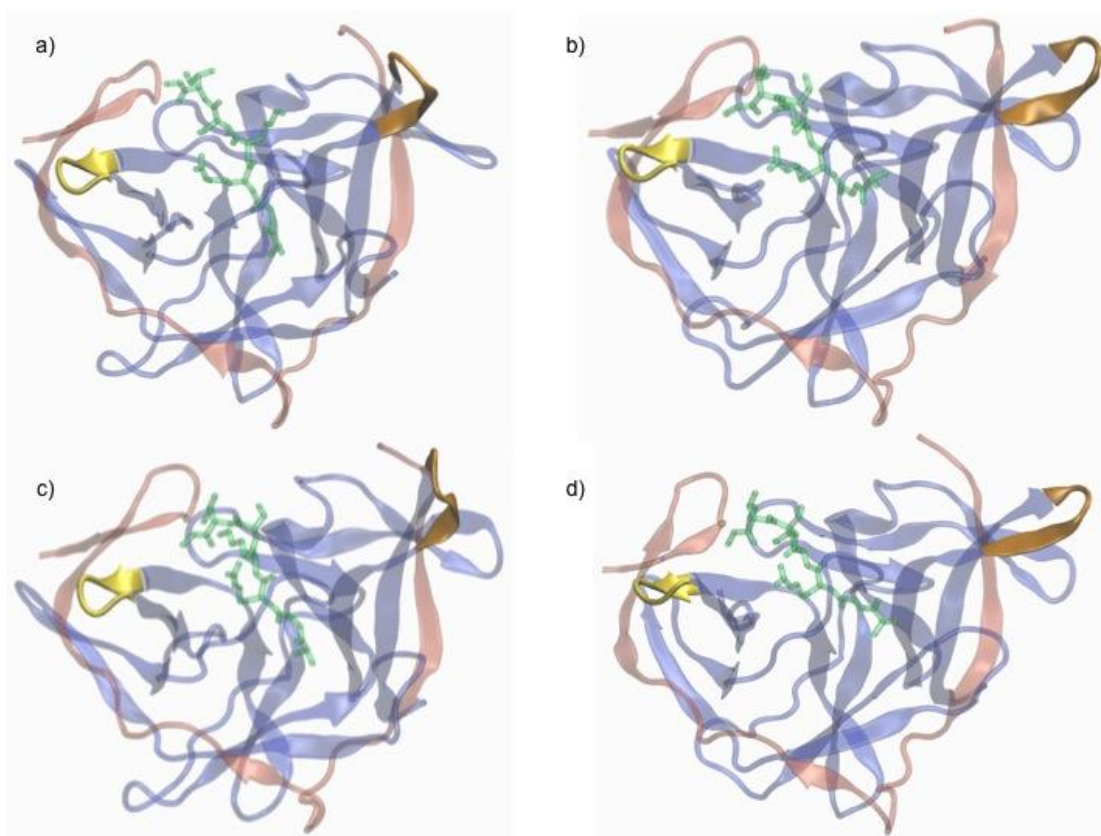


**Figura 68:** Distância entre os resíduos ILE12 e SER140.

A distância (Figura 68) entre os resíduos ILE12 (amarelo na Figura 66) e SER140 (laranja na Figura 66) apresentou variação de até 21,0 Å. Isso indica que durante a dinâmica molecular essas regiões fizeram um movimento de abrir e fechar. Como as duas regiões ficam próximas a região de acesso ao sítio ativo, uma molécula maior pode ter maior dificuldade de aproximação ao sítio ativo.

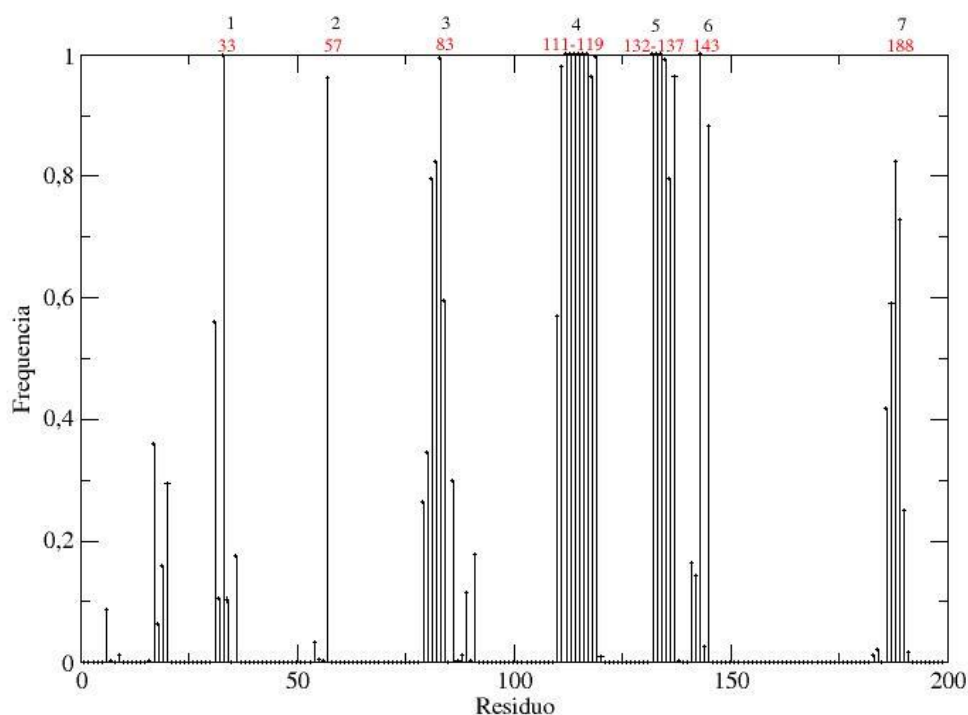
Esse resultado é de grande importância, especialmente em trabalhos de *docking molecular*, pois a utilização de uma única estrutura do alvo molecular pode não ser a melhor opção. Talvez haja a necessidade do uso de um conjunto de estruturas que melhor representem a dinâmica dessa proteína. Isso foi também indicado no trabalho de De Almeida et al., onde distinguiram quatro estados conformacionais diferentes para a NS3 do vírus da Dengue, através de dinâmica molecular (DE ALMEIDA et al., 2013). A Figura 69 mostra algumas dessas estruturas que apresentam essa mudança estrutural referente ao movimento das regiões amarela e laranja da Figura 66. Fica claro que o movimento observado na região amarela e laranja é acompanhado por outros movimentos da proteína. Possivelmente o substrato, por ter mais resíduos que o inibidor (caso DENV2f-INIB), provoque maiores perturbações no sistema, e apresente maiores dificuldades para alcançar sua posição ideal no sítio ativo da proteína.





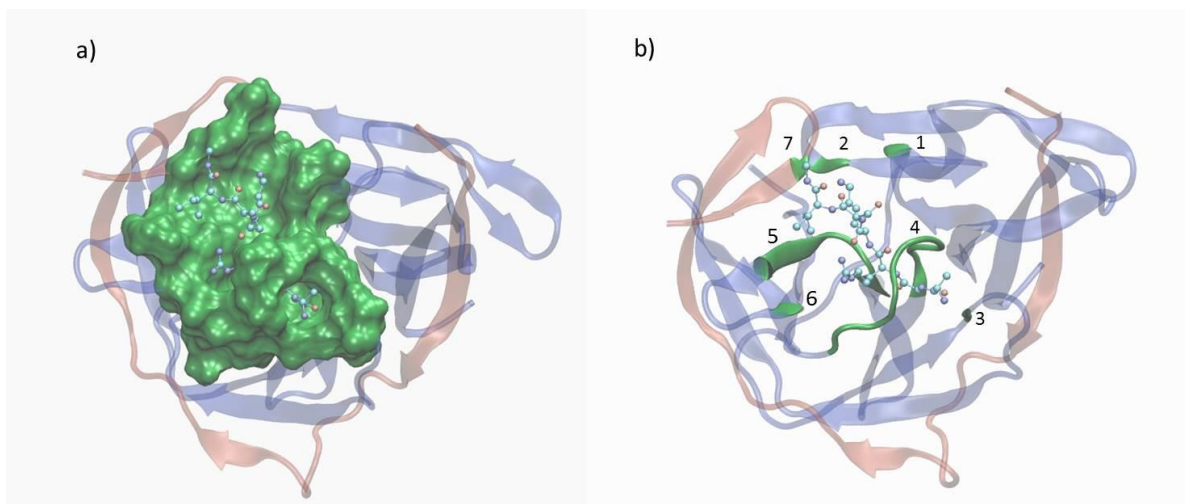
**Figura 69:** Ilustração de algumas estruturas da DENV2f-SUBS em que as regiões dos resíduos 112-116 e os 138-141 estão fazendo um movimento de se afastar e se aproximar uns dos outros, relativamente. Distância relativa entre as regiões amarela e laranja a) 45,3 Å; b) 35,5 Å; c) 44,5 Å e d) 34,1 Å. Em azul a NS3; vermelho a NS2B; verde o substrato e amarelo os resíduos 112-116 e, laranja os resíduos 138-141.

Com relação às possíveis interações que substrato possa fazer, observamos os resíduos da protease de serina que mais aparecem num raio de 5.0 Å em torno do substrato, contando os com maior frequência de aparecimento neste raio (Figura 70). Os resíduos do sítio ativo aparecem próximos ao substrato numa frequência maior que 90%, assim como os resíduos 83,111-119, 132-137, 143, o único resíduo da NS2B que aparece com mais de 70% é o resíduo 188, mostrando que há pouca interação entre o substrato e a NS2B, especialmente a porção grampo- $\beta$  terminal, que no caso anterior (DENV2f-INIB) foi responsável por várias interações com o inibidor. Experimentalmente se observa que em inibidores grandes não há interação com a NS2B.(NOBLE et al., 2012) Ou seja, esse resultado está de acordo com dados já observados para moléculas grandes (mais de seis resíduos).(NOBLE et al., 2012)



**Figura 70:** Resíduos que estão num raio de 5.0 Å dos resíduos do substrato. Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.

Na Figura 71 é possível verificar a região dos resíduos que estão próximos ao substrato-1 num raio de 5.0 Å. Uma das regiões (região 4 na Figura 71 (b)) é aquela vista na Figura 66 (cor azul), uma região bastante flexível. Uma análise das ligações de hidrogênio será capaz de tirar algumas dúvidas sobre o tipo de interação do substrato com a proteína.



**Figura 71:** Ilustração dos resíduos que estão num raio de 5.0 Å dos resíduos do substrato. a) os resíduos estão representados por uma superfície, e o substrato está sobre essa superfície. ; b) os resíduos estão representados em verde na imagem, e a numeração se refere à região que está representada no gráfico da **Figura 70**. Representado em azul a NS3, em vermelho a NS2B, em bolas o substrato, e em verde os resíduos que estão num raio de 5.0 Å do substrato. No substrato não há a representação dos hidrogênios para a melhor visualização.

A análise das ligações de hidrogênio entre substrato e a proteína (DENV2f) (Tabela 3) mostra que existem poucas interações de natureza forte entre essas moléculas, o que pode indicar que a permanência do substrato na região do sítio ativo seja devida a efeitos estéricos, interações de carga e interações de van der Waals (vdW), o que já foi observado experimentalmente para inibidores com mais de quatro resíduos. (NOBLE et al., 2012)

Apenas três dos sete resíduos do substrato interagem através de ligações de hidrogênio com a proteína, dois resíduos com carga positiva (ARG195 e LYS196) e um de carga neutra e apolar (ILE198) (Tabela 3).

Os resíduos da proteína que interagem por ligação de hidrogênio com o substrato foram os ASP111, PHE112, GLY115, GLY133 e TYR143 (Tabela 3), ou seja, interações observadas também no caso DENV2f-INIB (Tabela 2 Página 107), exceto para os resíduos do sítio ativo que não aparecem no caso com o substrato.

Os resíduos responsáveis pela cavidade oxianion também são observados interagindo com o substrato, ASP111, GLY133 e TYR143, o que é importante para a etapa da reação na estabilização de uma das etapas.

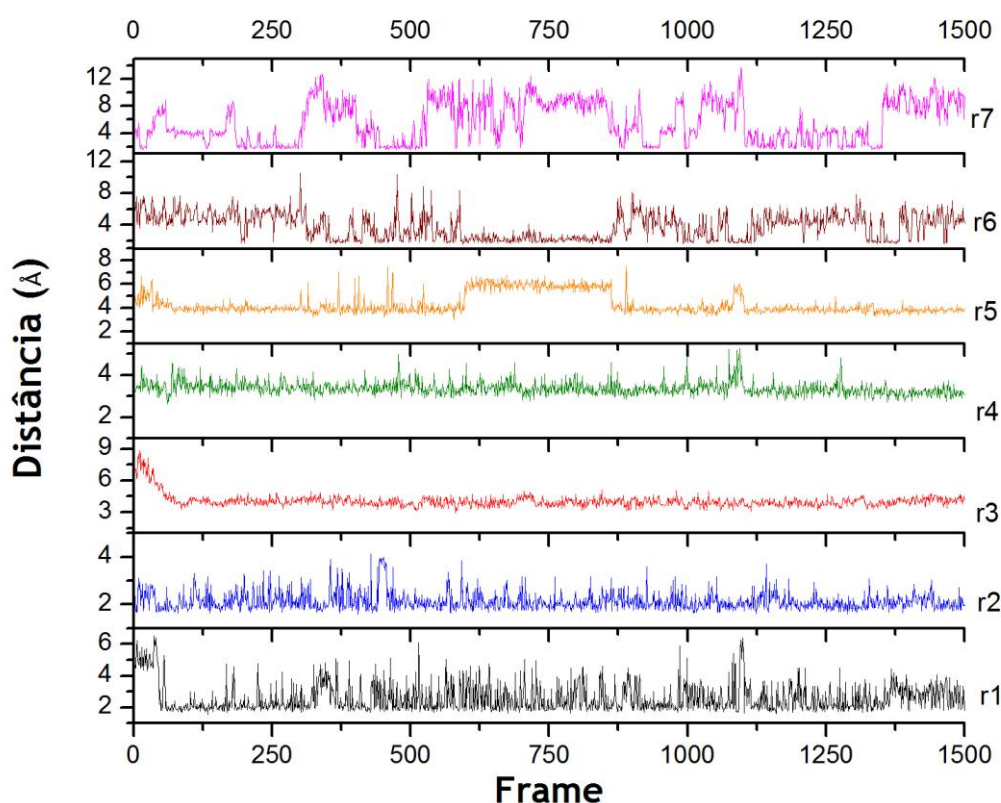
Uma maneira de melhor visualizar o comportamento dessas ligações de hidrogênio durante a simulação é acompanhando as distâncias dessas ligações durante a simulação de dinâmica molecular (Figura 72).



**Tabela 3:** Dados das ligações de hidrogênio entre a DENV2f e o substrato-1. Resíduos do substrato estão destacados em negrito.

| Aceitador     | Doador        | % Tempo | Média da distância (Å) |
|---------------|---------------|---------|------------------------|
| GLY133        | <b>LYS196</b> | 61      | 2,87                   |
| <b>ARG195</b> | GLY115        | 56      | 2,85                   |
| PHE112        | <b>ARG195</b> | 29      | 2,84                   |
| ASP111        | <b>ARG195</b> | 24      | 2,82                   |
| <b>ILE198</b> | TYR143        | 17      | 2,80                   |

Na Figura 72 é possível verificar que só r1 e r2 passam maior parte do tempo com a distância da ligação de hidrogênio menor que 2,5 Å. Essas duas ligações de hidrogênio são entre resíduos do sítio ativo, r1 (SER117-HIS33) e r2 (ASP57-HIS33).



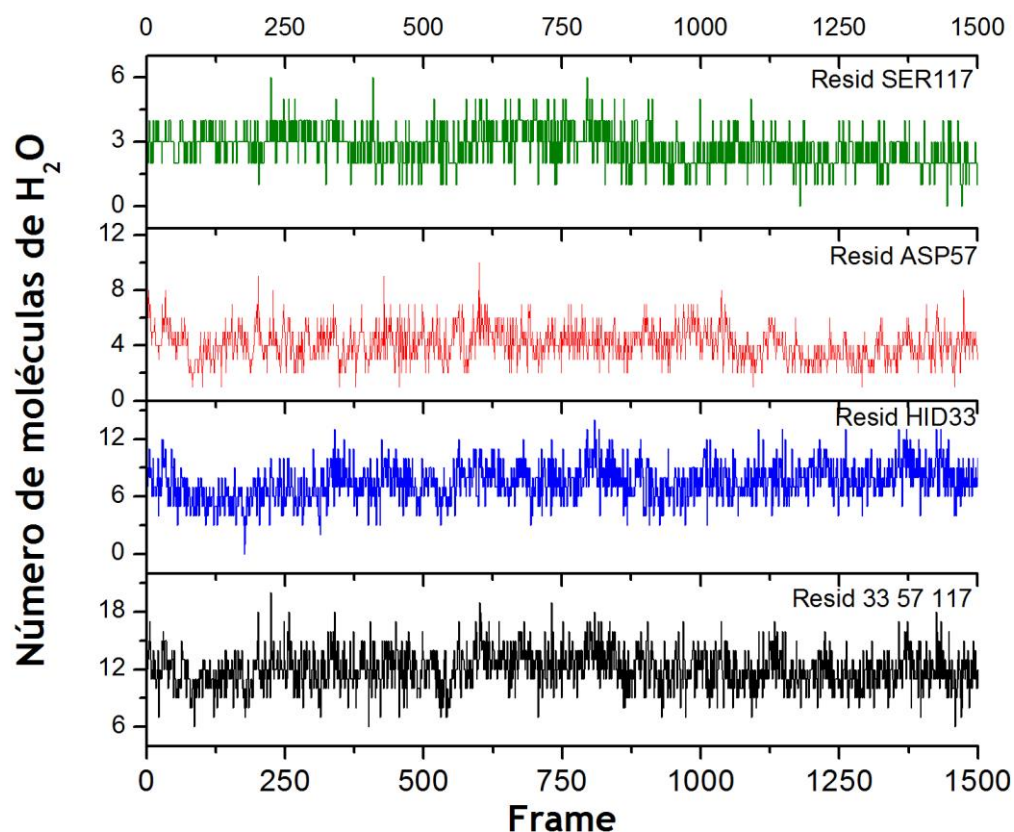
**Figura 72:** Distâncias de ligação de hidrogênio versus os frames da simulação. As distâncias r1 (SER117-HIS33); r2 (ASP57-HIS33); r3 (GLY133-LYS196); r4 (ARG195-GLY115); r5 (PHE112-ARG195); r6 (ASP11-ARG195); r7 (AILE198-TYR143).

