

UNIVERSIDADE DE SÃO PAULO
INSITUTO DE FÍSICA DE SÃO CARLOS

AMANDA SOUZA CÂMARA

Uma transição assimétrica entre estados simétricos:
o alosterismo da Glucosamina 6-fosfato Desaminase

São Carlos
2013

AMANDA SOUZA CÂMARA

Uma transição assimétrica entre estados simétricos:
o alosterismo da Glucosamina 6-fosfato Desaminase

Dissertação apresentada ao
Programa de Pós- Graduação em
Física do Instituto de Física de São
Carlos da Universidade de São Paulo
para obtenção do título de Mestre em
Ciências.

Área de concentração: Física
Aplicada
Opção: Física Biomolecular
Orientador: Prof. Dr. Eduardo
Horjales Reboredo

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos
2013

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Souza Câmara, Amanda

Uma transição assimétrica entre estados simétricos: o alosterismo da Glucosamina 6-fosfato Desaminase / Amanda Souza Câmara; orientador Eduardo Horjales Reboredo - versão corrigida -- São Carlos, 2013.

118 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2013.

1. Alosterismo. 2. Simetria. 3. Dinâmica molecular. 4. Análise de modos normais. 5. Tensor de inércia. I. Horjales Reboredo, Eduardo, orient. II. Título.

A meus pais e minha irmã.

AGRADECIMENTOS

Ao Prof. Dr. Eduardo Horjales parece-me pouco agradecer. É uma honra participar deste projeto, um privilégio tentar contribuir em suas discussões científicas e uma experiência única dividir a sala com você! Obrigada pela oportunidade, pelos conselhos acadêmicos e pessoais, pelas histórias, pela dedicação e pela confiança! Um grande abraço!

Ao meu grupo de trabalho, Michel, Valquiria e David, além dos que por aqui passaram, Maycou, Heline e Francesco, pela companhia no dia-a-dia e por me ajudarem a decifrar o que chefe queria!

Aos amigos da cristal, pela companhia, pelos “bom dia”s e por sanarem minhas dúvidas de novata.

Aos técnicos, meus exemplos no laboratório, Humberto, Susana, Bel e Andressa pela paciência e consideração!

Aos professores Claudia Munte e Rinaldo Montalvão por toda a atenção e paciência com minhas dúvidas.

Agradeço aos meus amigos companheiros de simulação: Ricardo; Preta; Bachega, por todo esse brilho; Militar, pelas conversas produtivas; Crusca, por confiar no meu trabalho; Micha, pelos insights; e Sergio por tudo, você já é um ótimo professor!

Às meninas da biblioteca por serem tão prestativas!

Ao pessoal da pós-graduação, Ricardo, Patricia e Silvio por me guiarem nos trâmites da burocracia com tanta eficiência e cuidado.

Ao pessoal da secretaria da cristal, por permitirem que eu desenvolvesse esse trabalho com tranquilidade.

À equipe da Scinfor, em especial ao Gilmar, que cuidou de todas as ferramentas que precisei durante este trabalho.

A todo o corpo docente e funcionários do IFSC que possibilitaram um ótimo ambiente de trabalho.

Agradeço a Capes pelo apoio financeiro para realizar este projeto de mestrado.

Aos amigos da biofísica, Paty, Val, Ana, Militar, Joci, Crusca, Débs, Luis, Sabrina, Jú, Helton, Fábio, Fernando, Andressa, Bel, Júlio, Ana Isabel, Sussu, Zé, Bisonho, Camilo, Rodrigo e Mexicano por me acolherem.

À Paty, por todo o cuidado enquanto morou comigo. Você tem sido praticamente uma família inteira pra mim! =)

Aos amigos da graduação que continuaram por perto Preta, Esperto, Lesa, Gizmo, Carpete, Paraíba e Gringo.

Aos amigos da escola que me conhecem há tanto tempo, por nunca me deixarem ir embora quando eu achava que precisava trabalhar... À Luda e à Linosa por me acompanharem sempre!

À minha professora de piano, Simone, por me apresentar essa rota de fuga.

Agradeço a meus pais por serem pacientes e compreensivos, mesmo quando aflitos com meus caminhos tão tortos. E a minha irmã por acreditar tão sinceramente que eu mereça tantas coisas boas! A vocês, minhas desculpas pela ausência. Espero que ela se justifique com este trabalho.

Agradeço a Fufus pelo apoio financeiro imediato e incondicional!

Agradecer é um alívio ao coração, mas uma tortura a minha memória! Espero ter-me mostrado grata, quando o momento exigia, mais que esses agradecimentos possam demonstrar.

*“O olhar que assusta anda morto
O olhar que avisa anda aceso
Mas quando eu chego eu me perco
Nas tramas do teu segredo”*

Dorival Caymmi e Paulo César Coelho - Desenredo

RESUMO

CÂMARA, A.S. **Uma transição assimétrica entre estados simétricos: o alosterismo da Glucosamina 6-fosfato Desaminase**. 2013. 118p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Sistemas alostéricos são característicos de proteínas com um ou mais estados de equilíbrio. Nesse sentido, uma enzima passa por modificações de sua atividade quando um substrato cooperativo se liga a um estado ou outro (1). Estes estados são reconhecidos por possuírem uma conformação mais estável e coexistirem num ensemble. Este trabalho sustenta que tais proteínas oscilem naturalmente entre esses estados. Experimentos de difração de raios-X e RMN, que proporcionam parâmetros de deslocamento anisotrópicos e tempos de relaxação de spin nuclear, já demonstram a coexistência de ambos estados em solução e descrevem o movimento como uma mudança de equilíbrio populacional dos confórmeros (2). Também é possível desenvolver métodos numéricos, como o cálculo de modos normais e a simulação de dinâmica molecular, para associar a geometria proteica a um movimento sobre determinado potencial de campos de força. O sistema adotado para o desenvolvimento desses estudos é a enzima alostérica Glucosamina-6-fosfato Desaminase. Características que defendem seu uso são sua reversibilidade catalítica, rápido equilíbrio cinético e muito baixa afinidade do estado T por ligantes. Sua estrutura também já foi resolvida por experimentos de cristalografia, identificando ambos estados alostéricos. E a caracterização das mudanças estruturais entre os estados T e R está bem estabelecida, identificando diferentes subunidades a distintos graus de rotação e prevendo uma oscilação de baixa frequência entre eles (3). Resultados obtidos neste projeto constituem: (a) uma dinâmica de 100ns partindo do estado T de toda a proteína (hexamérica) solvatada explicitamente, formando um ensemble NVT de 92000 átomos através do programa NAMD, usando o campo de forças CHARMM; (b) análise de componentes principais aproveitando esta dinâmica e usando algoritmos do programa Gromacs; (c) e análise de modos normais, em que os cálculos de minimização de energia foram feitos pelo programa Gromacs sob o campo de forças ENCADV, no vácuo. Análises desses resultados envolvem cálculos de RMSDs e flutuações, trajetórias calculadas para os autovetores oriundos de NMA ou de PCA, fatores de Debye-Waller e a confirmação visual (e gráficos de distância entre resíduos) de aproximação a um estado ou outro. Como a prévia caracterização da movimentação alostérica, identificava duas regiões para cada monômero como representativas de corpos rígidos, também é desenvolvida uma análise por tensores de

inércia. Espera-se que, ao longo do tempo, essas subunidades se comportem como corpos quase rígidos e os movimentos destas regiões rígidas correspondam a uma maior representatividade da transição alostérica. Assim, a caracterização dos tensores seria capaz de filtrar movimentos de mais alta frequência que constituem ruído em relação a movimentos funcionais da proteína. - Algoritmos para cálculos matriciais dos tensores foram escritos em Fortran e em TCL.

Palavras-chave: Alostérico. Simetria. Dinâmica molecular. Análise de modos normais. Tensor de inércia.

ABSTRACT

CÂMARA, A.S. **An asymmetric transition between symmetric states: the Glucosamine 6-phosphate Deaminase allostery.** 2013 118p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Allosteric systems are characteristic of proteins with one or more equilibrium states. Such an enzyme experiences a modification of its activity when a cooperative substrate binds to a state or another, thus, establishing a change in population equilibrium (1). These states are recognized by having a more stable conformation, and they coexist in an ensemble. X-ray diffraction and NMR experiments already demonstrated this dynamic equilibrium, and simulation methods, as molecular dynamics and normal mode analysis, generally provide a more complete proof (2). The allosteric enzyme Glucosamine-6-Phosphate Deaminase appeared to be a good model to better understand the equilibrium dynamics as essential to the protein function, given its reversibility of the catalysis and rapid-equilibrium kinetic mechanism. It also has the structure elucidated for both its conformers (3). A computational approach would now give better perspective on how the conformational changes occur. A set of results of this latter kind were obtained: (a) a 100ns dynamic starting at the hexameric T conformer, explicitly solvated, building a NVT ensemble using NAMD program and CHARMM force field; (b) a principal components analysis making use of the calculated dynamic and of the Gromacs algorithms; (c) and normal mode analysis of the T conformer structure (pdb code 1fsf) minimized with Gromacs program using ENCADV vacuum force field. Not only the conventional analyses for these results (fluctuations and projections) were taken, but also an inertia tensor analysis was developed. As the allosteric conformational change, for this protein, was described by the displacement of only two rigid body subunits³, its description by inertia tensors should act as a filter for the high frequency and functionally uninteresting motions, which normally constitute only noise.

Keywords: Allostery. Symmetry. Molecular dynamics. Normal mode analysis. Inertia tensor.

LISTA DE FIGURAS

Figura 1.1	Superfície eletrônica da enzima Glucosamina 6-fosfato Desaminase (Gln6PD) e esquema da reação catalisada por ela. ..	25
Figura 3.1	Esquema de distribuição de cargas parciais usada para a parametrização do ligante H_1PO_4	41
Figura 3.2	Representação das divergências entre alinhamento local e global. ..	46
Figura 3.3	Representação da movimentação alostérica sugerida pela sobreposição dos estados cristalográficos.	47
Figura 3.4	Sítio alostérico formado entre dois monômeros do estado T (esquerda) e do estado R (direita), resolvidos por cristalografia. ..	49
Figura 3.5	Trímero das estruturas cristalográficas dos estados T (preto) e R (verde) alinhados pela região interna.	50
Figura 3.6	Análise da movimentação de encolhimento e alongamento de uma hélice α	52
Figura 3.7	Representação da projeção da variação dos autovetores do tensor de inércia em função do tempo.	53
Figura 3.8	Gráficos das variações em relação à média de cada um dos autovetores da hélice α em movimento de encolhimento e alongamento projetadas no plano perpendicular ao vetor médio. .	54
Figura 3.9	Análise dos tensores de inércia para os estados T (preto) e R (verde) cristalográficos.	54
Figura 3.10	Representação de como a deformação do corpo altera os tensores de inércia.	55
Figura 4.1	Superposição de trímeros das conformações T, em preto, e R, em verde, alinhadas estruturalmente pela região interna.	59
Figura 4.2	Representação de um monômero da GlcN6PD com algumas regiões destacadas.	62
Figura 4.3	Superposição de quadros do movimento calculado como primeira componente principal da transição entre as conformações R e T.	62
Figura 4.4	Seis gráficos que mostram o deslocamento angular do centro de massa durante a equilibração da dinâmica 1.	64
Figura 4.5	Representação dos parâmetros de deslocamento anisotrópicos para cada carbono α como um elipsoide. Valores resolvidos por	

Figura 4.6	Parâmetros de deslocamento anisotrópicos representados em elipsoides para cada carbono α . Calculados como uma média do deslocamento durante a trajetória da dinâmica 3.	69
Figura 4.7	Seis gráficos, que mostram o deslocamento angular do centro de massa. Dados calculados para a trajetória da dinâmica 3.	70
Figura 4.8	Seis gráficos que mostram o comprimento da ligação entre o nitrogênio da metionina N terminal e o carbono β da arginina 158, e entre o mesmo nitrogênio da metionina e o carbono β da lisina 160. Cálculo feito para a dinâmica 3.	72
Figuras 4.9/4.10	Representações dos tensores de inércia em função do tempo para o protômero A.	74/75
Figuras 4.11/4.12	Representações dos tensores de inércia em função do tempo para o protômero B.	76/77
Figuras 4.13/4.14	Representações dos tensores de inércia em função do tempo para o protômero C.	78/79
Figuras 4.15/4.16	Representações dos tensores de inércia em função do tempo para o protômero D.	80/81
Figuras 4.17/4.18	Representações dos tensores de inércia em função do tempo para o protômero E.	82/83
Figuras 4.19/4.20	Representações dos tensores de inércia em função do tempo para o protômero F.	84/85
Figura 4.21	Representação do movimento do primeiro modo normal (modo 7) encontrado.	90
Figura 4.22	Representação do movimento encontrado para o primeiro modo normal (modo 7) para a cadeia B em destaque.	90
Figura 4.23	Resultados oriundos do uso do programa bio3d para cálculos de Análise de Componentes Principais. Flutuações médias calculadas para cada resíduo do hexâmero completo, considerando somente carbonos α , para a trajetória da dinâmica 3.	93
Figura 4.24	Resultados oriundos do uso do programa bio3d para cálculos de Análise de Componentes Principais. Os gráficos mostram a contribuição das primeiras 20 componentes na descrição da trajetória.	95
Figura 4.25	Representação do movimento da primeira componente principal encontrada, quando considerado cada monômero	

	individualmente.	96
Figura 4.26	Seis gráficos, um para cada cadeia do hexâmero, que mostram o deslocamento angular do centro de massa. Dados calculados para a trajetória da dinâmica 2.	99
Figura 4.27	Seis gráficos que mostram o comprimento da ligação entre o nitrogênio da metionina N terminal e o carbono β da arginina 158 em vermelho, e entre o mesmo nitrogênio da metionina e o carbono β da lisina 160 em preto. Cálculo feito para a dinâmica F.	100

LISTA DE TABELAS E GRÁFICOS

Tabela 3.1	Alguns parâmetros de interações eletrostáticas e do método numérico de busca pelo vizinho mais próximo usados em duas etapas diferentes da minimização.	37
Tabela 3.2	Estado de protonação das histidinas.	39
Tabela 3.3	Relação de dissimilaridades entre as três dinâmicas realizadas.	40
Tabela 3.4	Comprimentos de ligação entre três resíduos chaves na formação do sítio alostérico, calculadas sempre entre o Nitrogênio terminal da Metionina e os carbonos beta da Arginina 158 ou da Lisina 160. Medidas tomadas em um instante conveniente da dinâmica 1.	41
Tabela 3.5	Distâncias entre o N terminal da Metionina e o Carbono beta da Arginina 158 ou da Lisina 160 para os dois estados cristalográficos T e R, de código 1fsf e 1dea, respectivamente.	49
Tabela 3.6	Lista de átomos e suas respectivas massas atômicas usadas para o cálculo de tensores de inércia.	57
Tabela 3.7	Formatação do arquivo que descreve os tensores de inércia.	58
Tabela 4.1	Lista de resíduos dos confórmeros T e R que se alinham, ou seja, cujos carbonos α ficaram a uma distância listada na terceira coluna.	60
Tabela 4.2	Energia potencial, força máxima encontrada para um átomo e força média encontrada para o <i>ensemble</i> de átomos no último instante das minimizações.	88
Tabela 4.3	Autovalores de ambas as minimizações para os dez primeiros modos normais encontrados.	88
Tabela 4.4	Autovalores e números de onda calculados pelo programa Gromacs com a rotina <i>g_nmeig</i> , autofrequência e período calculados manualmente correspondentes para os dez primeiros modos normais encontrados, para a minimização 2.	89
Gráfico 3.1	Variação em porcentagem dos autovetores de uma hélice α , durante o movimento de encolhimento e alongamento.	52
Gráfico 4.1	Flutuação média de cada resíduo do hexâmero, para o movimento da primeira componente principal calculada para a transição entre as conformações R e T.	61
Gráfico 4.2	Parâmetros de deslocamento atômico isotrópicos, representando cada protômero na trajetória escrita para a transição direta entre as conformações T e R.	63
Gráfico 4.3	Energias total e cinética do sistema durante a equilibração da dinâmica 1.	65
Gráfico 4.4	Energias total e cinética do sistema durante a equilibração da dinâmica 3.	65

