

# RESSALVA

Atendendo solicitação do autor,  
o texto completo desta dissertação  
será disponibilizado somente a partir  
de 15/12/2022



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Câmpus de São José do Rio Preto

---

Instituto de Biociências, Letras e Ciências Exatas

## Dissertação de Mestrado

Murilo Nogueira Sanches

Departamento de Física  
Instituto de Biociências, Letras e Ciências Exatas (IBILCE)  
Universidade Estadual Paulista (UNESP)  
São José do Rio Preto – SP

2021

**Murilo Nogueira Sanches**

**Representação detalhada do relevo de energia de monômeros de  
 $\beta$ -amilóide e de proteínas de repetições**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, câmpus de São José do Rio Preto.

Agência financiadora: CNPq – Proc. 130147/2020-6

Orientador: Prof. Dr. Vitor Barbanti Pereira Leite

São José do Rio Preto - SP

2021

S211r	<p>Sanches, Murilo Nogueira</p> <p>Representação detalhada do relevo de energia de -amiloide e de proteínas de repetições / Murilo Nogueira Sanches. -- São José do Rio Preto, 2021</p> <p>50 f.</p> <p>Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto</p> <p>Orientador: Vitor Barbanti Pereira Leite</p> <p>1. Superfícies de energia. 2. -amiloides. 3. Proteínas de Repetições.</p> <p>I. Título.</p>
-------	--

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

**Murilo Nogueira Sanches**

**Representação detalhada do relevo de energia de monômeros de  
 $\beta$ -amilóide e de proteínas de repetições**

**Comissão examinadora**

Prof. Dr. Vitor Barbanti Pereira Leite

Instituto de Biociência, Letras e Ciências Exatas, UNESP – São José do Rio Preto

Orientador

Profa. Dra. Thereza Amélia Soares da Silva

Instituto de Física, USP – São Paulo

Prof. Dr. Sidney Jurado de Carvalho

Instituto de Biociência, Letras e Ciências Exatas, UNESP – São José do Rio Preto

São José do Rio Preto - SP

15 de dezembro de 2021

2021

## Agradecimentos

Agradeço primeiramente à minha família, especialmente meus pais, que me proveram ótimas condições de estudo e sempre me auxiliaram incentivando meus estudos. Agradeço também à minha namorada Isabela Brunozi de Oliveira (Unesp), por sempre me apoiar e incentivar, por todo o carinho, zelo, paciência, compreensão, auxílio por sempre estar presente quando precisei, e tornando meus dias mais alegres.

Agradeço ao meu orientador Dr. Vitor Barbanti Pereira Leite (Unesp) por me aceitar em seu grupo de pesquisa e sempre auxiliar em meu projeto, assim como agradeço pelos ensinamentos transmitidos e pela confiança em mim depositada. Agradeço também a todo o grupo de pesquisa pela parceria e apoio.

Por fim, agradeço ao apoio do Instituto de Biociências, Letras e Ciências Exatas (IBILCE - Unesp) pela infraestrutura fornecida e agradeço ao CNPq que permitiu que o projeto fosse desenvolvido com todo suporte e apoio financeiro oferecido.

## Resumo

O processo pelo qual uma proteína assume sua configuração funcional é denominado enovelamento e corresponde a um problema que intriga os pesquisadores desde a década de 50. Devido à importância que essa classe de macromoléculas possui e às doenças que sua má atividade acarretam, compreender a formação de sua estrutura tridimensional é uma questão fundamental. Contudo, devido à sua complexidade, o processo ocorre em um espaço de fase multidimensional, de forma que sua análise não é trivial. Esse problema se torna ainda mais complicado devido ao fato de existirem classes especiais de proteínas que apresentam certas particularidades, como é o caso das proteínas intrinsecamente desordenadas e das proteínas de repetições. Enquanto a primeira classe apresenta alta flexibilidade e a ausência de um estado nativo de referência, a segunda classe corresponde à estruturas terciárias compostas pela repetição de um mesmo domínio, e que não enovelam de forma independente mas sim por meio de um mecanismo de cooperação. Neste trabalho foram analisados dois sistemas distintos. O primeiro consiste nas dinâmicas moleculares de monômeros de  $\beta$ -amilóides, visando compreender suas diferentes taxas de agregações. O segundo sistema corresponde às simulações de proteínas da família de repetições de anquirina, buscando a compreensão do mecanismo de formação dessas repetições. As análises foram realizadas por meio do *Energy Landscape Visualization Method* que, usando uma métrica baseada na distância interna entre os resíduos de cada conformação, é capaz de projetar a superfície de energia em espaços 2D e 3D. Os resultados demonstraram que a propensão de agregação das diferentes espécies de  $\beta$ -amilóide está contida no relevo de energia livre dos monômeros, enquanto que as repetições de anquirina apresentaram diferentes rotas de enovelamento com distintas probabilidades de ocorrência.

