

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

LANDERSON MATHEUS BARBOSA PEREIRA

DINÂMICA MOLECULAR DIRIGIDA (SMD) NA INVESTIGAÇÃO DA ESTABILIDADE
DE VARIANTES DA LIPASE CALB

RIO DE JANEIRO

2025

Landerson Matheus Barbosa Pereira

DINÂMICA MOLECULAR DIRIGIDA (SMD) NA INVESTIGAÇÃO DA ESTABILIDADE
DE VARIANTES DA LIPASE CALB

Trabalho de Conclusão de Curso
apresentado ao Curso de Química -
Atribuições Tecnológicas, da
Universidade Federal do Rio de
Janeiro, como requisito parcial para a
obtenção do título de Bacharel em
Química - Atribuições Tecnológicas.

Orientador: Rafael Dias Mesquita

Co-orientador: Nilton Silva Júnior

Rio de Janeiro

2025

CIP - Catalogação na Publicação

P436d Pereira, Landerson Matheus Barbosa
DINÂMICA MOLECULAR DIRIGIDA (SMD) NA
INVESTIGAÇÃO DA ESTABILIDADE DE VARIANTES DA LIPASE
CALB / Landerson Matheus Barbosa Pereira. -- Rio de
Janeiro, 2025.
72 f.

Orientador: Rafael Dias Mesquita.
Coorientador: Nilton Silva Júnior.
Trabalho de conclusão de curso (graduação) -
Universidade Federal do Rio de Janeiro, Instituto
de Química, Bacharel em Química, 2025.

1. lipase. 2. CalB. 3. simulação molecular. 4.
SMD. 5. estabilidade estrutural. I. Mesquita,
Rafael Dias, orient. II. Júnior, Nilton Silva,
coorient. III. Título.

AGRADECIMENTOS

Primeiramente, sou muito grato a Deus. Sem Ele, eu nem estaria aqui, e este trabalho não teria passado da capa. Também sou profundamente grato à minha família, por todo o suporte — direto e indireto — que tornou possível a realização deste trabalho.

Sou extremamente grato ao meu orientador e ao meu co-orientador, por todas as instruções, pelo apoio e pela paciência que tiveram comigo. Meus dois heróis não usam capa: orientam trabalhos acadêmicos com atenção, cuidado e dedicação. E muito grato à minha gerente no trabalho, que permitiu que eu pudesse estar no laboratório para a escrita deste trabalho.

Se eu citar nomes, seria injusto, pois corro o risco de esquecer alguém. Mas sou muitíssimo grato aos meus amigos e colegas da faculdade, pelo apoio — seja por meio de ajuda direta, dicas ou uma simples palavra de encorajamento. Eu não teria chegado até aqui sem vocês.

Tenho também uma gratidão especial pelos meus colegas de laboratório, que me ajudaram inúmeras vezes a superar os diversos obstáculos encontrados durante a realização deste trabalho. São pessoas gentis e queridas.

Sou igualmente grato aos meus professores, que me transmitiram habilidades e competências fundamentais para a concretização deste projeto. Também agradeço ao Wolfram Alpha, Symbolab e GeoGebra, pelo auxílio nos cálculos, e ao ChatGPT e Gemini, pelo suporte nas mais diversas disciplinas ao longo da graduação.

Não poderia deixar de mencionar minha gratidão aos motoristas de ônibus — sem eles, eu literalmente não teria chegado até aqui — e aos muitos funcionários, colaboradores e prestadores de serviço da UFRJ: técnicos de laboratório, equipe da limpeza, segurança, biblioteca, secretaria, bandejão, quiosques, xerox, entre outros. Graças a eles, tive a infraestrutura necessária para trilhar este caminho.

Sou grato a todas essas e a muitas outras pessoas, pois, se cheguei até aqui, foi com muita ajuda e apoio.

LISTA DE ILUSTRAÇÕES

Figura 1 - Reações genéricas catalisadas por lipases.....	11
Figura 2 - Lipase B de Candida antarctica (CalB). Imagem gerada utilizando o software VMD a partir do PDB do cristal 5A71.....	12
Figura 3 - Esquema da mutante R249L.....	13
Figura 4 - Esquema da mutante A281E.....	14
Figura 5 - Energia vs tempo ao longo da minimização e equilibração.....	24
Figura 6 - Evolução temporal da SMD de CalB selvagem (átomo fixo em azul e átomo tracionado em vermelho). (A) 12,5 ns. (B) 25 ns. (C) 37,5 ns. (D) 50 ns.....	26
Figura 7 - Evolução temporal da SMD da mutante A281E (átomo fixo em azul e átomo tracionado em vermelho). (A) 12,5 ns. (B) 25 ns. (C) 37,5 ns. (D) 50 ns.....	27
Figura 8 - Evolução temporal da SMD da mutante R249L (átomo fixo em azul e átomo tracionado em vermelho). (A) 12,5 ns. (B) 25 ns. (C) 37,5 ns. (D) 50 ns.....	28
Figura 9 - Gráficos suavizados de força vs tempo. (A) CalB_WT. (B) A281E. (C) R249L. (D) Médias das réplicas.....	30
Figura 10 - Gráfico de estrutura secundária vs tempo da réplica CalB_WT_r1.....	31
Figura 11 - Gráfico de estrutura secundária vs tempo da réplica CalB_WT_r2.....	32
Figura 12 - Gráfico de estrutura secundária vs tempo da réplica CalB_WT_r3.....	33
Figura 13 - Gráfico de estrutura secundária vs tempo da réplica A281E_r1.....	34
Figura 14 - Gráfico de estrutura secundária vs tempo da réplica A281E_r2.....	35
Figura 15 - Gráfico de estrutura secundária vs tempo da réplica A281E_r3.....	36
Figura 16 - Gráfico de estrutura secundária vs tempo da réplica R249L_r1.....	37
Figura 17 - Gráfico de estrutura secundária vs tempo da réplica R249L_r2.....	38
Figura 18 - Gráfico de estrutura secundária vs tempo da réplica R249L_r3.....	39
Figura 19 - Gráficos suavizados do número de ligações de hidrogênio vs tempo. (A) CalB_WT. (B) A281E. (C) R249L. (D) Médias das réplicas.....	41
Figura 20 - Estruturas dos aminoácidos envolvidos nas mutações.....	42

LISTA DE TABELAS

Tabela 1 - Cálculo do trabalho realizado no desenovelamento mecânico das proteínas.....	43
--	-----------

LISTA DE ABREVIATURAS E SIGLAS

CalB – Lipase B de *Candida antarctica*

MD – Dinâmica Molecular (*Molecular Dynamics*)

SMD – Dinâmica Molecular Dirigida (*Steered Molecular Dynamics*)

GBIS – *Generalized Born Implicit Solvent*

PDB – *Protein Data Bank*

VMD – *Visual Molecular Dynamics*

RESUMO

As lipases são enzimas versáteis que catalisam reações de hidrólise, esterificação e transesterificação, sendo amplamente aplicadas na indústria. A lipase B de *Candida antarctica* (CalB) destaca-se por sua estabilidade e seletividade, características que motivam o estudo de variantes enzimáticas com potencial desempenho aprimorado. Este trabalho investigou o comportamento mecânico de desenovelamento da CalB selvagem e de duas variantes (R249L e A281E), por meio de simulações de Dinâmica Molecular Dirigida (SMD), utilizando o software NAMD 2.14 com o campo de força CHARMM36m e solvente implícito GBIS. O software VMD foi empregado para visualização estrutural e análise quantitativa. Os resultados mostraram que a variante R249L apresentou maior resistência mecânica, enquanto a A281E exibiu maior variabilidade conformacional durante o desenovelamento, o que está de acordo com dados experimentais da literatura. Além disso, foi observada a presença de um possível estado intermediário ao longo do processo, reforçando a capacidade da SMD em capturar eventos conformacionais relevantes.

Palavras-chave: lipase; CalB; simulação molecular; SMD; estabilidade estrutural.

ABSTRACT

Lipases are versatile enzymes that catalyze hydrolysis, esterification, and transesterification reactions, being widely used in industry. *Candida antarctica* lipase B (CalB) stands out for its stability and selectivity, characteristics that drive the study of enzyme variants with potentially improved performance. This work investigated the mechanical unfolding behavior of wild-type CalB and two variants (R249L and A281E) through Steered Molecular Dynamics (SMD) simulations, using NAMD 2.14 software with the CHARMM36m force field and GBIS implicit solvent. The VMD software was employed for structural visualization and quantitative analysis. The results showed that the R249L variant exhibited greater mechanical resistance, while A281E displayed higher conformational variability during unfolding, which is consistent with experimental data from the literature. Furthermore, the presence of a possible intermediate state was observed during the process, highlighting the ability of SMD to capture relevant conformational events.

Keywords: lipase; CalB; molecular simulation; SMD; structural stability.

SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 LIPASE B DE CANDIDA ANTARCTICA (CALB).....	11
1.2 VARIANTES DA CALB.....	13
1.3 DINÂMICA MOLECULAR DIRIGIDA (SMD).....	14
2 OBJETIVOS.....	16
2.1 OBJETIVO GERAL.....	16
2.2 OBJETIVOS ESPECÍFICOS.....	16
3 METODOLOGIA.....	17
3.1 SOFTWARES E MODELOS UTILIZADOS.....	17
3.2 PREPARAÇÃO.....	18
3.3 MINIMIZAÇÃO E EQUILIBRAÇÃO.....	19
3.4 FIXAÇÃO DOS ÁTOMOS.....	19
3.5 SIMULAÇÕES POR DINÂMICA MOLECULAR DIRIGIDA (SMD).....	20
3.6 ANÁLISE DE DADOS.....	20
3.6.1 Minimização e Equilíbrio.....	20
3.6.2 Imagens e Vídeos das Simulações.....	20
3.6.3 Confecção dos gráficos de força vs tempo.....	21
3.6.4 Confecção dos gráficos de estrutura secundária vs tempo.....	21
3.6.5 Confecção dos gráficos de número de ligações de hidrogênio vs tempo.....	21
3.6.7 Cálculo da área de superfície acessível à água.....	22
3.6.8 Identificação de resíduos vizinhos.....	22
3.6.6 Confecção dos gráficos de força vs deslocamento e cálculo de trabalho.....	22
4 RESULTADOS E DISCUSSÃO.....	23
4.1 PREPARAÇÃO.....	23
4.2 MINIMIZAÇÃO E EQUILIBRAÇÃO.....	23
4.3 SIMULAÇÕES POR DINÂMICA MOLECULAR DIRIGIDA (SMD).....	25
5 CONCLUSÃO.....	46
6 REFERÊNCIAS.....	47
ANEXO A - Script “CalB_GBiS.pgn”.....	49
ANEXO B - Script “minim_equil.conf”.....	50
ANEXO C - Script “Fix_SMD.tcl”.....	53
ANEXO D - Script “CalB_SMD_NAMD.conf”.....	54
ANEXO E - Script “namdstats.tcl”.....	58
ANEXO F - Script “smd_final.tcl”.....	63
ANEXO G - Script “ft.tcl”.....	65
ANEXO H - Código para cálculo da área de superfície acessível à água.....	66
ANEXO I - Códigos para a Identificação de resíduos vizinhos.....	67
ANEXO J - Script “ft_deslocamento.tcl”.....	68
ANEXO K - Script “trabalho.py”.....	69

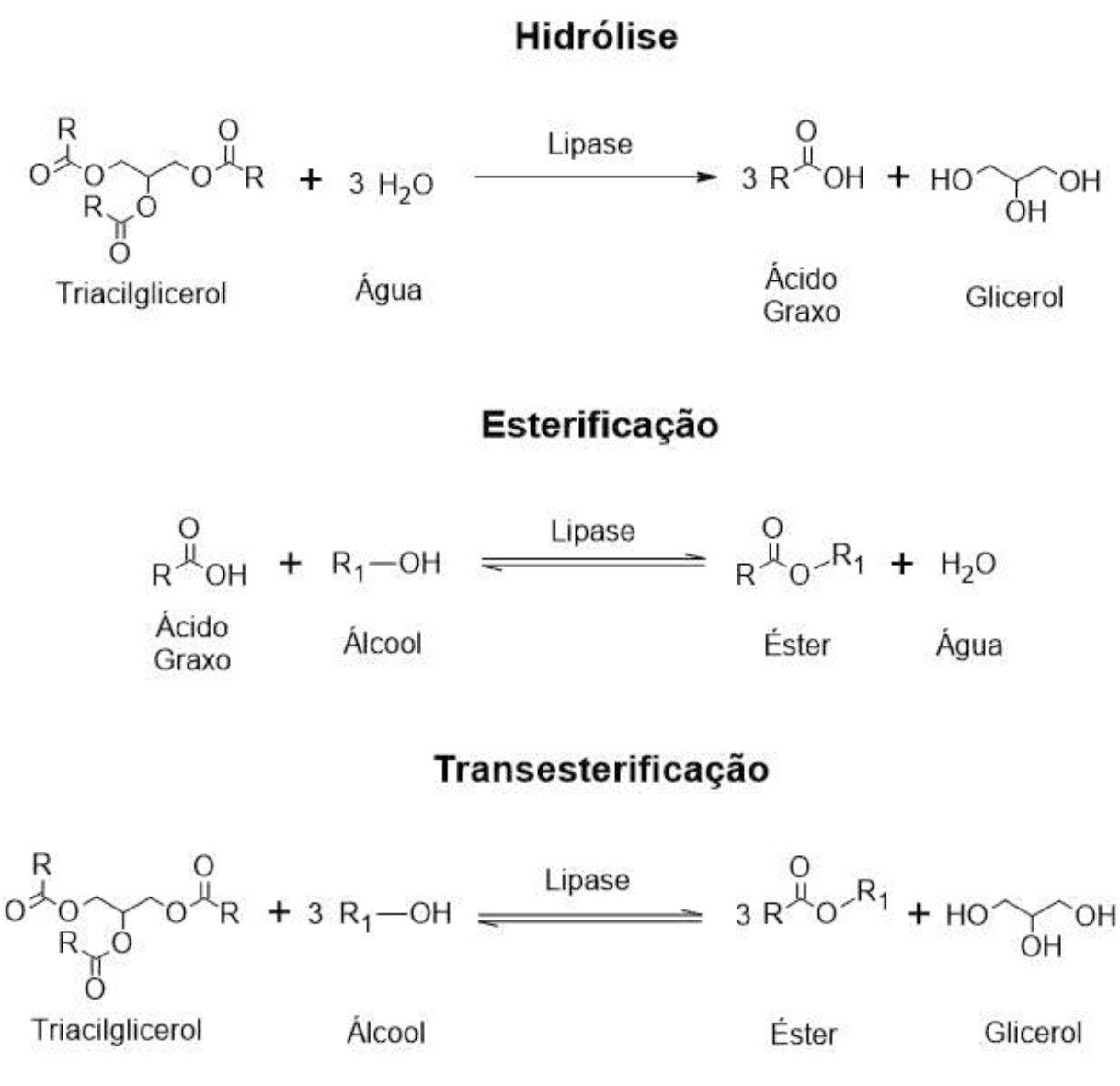
ANEXO L - Gráficos com os dados brutos de força vs tempo. (A) CalB_WT. (B) A281E. (C) R249L.....	71
ANEXO M - Gráficos com os dados brutos do número de ligações de hidrogênio vs tempo (ns). (A) CalB_WT. (B) A281E. (C) R249L.....	72
ANEXO N - Gráficos com os dados brutos de força vs deslocamento. (A) CalB_WT_r1. (B) CalB_WT_r2. (C) CalB_WT_r3. (D) A281E_r1. (E) A281E_r2. (F) A281E_r3. (G) R249L_r1. (H) R249L_r2. (I) R249L_r3.....	73

1 INTRODUÇÃO

1.1 LIPASE B DE CANDIDA ANTARCTICA (CALB)

As lipases são enzimas versáteis que catalisam reações de hidrólise, esterificação e transesterificação, sendo amplamente aplicadas na indústria (BRITO E CUNHA et al., 2019; KUNDYS et al., 2017). Reações genéricas de hidrólise, esterificação e transesterificação podem ser vistas na Figura 1.

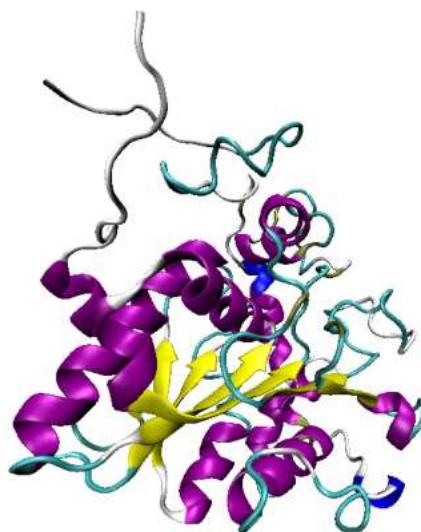
Figura 1 - Reações genéricas catalisadas por lipases.



Fonte: Autor

A lipase B de *Candida antarctica* (CalB) é uma das lipases mais utilizadas comercialmente devido à sua estabilidade, seletividade e ampla faixa de atuação sobre diferentes substratos (RABBANI et al., 2015; UPPENBERG et al., 2002). A estrutura da CalB pode ser vista na Figura 2.

Figura 2 - Lipase B de *Candida antarctica* (CalB). Imagem gerada utilizando o software VMD a partir do PDB do cristal 5A71.