Gráfico 4.5	Energias total e cinética do sistema durante a dinâmica 3.	66
Gráfico 4.6	Energia potencial do sistema durante a dinâmica 3.	66
Gráfico 4.7	Parâmetros de deslocamento atômico isotrópicos representando cada protômero na trajetória escrita para a dinâmica 3.	67
Gráfico 4.8	Parâmetros de deslocamento atômico isotrópicos representando cada protômero na trajetória calculada para o primeiro modo normal.	91
Gráfico 4.9	Contribuição das primeiras 20 componentes na descrição da dinâmica 3. ..	92
Gráfico 4.10	Parâmetros de deslocamento atômico isotrópicos representando cada protômero na trajetória calculada para a primeira componente principal. ..	94
Gráfico 4.11	Energias total e cinética do sistema durante a dinâmica 2.	98
Gráfico 4.12	Pressão calculada para o sistema durante a dinâmica 2.	98

SUMÁRIO

1. Introdução	23
2. Objetivos	31
3. Métodos de simulação computacional para o estudo da transição alostérica	33
3.1 <i>Análise de Modos Normais</i>	33
3.2 <i>Dinâmica molecular</i>	38
3.3 <i>Dinâmica molecular com ligante</i>	40
3.4 <i>Técnicas de análise das dinâmicas</i>	42
3.4.1 <i>Análise de Componentes Principais</i>	42
3.4.2 <i>Fatores de temperatura ou parâmetros de deslocamento atômico</i>	44
.....	
3.4.3 <i>Alinhamento e análise de corpos rígidos</i>	45
3.4.4 <i>Existência do sítio alostérico</i>	48
3.4.5 <i>Deslocamento do centro de massa</i>	50
3.4.6 <i>Tensores de inércia</i>	51
3.4.7 <i>Correlação de oscilações harmônicas</i>	56
3.4.8 <i>Algoritmos para análises estruturais</i>	56
4. Resultados e Discussões	59
5. Mais discussões e Conclusões	103
6. Perspectivas	111
Referências	113

1 Introdução

“(…) *DNA has told us about the molecular basis of life, but in fact DNA's and most RNA's are chemically inert, whereas proteins are the workhorses of the living cell.*” (4). E a funcionalidade delas está estritamente ligada a sua estrutura. Com apenas 20 aminoácidos elas se organizam em variados arranjos sequenciais e espaciais para se dividir em classes de trabalhadores que vem a ser reguladores, bombas, motores, receptores, tradutores, armazenadores, transportadores e inúmeros outros (5). A versatilidade de funções que elas possuem e a intrincada rede de reações bioquímicas em que atuam são tamanhas que se faz necessária fina administração e regulação das atividades.

Uma das formas de regular a funcionalidade de proteínas é pelo alosterismo, cuja teoria foi formalizada por Monod *et al.* em 1965 (1). Ela estabelece que para proteínas com duas ou mais estruturas em equilíbrio pode haver um ligante cooperativo ou metabólito que altere a atividade, mas que não possua relação estereoquímica com o substrato ou o produto (6). Em outras palavras, deve haver um segundo ligante que tenha afinidade por outro sítio, diferente do ativo, o alostérico. E esse ligante alostérico reestabeleceria novo equilíbrio entre as diferentes conformações que a proteína possa adotar. Como essas duas conformações teriam afinidades diferentes pelo ligante natural (não o alostérico), quando uma estrutura passa a ser mais presente que a outra numa solução a funcionalidade da proteína é modificada. Ou seja, é a mudança de equilíbrio populacional entre essas conformações que regula a função a ser exercida pela proteína.

Em termos de arranjo espacial, essas proteínas são constituídas por subunidades idênticas dispostas geometricamente de forma a manter uma simetria, estabelecendo oligômeros*. A diferença estrutural entre as duas conformações estaria nas interações entre protômeros. Em uma, essas interações seriam tais que não permitiriam uma ligação forte do substrato, chamada de conformação T, enquanto a outra, conformação R, já não estaria constricta dessa maneira (1).

Monod ainda prediz, nesta teoria, que durante a transição de um estado a outro, de uma conformação a outra, as interações interprotoméricas seriam refeitas também de maneira simétrica, constituindo um movimento concertado – modelo MWC ou de simetria concertada. Já Koshland em 1966 propõe um modelo de transição sequencial, em que cada subunidade é livre para mudar sua estrutura terciária, do estado T para o R, com a ligação do substrato e então

alterar a atividade das subunidades vizinhas – modelo KNF ou de interação sequencial (7). Mas o modelo de simetria e movimento concertado pode se apoiar no teorema que diz que qualquer movimento de um corpo pode ser descrito por uma rotação e uma translação ao longo de um eixo conveniente (4, 8-9). Já existindo um eixo que mantenha a simetria interna do oligômero, é natural supor que o mesmo eixo coordenaria o movimento e a consequente simetria dele, por exemplo, uma rotação dos protômeros em torno deste eixo.

Uma vantagem biológica do alosterismo é a não necessidade de interação entre o ligante alostérico e o substrato, pois a regulação se daria somente por uma mudança estrutural na proteína estabilizada ou induzida por seu ligante alostérico (4).

A estabilização ou indução da mudança estrutural pelo ligante envolve uma questão primordial (10). Seria a ligação dele que forneceria energia suficiente para promover a mudança conformacional, ou a proteína já possuiria a energia necessária para transitar espontaneamente entre os estados, sendo o ligante responsável apenas por estabilizar um equilíbrio entre conformações em solução, um equilíbrio populacional?

De qualquer forma, é sempre a proporção entre conformações presentes em solução que estabelece o caminho da reação ou função. E a diferença energética envolvida para mudar essa proporção, esse equilíbrio populacional, parece não ser em vão, posto é que necessária para regular as atividades celulares (11-13). Max Perutz quando discute o alosterismo em seu livro *Mechanisms of Cooperativity and Allosteric Regulation in Proteins*, termina seu prefácio com a crença de que “*one day we shall learn how chemical energy is generated and turned into motion*”. Em muitos aspectos essa questão já foi respondida, como para a contração muscular (14), mas para o alosterismo ainda não se sabe a forma do potencial, como função de conformações, para quê a transição ocorra e muito menos a quantidade de energia necessária para isso (15-16).

Buscando responder à questão de como ocorre a transição alostérica, a enzima Glucosamina 6-fosfato desaminase de *E. coli* aparece como um modelo conveniente (3). A figura 1.1 mostra sua estrutura representada por sua superfície eletrônica e a reação em que atua. Ela catalisa a isomerização e desaminação reversível de D-glucosamina 6-fosfato, resultando em D-frutose 6-fosfato e amônia (17-18).

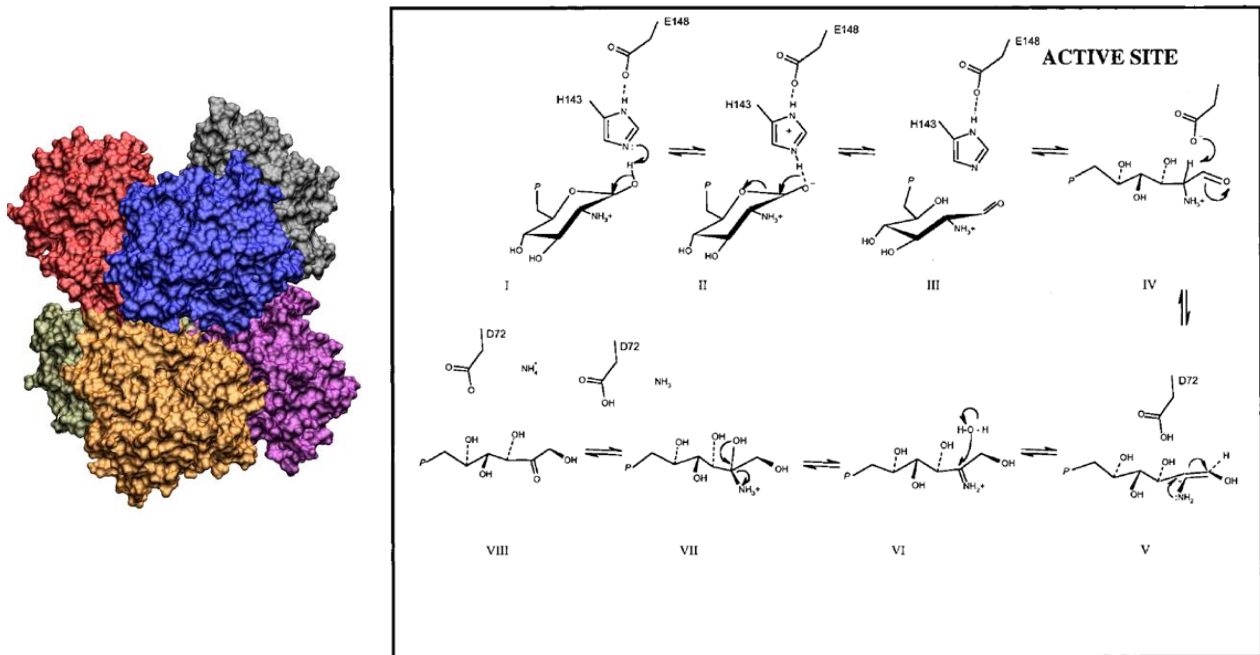


Figura 1.1: À esquerda figura da superfície eletrônica da enzima Glucosamina 6-fosfato Desaminase (Gln6PD). À direita esquema da reação catalisada por ela, que ocorre em duas partes: a primeira é a abertura do anel de frutose e a segunda é sua desaminação (19).

Esse é um passo no catabolismo de açúcares aminados, usados como fonte de carbono. No entanto, esses açúcares, como glucosamina ou N-acetil-D-glucosamina, também são constituintes de lipossacarídeos e proteoglicanos presentes na parede celular bacteriana. Quando não há disponibilidade desses, a reação reversa ocorre pela ação da enzima glucosamina 6-fosfato sintetase, usando Frutose 6-fosfato e L-glutamina como substratos. A atividade de ambas as enzimas é muito bem regulada não só pela inibição do RNA mensageiro, mas também por alosterismo de uma delas (20).

Assim, a proteína GlcN6P desaminase é também uma enzima alostérica e possui as características estruturais descritas na teoria de Monod. Ela é hexamérica formada por subunidades idênticas de 266 aminoácidos (21), como mostra a figura 1.1. Ela também pode ser descrita como um dímero de trímeros (um trímero seria os três protômeros de cima e outro os três de baixo), ou um trímero de dímeros (cada dímero é composto por um protômero e seu imediato abaixo mais próximo). Apresenta-se em solução em dois estados conformacionais reversíveis T e R, que, no cristal, possuem simetria $P6_322$ e $R32$, respectivamente. Ela possui um ativador alostérico, o açúcar aminado N-acetilglucosamina 6-fosfato, e dois substratos, glucosamina 6-

fosfato ou frutose 6-fosfato, a depender do sentido da reação. Nenhum desses tem afinidade pela conformação T. Quando o ativador GlcNAc6P se liga no sítio alostérico, na conformação R, ele a estabiliza na forma R, e os substratos podem se ligar no sítio ativo. Um inibidor competitivo também pode se ligar ao sítio ativo. Neste caso ocorre um impedimento da reação por este protômero, mas todos os outros também se estabilizam na forma R e são ativos, tratando-se de uma ativação paradoxal. Ainda, os substratos também podem atuar como ativadores. Ao ligar em um sítio ativo, os outros 5 sítios do hexâmero ficam em conformação R, ativados (22- 23). Esta ativação chama-se homotrópica (1).

Dados experimentais de ativação enzimática mostram que o ativador liga-se exclusivamente à conformação R, e todos os outros ligantes têm baixíssima afinidade pela forma T. Ambas as ativações, homotrópica e heterotrópica, seguem o modelo MWC de alosterismo. Já na ausência de ligantes o estado predominante é na conformação T, sendo encontradas entre 10.000 e 100.000 estruturas deste para cada uma do estado R (20).

Ainda que em proporções discrepantes, ambos estados são encontrados num *ensemble*, coexistindo em equilíbrio mesmo na ausência de ligantes. Isso indica que é possível encontrar uma movimentação em que espontaneamente haja a transição entre os dois confórmeros. Essa movimentação já foi sugerida baseando-se em estruturas resolvidas por cristalografia, e propõe a divisão de um protômero em unidades menores, seguindo critérios de funcionalidade e rigidez da região, que se movimentariam como corpos rígidos apenas por rotações e translações (3). A exposição exata dessa divisão e da movimentação consta na seção de metodologia desta dissertação.

A descrição da transição alostérica como uma mudança conformacional de corpos rígidos é um modelo bastante simplificado, mas que pode incrementar informações durante a etapa de refinamento de resolução de estrutura por difração de raios x (24). De fato, descrever a proteína como composta por corpos quase rígidos adiciona um termo de contribuição (o tensor de TLS) para o fator de temperatura, ou parâmetros de deslocamento atômico e diminui o valor do fator R, valor que indica o quanto o modelo cristalográfico se distancia dos dados observados (25).

Uma forma de caracterizar a movimentação e a mudança conformacional por que passam corpos quase rígidos é pelo cálculo do tensor de TLS. Explicitamente, TLS representa movimentos de translação, libração ou respiração, e torção – que no inglês se traduzem para translation, libration e screw. Ele é uma contribuição para a flexibilidade da proteína, além de

deslocamentos isotrópicos ou térmicos e dos anisotrópicos, que encontram uma direção preferencial na oscilação da molécula. Estas são três descrições alternativas. Deslocamentos isotrópicos são descritos por esferas para cada átomo, anisotrópicos descrevem um elipsoide para cada átomo de forma independente de átomos vizinhos, enquanto que tensores de TLS também são descritos por elipsoides, mas átomos vizinhos estão fortemente correlacionados. Todos esses movimentos são detectados pela técnica de cristalografia, mas é necessária alta resolução (26).

Ao contrário do senso comum, essa técnica ainda carrega muita informação sobre a flexibilidade da proteína, ainda que restrita pelo empacotamento cristalino. O processo de cristalização, feito a temperatura ambiente, possibilita que a proteína mantenha uma movimentação natural permitida pela flexibilidade de sua estrutura. E o empacotamento de milhares de unidades proteicas resulta num grande *ensemble* de conformações. Ao ser repentinamente resfriado, para as medições de difração de raios x, esse *ensemble* se mantém - no mínimo parcialmente, naquelas conformações de mudança mais lenta (transições de tempos maiores a milissegundos). E como numa fotografia de curta exposição, cada unidade é detectada com sua conformação exatamente no instante de congelamento. A depender da resolução é possível encontrar, além das posições médias de cada átomo do conjunto de proteínas, o quanto cada conformação se diferencia uma da outra pelos parâmetros de deslocamento atômicos.

Assim, o mesmo espaço amostral presente num cristal permite encontrar posições médias a partir da simetria da cela unitária e do grupo espacial, e também de deslocamentos atômicos, o quanto cada conformação presente no *ensemble* se distancia de uma estrutura média.

Como um cristal deve compreender por volta de 10^{15} unidades de proteínas (considerando-se os volumes da cela cristalina e do cristal), e, em solução, para cada 10.000 conformações no estado T existe uma no estado R (no caso da GlcN6PD), é possível que o *ensemble* de um cristal consiga abranger algumas conformações no estado R (ainda que numa proporção menor que em solução), quando a cristalização é feita na ausência de ligantes. Assim também é detectada a grande flexibilidade de pequenas regiões da proteína, como a tampa do sítio ativo da GlcN6PD no estado T, que apresenta um movimento desordenado e, conseqüentemente, inúmeras conformações distintas no espaço amostral do cristal (27).