Palavras-chaves: Superfícies de energia;  $\beta$ -amilóide; proteínas de repetições.

# Abstract

The process by which a protein assumes its functional configuration is called folding and corresponds to a problem that has intrigued researchers since the 1950s. Due to the importance of this class of macromolecules and the diseases that their harmful activity entail, understanding the formation of its three-dimensional structure is a fundamental question. However, due to its complexity, the process takes place in a multidimensional phase space, so its analysis is not trivial. This problem becomes even more complicated due to the fact that there are special classes of proteins that have certain particularities, such as intrinsically disordered proteins and repeat proteins. While the first class presents high flexibility and the absence of a native state of reference, the second class corresponds to tertiary structures composed by the repetition of the same domain, and which does not fold independently but through a cooperation mechanism. In this work, two distinct systems were analyzed. The first consists of the molecular dynamics of  $\beta$ -amyloid monomers, aiming to understand their different aggregation rates. The second system corresponds to the simulations of proteins from the ankyrin repeat family, in order to understand the mechanism of formation of these repeats. The analyzes were performed using the Energy Landscape Visualization Method. This method uses a metric based on the internal distance between the residues of each conformation, and it is able to represent the ensemble of molecular conformations in an effective 2D and 3D phase space. The results show that the aggregation propensities of different species of amyloid- $\beta$  are encoded in the free energy spectrum of its monomers, while the ankyrin repetitions presented different folding routes with distinct occurrence probability.

Keywords: Energy landscape; amyloid- $\beta$ ; repeat proteins.



## Lista de Figuras

1	Esquema do funil de enovelamento. A coordenada de reação ou parâmetro de ordem, $Q$ , é correlacionada com a energia e indica o grau de similaridade dos estados com o estado nativo (Figura adaptada de [1]). . . . .	13
2	Representação da estrutura terciária das proteínas de repetições de anquirina, com a 3ANK correspondendo ao PDB 1N0QA, 4ANK ao PDB 1N0R e 6ANK ao PDB 1NFIE. . . . .	14
3	Exemplo de duas estruturas ( $k$ e $l$ ) onde são indicadas as distâncias internas ( $r_{i,j}^{k(l)}$ ) entre os aminoácidos (1, 7, 13, 33 e 60). As diferenças estruturais estão relacionadas com as variações destas distâncias entre todos os aminoácidos, de acordo com a Equação 1. . . . .	18
4	Espaço de fase conformacional das estruturas obtidas por dinâmica molecular visualizado através do ELViM com a) distinção das espécies presentes na projeção com A $\beta$ 40 em azul, A $\beta$ 42 em vermelho e as mutações da A $\beta$ 40 em verde; b) coloração em função do raio de giro, com as conformações significativas de cada região exibidas ao redor. . . . .	26
5	Projeção bidimensional do ELViM como função da coordenada $q$ com o monômero de fibra como referência (estrutura em vermelho, PDB 2LMN). A linha vermelha indica a divisão entre as regiões de alto e baixo $q_{fiber}$ na projeção. . . . .	27
6	(a) Perfil de energia livre em função do $q_{fiber}$ para a A $\beta$ 40 e A $\beta$ 42. (b) Perfil de energia livre em função do $q_{fiber}$ para os mutantes da A $\beta$ 40. . . . .	27
7	(a) Distribuição de energia livre das estruturas da A $\beta$ 40 (em cima) e da A $\beta$ 42 (em baixo) na projeção do ELViM, com a região de menor energia indicada pelo G em vermelho e a região de alto $q_{fiber}$ indicada pela elipse com o F em vermelho. (b) Distribuição média do $q_{fiber}$ da Figura 5. . . . .	29
8	Distribuição de energia livre das mutações de A $\beta$ 40 com o caminho de menor energia livre destacado em amarelo. . . . .	30
9	Diferença entre os mapas da média de contato para as regiões de menor energia livre com as regiões de alto $q_{fiber}$ para as estruturas de A $\beta$ 40, A $\beta$ 42 e E22Q. . . . .	30