Já as outras ligações de hidrogênio passam pouco tempo com distância menor que 2,5 Å, mostrando os casos r3, r4, r5, r6, e r7 com valores acima de 3,0 Å, ou seja, interações mais fracas entre esses resíduos. Isto evidencia que na presença desse substrato, os resíduos

do sítio ativo tenderam a interagir entre si ao invés de com o substrato, e mostra que as interações entre proteína e substrato não são fortes. Portanto, o que deve manter o substrato nessa região são efeitos estéricos, interações de carga e van der Waals (vdW) .

Os efeitos estéricos podem estar relacionados aos resíduos da Figura 66 (Página 118), em que os resíduos 112-116 ficam sobre parte do substrato, os resíduos 10-14 e resíduos 138-141 que fazem um movimento do tipo abertura e fechamento. Isso tudo pode estar relacionado ao tamanho da molécula do substrato (sete resíduos) que induz outro tipo de comportamento na proteína (DENV2f).

Nesse sistema também avaliamos o número médio de moléculas de água próximas ao sítio ativo (HIS33, ASP57 e SER117), pois a água pode participar de uma das fases da reação de quebra da ligação peptídica do substrato.(HEDSTROM, 2002b) Analisamos também o número de moléculas de água próximas a cada resíduo do sítio ativo separadamente, avaliando assim quais dos resíduos interagem mais com moléculas de água. Ao observar a Figura 73, pode-se ver que existe um grande número de moléculas de água próximas ao sítio ativo (média de 12 moléculas de H<sub>2</sub>O), sendo o resíduo HIS33 (média de 7 moléculas de H<sub>2</sub>O) o que apresenta maior quantidade quando comparada com os demais resíduos do sítio ativo, ASP57 e SER117, para os quais notamos uma média de 4 e 3 moléculas de água próximas, respectivamente. No caso do DENV2f-INIB, o resíduo SER117 aparece com mais duas moléculas de água próximas, ou seja, este resíduo tende a interagir mais com as moléculas de água do que com o substrato, além da interação com o resíduo HIS33 que aparece, nesse caso, mais forte. As moléculas de água próximas aos resíduos HIS33 e SER117 não interferiram na interação entre os mesmos (Figura 72).



**Figura 73:** Número de moléculas de H<sub>2</sub>O que estão a uma distância de 5.0 Å dos resíduos do sítio ativo. Em preto são as águas que estão próximas a tríade do sítio ativo, em azul ao resíduo HIS33, em vermelho ao resíduo ASP57 e verde ao resíduo SER117.

Os resultados desse sistema mostraram que o modelo DENV2f está de acordo com dados experimentais da proteína na presença de um inibidor com maior número de resíduos (mais que seis). (NOBLE et al., 2012)

Entretanto, como não existe uma estrutura cristalina resolvida dessa protease com seu substrato, portanto, era necessário um sistema mais comportado do ponto de vista estrutural, e com dados cristalográficos para o modelo DENV2f. Deste modo, necessitava-se de um estudo com um modelo de substrato com um número menor de resíduos, pois o DENV2f-INIB, onde o inibidor tem quatro resíduos poderia servir de modelo DENV2f com um novo modelo de substrato.

### 5.5 MODELO DENV2f COM SUBSTRATO 2 (DENV2f-SUBS2)

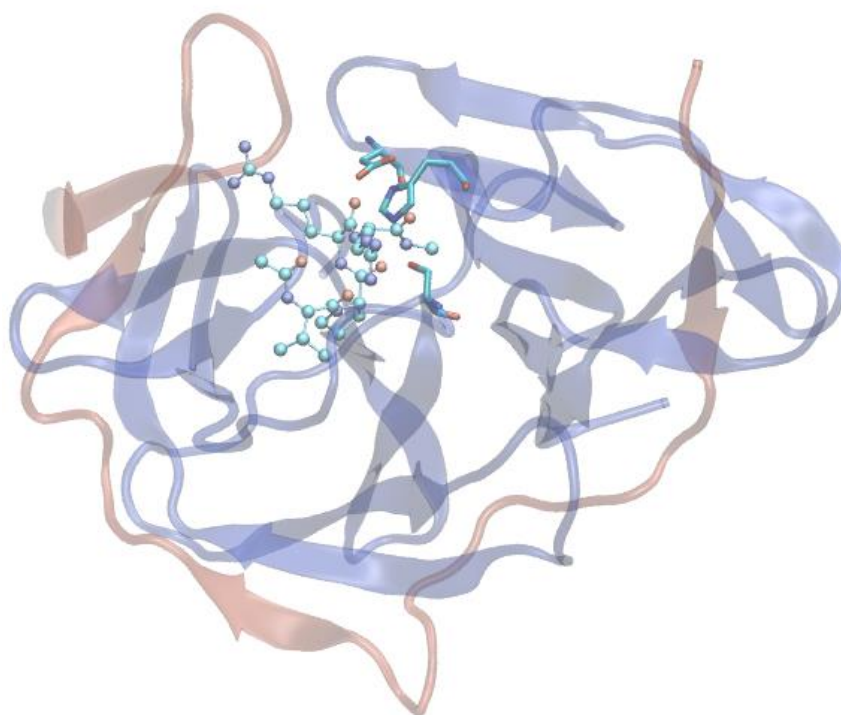
Observamos, na simulação anterior do modelo DENV2f com o substrato-1 (ALA-GLY-ARG-LYS-SER-ILE-NME), que era necessário um novo modelo de substrato para o estudo da protease DENV2f.

A sequência de aminoácidos do novo modelo de substrato foi obtida de um artigo experimental que testou várias sequências de aminoácidos de oito resíduos, (SHIRYAEV et al., 2007) mostrando que a sequência que apresentava maior eficiência catalítica para a protease de serina (NS2B-NS3) do vírus da Dengue era GKKRR↓PVR (a seta para baixo indica a região da quebra da ligação entre os resíduos). (SHIRYAEV et al., 2007) . Desta maneira, decidimos estudar a DENV2f com um novo modelo de substrato baseado nesses dados.

Então, escolhemos um segundo modelo para o substrato, aqui referenciado de substrato-2, com os quatro aminoácidos RRVP ( $\text{CH}_3\text{NH-R-R-P-V-COCH}_3$ ), pois evitaria assim problemas com estrutura desse substrato, onde  $\text{CH}_3\text{NH}$  e  $\text{COCH}_3$  são resíduos para completar a valência.

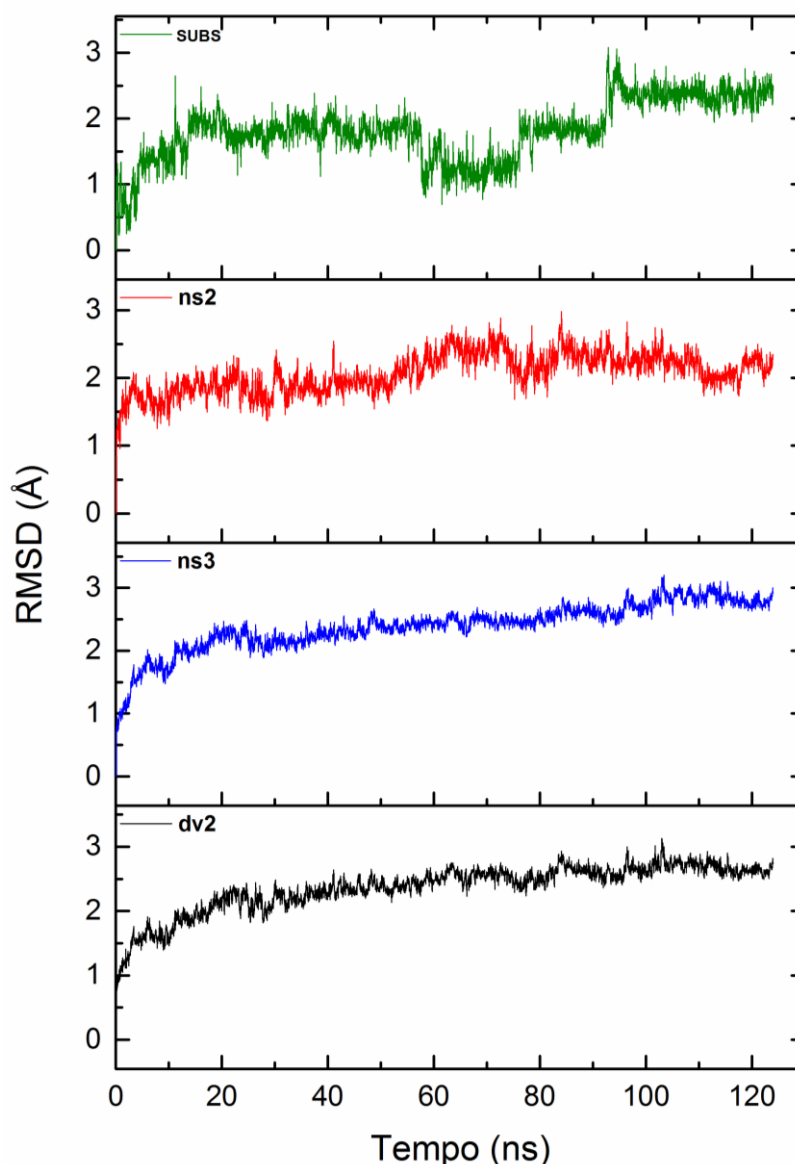
Alinhamos os resíduos desse novo substrato na estrutura de raio-X do inibidor que também contém quatro aminoácidos, mantendo a díade R-P próxima a região da serina do sítio ativo (Figura 74). Isso é fundamental para o estudo do mecanismo de reação que será mostrado na próxima seção (Página 133). Desta maneira, fizemos o estudo de dinâmica molecular desse novo sistema (DENV2f-SUBS2).

A Figura 75 mostra o RMSD obtido para a simulação de 125 ns de dinâmica molecular para a estrutura DENV2f-SUBS2. Nos primeiros 5 ns de simulação, onde o substrato ( $\text{CH}_3\text{NH-R-R-P-V-COCH}_3$ ) estava preso por uma restrição com constante de força de  $200 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ , o seu RMSD (gráfico verde na Figura 75), em relação a estrutura inicial, se mantém numa região com pouca oscilação. Depois de retirada a restrição, o substrato se estabiliza oscilando numa região de  $1,8 \text{ \AA}$ , entretanto, na região de 80-100 ns percebe-se uma maior oscilação, provavelmente devido ao molde usado para o substrato-2, em que os resíduos do inibidor são diferentes, e o alinhamento só se deu pelo *backbone* da estrutura, portanto, as cadeias laterais acabam se movendo mais para se adaptar as regiões da protease. No entanto, o comportamento da protease de serina (DENVf2 na Figura 75), assim como das regiões NS3 e NS2B é de um RMSD de aproximadamente  $2,4 \text{ \AA}$ , mas são estabilizados depois dos 40 ns de dinâmica molecular.



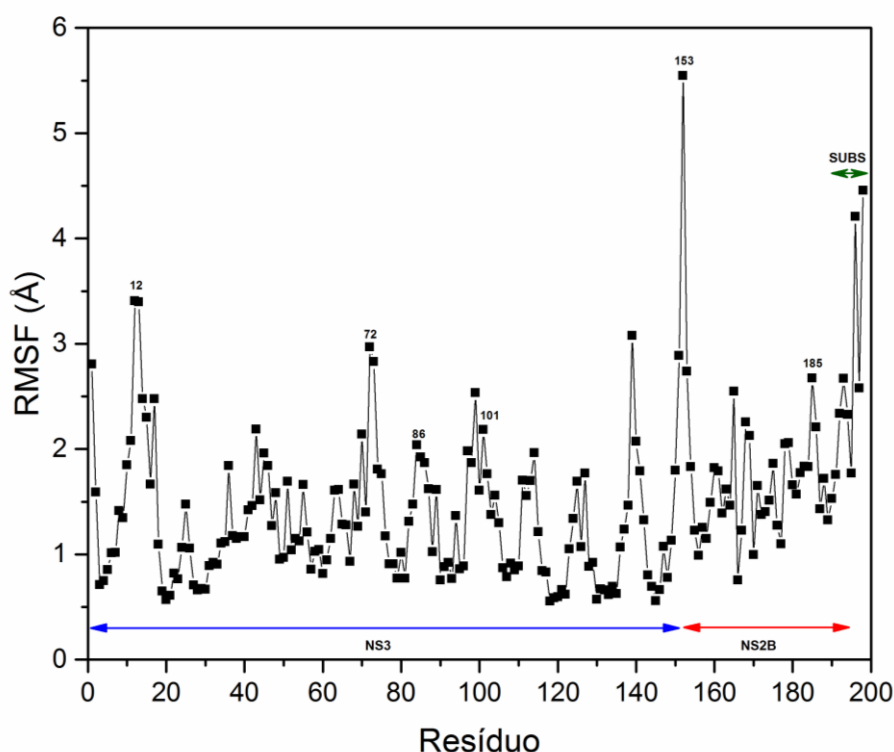
**Figura 74:** Uma conformação da DENV2f-SUBS2 retirada da simulação de dinâmica molecular numa *ensemble* NPT. As moléculas de água da caixa de simulação não estão representadas na figura para melhor visualização. Em Azul NS3, vermelho a NS2B, representado em canudo estão os resíduos do sítio ativo (HIS33, ASP57 e SER117), e a representação em bola são do inibidor sem os hidrogênios para a melhor visualização.

Observa-se que existe uma variação menor em torno da média do RMSD quando comparamos com o caso anterior (DENV2f-SUBS1), mostrando que um substrato menor interfere menos na estrutura da proteína.



**Figura 75:** RMSD para DENV2f (preto), assim como, o RMSD só para a NS3 resíduos 1 a 152 em azul, NS2B resíduos 153 a 192 em vermelho e para o substrato (2) em verde.

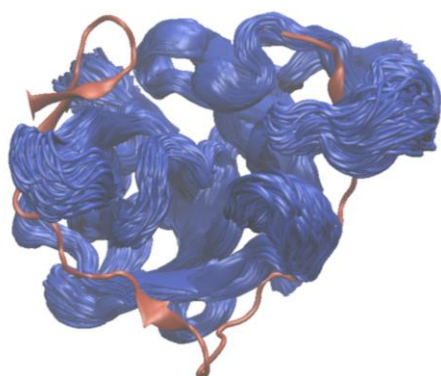
Analisando o gráfico de RMSF (Figura 76), observa-se que a mobilidade dos resíduos é menor do que no sistema DENV2f-SUBS1 (Figura 63, Página 115). Exceto o resíduo 153, pois é um resíduo terminal e tem grande mobilidade em todos os sistemas estudados até aqui. Ou seja, o número menor de resíduos no substrato causou uma menor perturbação no sistema, desse modo, fez com que o sistema se ajustasse mais rapidamente que no caso anterior (DENV2f-SUBS1). Ao compararmos com o sistema DENV2f-INIB (Figura 52, Página 104), percebe-se que o sistema DENV2f-SUBS2 (Figura 76) apresenta maior mobilidade dos seus resíduos, indicando que o substrato provocou uma perturbação maior no sistema.



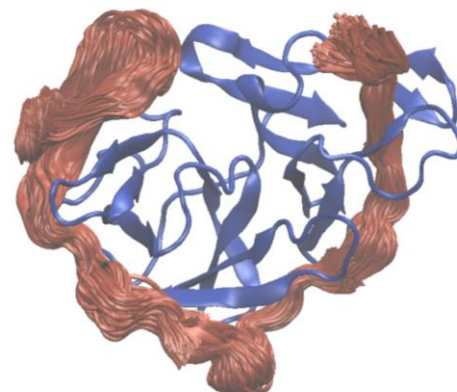
**Figura 76:** RMSF para os resíduos da DENV2f-SUBS, região delimitada com seta azul são os resíduos da NS3, a seta em vermelho são os resíduos da NS2B, já a seta em verde delimita os resíduos do substrato. A numeração é para os resíduos que apresentam maior mobilidade.