Essas diferenças conformacionais existem por que a proteína não possui apenas alguns estados estáticos. Pelo contrário, sua estrutura possibilita flexibilidade e dinâmica tais que lhe confirmam determinada função (28). Um dos objetivos desta dissertação é determinar se as

simulações computacionais podem descrever essa flexibilidade em termos compatíveis com os dados bioquímicos (ou de atividade catalítica). Segundo a teoria ergódica há uma relação entre médias temporais e conformações do *ensemble*, na medida em que dado um período suficientemente longo todos os estados de um sistema variável no tempo serão visitados (29). Ou seja, permitindo que a proteína varie sua conformação, após um tempo suficientemente longo, todas as conformações acessíveis a sua estrutura (em determinadas temperatura e pressão, por exemplo) terão sido visitadas. Isso se cumpre desde que se mantenha a mesma energia total (soma das cinética e potencial) quando um *ensemble* micro canônico, e mesma temperatura quando um *ensemble* canônico.

Uma dinâmica molecular, da mesma forma, poderá gerar um *ensemble* de conformações e também proporcionará parâmetros de deslocamento comparáveis aos dados pela cristalografia. No entanto, cada conjunto, ou espaço amostral resultante de cada técnica (seja experimental ou de simulação computacional) terá suas diferenças. A dinâmica pode simular temperaturas mais altas, aumentando a flexibilidade da proteína, enquanto a cristalização geralmente é realizada a temperatura ambiente e a mobilidade ainda é reduzida pelo empacotamento. Isso leva a acreditar que a dinâmica alcançaria conformações não previstas pela cristalografia, ou amplitudes maiores de movimento – resultado também das simplificações do potencial interatômico. Todavia, o *ensemble* de conformações que compõe o cristal é muito maior que o factível para a dinâmica (limitada pelo tempo de cálculo para longas trajetórias e também pela dependência temporal entre um passo na dinâmica e outro). Isto considerando, por exemplo, os experimentos feitos aqui para a GIN6PD, cujo *ensemble* cristalino possuía por volta de 10^{15} conformações, enquanto a dinâmica calculada gerou uma amostragem de 10^5 conformações.

Parâmetros de deslocamento atômicos calculados para a dinâmica, refletindo a flexibilidade proporcionada por esta técnica (quando comparado com a cristalografia), aparentam tender a formar elipsoides mais finos. Eles podem até alcançar as mesmas amplitudes ou maiores que da cristalografia, mas não são tão espalhados (não tem a mesma variabilidade de conformações), como consequência do tempo da dinâmica e de uma trajetória direcionada pelas condições iniciais. Estas características gerais da dinâmica serão observadas em resultados particulares deste trabalho.

Dinâmicas moleculares também são dependentes da escolha do campo de forças, do

potencial, além das condições iniciais e do período simulado. Campos de forças podem descrever com muitos detalhes a natureza, aproximando-a da realidade como a compreendemos, mas representariam um custo de cálculo computacional ainda inviável. Escolher significa selecionar que potenciais provavelmente interferem no experimento que se deseja realizar (30-31).

Para analisar uma transição alostérica é preciso considerar alguns pontos como longos tempos de simulação, pois se espera um movimento global de baixa frequência (32); sistemas atomísticos, por conta de interações intermonoméricas que caracterizam um estado alostérico ou outro (entre outras características, os confórmeros T e R se diferenciam por ligações interatômicas que os estabilizam, mudando a estrutura quaternária da proteína (1)); e sistemas grandes, visto que muitas proteínas alostéricas são oligoméricas. Assim, o estudo do alosterismo, apenas por sua natureza, já implica alto custo computacional.

A possibilidade de descrever tais sistemas como compostos por corpos quase rígidos diminuiria esse custo, mas essa suposição ainda necessita validação e respaldo experimental – um dos alvos deste trabalho.

Experimentalmente, a descrição por corpos quase rígidos já se mostrou benéfica na cristalografia e Monod também previa, em sua teoria sobre o alosterismo, que os protômeros passariam apenas por movimentos de rotação, cada um como um corpo rígido, na transição alostérica. Além de simplificar as mudanças conformacionais da proteína facilitando seu entendimento, tal desenvolvimento seria sensível apenas a movimentos globais. Isso se torna uma vantagem na medida em que, geralmente, são estes os movimentos funcionais, enquanto a movimentação de cadeias laterais, por exemplo, apenas acarreta ruído às análises.

Neste sentido, um filtro de movimento global proposto nesta dissertação é a análise de regiões, entendidas como corpos quase rígidos, pelo cálculo de tensores de inércia. Pela mudança na direção de seus autovetores é possível discutir a rotação do corpo, enquanto a variação dos autovalores expressaria mudanças conformacionais, como será explicado e demonstrado adiante.

Outros filtros são a Análise de Modos Normais e a Análise de Componentes Principais. Eles supõem que o movimento total da proteína possa ser descrito como uma soma de funções, oscilações, harmônicas e assim o decompõem (33). O primeiro método é capaz de encontrar as oscilações de mais baixa frequência, mais globais, enquanto o segundo encontra uma relação entre o modo calculado e sua representatividade numa dinâmica simulada.

Ambas as análises remetem à discussão da harmonicidade do movimento estudado. ANM

pressupõe o potencial energético da proteína como uma função harmônica para o desenvolvimento do cálculo (34). PCA já consegue considerar como uma soma de harmônicos espaciais (em diferentes eixos ortogonais) (35), mas não consegue descrever o movimento como um todo quando há componentes anarmônicas.

Essa descrição é muito válida na medida em que encontrar uma função, uma frequência de oscilação, descritiva do movimento funcional dá uma noção da quantidade de energia necessária para esta oscilação. Quantificar essa energia significa poder averiguar sobre sua origem, sobre a factibilidade de a proteína realizar sua função e, finalmente, entender como energia química é transformada em movimento.

Assim, caracterizando a dinâmica molecular é possível encontrar a conexão entre a estrutura proteica e sua função biológica. Além disso, seria viável inferir sobre por que algumas relações estrutura-função são energeticamente favoráveis em detrimento de outras. E, aos poucos, a intrincada rede de regulações e caminhos de reação que coordenam a célula seria desvendada.

2 Objetivos

Este trabalho tem por objetivo uma melhor compreensão do método de regulação alostérica da enzima Glucosamina 6-fosfato Desaminase. Para isto propõe-se:

- Estipular uma metodologia para caracterização da transição alostérica;
- Aplicar e modelar esta metodologia usando a proteína GlcN6PD;

A metodologia que se pretende desenvolver deverá basear-se em experimentos de simulação computacional, usando as técnicas de Dinâmica Molecular e Análise de Modos Normais, e na comparação destes com resultados experimentais de cristalografia pré-existent.

A comparação será feita através de cálculos como a Análise de Componentes Principais e outros algoritmos a serem escritos que analisem mudanças conformacionais da proteína em questão.

As análises estruturais que se pretende desenvolver com esses algoritmos são de:

- Reconhecimento de regiões que se comportem como corpos quase rígidos, ou elásticos;
- Formação do sítio alostérico através de distâncias interatômicas;
- Flexibilidade caracterizada por parâmetros de deslocamento atômicos, ou fatores de temperatura, isotrópicos e anisotrópicos;
- Rotação e mudanças conformacionais caracterizadas por tensores de inércia em função do tempo.

3 Métodos de simulação computacional para o estudo da transição alostérica

Simular um sistema computacionalmente permite sua submissão a condições irreais, possibilitando análises de experimentos em que poucos parâmetros influenciam. Pode-se encontrar, por exemplo, que o movimento de uma proteína seja descrito por uma composição de oscilações. Oscilações de mais baixa frequência representam mudanças conformacionais globais, envolvendo alterações nas estruturas quaternária e terciária, de maior amplitude e menor energia. Assim, estas seriam, provavelmente, mais representativas do movimento funcional da proteína.

A hipótese de que uma proteína possa transitar livremente entre dois estados alostéricos sem necessariamente ser induzida por um ligante, mas que ele venha para estabilizar um equilíbrio populacional de um dos estados, já tem sido observada por experimentos de cristalografia e de ressonância magnética nuclear (36).

Este trabalho procura observar se a dinâmica intrínseca à estrutura proteica representaria tal transição para o caso de proteínas alostéricas. Duas técnicas principais de simulação computacional foram utilizadas para verificar o suposto à cima, seguidas de algumas rotinas computacionais de análise. Essas técnicas são a Análise de Modos Normais e Dinâmica Molecular.

3.1 Análise de Modos Normais

A priori, esta técnica considera que o movimento em questão seja harmônico. O que significa ser possível descrever somente até o segundo termo da expansão de Taylor de sua energia potencial sem recorrer a mais termos, que confeririam uma anarmonicidade do movimento. A equação (1) mostra a expansão de Taylor para a energia potencial em uma dimensão.

$$V(r) = V_0 + \frac{\partial V}{\partial r} \Big|_0 (r - r_0) + \frac{1}{2} \frac{\partial^2 V}{\partial^2 r} \Big|_0 (r - r_0)^2 \quad (1)$$

Ainda que simplificada, são três termos a serem encontrados para uma completa descrição. No entanto, ao considerar que a conformação média do movimento estaria num mínimo de potencial, ambos dois primeiros termos se anulam (o primeiro por ser uma constante aditiva e que pode ser escolhida como zero sem perda de informações sobre o movimento do corpo, e o segundo por ser a derivada num ponto de mínimo), deixando a função descrita somente pela matriz Hessiana. Explicitamente, para cada átomo i da proteína e em coordenadas generalizadas, $q_i = m_i^{1/2} \Delta x_i$, a matriz Hessiana $3N \times 3N$ é (2):

$$F = \frac{1}{2} \sum_{i,j=1}^{3N} \frac{\partial^2 V}{\partial q_i \partial q_j} q_i q_j \quad (2)$$

Diagonalizar esta função, como na equação (3), onde w_j são os autovetores e ω_j são autovalores e j varia de 1 a $3N$, sendo N o número de átomos, caracterizando um espaço de $3N$ dimensões, significa encontrar um conjunto de autovetores independentes, e seus correspondentes autovalores, que verificam a equação 4, escrita como um conjunto de funções harmônicas de direção dada pelos autovetores e frequência relacionada aos autovalores, onde A_j representa a amplitude de oscilação e ε_j a fase.

$$F w_j = \omega_j^2 w_j \quad (3)$$

$$Q_j = \sum_{i=1}^{3N} w_{ij} q_i = A_j \cos(\omega_j t + \varepsilon_j) \quad (4)$$

Reescrevendo as coordenadas como coordenadas de cada átomo consideradas as massas, tem-se uma função de movimento para cada átomo i da estrutura e cada modo normal j , como na equação 5, que mostra a solução harmônica reescrita para o espaço cartesiano.

$$\Delta x_{ij} = \frac{w_{ij}}{\sqrt{m_i}} A_j \cos(\omega_j t + \varepsilon_j) \quad (5)$$

Os 6 primeiros autovalores, que representam movimentos de rotação e translação

(33), devem ser próximos de zero (validando a diagonalização). Além destes, a solução de mais baixa frequência seria o primeiro modo normal da proteína, e o que melhor descreve movimentos globais. Enquanto que frequências maiores corresponderiam a movimentos locais, de poucos resíduos (33).

Como a transição alostérica é, principalmente, uma deformação na estrutura quaternária, espera-se que seja este o movimento de mais baixa frequência, ou uma combinação de movimentos composta pelos modos normais de mais baixa frequência.

Encontrar estes modos normais exige então, primeiramente, a definição de uma função energia potencial a que se sujeita a proteína; em segundo lugar uma rigorosa minimização de energia da estrutura, para que a derivada do potencial neste ponto possa ser considerada nula; e por fim a diagonalização da matriz hessiana.

O campo de forças usado para este método foi o ENCADV (37), que simula a solvatação através de reescalonamento das constantes dielétricas. A teoria de modos normais é elaborada para sistemas isolados, com energia constante. Assim, o sistema não pode ser periódico e as soluções achadas para sistemas isolados com solvente explícito (com constante dielétrica com incremento radial, casca esférica com posições restringidas, entre outras simplificações) resultam aproximações menos apropriadas que o potencial com solvente implícito. De forma que outros campos de forças, que poderiam usar de solvatação implícita, não resultaram em boa minimização de energia. O campo ENCADV, disponível no pacote do programa Gromacs (38), ainda que considerado ultrapassado, foi o único (dentre os testados - os inclusos no pacote do Gromacs) que proporcionou uma minimização adequada. Importante ressaltar que uma simulação da proteína em vácuo retira termos de fricção com o solvente que poderiam acarretar em componentes anarmônicas do movimento, e a falta de condições periódicas de contorno impede que sejam consideradas translação e rotação como termos harmônicos.

A minimização requereu alguns ajustes nos parâmetros de raio de corte e switch nos potenciais de Coulomb e Van der Waals, como recomendado pelo programa Gromacs, além de cálculos sempre com dupla precisão. Também foi necessário um alto valor para rlist, por

se tratar de uma proteína consideravelmente grande.

Três métodos de minimização estão disponíveis no Gromacs. Steepest Descent, Gradientes Conjugados e L-BFGs.

O método de Steepest Descent calcula (seguindo a equação 6) novas coordenadas a partir de uma original, na direção do negativo do gradiente do potencial, ou seja, na direção da força, como na equação (6), onde Max se refere ao maior valor encontrado da maior força em um átomo entre todos eles, e h_n é o máximo deslocamento permitido.

$$r_{n+1} = r_n + \frac{F_n}{\max(F_n)} h_n \quad (6)$$

Se o potencial das novas coordenadas for menor que original, essas passam a ser o novo ponto de partida para o cálculo e h_n aumenta 20%, do contrário ele diminui para 20% (39). No entanto, para valores muito baixos de força máxima, sua convergência é lenta, necessitando um número grande de iterações (40). Assim, serve apenas como aproximação inicial para a análise de modos normais.

Um problema que esse método enfrenta é frequentemente calcular passos na mesma direção que em passos anteriores (41). Gradientes conjugados é capaz de evitar isso, fazendo com que cada passo seguinte seja dado na direção do vetor conjugado ao vetor anterior. É um método inicialmente mais lento - por obrigatoriamente tomar direções ortogonais, levando a um caminho em “zig-zag” -, mas mais eficiente em regiões mais próximas do mínimo de potencial (42).

Diferente dos métodos anteriores, ditos de primeira ordem, L-BFGS é de segunda ordem. Para calcular a direção de minimização, ele considera não somente a direção de menor declive, como também uma aproximação da curvatura em torno do ponto de partida. Como longe do mínimo a curvatura tende a mudar muito este método não é recomendado para passos iniciais, mas próximo a esse ponto, ele se torna um método muito mais eficiente que os anteriores (39).

Uma primeira abordagem da minimização de energia foi feita por Gradientes Conjugados, seguida do método L-BFGS, por seu caráter refinado. Alguns parâmetros de eletrostática e para o método numérico de vizinho mais próximo estão listados na tabela a

seguir (tabela 3.1) para cada método de minimização utilizado.

Tabela 3.1- Alguns parâmetros (como denominados pelo programa Gromacs) de interações eletrostáticas e do método numérico de busca pelo vizinho mais próximo usados em duas etapas diferentes da minimização.

	Gradientes Conjugados	L-BFGS
emtol (kJ mol ⁻¹ nm ⁻¹)	0.01	0.0000001
emstep (nm)	0.01	0.01
nstlist (passos)	1	1
ns_type	Grid	grid
rlist (nm)	1.6	2.0
coulombtype	Encad-shift	Reaction-Field-zero
rcoulomb_switch (nm)	0.8	0.8
rcoulomb (nm)	1.0	1.4
epsilon_rf	0	0
Vdwtype	Encad-shift	switch
rvdw (nm)	1.0	1.4
rvdw_switch (nm)	0.8	0.8
optimize_fft	Yes	yes

A qualidade da minimização pode ser avaliada de duas formas (33). Uma é garantir que a derivada primeira da energia, o negativo da força, seja a mais próxima de zero possível. Considera-se que a força máxima, a maior força em um átomo dentre todos, deva ser da ordem de 10^{-6} - considerando que a força resultante em um átomo é a soma de inúmeras forças provenientes de interações entre pares de átomos. Outra maneira é verificar se após a diagonalização, os seis primeiros autovalores são próximos de zero e não negativos. A proximidade de zero se dá por a rotação e a translação não contribuírem para o movimento, enquanto valores negativos indicam uma curvatura na energia correspondente a um ponto máximo. Esses pontos de máximo associados a pontos de mínimo (descritos por outros autovalores positivos) caracterizam um ponto de sela no espaço de 3N direções, o que significa insuficiente minimização.