Fonte: Autor

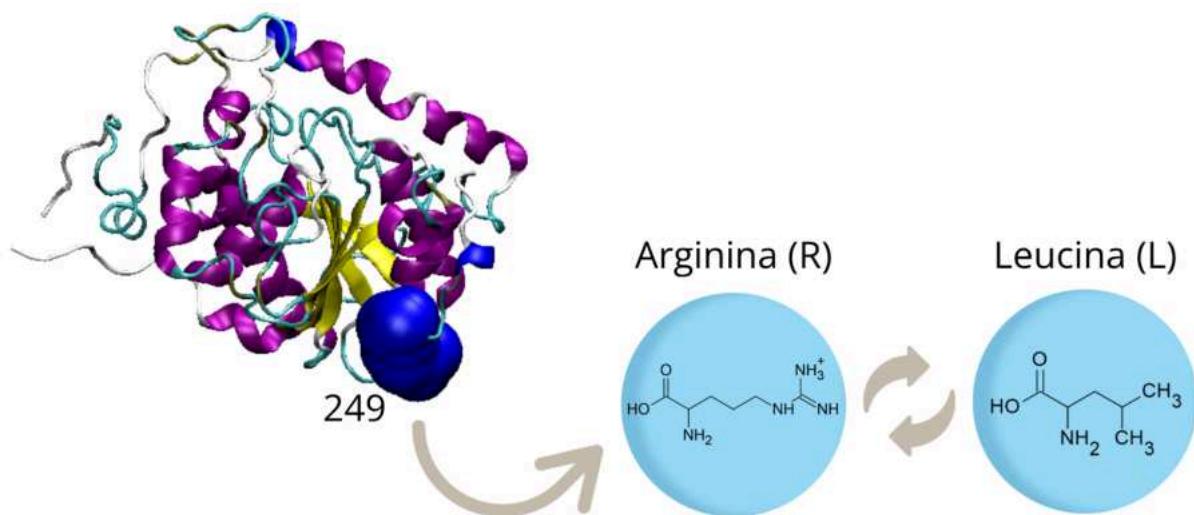
Sua estrutura contém uma tríade catalítica típica das hidrolases, composta por serina 105 (Ser105), ácido aspártico 187 (Asp187) e histidina 224 (His224), que atuam no mecanismo de clivagem do substrato (POOJARI & CLARSON, 2013). Essa tríade está localizada em uma fenda próxima a uma região móvel conhecida como tampa, que regula o acesso ao sítio ativo. A CalB apresenta uma tampa curta, que não bloqueia completamente o sítio ativo, contribuindo para seu comportamento catalítico singular. Outro ponto importante é o fenômeno da ativação interfacial, no qual a enzima muda sua conformação ao interagir com superfícies hidrofóbicas, facilitando o acesso ao sítio ativo (POOJARI & CLARSON, 2013).

Os substratos preferenciais da CalB são ésteres de ácidos graxos e álcoois de cadeias pequenas a médias, o que a torna aplicável na síntese de biodiesel, produção de polímeros biodegradáveis, cosméticos e alimentos (BRITO E CUNHA et al., 2019; KUNDYS et al., 2017).

1.2 VARIANTES DA CALB

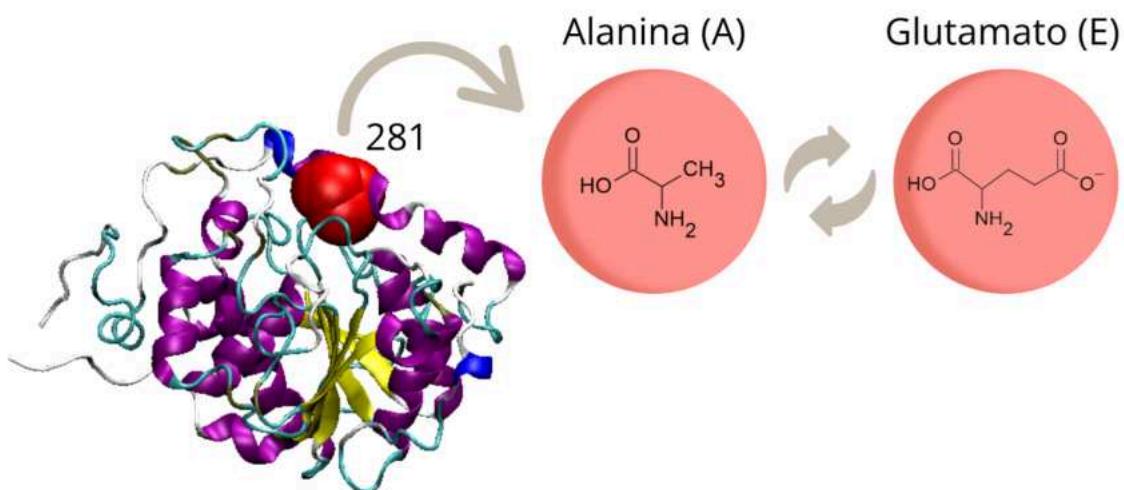
Variantes da CalB vêm sendo desenvolvidas com o objetivo de aprimorar propriedades como termoestabilidade, resistência a solventes e seletividade enzimática (XU et al., 2019). Mutações pontuais, como R249L e A281E, já foram estudadas com foco em sua estabilidade térmica, sendo que a primeira demonstrou maior termoestabilidade, enquanto a segunda se mostrou mais sensível ao calor (KIM et al., 2010). Esquemas das mutantes R249L e A281 podem ser vistos nas Figuras 3 e 4, respectivamente.

Figura 3 - Esquema da mutante R249L



Fonte: Autor

Figura 4 - Esquema da mutante A281E



Fonte: Autor

A ausência de glicosilação na CalB impacta de maneira relevante sua estabilidade térmica. Variantes deglicosiladas da enzima demonstraram redução na estabilidade quando submetidas a temperaturas elevadas, perdendo parte de sua atividade residual em comparação com a enzima glicosilada do tipo selvagem. (HÖCK et al., 2018).

Além disso, já foram descritas sequências homólogas à CalB com potencial biotecnológico, oriundas de outros microrganismos, o que reforça o interesse científico e industrial por essa classe de enzimas (FARIA et al., 2024).

1.3 DINÂMICA MOLECULAR DIRIGIDA (SMD)

A dinâmica molecular (MD) é uma técnica computacional amplamente utilizada para simular o comportamento de sistemas moleculares ao longo do tempo.

Por meio da resolução de equações de mecânica clássica, a MD permite prever os movimentos dos átomos e moléculas em resposta às forças internas e externas, fornecendo informações detalhadas sobre estabilidade, flexibilidade e interações moleculares. Essa abordagem é especialmente útil no estudo de proteínas, pois permite investigar, em escala atômica, como ocorrem mudanças conformacionais e como essas alterações impactam na função biológica ou estabilidade estrutural da macromolécula (PHILLIPS et al., 2020).

A Dinâmica Molecular Dirigida (Steered Molecular Dynamics – SMD) é uma variação da MD clássica que incorpora a aplicação de uma força externa controlada a átomos ou regiões específicas da molécula. Essa força é aplicada ao longo do tempo, simulando processos como o desenovelamento de uma proteína (PATEL et al., 2014).

No contexto deste trabalho, a SMD foi utilizada para avaliar a resistência mecânica da CalB e de suas variantes, fornecendo informações sobre sua estabilidade estrutural frente à aplicação de força. Além disso, a SMD possibilita identificar estados conformacionais intermediários, como os *molten globules*, que representam estruturas parcialmente dobradas e que podem surgir ao longo do processo de desenovelamento (LU & SCHULTEN, 1999).

2 OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar o comportamento mecânico de desenovelamento da lipase B de *Candida antarctica* (CalB) e de duas de suas variantes, por meio de Dinâmica Molecular Dirigida (SMD).

2.2 OBJETIVOS ESPECÍFICOS

- Realizar simulações Dinâmica Molecular Dirigida (SMD) da CalB e de duas de suas variantes utilizando o *software* NAMD 2.14.
- Analisar características químicas e estruturais das proteínas ao longo das simulações de SMD.
- Verificar se há correlação entre a resistência mecânica simulada e a estabilidade experimental das variantes.

3 METODOLOGIA

Para a condução deste estudo, a metodologia foi estruturada em diferentes etapas, desde o preparo dos arquivos com as coordenadas atômicas das proteínas, passando pelas simulações por Dinâmica Molecular Dirigida (SMD), até a análise dos dados. Foram examinadas três formas da enzima CalB: a forma selvagem e duas variantes geradas computacionalmente, A281E (termoinstável) e R249L (termoestável).

Por motivos práticos, o presente estudo foi conduzido com a forma deglicosilada da CalB e das variantes.

3.1 SOFTWARES E MODELOS UTILIZADOS

As simulações foram realizadas com o *software* NAMD 2.14. A metodologia das simulações segue, de forma adaptada, os princípios apresentados no tutorial de SMD do NAMD, desenvolvido pelo grupo Theoretical and Computational Biophysics Group da Universidade de Illinois (PHILLIPS et al., 2017).

Além do NAMD, foi utilizado o *software* VMD (*Visual Molecular Dynamics*), versão 1.9.4a53, empregado tanto para o preparo quanto para a análise visual e quantitativa das simulações moleculares, sendo uma ferramenta amplamente integrada ao NAMD para sistemas biológicos (HUMPHREY et al., 1996).

Adicionalmente, empregou-se o modelo de solvente implícito (*Generalized Born Implicit Solvent* - GBIS), que permite simular os efeitos do meio aquoso sem representar explicitamente cada molécula de água. Esse modelo reduz o custo computacional, sendo adequado para estudos de estabilidade mecânica como os realizados neste trabalho (PHILLIPS et al., 2020).

Utilizou-se o campo de força CHARMM36m, uma versão refinada do tradicional CHARMM36 que oferece maior precisão na representação da estrutura de proteínas enoveladas e desordenadas (HUANG et al., 2016). Desenvolvido com base em dados experimentais, o CHARMM36m ajusta parâmetros estruturais para

refletir com maior fidelidade os comportamentos conformacionais das proteínas, sendo especialmente indicado para simulações de estabilidade estrutural e Dinâmica Molecular.

Os *scripts* foram executados a partir de arquivos de comando (*command*), que indicavam o programa a ser executado (VMD ou NAMD), o número de CPUs utilizadas e o *script* de entrada correspondente. Essa estrutura foi empregada em todas as etapas descritas a seguir, exceto para o cálculo do trabalho. Todas as simulações foram conduzidas utilizando o *software* NAMD 2.14.

As etapas descritas a seguir, com exceção da preparação, foram realizadas três vezes para cada uma das proteínas (CalB selvagem, mutante R249L e mutante A281E). As réplicas foram nomeadas com os sufixos “_r1”, “_r2” e “_r3”, respectivamente.

3.2 PREPARAÇÃO

Os arquivos em formato PDB das mutantes A281E e R249L foram gerados utilizando o *software* UCSF Chimera, com a biblioteca padrão de mutações Dunbrack, sendo nomeados, respectivamente, como "CalB_TERMO_A281E.pdb" e "CalB_TERMO_R249L.pdb". Já o arquivo em formato PDB da CalB selvagem foi obtido do banco de dados *Protein Data Bank* (<http://www.rcsb.org/pdb/>), sob o código 5A71, e foi nomeado como "CalB_WT_base.pdb".

Na etapa de preparação do sistema, foi executado o *script* "CalB_GBIS.pgn", conforme indicado no Anexo A. Os arquivos de entrada foram o respectivo arquivo PDB da proteína e os arquivos de topologia do campo de força CHARMM36m. Os arquivos de saída foram um arquivo com as topologias (.psf) e outro com as coordenadas dos átomos do sistema (.pdb).

3.3 MINIMIZAÇÃO E EQUILIBRAÇÃO

Na etapa de minimização e equilibração, foi executado o *script* "minim_equil.conf", presente no Anexo B. O tempo de simulação da minimização foi de 100 ps, e o tempo total da simulação, ou seja, da minimização e equilibração juntas, foi de 1000 ps (1 ns). Os arquivos de entrada foram os gerados na etapa anterior, juntamente com os arquivos de parâmetros do campo de força CHARMM36m. Os arquivos de saída foram os de coordenadas (.coor), velocidades (.vel), sistema estendido (.xsc) e o registro da simulação (.log).

A simulação seguinte foi conduzida a uma temperatura de 298,15 K, com o uso de solvente implícito GBIS, aplicando no modelo uma constante dielétrica de 78,5 (água) e concentração salina de 0,15 mol L⁻¹. A minimização foi realizada por 100.000 passos utilizando o algoritmo de gradiente conjugado, seguida por uma equilibração de 1.000.000 passos (equivalente a 1 ns), com passo de integração de 1 fs. As interações eletrostáticas e de Van der Waals foram tratadas com *cutoff* de 18,0 Å, *switchdist* de 17,0 Å e *pairlistdist* de 20,0 Å. As forças foram calculadas com exclusão *scaled* 1-4, e a dinâmica foi controlada por termostato de Langevin com coeficiente de amortecimento de 1 ps⁻¹. A temperatura alvo foi mantida constante durante toda a simulação e as velocidades foram reinicializadas antes do início da equilibração.

3.4 FIXAÇÃO DOS ÁTOMOS

Na etapa de fixação dos átomos para a SMD, executou-se o *script* "Fix_SMD.tcl", descrito no Anexo C. Os arquivos de entrada foram o arquivo de topologia (.psf) gerado na primeira etapa e o arquivo de coordenadas (.coor) da etapa anterior. O arquivo de saída foi um arquivo de referência (.ref), que forneceu ao NAMD a informação necessária para identificar o átomo fixo e o átomo a ser tracionado, que foram os carbonos alfa (Ca) dos resíduos localizados nas extremidades N-terminal (Leu1) e C-terminal (Pro317) da CalB, respectivamente.

3.5 SIMULAÇÕES POR DINÂMICA MOLECULAR DIRIGIDA (SMD)

Na etapa da Dinâmica Molecular Dirigida (SMD), foi executado o *script* descrito no Anexo D. O tempo de simulação da SMD foi de 100 ns. Os arquivos de entrada foram os gerados nas etapas anteriores, e os arquivos de saída incluíram arquivos de trajetória (.dcd e .xst), coordenadas (.coor), velocidades (.vel), sistema estendido (.xsc), registro da simulação (.log) e arquivos de reinício (*restart*).

3.6 ANÁLISE DE DADOS

Para a análise de dados, foram construídos gráficos, vídeos e imagens com os resultados obtidos ao longo das etapas anteriormente citadas.

3.6.1 Minimização e Equilibração

A análise de dados incluiu a obtenção dos dados de minimização e equilibração. Para isso foi utilizado o *script* "namdstats.tcl", presente no Anexo E. A entrada foi o arquivo de registro da simulação (.log) gerado pelo NAMD durante a minimização e equilibração, e a saída foi um arquivo de dados (.dat) contendo as informações de energia por *frame*.

3.6.2 Imagens e Vídeos das Simulações

As imagens das proteínas ao longo da simulação foram geradas com o VMD 1.9.4a53, o *script* utilizado para a geração de imagens pode ser visto no Anexo F.

Para a geração dos vídeos das simulações, após ter sido utilizado o *script* presente no Anexo F, acessou-se o menu *Display* e foi ativado o modo *Orthographic*, que garante uma projeção sem distorções de perspectiva. Acessou-se o menu *Extensions > Visualization > Movie Maker*. Na seção *Movie Settings*, foi selecionada a opção *Trajectory* como tipo de visualização. Em seguida, na aba *Format*, escolheu-se o formato de saída de imagem animada (.gif). No campo *Trajectory step size*, foi inserido o valor 7, a fim de reduzir o número de *frames* e otimizar a fluidez do vídeo. Por fim, clicou-se em *Make Movie*, o que resultou na geração do arquivo.

3.6.3 Confecção dos gráficos de força vs tempo

Para a obtenção dos dados do gráfico de força em função do tempo, utilizou-se o *script* "ft.tcl", presente no Anexo G. Sua entrada foi o arquivo de registro da simulação (.log) gerado na etapa da SMD, e a saída foi um arquivo de dados (.dat) com os valores de força (em pN) versus tempo (em ns).

Os gráficos foram construídos com o auxílio do software *GraphPad Prism* 8, a suavização dos gráficos foi feita utilizando o método *Smooth*, com ordem zero e 10 vizinhos.

3.6.4 Confecção dos gráficos de estrutura secundária vs tempo

Para a geração dos gráficos de estrutura secundária em função do tempo utilizou-se o VMD, tendo sido carregados o arquivo de topologia (.psf) e o arquivo de trajetória (.dcd) da respectiva réplica, acessou-se o menu *Extensions > Analysis > Timeline*. Na janela do *Timeline*, foi selecionada a aba *Analysis*, na qual o parâmetro *Frame Range* foi ajustado para um valor máximo de 5000 *frames*. Em seguida, na aba *Calculate*, clicou-se em *Calc. Sec. Struct*, o que permitiu o cálculo de estrutura secundária a cada *frame* da simulação e a obtenção do gráfico.

3.6.5 Confecção dos gráficos de número de ligações de hidrogênio vs tempo

Para o cálculo do número de ligações de hidrogênio utilizou-se o VMD, tendo sido carregados o arquivo de topologia (.psf) e o arquivo de trajetória (.dcd) da respectiva réplica, foi utilizada a ferramenta de análise *Hydrogen Bonds* para calcular o número de ligações de hidrogênio ao longo da simulação. Os parâmetros utilizados foram os padrão do programa (*Update selections every frame* marcado; *Donor-Acceptor distance (A)*: 3,0; *Angle cutoff (degrees)*: 20). Os dados obtidos foram salvos (formato .agr).