10	Diferença entre os mapas da média de contato para as regiões de menor energia livre com as regiões de alto $q_{fiber}$ para as estruturas das mutações de A $\beta$ 40. . . . .	31
11	Análises das trajetórias enviesadas da anquirina com três repetições. A) Perfil de energia livre em função da coordenada $Qo$ . B) Grau de enovelamento de cada resíduo em função de $Qo$ com a cor indicando o $Qo$ médio local para um resíduo em relação ao conjunto de estruturas do $Qo$ global. C) Espaço de fase conformacional projetado através do ELViM em função de $Qo$ . D) Projeção tridimensional das trajetórias enviesadas da 3ANK em função da energia total da dinâmica com as estruturas não enviesadas destacadas em preto e o eixo Z correspondendo à coordenada $Qo$ . . . . .	33
12	Projeção bidimensional do espaço de fase conformacional da 3ANK em (A) função do raio de giro e (B) em função da coordenada $Qw$ . (C) Projeção destacando as estruturas não enviesadas em vermelho com as conformações enviesadas em cinza. . . . .	33
13	Mapa de contato médio para as estruturas da 3ANK que correspondem as diferentes regiões da projeção do ELViM (destacadas em vermelho). . . . .	34
14	Mapa de contato médio para os estados de transição da 3ANK com os mesmos valores de $Qo$ global. A diferença entre essas conformações combinada com o fato de que possuem valores similares de $Qo$ é a razão por trás do aparente <i>backtracking</i> observado na Figura 11B. . . . .	34
15	A) Perfil de energia livre em função da coordenada $Qo$ para a 4ANK. B) Grau de enovelamento de cada resíduo em função de $Qo$ com a cor indicando o $Qo$ médio local para um resíduo em relação ao conjunto de estruturas do $Qo$ global. C) Projeção bidimensional do espaço de fase conformacional da 4ANK em função da coordenada $Qo$ . D) Projeção 3D das trajetórias enviesadas em função da energia total da dinâmica com as trajetórias não enviesadas destacadas em preto. . . . .	36
16	Projeção bidimensional do espaço de fase conformacional da 4ANK em A) função do raio de giro e B) em função da coordenada $Qw$ . C) Projeção amostrando as estruturas obtidas com os dois tipos de simulações: as conformações não enviesadas em vermelho e as enviesadas em cinza. . . . .	37

17	Mapa de contato médio das estruturas de 4ANK correspondendo às diferentes regiões da projeção do ELViM (destacadas em vermelho). . . . .	37
18	Mapa de contato médio dos estados de transição da 4ANK com valores semelhantes de $Q_o$ global. A diferença entre essas conformações combinada com o fato de que apresentam valores próximos de $Q_o$ é o motivo por trás do aparente <i>backtracking</i> . . . . .	38
19	Taxa de enovelamento para a região I e região II da 4ANK para cada temperatura diferente. . . . .	39
20	Análises das trajetórias da 6ANK. A) Perfil de energia livre em função da coordenada $Q_o$ . B) Grau de enovelamento para cada resíduo em função de $Q_o$ , com a cor indicando a média do $Q_o$ local para um resíduo sobre o conjunto de estruturas de um dado $Q_o$ global. C) Espaço de fase conformacional da 6ANK visualizado através do ELViM em função de $Q_o$ . D) Projeção tridimensional das trajetórias em função da energia total da dinâmica, com o eixo Z correspondendo à coordenada $Q_o$ . . . . .	40
21	Mapa de contato médio das estruturas da 6ANK correspondendo às diferentes regiões da projeção do ELViM (destacadas em vermelho). . . . .	41
22	Mapa de contato médio dos estados de transição da 6ANK com mesmos valores de $Q_o$ global. A diferença entre essas conformações combinada com o fato de que apresentam valores próximos de $Q_o$ é o motivo por trás do aparente <i>backtracking</i> . . . . .	41
23	A) Grau de enovelamento de cada resíduo em função de $Q$ . B) Projeções bidimensionais para as trajetórias enviesadas da 3ANK em função de $Q^{Rn}$ para cada repetição. C) Projeções bidimensionais para as trajetórias não enviesadas da 3ANK em função de $Q^{Rn}$ para cada repetição. . . . .	42
24	A) Grau de enovelamento de cada resíduo em função de $Q_o$ . B) Projeções bidimensionais para as trajetórias enviesadas da 4ANK em função de $Q^{Rn}$ para cada repetição. C) Projeções bidimensionais para as trajetórias não enviesadas da 4ANK em função de $Q^{Rn}$ para cada repetição. . . . .	42
25	A) Grau de enovelamento de cada resíduo em função de $Q_o$ . B) Projeções bidimensionais para as trajetórias enviesadas da 6ANK em função de $Q^{Rn}$ para cada repetição. . . . .	43

# Sumário

<b>Trabalhos apresentados e publicações</b>	<b>12</b>
<b>1 Introdução</b>	<b>12</b>
1.1 $\beta$ -amilóides . . . . .	13
1.2 Proteínas de repetições . . . . .	14
<b>2 Objetivos</b>	<b>16</b>
<b>3 Metodologia</b>	<b>17</b>
3.1 ELViM: Energy Landscape Visualization Method . . . . .	17
3.2 Método dos Múltiplos Histogramas . . . . .	20
3.3 Modelo AWSEM . . . . .	20
3.3.1 Dinâmicas com $\beta$ -Amilóide . . . . .	22
3.3.2 Dinâmicas com Proteínas de repetições . . . . .	22
<b>4 Resultados e discussões</b>	<b>25</b>
4.1 $\beta$ -Amilóide . . . . .	25
4.2 Proteínas de repetições . . . . .	29
4.2.1 3 repetições . . . . .	30
4.2.2 4 repetições . . . . .	35
4.2.3 6 repetições . . . . .	37
<b>5 Conclusões</b>	<b>43</b>
<b>Referências</b>	<b>45</b>
<b>A Descrição da representação <i>coarse-grained</i></b>	<b>49</b>
<b>B Campo de força do AWSEM</b>	<b>49</b>