Aplicando os dois métodos como mostrado na tabela 3.1, a minimização alcançou

uma energia potencial e Força máxima que poderiam ser suficientes para a análise de modos normais. No entanto, análises inconclusivas (considerando que as frequências obtidas eram muito mais altas que as previstas) obrigaram a continuação da minimização. Essa continuação foi feita com alguns ciclos alternados dos três métodos (Gradientes Conjugados, Steepest Descent, e L-BFGS), sendo finalizada com o L-BFGS, alcançando melhores valores de Energia potencial e Força máxima, como será mostrado adiante. Nestas últimas etapas os parâmetros utilizados foram os mesmos mostrados para o L-BFGS na tabela 3.1.

Terminada a minimização, a construção da matriz Hessiana foi feita por um cálculo de modos normais, usando o integrador nm do Gromacs. Os parâmetros de busca por vizinho mais próximo e eletrostática, foram mantidos assim como na minimização de energia por L-BFGS.

E por fim, a diagonalização da matriz foi calculada com a rotina `g_nmeig`, também do Gromacs, com dupla precisão.

3.2 Dinâmica Molecular

Três dinâmicas simples foram simuladas, com o programa NAMD (43), para verificar a espontaneidade da transição alostérica.

A estrutura inicial foi retirada do Protein Data Bank (44). O arquivo pdb usado tem código 1fsf (22) e corresponde a um monômero da Glucosamina 6-fosfato Desaminase na sua forma T, resolvido por técnica de cristalografia e difração de raios-x, com resolução de 1,9Å. Os outros 5 monômeros foram gerados pela simetria de grupo espacial do cristal, usando o programa Coot (45).

Usando o programa NAMD, dado um arquivo pdb é preciso gerar um arquivo de estruturas psf para que se prossigam os cálculos de dinâmica. Esse arquivo foi gerado pela rotina `psfgen` e considera cada cadeia separadamente, além de protonar como na tabela seguinte (tabela 3.2) todas as histidinas, por critérios de proximidade com resíduos ácidos.

Tabela 3.2 - Estado de protonação das histidinas. Ela foi feita individualmente de acordo com a distância de cada uma a resíduos ácidos. Na coluna da direita encontram-se os resíduos de histidina existentes na cadeia monomérica, e na coluna da esquerda que nitrogênios foram protonados.

Histidina	protonação
19	N ϵ
56	N ϵ
65	N ϵ e N δ
81	N ϵ e N δ
90	N ϵ e N δ
96	N ϵ e N δ
132	N δ
143	N δ
164	N ϵ e N δ
222	N ϵ e N δ
232	N ϵ e N δ

O sistema foi montado de forma a solvatar a proteína em água (de parametrização TIP3P (46)) sem adição de sais, numa caixa triclinica com todos os ângulos iguais a 90°, deixando uma camada de 0.5 nm de solvente a mais em cada extremidade da proteína nas três coordenadas cartesianas, feita com o programa VMD (47).

Este sistema foi calculado nas seguintes etapas considerando-se o campo de forças charmm (48), disponível no programa NAMD, e condições periódicas de contorno.

Inicialmente o sistema é solvatado sem íons deixando 5Å de camada de solvatação além das extremidades da proteína. Assim ele é minimizado, ou seja, as posições de cada átomo são recalculadas para ficarem as mais próximas possíveis de um mínimo de potencial. Evitando, assim, que átomos com baixa resolução (dadas as limitações da técnica de difração de raios-X) fiquem mal posicionados e iniciem a dinâmica com alta energia. Além disso, a resolução de estruturas por cristalografia implica encontrar a posição média do *ensemble* de proteínas no cristal, o que não é necessariamente a conformação de menor energia, e muitas

vezes leva a existência de interações químicas irreais.

Esta minimização foi feita com o algoritmo de gradientes conjugados. As interações eletrostáticas foram tratadas por Particle Mesh Ewald (PME), com um raio de corte de 1.0 nm, distância mínima entre pares de átomos de 1.3 nm, com uma distância de switch de 0.9 nm.

Após a minimização, é feita uma equilibração. Nesta etapa todos os átomos recebem uma velocidade sendo colocados em banho térmico de 310K, sob o algoritmo de langevin. Isso permite uma suave elevação de energia do sistema. Quando estável - analisados parâmetros de energias total e cinética, pressão, volume e temperatura - o sistema está pronto para iniciar uma trajetória realista, visto que tais parâmetros se aproximam em simular um ensemble real. A pressão é controlada pelo método de Nose-Hoover (49), usando o pistão de Langevin (50), e a temperatura controlada como numa dinâmica de Langevin, com um decaimento e forças aleatórias adicionais (51).

Três dinâmicas foram realizadas, com tempos de equilibração e velocidades iniciais, como mostra a tabela 3.3.

Tabela 3.3: Relação de dissimilaridades entre as três dinâmicas realizadas.

Dinâmica	Tempo de equilibração	Velocidades iniciais
Dinâmica 1	800ps	Aleatórias
Dinâmica 2	800ps	Simétricas em relação à cadeia A
Dinâmica 3	20ps	Aleatórias

De modo geral, todas foram dinâmicas de ensemble NVT, sem controle de pressão, mas mantendo-se o volume constante, e a temperatura controlada pelo método de Langevin.

3.3 *Dinâmica Molecular com ligante*

Como um dos objetivos do estudo é verificar se a movimentação espontânea da proteína é capaz de alcançar um estado conformacional que possa se ligar a um ligante

alostérico (confórmero R) partindo de uma estrutura que não possui ligantes (confórmero T). Uma dinâmica da proteína com um fosfato em um de seus sítios ativos foi realizada.

A conformação inicial da proteína foi retirada de um dos passos das dinâmicas sem ligante descritas anteriormente. A conformação foi escolhida de tal maneira que apenas um dos sítios alostéricos estivesse formado como num estado R, enquanto todos os outros estivessem mais próximos do estado T. Para isso foram medidas as distâncias entre o N terminal de uma cadeia e os carbonos beta da Arginina 158 e Lisina 160 da cadeia vizinha, três aminoácidos chaves na formação do sítio alostérico. A conformação escolhida possui as distâncias indicadas na tabela 3.4.

Tabela 3.4 - Comprimentos de ligação entre três resíduos chaves na formação do sítio alostérico, calculadas sempre entre o Nitrogênio terminal da Metionina e os carbonos beta da Arginina 158 ou da Lisina 160. Na vertical, estão as distâncias de ligação que definem o mesmo sítio, enquanto na horizontal vê-se o comprimento das ligações entre resíduos de mesmo número, mas em sítios diferentes. Medidas tomadas em um instante conveniente da dinâmica 1.

resíduos	Cadeias que formam o sítio alostérico com o N terminal ou a arginina158 e a lisina160, respectivamente.					
	A e C	B e A	C e B	E e F	F e D	D e E
Met1 e Arg158	12,129 Å	15,448 Å	11,551 Å	13,482 Å	13,191 Å	10,852 Å
Met1 e Lys160	13,481 Å	17,002 Å	13,344 Å	13,466 Å	13,749 Å	10,973 Å

Para a inclusão do fosfato no sistema, foi preciso parametrizá-lo. A pH 8, situação em que se estabiliza o confórmero R, o fosfato se encontra monoprotonado. Partindo-se de um fosfato duplamente protonado, H_2PO_4 , já parametrizado e disponível no programa Gromacs, foi calculada nova distribuição de cargas (conforme mostra figura 3.1) e assim feita nova parametrização para o ligante H_1PO_4 , além de novo arquivo de coordenadas. Os grupos de cargas foram calculados como mostra o esquema a seguir, totalizando uma carga de -3.

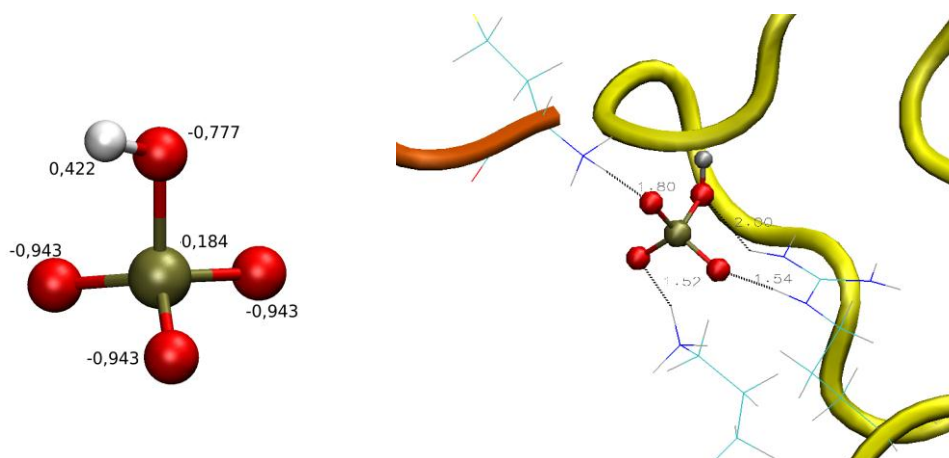


Figura 3.1 – À esquerda, esquema de distribuição de cargas parciais usada para a parametrização do ligante H₁PO₄. À direita, localização do fosfato no sítio alostérico entre as cadeias E (amarela) e D (laranja).

Obtida a estrutura do fosfato, ele foi colocado em situação favorável no único sítio alostérico que se aproximava da conformação R (o sítio entre as cadeias E e D, como indicado na tabela 3.4), aplicando as ferramentas de rotação e translação, somente no fosfato, disponíveis no programa Coot.

Usando o programa Gromacs, foram incluídas no arquivo de topologia, contendo parametrizações da proteína, as parametrizações do fosfato monoprotonado. O sistema todo foi solvatado com moléculas de água de parametrização TIP3P, numa caixa triclínica de forma a possuir uma camada de solvente de 0.5 nm a mais para as extremidades da proteína nas três coordenadas cartesianas.

Sob o campo de forças charmm27, disponível no pacote do Gromacs, todo o sistema foi minimizado, equilibrado deixando somente as águas livres para movimentação, seguido de uma equilibração deixando-se tudo livre para movimentação como um NVT e depois como um NPT, e por último foi iniciada uma dinâmica NPT com o algoritmo V-rescale (52) para acoplamento de temperatura e o Parrinello-Rahman para acoplamento de pressão (53).

3.4 Técnicas de análise das dinâmicas

Algumas análises foram feitas para averiguar sobre mudanças nas estruturas quaternária e terciária da proteína ao longo das trajetórias simuladas. Elas estão listadas a seguir e foram usadas convenientemente para cada simulação realizada (três dinâmicas simples, uma dinâmica com ligante e análise de modos normais). Excetuando a Análise de Componentes Principais, feita com o programa Bio3D (54), para todas as outras análises foram escritas rotinas em Fortran ou Tcl. As estruturas cristalográficas usadas para comparação foram as de código pdb 1fsf (22) e 1dea (23), respectivamente, das conformações T e R da GIN6PD.

3.4.1 *Análise de Componentes Principais*

Em semelhança com a Análise de Modos Normais, este método também tem por objetivo decompor um movimento em oscilações harmônicas com diferentes eixos ortogonais e identificar aqueles de maior representatividade (35).

Diferente de NMA, PCA não pressupõe que haja relações temporais entre as estruturas do *ensemble* fornecido, ou seja, não considera um potencial, uma função, que as relacione, mas também encontra um espaço de $3N$ dimensões, onde N é o número de átomos, em que se distribuem as conformações. A cada eixo é atribuído um movimento de todos os átomos nesta direção, englobando as conformações que são distinguíveis neste. Encontrados os eixos ortogonais de maior representatividade desta distribuição, ou seja, que melhor distinguem as conformações do ensemble, foi possível inferir sobre quantos deles são suficientes para descrever todas as conformações.

Caso as estruturas fornecidas se relacionem por um movimento harmônico, esta análise deve encontrar uma componente principal que o descreva, ou uma soma de componentes que componha o movimento. Denominam-se anarmônicos os modos que compõem essa soma (55). Assim, a soma de todos os modos converge para descrever 100% das estruturas fornecidas. Quando a análise não encontra uma convergência na soma dos modos anarmônicos, pressupõe-se que não seja possível relacionar todas as estruturas

fornecidas por um movimento harmônico.

A técnica se reduz a encontrar a correlação entre as estruturas dadas, ou, neste caso, entre as distintas conformações encontradas no decorrer da trajetória calculada na dinâmica (33). Para isso é calculada uma matriz de covariância (equação 7), de tamanho $3N \times 3N$, em que existem $3N$ coordenadas r_i - onde o índice i varia de 1 a $3N$, correspondendo às três coordenadas cartesianas, x , y , e z , de cada átomo dentre os N da estrutura da proteína -, para cada estrutura j do *ensemble* fornecido, do qual são calculadas as médias de acordo com a equação (7), onde o índice j é relativo ao número de conformações fornecidas, e as médias são feitas para o *ensemble* de conformações j .

$$C = \left\langle (r_{ij} - \langle r_{ij} \rangle) (r_{ij} - \langle r_{ij} \rangle)^T \right\rangle \quad (7)$$

Essa matriz é diagonalizada para resultar em autovetores que indicam o espaço de uma movimentação e seus correspondentes autovalores, que indicam, nesse caso, a variação da estrutura nesse espaço. Os autovalores e autovetores se relacionam como na equação (8),

$$p_k = \mu_k \cdot (x_i - \langle x_i \rangle) \quad (8)$$

enquanto que o deslocamento de cada átomo para um modo, ou, componente k é dado pela equação (9), onde o índice i varia de 1 a $3N$, correspondendo às três coordenadas cartesianas, x , y , e z , de cada átomo dentre os N da estrutura da proteína,

$$r_{ik} = p_k \cdot \mu_k + \langle r_{ij} \rangle \quad (9)$$

resultando em k trajetórias que compõem o movimento inicial.

É fundamental para este método que todas as estruturas estejam alinhadas, superpostas entre si, removendo termos de rotação e translação da proteína como um todo.

Em termos gerais, qualquer conjunto de estruturas pode ser dado como informação inicial e por isso não há relação com o tempo de dinâmica e não é possível, somente com esta análise, inferir sobre as frequências dos movimentos. Mas isso também possibilita ampla utilização técnica, que neste projeto foi usada para analisar não só as variadas conformações

durante a trajetória da dinâmica, mas também as encontradas para uma trajetória calculada para o primeiro modo normal, e as diferenças entre as estruturas cristalográficas T e R.

3.4.2 Fatores de temperatura ou parâmetros de deslocamento atômico

O fator de temperatura é um parâmetro experimental que pode ser obtido da técnica de difração de raios-x. Ele reflete a flexibilidade de cada átomo, podendo ser descrito, em sua forma mais abrangente, como uma matriz de deslocamentos anisotrópicos. A equação (10) mostra a matriz assim como definida no Protein Data Bank.

$$U = 10000 \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix} \quad (10)$$

No entanto, tal dado exige alta resolução da estrutura e muitas vezes ele é encontrado somente como o traço desta matriz. De fato, o fator B, fator de temperatura isotrópico, é dado pela equação (11), sendo chamado também de fator de Debye-Waller ou parâmetro de deslocamento atômico isotrópico.

$$B = 8\pi^2 (U_{11} + U_{22} + U_{33})/3 \quad (11)$$

Assim, o fator isotrópico, simplificado, reflete somente a amplitude do movimento, enquanto que a representação dos outros seis termos anisotrópicos acrescenta informações sobre a direção do movimento de cada átomo.