Os gráficos foram construídos com o auxílio do software *GraphPad Prism* 8, a suavização dos gráficos foi feita utilizando o método *LOWESS*, com 10 pontos.

3.6.7 Cálculo da área de superfície acessível à água

Para o cálculo da área de superfície acessível à água utilizou-se o VMD, tendo sido carregados o arquivo de topologia (.psf) e o arquivo de trajetória (.dcd) da proteína, o cálculo foi feito na extensão *Tk Console*, que fica no menu *Extensions* do VMD, mais detalhes estão disponíveis no Anexo H.

3.6.8 Identificação de resíduos vizinhos

Para a identificação dos resíduos vizinhos, utilizou-se o VMD, tendo sido carregados o arquivo de topologia (.psf) e o arquivo de trajetória (.dcd) da proteína, a identificação foi feita com o auxílio da extensão *Tk Console*, que fica no menu *Extensions* do VMD, mais detalhes estão disponíveis no Anexo I.

3.6.6 Confecção dos gráficos de força vs deslocamento e cálculo de trabalho

Para a obtenção dos dados do gráfico de força em função do deslocamento da CalB e mutantes, utilizou-se o *script* "ft_deslocamento.tcl", descrito no Anexo J. Sua entrada foi o arquivo de registro da simulação (.log) da SMD, e a saída foi um arquivo de dados (.dat) contendo os valores de força (em pN) versus deslocamento (em Å). Os gráficos foram construídos com o auxílio do *software GraphPad Prism 8*.

Por fim, o cálculo do trabalho realizado durante a SMD foi feito com a execução do *script* "trabalho.py" em Python 3, presente no Anexo K. Esse *script* permitiu o cálculo da área sob a curva do gráfico de força em função do deslocamento, de 0 a 750 Å, por meio do método trapezoidal. Sua entrada foi o arquivo de dados (.dat) gerado na etapa anterior, e a saída foi o valor de trabalho em kcal/mol.

4 RESULTADOS E DISCUSSÃO

Nesta seção, são apresentados e discutidos os principais resultados obtidos ao longo do desenvolvimento do trabalho, com base nas simulações de Dinâmica Molecular Dirigida (SMD) aplicadas à CalB selvagem e às variantes R249L e A281E.

4.1 PREPARAÇÃO

O arquivo em formato PDB da CalB selvagem foi obtido do banco de dados RCSB PDB. Tipicamente, esse tipo de estrutura é determinado por cristalografia de raios X, uma técnica que, devido à baixa densidade eletrônica dos átomos de hidrogênio, geralmente não permite sua detecção direta (PHILLIPS et al., 2017). Além disso, como a proteína é cristalizada em condições não fisiológicas, a estrutura obtida pode apresentar átomos e resíduos ausentes, estar em um estado de alta energia conformacional e não incluir o ambiente aquoso em que a proteína normalmente se encontra. Por esse motivo, a etapa de preparação foi fundamental: nela, resíduos selecionados foram protonados a pH 7,4, pontes dissulfeto específicas foram formadas entre pares de cisteínas, os ângulos e diedros ausentes foram reconstruídos, e coordenadas padrão foram atribuídas aos átomos faltantes. Adicionalmente, a proteína foi, simulada em uma concentração salina fisiológica.

A etapa de preparação gerou um arquivo que contém a posição (coordenadas) dos átomos (.pdb), não contendo informações de conectividade. Ela também gerou um arquivo de topologia (.psf), que é um arquivo que guarda a informação estrutural da proteína (como os vários tipos de ligações e interações) (PHILLIPS et al., 2020).

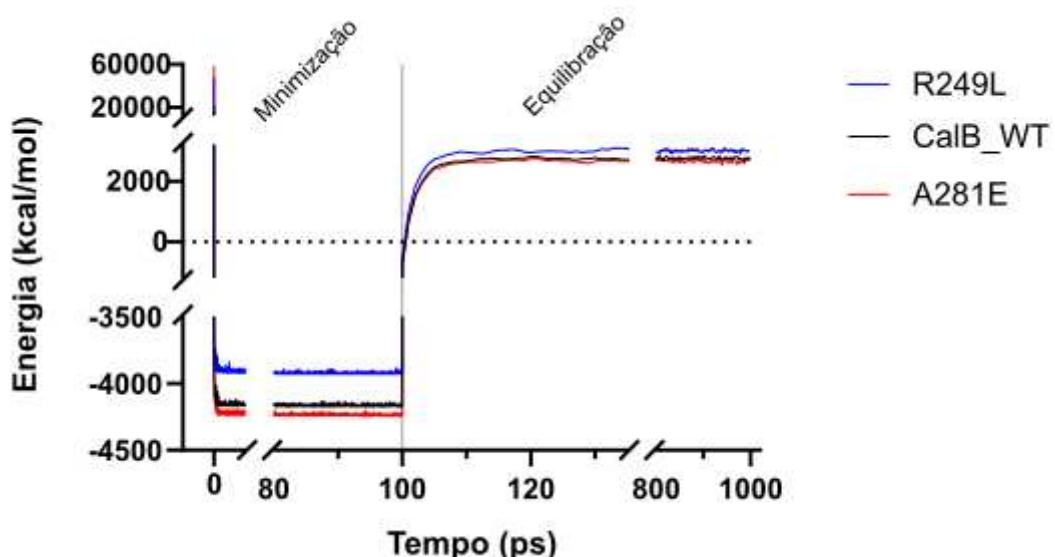
4.2 MINIMIZAÇÃO E EQUILIBRAÇÃO

A minimização de energia tem como objetivo encontrar um mínimo local de energia, correspondente a um estado em que a molécula se encontra relaxada. Para isso, as posições dos átomos são ajustadas gradualmente enquanto a energia do

sistema é recalculada até que se atinja esse ponto mínimo. A equilbração, por outro lado, envolve a aplicação da Dinâmica Molecular, na qual a Segunda Lei de Newton é resolvida para cada átomo do sistema, permitindo determinar suas trajetórias. O estado de equilíbrio é avaliado com base na análise da distribuição das velocidades, temperatura, pressão e outros parâmetros ao longo do tempo. A minimização e a equilbração são importantes, pois permitem que a proteína alcance uma conformação mais estável para que seja feita a dinâmica (PHILLIPS et al., 2017).

O gráfico de energia vs tempo obtido na equilbração e minimização pode ser visto na Figura 5.

Figura 5 - Energia vs tempo ao longo da minimização e equilbração.



Fonte: Autor

Como pode ser observado na Figura 5, na etapa de minimização foi obtido um mínimo local de energia para as três proteínas, e na etapa de equilbração, elas alcançaram um patamar de energia praticamente constante, portanto, a minimização e a equilbração foram alcançadas com êxito.

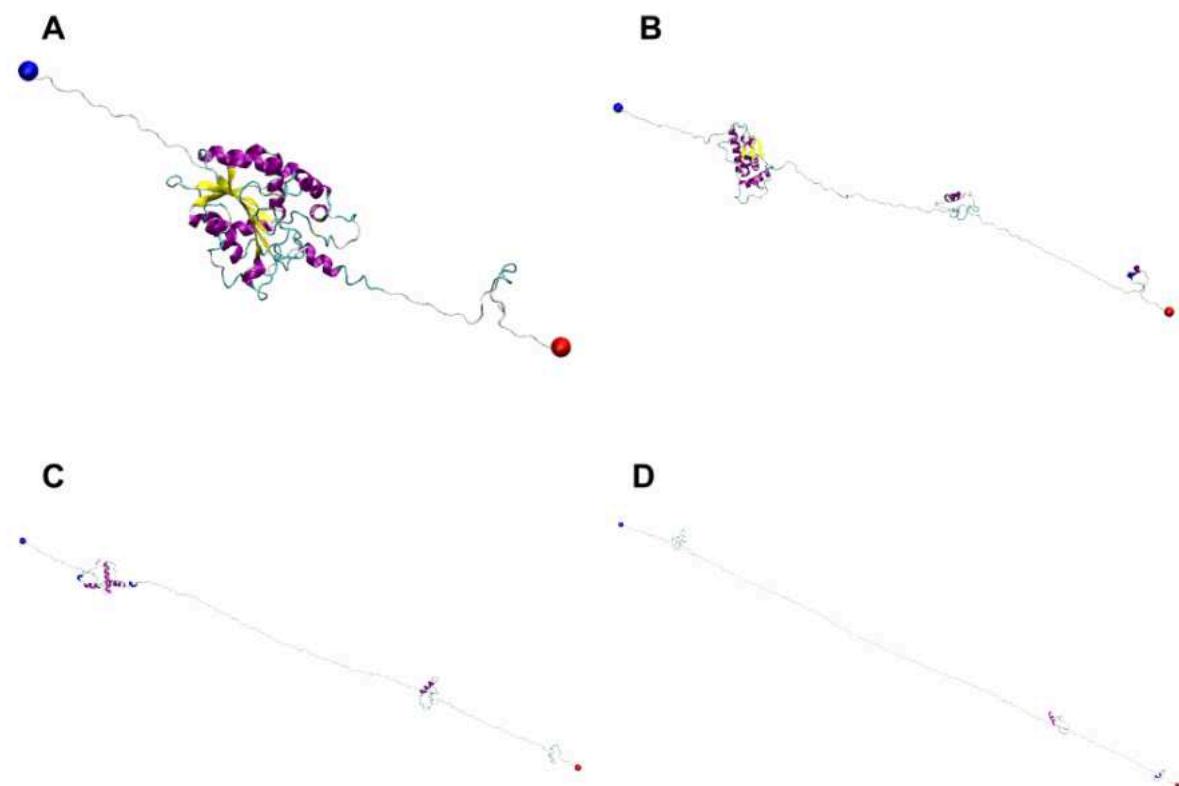
As mutantes partiram de um ponto de maior energia (~ 60.000 kcal/mol) que a selvagem (~20.000 kcal/mol). Isso se deve ao fato de seus arquivos PDB terem sido gerados computacionalmente e não serem da mesma origem que o PDB da selvagem, ou seja oriundos de uma cristalização. Por outro lado pode ser observado

que todas as três proteínas acabam a equilíbrio com valores muito próximos (~2.500 kcal/mol).

4.3 SIMULAÇÕES POR DINÂMICA MOLECULAR DIRIGIDA (SMD)

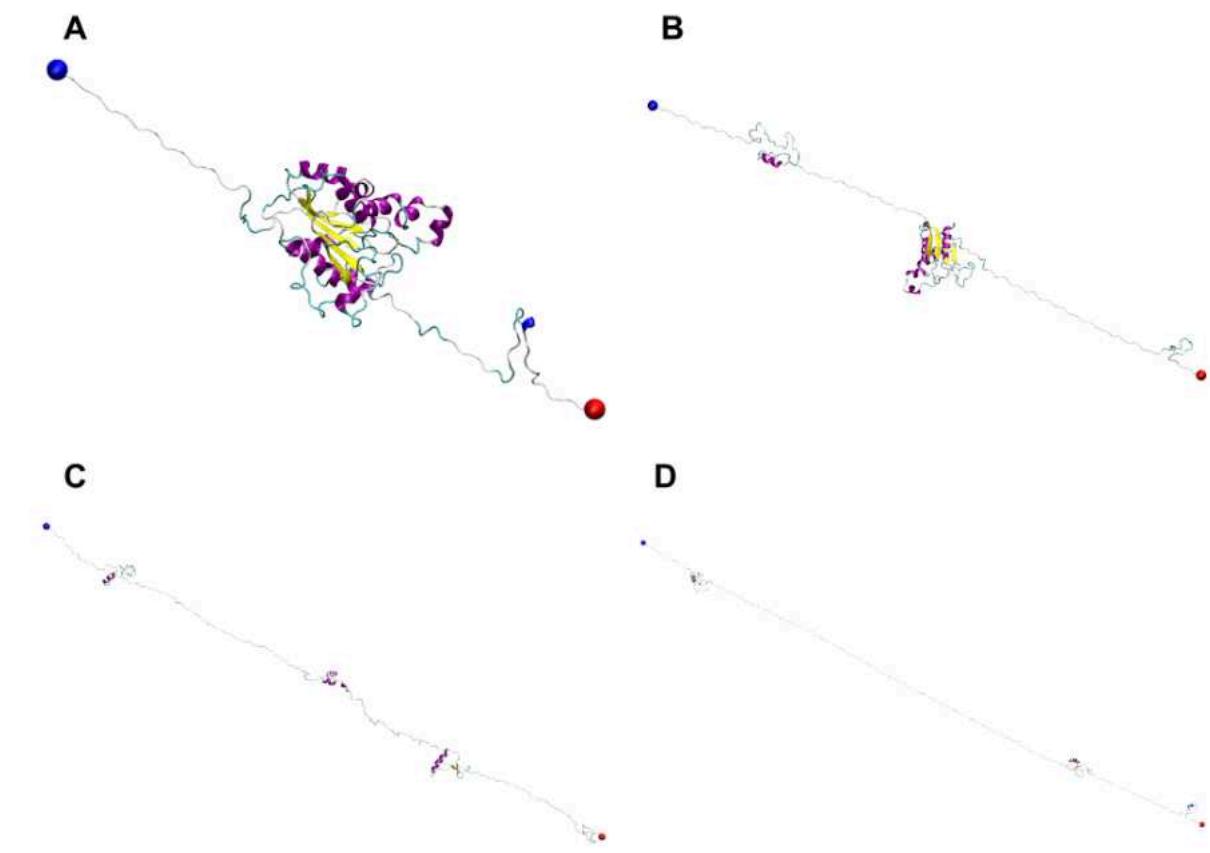
A análise de SMD calcula as coordenadas atômicas identificando as modificações estruturais consequentes da força aplicada propositalmente para desnaturar a proteína, e elas são gravadas no arquivo várias vezes ao longo de uma simulação. Assim, ele fornece uma trajetória do sistema durante o tempo de execução. Um outro arquivo gerado foi um arquivo de registro da simulação (.log), esse arquivo contém as informações da dinâmica. Essas linhas incluem: o passo de tempo atual, a posição atual do centro de massa dos átomos restringidos e a força atual aplicada ao centro de massa (PHILLIPS et al., 2020). A evolução temporal da SMD de uma réplica de cada enzima pode ser vista nas Figuras 6, 7 e 8.

Figura 6 - Evolução temporal da SMD de CalB selvagem (átomo fixo em azul e átomo tracionado em vermelho). (A) 12,5 ns. (B) 25 ns. (C) 37,5 ns. (D) 50 ns.



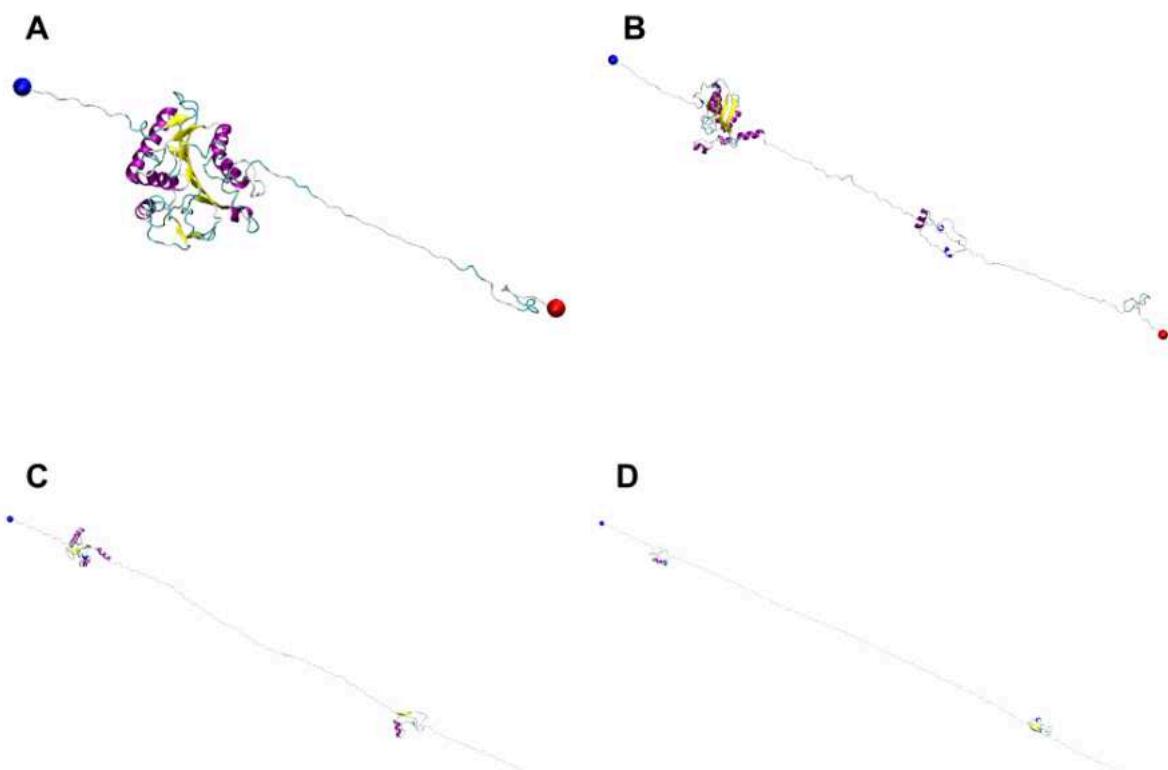
Fonte: Autor

Figura 7 - Evolução temporal da SMD da mutante A281E (átomo fixo em azul e átomo tracionado em vermelho). (A) 12,5 ns. (B) 25 ns. (C) 37,5 ns. (D) 50 ns.