# Trabalhos apresentados e publicações

## Apresentações em painel

- "Comparação de potenciais não específicos com potenciais baseados em estrutura no estudo de enovelamento de proteínas" - 9<sup>a</sup> Escola de Modelagem Molecular em Sistemas Biológicos, Petrópolis de 20 a 24/08/2018.

- "Análise de enovelamento de proteínas por diferentes potenciais" - 1<sup>a</sup> fase do 30<sup>o</sup> Congresso de Iniciação Científica da UNESP, São José do Rio Preto de 30 a 31 de outubro de 2018.

- "Análise de enovelamento de proteínas por diferentes potenciais" - 2<sup>a</sup> fase do 30<sup>o</sup> Congresso de Iniciação Científica da UNESP, São Paulo de 21 a 22 de novembro de 2018.

- "Estudo energético das conformações de  $\beta$ -amilóide" - 1<sup>a</sup> fase do 31<sup>o</sup> Congresso de Iniciação Científica da UNESP, São José do Rio Preto, de 3 a 5 de setembro de 2019.

- "Estudo energético das conformações de  $\beta$ -amilóide" - 2<sup>a</sup> fase do 31<sup>o</sup> Congresso de Iniciação Científica da UNESP, Sorocaba, de 25 a 26 de novembro de 2019.

- "Probing propensities of amyloid- $\beta$  fiber formation through analysis of its energy landscape" - Encontro de Outono da Sociedade Brasileira de Física, de 21 a 25 de junho de 2021.

## Apresentações orais

- "Caracterização conformacional e visualização do perfil energético em estudos de proteínas intrinsecamente desordenadas" - Semana da Física, São José do Rio Preto, de 4 a 11 de novembro de 2019.

## Artigo publicado

B. da Silva, F.; M. de Oliveira, V.; Sanches, M. N.; Contessoto, V. G.; Leite, V. B. "Rational Design of Chymotrypsin Inhibitor 2 by Optimizing Non-Native Interactions" Journal of Chemical Information and Modeling 2019, 60, 982–988.

## Artigos submetidos

Sanches, M. N.; Knapp, K.; de Oliveira, A. B.; Wolynes, P. G.; Onuchic, J. N.; Leite, V. B. "Examining the Ensembles of Amyloid- $\beta$  Monomer Variants and their Propensities to Form Fibers Using an Energy Landscape Visualization Method" Journal of Physical Chemistry B.

Sanches, M. N.; Parra, R. G.; Viegas, R. G.; de Oliveira, A. B.; Wolynes, P. G.; Ferreira, D. U.; Leite, V. B.; "Resolving the Fine Structure in the Energy Landscapes of Repeat Proteins" Quarterly Reviews of Biophysics Discovery.

# 1 Introdução

Compreender os mecanismos biológicos é um dos principais objetivos e desafios da ciência moderna. Nesse intuito, a principal classe de macromoléculas a ser estudadas são as proteínas. Responsáveis por inúmeras funções essenciais para a manutenção da vida, grande parte de seus comportamentos permanece sem uma solução definitiva. O processo no qual a cadeia de aminoácidos assume uma estrutura funcional é denominado enovelamento, e é o principal alvo de estudos. Essas pesquisas procuram compreender como a informação presente em sua sequência de resíduos é capaz de guiar as proteínas ao seu estado nativo. Entender esse processo é de extrema importância, uma vez que diversas doenças estão diretamente relacionadas ao mal enovelamento de algumas proteínas, como a doença de Alzheimer [2]. Estudos revelaram que a superfície de energia dos estados assumidos pela proteína tem um papel fundamental para determinar o caminho até seu estado nativo [1]. Na década de 90, Onuchic propôs um modelo capaz de explicar o enovelamento de uma proteína por meio de sua superfície de energia [3]. Através de um mapeamento estatístico das possíveis conformações de uma proteína, a superfície apresenta o estado nativo como o mínimo global de energia, de forma que sua topologia seja afunilada, com múltiplos caminhos que levam a proteína ao estado estrutural nativo. Devido a esses caminhos, o funil apresenta rugosidades em sua superfícies, que correspondem a mínimos locais de energia, responsáveis por armadilhar essas macromoléculas em um estado parcialmente enovelado [4, 5].