Como, nestes estudos em particular a Glucosamina 6-fosfato desaminase foi resolvida para o confórmero T com uma alta resolução, 1,9 Å, seria possível também considerar os termos anisotrópicos e compará-los com os valores encontrados para a dinâmica, validando a simulação.

Calcular a matriz de deslocamentos anisotrópicos também é possível para uma dada trajetória, visto que ela pode ser reescrita como na equação 12, como calculada para cada

átomo em uma proteína, cuja estrutura possui diferentes conformações. A média é feita sobre o *ensemble* de estruturas fornecido, neste caso as conformações da trajetória.

$$U = 10000 \begin{bmatrix} (x - \langle x \rangle)(x - \langle x \rangle) & (x - \langle x \rangle)(y - \langle y \rangle) & (x - \langle x \rangle)(z - \langle z \rangle) \\ (y - \langle y \rangle)(x - \langle x \rangle) & (y - \langle y \rangle)(y - \langle y \rangle) & (y - \langle y \rangle)(z - \langle z \rangle) \\ (z - \langle z \rangle)(x - \langle x \rangle) & (z - \langle z \rangle)(y - \langle y \rangle) & (z - \langle z \rangle)(z - \langle z \rangle) \end{bmatrix} \quad (12)$$

O cálculo de parâmetros de deslocamento atômicos foi feito para analisar todas as trajetórias simuladas, não somente a dinâmica molecular, como o primeiro modo normal, a primeira componente principal e a transição encontrada por PCA entre os confórmeros T e R.

3.4.3 Alinhamento e análise de corpos rígidos

Análises de mudanças conformacionais, base deste trabalho, são altamente dependentes de um bom alinhamento, ou superposição, das estruturas em questão, desprezando, assim, termos de rotação e translação que interferem na quantificação de deslocamentos atômicos.

Ademais, a escolha consciente entre um alinhamento global e um alinhamento local pode simplificar a compreensão do movimento que está acontecendo, como exemplificado no esquema abaixo (figura 3.2). A diferença entre as figuras fica clara no alinhamento local, enquanto no global percebe-se apenas que existe alguma mudança.

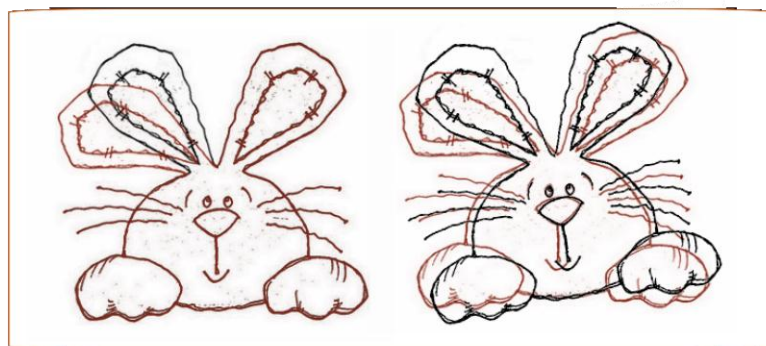


Figura 3.2 - Duas figuras que diferem apenas em uma região específica, a orelha direita do coelho. Na figura esquerda elas estão alinhadas localmente, usando a cabeça como base. Enquanto na figura da direita foi feito um alinhamento global, em que o desvio deve ser o menor possível para todos os pontos do desenho.

Essa ideia também é aplicável para proteínas. Em particular, para a Glucosamina 6-fosfato Desaminase, é claro que duas regiões para cada protômero, se movimentam de maneiras diferentes. Como mostra o esquema abaixo (figura 3.3), ela pode ser dividida em dois corpos quase rígidos que se movimentariam de maneira independente. A região interna se comporta como um corpo praticamente fixo, enquanto a região externa como um corpo quase rígido (que não passa por grandes deformações) que possui uma movimentação de rotação em torno de um eixo.

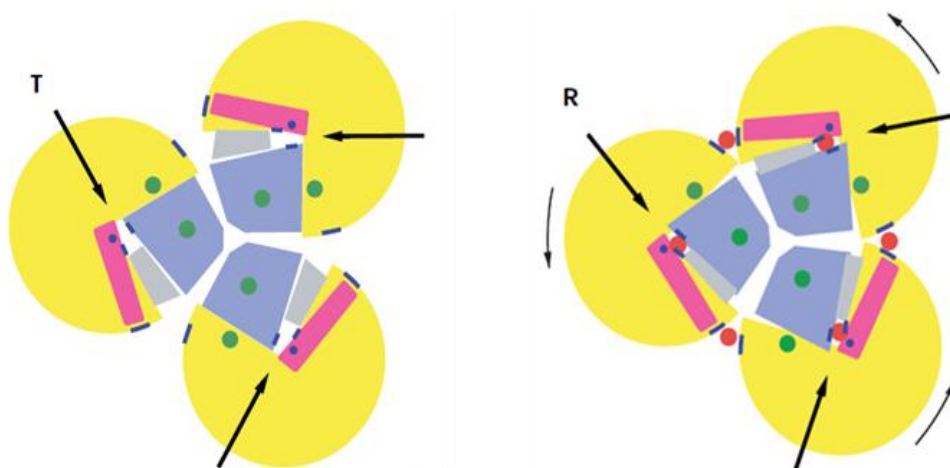


Figura 3.3 - Representação da movimentação alostérica sugerida pela sobreposição dos estados cristalográficos. A esquerda representa-se um trîmero estado T e à direita está um do estado R. Propõe-se a existência de duas subunidades quase rígidas, a interna em azul e a externa em amarelo. A externa passa por uma rotação em torno do ponto verde na região amarela de modo a possibilitar a ligação de grupos fosfatos (pontos em vermelho) tanto no sítio alostérico quanto no ativo. Figura retirada de (3).

Dada à relevância de um alinhamento cuidadoso, uma rotina em Fortran foi escrita

para o cálculo desses alinhamentos.

Baseado em algoritmo dentro do pacote LM-Mdanalysis (56), o alinhamento é feito retirando-se a diferença entre os centros de massa, e em um passo pelo cálculo da matriz quaternion seguida de sua diagonalização (pelo método Jacobi (57)), que resulta na matriz rotação a ser aplicada (58). Retirando a translação e a rotação dentre as diferenças nas coordenadas para o grupo de átomos considerados. Ressaltando que a região base para este cálculo pode ser definida apenas como trechos da sequência primária da proteína e também por tipo de átomo, considerando, por exemplo, apenas átomos da cadeia principal.

Foi implementada a este algoritmo uma rotina de iterações que reescolhe resíduos da região base para o alinhamento de acordo com as distâncias dos carbonos α , depois de superimpostos, de forma a convergir à lista de átomos da região base, para somente aqueles pertencentes a um core rígido. Isso refina o alinhamento local, ainda que a escolha inicial de átomos não represente adequadamente um corpo quase rígido. Ou seja, as sucessivas iterações são capazes de modificar a lista de átomos a serem alinhados de acordo com a distância entre eles após cada alinhamento. Assim, dada uma distância máxima de proximidade entre pares de átomos para incluí-los na lista, o alinhamento final encontra regiões que não se distanciam mais que a dada distância. Essas regiões são denominadas de core rígido, possuindo diferenças de rotação e translação entre as estruturas fornecidas, mas sem muita flexibilidade conformacional.

3.4.4 *Existência do sítio alostérico*

Uma das características principais de proteínas alostéricas é sua capacidade de serem reguladas por um ligante que interaja com um sítio da proteína distinto do ativo, o alostérico. Essa interação não possui diferentes produtos, mas provoca ou estabiliza uma mudança conformacional que permite a ativação do sítio ativo, que em outra conformação não consegue ligar o substrato.

Para a Glucosamina 6-fosfato desaminase, a transição alostérica significa a mudança de um estado sem capacidade de ligar ligantes nem ativadores (o estado T), para outro que

possui alta afinidade por ligantes tanto no sítio ativo quanto no alostérico (o estado R). Esta alta afinidade é garantida por uma aproximação maior entre resíduos do sítio alostérico, em particular do N terminal de um protômero e da arginina 158 e lisina 160 do protômero vizinho, como mostra a figura 3.4. Essa aproximação é estável na presença de um fosfato, sendo ele suficiente para manter a conformação R, ativando o sítio ativo. Assim, tal parâmetro, de distâncias no sítio ativo, é condição necessária e suficiente para caracterizar a conformação ativa, o estado R.

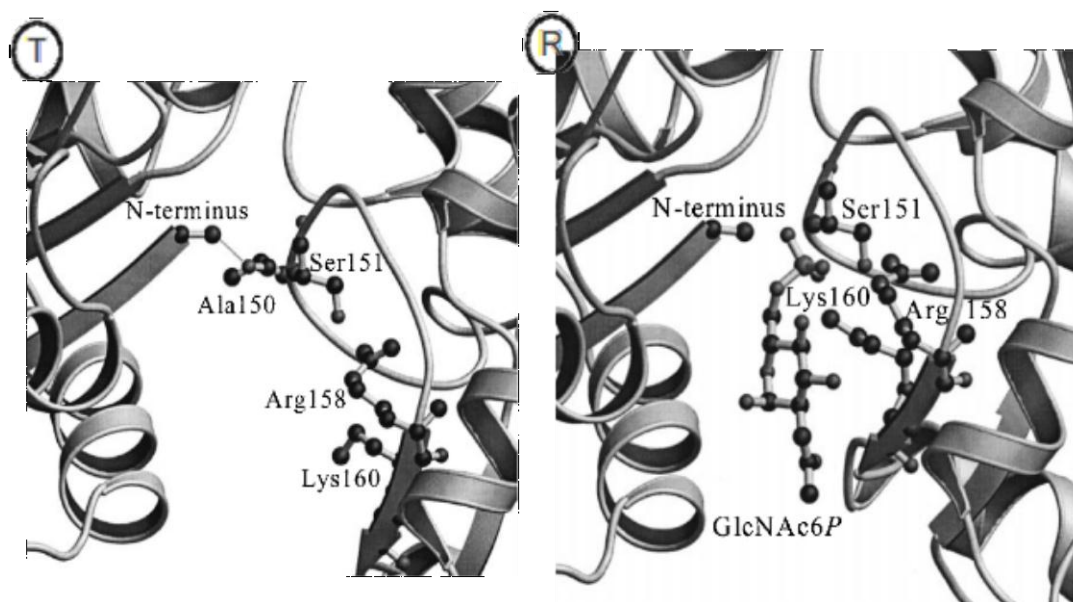


Figura 3.4 - Sítio alostérico formado entre dois monômeros do estado T (esquerda) e do estado R (direita), resolvidos por cristalografia. No estado T a arginina 158 e a lisina 160 ficam distantes do N terminal do monômero vizinho. Enquanto que no estado R esses três resíduos estão próximos o suficiente para fixar um ligante sítio, o GlcNAc6P, neste cristal. Figura retirada de (59).

A tabela 3.5 mostra essas distâncias para os dois conformêros, sendo considerada apenas a distância entre N terminal e carbono beta dos outros dois resíduos, a fim de retirar influências da alta flexibilidade da cadeia lateral.

Tabela 3.5 - Distâncias entre o N terminal da Metionina e o Carbono beta da Arginina 158 ou da Lisina 160 para os dois estados cristalográficos T e R, de código 1fsf e 1dea, respectivamente.

Distâncias cristalográficas entre N terminal e Carbono β		
Resíduos	Estado T	Estado R
Met1 e Arg158	14,2Å	10,2Å
Met1 e Lys160	13,3Å	10,4Å

Essas distâncias foram monitoradas durante a dinâmica pra analisar a formação do sítio alostérico no decorrer da dinâmica, supondo que ele possa se formar espontaneamente apenas por uma oscilação natural da estrutura proteica.

3.4.5 *Deslocamento do centro de massa*

Ainda que seja um bom parâmetro para avaliar a transição alostérica, distâncias interatômicas são carregadas de ruído devido à alta flexibilidade local das cadeias laterais, e podem não revelar com clareza sobre movimentações globais.

Como o movimento em estudo já foi caracterizado como a rotação da região externa (3), que se mantém como um corpo quase rígido, um parâmetro avaliado aqui é a movimentação do centro de massa dessa região. Sem considerar a tampa do sítio ativo, essa região compreende os resíduos de 1 a 161, 185 a 206, 233 a 243 e de 251 a 266.

Assim como para as distâncias entre resíduos no sítio alostérico, os confórmeros R e T (cristalográficos) diferem na posição do centro de massa da região externa. Centralizando toda a estrutura na origem, o ângulo formado entre este centro de massa, a origem e o eixo x, apresenta uma mudança de 10° quando comparados os estados T e R, como mostra a figura 3.5 abaixo.

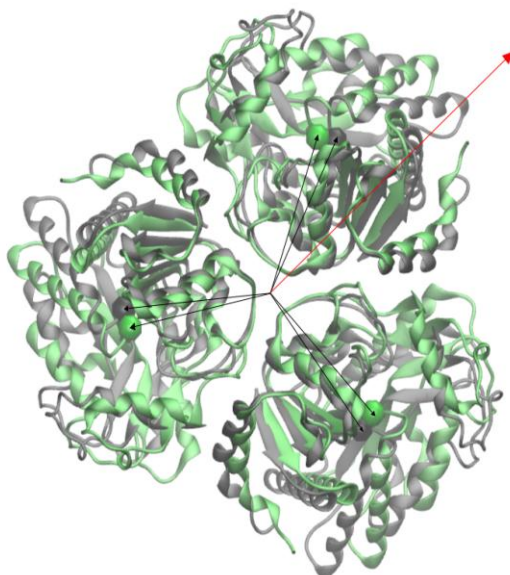


Figura 3.5 - Trímero das estruturas cristalográficas dos estados T (preto) e R (verde) alinhados pela região interna. Os centros de massa estão representados por esferas de mesma cor. Em vermelho está representado o eixo x, que, junto com as setas pretas, forma os ângulos monitorados durante a dinâmica. Cada par de setas vermelhas formam um ângulo de 10° entre si, representado a mudança angular de centro de massa que ocorre na transição alostérica.

Tal deslocamento angular foi também monitorado durante a dinâmica.

3.4.6 *Tensores de inércia*

Outro parâmetro que avalia sobre deformações e movimentações globais é o tensor de inércia (8). Dependente da posição dos átomos e da massa deles, como na definição a seguir, na equação (13), onde o índice i representa cada componente, ou átomo para o caso de proteínas.

$$I = \sum_{i=1}^N m_i \begin{bmatrix} y_i^2 + z_i^2 & x_i y_i & x_i z_i \\ y_i x_i & x_i^2 + z_i^2 & y_i z_i \\ z_i x_i & z_i y_i & x_i^2 + y_i^2 \end{bmatrix} \quad (13)$$

A diagonalização desse tensor resulta em três autovetores e três autovalores, como na equação 14, onde os I_i representam os autovalores, enquanto as colunas da matriz R são os

autovetores, com suas direções dadas pelas três coordenadas cartesianas, x, y e z, de cima para baixo em cada coluna, respectivamente. Quando o corpo passa por uma rotação a direção dos autovetores muda, enquanto que uma mudança nos autovalores reflete uma deformação do corpo.

$$I_D = RIR^{-1} = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \quad (14)$$

Autovalores que variam pouco são característicos de corpos quase rígidos.

O movimento de encolhimento e alongamento de uma hélice alfa, por exemplo, apresenta as mudanças de direção e amplitude para o tensor de inércia, como indicado na figura 3.6.