Fonte: Autor

Figura 8 - Evolução temporal da SMD da mutante R249L (átomo fixo em azul e átomo tracionado em vermelho). (A) 12,5 ns. (B) 25 ns. (C) 37,5 ns. (D) 50 ns.



Fonte: Autor

Nas Figuras 6, 7 e 8, observa-se que as pontes dissulfeto permanecem intactas. Isso ocorre porque são ligações covalentes, as quais não se rompem durante a SMD.

Os vídeos das simulações de todas as réplicas podem ser vistos nos *links* abaixo.

CalB_WT_r1: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/CalB_WT_r1.gif

CalB_WT_r2: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/CalB_WT_r2.gif

CalB_WT_r3: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/CalB_WT_r3.gif

A281E_r1: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/A281E_r1.gif

A281E_r2: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/A281E_r2.gif

A281E_r3: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/A281E_r3.gif

R249L_r1: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/R249_r1.gif

R249L_r2: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/R249_r2.gif

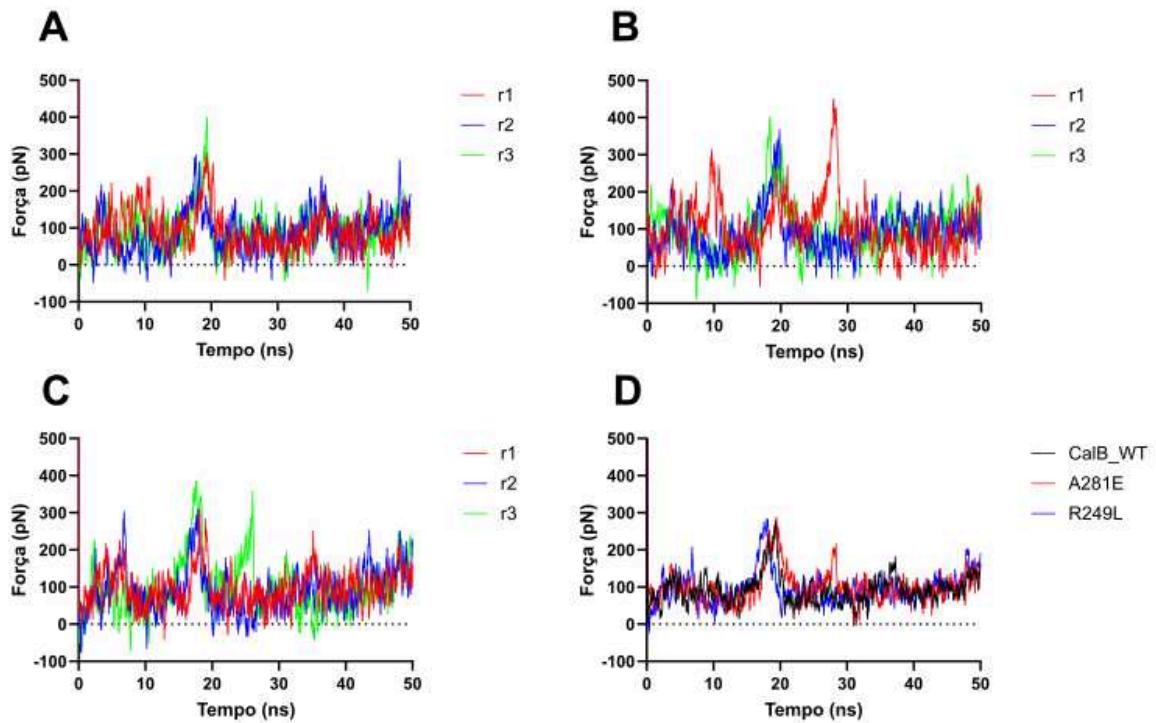
R249L_r3: https://www.bioinfo.iq.ufrj.br/permanent/calb/smd/R249_r3.gif

As simulações de SMD foram realizadas a uma temperatura de 298,15 K, com o uso de solvente implícito GBIS, aplicando no modelo uma constante dielétrica de 78,5 (água) e concentração salina de 0,15 mol L⁻¹, visando reduzir o custo computacional. Embora essas condições sejam mais simples — incluindo a ausência de tampão HEPES, NaCl e controle preciso de temperatura, elas foram suficientes para avaliar a resistência mecânica relativa das variantes, permitindo comparações estruturais coerentes com a estabilidade observada em bancada (KIM et al., 2010).

4.4 ANÁLISE DE DADOS

Com os resultados obtidos das simulações de SMD, foram construídos gráficos de força vs tempo, os gráficos com os dados suavizados podem ser vistos na Figura 9, os gráficos com os dados brutos podem ser vistos no Anexo L.

Figura 9 - Gráficos suavizados de força vs tempo. (A) CalB_WT. (B) A281E. (C) R249L. (D) Médias das réplicas.

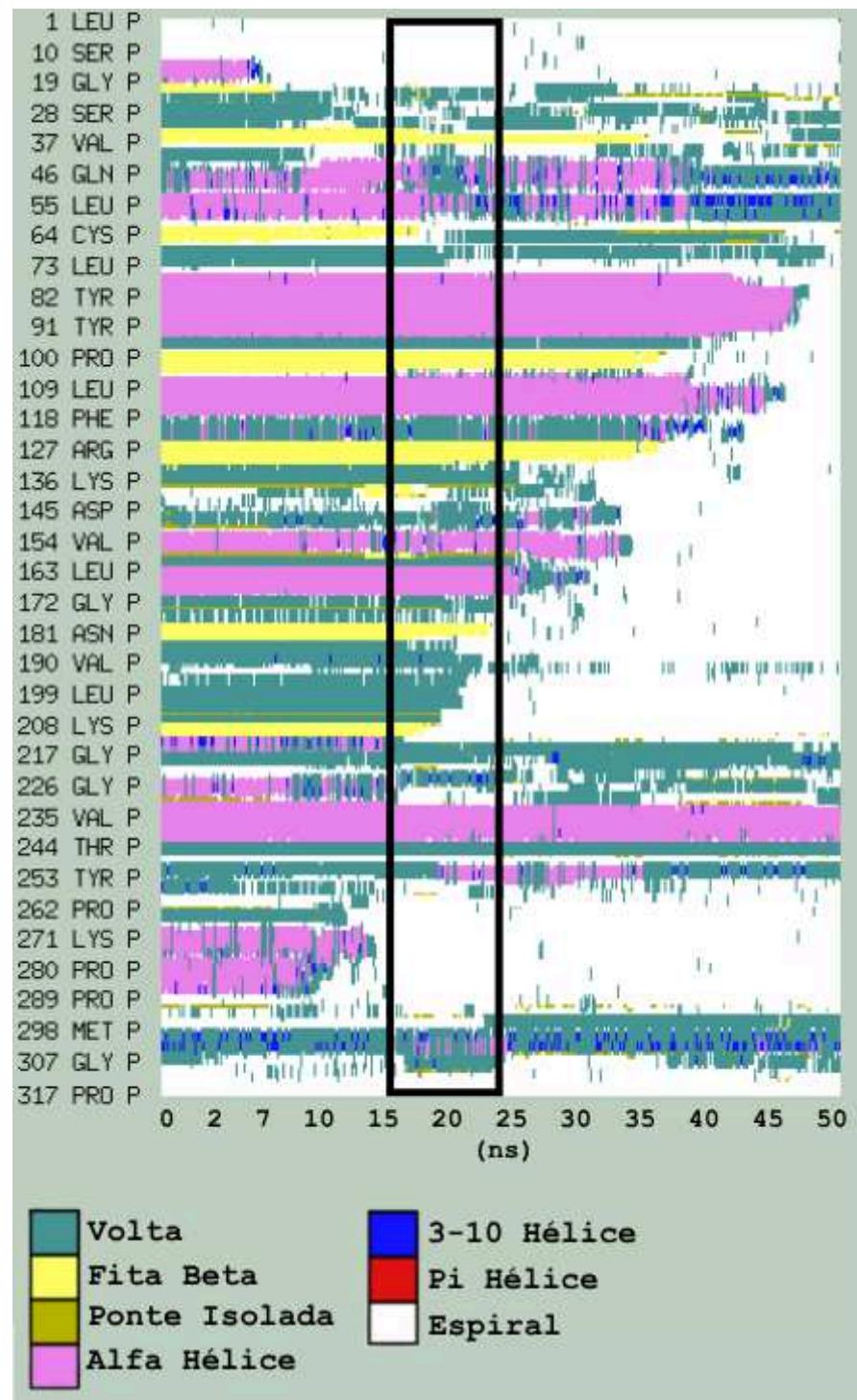


Fonte: Autor

A análise dos gráficos de força em função do tempo revelou um padrão comum entre as três variantes da CalB: um pico pronunciado de força, seguido de uma queda abrupta e posterior estabilização em níveis de força mais moderados.

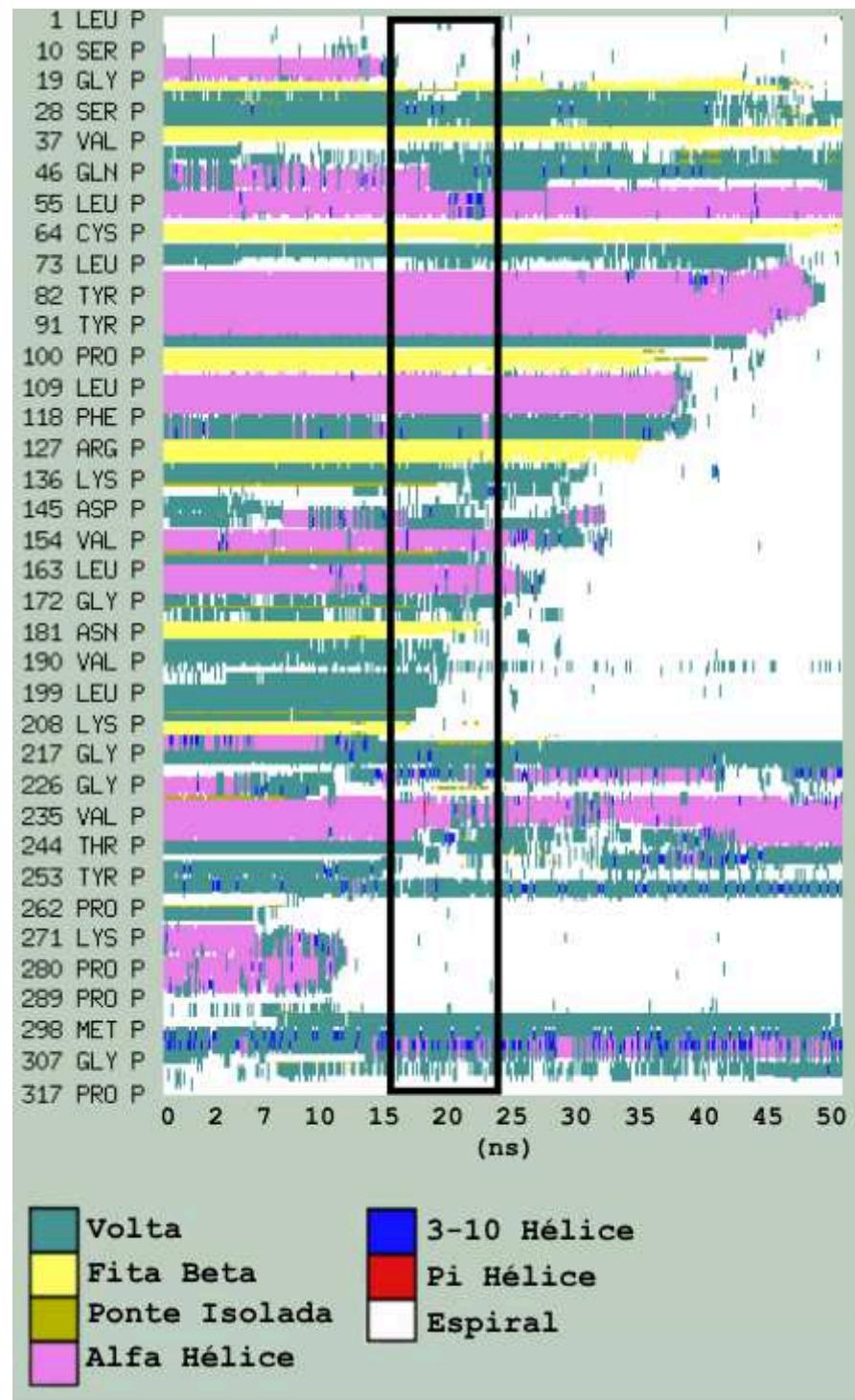
Com o intuito de averiguar o que ocorria com as proteínas no intervalo em que foi observado esse pico de força no gráfico, foram construídos gráficos de estrutura secundária vs tempo, sendo destacado por um retângulo preto o intervalo no qual foi observado esse pico em comum (entre 16 e 23 ns), os gráficos podem ser vistos nas Figuras de 10 a 18.

Figura 10 - Gráfico de estrutura secundária vs tempo da réplica CalB_WT_r1.



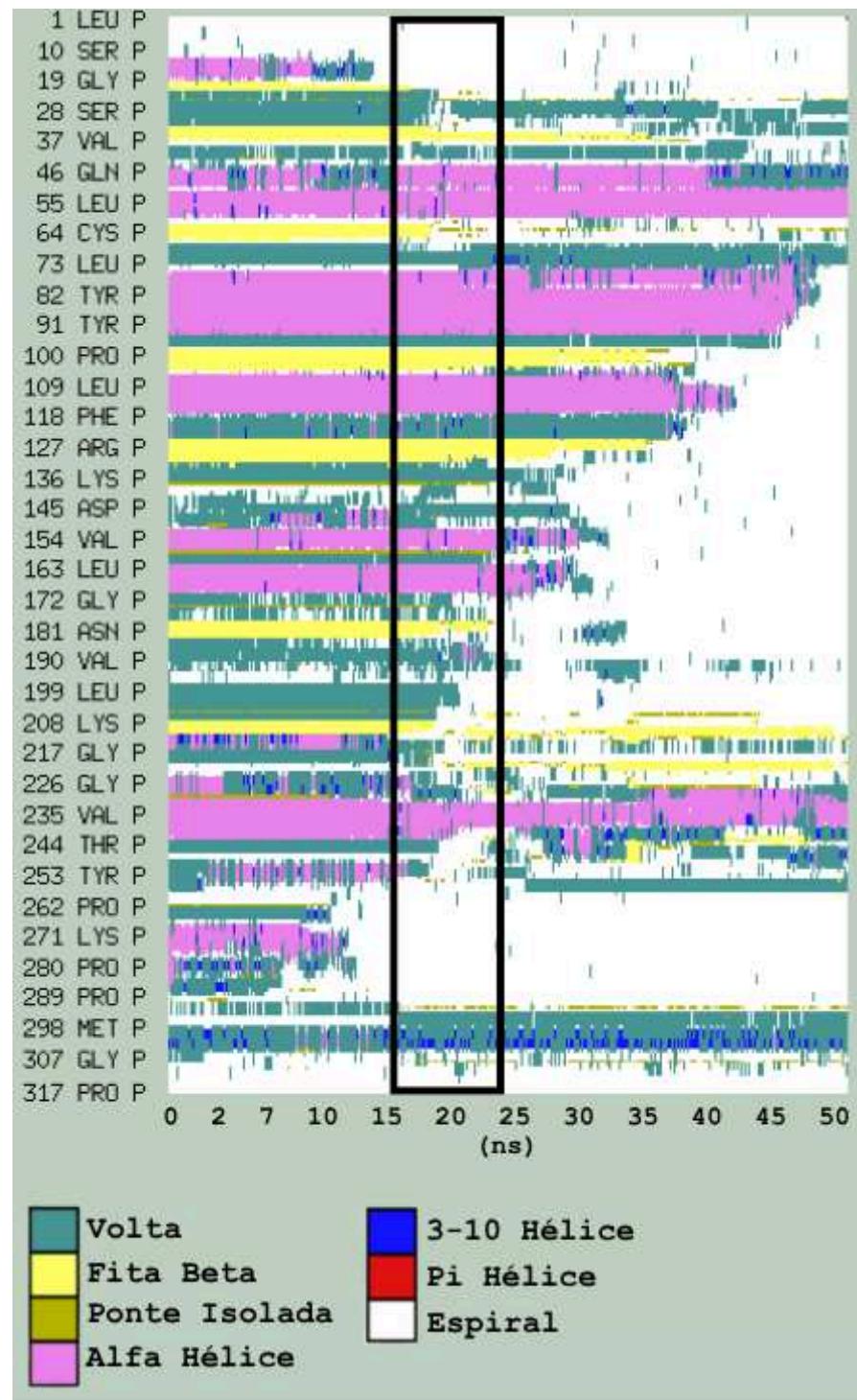
Fonte: Autor

Figura 11 - Gráfico de estrutura secundária vs tempo da réplica CalB_WT_r2.