Na Figura 77 observa-se a superposição de 700 frames de simulação, permitindo analisar de maneira visual a mobilidade da cadeia principal de cada parte da proteína, NS2B e NS3. Quando o substrato-2 está posicionado na região do sítio ativo, a estrutura se mostra com pouca variação da posição dos átomos da cadeia principal para NS3 e NS2B. Não foi o que percebeu-se na secção anterior. Isso indica que esse substrato com quatro aminoácidos é um modelo melhor para o estudo do mecanismo de reação.

a)

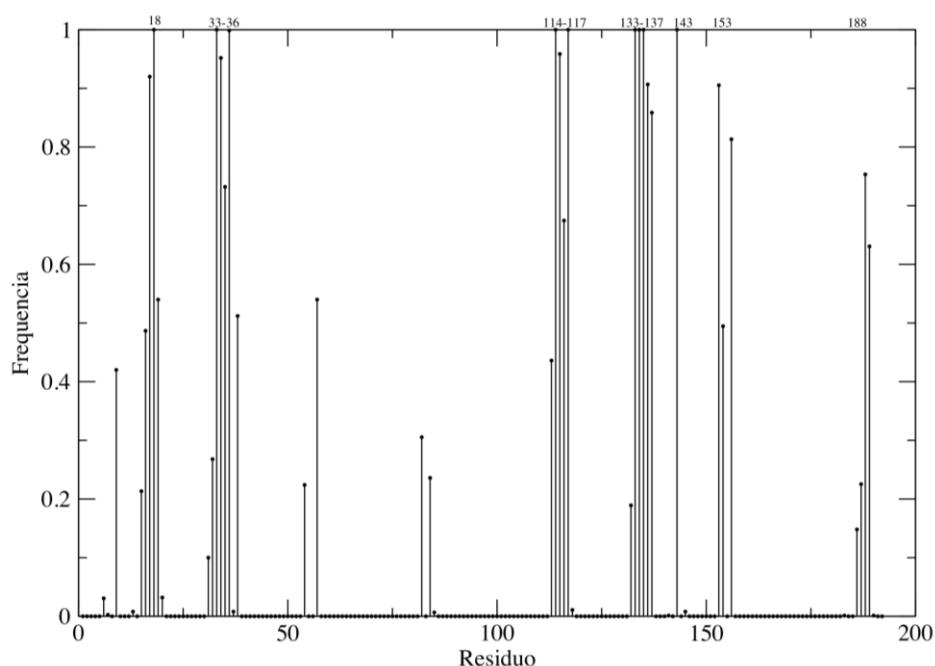


b)



**Figura 77:** Ilustração da superposição para as 750 conformações retiradas da simulação de dinâmica molecular NPT da DENV2f-SUBS2. a) superposição da NS3; b) superposição da NS2B. A região da NS3 (azul), em vermelho a NS2B.

A Figura 78 mostra um histograma dos resíduos da protease que estão a aproximadamente 5,0 Å do substrato. Dois dos resíduos da tríade catalítica (HIS33 e SER117) aparecem em todos os frames dentro dessa região, e o outro resíduo da tríade (ASP57) aparece em 60 % dos frames. Só um resíduo SER188, da região NS2B aparece em mais de 50 % dos frames. A maior interação do substrato é com resíduos da NS3. As interações por podem ser ligações de hidrogênio, eletrostáticas ou de van der Waals.



**Figura 78:** Gráfico dos resíduos que estão num raio de 5.0 Å dos resíduos do substrato (2). Os resíduos vizinhos estão no eixo da abscissa, enquanto a frequência da contagem dos frames que aparecem está no eixo da ordenada.

Observando a Tabela 4, identifica-se quais são os resíduos da protease que mais fazem ligações de hidrogênio com o substrato-2.

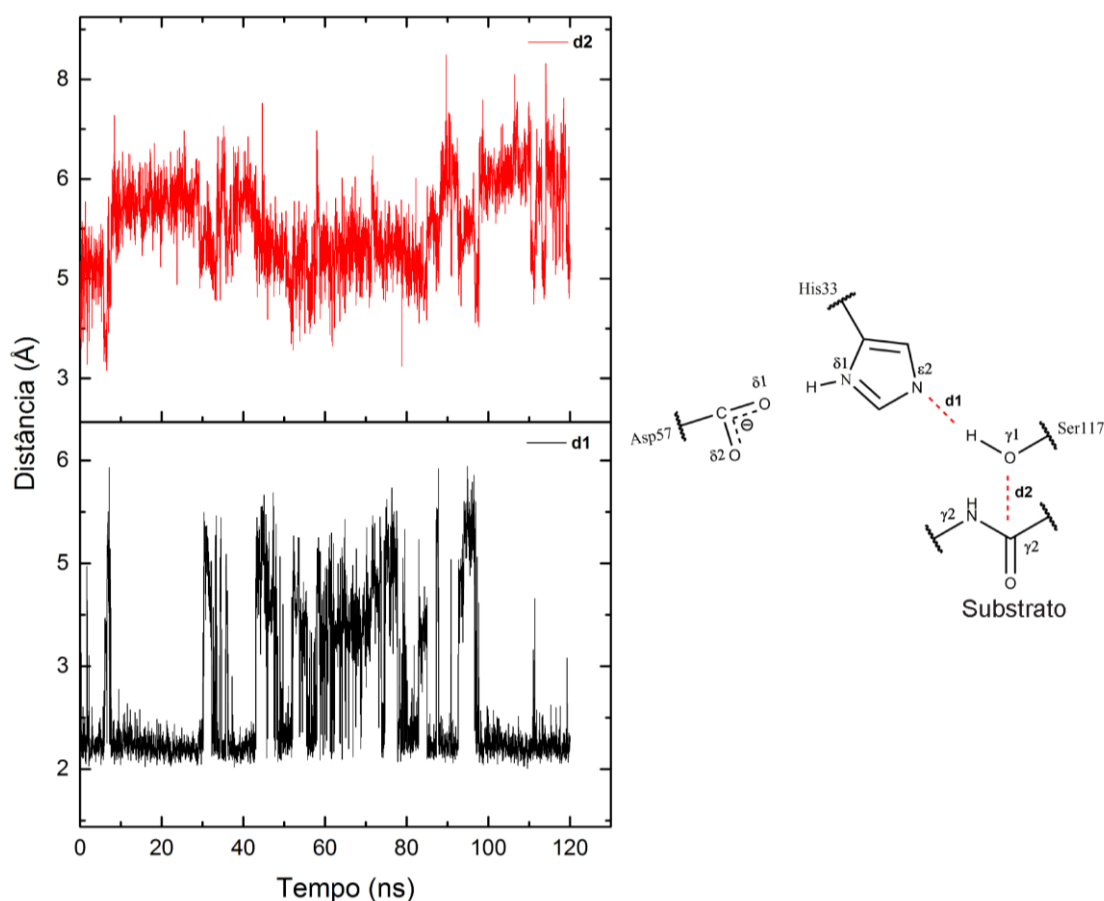
**Tabela 4:** Dados das ligações de hidrogênio entre a DENV2f e o substrato (2). Resíduos do substrato estão destacados em negrito.

| Aceitador     | Doador        | % Tempo | Média da distância (Å) |
|---------------|---------------|---------|------------------------|
| GLU153        | <b>ARG197</b> | 29      | 2,81                   |
| VAL34         | <b>ARG197</b> | 28      | 2,82                   |
| PRO114        | <b>ARG196</b> | 11      | 2,88                   |
| <b>ARG197</b> | TYR143        | 10      | 2,84                   |



Os resíduos da proteína que mais interagem por ligações de hidrogênio com substrato-2 foram GLU153, VAL34, PRO114 e TYR143. Esses resíduos diferem do caso com o substrato-1 (Tabela 3, Página 124), exceto pelo resíduo TYR143. Sabemos que o resíduo TYR143 pode ser importante para estudo do mecanismo, pois o mesmo é forma uma cavidade oxianion que estabiliza um dos intermediários da reação.

Para podermos começar o estudo da reação foi necessário avaliar duas distâncias específicas. Uma é definida entre o oxigênio da serina (SER117) e o carbono carboxílico de um dos resíduos do substrato, o que corresponde ao ataque nucleofílico da serina. A outra distância é definida entre o hidrogênio da serina e o nitrogênio épsilon da histidina, ambos resíduos do sítio ativo.



**Figura 79:** Gráfico de distâncias e ilustração das mesmas na imagem ao lado. A distância d1 é definida entre o hidrogênio  $H_{\gamma1}$  da serina e o nitrogênio  $N_{\epsilon2}$  da histidina, e d2 é a distância entre o oxigênio  $O_{\gamma1}$  da serina e o carbono carboxílico  $C_{\gamma2}$  do substrato-2.

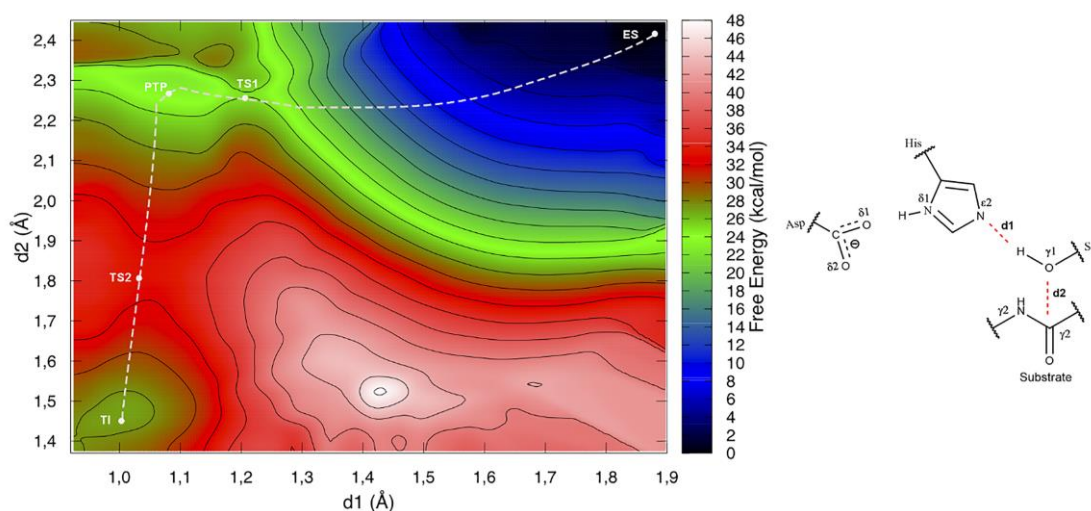
A média dessas duas distâncias são  $d1 = 2,75 (\pm 1,18) \text{ \AA}$  e  $d2 = 5,65 (\pm 0,72) \text{ \AA}$  (Figura 79). Como veremos na seção seguinte, essas distâncias d1 e d2 são de fundamental importância para o estudo do mecanismo de reação catalisado pela protease de serina.

Observa-se que o substrato em alguns frames fica aproximadamente a 3.0 Å da serina (SER117) do sítio ativo (d2). Assim como o hidrogênio da serina está sempre muito próximo ao nitrogênio épsilon da histidina (d1). Para a próxima etapa do trabalho, foi necessário escolhermos uma estrutura boa para começar o estudo da reação de quebra da ligação peptídica entre resíduos do substrato. Escolhemos a estrutura que apresentou menores valores d1(3,0 Å) e d2 (1,9 Å), simultaneamente.

## 5.6 MECANISMO DE REAÇÃO

### 5.6.1 Etapa de Acliação

O perfil de energia livre (Potencial de Força Média, PMF) para o processo de formação do intermediário tetraédrico, calculado usando *Umbrella Sampling* com duas dimensões, está mostrado na Figura 80. A linha tracejada indica o caminho de menor energia para a formação do intermediário tetraédrico. Os valores de energia livre nos pontos críticos referentes a esse caminho estão na Figura 81.



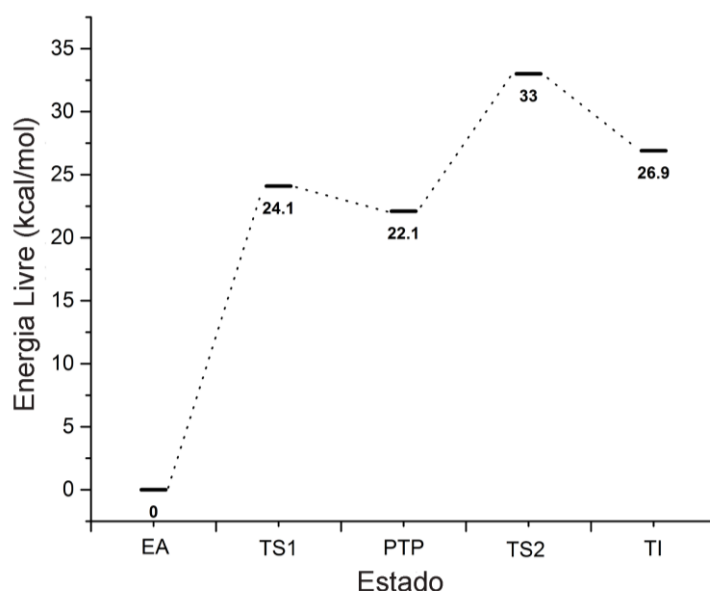
**Figura**

**80:** Perfil de energia livre bidimensional (ou PMF) QM(PM3PDDG)/MM para a formação do intermediário tetraédrico (TI) pela reação da NS2B-NS3 protease com o substrato (VPRR). O zero da energia corresponde à formação do complexo enzima-substrato (ES). A linha tracejada representa o caminho de menor energia para a formação do intermediário tetraédrico. Ilustração das mesmas na imagem ao lado. A distância d1 é definida entre o hidrogênio H<sub>γ1</sub> da serina e o nitrogênio N<sub>ε2</sub> da histidina, e d2 é a distância entre o oxigênio O<sub>γ1</sub> da serina e o carbono carboxílico C<sub>γ2</sub> do substrato-2.

Essa primeira etapa da reação envolve dois eventos: a transferência do próton da SER117 para HIS33, representada pela coordenada d1, e o ataque nucleofílico do oxigênio da SER117 ao carbono carboxílico do substrato. A literatura mostra duas possibilidades: Warshel e colaboradores sugerem uma reação em etapas com a transferência de próton da serina para histidina, anterior ao ataque nucleofílico.(BENTZIEN et al., 1998; WARSHEL, 1991) Outros

apoiam que a transferência do próton entre SER e HIS ocorre de maneira concertada com o ataque nucleofílico da serina ao substrato.(ISHIDA; KATO, 2003b)

No caso específico da quebra da ligação peptídica pela NS2B-NS3 do vírus da Dengue, nossos resultados mostram que o caminho de menor energia passa por um mínimo relacionado à esta transferência do próton entre a serina e a histidina. A barreira para a formação desse estado intermediário (PTP, produto da transferência de próton), é de 24,1 kcal/mol. Entretanto, esse estado fica em um poço energético muito raso (Figura 81), indicativo de tempo de vida curto para a espécie, o que pode explicar a controvérsia sobre o mecanismo. Uma vez que o próton é transferido para a histidina, a reação segue para a formação do intermediário tetraédrico com uma barreira de 10,9 kcal/mol.



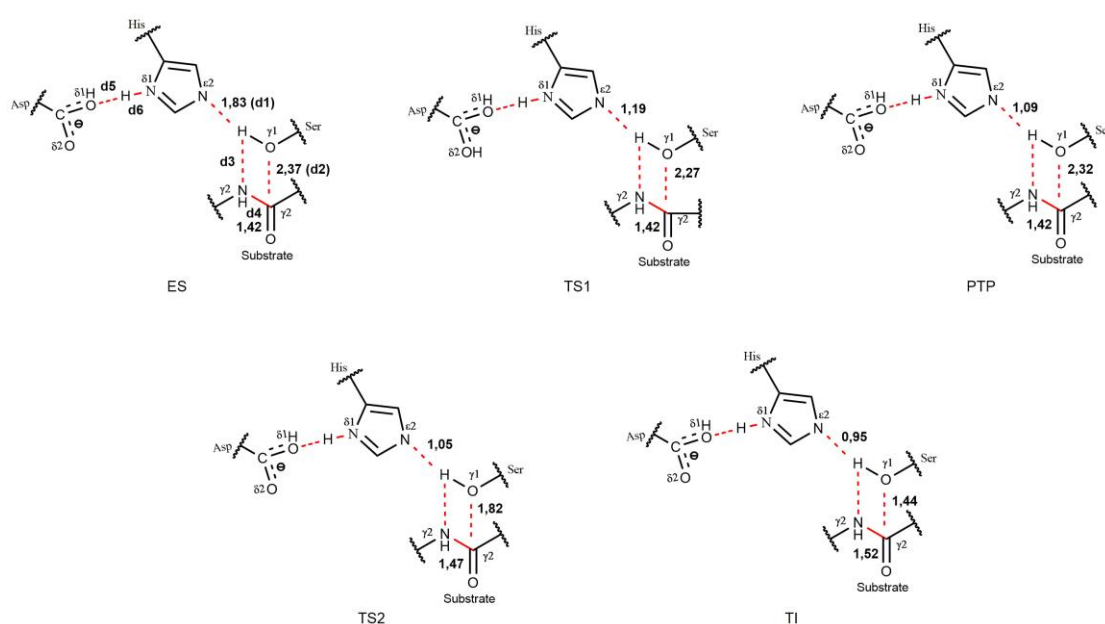
**Figura 81:** Ilustração da diferença de energia livre dos pontos estacionários em relação ao complexo enzima-substrato (ES) que representa o caminho de menor energia observado na **Figura 80**. Deste modo, apresentando o patamar de energia para cada ponto estacionário e, uma melhor visualização das barreiras de energia entre os estados.

O estado de transição para a transferência do próton ( $H_{\gamma 1}$ ) da serina para o nitrogênio ( $N_{\epsilon 2}$ ) da histidina ocorre em torno de  $d1(H_{\gamma 1} - N_{\epsilon 2}) = 1.19 \text{ \AA}$  e  $d2(O_{\gamma 1} - C_{\gamma 2}) = 2.27 \text{ \AA}$  (Tabela 5). Isso mostra apenas uma pequena aproximação entre  $O_{\gamma 1}$  da serina ao  $C_{\gamma 2}$  do substrato com relação a estrutura inicial. Na formação do intermediário (PTP),  $d1$  é encurtada enquanto  $d2$  permanece a mesma ( $d1 = 1.09 \text{ \AA}$ ,  $d2 = 2.27 \text{ \AA}$ ) (Tabela 5), indicando a transferência do próton antecedendo o ataque nucleofílico. Esse resultado é semelhante ao obtido por Kato e Ishida.(ISHIDA; KATO, 2003b) Estudando o mecanismo de reação da serina protease tripsina, eles observam também que, após uma pequena aproximação inicial do substrato, o ataque nucleofílico ocorre apenas após a transferência do próton.