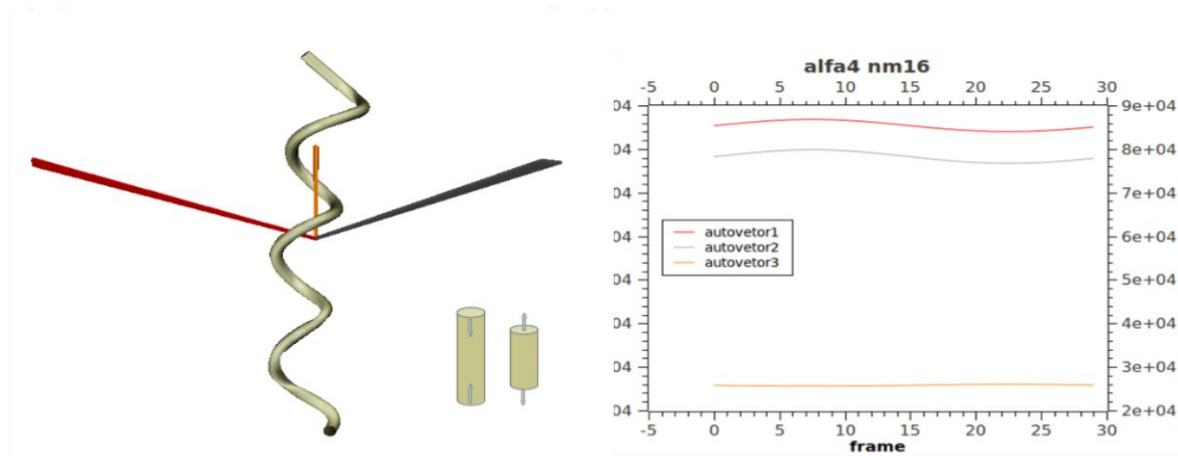


Figura 3.6 - Análise da movimentação de encolhimento e alongamento de uma hélice α , como ilustrado no esquema central. A esquerda está representada a hélice analisada e os três momentos de inércia que compõem o tensor. Está desenhado um tensor para cada passo trajetória que compõe o movimento. A direita está graficada a amplitude, o autovalor, correspondente a cada autovetor, durante os passos da trajetória.

A amplitude dos autovetores é bastante informativa quanto à estrutura do corpo estudado. A hélice da figura 3.6, por exemplo, possui o autovetor ao longo de seu eixo muito menor se comparado os autovetores ortogonais. No entanto, é a variação desses autovalores que carrega informações sobre as deformações do corpo durante essa trajetória particular. Essa variação nos autovalores é dada pela equação (15), onde o índice i varia de 1 a 3 como indicado na equação 14 para os autovalores.

$$\Delta I_i(t) = \frac{I_i(t) - \langle I_i(t) \rangle}{\langle I_i(t) \rangle} \quad (15)$$

Essa variação dos autovalores pode ser graficada e a mudança torna-se muito mais clara que apenas representando a amplitude contra o tempo, como mostrado no gráfico 3.1.

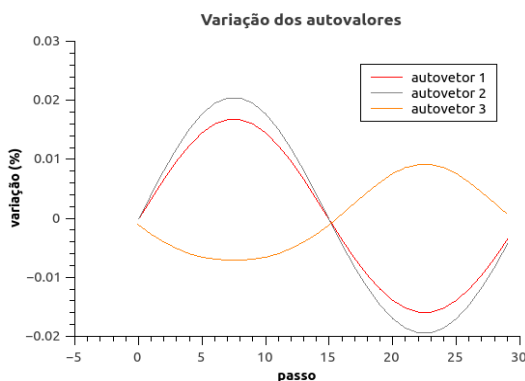


Gráfico 3.1 - Variação em porcentagem dos autovetores de uma hélice α , durante o movimento de encolhimento e alongamento.

A análise da direção dos autovetores também se torna mais clara quando calculada sua variação com a trajetória em relação a uma direção média. Essa variação pode ser dada pela projeção dos autovetores no plano perpendicular a um autovetor médio, como mostra a figura 3.7. Esse novo plano é definido, então, por dois novos versores, p e q.

$$\Delta R_i(t) = \frac{R_i(t) \times \langle R_i(t) \rangle}{\langle R_i(t) \rangle}$$

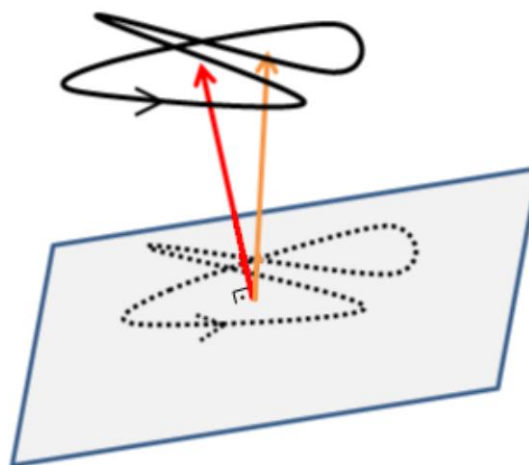


Figura 3.7 - Representação da projeção da variação dos autovetores do tensor de inércia em função do tempo. À esquerda está a equação que define o cálculo da variação dos autovetores, sendo i índice que varia de 1 a 3 como indicado na equação 14 para os autovalores e seus respectivos autovetores. À direita encontra-se um esquema representando a projeção dos autovetores durante a trajetória sobre um

plano perpendicular ao vetor médio dos autovetores (que possuem a mesma origem) ao longo do tempo da dinâmica. No desenho a média, $\langle R_i(t) \rangle$, está representada em vermelho e em azul está o plano perpendicular a ela. Em laranja está o autovetor variável no tempo $R_i(t)$, que descreve uma trajetória desenhada em preto, e projetada no plano como uma linha tracejada.

Tomando, mais uma vez, a hélice α em movimento de encolhimento e alongamento, a projeção da variação de cada um de seus autovetores está graficada na figura 3.8. Comparando com o esquema da figura 3.7, os desenhos formado nos três gráficos são análogos a trajetória pontilhada no plano azul.

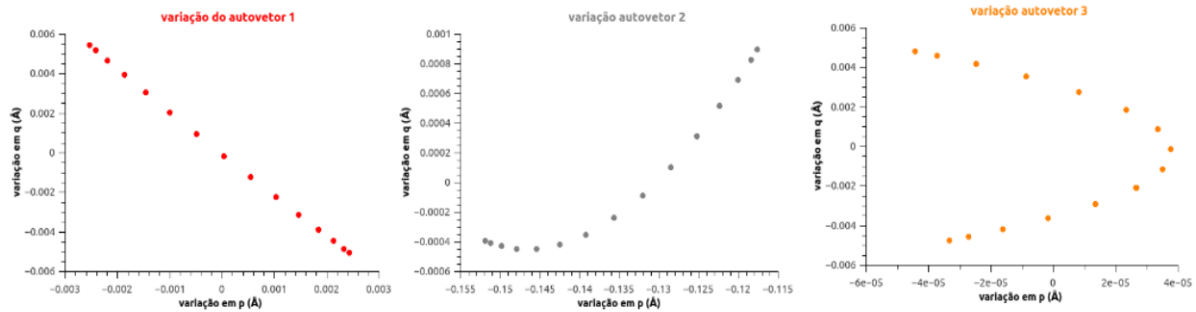


Figura 3.8 - Gráficos das variações em relação à média de cada um dos autovetores da hélice α em movimento de encolhimento e alongamento projetadas no plano perpendicular ao vetor médio.

Para os confórmeros T e R, também é encontrada uma variação no tensor de inércia calculado para a região externa, como mostra a figura 3.9.

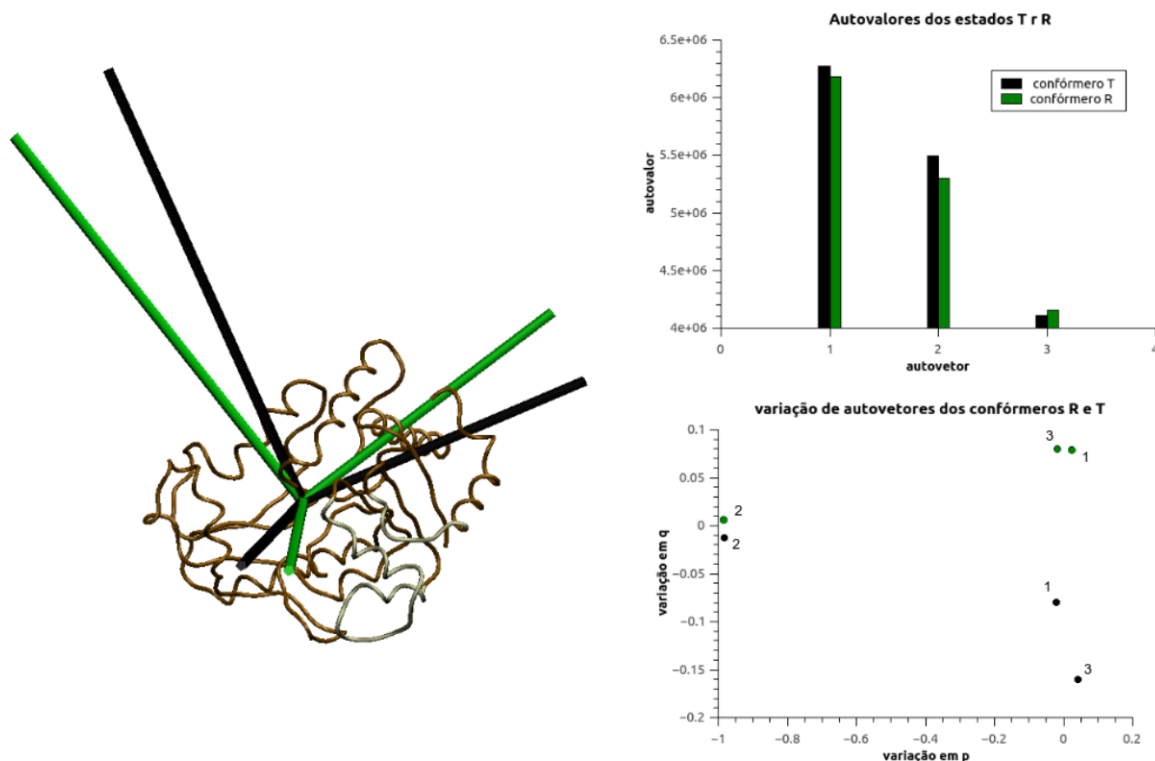


Figura 3.9 - Análise dos tensores de inércia para os estados T (preto) e R (verde) cristalográficos. Os cálculos foram feitos somente para a região externa do monômero, ressaltada em marrom no desenho a esquerda. O gráfico à direita e acima compara os autovalores do T com o R. E o gráfico abaixo é a projeção de cada autovetor num plano perpendicular a um autovetor médio entre as duas conformações.

Ao longo da dinâmica, esses tensores também foram calculados, fazendo-se as análises aqui exemplificadas.

A validação sobre a definição de algumas regiões como corpos quase rígidos também pode ser feita com os tensores de inércia. Quão rígida a região escolhida é está relacionada com o valor dos momentos de inércia, ou seja, com o a variação da amplitude dos autovalores, que não existe no caso de um rígido. Já a variação da direção dos autovetores é causada pela rotação da região.

A figura 3.10 contrasta resultados da dinâmica de um monômero completo, com somente a região externa, sem a tampa do sítio ativo, do mesmo monômero.

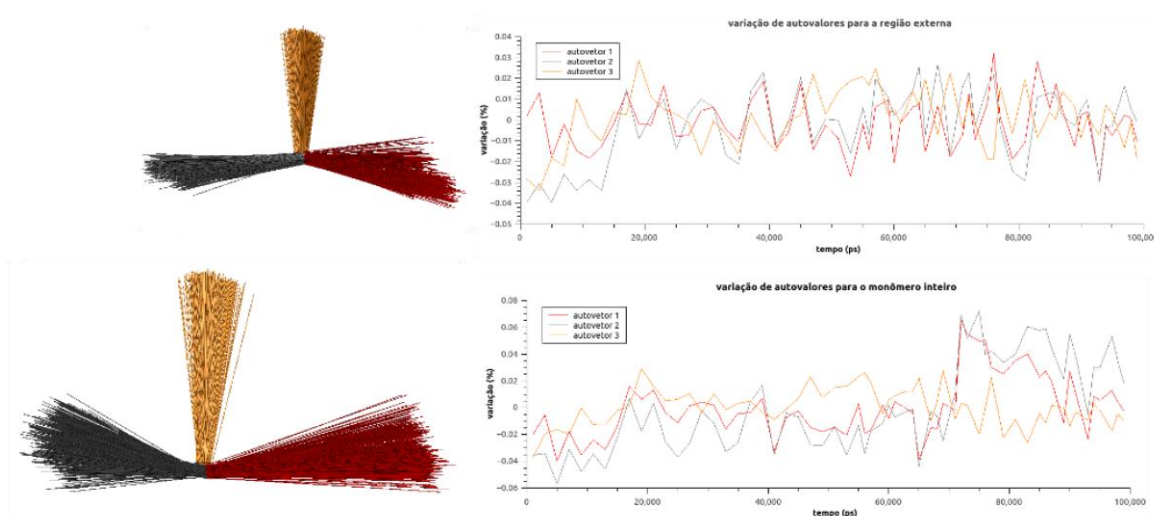


Figura 3.10 - Representação de como a deformação do corpo altera os tensores de inércia. A cima está o desenho dos tensores de inércia da região externa de um monômero ao longo da trajetória e o gráfico da variação de seus autovalores. Abaixo estão as mesmas análises para o monômero completo, em mesma escala para os desenhos dos autovetores.

O exemplo da figura 3.10, mostra que os autovalores são muito mais variantes para o monômero inteiro que para somente a região externa, considerada um corpo quase rígido. Também os cones formados pelos tensores ao longo da dinâmica ficam consideravelmente mais fechados para o cálculo somente da região externa. Isso exemplifica o porquê da análise de regiões quase rígidas no lugar de análises globais, assim como no caso de alinhamentos locais e globais.

3.4.7 Correlação de oscilações harmônicas

Como durante o projeto muitas trajetórias que representavam uma oscilação harmônica na direção de um autovetor, encontrado por Análise de Modos Normais ou por Análise de Componentes Principais, foram simuladas, uma ferramenta usada para comparar estes modos foi o produto escalar dos vetores que a definia.

Ainda que os programas de análise utilizados (Gromacs e Bio3d) calculem os autovetores, eles não são fornecidos de forma simples para leitura. Assim, eles foram

calculados como versores de dimensão $3N$, três para cada coordenada cartesiana de cada átomo (foram considerados somente carbonos α), dados pelas coordenadas de uma das duas conformações de maior amplitude, subtraídas as coordenadas da conformação média do *ensemble* da dinâmica, divididas pelo módulo desta subtração.

O produto escalar dos versores é então calculado e sua proximidade à unidade reflete um movimento em direções semelhantes, enquanto valores próximos de zero significam movimentos em direções ortogonais. Como uma validação do cálculo, foi feito o produto de escalar de um movimento com ele mesmo, o que resultou num escalar próximo a um, com erro da ordem de 10^{-6} .

No entanto, como PCA e NMA resultam em $3N$ autovetores ortogonais, a probabilidade de se encontrar autovetores de direções paralelas é muito pequena, de forma a se esperar sempre valores muito próximos de zero. Ainda assim, quão próximos de zero pode ser uma análise relevante, visto que o erro de cálculo está na ordem de 10^{-6} , enquanto os resultados que serão mostrados adiante estão da ordem de 10^{-2} ou 10^{-3} .

3.4.8 Algoritmos para análises estruturais

Com exceção do PCA e do NMA, todos os outros algoritmos para análise estrutural usados neste trabalho são de autoria própria. Alguns deles ainda fazem uso da sub-rotina para leitura de arquivos em formato dcd do pacote LM-Mdanalysis (56), ou da sub-rotina Jacobi para diagonalização de matriz retirada do livro *The art of scientific computing* de William H. Press (57).

Todos eles foram escritos em linguagem Fortran90. Aqueles que necessitavam representação gráfica foram escritos em Tk/Tcl para execução e visualização pelo programa VMD. Também foi usado o programa Inkscape (60) para a edição de imagens e o Qtiplot (61) pra montagem dos gráficos. A seguir está descrito o algoritmo feito para cálculo dos tensores de inércia.