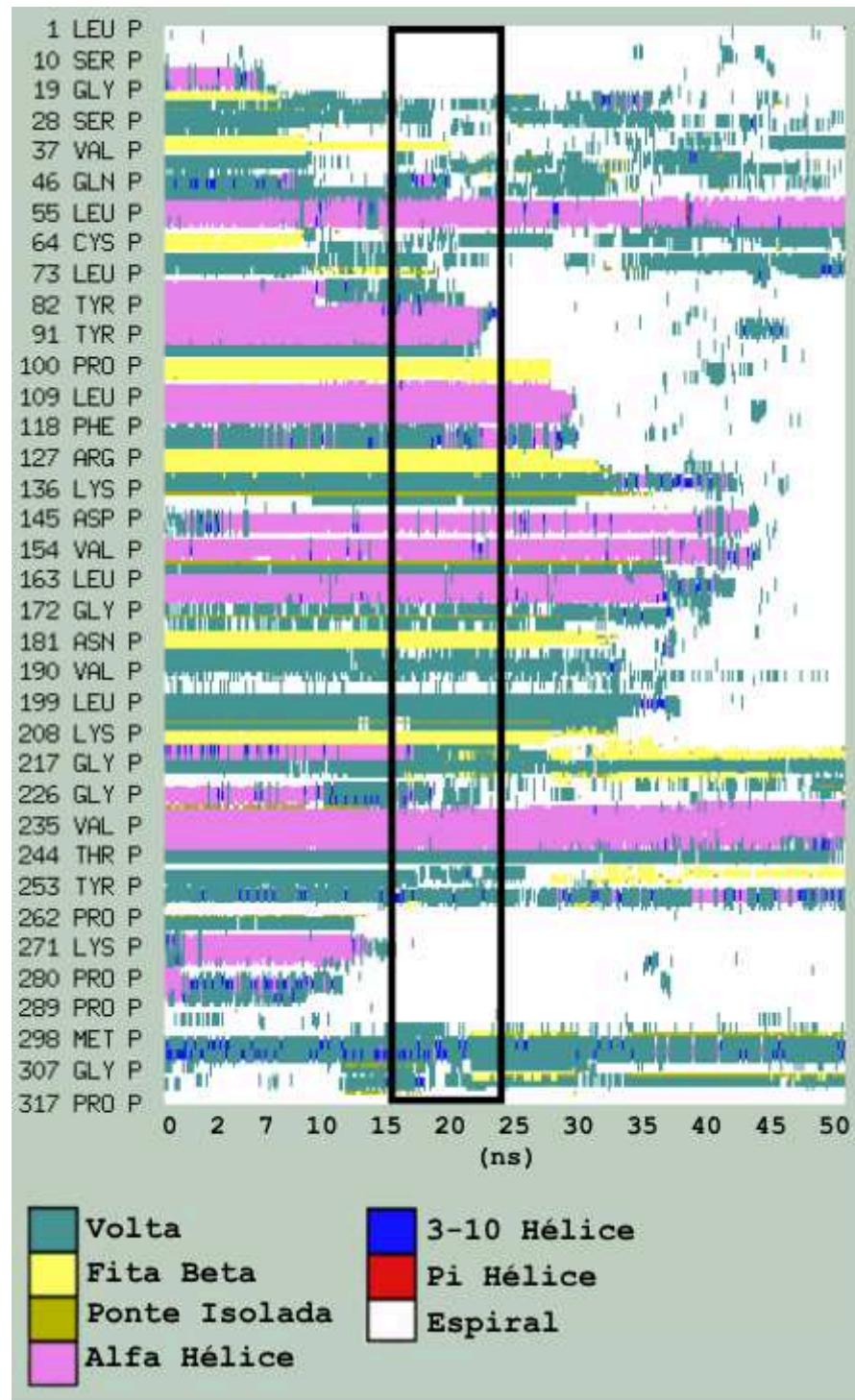
Fonte: Autor

Figura 12 - Gráfico de estrutura secundária vs tempo da réplica CalB_WT_r3.



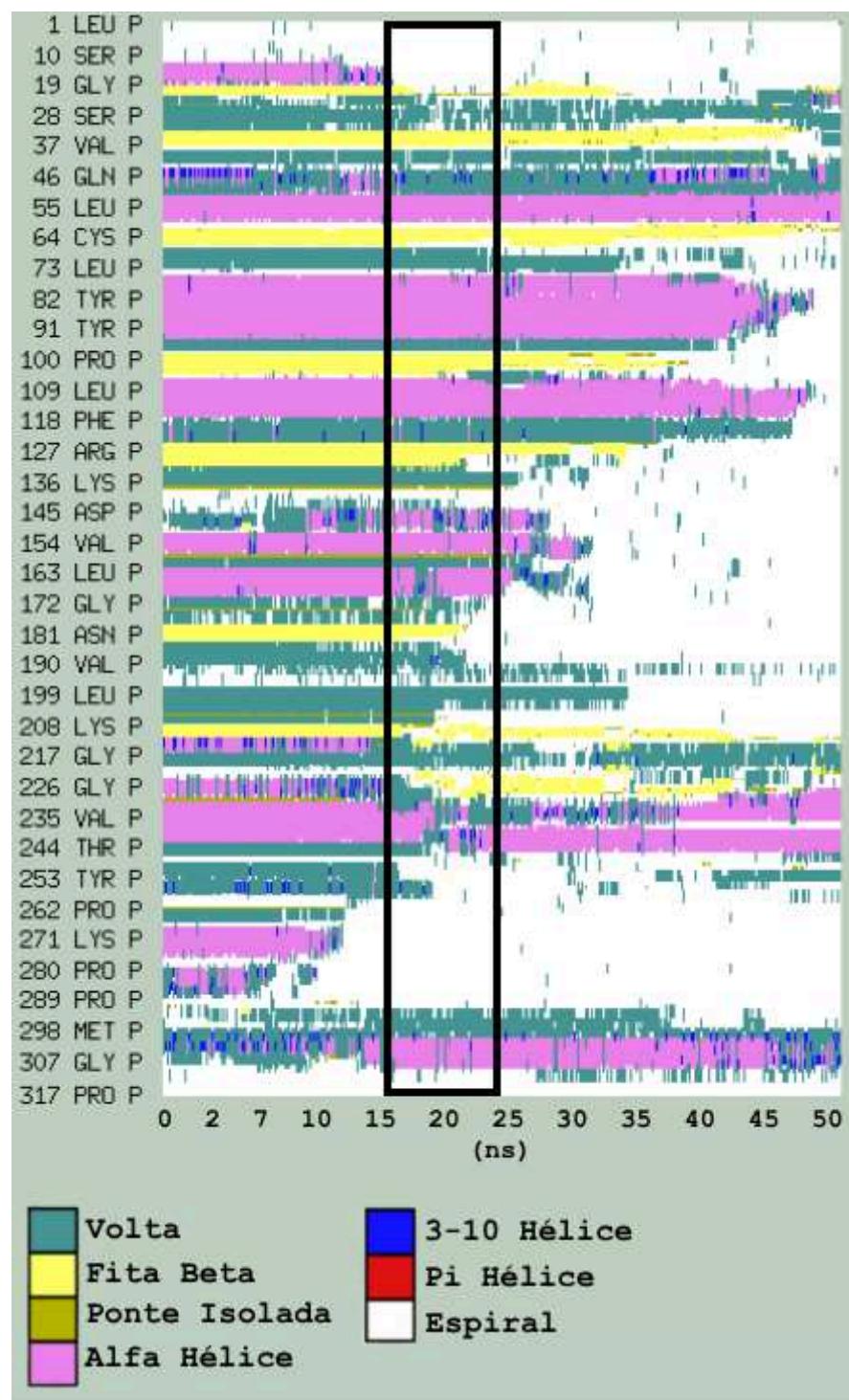
Fonte: Autor

Figura 13 - Gráfico de estrutura secundária vs tempo da réplica A281E_r1.



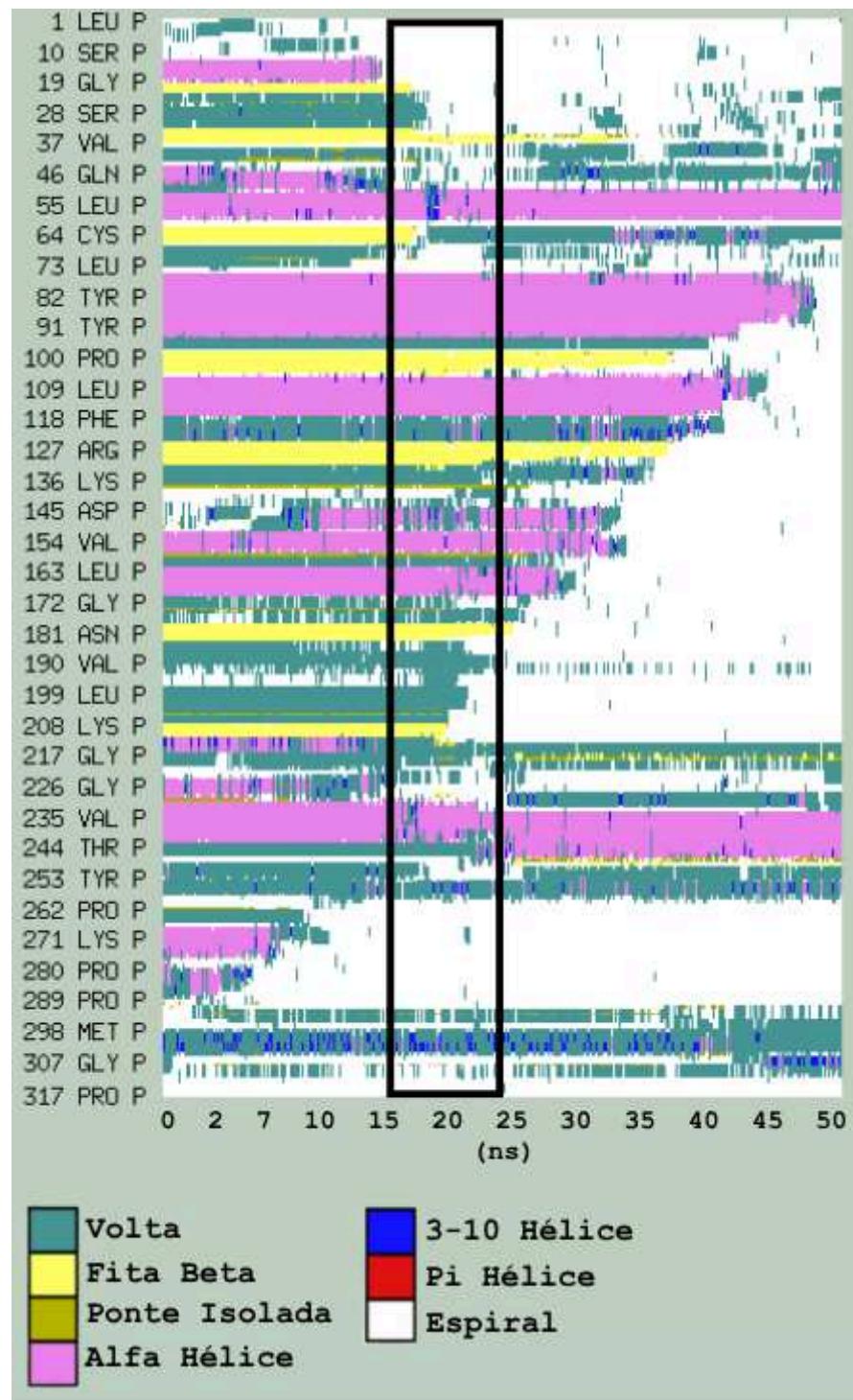
Fonte: Autor

Figura 14 - Gráfico de estrutura secundária vs tempo da réplica A281E_r2.



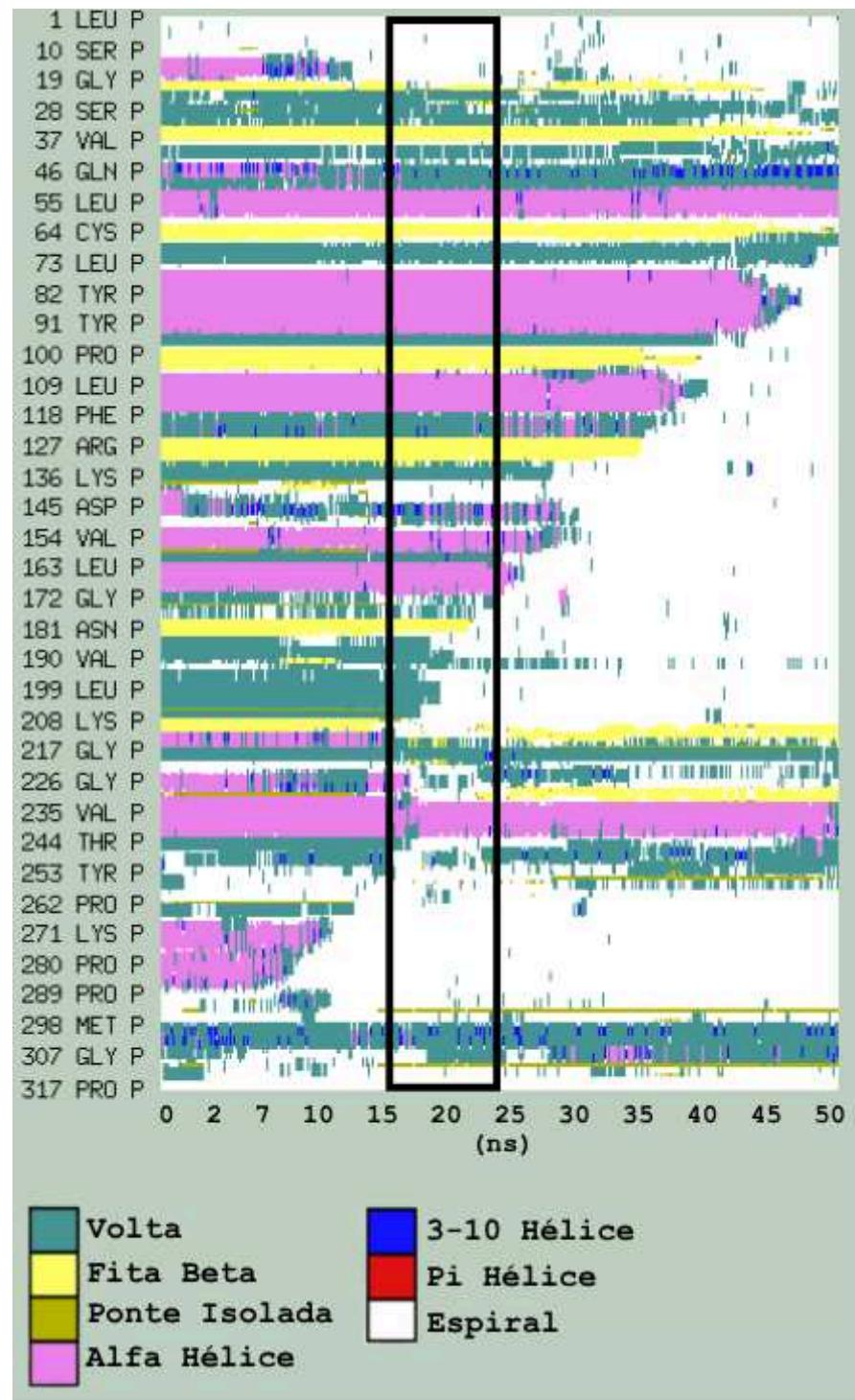
Fonte: Autor

Figura 15 - Gráfico de estrutura secundária vs tempo da réplica A281E_r3.



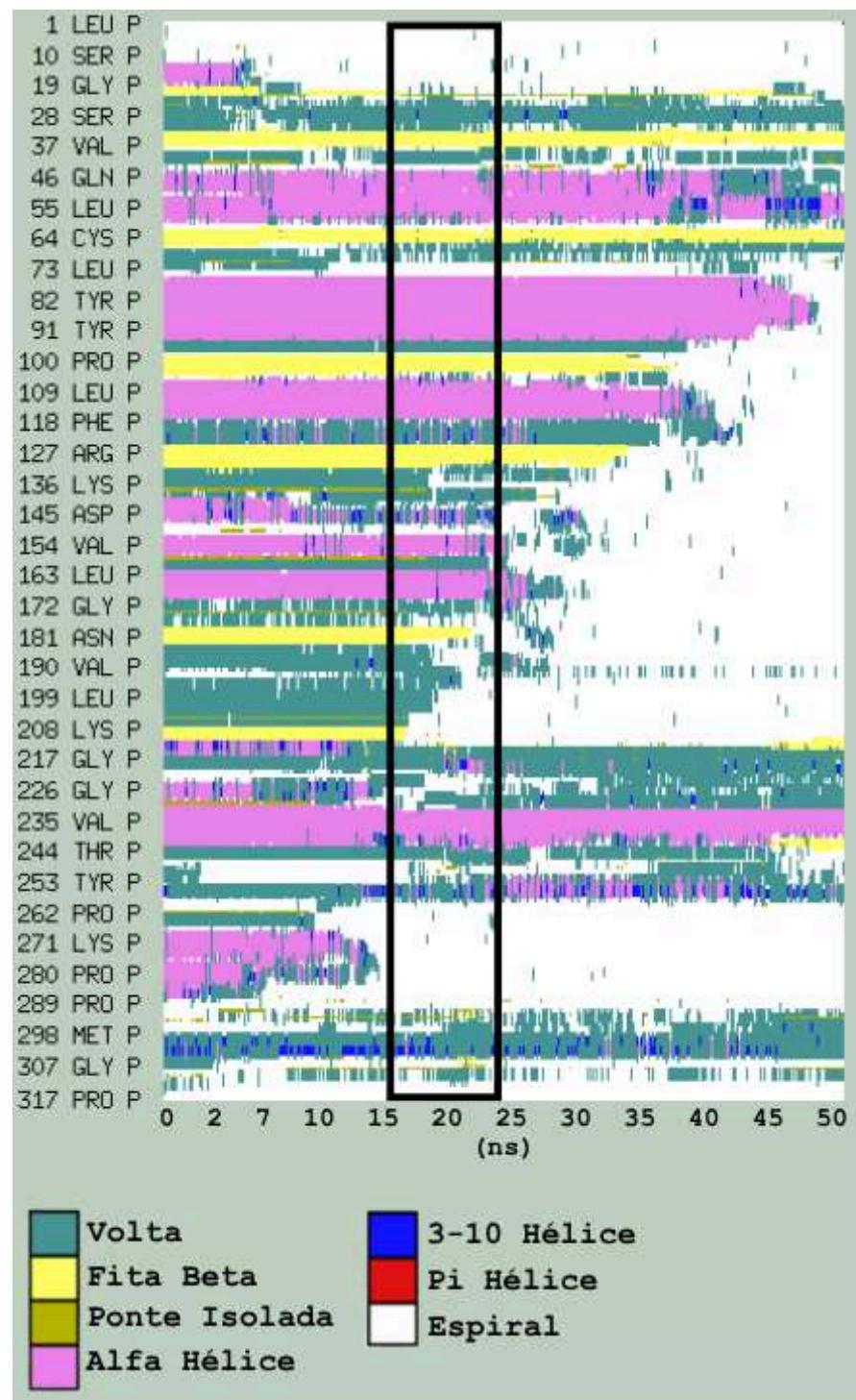
Fonte: Autor

Figura 16 - Gráfico de estrutura secundária vs tempo da réplica R249L_r1.