**Tabela 5:** Valores médios e desvios padrão das distâncias em cada um dos pontos estacionários, na qual estão identificadas na **Figura 80**. Médias obtidas sobre 50 ps de simulação em cada um dos pontos estacionários. Valores em Å.

| Estado | Distância       |                 |             |             |             |             |
|--------|-----------------|-----------------|-------------|-------------|-------------|-------------|
|        | d1 <sup>a</sup> | d2 <sup>a</sup> | d3          | d4          | d5          | d6          |
| ES     | 1,83 ± 0,12     | 2,37 ± 0,01     | 3,88 ± 0,68 | 1,42 ± 0,03 | 0,99 ± 0,03 | 3,17 ± 1,28 |
| TS1    | 1,19 ± 0,02     | 2,27 ± 0,01     | 3,14 ± 0,18 | 1,42 ± 0,03 | 1,01 ± 0,03 | 1,66 ± 0,07 |
| PTP    | 1,09 ± 0,02     | 2,32 ± 0,01     | 3,26 ± 0,19 | 1,42 ± 0,03 | 1,01 ± 0,03 | 1,67 ± 0,07 |
| TS2    | 1,05 ± 0,02     | 1,82 ± 0,01     | 2,89 ± 0,19 | 1,47 ± 0,03 | 1,01 ± 0,03 | 1,69 ± 0,14 |
| TI     | 0,95 ± 0,02     | 1,44 ± 0,01     | 3,55 ± 0,37 | 1,52 ± 0,04 | 0,98 ± 0,04 | 3,29 ± 1,31 |

<sup>a</sup> Coordenadas d1 e d2 estão restritas durante esta etapa.



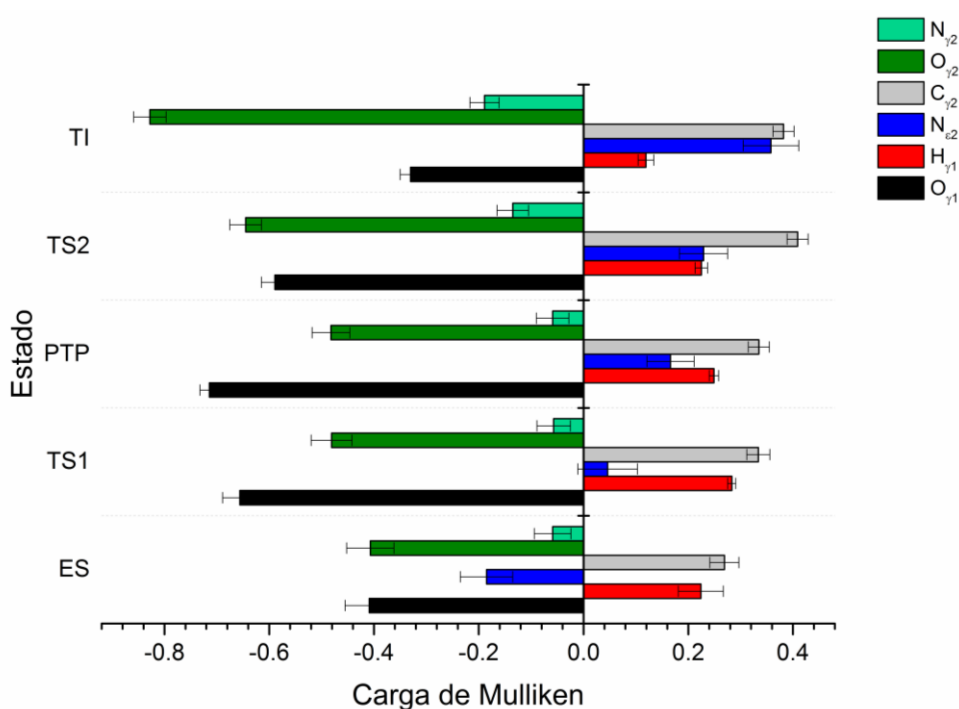
**Figura 82:** Esquema para representar algumas das distâncias médias que estão na **Tabela 5**. Os valores das distâncias estão em Å.

#### 5.6.1.1 O papel da ASP57 no mecanismo.

Alguns trabalhos sugerem que o aumento da nucleofilicidade da serina (SER117) seja devido não a uma, mas a duas transferências de prótons concertadas: da SER117- $O_{\gamma 1} \rightarrow HIS33-N_{\epsilon 2}$  e  $HIS33-N_{\delta 1} \rightarrow ASP57-O_{\delta 1}$ . (HUNKAPILLER et al., 1973; SCHOLTEN; HOGG; RAUSHEL, 1988; WARSHEL; RUSSELL, 1986) No nosso caso isso também é observado. Na Tabela 5 observa-se que, mesmo antes de ocorrer a transferência de próton  $SER117-O_{\gamma 1} \rightarrow HIS33-N_{\epsilon 2}$ , o hidrogênio delta ( $\delta$ ) da HIS33 ( $H_{\delta 1}$ ) já foi transferido para o ASP57 ( $O_{\delta 1}$ ). Essa transferência aumenta a nucleofilicidade da posição épsilon ( $\epsilon$ ) da HIS33, permitindo a transferência  $SER117-O_{\gamma 1} \rightarrow HIS33-N_{\epsilon 2}$ , ativando a SER117 para o ataque nucleofílico ao substrato.

### 5.6.1.2 Cargas atômicas

A mudança no caráter básico no nitrogênio HIS33- $N_{\epsilon 2}$  pode ser acompanhada pela variação nas cargas de Mulliken. A Figura 83 mostra as cargas nos pontos estacionários da reação para os átomos envolvidos. É possível notar a mudança de caráter básico do átomo  $N_{\epsilon 2}$  para uma natureza ácida após a transferência do próton  $SER117-O_{\gamma 1} \rightarrow HIS33-N_{\epsilon 2}$ . Essa mudança de característica é fundamental para etapa de quebra da ligação peptídica, pois esse hidrogênio ( $H_{\gamma 1}$ ) será transferido para o nitrogênio do substrato ( $N_{\gamma 2}$ ), levando à quebra da ligação peptídica.



**Figura 83:** Cargas de Mulliken para alguns dos átomos da região de reação no processo de formação do intermediário tetraédrico. As barras coloridas são os valores médios para cada um dos pontos estacionários. O desvio padrão é mostrado pela barra em preto em cada uma das cargas. Isso corresponde a 50 ps em cada um dos pontos estacionários.

A carga negativa sobre o  $SER117-O_{\gamma 1}$  aumenta após a transferência do próton  $SER117-O_{\gamma 1} \rightarrow HIS33-N_{\epsilon 2}$  formando o intermediário PTP, momento em que apresenta seu valor mais baixo de -0.714. Esse aumento de carga é importante para a ativação da  $SER117$ , que ataca o carbono  $C_{\gamma 2}$  do substrato para formar o intermediário tetraédrico. Após esse ataque, a carga negativa do  $O_{\gamma 1}$  volta a cair com a redistribuição de carga para o  $O_{\gamma 2}$  e  $N_{\gamma 2}$ , já preparando esse último para a próxima etapa.

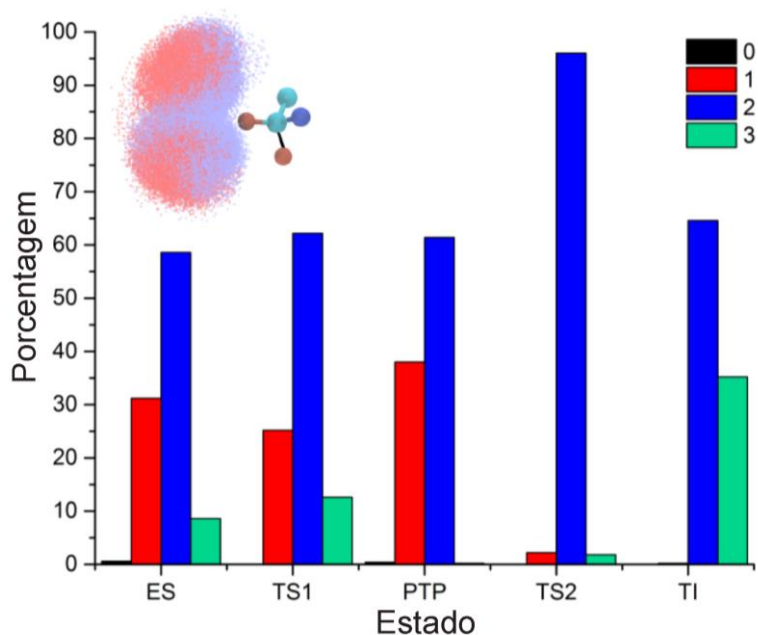
No processo, o oxigênio  $O_{\gamma 2}$  (substrato) tem sua carga cada vez mais negativa durante o processo de ataque nucleofílico, chegando a um valor de -0.828, quando o intermediário

tetraédrico (TI) é formado. Essa carga negativa é estabilizada por moléculas de água próximas (a pelo menos 3.0 Å) do átomo  $O_{\gamma 2}$  do substrato (Figura 83).

As cargas formais de Mulliken dos nitrogênios  $N_{\epsilon 2}$  e  $N_{\gamma 2}$  (Figura 83) mostram que, enquanto o nitrogênio HIS33- $N_{\epsilon 2}$  aumenta sua natureza ácida, o nitrogênio  $N_{\gamma 2}$  (substrato) fica cada vez mais básico. Isto é crucial para a etapa seguinte de clivagem da ligação peptídica, pois a transferência do  $H_{\gamma 1}$  entre esses nitrogênios ( $HIS33-N_{\epsilon 2} \rightarrow Subs-N_{\gamma 2}$ ) induz a quebra da ligação peptídica ( $d4, C_{\gamma 2} - N_{\gamma 2}$ ).

#### 5.6.1.3 A cavidade oxiânion (oxianion hole) na NS2B-NS3 é formada por moléculas de água.

Na Figura 84 é possível notar que, nos pontos estacionários, há duas moléculas de água dentro de um raio de até 3.0 Å de  $O_{\gamma 2}$  em 60% dos frames, ajudando a estabilizar a carga negativa. Durante o ataque nucleofílico (TS2) essa porcentagem chega a 95% dos frames. Quando o intermediário tetraédrico (TI) se forma completamente, essa quantia retorna a 65%, enquanto a porcentagem de frames onde três moléculas de água se aproximam de  $O_{\gamma 2}$  aumenta, chegando a 35% dos frames. Como nesse sistema não há nenhum resíduo da proteína próximo o suficiente (3.5 Å) do oxigênio  $O_{\gamma 2}$ , a cavidade oxiânion (*oxianion hole*) é formada apenas por moléculas de água, podendo ser duas ou três moléculas de água estabilizando a carga negativa sobre  $O_{\gamma 2}$ . Daggett e colaboradores, por cálculos teóricos usando os métodos semi-empíricos AM1 e PM3, para modelos do sítio ativo da tripsina reagindo com compostos amida e ester na formação de complexos tetraédricos, mostraram a importância da cavidade oxiânion na estabilização do intermediário tetraédrico em até 14 kcal/mol, utilizando duas moléculas de água na região. (SCHROEDER; DAGGETT; KOLLMAN, 1991)



**Figura 84:** Número de moléculas de água em percentagem dos frames de cada ponto estacionário. Cada uma das cores representa um número específico de moléculas de água. Esse dado foi obtido em 50 ps de cada um dos pontos estacionários.

### 5.6.2 Etapa de Quebra da Ligação Peptídica

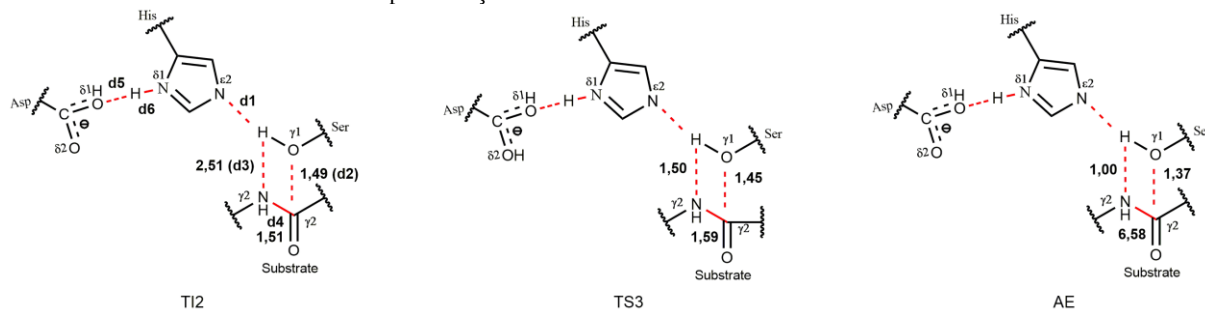
Para esta etapa, partimos de uma das estruturas da simulação do ponto estacionário referente ao estado TI. Na simulação da etapa anterior, as coordenadas d1 e d2 estavam restritas. Dentre as estruturas geradas, escolhemos a que apresentava o menor valor para a coordenada que representa a transferência do próton ( $H_{\gamma 1}$ ) da His51 para o nitrogênio ( $N_{\gamma 2}$ ) do grupo de saída do substrato ( $H_{\gamma 1}-N_{\gamma 2}$ ,  $d3=2.51$  Å). Essa estrutura foi submetida à uma minimização de energia na qual apenas d3 foi restrito ao seu valor inicial, e todas as demais distâncias foram deixadas livres. À estrutura final obtida nos referimos por TI2 (intermediário tetraédrico 2).

A segunda etapa da reação foi simulada controlando-se apenas a transferência de próton (d3), o que levou automaticamente à clivagem da ligação peptídica (d4). A estrutura escolhida como ponto de partida foi submetida à minimização de energia com restrição apenas na coordenada ( $d3=2.51$  Å). Inicialmente, utilizamos 5 janelas para reduzir d3 até a distância de 2.05 Å. A partir desse ponto utilizamos 11 janelas para reduzir d3 até o valor final de 0.95 Å.

**Tabela 6:** Valores médios e desvios padrão das distâncias em cada um dos pontos estacionários, na qual estão identificadas na. Médias obtidas sobre 30 ps de simulação em cada um dos pontos estacionários. Valores em Å.

| Estado     | Distância   |             |                 |             |             |             |
|------------|-------------|-------------|-----------------|-------------|-------------|-------------|
|            | d1          | d2          | d3 <sup>a</sup> | d4          | d5          | d6          |
| <b>TI2</b> | 1,03 ± 0,03 | 1,49 ± 0,04 | 2,51 ± 0,03     | 1,51 ± 0,04 | 1,00 ± 0,04 | 1,72 ± 0,19 |
| <b>TS3</b> | 1,20 ± 0,09 | 1,45 ± 0,04 | 1,50 ± 0,04     | 1,59 ± 0,05 | 1,61 ± 0,06 | 1,06 ± 0,04 |
| <b>AE</b>  | 3,58 ± 0,48 | 1,37 ± 0,02 | 1,00 ± 0,02     | 6,58 ± 0,67 | 1,66 ± 0,06 | 1,04 ± 0,03 |

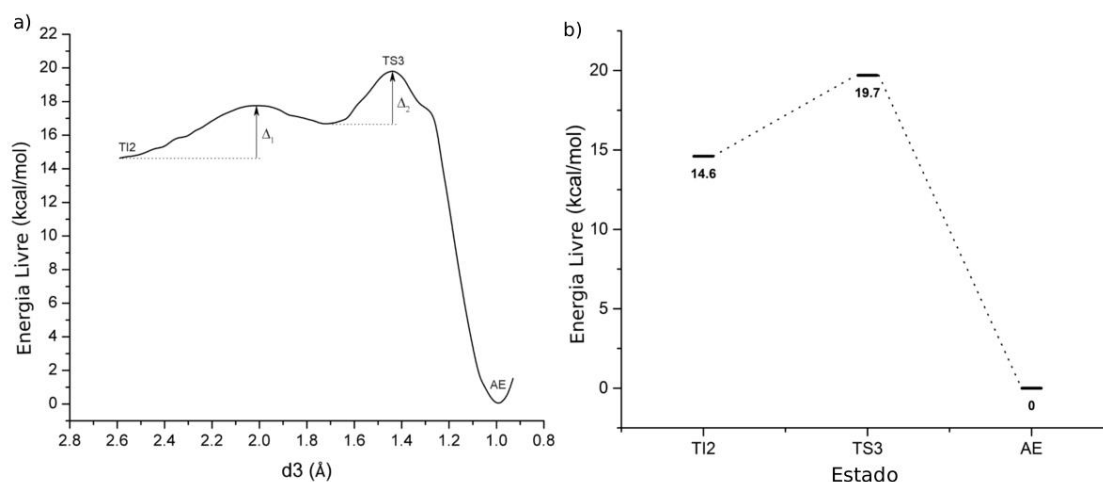
<sup>a</sup> Coordenada restrita durante essa etapa da reação.



**Figura 85:** Esquema para representar algumas das distâncias médias que estão na **Tabela 6**. Os valores das distâncias estão em Å.