Todavia, existem diversas famílias de proteínas que apresentam especificidades que tornam seu enovelamento de difícil compreensão. Seja por não possuírem um estado nativo ou por enovelarem através de interações específicas entre determinadas regiões, suas superfícies de energia são demasiadamente complexas. É nessa classe que se encontram dois tipos de moléculas de grande interesse: as proteínas intrinsecamente desordenadas (IDPs) e as proteínas de repetições. Enquanto que as IDPs não possuem um estado nativo ou de referência, as proteínas de repetições apresentam a união de domínios idênticos que se formam através de um processo cooperativo.

Para analisar essas superfícies de forma convencional, é necessário o uso de inúmeras coordenadas de reação, o que acaba por dificultar ou até mesmo inviabilizar os estudos. É nesse contexto que o *Energy Landscape Visualization Method* (ELViM) surge como uma alternativa que possibilita a análise e o entendimento do modelo estudado. Nesse trabalho, foram estudadas as mudanças conformacionais nos monômeros de  $\beta$ -amilóides e

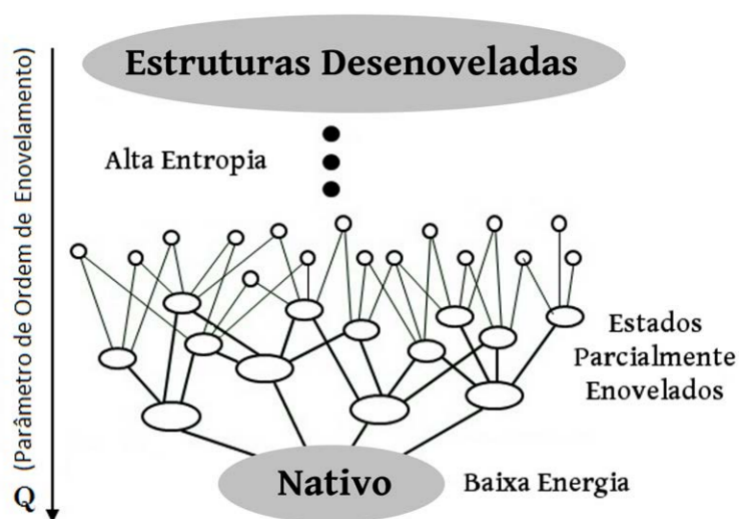


Figura 1: Esquema do funil de enovelamento. A coordenada de reação ou parâmetro de ordem,  $Q$ , é correlacionada com a energia e indica o grau de similaridade dos estados com o estado nativo (Figura adaptada de [1]).

o enovelamento das repetições de anquirina. Esse estudo foi realizado em colaboração com a *Rice University*, o *Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales*, e o *European Molecular Biology Laboratory*.

## 1.1 $\beta$ -amilóides

Nos últimos 20 anos, o entendimento do processo de agregação das proteínas que causam a doença de Alzheimer tem sido um grande foco de estudo. Essa doença possui uma relação direta com a proteína  $\beta$ -amilóide ( $A\beta$ ), uma IDP que surge de forma mais abundante como um polímero de 40 ou 42 resíduos através da clivagem da proteína amilóide precursora [6, 7]. A relação dessa IDP com a doença de Alzheimer reside na tendência de agregação, formando placas que podem bloquear a sinalização entre as células nas sinapses [2]. Contudo, estudos recentes sugerem que essa tendência pode estar contida no espectro energético das formas monoméricas [8]. Um dos indícios está no fato de que, apesar de diferirem em apenas dois aminoácidos, as duas sequências apresentam agregações distintas. Enquanto que a abundância natural da  $A\beta_{40}$  é de 10/1 em relação à  $A\beta_{42}$ , ela somente está presente em 1/3 das placas formadas, de modo que a agregação da  $A\beta_{42}$  ocorre de forma muito mais rápida e espontânea [9, 10]. Esse resultado corrobora com estudos anteriores que indicam que a taxa de agregação da  $A\beta_{42}$  está relacionada com o

conteúdo de folhas- $\beta$  presente em sua forma monomérica [11]. Junto a esse estudo também foram identificadas mutações na sequência de 40 aminoácidos que apresentam um aumento na taxa de agregação [12–15].

O ELViM foi utilizado nesse trabalho para analisar as mudanças conformacionais dos monômeros de A $\beta$ 40, A $\beta$ 42 e de seis mutações da A $\beta$ 40: A21G, D7N, D23N, E22G, E22K e E22Q, a fim de compreender suas taxas de agregação. Essa técnica de visualização da superfície de energia já se provou capaz de discernir essas superfícies em IDPs, como demonstrado em trabalho com a PAGE4 [16].