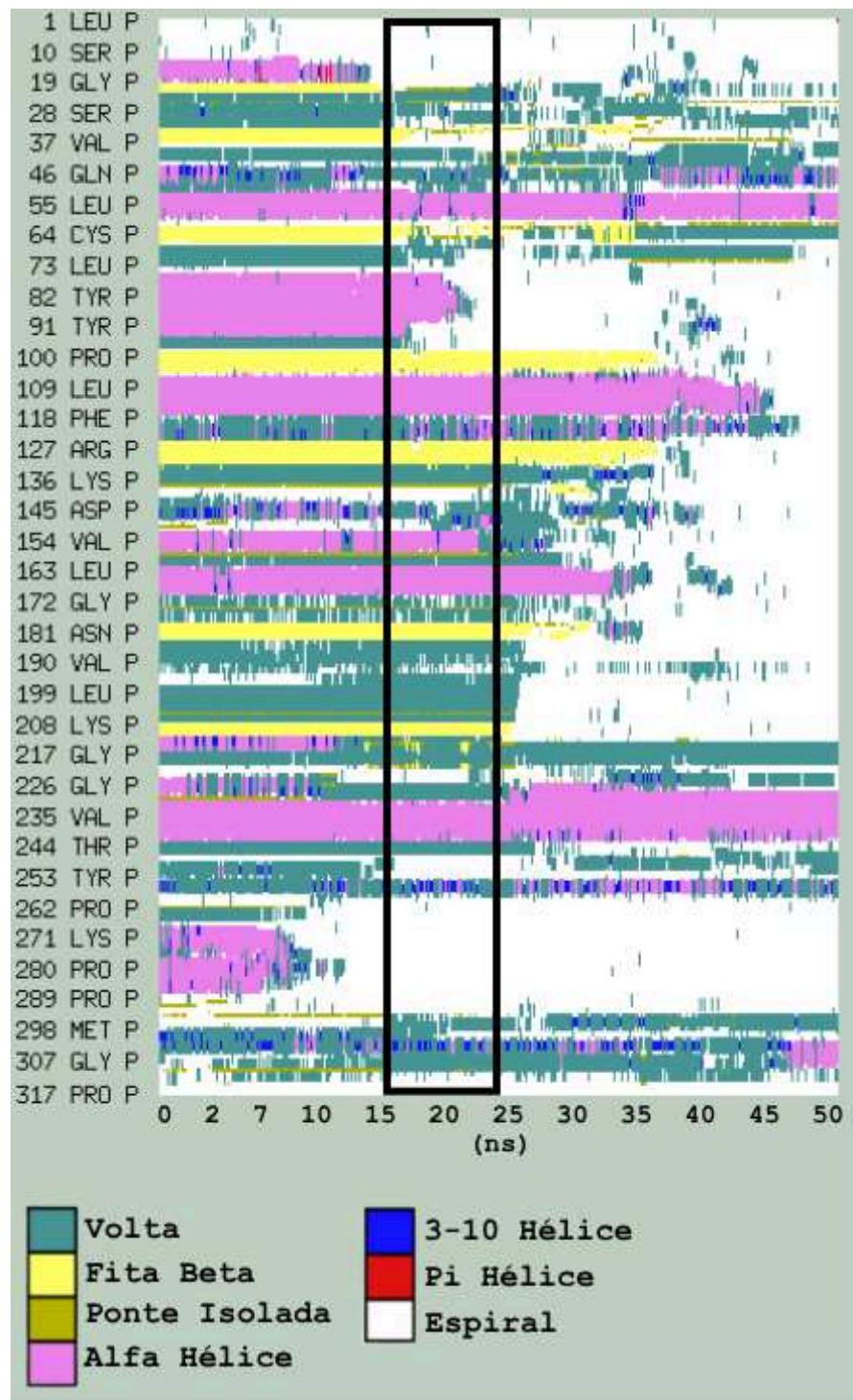
Fonte: Autor

Figura 17 - Gráfico de estrutura secundária vs tempo da réplica R249L_r2.



Fonte: Autor

Figura 18 - Gráfico de estrutura secundária vs tempo da réplica R249L_r3.



Fonte: Autor

A partir dos gráficos de estrutura secundária em função do tempo, foi possível verificar que no intervalo entre 16 e 23 ns em que foi observado o pico em comum, as estruturas secundárias que foram desenoveladas estavam entre os resíduos Gly172 e Lys208 (volta, ponte isolada e fita beta), para todas as réplicas de CalB selvagem e para as réplicas 1 e 2 da mutante R249L.

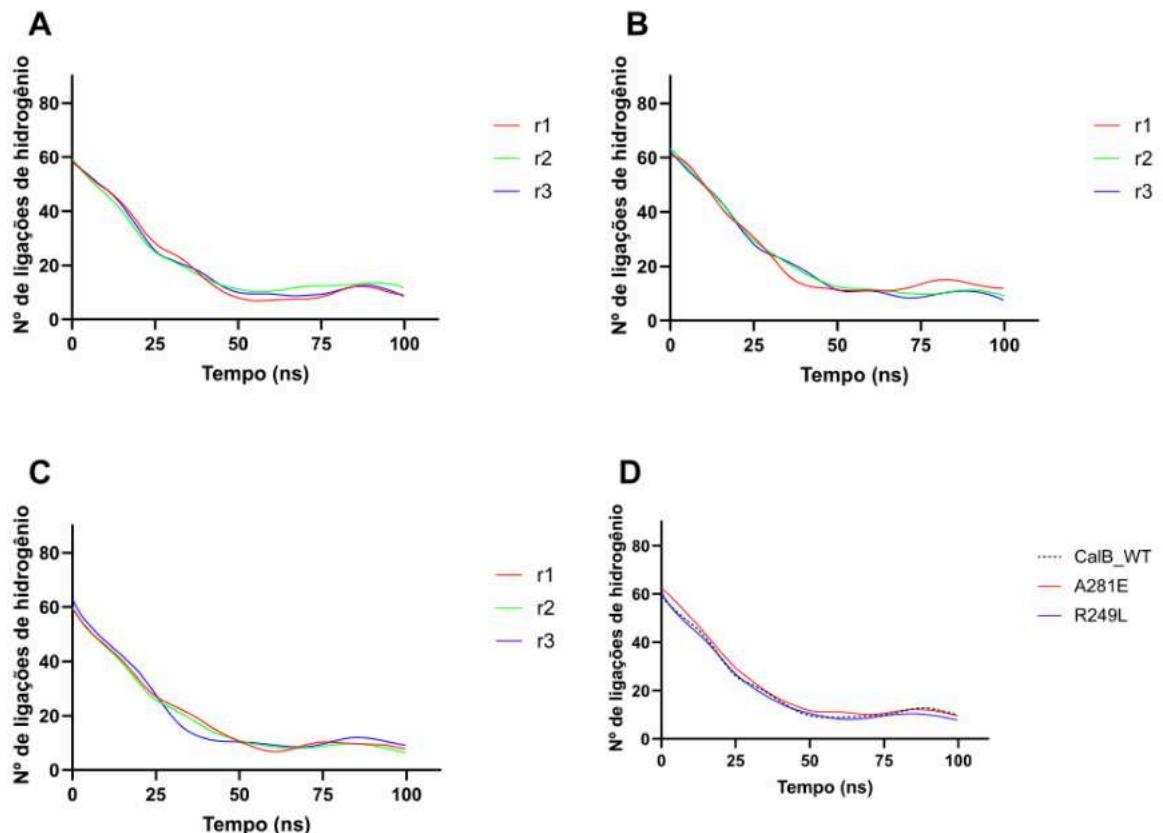
Para a réplica 3 da mutante R249L, o desenovelamento dessa região entre os resíduos Gly172 e Lys208 ocorre posteriormente, em torno de 25 ns (Figura 18) e há um pico de força em 25 ns (Figura 9C).

As réplicas da mutante A281E, no intervalo entre 16 e 23 ns, apresentaram modos de desenovelamento distintos desse padrão, e distintos entre si, o que sugere alta variabilidade conformacional.

Esse pico de força sugere a existência de um estado intermediário estrutural que exige uma maior força para ser desfeito durante o processo de desenovelamento, possivelmente correspondente a um *molten globule*. No contexto da Dinâmica Molecular Dirigida, esse tipo de transição é interpretado como a ruptura de um núcleo estrutural estável, com acúmulo temporário em uma conformação parcialmente organizada antes da desnaturação completa. Esse fenômeno já foi descrito na literatura, onde foram observados estados intermediários em simulações de desenovelamento mecânico por SMD, evidenciando quedas de força seguidas de regiões estáveis como indicativo de estruturas metaestáveis ao longo da tração. A persistência da força mecânica em níveis intermediários após o pico principal reforça a ideia de que a proteína não se torna totalmente desenovelada imediatamente (LU & SCHULTEN, 1999).

Com os resultados obtidos do cálculo do número de ligações de hidrogênio foram construídos gráficos. Os gráficos com os dados suavizados podem ser vistos na Figura 19 e os gráficos brutos podem ser vistos no Anexo M.

Figura 19 - Gráficos suavizados do número de ligações de hidrogênio vs tempo. (A) CalB_WT. (B) A281E. (C) R249L. (D) Médias das réplicas.



Fonte: Autor

De maneira geral, as réplicas e as mutantes foram semelhantes entre si, sendo que após 50 ns, o número de ligações de hidrogênio praticamente não varia.

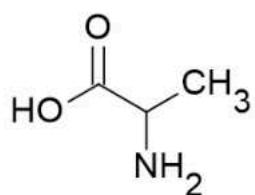
Na Figura 19D, é possível notar que o número de ligações de hidrogênio da mutante A281E é um pouco maior que o das outras mutantes, isso pode ser explicado pelo fato da mutante A281E ter sua mutação originada da troca de uma alanina (A), um aminoácido de cadeia lateral apolar e muito curta, pelo glutamato (E), um aminoácido carregado negativamente e cadeia lateral mais longa, sendo que o glutamato permite fazer um número maior de ligações de hidrogênio (KIM et al., 2010). Embora à primeira vista pareça contraditório que a mutante termoinstável apresente um número maior de ligações de hidrogênio, é importante destacar que a diferença no número absoluto de ligações observadas é relativamente baixa.

O Glu281 da A281E encontra-se próximo a outros resíduos igualmente carregados negativamente (Glu81, Glu188, Glu269 e Glu294), o que pode gerar repulsão eletrostática significativa. Essa concentração de cargas negativas em uma mesma região da proteína pode comprometer o empacotamento local e, consequentemente, reduzir a estabilidade estrutural da enzima. Assim, ainda que a mutante apresente um número levemente maior de ligações de hidrogênio, essas repulsões contribuem para a termoinstabilidade observada experimentalmente (KIM et al., 2010).

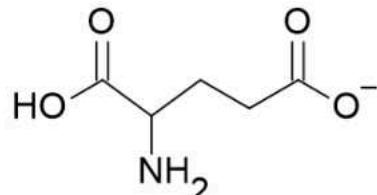
A análise da área de superfície acessível à água indicou que os resíduos Glu281 e Leu249 apresentam alta acessibilidade ao solvente ($273,29 \text{ \AA}^2$ e $285,63 \text{ \AA}^2$, respectivamente), o que sugere que ambas as mutações ocorrem em regiões expostas da proteína. Isso corrobora a hipótese de que a mutação A281E introduz uma carga negativa em uma região já acessível, potencializando efeitos desestabilizadores por repulsão eletrostática local. As estruturas dos aminoácidos citados podem ser vistas na Figura 20.

Figura 20 - Estruturas dos aminoácidos envolvidos nas mutações.

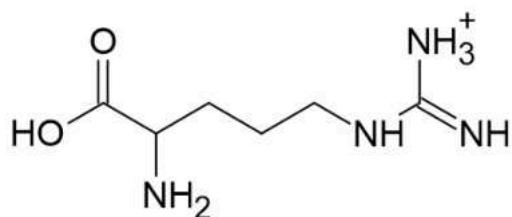
A - Alanina (A)



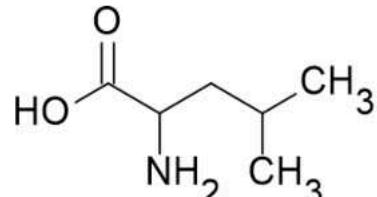
B - Glutamato (E)



C - Arginina (R)



D - Leucina (L)



Fonte: Autor

Os resultados dos valores de trabalho podem ser vistos na Tabela 1, eles foram obtidos de gráficos de força vs deslocamento (disponíveis no Anexo N), pelo processamento da área sob a curva, de 0 a 750 Å, por meio do método trapezoidal.

Tabela 1 - Cálculo do trabalho realizado no desenovelamento mecânico das proteínas.

Variante	Réplica 1 (kcal/mol)	Réplica 2 (kcal/mol)	Réplica 3 (kcal/mol)	Média (kcal/mol)	Desvio Padrão (kcal/mol)
CalB_WT	1169,75	1166,01	1172,95	1169,56	3,46
A281E	1298,77	1149,64	1217,80	1222,07	74,67
R249L	1243,61	1197,65	1323,19	1254,82	63,51

Fonte: Autor

Como pode ser observado na Tabela 1, a CalB selvagem foi a que possuiu menor variação entre suas réplicas (menor desvio padrão) e a que requereu a menor quantidade de trabalho necessário para desenovelamento (1169,56 kcal/mol). Em contrapartida, a mutante R249L foi a que requereu a maior quantidade de trabalho para ser desenovelada (1254,82 kcal/mol), e a mutante A281E foi a que apresentou a maior variabilidade entre suas réplicas (maior desvio padrão). As diferenças nos valores médios e nos desvios padrão sugerem alterações estruturais relevantes entre as formas enzimáticas.

A CalB selvagem apresentou a menor média de trabalho (1169,56 kcal/mol) e o menor desvio padrão (3,46 kcal/mol), indicando consistência entre as réplicas e sugerindo uma estrutura relativamente estável e flexível. Essas características podem estar associadas a uma conformação mais acessível e menos compacta, o que está de acordo com seu bom desempenho catalítico já conhecido na literatura (UPPENBERG et al., 2002). Embora a SMD não simule diretamente a atividade enzimática, a resposta mecânica observada pode refletir características estruturais compatíveis com sua função biológica.

O fato da variante R249L ter apresentado a maior média de trabalho (1254,82 kcal/mol) é coerente com a sua termoestabilidade observada experimentalmente,

evidenciando de que se trata de uma enzima mais estável frente à desnaturação tanto térmica quanto mecânica (KIM et al., 2010).

A mutante A281E, por sua vez, ter apresentado o maior desvio padrão dentre as três, indica uma maior variabilidade conformacional entre as réplicas. Essa variabilidade conformacional observada se mostra coerente com a sua termoinstabilidade observada experimentalmente (KIM et al., 2010).

Apesar da média de trabalho da mutante A281E ser superior à da CalB selvagem, o seu alto desvio padrão faz com que os intervalos de variação se sobreponham. Isso indica que, em algumas simulações, a A281E pode exigir menos trabalho para ser desenovelada do que a CalB. Com um número maior de réplicas, seria possível uma inferência mais precisa sobre essa relação.

É válido destacar que as simulações foram realizadas com versões deglicosiladas das enzimas, o que pode ter influenciado os resultados de estabilidade estrutural observados. A literatura aponta que a ausência de glicosilação em variantes da CalB reduz sua estabilidade térmica, resultando em perda significativa de atividade residual após aquecimento, quando comparadas à forma glicosilada (HÖCK et al., 2018). Dessa forma, é possível que a resistência mecânica estimada nas simulações tenha sido inferior à que seria observada em proteínas glicosiladas. No entanto, como todas as variantes simuladas estavam igualmente deglicosiladas, é provável que a ausência de glicosilação não tenha comprometido a validade da comparação entre os resultados obtidos.