A Figura 86 mostra a energia livre para a reação da transferência do próton  $H_{\gamma1}$  para o nitrogênio  $N_{\gamma2}$ . Conforme o próton  $H_{\gamma1}$  é transferido para o nitrogênio  $N_{\gamma2}$  do substrato (d3 diminui), ocorre a quebra da ligação peptídica (d4 aumenta). Há uma pequena barreira de  $\Delta_1 = 3,1$  kcal/mol, seguido de um pequeno mínimo (0,9 kcal/mol mol) e uma nova barreira de  $\Delta_2 = 2,9$  kcal/mol. A primeira barreira está relacionada com um evento de rearranjo que acontece longe da coordenada de reação d3, onde o ASP75 momentaneamente se afasta da HIS33, e depois aproxima-se novamente da HIS33. Como o vale entre os picos é raso, e se refere a um evento indireto, consideramos como barreira de energia para esse processo a diferença de energia entre os estados TI2 e TS3, que é 5,1 kcal/mol. Este valor está abaixo das barreiras de energia para a formação de TI, de tal forma que a formação de TI deve ser o passo determinante da velocidade para a reação. Com a quebra da ligação peptídica, a ligação entre a SER117 e o substrato se fortalece. No estado acil-enzima (AE), a ligação acil-enzima apresenta seu menor valor  $d2 = 1,37$  Å (Tabela 6). Ao mesmo tempo, o próton  $H_{\delta1}$  é devolvido para a HIS33, o que pode ser notado pelo aumento de d5 e diminuição de d6.

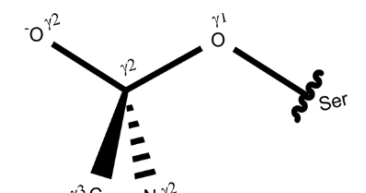




**Figura 86:** a) Perfil de energia livre para a reação de quebra da ligação peptídica. b) Esquema da diferença de energia livre dos pontos estacionários em relação ao produto AE. Ambos correspondendo a coordenada de reação unidimensional para a transferência do hidrogênio ( $H_{\gamma 1}$ ) da histidina ( $N_{\epsilon 2}$ ) para o nitrogênio ( $N_{\gamma 2}$ ) do grupo de saída do substrato (ARG).

Ao observar o comportamento dos ângulos (Tabela 7), nota-se que esses sofrem maior mudança quando a ligação peptídica é quebrada e o intermediário tetraédrico é desfeito. O carbono  $C_{\gamma 2}$ , que apresentava uma hibridização  $sp^3$ , passa para uma hibridização  $sp^2$  na formação dos produtos (AE). Os ângulos que sofrem maior mudança são os que incluem o átomo de nitrogênio ( $N_{\gamma 2}$ ), pertencente à arginina do grupo de saída.

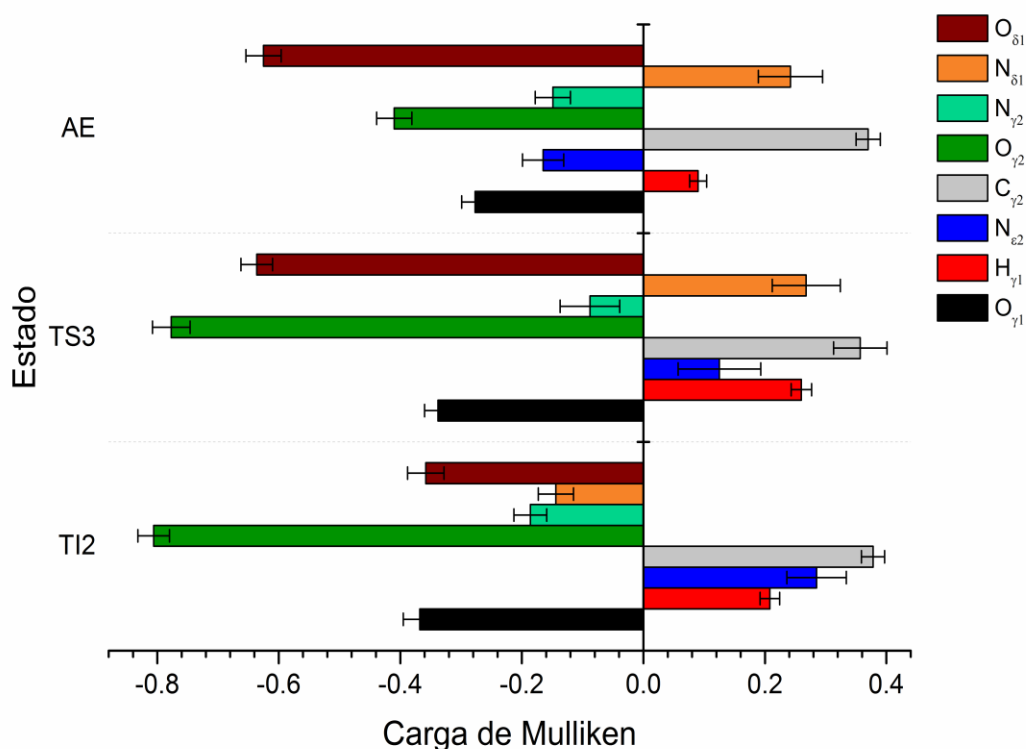
**Tabela 7:** Valores médios e desvios padrão dos ângulos (seis ângulos que definem uma estrutura tetraédrica) em cada um dos pontos estacionários, na qual estão identificadas na **Figura 86**. Médias obtidas sobre 30 ps de simulação em cada um dos pontos estacionários. O carbono alpha  $C_{\gamma 3}$  pertence ao resíduo da prolina.

| Ângulo<br>(°) | Estado        |               |               |  |
|---------------|---------------|---------------|---------------|---------------------------------------------------------------------------------------|
|               | TI2           | TS3           | AE            |                                                                                       |
| a1            | 108,29 (3,12) | 109,76 (3,66) | 119,35 (2,99) | a1: $O_{\gamma 1} - C_{\gamma 2} - O_{\gamma 2}$                                      |
| a2            | 108,29 (3,37) | 107,74 (3,60) | 114,09 (3,39) | a2: $O_{\gamma 1} - C_{\gamma 2} - C_{\gamma 3}$                                      |
| a3            | 94,88 (3,35)  | 96,37 (4,04)  | 98,19 (17,61) | a3: $O_{\gamma 1} - C_{\gamma 2} - N_{\gamma 2}$                                      |
| a4            | 113,13 (2,78) | 112,64 (3,20) | 83,81 (8,14)  | a4: $C_{\gamma 3} - C_{\gamma 2} - N_{\gamma 2}$                                      |
| a5            | 114,68 (3,09) | 111,52 (3,24) | 90,49 (17,74) | a5: $N_{\gamma 2} - C_{\gamma 2} - O_{\gamma 2}$                                      |
| a6            | 115,11 (3,39) | 116,55 (3,30) | 126,22 (3,59) | a6: $C_{\gamma 3} - C_{\gamma 2} - O_{\gamma 2}$                                      |

### 5.6.2.1 Cargas atômicas

A mudança da carga atômica (Figura 87) em alguns átomos importantes para a reação de quebra da ligação peptídica, evidencia a importância do ASP57 na estabilização do sistema reacional. A carga negativa do oxigênio ASP57- $O_{\delta 1}$  aumenta quando a estrutura tetraédrica é quebrada. Nesse momento,  $d5 (H_{\delta 1} - O_{\delta 1}) = 1.66 \text{ \AA}$  e  $d6 (H_{\delta 1} - N_{\delta 1}) = 1.04 \text{ \AA}$ , indicando que o próton  $H_{\delta 1}$  volta ao HIS33- $N_{\delta 1}$ , regenerando a histidina. Com a clivagem da ligação peptídica,

o oxigênio Subs- $O_{\gamma 2}$  volta a ter uma carga aproximadamente -0.4 (Figura 87), igual ao complexo ES do começo do processo (Figura 83).



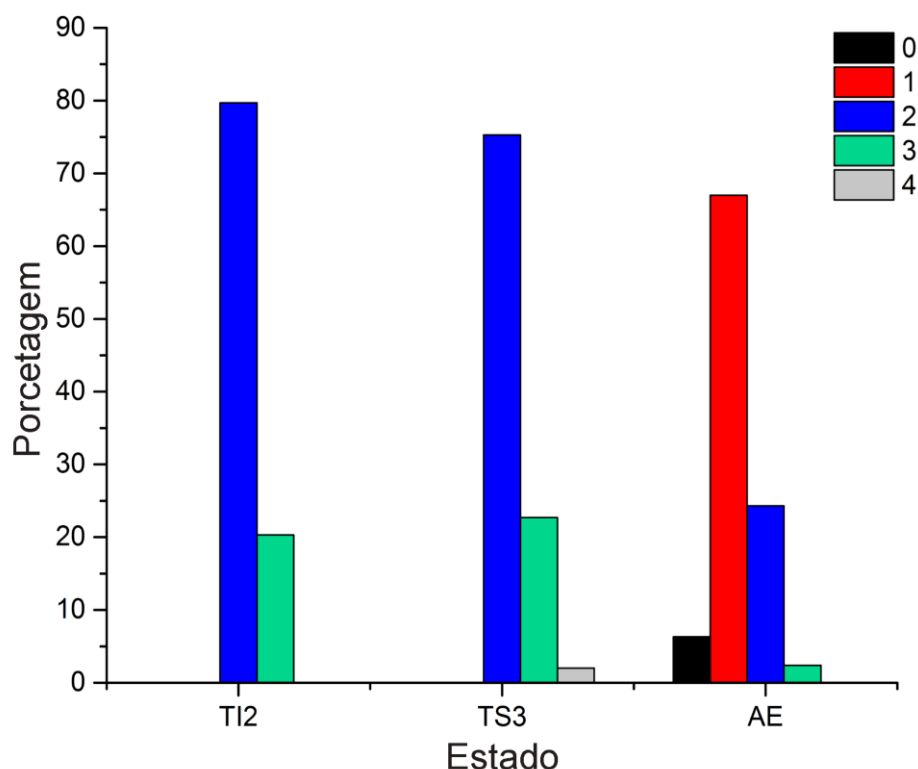
**Figura 87:** Cargas de Mulliken para alguns dos átomos da região de reação durante o processo de quebra da ligação peptídica. As barras coloridas são os valores médios para cada um dos pontos estacionários. O desvio padrão é mostrado pela barra em preto em cada uma das cargas. Isso corresponde a 30 ps em cada um dos pontos estacionários.

O nitrogênio HIS33- $N_{\epsilon 2}$  da histidina, também muda de característica com relação a carga, que está positiva (ácido) no estado TI2 e, ao final da reação com a quebra da ligação peptídica, volta a sua característica básica inicial (ES), com carga negativa.

#### 5.6.2.2 Cavidade Oxiânion

Nesta etapa da reação também investigamos a presença de moléculas de água dentro de um raio de 3,0 Å do oxigênio Subs- $O_{\gamma 2}$  (Figura 88). Observa-se que o número de moléculas de água nessa região diminui quando a ligação peptídica é clivada e o intermediário tetraédrico é desfeito. Enquanto que no início de coordenada de reação (TI2) a maioria dos frames apresentavam duas ou três moléculas de água nessa região, no final (AE) há predominantemente apenas uma molécula de água próxima ao Subs- $O_{\gamma 2}$ , e uma parte significativa dos frames (~5 %) não apresentam nenhuma molécula de  $H_2O$  na região. Isso

evidencia o papel das moléculas de água funcionando como a cavidade oxiânion para estabilizar o estado TI2. Depois de desfeito o TI2, a carga do Subs-O<sub>γ2</sub> volta ao seu estado normal e a cavidade oxiânion é desfeita.



**Figura 88:** Número de moléculas de água em percentagem dos frames de cada ponto estacionário. Cada uma das cores representa um número específico de moléculas de água. Isso corresponde a 30 ps em cada um dos pontos estacionários.

As ligações de hidrogênio no sítio ativo são fundamentais para a catálise da reação de hidrólise da ligação peptídica pela protease de serina (NS2B-NS3), pois a transferência de próton entre a serina (SER) e a histidina (HIS) é viável devido ao ácido aspártico (ASP) que estabiliza a histidina através de uma ligação de hidrogênio, proporcionando assim que a histidina haja como base na primeira etapa da reação, assim como, agindo como ácido na etapa de quebra da ligação peptídica. As ligações de hidrogênio são também importantes para a estabilização do intermediário tetraédrico na região da cavidade oxiânion. Portanto, ao usar o método aproximado PDDG-PM3, é esperado que o valor da energia de ativação para reação seja superestimado (3.7.1.5 Estudo com proteases, Página 69). Pois, em estudo anterior comparando os métodos semi-empíricos AM1 e PM3 em reações de hidrólise de amidas e ésteres catalisada por protease de serina, tripsina, foi mostrado que o método PM3 obteve

resultados superiores para geometrias de ligação de hidrogênio, obtendo-se geometrias semelhantes aos encontrados na estrutura do cristal (SCHROEDER; DAGGETT; KOLLMAN, 1991). No entanto, as barreiras de energia para sistemas de transferência de prótons que contêm ligações de hidrogênio são normalmente superestimadas com PM3 (MORPURGO; BOSSA; MORPURGO, 1998). A modificação PDDG-PM3 fornece uma descrição melhor das interações de atração de van der Waals entre átomos, produzindo energias de ligação de hidrogênio compatíveis para um conjunto de dímeros de água e dímeros de benzeno (REPASKY; CHANDRASEKHAR; JORGENSEN, 2002).

Deste modo, o método é capaz de indicar as tendências e estimar qual etapa é determinante para a cinética da reação.

## 6. CONSIDERAÇÕES FINAIS

---

Conforme com o que foi proposto na seção de objetivos específicos, todos os sistemas (DENVa2, DENV2f, DENV2f-INIB, DENV2f-SUBS1 e DENV2f-SUBS2) foram simulados por dinâmica molecular, e apresentaram concordância com dados experimentais.

Os cálculos mostram que a protease NS2B-NS3, apesar de existir predominantemente na forma aberta (inativa), pode existir também na conformação fechada (ativa) em solução, mesmo na ausência de substrato ou inibidor.

A simulação com o modelo DENV2f mostrou maior rigidez na sua estrutura quando o inibidor estiver aprisionado a ela. Entretanto, o mesmo comportamento não é observado no caso com o substrato-1, no qual a DENV2f apresenta maior mobilidade. Essa maior mobilidade da proteína quando na presença de moléculas maiores (mais de cinco resíduos) é altamente relevante, especialmente em trabalhos de *docking molecular* com uso de um conjunto de estruturas para representar a dinâmica da proteína. Ao simularmos com um substrato com quatro resíduos ( $\text{CH}_3\text{NH-RRPV-COCH}_3$ ) obtivemos resultados satisfatórios com relação a estabilidade estrutural da DENV2f, desta maneira pudemos usar o sistema DENV2f-SUBS2 para a etapa do mecanismo catalítico de quebra da ligação peptídica do substrato-2.

O modelo DENV2f proposto se mostrou capaz de ser usado para o estudo reacional de quebra da ligação peptídica entre resíduos do substrato da NS2B-NS3 protease da Dengue, bem como, para a modelagem na triagem de inibidores do vírus da Dengue.

A parte fundamental desse trabalho foi o estudo do mecanismo da reação catalisada pela serina protease (NS2B-NS3) do vírus da Dengue, usando simulações de dinâmica molecular com método híbrido QM/MM, e a técnica de amostragem “*Umbrella Sampling*” para obter a energia livre do processo. O perfil energético para a reação de acilação, ou seja, formação do intermediário tetraédrico, está dividido em duas partes. A primeira parte é a ativação da SER117, através da transferência do próton ( $\text{H}_{\gamma 1}$ ) da serina para o nitrogênio épsilon da HIS33, com energia livre de 24,1 kcal/mol. Para que essa etapa ocorra, é necessário que o hidrogênio delta da SER117 seja transferido para o ASP57, aumentando a basicidade da posição épsilon da SER117. A segunda etapa é o ataque nucleofílico da serina ao carbono carboxílico do substrato ( $\text{C}_{\gamma 2}$ ), com uma barreira de energia livre de 10,9 kcal/mol. O TI formado é estabilizado pela região chamada de cavidade oxiânion (*oxyanion hole*), que neste caso é formado por duas a três moléculas de água.

Na etapa seguinte da reação o próton  $H_{\gamma 1}$  é transferido para o nitrogênio  $N_{\gamma 2}$  do substrato, levando diretamente à quebra da ligação peptídica. O TI é mantido até que ocorra a quebra da ligação peptídica (-ARG-PRO-) entre  $N_{\gamma 2}$  e  $C_{\gamma 2}$ , com barreira de energia livre de 5,1 kcal/mol. Quando a estrutura tetraédrica é desfeita a região da cavidade oxiânion também se desfaz.

Durante toda a reação a redistribuição de carga entre átomos do sítio ativo, bem como o acesso de moléculas de água permitindo a estabilização do intermediário tetraédrico, são de fundamental importância para o processo catalítico.

Os cálculos apresentados mostram que o passo determinante para o processo é a formação do intermediário tetraédrico, pois o processo de decomposição deste intermediário tem um valor pequeno para a barreira de energia livre. A energia livre de ativação calculada é 24,1 kcal/mol, em boa concordância com os valores estimados experimentalmente e teoricamente para outras proteases de serina.(ISHIDA; KATO, 2003b; MARTÍNEZ-GONZÁLEZ et al., 2015; ZHANG et al., 1997)

A partir do conhecimento das estruturas de transição para a catalise enzimática é possível modelar inibidores análogos a esses estados, interferindo na reprodução do vírus da Dengue.

## 7. PERSPECTIVAS

---

Análogos de estado de transição são compostos que se assemelham ao estado de transição de uma reação catalisada. Análogos do estado de transição entram no sítio catalítico da mesma maneira que o substrato, e interagem com sítio catalítico até localizarem uma estrutura intimamente relacionada com o estado de transição, mas que é quimicamente estável. Esses compostos podem ser inibidores potentes para processos enzimáticos.