## 1.2 Proteínas de repetições

Em geral, um domínio proteico estrutural é um módulo relacionado à estabilidade e ao processo de enovelamento de uma proteína. Contudo, existe uma classe especial que apresenta repetições de um mesmo domínio que não enovelam de forma independente, mas sim através de uma cooperação que leva ao estado nativo por um caminho específico [17]. Dentre essas proteínas, destacam-se as chamadas de solenoides, compostas em geral por 20 a 40 trechos de aminoácidos semelhantes que se enovelam em estruturas alongadas [18]. A Figura 2 ilustra três exemplos de solenoides para a família de repetições de anquirina.



Figura 2: Representação da estrutura terciária das proteínas de repetições de anquirina, com a 3ANK correspondendo ao PDB 1N0QA, 4ANK ao PDB 1N0R e 6ANK ao PDB 1NFIE.

Uma das particularidades dessas proteínas é o fato de que essas matrizes são estabilizadas apenas por interações dentro de cada repetição ou entre repetições adjacentes, sem a interação entre os contatos de resíduos muito distantes na cadeia [19, 20]. Essa arquitetura contrasta com a típica de proteínas globulares, nos quais a cadeia se dobra para formar interações entre resíduos que se encontram distantes na sequência. Dessa



forma, o processo de enovelamento, bem como a superfície de energia, das proteínas de repetições difere do enovelamento de proteínas globulares, uma vez que as interações dos resíduos distantes em sequência frequentemente desempenham um papel central em definir a topologia da proteína. [21].

Nesse trabalho foi analisado o enovelamento das matrizes de repetições por meio do ELViM [22], utilizado para explorar as superfícies de energia da proteína anquirina de três, quatro e seis repetições (3ANK, 4ANK e 6ANK), assim como demonstrar a cooperação entre as diferentes repetições e a forma como essas proteínas varrem o espaço conformacional.

## 5 Conclusões

As superfícies de energia foram analisadas através do método ELViM para dois sistemas diferentes: A IDP  $\beta$ -amilóide e as repetições de anquirina, especificamente de três, quatro e seis repetições.

A natureza desordenada e a alta flutuação das IDPs resulta em uma superfície complexa de alta dimensionalidade, além do fato de que a falta de estrutura nativa também implica na falta de uma coordenada de reação adequada. Por meio do ELViM foi possível explorar as mudanças conformacionais nas formas monoméricas das  $A\beta$ , o que revelou que uma pré disposição para agregação possivelmente possui uma relação direta com o conteúdo de folhas- $\beta$  dos monômeros, uma vez que a  $A\beta_{42}$  e os mutantes, que possuem altas taxas de agregação, apresentam em suas rotas uma barreira de energia menor que a presente na  $A\beta_{40}$ . Além disso, a localização dos estados fundamentais nas projeções revelam que a maioria dos contatos necessários para iniciar a formação de fibras já estão formados na  $A\beta_{42}$  e nas mutações, enquanto que a  $A\beta_{40}$  precisa desfazer suas hélices primeiro.

Já os resultados para as proteínas de repetição confirmaram as expectativas de que suas superfícies são de fato afuniladas em direção ao estado nativo. Contudo, a

ocorrência de estados intermediários e a localização dos estados de transição são governadas pela topologia mais detalhada de cada proteína. Vemos também que, mesmo quando a topologia parece ser altamente simétrica, a população de rotas de enovelamento acaba sendo fortemente assimétrica. Todos os três sistemas parecem retroceder a formação de alguns elementos quando monitorados utilizando somente a média do  $Qo$  local, efeito que pode ser entendido como decorrente de serem estruturas diferentes de rotas de enovelamento paralelas que possuem o mesmo  $Qo$  global. É provável que à medida que essas matrizes de repetição crescem, a cooperatividade entre as repetições decai, conforme a energia das interfaces de tamanho finito deve ficar cada vez mais forte para superar a entropia de incorporar uma nova repetição [35]. Isso é observado para o maior sistema que foi analisado nesse trabalho, a  $I\kappa B\alpha$ , que tem uma região bastante plana da superfície que consiste em um subdomínio enovelado de cerca de 3 unidades de repetição consecutivas, o que é consistente com o conjunto estrutural proposto para esta proteína no estado livre não ligado e que está fortemente relacionado ao seu mecanismo funcional [36].

Através do *Energy Landscape Visualization Method* foi possível analisar as superfícies de energia de sistemas que até então apresentavam empecilhos em seu entendimento devido às suas naturezas. Os resultados obtidos confirmaram algumas hipóteses já existentes e revelaram novos mecanismos pelos quais esses sistemas atuam.