O uso do campo de força CHARMM36m foi essencial para garantir a fidelidade estrutural das simulações. Essa versão refinada do CHARMM36 foi desenvolvida para representar com maior precisão proteínas tanto enoveladas quanto desordenadas, possibilitando uma modelagem mais realista dos movimentos conformacionais induzidos durante a SMD (HUANG et al., 2016). Assim, os parâmetros empregados foram adequados para investigar alterações estruturais associadas à estabilidade das variantes da CalB.

A adoção do modelo de solvente implícito GBIS contribuiu para reduzir o custo computacional das simulações, mantendo uma representação eficiente do ambiente aquoso. Embora esse modelo não capture interações específicas com

moléculas individuais de água, ele é considerado apropriado para simulações que envolvem análise da estabilidade mecânica e desenovelamento proteico, como a SMD, proporcionando um bom equilíbrio entre precisão e eficiência (PHILLIPS et al., 2020).

A aplicação de SMD demonstrou ser útil para revelar tendências estruturais relacionadas à estabilidade e flexibilidade de enzimas e suas variantes (PATEL et al., 2014), fornecendo dados em nível atômico que complementam informações experimentais (PHILLIPS et al., 2020). No entanto, é importante destacar que as conclusões apresentadas são limitadas pelo número de réplicas executadas. A inclusão de um maior número de repetições seria essencial para reforçar a confiabilidade estatística dos resultados e minimizar a influência de flutuações pontuais.

5 CONCLUSÃO

Diante do exposto, os resultados obtidos por meio das simulações de Dinâmica Molecular Dirigida permitiram identificar diferenças no comportamento mecânico de desenovelamento entre a CalB selvagem e suas variantes R249L e A281E. A variante R249L apresentou maior resistência mecânica, em consonância com sua maior termoestabilidade experimental, enquanto a A281E mostrou maior variabilidade conformacional, compatível com sua instabilidade térmica. A detecção de um possível estado intermediário durante o desenovelamento reforça o potencial da SMD para investigar características estruturais metaestáveis. Os dados obtidos demonstram que a SMD pode ser uma ferramenta útil na avaliação da estabilidade de enzimas mutantes, oferecendo informações complementares às análises experimentais.

6 REFERÊNCIAS

- BRITO E CUNHA, D. A. et al. Structural differences of commercial and recombinant lipase B from *Candida antarctica*: An important implication on enzymes thermostability. *International Journal of Biological Macromolecules*, v. 140, p. 761–770, 1 nov. 2019.
- FARIA, P. E. et al. Unveiling six novel CALB-like lipases using genome-centric and patent-driven prospection. *Enzyme and Microbial Technology*, v. 181, p. 110525, 1 dez. 2024.
- HÖCK, H. et al. Comparison of *Candida Antarctica* Lipase B Variants for Conversion of ε-Caprolactone in Aqueous Medium—Part 2. *Polymers*, v. 10, n. 5, p. 524, 14 maio 2018.
- HUANG, J. et al. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nature Methods*, v. 14, n. 1, p. 71–73, 7 nov. 2016.
- HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, v. 14, n. 1, p. 33–38, 1 fev. 1996.
- KIM, H. S.; LE, Q. A. T.; KIM, Y. H. Development of thermostable lipase B from *Candida antarctica* (CalB) through in silico design employing B-factor and RosettaDesign. *Enzyme and Microbial Technology*, v. 47, n. 1-2, p. 1–5, 5 jul. 2010.
- KUNDYS, A. et al. *Candida Antarctica* Lipase B as Catalyst for Cyclic Esters Synthesis, Their Polymerization and Degradation of Aliphatic Polyesters. *Journal of Polymers and the Environment*, v. 26, n. 1, p. 396–407, 24 jan. 2017.
- LU, H.; SCHULTEN, K. Steered Molecular Dynamics Simulations of Force-Induced Protein Domain Unfolding. *Proteins: Structure, Function, and Bioinformatics*, v. 35, n. 4, p. 453–463, 1 jun. 1999.
- PATEL, J. S. et al. Steered Molecular Dynamics Simulations for Studying Protein–Ligand Interaction in Cyclin-Dependent Kinase 5. *Journal of Chemical Information and Modeling*, v. 54, n. 2, p. 470-480, 28 jan. 2014.
- PHILLIPS, J. C. et al. (2017). *NAMD Tutorial*. Disponível em: <<https://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/index.html>> Acesso em: 9 jun. 2025.
- PHILLIPS, J. C. et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *The Journal of Chemical Physics*, v. 153, n. 4, p. 044130, 30 jul. 2020.
- POOJARI, Y.; CLARSON, S. J. Thermal stability of *Candida antarctica* lipase B immobilized on macroporous acrylic resin particles in organic media. *Biocatalysis and Agricultural Biotechnology*, v. 2, n. 1, p. 7–11, 1 jan. 2013.
- RABBANI, G. et al. Impact of Structural Stability of Cold Adapted *Candida Antarctica* Lipase B (CaLB): In Relation to pH, Chemical and Thermal Denaturation. *RSC*

Advances, v. 5, n. 26, p. 20115–20131, 17 fev. 2015.

UPPENBERG, J. et al. Crystallographic and Molecular-Modeling Studies of Lipase B from *Candida Antarctica* Reveal a Stereospecificity Pocket for Secondary Alcohols. 1 maio 2002.

XU, J. et al. Stereodivergent Protein Engineering of a Lipase To Access All Possible Stereoisomers of Chiral Esters with Two Stereocenters. *Journal of the American Chemical Society*, 26 abr. 2019.

ANEXO A - Script “CalB_GBiS.pgn”

```
# VMD script to build CalB topology containing all post-translational
modifications

package require psfgen

topology ../CHARMM36m/top_all36_prot.rtf
topology ../CHARMM36m/top_all36_carb.rtf
topology ../CHARMM36m/stream/carb/toppar_all36_carb_glycopeptide.str

pdbalias residue HIS HSP
pdbalias atom ILE CD1 CD
pdbalias atom PRO OXT OT2

segment PROA {pdb CalB_WT_base.pdb}
coordpdb CalB_WT_base.pdb PROA

patch ASPP PROA:134 ; # Protonation patches
patch GLUP PROA:294

patch DISU PROA:22 PROA:64 ; # Disulfide bridges
patch DISU PROA:216 PROA:258
patch DISU PROA:293 PROA:311

regenerate angles dihedrals
guesscoord

writepdb CalB_WT_base_NAMD.pdb
writepsf CalB_WT_base_NAMD.psf

package require solvate
exit
```

ANEXO B - Script “minim_equil.conf”

```

#####
## JOB DESCRIPTION ##

#####
# Minimization and equilibration

#####
## ADJUSTABLE PARAMETERS ##

#####
structure ..../1Preparation/CalB_WT_base_NAMD.psf
coordinates ..../1Preparation/CalB_WT_base_NAMD.pdb
set outname CalB_equil
firsttimestep      0
set temperature     298.15

#####
## SIMULATION PARAMETERS ##

#####

# Input
paraTypeCharmm      on
parameters ..../CHARMM36m/toppar_water_ions_namd.str
parameters ..../CHARMM36m/stream/carb/toppar_all36_carb_glycopeptide.str
parameters ..../CHARMM36m/par_all35_ethers.prm
parameters ..../CHARMM36m/par_all36_carb.prm
parameters ..../CHARMM36m/par_all36_cgenff.prm
parameters ..../CHARMM36m/par_all36_lipid_ljpme.prm
parameters ..../CHARMM36m/par_all36m_prot.prm
parameters ..../CHARMM36m/par_all36_na.prm

```

```

temperature           $temperature

# Implicit Solvent
gbis                 on
solventDielectric    78.5
ionConcentration     0.15 # Salt concentration, in M
alphaCutoff          16.0

# Force-Field Parameters
exclude              scaled1-4
1-4scaling           1.0
cutoff               18.0
switching            on
switchdist           17.0
pairlistdist         20.0

# Integrator Parameters
timestep             1.0  # 2fs/step
rigidBonds           all  # needed for 2fs steps
nonbondedFreq        1
fullElectFrequency   2
stepspercycle        10

# Output control
outputname           $outname
restartfreq          0
dcdfreq               0
xstFreq               0
outputPressure        0
outputEnergies        1000

# Equilibration protocol

```

```
langevin on
langevinDamping 1
langevinTemp $temperature # Target temperature
langevinHydrogen off

# Minimize and equilibrate system in one single run
minimize 100000
reinitvels $temperature
run 1000000 ; # 1 ns
```

ANEXO C - Script “Fix_SMD.tcl”

```
#carregando moleculas no VMD
mol load psf ../1Preparation/CalB_WT_base_NAMD.psf
mol addfile ../2MinEq_r1/CalB_equil.coor

#Seting Fixed atom (resid 1 and name CA)
set allatoms [atomselect top all]
$allatoms set beta 0
set fixedatom [atomselect top "resid 1 and name CA"]
$fixedatom set beta 1

#getting last resid number
set resids [$allatoms get resid]
set lastRESID [lindex $resids end]

#Seting SMD atom
$allatoms set occupancy 0
set smdatom [atomselect top "resid 317 and name CA"]
$smdatom set occupancy 1

#save protein to PDBfile
$allatoms writepdb ./CalB_SMD.ref

exit
```

ANEXO D - *Script “CalB_SMD_NAMD.conf”*

```

#####
## JOB DESCRIPTION ##

#####
# Constant Velocity Pulling N- C- Termini

#####

## ADJUSTABLE PARAMETERS ##

#####

outputname CalB_SMD_NAMD_pcv
set fixedname ../3Fixed_SMD_atoms_r1/CalB_SMD.ref

structure      ../1Preparation/CalB_WT_base_NAMD.psf
coordinates    ../1Preparation/CalB_WT_base_NAMD.pdb

set restart 1
set temperature 298.15

proc get_first_ts { xscfile } {
    set fd [open $xscfile r]
    gets $fd
    gets $fd
    gets $fd line
    set ts [lindex $line 0]
    close $fd
    return $ts
}

```

```

if {$restart == 1} {
    bincoordinates ../2MinEq_r1/CalB_equil.coor
    binvelocities ../2MinEq_r1/CalB_equil.vel ;# remove the "temperature"
entry if you use this!

#extendedSystem ../2MinEq_r1/CalB_equil.xsc
set currenttimestep [get_first_ts ../2MinEq_r1/CalB_equil.xsc]
} else {
    #temperature $temperature
    set currenttimestep 0
}

firsttimestep $currenttimestep

#####
## SIMULATION PARAMETERS ##

#####
# Force-Field Parameters
paraTypeCharmm      on
parameters ../CHARMM36m/toppar_water_ions_namd.str
parameters ../CHARMM36m/stream/carb/toppar_all36_carb_glycopeptide.str
parameters ../CHARMM36m/par_all35_ethers.prm
parameters ../CHARMM36m/par_all36_carb.prm
parameters ../CHARMM36m/par_all36_cgenff.prm
parameters ../CHARMM36m/par_all36_lipid_ljpme.prm
parameters ../CHARMM36m/par_all36m_prot.prm
parameters ../CHARMM36m/par_all36_na.prm

gbis      on
solventDielectric 78.5
ionConcentration 0.15 # Salt concentration, in M

```

```

alphaCutoff          16.0

# Force-Field Parameters

exclude           scaled1-4
1-4scaling        1.0
cutoff            18.0
switching         on
switchdist        17.0
pairlistdist     20.0

# Integrator Parameters

timestep          1.0 ;# 1fs/step
rigidBonds        all
nonbondedFreq     1
fullElectFrequency 2
stepspercycle    10

# Constant Temperature Control

# The temperature control should be disabled in order to disturb the movement
# of the atoms as little as possible in steered molecular dynamics.

# Switch off the Constant Temperature Control by changing:

langevin          on ;# do (on) or do not do (off) langevin dynamics
langevinDamping   1      ;# damping coefficient (gamma) of 5/ps
langevinTemp      $temperature
langevinHydrogen  no    ;# don't couple langevin bath to hydrogens

# Fixed Atoms Constraint (set PDB beta-column to 1)

# However, we need enable the Fixed Atoms Constraint by changing the
# following lines:

fixedAtoms        on
fixedAtomsFile    $fixedname
fixedAtomsCol     B

#####

```

```

## EXTRA PARAMETERS ##

#####
restartfreq      10000 # Original: 10000
dcdfreq          10000
xstFreq          10000
outputEnergies   10000
outputPressure   10000

SMD on
SMDFile ./$fixedname
SMDk 7 # ainda precisa ser encontrado o valor ideal, depende do sistema e
está atrelada à força do SMD
SMDVel 0.000013663 # per timestep
SMDDir 0 0 1
SMDOutputFreq 10000 # Original: 10000

#####

## EXECUTION SCRIPT ##

#####
run 100000000 ; # (100 ns)

```

ANEXO E - Script “namdstats.tcl”

```

puts "Usage: data_avg <logfile> \[<first timestep> <last timestep>\]"
puts "    <first timestep> and <last timestep> may be entered as numbers or"
puts "    <first timestep> = 'first' will start at the beginning of the
simulation"
puts "    <last timestep> = 'last' will go to the end of the simulation"
puts "Usage: data_time <data stream> <logfile> \[<first timestep> <last
timestep>\]"
puts "    <data stream> = BOND, ANGLE, DIHED, IMPRP, ELECT, VDW, BOUNDARY,
MISC, KINETIC, TOTAL, TEMP, TOTAL2, TOTAL3, TEMPAVG"

proc data_avg {logfile {first 0} {last -1}} {
    set file [open $logfile r]
    while { [gets $file line] != -1 } {
        if [regexp "ETITLE:" $line] {set etitles $line}
    }
    close $file
    puts "Calculating averages..."

    set file [open $logfile r]
    set ener ""
    while { [gets $file line] != -1 } {
        if [regexp "ENERGY: " $line] {set ener "$ener $line"}
    }
    close $file

    set l [llength $etitles]
    set lc [expr [llength $ener]/$l]

    for {set k 0} {$k < $l} {incr k} {
        set tstest [lindex $etitles $k]
        if {$tstest == "TS"} {set tsnum $k}
    }
}

```

```

}

for {set k 0} {$k < $lc} {incr k} {
    set ts([expr $k+1]) [lindex $ener [expr $l*$k+$tsnum]]
}
set lastts $ts($k)
set lastline $lc
set firstts $ts(1)
set firstline 1

if {$first != "first"} {
    set k 1
    while {$first > $ts($k)} {
        set firstts $ts([expr $k+1])
        set firstline [expr $k+1]
        incr k
    }
}

if {$last != "last"} {
    set k 1
    while {$last >= $ts($k)} {
        set lastts $ts($k)
        set lastline $k
        incr k
    }
}

puts "CALCULATING DATA FROM Timestep $firstts TO $lastts:"

set k 0
for {set i 1} {$i <= $lc} {incr i} {
    for {set j 1} {$j <= [llength $etitles]} {incr j} {

```

```

    set e($i,$j) [lindex $ener $k]
    incr k
}
}

for {set j 3} {$j <= [llength $etitles]} {incr j} {
    set total 0
    for {set i $firstline} {$i <= $lastline} {incr i} {
        set total [expr $total + $e($i,$j)]
    }
    set avg [expr $total/(1+$lastline-$firstline)]
    puts "[lindex $etitles [expr $j-1]]: $avg"
}

unset ener
unset ts
}

proc data_time {data logfile {first 0} {last -1}} {
    set file [open $logfile r]
    while { [gets $file line] != -1 } {
        if [regexp "ETITLE:" $line] {set etitles $line}
    }
    close $file
    puts "Getting $data data..."

    set file [open $logfile r]
    set ener ""
    while { [gets $file line] != -1 } {
        if [regexp "ENERGY: " $line] {set ener "$ener $line"}
    }
    close $file
}

```

```

set lc [expr [llength $ener]/[llength $etitles]]

set k 0
for {set i 1} {$i <= $lc} {incr i} {
    for {set j 1} {$j <= [llength $etitles]} {incr j} {
        set e($i,$j) [lindex $ener $k]
        incr k
    }
}

set lastts $e($lc,2)
set lastline $lc
set firstts $e(1,2)
set firstline 1

if {$first != "first"} {
    set k 1
    while {$first > $e($k,2)} {
        set firstts $e([expr $k+1],2)
        set firstline [expr $k+1]
        incr k
    }
}

if {$last != "last"} {
    set k 1
    while {$last >= $e($k,2)} {
        set lastts $e($k,2)
        set lastline $k
        incr k
    }
}

```

```
set j 0

switch $data {
    BOND {set j 3}
    ANGLE {set j 4}
    DIHED {set j 5}
    IMPRP {set j 6}
    ELECT {set j 7}
    VDW {set j 8}
    BOUNDARY {set j 9}
    MISC {set j 10}
    KINETIC {set j 11}
    TOTAL {set j 12}
    TEMP {set j 13}
    TOTAL2 {set j 14}
    TOTAL3 {set j 15}
    TEMPAVG {set j 16}
    default {
        puts "Invalid data stream selection"
    }
}

if {$j != 0} {
    set file [open $data.dat w]
    for {set i $firstline} {$i <= $lastline} {incr i} {
        puts $file "$e($i,2) $e($i,$j)"
    }
    close $file
}

unset ener
puts "Done"
}
```