Portanto, a partir do conhecimento detalhado do mecanismo enzimático e das estruturas do estado de transição, da NS2B-NS3 protease da DENV2, será possível modelar inibidores analógico a esses estados, inibindo desse modo a enzima e interferindo no processo de replicação viral do vírus do dengue. Deste modo, é necessário gerar ainda mais configurações desses estados para podermos buscar nessas estruturas compostos que sejam quimicamente similares as essas interações, isso através de um *Virtual Screening* a partir de um banco de dados de moléculas.

Um outro ponto que ainda não está esclarecido na literatura, é se a estrutura aberta da NS2B-NS3 protease da Dengue também é capaz de catalisar o processo de replicação viral. Assim, pretendemos estudar essa mesma reação na estrutura aberta da NS2B-NS3 protease da Dengue (2FOM), verificando se há alguma mudança nas barreiras de energia para o processo.

O vírus da Zika faz parte da família *Flaviviridae* que incluindo o vírus da febre-amarela (YFV), vírus da hepatite C (VHC), e vírus da encefalite japonesa (JEV), assim como o vírus da Dengue. Portanto, a NS2B-NS3 protease também é responsável pela replicação desse vírus, deste modo, será possível o estudo de possíveis inibidores através dessa metodologia.

## REFERÊNCIAS

- AHLQUIST, P. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. **Nature Reviews Microbiology**, v. 4, n. 5, p. 371–382, 3 maio 2006.
- ALESHIN, A. E. et al. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. **Protein science : a publication of the Protein Society**, v. 16, n. 5, p. 795–806, maio 2007.
- ALLINGER, N. L. Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. **Journal of the American Chemical Society**, v. 99, n. 25, p. 8127–8134, dez. 1977.
- AMARA, P. et al. The generalized hybrid orbital method for combined quantum mechanical/molecular mechanical calculations: formulation and tests of the analytical derivatives. **Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)**, v. 104, n. 5, p. 336–343, 16 ago. 2000.
- ARAKAKI, T. L. et al. Catalytically active Dengue virus NS3 protease forms aggregates that are separable by size exclusion chromatography. **Protein expression and purification**, v. 25, n. 2, p. 241–7, jul. 2002.
- ARAVAPALLI, S. et al. Inhibitors of Dengue virus and West Nile virus proteases based on the aminobenzamide scaffold. **Bioorganic & medicinal chemistry**, v. 20, n. 13, p. 4140–8, jul. 2012.
- ARNOLD, K. et al. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. **Bioinformatics (Oxford, England)**, v. 22, n. 2, p. 195–201, 15 jan. 2006.
- ASH, E. L. et al. Unusual <sup>1</sup>H NMR chemical shifts support (His) C(epsilon) 1...O==C H-bond: proposal for reaction-driven ring flip mechanism in serine protease catalysis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 97, n. 19, p. 10371–6, set. 2000.
- BACHOVCHIN, W. W. Review: Contributions of NMR spectroscopy to the study of hydrogen bonds in serine protease active sites. **Magnetic Resonance in Chemistry**, v. 39, n. S1, p. S199–S213, dez. 2001.
- BARRETO, M. L.; TEIXEIRA, M. G. Dengue no Brasil: situação epidemiológica e



contribuições para uma agenda de pesquisa. **Estudos Avançados**, v. 22, n. 64, p. 53–72, dez. 2008.

BAYLY, C. I. et al. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. **The Journal of Physical Chemistry**, v. 97, n. 40, p. 10269–10280, out. 1993.

BENEDICT, M. The first releases of transgenic mosquitoes: an argument for the sterile insect technique. **Trends in Parasitology**, v. 19, n. 8, p. 349–355, ago. 2003.

BENTZIEN, J. et al. Hybrid ab Initio Quantum Mechanics/Molecular Mechanics Calculations of Free Energy Surfaces for Enzymatic Reactions: The Nucleophilic Attack in Subtilisin. **The Journal of Physical Chemistry B**, v. 102, n. 12, p. 2293–2301, mar. 1998.

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

BHATT, S. et al. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504–7, 25 abr. 2013.

BLOW, D.; BIRKTOFT, J.; HARTLEY, B. Role of a buried acid group in the mechanism of action of chymotrypsin. **Nature**, v. 221, n. 5178, p. 337–340, 25 jan. 1969a.

BLOW, D. M. Structure and mechanism of chymotrypsin. **Accounts of Chemical Research**, v. 9, n. 4, p. 145–152, abr. 1976.

BLOW, D. M.; BIRKTOFT, J. J.; HARTLEY, B. S. Role of a Buried Acid Group in the Mechanism of Action of Chymotrypsin. **Nature**, v. 221, n. 5178, p. 337–340, 25 jan. 1969b.

BODENREIDER, C. et al. A fluorescence quenching assay to discriminate between specific and nonspecific inhibitors of dengue virus protease. **Analytical Biochemistry**, v. 395, n. 2, p. 195–204, 2009.

BORDOLI, L. et al. Protein structure homology modeling using SWISS-MODEL workspace. **Nature protocols**, v. 4, n. 1, p. 1–13, 2009.

CASE, D. et al. **Amber** University of California, San Francisco, 2010a.

CASE, D. A. et al. **AMBER 11** San Francisco University of California, , 2010b.

CHAMBERS, T. J. et al. Flavivirus Genome Organization, Expression and Replication. **Annual Reviews in Microbiology**, v. 44, p. 649–688, 1990.

CHAPPELL, K. J. et al. Mutagenesis of the West Nile virus NS2B cofactor domain reveals

two regions essential for protease activity. **The Journal of general virology**, v. 89, n. Pt 4, p. 1010–4, abr. 2008.

CHEN, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta crystallographica. Section D, Biological crystallography**, v. 66, n. Pt 1, p. 12–21, jan. 2010.

CLELAND, W. W.; KREEVOY, M. M. Low-barrier hydrogen bonds and enzymic catalysis. **Science (New York, N.Y.)**, v. 264, n. 5167, p. 1887–90, 24 jun. 1994.

CLUM, S. Cotranslational Membrane Insertion of the Serine Proteinase Precursor NS2B-NS3(Pro) of Dengue Virus Type 2 Is Required for Efficient in Vitro Processing and Is Mediated through the Hydrophobic Regions of NS2B. **Journal of Biological Chemistry**, v. 272, n. 49, p. 30715–30723, dez. 1997.

CORNELL, W. D. et al. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. **Journal of the American Chemical Society**, v. 117, n. 19, p. 5179–5197, maio 1995.

COSTA, C. N. Dengue : twenty-five years since reemergence in Brazil Dengue : vinte e cinco anos da reemergência no Brasil. p. 7–18, 2009.

COUGHLIN, S. R. Thrombin signalling and protease-activated receptors. **Nature**, v. 407, n. 6801, p. 258–264, 14 set. 2000.

CRAMER, C. J. **Essentials of computational chemistry : theories and models**. [s.l.] Wiley, 2004.

DAGGETT, V.; SCHROEDER, S.; KOLLMAN, P. Catalytic pathway of serine proteases: classical and quantum mechanical calculations. **Journal of the American Chemical Society**, v. 113, n. 14, p. 8926–8935, 1991.

DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. **The Journal of Chemical Physics**, v. 98, n. 12, p. 10089, 1993.

DE ALMEIDA, H. et al. New binding site conformations of the dengue virus NS3 protease accessed by molecular dynamics simulation. **PloS one**, v. 8, n. 8, p. e72402, jan. 2013.

DEWAR, M. J. S. et al. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. **Journal of the American Chemical Society**, v. 107, n. 13, p. 3902–3909, jun. 1985.

- DEWAR, M. J. S.; THIEL, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. **Journal of the American Chemical Society**, v. 99, n. 15, p. 4899–4907, jun. 1977.
- EKICI, O. D.; PAETZEL, M.; DALBEY, R. E. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. **Protein science : a publication of the Protein Society**, v. 17, n. 12, p. 2023–37, dez. 2008.
- ERAMIAN, D. et al. How well can the accuracy of comparative protein structure models be predicted? **Protein science : a publication of the Protein Society**, v. 17, n. 11, p. 1881–93, nov. 2008.
- ERBEL, P. et al. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. **Nature Structural & Molecular Biology**, v. 13, n. 4, p. 372–373, 12 mar. 2006.
- ERLANGER, B. F. et al. Phenothiazine-N-carbonyl chloride, a specific inactivator of chymotrypsin. **The Biochemical journal**, v. 118, n. 3, p. 421–5, jul. 1970.
- FALGOUT, B.; MARKOFF, L. Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum . Evidence that Flavivirus NS1-NS2A Cleavage Is Mediated by a Membrane-Bound Host Protease in the Endoplasmic Reticulum. **Microbiology**, v. 69, n. 11, 1995.
- FALGOUT, B.; MILLER, R. H.; LAI, C. Deletion Analysis of Dengue Virus Type 4 Nonstructural Protein NS2B : Identification of a Domain Required for NS2B-NS3 Protease Activity CATATGGCTAGCTCTTGGCCTC-TCTTAACGAGGGCATAATGGCTGTAGQGCCGTTAGTCTC. v. 67, n. 4, p. 2034–2042, 1993.
- FARRELL, C. &. **Biochemistry, 6th Edition**. [s.l.] Thomas-Brooks / Cole, 2008.
- FEIG, M.; BROOKS, C. L. Recent advances in the development and application of implicit solvent models in biomolecule simulations. **Current Opinion in Structural Biology**, v. 14, n. 2, p. 217–224, abr. 2004.
- FERSHT, A. R.; REQUENA, Y. Free energies of hydrolysis of amides and peptides in aqueous solution at 25 degrees. **Journal of the American Chemical Society**, v. 93, n. 14, p. 3499–504, 14 jul. 1971.
- FIELD, M. J. **A Practical Introduction to the Simulation of Molecular Systems**. [s.l.]

Cambridge University Press, 2007.

FODOR, K. et al. Enzyme:substrate hydrogen bond shortening during the acylation phase of serine protease catalysis. **Biochemistry**, v. 45, n. 7, p. 2114–21, mar. 2006.

FRANCO, C. et al. The dengue threat to the United States. **Biosecurity and bioterrorism : biodefense strategy, practice, and science**, v. 8, n. 3, p. 273–6, set. 2010.

FRANZ, A. W. E. et al. Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 11, p. 4198–203, 14 mar. 2006.

FRENKEL, D.; SMIT, B. **Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science)**. [s.l.] Academic Press, 2001.

FREY, P. A.; WHITT, S. A.; TOBIN, J. B. A low-barrier hydrogen bond in the catalytic triad of serine proteases. **Science (New York, N.Y.)**, v. 264, n. 5167, p. 1927–30, 24 jun. 1994.

GAO, J. et al. A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. v. 5639, n. 98, p. 4714–4721, 1998.

GARCIA-VILOCA, M.; GAO, J. Regular article Generalized hybrid orbital for the treatment of boundary atoms in combined quantum mechanical and molecular mechanical calculations using the semiempirical parameterized model 3 method. p. 280–286, 2004.

HADINEGORO, S. R. et al. Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease. 23 set. 2015.

HARPER, E.; BERGER, A. On the size of the active site in proteases: pronase. **Biochemical and biophysical research communications**, v. 46, n. 5, p. 1956–60, 10 mar. 1972.

HARTLEY, B. S. AMINO-ACID SEQUENCE OF BOVINE CHYMOTRYPSINOGEN-A. **Nature**, v. 201, p. 1284–7, mar. 1964.

HARTLEY, B. S.; KILBY, B. A. The reaction of p-nitrophenyl esters with chymotrypsin and insulin. **The Biochemical journal**, v. 56, n. 2, p. 288–97, fev. 1954.

HEDSTROM, L. Introduction: Proteases. **Chemical Reviews**, v. 102, n. 12, p. 4429–4430, dez. 2002a.

HEDSTROM, L. Serine protease mechanism and specificity. **Chemical reviews**, v. 102, n. 12, p. 4501–24, dez. 2002b.

HEDSTROM, L. Serine protease mechanism and specificity. **Chemical reviews**, v. 102, n.

12, p. 4501–24, dez. 2002c.

HENCHAL, E. A; PUTNAK, J. R. The dengue viruses. **Clinical microbiology reviews**, v. 3, n. 4, p. 376–96, out. 1990.

HESS, G. P. et al. Studies of the Activity of Chymotrypsin. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 257, n. 813, p. 89–104, fev. 1970a.

HESS, G. P. et al. Studies of the Activity of Chymotrypsin. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 257, n. 813, p. 89–104, 12 fev. 1970b.

HIROARA, H.; BENDER, M. L.; STARK, R. S. Acylation of alpha-chymotrypsin by oxygen and sulfur esters of specific substrates: kinetic evidence for a tetrahedral intermediate.

**Proceedings of the National Academy of Sciences of the United States of America**, v. 71, n. 5, p. 1643–7, maio 1974.

HORNAK, V. et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. **Proteins: Structure, Function, and Bioinformatics**, v. 65, n. 3, p. 712–725, 15 nov. 2006.

HUB, J. S.; DE GROOT, B. L.; VAN DER SPOEL, D. G-whams-a free Weighted Histogram Analysis implementation including robust error and autocorrelation estimates. **Journal of Chemical Theory and Computation**, v. 6, n. 12, p. 3713–3720, 2010.

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

HUNKAPILLER, M. W. et al. Carbon nuclear magnetic resonance studies of the histidine residue in  $\alpha$ -lytic protease. Implications for the catalytic mechanism of serine proteases. **Biochemistry**, v. 12, n. 23, p. 4732–4743, nov. 1973.

ISHIDA, T. Low-barrier hydrogen bond hypothesis in the catalytic triad residue of serine proteases: correlation between structural rearrangement and chemical shifts in the acylation process. **Biochemistry**, v. 45, n. 17, p. 5413–20, maio 2006.

ISHIDA, T.; KATO, S. Theoretical perspectives on the reaction mechanism of serine proteases: the reaction free energy profiles of the acylation process. **Journal of the American Chemical Society**, v. 125, n. 39, p. 12035–48, out. 2003a.

ISHIDA, T.; KATO, S. Theoretical perspectives on the reaction mechanism of serine proteases: the reaction free energy profiles of the acylation process. **Journal of the American Chemical Society**, v. 125, n. 39, p. 12035–48, 1 out. 2003b.

- ISHIDA, T.; KATO, S. Role of Asp102 in the catalytic relay system of serine proteases: a theoretical study. **Journal of the American Chemical Society**, v. 126, n. 22, p. 7111–8, jun. 2004.
- JENSEN, F. **Introduction to computational chemistry**. [s.l.] John Wiley & Sons, 2007.
- JOHN, B.; SALI, A. Comparative protein structure modeling by iterative alignment, model building and model assessment. **Nucleic acids research**, v. 31, n. 14, p. 3982–92, 15 jul. 2003.
- JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. **The Journal of Chemical Physics**, v. 79, n. 2, p. 926, 1983.
- KARPLUS, M.; KURIYAN, J. Molecular dynamics and protein function. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 19, p. 6679–85, 10 maio 2005.
- KARPLUS, M.; MCCAMMON, J. A. Molecular dynamics simulations of biomolecules. **Nature structural biology**, v. 9, n. 9, p. 646–52, set. 2002.
- KÄSTNER, J. Umbrella sampling. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, v. 1, n. December, p. n/a-n/a, 26 maio 2011.
- KIM, Y. M. et al. NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. **The Journal of biological chemistry**, v. 288, n. 18, p. 12891–900, 3 maio 2013.
- KLINE, K. et al. Neglected Tropical Diseases of Oceania: Review of Their Prevalence, Distribution, and Opportunities for Control. **PLoS Neglected Tropical Diseases**, v. 7, n. 1, p. e1755, 31 jan. 2013.
- KOLLMAN, P. Molecular Modeling. **Annual Review of Physical Chemistry**, v. 38, n. 1, p. 303–316, out. 1987.
- KOLLMAN, P. A.; HAYES, D. M. Theoretical calculations on proton-transfer energetics: studies of methanol, imidazole, formic acid, and methaneethiol as models for the serine and cysteine proteases. **Journal of the American Chemical Society**, v. 103, n. 11, p. 2955–2961, jun. 1981.
- KONISHI, E. Issues related to recent dengue vaccine development. **Tropical medicine and health**, v. 39, n. 4 Suppl, p. 63–71, dez. 2011.

- KOSSIAKOFF, A. A; SPENCER, S. A. Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: neutron structure of trypsin. **Biochemistry**, v. 20, n. 22, p. 6462–6474, 27 out. 1981.
- KRAEMER, M. U. et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. **eLife**, v. 4, n. JUNE2015, p. 1–18, 30 jun. 2015.
- KRAUT, J. Serine Proteases: Structure and Mechanism of Catalysis. **Annual Review of Biochemistry**, v. 46, n. 1, p. 331–358, jun. 1977.
- KUMAR, S. et al. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. **Journal of Computational Chemistry**, v. 13, n. 8, p. 1011–1021, 1992.
- LA RUCHE, G. et al. First two autochthonous dengue virus infections in metropolitan France, September 2010. **Euro surveillance : bulletin Européen sur les maladies transmissibles = European communicable disease bulletin**, v. 15, n. 39, p. 19676, set. 2010.
- LEACH, A. R. **Molecular modelling : principles and applications**. [s.l.] Prentice Hall, 2001.
- LI, H. et al. The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 converge within a region of 20 amino acids. **Journal of virology**, v. 73, n. 4, p. 3108–16, abr. 1999.
- LI, J. et al. Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. **The Journal of biological chemistry**, v. 280, n. 31, p. 28766–74, 5 ago. 2005.
- LII, J. H.; ALLINGER, N. L. Molecular mechanics. The MM3 force field for hydrocarbons. 3. The van der Waals' potentials and crystal data for aliphatic and aromatic hydrocarbons. **Journal of the American Chemical Society**, v. 111, n. 23, p. 8576–8582, nov. 1989.
- LIPMAN, D. J.; PEARSON, W. R. Rapid and sensitive protein similarity searches. **Science**, v. 227, n. 4693, p. 1435–1441, 1985.
- M. J. FRISCH, G. W. TRUCKS, H. B. SCHLEGEL, G. E. SCUSERIA, M. A. ROBB, J. R. CHEESEMAN, G. SCALMANI, V. BARONE, B. MENNUCCI, G. A. PETERSSON, H. NAKATSUJI, M. CARICATO, X. LI, H. P. HRATCHIAN, A. F. IZMAYLOV, J. BLOINO, G. ZHENG, J. L. SONNENBERG, M. HAD, AND D. J. F. **Gaussian09** Wallingford CT Gaussian, Inc., , 2009. Disponível em:

<[http://www.gaussian.com/g\\_tech/g\\_ur/m\\_citation.htm](http://www.gaussian.com/g_tech/g_ur/m_citation.htm)>. Acesso em: 16 jan. 2014

MACKERELL, A. D. Empirical force fields for biological macromolecules: Overview and issues. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1584–1604, out. 2004.