## Referências

- [1] Vinícius de Godoi Contessoto, Antonio Bento de Oliveira Junior, Jorge Chahine, Ronaldo Junio de Oliveira, and Vitor Barbanti Pereira Leite. Introdução ao problema de enovelamento de proteínas: uma abordagem utilizando modelos computacionais simplificados. *Revista Brasileira de Ensino de Física*, 40(4):e4307, 2018.
- [2] John Hardy and Dennis J Selkoe. The amyloid hypothesis of alzheimer’s disease: progress and problems on the road to therapeutics. *science*, 297(5580):353–356, 2002.
- [3] P. E. Leopold, M. Montal, and J. N. Onuchic. Protein folding funnels: a kinetic approach to the sequence-structure relationship. 89(18):8721–8725.
- [4] José Nelson Onuchic, Nicholas D Socci, Zaida Luthey-Schulten, and Peter G Wolynes. Protein folding funnels: the nature of the transition state ensemble. *Folding and Design*, 1(6):441–450, 1996.
- [5] José Nelson Onuchic, Zaida Luthey-Schulten, and Peter G Wolynes. Theory of protein folding: the energy landscape perspective. *Annual review of physical chemistry*, 48(1):545–600, 1997.
- [6] Cheick Ahmed Tidiane Sissoko. *Profiling tau hyperphosphorylation in formalin-fixed paraffin embedded human brain with liquid chromatography tandem mass spectrometry*. PhD thesis, Icahn School of Medicine at Mount Sinai, 2020.
- [7] Weihua Zheng, Min-Yeh Tsai, Mingchen Chen, and Peter G Wolynes. Exploring the aggregation free energy landscape of the amyloid- $\beta$  protein (1–40). *Proceedings of the National Academy of Sciences*, 113(42):11835–11840, 2016.
- [8] Debayan Chakraborty, John E Straub, and Dave Thirumalai. Differences in the free energies between the excited states of a $\beta$ 40 and a $\beta$ 42 monomers encode their aggregation propensities. *Proceedings of the National Academy of Sciences*, 117(33):19926–19937, 2020.
- [9] Weihua Zheng, Min-Yeh Tsai, and Peter G Wolynes. Comparing the aggregation free energy landscapes of amyloid beta (1–42) and amyloid beta (1–40). *Journal of the American Chemical Society*, 139(46):16666–16676, 2017.
- [10] Bogdan Barz, Qinghua Liao, and Birgit Strodel. Pathways of amyloid- $\beta$  aggregation depend on oligomer shape. *Journal of the American Chemical Society*, 140(1):319–327, 2018.
- [11] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.
- [12] Simona Tomaselli, Veronica Esposito, Paolo Vangone, Nico AJ van Nuland, Alexandre MJJ Bonvin, Remo Guerrini, Teodorico Tancredi, Piero A Temussi, and Delia Picone. The  $\alpha$ -to- $\beta$  conformational transition of alzheimer’s a $\beta$ -(1–42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of  $\beta$  conformation seeding. *ChemBioChem*, 7(2):257–267, 2006.

- [13] Subramanian Vivekanandan, Jeffrey R Brender, Shirley Y Lee, and Ayyalusamy Ramamoorthy. A partially folded structure of amyloid-beta (1–40) in an aqueous environment. *Biochemical and biophysical research communications*, 411(2):312–316, 2011.
- [14] Aida Attar, Derya Meral, Brigita Urbanc, and Gal Bitan. Assembly of amyloid  $\beta$ -protein variants containing familial alzheimer’s disease-linked amino acid substitutions. In *Bio-nanoimaging*, pages 429–442. Elsevier, 2014.
- [15] Asa Hatami, Sanaz Monjazebe, Saskia Milton, and Charles G Glabe. Familial alzheimer’s disease mutations within the amyloid precursor protein alter the aggregation and conformation of the amyloid- $\beta$  peptide. *Journal of Biological Chemistry*, 292(8):3172–3185, 2017.
- [16] Antonio B Oliveira Junior, Xingcheng Lin, Prakash Kulkarni, José N Onuchic, Susmita Roy, and Vitor BP Leite. Exploring energy landscapes of intrinsically disordered proteins: Insights into functional mechanisms. *Journal of Chemical Theory and Computation*, 2021.
- [17] Lisanna Paladin, Martina Bevilacqua, Sara Errigo, Damiano Piovesan, Ivan Mičetić, Marco Necci, Alexander Miguel Monzon, Maria Laura Fabre, Jose Luis Lopez, Juliet F Nilsson, Javier Rios, Pablo Lorenzano Menna, Maia Cabrera, Martin Gonzalez Buitron, Mariane Gonçalves Kulik, Sebastian Fernandez-Alberti, Maria Silvina Fornasari, Gustavo Parisi, Antonio Lagares, Layla Hirsh, Miguel A Andrade-Navarro, Andrey V Kajava, and Silvio C E Tosatto. RepeatsDB in 2021: improved data and extended classification for protein tandem repeat structures. *Nucleic Acids Research*, 49(D1):D452–D457, 11 2020.
- [18] R Gonzalo Parra, Rocío Espada, Nina Verstraete, and Diego U Ferreiro. Structural and energetic characterization of the ankyrin repeat protein family. *PLoS computational biology*, 11(12):e1004659, 2015.
- [19] R Gonzalo Parra, Rocío Espada, Ignacio E Sánchez, Manfred J Sippl, and Diego U Ferreiro. Detecting repetitions and periodicities in proteins by tiling the structural space. *The Journal of Physical Chemistry B*, 117(42):12887–12897, 2013.
- [20] Rocío Espada, R. Gonzalo Parra, Manfred J. Sippl, Thierry Mora, Aleksandra M. Walczak, and Diego U. Ferreiro. Repeat proteins challenge the concept of structural domains. *Biochemical Society Transactions*, 43:844–849, 2015.
- [21] Diego U. Ferreiro, Aleksandra M. Walczak, Elizabeth A. Komives, and Peter G. Wolynes. The energy landscapes of repeat-containing proteins: Topology, cooperativity, and the folding funnels of one-dimensional architectures. *PLOS Computational Biology*, 4(5):1–13, 05 2008.
- [22] Antonio B. Oliveira, Huan Yang, Paul C. Whitford, and Vitor B. P. Leite. Distinguishing biomolecular pathways and metastable states. *Journal of Chemical Theory and Computation*, 15(11):6482–6490, 2019. PMID: 31618581.