Para calcular os tensores de inércia de uma proteína em todos os quadros de uma

trajetória simulada é preciso fornecer como arquivos de entrada a trajetória em formato dcd e um arquivo de estrutura no formato pdb. Para uma análise correta é preciso que o arquivo de trajetória tenha seus quadros previamente alinhados por uma região distinta a que se deseja analisar.

Duas informações serão requisitadas, quais são os seguimentos que compõem a região a ser analisada e a que cadeia pertencem. Assim, por uma busca comparativa no arquivo pdb, os números dos átomos a serem considerados serão encontrados.

Todos os átomos são levados em conta para o cálculo e a massa usada para cada tipo se encontra na tabela 3.6.

Tabela 3.6 - Lista de átomos e suas respectivas massas atômicas usadas para o cálculo de tensores de inércia.

Átomo	Massa (u)
O	15,9994
N	14,0067
C	12,0110
S	32,0600
H	1,0080

O centro de massa da região selecionada é então calculado, para uma conformação escolhida, e todos os átomos desta mesma região, mas de outras conformações, são recentralizados para este ponto. Uma matriz 3x3 de tensor de inércia é então calculada como definida na equação 13, somente para os átomos selecionados. Na sequência, ela é diagonalizada pelo método de Jacobi retornando três autovalores e três autovetores de dimensão 3. Estes são organizados de forma que o autovetor 1 seja o de maior autovalor, o autovetor 2, seja o de segundo maior autovalor, e o autovetor 3 o de menor.

Os resultados obtidos são, por fim, organizados em um arquivo formatado. Essa formatação é aqui denominada de tin, e possui a ordem explicada na tabela 3.7.

Tabela 3.7 - Formatação do arquivo que descreve os tensores de inércia, denominada tin.

Colunas	Tipo de dado	Campo
1-6	Inteiro	Número do quadro
8	Caractere	Nome da cadeia
10-19	Real (10.5)	Componente x do centro de massa
21-30	Real (10.5)	Componente y do centro de massa
32-41	Real (10.5)	Componente z do centro de massa
43-62	Real (20.2)	Autovalor
64-73	Real (10.7)	Componente x do autovetor
75-84	Real (10.7)	Componente y do autovetor
86-95	Real (10.7)	Componente z do autovetor

Cada linha corresponde aos dados de um autovetor, centralizado no centro de massa da região escolhida. São necessárias três linhas assim formatadas para descrever o tensor de inércia em cada instante de tempo, ou cada quadro da trajetória. Elas aparecem em sequência neste arquivo, de forma que o número do quadro apenas muda a cada três linhas.

Esse arquivo é usado em outros algoritmos de análise do tensor de inércia, como para a representação gráfica ou a projeção dos autovalores e autovetores.

4 Resultados e Discussão

Os resultados a seguir são de experimentos, análises de simulações, que focam na observação da transição da proteína entre seus dois estados alostéricos. Ambos estados têm suas estruturas resolvidas por cristalografia, e são as de código PDB 1fsf e 1dea que servem

como parâmetro de comparação para os estados T e R respectivamente. A análise estrutural dos dois estados é feita sempre partindo de suas conformações superpostas pela região interna, fazendo um alinhamento local. A figura 4.1 mostra os dois conformémeros cristalográficos superpostos, ressaltando a região interna em azul.

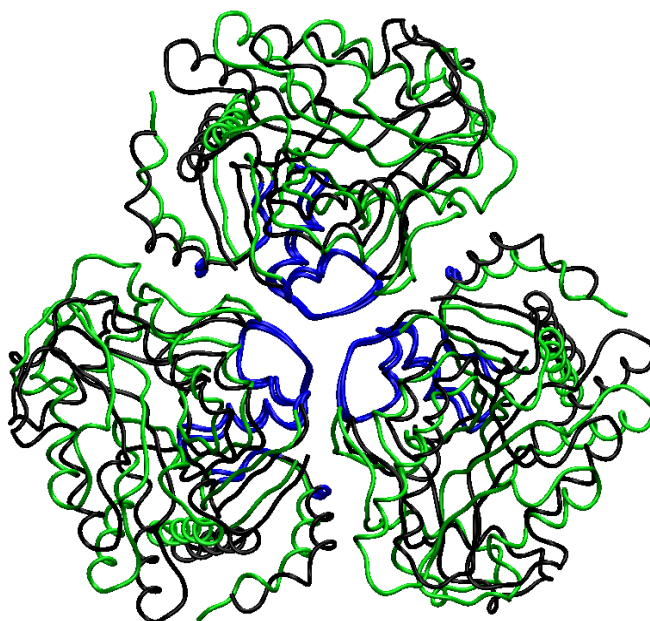


Figura 4.1 - Superposição de trimeros das conformações T, em preto, e R, em verde, alinhadas estruturalmente pela região interna, em azul, que abrange os resíduos de 207 a 233 e de 244 a 250. Está representado apenas um trímero de cada conformação, cujos protômeros se relacionam por simetria 3 em torno do eixo z. O outro trímero se relaciona a este por uma rotação e inversão de ordem 2.

Da figura 4.1 percebe-se que a região interna, que compreende os resíduos de 207 a 233 e de 244 a 250, de fato não mostra grandes desvios entre as duas estruturas. A tabela 4.1 mostra as distâncias entre carbonos alfas que se superpõem na região interna, e não há distância maior que 0,55Å.

Tabela 4.1 - Lista de resíduos dos conformémeros T e R que se alinharam, ou seja, cujos carbonos α ficaram a uma distância listada na terceira coluna.

Resíduos		Distância entre C α (Å)
Confórmero T	Confórmero R	
212	212	0,542
213	213	0,455
214	214	0,518
215	215	0,531
216	216	0,489
217	217	0,417
218	218	0,343
219	219	0,268
220	220	0,260
221	221	0,369
222	222	0,283
223	223	0,412
224	224	0,169
225	225	0,540
226	226	0,173
227	227	0,442
244	244	0,316
245	245	0,530
246	246	0,458
247	247	0,191
248	248	0,273
249	249	0,273
250	250	0,384
251	251	0,491
252	252	0,373
253	253	0,441

Uma oscilação harmônica, entre as duas conformações foi calculado pelo método de Análise de Componentes Principais, PCA. Essa trajetória será chamada aqui, por questões de

conveniência, de transição ideal. O gráfico (gráfico 4.1) a seguir mostra a flutuação média em Å de cada resíduo para este movimento.

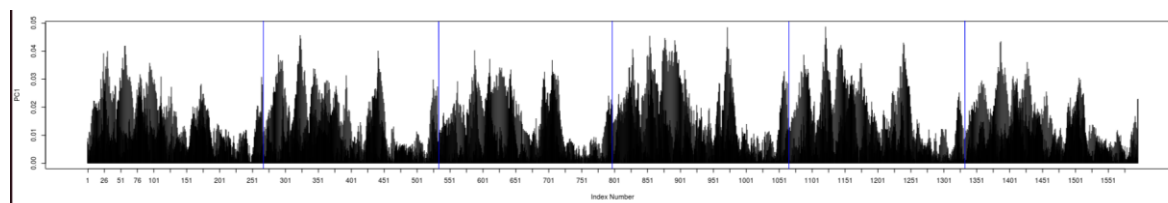


Gráfico 4.1 - Gráfico gerado pelo programa bio3d. Representa a flutuação média em Å de cada resíduo do hexâmero, considerando somente os carbonos α , para o movimento da primeira componente principal calculada para a transição entre as conformações R e T. As linhas azuis verticais indicam o início e o fim de cada monômero. Da esquerda para a direita estão mostradas as cadeias A, B, C, D, E e F.

O gráfico 4.1 mostra uma flutuação similar que se repete seis vezes para toda a estrutura. Essa repetição corresponde a cada protômero do hexâmero. O movimento para cada protômero, ainda que muito semelhante, não é idêntico. A sequência de picos é praticamente a mesma, mas eles possuem amplitudes diferentes. Ou seja, não foi descrito por esta técnica, um movimento simétrico entre as cadeias. É possível que isto seja uma consequência das diferentes simetrias nos cristais dos dois conformêros. O conformêro T, um monômero como unidade assimétrica, possui grupo espacial de simetria $P6_322$, enquanto o conformêro R, um dímero como unidade assimétrica, possui grupo espacial $H32$.

Observa-se ainda, no gráfico 4.1, que a região correspondente à tampa do sítio ativo, que compreende os resíduos de 163 a 183, e boa parte da região externa têm as mais altas amplitudes, e a região interna é bastante rígida. Para algumas cadeias, observam-se algumas peculiaridades. Por exemplo, as cadeias D e E apresentam o loop do resíduo 80 com uma flutuação maior que para as outras cadeias, enquanto a hélice α antes do C terminal se mostra mais flexível nas cadeias C e E. Essas regiões estão destacadas na figura 4.2.

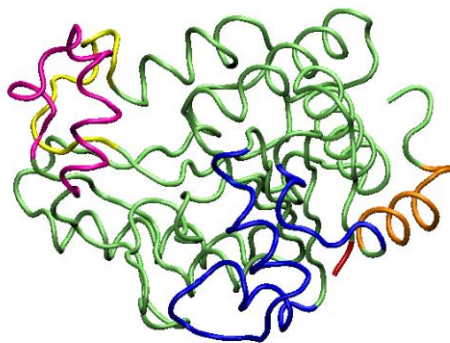


Figura 4.2 - Representação de um monômero da GlcN6PD com algumas regiões destacadas. Em rosa está a tampa do sítio ativo, que reúne os resíduos de 163 a 183; em amarelo está um loop que compõe o sítio ativo e engloba o resíduo 80; em laranja está a hélice do C terminal; em azul está a região interna; e em vermelho está ressaltado o resíduo N terminal.

A figura 4.3 mostra a cadeia A passando pela movimentação proposta pelo PCA para a transição entre os conformémeros T e R.

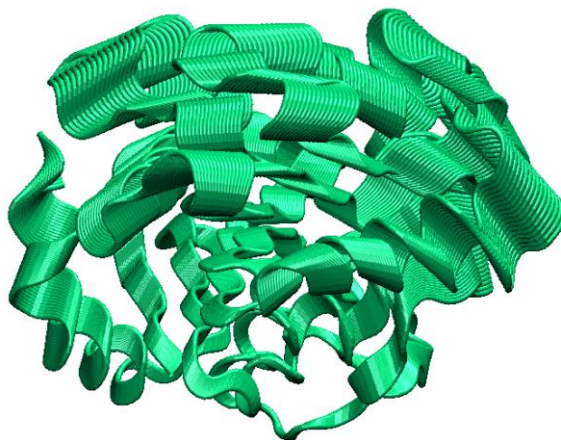


Figura 4.3- Superposição de quadros do movimento calculado como primeira componente principal da transição entre as conformações R e T. Apenas um monômero está representado, mas a transição foi calculada para o hexâmero e todas as outras cadeias mostram comportamento semelhante.

Mais uma vez, a figura 4.3 deixa claro que a movimentação maior se dá na região externa, como que numa movimentação em torno de um eixo paralelo ao z de simetria 6_3 . E o N terminal tem pouca flexibilidade, enquanto que a hélice do C terminal é bastante flexível.

Outra análise de flexibilidade estrutural calculada para esta transição foi o Parâmetro de Deslocamento Isotrópico, usando como base a trajetória gerada pelo PCA. Este dado consta no gráfico 4.2.

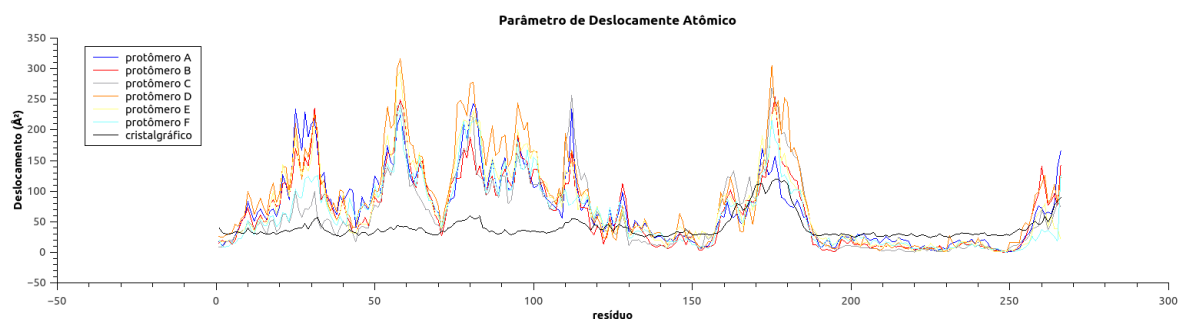


Gráfico 4.2 - Parâmetros de deslocamento atômico isotrópicos, ou fator de temperatura isotrópico, calculados para cada Carbono α . As seis linhas coloridas representam cada protômero na trajetória escrita para a transição direta entre as conformações T e R calculada por Análise de Componentes principais com o uso do programa bio3d. A linha preta representa o mesmo parâmetro encontrado por cristalografia para a estrutura de código 1fsf, um monômero na conformação T.

Ainda que seja um parâmetro muito semelhante ao de flutuações já mostrado, este possui a vantagem de poder ser comparado com um dado experimental, que é o fator de temperatura do conformero T. Comparando o deslocamento das cadeias durante a transição gerada com o dado cristalográfico, vê-se que a região interna de fato é que menos se movimenta em ambos os experimentos. As amplitudes, no entanto, apresentam-se muito maiores na simulação que no cristal, o que pode ser explicado pela natureza de cada técnica. Os átomos em um cristal têm sua flexibilidade diminuída por conta do empacotamento e da baixa temperatura. Enquanto um movimento gerado por PCA, possui uma amplitude irreal, e provoca deslocamentos em átomos que poderiam até ser inviáveis biologicamente. Ainda assim, em ambas as técnicas vê-se a diferença de flexibilidade entre a região externa e a interna. No entanto, a tampa do sítio ativo e a região externa possuem amplitudes semelhantes na simulação, mas muito diferentes no cristal. Provavelmente, a tampa possui uma movimentação independente da transição alostérica, e influencia tanto para esta quanto qualquer outra parte da região externa.

Uma vez identificado um movimento ideal que faria a transição do conformero T ao R, e vice-versa, foi feita uma dinâmica para avaliar a espontaneidade desse movimento. Essa dinâmica foi de um sistema simples, de proteína solvatada em água, de duração de 100ns, descrita na metodologia como a dinâmica 1. Essa dinâmica mostrou-se bastante estável durante toda a trajetória, tanto para a energia quanto para a pressão, como será mostrado mais adiante. Análises dela indicaram, no entanto, que a equilibração foi tão longa que, possivelmente, encobriu a transição do estado T inicial para o estado R. Isso fica claro na figura 4.4, que mostra o deslocamento angular do centro de massa da região externa de cada

protômero durante a equilibração de 400ps.