MARTÍ-RENOM, M. A. et al. Comparative protein structure modeling of genes and genomes. **Annual review of biophysics and biomolecular structure**, v. 29, p. 291–325, 2000.

MARTÍNEZ-GONZÁLEZ, J. Á. et al. Theoretical Study of the Free Energy Surface and Kinetics of the Hepatitis C Virus NS3/NS4A Serine Protease Reaction with the NS5A/5B Substrate. Does the Generally Accepted Tetrahedral Intermediate Really Exist? **ACS Catalysis**, v. 5, n. 1, p. 246–255, 2 jan. 2015.

MATTHEWS, B. W. et al. Three-dimensional structure of tosyl-alpha-chymotrypsin. **Nature**, v. 214, n. 5089, p. 652–6, maio 1967a.

MATTHEWS, B. W. et al. Three-dimensional structure of tosyl-alpha-chymotrypsin. **Nature**, v. 214, n. 5089, p. 652–6, 13 maio 1967b.

MCCAMMON, J. A.; GELIN, B. R.; KARPLUS, M. Dynamics of folded proteins. **Nature**, v. 267, n. 5612, p. 585–90, 16 jun. 1977.

MORPURGO, S.; BOSSA, M.; MORPURGO, G. O. **Critical test of PM3-calculated proton transfer activation energies: a comparison with ab initio and AM1 calculations** **Journal of Molecular Structure: THEOCHEM**, 1998. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0166128097002194>>

MURRELL, S.; WU, S.-C.; BUTLER, M. Review of dengue virus and the development of a vaccine. **Biotechnology advances**, v. 29, n. 2, p. 239–47, 2011.

NANDA, D. N.; JUG, K. SINDO1. A semiempirical SCF MO method for molecular binding energy and geometry I. Approximations and parametrization. **Theoretica Chimica Acta**, v. 57, n. 2, p. 95–106, 1980.

NATARAJAN, S. NS3 protease from flavivirus as a target for designing antiviral inhibitors against dengue virus. **Genetics and Molecular Biology**, v. 33, n. 2, p. 214–219, 2010.

NEURATH, H. Evolution of proteolytic enzymes. **Science (New York, N.Y.)**, v. 224, n. 4647, p. 350–7, 27 abr. 1984.

NITSCHKE, C. et al. Retro peptide-hybrids as selective inhibitors of the Dengue virus NS2B-



NS3 protease. **Antiviral research**, v. 94, n. 1, p. 72–9, abr. 2012.

NITSCHKE, C.; STEUER, C.; KLEIN, C. D. Arylcianoacrylamides as inhibitors of the Dengue and West Nile virus proteases. **Bioorganic & medicinal chemistry**, v. 19, n. 24, p. 7318–37, dez. 2011.

NOBLE, C. G. et al. Ligand-bound structures of the dengue virus protease reveal the active conformation. **Journal of virology**, v. 86, n. 1, p. 438–46, jan. 2012.

NOBLE, C. G.; SHI, P.-Y. Structural biology of dengue virus enzymes: Towards rational design of therapeutics. **Antiviral research**, v. 96, n. 2, p. 115–126, 17 set. 2012.

PARKINSON, T.; PRYDE, D. C. Small molecule drug discovery for Dengue and West Nile viruses: applying experience from hepatitis C virus. **Future medicinal chemistry**, v. 2, n. 7, p. 1181–203, jul. 2010.

PETREY, D. et al. Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. **Proteins**, v. 53 Suppl 6, p. 430–5, 2003.

POLGÁR, L. The catalytic triad of serine peptidases. **Cellular and molecular life sciences : CMLS**, v. 62, n. 19–20, p. 2161–72, out. 2005.

QI, R.; ZHANG, L.; CHI, C. Biological characteristics of dengue virus and potential targets for drug design. **Acta Biochimica et Biophysica Sinica**, v. 40, n. 2, p. 91–101, fev. 2008a.

QI, R.; ZHANG, L.; CHI, C. Biological characteristics of dengue virus and potential targets for drug design. **Acta Biochimica et Biophysica Sinica**, v. 40, n. 2, p. 91–101, fev. 2008b.

REPASKY, M. P.; CHANDRASEKHAR, J.; JORGENSEN, W. L. PDDG/PM3 and PDDG/MNDO: improved semiempirical methods. **Journal of computational chemistry**, v. 23, n. 16, p. 1601–22, dez. 2002.

ROBILLARD, G.; SHULMAN, R. G. High resolution nuclear magnetic resonance study of the histidine–aspartate hydrogen bond in chymotrypsin and chymotrypsinogen. **Journal of molecular biology**, v. 71, n. 2, p. 507–11, nov. 1972.

ROCHA, G. B. et al. RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. **Journal of Computational Chemistry**, v. 27, n. 10, p. 1101–1111, 30 jul. 2006.

ROOTHAAN, C. C. J. New Developments in Molecular Orbital Theory. **Reviews of Modern Physics**, v. 23, n. 2, p. 69–89, 1 abr. 1951.

ROUX, B.; SIMONSON, T. Implicit solvent models. **Biophysical Chemistry**, v. 78, n. 1–2,

p. 1–20, abr. 1999.

RYCKAERT, J.-P.; CICCOTTI, G.; BERENDSEN, H. J. . Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. **Journal of Computational Physics**, v. 23, n. 3, p. 327–341, mar. 1977.

SALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of molecular biology**, v. 234, n. 3, p. 779–815, 5 dez. 1993.

SCHEINER, S.; LIPSCOMB-T, W. N. Molecular orbital studies of enzyme activity : Catalytic mechanism of serine proteinases \* *Biochemistry* : v. 73, n. 2, p. 432–436, 1976.

SCHOLTEN, J. D.; HOGG, J. L.; RAUSHEL, F. M. Methyl chymotrypsin catalyzed hydrolyses of specific substrate esters indicate multiple proton catalysis is possible with a modified charge relay triad. **Journal of the American Chemical Society**, v. 110, n. 24, p. 8246–8247, nov. 1988.

SCHRAMM, V. L. Enzymatic Transition States , Transition-State Analogs , Dynamics , Thermodynamics , and Lifetimes. 2011.

SCHROEDER, S.; DAGGETT, V.; KOLLMAN, P. A comparison of the AM1 and PM3 semiempirical models for evaluating model compounds relevant to catalysis by serine proteases. **Journal of the American Chemical Society**, v. 113, n. 23, p. 8922–8925, nov. 1991.

SHEPARD, D. S. et al. The global economic burden of dengue: a systematic analysis. **The Lancet Infectious Diseases**, abr. 2016.

SHIRYAEV, S. A et al. Cleavage preference distinguishes the two-component NS2B-NS3 serine proteinases of Dengue and West Nile viruses. **The Biochemical journal**, v. 401, n. 3, p. 743–52, fev. 2007.

SHIRYAEV, S. A et al. Virtual ligand screening of the National Cancer Institute (NCI) compound library leads to the allosteric inhibitory scaffolds of the West Nile Virus NS3 proteinase. **Assay and drug development technologies**, v. 9, n. 1, p. 69–78, fev. 2011.

SLATER, J. C. **Quantum theory of matter**. [s.l.] McGraw-Hill, 1968.

SPENCER, T.; STURTEVANT, J. M. The Mechanism of Chymotrypsin-catalyzed Reactions. III. **Journal of the American Chemical Society**, v. 81, n. 8, p. 1874–1882, abr. 1959.

SPRANG, S. et al. **The three-dimensional structure of Asn102 mutant of trypsin: role of**

**Asp102 in serine protease catalysis.** *Science (New York, N.Y.)*, ago. 1987.

STEUER, C. et al. Synthesis and biological evaluation of  $\alpha$ -ketoamides as inhibitors of the Dengue virus protease with antiviral activity in cell-culture. **Bioorganic & medicinal chemistry**, v. 19, n. 13, p. 4067–74, jul. 2011.

STEWART, J. J. P. Optimization of parameters for semiempirical methods I. Method. **Journal of Computational Chemistry**, v. 10, n. 2, p. 209–220, mar. 1989.

STROUD, R. M.; KAY, L. M.; DICKERSON, R. E. The structure of bovine trypsin: electron density maps of the inhibited enzyme at 5 Angstrom and at 2-7 Angstrom resolution. **Journal of molecular biology**, v. 83, n. 2, p. 185–208, fev. 1974.

SU, X.-C. et al. NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B-NS3 protease. **PLoS neglected tropical diseases**, v. 3, n. 12, p. e561, jan. 2009.

THIEL, W. Chapter 21 - Semiempirical quantum-chemical methods in computational chemistry. **Theory and Applications of Computational Chemistry**, p. 559–580, 2005.

TIEW, K.-C. et al. Inhibition of Dengue virus and West Nile virus proteases by click chemistry-derived benz[d]isothiazol-3(2H)-one derivatives. **Bioorganic & medicinal chemistry**, v. 20, n. 3, p. 1213–21, fev. 2012.

TOPF, M.; RICHARDS, W. G. Theoretical studies on the deacylation step of serine protease catalysis in the gas phase, in solution, and in elastase. **Journal of the American Chemical Society**, v. 126, n. 44, p. 14631–41, 10 nov. 2004.

TOPF, M.; VÁRNAI, P.; RICHARDS, W. G. Ab initio QM/MM dynamics simulation of the tetrahedral intermediate of serine proteases: Insights into the active site hydrogen-bonding network. **Journal of the American Chemical Society**, v. 124, n. 49, p. 14780–14788, 2002a.

TOPF, M.; VÁRNAI, P.; RICHARDS, W. G. Ab Initio QM/MM Dynamics Simulation of the Tetrahedral Intermediate of Serine Proteases: Insights into the Active Site Hydrogen-Bonding Network. **Journal of the American Chemical Society**, v. 124, n. 49, p. 14780–14788, dez. 2002b.

TORRIE, G. M.; VALLEAU, J. P. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. **Chemical Physics Letters**, v. 28, n. 4, p. 578–581, out. 1974.

TORRIE, G. M.; VALLEAU, J. P. Nonphysical sampling distributions in Monte Carlo free-

energy estimation: Umbrella sampling. **Journal of Computational Physics**, v. 23, n. 2, p. 187–199, 1977.

TUBERT-BROHMAN, I. et al. Extension of the PDDG/PM3 and PDDG/MNDO semiempirical molecular orbital methods to the halogens. **Journal of Computational Chemistry**, v. 25, n. 1, p. 138–150, 15 jan. 2004.

TURK, B. Targeting proteases: successes, failures and future prospects. **Nature Reviews Drug Discovery**, v. 5, n. 9, p. 785–799, set. 2006.

UBERUAGA, B. P.; ANGHEL, M.; VOTER, A. F. Synchronization of trajectories in canonical molecular-dynamics simulations: observation, explanation, and exploitation. **The Journal of chemical physics**, v. 120, n. 14, p. 6363–74, 8 abr. 2004.

UMEYAMA, H.; HIRONO, S.; NAKAGAWA, S. Charge state of His-57-Asp-102 couple in a transition state analogue-trypsin complex: a molecular orbital study. **Proceedings of the National Academy of Sciences of the United States of America**, v. 81, n. 20, p. 6266–70, out. 1984.

UMEYAMA, H.; IMAMURA, A.; NAGATA, C. A molecular orbital study on the enzymic reaction mechanism of alpha-chymotrypsin. **Journal of theoretical biology**, v. 41, n. 3, p. 485–502, out. 1973.

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

VIGILÂNCIA, S. DE. **Boletim Epidemiológico Secretaria de Vigilância em Saúde – Ministério da Saúde** Boletim Epidemiol. [s.l: s.n.]. Disponível em:

<<http://u.saude.gov.br/images/pdf/2015/marco/25/Boletim-tuberculose-2015.pdf>>. Acesso em: 21 jun. 2016.

VOTANO, J.; PARHAM, M.; HALL, L. **A Practical Introduction to the Simulation of Molecular Systems**. [s.l: s.n.].

WAHLGREN, W. Y. et al. The catalytic aspartate is protonated in the Michaelis complex formed between trypsin and an in vitro evolved substrate-like inhibitor: a refined mechanism of serine protease action. **The Journal of biological chemistry**, v. 286, n. 5, p. 3587–96, fev. 2011a.

WAHLGREN, W. Y. et al. The catalytic aspartate is protonated in the Michaelis complex formed between trypsin and an in vitro evolved substrate-like inhibitor: a refined mechanism

of serine protease action. **The Journal of biological chemistry**, v. 286, n. 5, p. 3587–96, 4 fev. 2011b.

WALKER, R. C.; CROWLEY, M. F.; CASE, D. A. The implementation of a fast and accurate QM/MM potential method in Amber. **Journal of computational chemistry**, v. 29, n. 7, p. 1019–31, maio 2008.

WALSH, K. A.; NEURATH, H. TRYPSINOGEN AND CHYMOTRYPSINOGEN AS HOMOLOGOUS PROTEINS. **Proceedings of the National Academy of Sciences of the United States of America**, v. 52, p. 884–9, out. 1964.

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

WANG, S.; HU, P.; ZHANG, Y. Ab initio quantum mechanical/molecular mechanical molecular dynamics simulation of enzyme catalysis: The case of histone lysine methyltransferase SET7/9. **Journal of Physical Chemistry B**, v. 111, n. 14, p. 3758–3764, 2007.

WARSHEL, A. **Computer Modeling of Chemical Reactions in Enzymes and Solution**. [s.l.] Wiley, 1991.

WARSHEL, A.; LEVITT, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. **Journal of molecular biology**, p. 227–249, 1976a.

WARSHEL, A.; LEVITT, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. **Journal of molecular biology**, p. 227–249, 1976b.

WARSHEL, A.; PAPAZYAN, A.; KOLLMAN, P. A. On low-barrier hydrogen bonds and enzyme catalysis. **Science**, v. 269, n. 5220, p. 102, 1995.

WARSHEL, A.; RUSSELL, S. Theoretical correlation of structure and energetics in the catalytic reaction of trypsin. **J. Am. Chem. Soc.**, v. 108, n. 21, p. 6569–6579, out. 1986.

WARSHEL, A.; AQVIST, J.; CREIGHTON, S. Enzymes work by solvation substitution rather than by desolvation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 86, n. 15, p. 5820–4, ago. 1989.

WEAVER, S. C.; VASILAKIS, N. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral

disease. **Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases**, v. 9, n. 4, p. 523–40, jul. 2009.

WEINER, S. J. et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. **Journal of the American Chemical Society**, v. 106, n. 3, p. 765–784, fev. 1984.

WEINER, S. J.; SINGH, U. C.; KOLLMAN, P. A. Simulation of formamide hydrolysis by hydroxide ion in the gas phase and in aqueous solution. **Journal of the American Chemical Society**, v. 107, n. 8, p. 2219–2229, abr. 1985.

**WHO | Dengue and severe dengue**. Disponível em:

<<http://www.who.int/mediacentre/factsheets/fs117/en/>>. Acesso em: 21 jun. 2016.

WICHAPONG, K. et al. Homology modeling and molecular dynamics simulations of Dengue virus NS2B/NS3 protease: insight into molecular interaction. **Journal of molecular recognition : JMR**, v. 23, p. 283–300, 2010a.

WICHAPONG, K. et al. Homology modeling and molecular dynamics simulations of Dengue virus NS2B/NS3 protease: insight into molecular interaction. **Journal of molecular recognition : JMR**, v. 23, n. 3, p. 283–300, maio 2010b.

WORLD HEALTH ORGANIZATION. **Dengue: guidelines for diagnosis, treatment, prevention and control**. [s.l.: s.n.].

WORLD HEALTH ORGANIZATION. **Dengue: guidelines for diagnosis, treatment, prevention and control**. [s.l.: s.n.].

YIN, Z. et al. Peptide inhibitors of dengue virus NS3 protease . Part 1 : Warhead. v. 16, p. 36–39, 2006a.

YIN, Z. et al. Peptide inhibitors of dengue virus NS3 protease . Part 2 : SAR study of tetrapeptide aldehyde inhibitors. v. 16, p. 40–43, 2006b.

YOUNG, D. **Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems**. [s.l.] Wiley, 2004. v. 9

YUSOF, R. et al. Purified NS2B / NS3 Serine Protease of Dengue Virus Type 2 Exhibits Cofactor NS2B Dependence for Cleavage of Substrates with Dibasic Amino Acids in Vitro \*. **Biochemistry**, v. 275, n. 14, p. 9963–9969, 2000.

ZERNER, B.; BENDER, M. L. The Kinetic Consequences of the Acyl-Enzyme Mechanism

for the Reactions of Specific Substrates with Chymotrypsin. **Journal of the American Chemical Society**, v. 86, n. 18, p. 3669–3674, set. 1964.