- [23] Antonio B Oliveira Jr, Huan Yang, Paul C Whitford, and Vitor BP Leite. Distinguishing biomolecular pathways and metastable states. *Journal of chemical theory and computation*, 15(11):6482–6490, 2019.
- [24] Corey Hardin, Michael P Eastwood, Zaida Luthey-Schulten, and Peter G Wolynes. Associative memory hamiltonians for structure prediction without homology: alpha-helical proteins. *Proceedings of the National Academy of Sciences*, 97(26):14235–14240, 2000.
- [25] Eduardo Tejada, Rosane Minghim, and Luis Gustavo Nonato. On improved projection techniques to support visual exploration of multi-dimensional data sets. *Information Visualization*, 2(4):218–231, 2003.
- [26] Alan M. Ferrenberg and Robert H. Swendsen. New monte carlo technique for studying phase transitions. 61(23):2635.
- [27] Shankar Kumar, John M Rosenberg, Djamal Bouzida, Robert H Swendsen, and Peter A Kollman. THE weighted histogram analysis method for free energy calculations on biomolecules. i. the method. 13(8):1011–1021.
- [28] Aram Davtyan, Nicholas P Schafer, Weihua Zheng, Cecilia Clementi, Peter G Wolynes, and Garegin A Papoian. Awsem-md: protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing. *The Journal of Physical Chemistry B*, 116(29):8494–8503, 2012.
- [29] Michael P Eastwood and Peter G Wolynes. Role of explicitly cooperative interactions in protein folding funnels: a simulation study. *The Journal of Chemical Physics*, 114(10):4702–4716, 2001.
- [30] Robert B. Best, Gerhard Hummer, and William A. Eaton. Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences*, 110(44):17874–17879, 2013.
- [31] Mona Habibi, Jörg Rottler, and Steven S Plotkin. As simple as possible, but not simpler: exploring the fidelity of coarse-grained protein models for simulated force spectroscopy. *PLoS computational biology*, 12(11):e1005211, 2016.
- [32] Sébastien Côté, Philippe Derreumaux, and Normand Mousseau. Distinct morphologies for amyloid beta protein monomer: A $\beta$ 1–40, a $\beta$ 1–42, and a $\beta$ 1–40 (d23n). *Journal of chemical theory and computation*, 7(8):2584–2592, 2011.
- [33] Samuel S. Cho, Yaakov Levy, and Peter G. Wolynes. P versus q: Structural reaction coordinates capture protein folding on smooth landscapes. *Proceedings of the National Academy of Sciences*, 103(3):586–591, 2006.
- [34] Diego U Ferreiro and Elizabeth A Komives. Molecular mechanisms of system control of nf- $\kappa$ b signaling by  $\kappa$ b $\alpha$ . *Biochemistry*, 49(8):1560–1567, 2010.
- [35] Ezequiel A Galpern, María I Freiberger, and Diego U Ferreiro. Large ankyrin repeat proteins are formed with similar and energetically favorable units. *Plos one*, 15(6):e0233865, 2020.

- [36] Stephanie ME Truhlar, Erika Mathes, Carla F Cervantes, Gourisankar Ghosh, and Elizabeth A Komives. Pre-folding  $\text{I}\kappa\text{B}\alpha$  alters control of  $\text{NF-}\kappa\text{B}$  signaling. *Journal of molecular biology*, 380(1):67–82, 2008.