ANEXO F - Script “smd_final.tcl”

```

import numpy as np

# Step

# Caminhos absolutos dos arquivos

set psffile
"/home/user/Downloads/view_mutants/A281E_r1/CalB_TERMO_A281E_NAMD.psf"

set dcdfilename
"/home/user/Downloads/view_mutants/A281E_r1/CalB_SMD_NAMD_pcv.dcd"

# Carrega estrutura e trajetória

mol new $psffile type psf waitfor all

mol addfile $dcdfilename type dcd waitfor all

# Avança um frame para garantir que as coordenadas estejam disponíveis

animate goto 0

display update

# Ativa estrutura secundária baseada em mudança de frame (via trace)

source "/home/user/Downloads/vmd-1.9.4a53/scripts/vmd/sscache.tcl"

start_sscache 0

# Remove representações anteriores, se houver

mol delrep 0 0

# Representação cartoon da proteína

mol representation NewCartoon

mol color Structure

mol selection "protein"

mol material Opaque

mol addrep 0

# Esfera VDW no resíduo inicial (resid 1)

```

```
mol representation VDW 2.0 12
mol color ColorID 0
mol selection "resid 1 and name CA"
mol material Opaque
mol addrep 0

# Esfera VDW no resíduo final (resid 317)
mol representation VDW 2.0 12
mol color ColorID 1
mol selection "resid 317 and name CA"
mol material Opaque
mol addrep 0

# Aplica suavização (smoothing) a todas as representações
for {set i 0} {$i < [molinfo top get numreps]} {incr i} {
    mol smoothrep 0 $i 10
}

# Configurações visuais (fundo branco, sombras etc.)
color Display Background white
display depthcue off
display ambientocclusion on
display shadows on
display aoambient 0.8
display aodirect 0.2
```

ANEXO G - Script “ft.tcl”

```

### Open the log file for reading and the output .dat file for writing
set file [open CalB_SMD_NAMD.log r]
set output [open ft.dat w]

### Gather input from user.
puts "Enter a value for n_x:"
set nx [gets stdin]
puts "Enter a value for n_y:"
set ny [gets stdin]
puts "Enter a value for n_z:"
set nz [gets stdin]

### Loop over all lines of the log file
set file [open CalB_SMD_NAMD.log r]
while { [gets $file line] != -1 } {

    ### Determine if a line contains SMD output. If so, write the
    ### timestep followed by f(dot)n to the output file
    if {[lindex $line 0] == "SMD"} {

        puts $output "[expr 0.000001 * ([lindex $line 1] - 1090000)] [expr
$nx*[lindex $line 5] + $ny*[lindex $line 6] + $nz*[lindex $line 7]]"
    }
}

### Close the log file and the output .dat file
close $file
close $output
exit

```

ANEXO H - Código para cálculo da área de superfície acessível à água

Para o cálculo da área de superfície acessível à água do resíduo 249 da R249L, na janela do *Tk Console*, foi inserido o comando:

```
measure sasa 1.4 [atomselect top "resid 249"]
```

Para o cálculo da área de superfície acessível à solvente do resíduo 281 da A281E, na janela do *Tk Console*, foi inserido o comando:

```
measure sasa 1.4 [atomselect top "resid 281"]
```

ANEXO I - Códigos para a Identificação de resíduos vizinhos

Para saber as posições dos resíduos vizinhos, na janela do *Tk Console*, foram inseridos os comandos:

```
set glus [atomselect top "resname GLU"]  
  
$glus get {resid chain}
```

Depois, para saber quais eram os resíduos vizinhos, na janela do *Tk Console*, foram inseridos os comandos:

```
foreach r {81 188 269 294} {  
  
set sel [atomselect top "resid $r and chain P and name CA"]  
  
puts "Resid $r: [lindex [$sel get resname] 0]"  
  
}
```

ANEXO J - Script “ft_deslocamento.tcl”

```

### Open the log file for reading and the output .dat file for writing
set file [open CalB_SMD_NAMD.log r]
set output [open ft_deslocamento.dat w]

### Gather input from user.
puts "Enter a value for n_x:"
set nx [gets stdin]
puts "Enter a value for n_y:"
set ny [gets stdin]
puts "Enter a value for n_z:"
set nz [gets stdin]

### Loop over all lines of the log file
set file [open CalB_SMD_NAMD.log r]
while { [gets $file line] != -1 } {

    ### Determine if a line contains SMD output. If so, write the
    ### timestep followed by f(dot)n to the output file
    if {[lindex $line 0] == "SMD"} {

        #puts $output "[lindex $line 1] [expr $nx*[lindex $line 5] + $ny*[lindex
        $line 6] + $nz*[lindex $line 7]]"

        #puts $output "[[1.3663e-05*expr[lindex $line 1]-1090000]] [expr
        $nx*[lindex $line 5] + $ny*[lindex $line 6] + $nz*[lindex $line 7]]"

        puts $output "[expr 1.3663e-5 * ([lindex $line 1] - 1090000)] [expr
        $nx*[lindex $line 5] + $ny*[lindex $line 6] + $nz*[lindex $line 7]]"

    }

}

### Close the log file and the output .dat file
close $file
close $output
exit

```

ANEXO K - Script “trabalho.py”

```
import numpy as np

# Step 1: Read the file, removing the first line and lines where x > 750

def read_data(filename):
    x = []
    y = []

    # Read the file line by line
    with open(filename, 'r') as f:
        lines = f.readlines()[1:] # Skip the first line

        for line in lines:
            values = line.split()
            x_val = float(values[0])

            # Skip the line if x > 750
            if x_val > 750:
                continue

            # Add the x and y values to their respective lists
            x.append(x_val)
            y.append(float(values[1]))

    return np.array(x), np.array(y)

# Step 2: Apply Trapezoidal Rule

def calculate_area(x, y):
    # Calculate the area using the Trapezoidal Rule
    area = np.trapz(y, x)

    return area
```

```
# Step 3: Main process

def main():

    filename = 'ft_deslocamento.dat' # Your file name here

    # Read the data from the file
    x, y = read_data(filename)

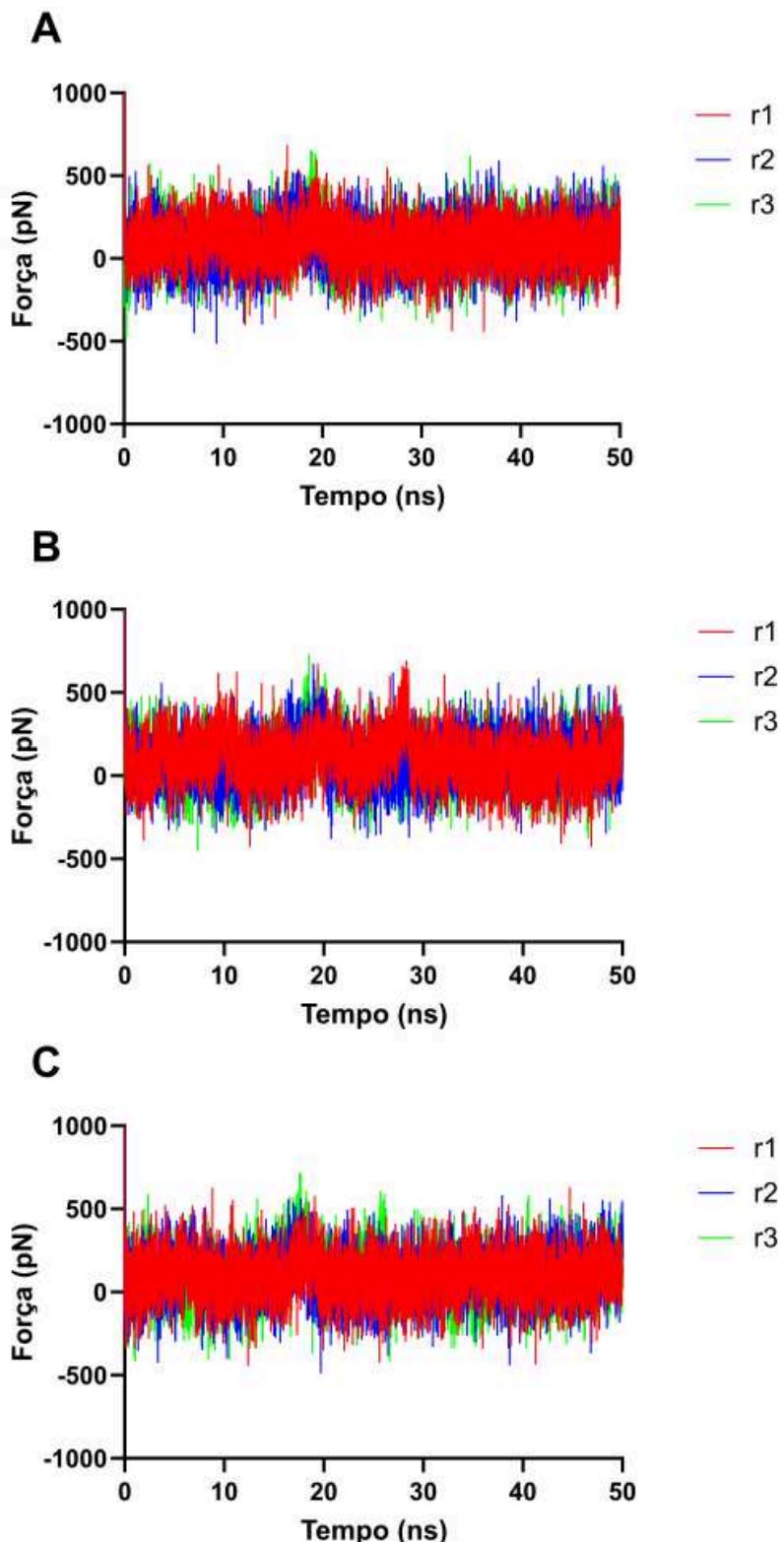
    #print (x)
    #print (y)

    # Calculate the area under the curve using the Trapezoidal Rule
    area = calculate_area(x, y)

    print("Area under the curve (Trapezoidal Rule):", area, "pN*A")
    print("Area under the curve (Trapezoidal Rule):", area*6.242e-4, "eV")
    print("Area under the curve (Trapezoidal Rule):", area*1.439e-3,
          "kcal/mol")

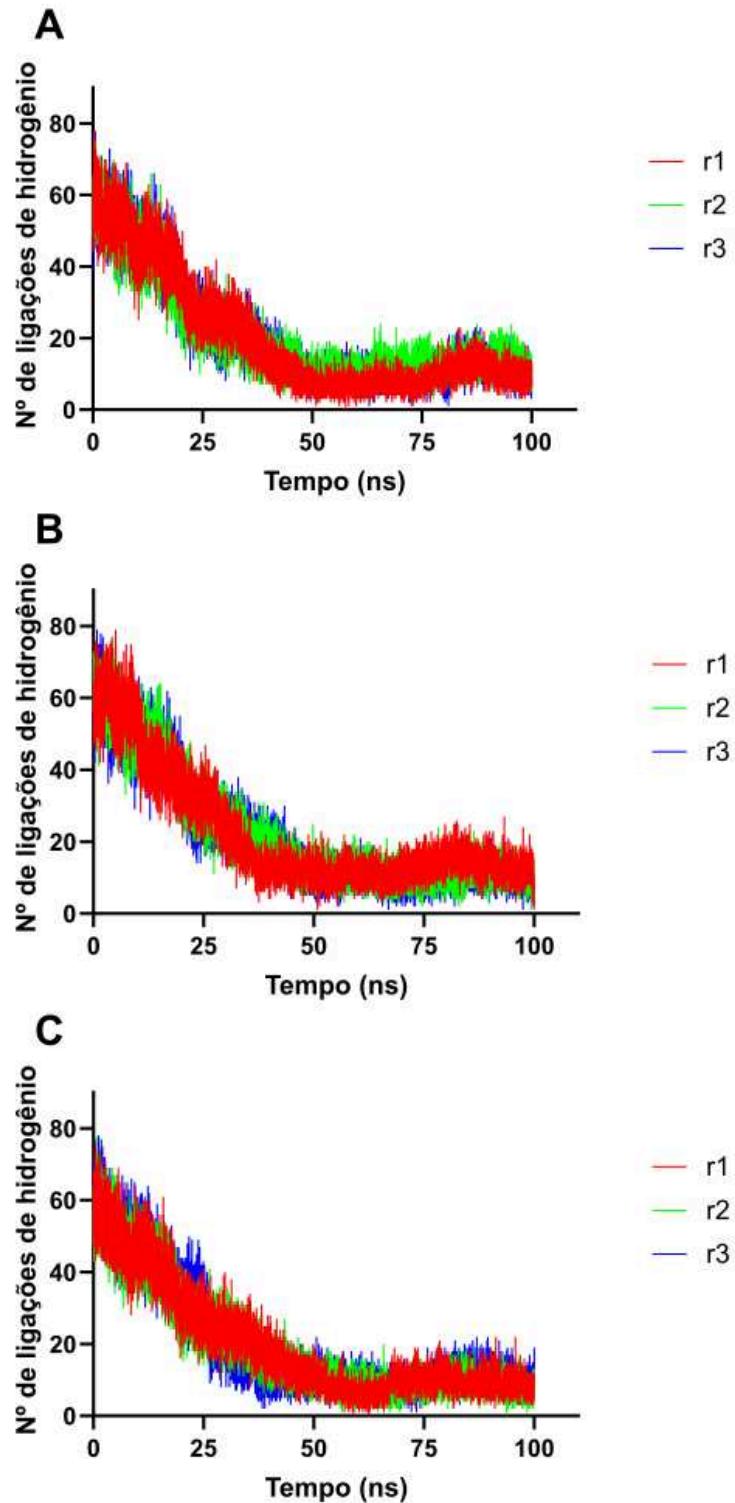
if __name__ == '__main__':
    main()
```

**ANEXO L - Gráficos com os dados brutos de força vs tempo. (A) CaIB_WT. (B)
A281E. (C) R249L.**



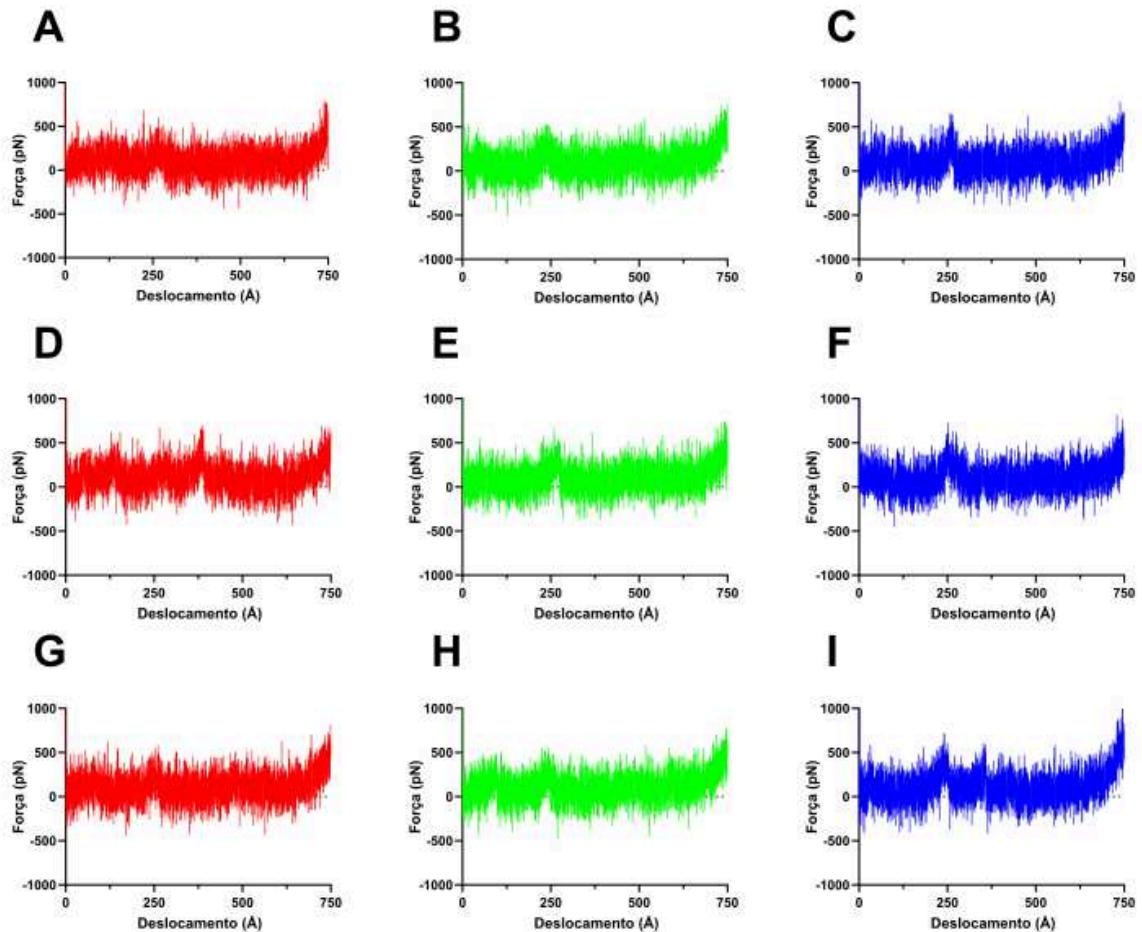
Fonte: Autor

ANEXO M - Gráficos com os dados brutos do número de ligações de hidrogênio vs tempo (ns). (A) CalB_WT. (B) A281E. (C) R249L.



Fonte: Autor

ANEXO N - Gráficos com os dados brutos de força vs deslocamento. (A) CalB_WT_r1. (B) CalB_WT_r2. (C) CalB_WT_r3. (D) A281E_r1. (E) A281E_r2. (F) A281E_r3. (G) R249L_r1. (H) R249L_r2. (I) R249L_r3.



Fonte: Autor