ZERNER, M. C. Semiempirical Molecular Orbital Methods. In: **Reviews in Computational Chemistry, Volume 2**. [s.l.] John Wiley & Sons, Inc., 1991. p. 313–365.

ZHANG, R. et al. Probing the substrate specificity of hepatitis C virus NS3 serine protease by using synthetic peptides. **Journal of virology**, v. 71, n. 8, p. 6208–13, ago. 1997.

ZHANG, X. et al. Cryo-EM structure of the mature dengue virus at 3.5-Å resolution. **Nature Structural & Molecular Biology**, v. 20, n. 1, p. 105–110, 16 dez. 2012.

ZHOU, Y.; ZHANG, Y. Serine protease acylation proceeds with a subtle re-orientation of the histidine ring at the tetrahedral intermediate. **Chemical communications (Cambridge, England)**, v. 47, n. 5, p. 1577–9, 7 fev. 2011a.

ZHOU, Y.; ZHANG, Y. Serine protease acylation proceeds with a subtle re-orientation of the histidine ring at the tetrahedral intermediate. **Chemical communications (Cambridge, England)**, v. 47, n. 5, p. 1577–9, 7 fev. 2011b.

## APÊNDICE A: Protocolos de Simulação

---

Aqui encontram-se reproduzidos os arquivos de controle das simulações.

### **Minimização – Etapa -1 – Minimização do solvente**

```
&cntrl
    imin=1,          (indica que será realizado cálculo de minimização)
    maxcyc=1000,     (número de passos para a minimização)
    ncyc=100,        (mudará o método de minimização após 1000 passos de “steepest descent” para
                     gradiente conjugado)
    ntp=1000,        (a cada 1000 passo o programa escreve no arquivo de saída)
    ntmin=1,         (comando sobre o método de minimização, neste caso, steepest descent e
                     gradiente conjugado, juntos)
    ntb=1,           (uso de condições periódicas de contorno à volume constante)
    cut=10.0,        (uso de um raio de corte, cutoff, para as interações não ligadas)
    restraint_wt=100.0, (valor da constante de força em Kcal/mol.Å2 para a
                     restrição das posição atômicas dadas em restraintmask)
    restraintmask=':1-197' (Restrição para que o soluto, proteína, fique fixa)
/
```

### **Minimização – Etapa -2 – Minimização do sistema**

```
&cntrl
    imin=1, ntx=1,
    maxcyc=1000, ncyc=100, ntp=1000,
    ntmin=1,
    ntb=1,
    cut=10.0,
/
```

### **NVT-1 (Termalização 0K – 100K)**

```
&cntrl
    ntx=1,          (as coordenadas são lidas no arquivo de entrada, e as velocidades iniciais são
                     geradas pela distribuição de Boltzmann)
    ntb=1           (uso de condições periódicas de contorno a volume constante)
    cut=10.0,
    ntp=1000,       (intervalo de passos para escrever o arquivo de saída)
    ntwx=1000,      (intervalo de passos para escrever o arquivo da trajetória)
    ntwr=10000,     (intervalo de passos para escrever o arquivo de coordenadas da última
                     configuração)
```



```

ntc=2,          (o algoritmo SHAKE é usado para restringir as ligações envolvendo
                hidrogênios)
ntf=2,          (relacionada a ntc, as ligações envolvendo hidrogênio são omitidas)
nscm=1000,      (remove os movimentos translacionais e rotacionais em intervalos regulares)
nstlim=50000,  (número total de passos que será executado na simulação)
dt=0.002,      (intervalo de tempo entre um passo e outro)
tempi=0.0,     (temperatura inicial derivada da energia cinética)
temp0=100.0,   (temperatura média final para a simulação)
ntt=3,         (dinâmica de Langevin usada para controlar a temperatura)
gamma_ln=1.0   (Relacionada a ntt=3, frequência de colisão de 1 ps-1)

```

/

### **NVT-2 (Termalização 100K - 300K)**

```
&cntrl
```

```

  irest=1,      (leitura de coordenadas e velocidades a partir de um arquivo da simulação
                anterior)
  ntx=5,        (estar relacionado com irest=1, coordenadas e velocidades serão lidos, o
                formatado do arquivo ASCII)
  ntb=1, cut=10.0,
  ntp=1000, ntwx=1000, ntwr=10000,
  ntc=2, ntf=2, nscm=1000,
  nstlim=50000, dt=0.002,
  ntt=3, tempi=100.0, temp0=300.0, gamma_ln=1.0,
  ig=22640      (semente para o geração de número pseudo-aleatório)

```

/

### **NPT (Produção)**

```
&cntrl
```

```

  irest=1, ntx=5,
  ntb=2,        (uso de condições periódicas de contorno a pressão constante)
  pres0=1.0,    (valor médio da pressão 1 atm)
  taup=1.0,     (tempo de relaxação da pressão, 1ps)
  cut=10.0,
  ntp=2000, ntwx=2000, ntwr=10000,
  ntc=2, ntf=2, nscm=1000,
  nstlim=5000000, dt=0.002,
  ntt=3, temp0=300.0, gamma_ln=1.0,
  ioutfm=1      (o arquivo da trajetória é escrito no formato NetCDF)

```

/

*QM/MM-MD**Minimização*

```

&cntrl
  imin=1, maxcyc=1000, ncyc=100, ntmin=1,
  cut=10, ntp=100, ntwr=50,
  iwrap=1, noshakemask='@501-521,898-913,1792-1806,2914-3003',
  ifqnt=1, nmropt=1,
/
&qmmm
  qmmask='@501-521,898-913,1792-1806,2914-3003',
  qmcharge=+1,
  qm_theory='PM3-PDDG',
  qmcut=10.0,
  printcharges=1,
  writepdb=1,
/
&wt type='DUMPFREQ', istep1=1,
/
&wt type='END',
/
DISANG=restraint.rst
DUMPAVE=restraint.rst.out
LISTIN=POUT
LISTOUT=POUT

```

*Aquecimento*

```

&cntrl
  imin=0, irest=0, ntx=1,
  ntb=1, ntt=3, gamma_ln=2.0,
  temp0=300.0, tempi=0.0, nmropt=1, cut=10,
  iwrap=1, ifqnt=1,
  ntc=2, ntf=2, ntp=0,
  nstlim=5000, dt=0.001,
  ntp=100, ntwx=100, ntwr=100,
  noshakemask='@501-521,898-913,1792-1806,2914-3003'
/
&qmmm
  qmmask='@501-521,898-913,1792-1806,2914-3003'
  qmcharge=+1,
  printcharges=1,
  qmshake=0,
  qm_theory='PM3-PDDG',
  qmcut=10.0,
  writepdb=1,
/
&wt

```

```

type='DUMPFREQ', istep1=1,
/
&wt
type='TEMP0', istep1=0, istep2=2000,
value1=0.0, value2=300.0,
/
&wt
type='END',
/
DISANG=restraint.rst
DUMPAVE=restraint.rst.out
LISTIN=POUT
LISTOUT=POUT

```

### *Produção*

```

&cntrl
  irest=0, ntx=1, iwrap=1,
  ntb=2, ntp=1, pres0=1.0, taup=1.0, cut=10.0,
  ntp=100, ntwx=100, ntwr=100,
  ntc=2, ntf=2, noshakemask='@501-521,898-913,1792-1806,2914-3003',
  nstlim=50000, dt=0.001,
  ntt=3, temp0=300.0, gamma_ln=2.0,
  ifqnt=1, igb=0, nmropt=1
/
&qmmm
qmmask='@501-521,898-913,1792-1806,2914-3003',
qmcharge=+1,
printcharges=1,
qmshake=0,
qm_theory='PM3-PDDG'
qmcut=10.0
printcharges=1,
writepdb=1,
/
&wt type='DUMPFREQ', istep1=1,
/
&wt type='END',
/
DISANG=restraint.rst
DUMPAVE=restraint.rst.out
LISTIN=POUT
LISTOUT=POUT

```

## APÊNDICE B: Artigos Publicados

Page 1 of 9

Physical Chemistry Chemical Physics

View Article Online  
DOI: 10.1039/C6CP03209E

Journal Name

ARTICLE

# Reaction Mechanism of the Dengue Virus Serine Protease: A QM/MM study

M.C.P. Lima<sup>a</sup> and G.M. Seabra<sup>a†</sup>Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

The dengue virus (DENV) is the causative agent of the viral infection dengue fever. In spite of all the efforts in preventing the spread of the disease, once it is contracted, there is no specific treatment for dengue and the WHO guidelines are limited to rest and symptomatic treatment. In its reproductive cycle, DENV utilizes the NS2B-NS3pro, a serine protease, to cleave the vira polyprotein into its constituents. This enzyme is essential for the virus lifecycle, and presents an attractive target for the development of specific Dengue treatments. Here we used QM/MM-MD and Umbrella Sampling to study the first step (acylation) of the reaction catalyzed by NS2B-NS3pro, using the PDDG/PM3 semi-empirical Hamiltonian for the QM subsystem, and Amber FF99SB for the MM subsystem. Our results indicate that the nucleophilic attack on the substrate by Ser135 occurs in a stepwise manner, where a proton transfer to His51 first activates Ser135, which only later attacks the substrate. The rate-determining step is the Ser135 activation, with a barrier of 24.1 kcal/mol. Water molecules completing the oxyanion hole stabilize the negative charge formed on the carbonyl oxygen of the substrate. The final step in the process is a proton transfer from His51 to the substrate's nitrogen, which happens with a lower barrier of 5.1 kcal/mol, and leads directly to the peptide bond breaking.

## Introduction

The dengue virus (DENV) is the causative agent of the viral infection dengue fever, transmitted to humans by *Aedes aegypti* and *Aedes albopictus* mosquitoes, and that can escalate to hemorrhagic dengue fever and dengue shock syndrome, leading to death.<sup>1</sup> It has been estimated to be around 390 million infections a year, from which only 96 million are apparently manifest, numbers up to three times greater than previously estimated by the World Health Organization.<sup>2</sup> A recent study estimates that, only in 2013, there have been over 58 million symptomatic DENV infections, including over 13,000 fatal cases, and a total annual cost of US\$ 8.9 billion.<sup>3</sup> The first dengue vaccine was developed in 2015 (DengVax, Sanofi-Pasteur), and has been licensed in Brazil, Mexico and the Philippines, despite having only 65.6% efficacy in preventing the disease.<sup>4–6</sup> Other efforts are trying to develop more effective vaccines, such as the in development by a partnership between the Oswaldo Cruz Foundation (FIOCRUZ) and Butantan Institute in Brazil.<sup>7,8</sup> Another line of action in testing

is the vector control through the release of genetically modified specimens to generate offspring that do not reach adulthood,<sup>9</sup> or sterilization of the mosquitoes by irradiation.<sup>10</sup> These procedures are still in the testing phase, and have no emergency efficiency in the rapid global spread of the virus.<sup>2</sup> In spite of all the efforts in preventing the spread of the disease, once it is contracted, there is no specific treatment for dengue and the WHO guidelines are limited to rest and symptomatic treatment.<sup>11</sup>

DENV is a member of the flavivirus family, together with the yellow fever (YFV), hepatitis C (HCV), Zika (ZICV), West Nile (WNV) and Japanese encephalitis (JEV) viruses.<sup>12</sup> The flaviviruses genome carries a single positive strand RNA that can be directly translated into a polypeptide chain (the polyprotein) by the host cell.

The DENV RNA encodes a large polyprotein precursor of 3,391 amino acids consisting of three structural proteins (C, prM, and E) and seven non-structural proteins (NS1, NS2a, NS2B, NS3, NS4A, NS4B and NS5).<sup>13,14</sup> The viral replication process, wherein the precursor polyprotein is cleaved into the virus proteins, is mediated by co- and posttranslational host proteases and two components of the virus itself, the protease domain of NS3 (NS3pro) and NS2B, the latter functioning as a cofactor.<sup>15,16</sup> NS3pro is a serine protease such as trypsin, with a catalytic triad composed of His51, Asp75, and Ser135.<sup>17</sup> The cofactor NS2B is a prerequisite for the catalytic activity, as demonstrated by *in vitro* studies with synthetic peptide substrates and the precursor polyprotein produced from viral

<sup>a</sup> Departamento de Química Fundamental, Centro de Ciências Exatas e da Natureza, Universidade Federal de Pernambuco, Av. Jornalista Aníbal Fernandes, s/nº, Cidade Universitária - Recife, PE - CEP 50.740-560

† Corresponding author. E-mail: gustavo.seabra@ufpe.br  
Electronic Supplementary Information (ESI) available: Sample input files for the simulations are available as supplementary information. See DOI: 10.1039/x0xx00000x



## The Role of the Conformational Dynamics of Glutathione S-Transferase Epsilon Class on Insecticide Resistance in *Anopheles gambiae*

Frederico J. S. Pontes,<sup>a</sup> Rafael T. Maia,<sup>a,b</sup> Maria Carolina P. Lima,<sup>a</sup>  
 Constância F. J. Ayres<sup>b</sup> and Thereza A. Soares<sup>\*a</sup>

<sup>a</sup>Departamento de Química Fundamental, Universidade Federal de Pernambuco,  
 Cidade Universitária, 50740-560 Recife-PE, Brazil

<sup>b</sup>Departamento de Entomologia, Centro de Pesquisas Aggeu Magalhães-FIOCRUZ,  
 Cidade Universitária, 50670-420 Recife-PE, Brazil

Glutathione S-transferases (GSTs) are enzymes capable of metabolizing cytotoxic compounds. The enzyme AgGSTE2, member of epsilon class GSTs (GSTE), is the most important GST conferring resistance to dichloro-diphenyl-trichloroethane (DDT) in *Anopheles gambiae*. We have investigated the conformational dynamics of three GSTE variants (GSTE2, GSTE2-I114T/F120L, GSTE5) from *A. gambiae*. Large-scale motions of helices H2 and H4 and conformational transition of the C-terminal governs the opening of the G-site and is expected to affect substrate binding and product release. This structural rearrangement places Glu116 (Glu120 in GSTE5) close of the thiol group of the tripeptide glutathione (GSH) cofactor, making this residue a candidate to act as a base in the activation of DDT. The structural rearrangement is noticeable for AgGSTE2-F120L, which has been shown to confer increased DDT-resistance. The other variants exhibit a more subtle rearrangement. These findings corroborate the hypothesis that the increase of the conformational dynamics of GST Epsilon class isoforms from *A. gambiae* promotes higher DDTase activity.

**Keywords:** molecular dynamics simulation, evolutionary constraint, positive and negative selection, metabolic resistance, malaria vector

### Introduction

Glutathione transferases (EC 2.5.1.18) are highly promiscuous proteins where broad functional promiscuity co-exists with highly conserved structural fold. Glutathione S-transferases (GSTs) constitute a large family of cytosolic and membrane-bound proteins that catalyze the nucleophilic addition of the tripeptide glutathione (GSH) to a variety of electrophilic toxins and drugs (xenobiotics), leading to their excretion.<sup>1,2</sup> Lately, several other activities have also been associated with GSTs, including steroid and leukotriene biosynthesis, peroxide degradation, double-bond *cis-trans* isomerization, dehydroascorbate reduction, Michael addition, and noncatalytic ligand binding and transport activity.<sup>3</sup> These polymorphic enzymes belong to supergene families whose members are generated by punctual mutations, gene duplication and alternative splicing.<sup>4</sup> The GST family is subdivided into several classes accordingly to its occurrence in different taxa.

The number of isoforms *per* class varies widely, ranging from one to forty.<sup>4</sup> A single GST isoform is capable of conjugating glutathione to several hydrophobic substrates. Such promiscuity when coupled with the large number of isozymes generates a large range of potential substrates.

Insects exhibit at least six classes of GSTs (Sigma, Omega, Theta, Zeta, Delta and Epsilon) with the first four classes present in almost all living organisms.<sup>4,5</sup> The Epsilon and Delta classes are arthropod specific and some members of these classes have been associated to insecticide resistance in culicid vectors.<sup>5,6</sup> The metabolism of toxic compounds in insects is conducted by a series of enzymes from different phases and GSTs act in phase II. GSTs metabolize these compounds in two ways: one has been cited above (through GSH conjugation) and the other mechanism is by sequestration.<sup>7</sup> In conjugation reaction, the active site residue interacts with the GSH sulfhydryl group generating the catalytically active thiolate anion (GS<sup>-</sup>). This nucleophilic thiolate anion attacks the electrophilic center of any lipophilic compound to form the corresponding GS<sup>-</sup> conjugate.<sup>8</sup> The conjugation neutralizes the electrophilic

\*e-mail: thereza.soares@ufpe.br

### APÊNDICE C: Apresentações em Eventos

---

- XVI Simpósio Brasileiro de Química Teórica (2011)- Ouro Preto-MG-Brasil - Replicação do Vírus da Dengue: Um Estudo Por Métodos Computacionais.
- II Simpósio de Planejamento e Desenvolvimento de Novos Fármacos para Doenças Negligenciadas (2011) – São Paulo –SP-Brasil - Molecular Dynamics Study of the Dengue Virus NS2B-NS3 Protease
- Biophysical Society 56th Annual Meeting (2012) – San Diego- CA-Estados Unidos - MOLECULAR DYNAMICS OF THE DENGUE VIRUS NS2B-NS3 PROTEASE.
- AMBER Workshop 2013 – Recife –PE-Brasil - MOLECULAR DYNAMICS OF TYPE-2 DENGUE VIRUS SERINE PROTEASE IN THE PRESENCE OF INHIBITOR AND SUBSTRATE.
- Biophysical Society 58th Annual Meeting (2014) – San Francisco- CA-Estados Unidos - MOLECULAR DYNAMICS OF THE DENGUE VIRUS NS2B-NS3 PROTEASE IN PRESENCE OF INHIBITOR OR SUBSTRATE.  
(Aceito