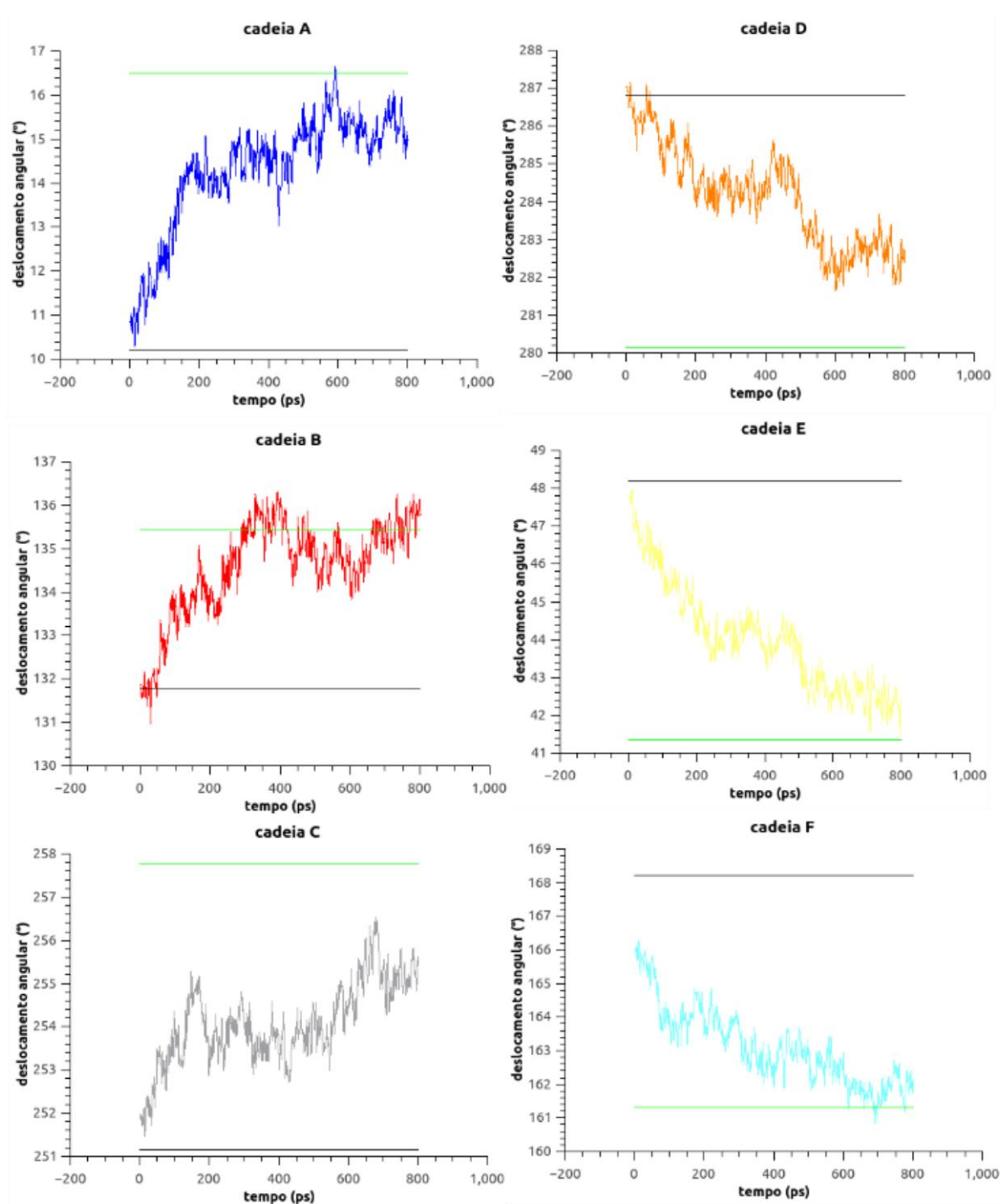


Figura 4.4 - Seis gráficos, um para cada cadeia do hexâmero, que mostram o deslocamento angular do centro de massa. As cadeias A, B e C formam um trîmero de simetria 3, assim como as cadeias D, E e F. Cálculo feito para a equilibração do sistema para a dinâmica 1. As linhas horizontais verde e preta representam o ângulo calculado para cada cadeia nas conformações R e T respectivamente.

Como essa equilibração mostrou-se mais que suficiente para estabilizar o sistema, como confirma o gráfico 4.3, outra dinâmica, a dinâmica 3, foi iniciada passando por uma equilibração de somente 20ps. Esta também se mostrou suficiente para estabilizar o sistema, como mostra o gráfico 4.4.

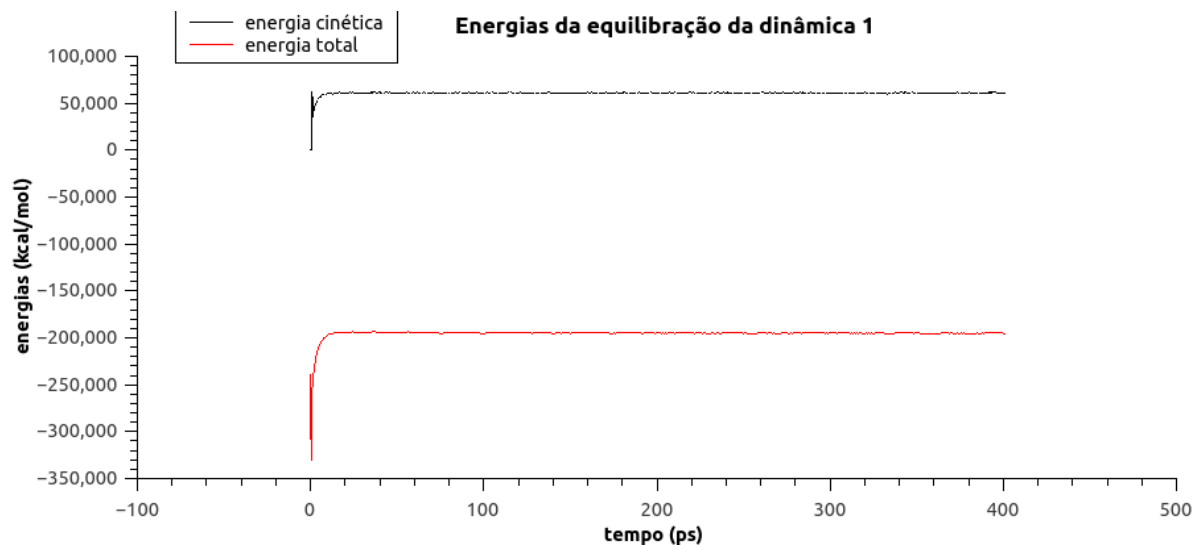


Gráfico 4.3 - Energias total (em vermelho) e cinética (em preto) do sistema, proteína em solvente, durante a equilibração da dinâmica 1, calculadas pelo programa NAMD.

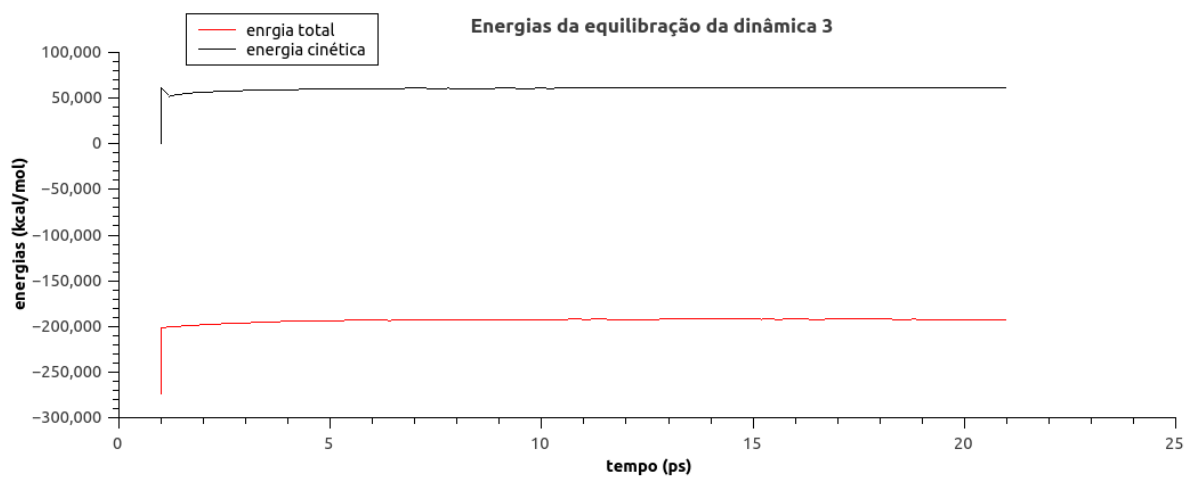


Gráfico 4.4 - Energias total (em vermelho) e cinética (em preto) do sistema, proteína em solvente, durante a equilibração da dinâmica 3, calculadas pelo programa NAMD.

Também essa dinâmica 3 se mostrou estável durante a trajetória, como mostram os gráficos 4.5 e 4.6.

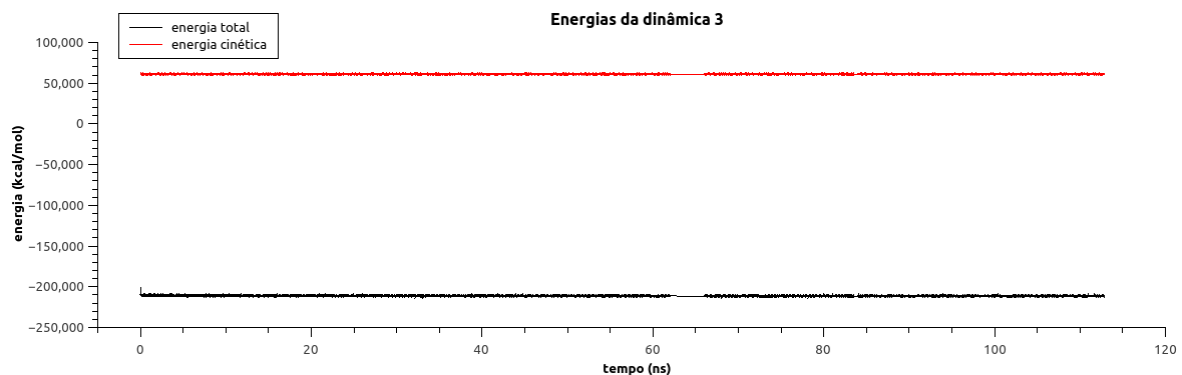


Gráfico 4.5 - Energias total (em vermelho) e cinética (em preto) do sistema, proteína em solvente, durante a dinâmica 3, calculadas pelo programa NAMD.

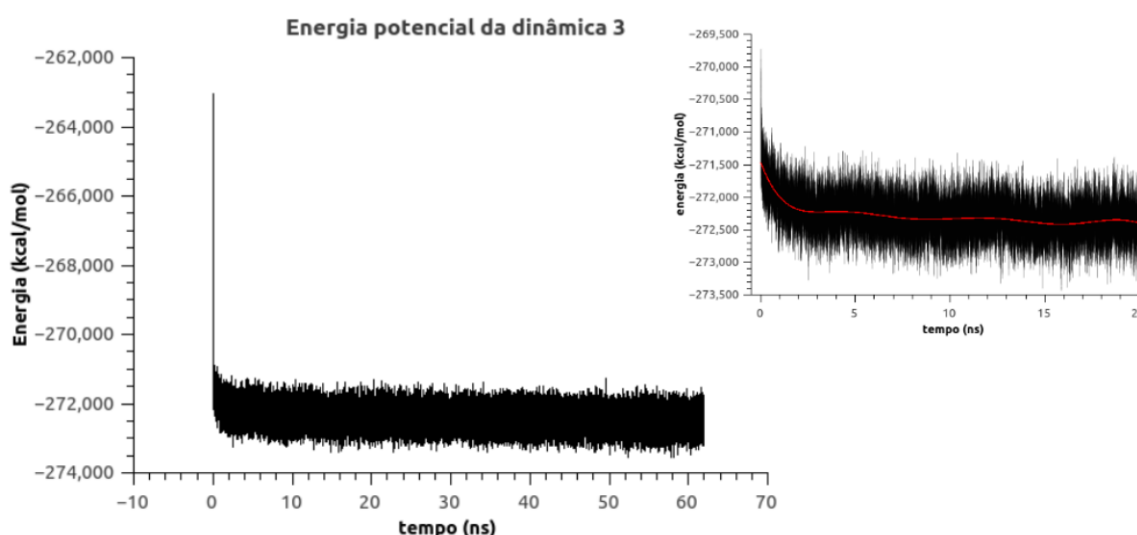


Gráfico 4.6 - Energia potencial do sistema durante a dinâmica 3, calculada como diferença das energias total e cinética. Em detalhe, no canto superior direito, está um gráfico em escala ampliada para o início da trajetória (de 0,05 a 20 ns). Em vermelho está representada a melhor curva ajustada para um polinômio de grau 9.

O gráfico 4.6 mostra que também a energia potencial permaneceu estável ao longo da dinâmica. O início dela, no entanto, é marcado por uma queda acentuada na energia que só termina próxima a 3ns, quando começa a oscilar com uma amplitude de 100 kcal/mol e uma frequência de 7ns, aproximadamente, como mostra a curva em vermelho, acompanhada de oscilações mais rápidas de amplitude 1Mcal/mol, que compõem a faixa preta. Mais exatamente, a energia potencial inicia em -263,1 Mcal/mol, cai para -271,8 Mcal/mol depois de 500ps, aos 3ns está em -272,2 Mcal e a oscilação se estabiliza em torno de -272,4 Mcal/mol depois de 10 ns, seguindo a média calculada como a curva vermelha.

Também a pressão do sistema permaneceu estável. Os 10 primeiros nano segundos estavam em -600bar com um desvio padrão de 124.73, enquanto os dez últimos nano

segundos ainda estavam em -616.6 bar com um desvio padrão de 122.94.

Apenas as estabilidades na energia e na pressão da dinâmica já conferem alguma confiabilidade à simulação. Ainda assim, a melhor forma de confirmar que a simulação representa uma situação real é se ela for capaz de reproduzir dados experimentais. Um parâmetro já medido experimentalmente para o conformero T, o mesmo usado para esta simulação, é o parâmetro de deslocamento atômico. O gráfico 4.7 compara o dado cristalográfico com o calculado para esta dinâmica, como uma média do deslocamento do carbono α de cada resíduo durante a trajetória simulada.

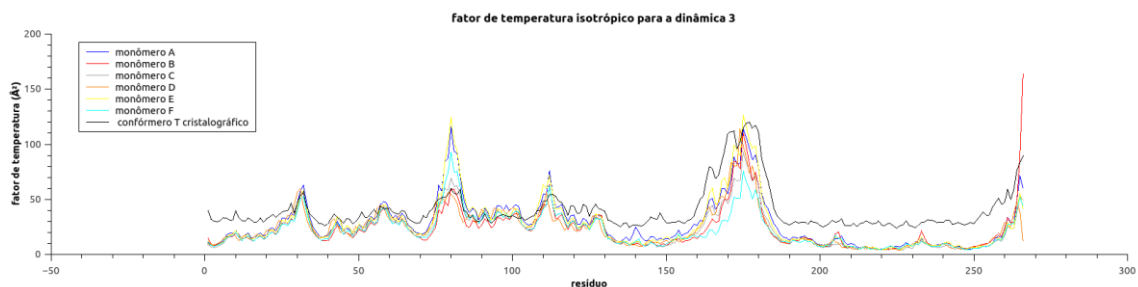


Gráfico 4.7 - Parâmetros de deslocamento atômico isotrópicos, ou fator de temperatura isotrópico, calculados para cada Carbono α . As seis linhas coloridas representam cada protômero na trajetória escrita para a dinâmica 3, calculada pelo programa NAMD. A linha preta representa o mesmo parâmetro encontrado por cristalografia para a estrutura de código 1fsf, um monômero na conformação T.

Os parâmetros de deslocamento atômico, mostrados no gráfico 4.7, já possuem uma amplitude mais compatível com o dado experimental que os calculados para a transição direta entre os conformeros, vistos no gráfico 4.2. Ainda assim, comparando as amplitudes da região externa com as da região interna, há uma diferença maior entre elas na dinâmica que na cristalografia. Talvez a própria dinâmica já permita maior flexibilidade, por simular uma temperatura maior, e é possível que o alinhamento local, feito sempre somente com a região interna como base, esteja influenciando para ressaltar o movimento da região externa. É interessante também que, pelo gráfico 4.7, três cadeias – A, E e F – apresentam o loop por volta do resíduo 80 tão flexível quanto a tampa do sítio ativo, característica que se assemelha à transição ideal construída por PCA.

Assim como o fator de temperatura isotrópico, outro parâmetro medido

experimentalmente é o deslocamento atômico anisotrópico. A medida deste parâmetro por difração de raios-X só é possível quando a estrutura resolvida possui uma resolução alta, e por isso existe somente para o conformero T, resolvido a $1,9\text{\AA}$, e mesmo assim carrega algum erro que a uma resolução maior seria menor. Não obstante, já possui maiores informações sobre a direção do movimento que somente o parâmetro isotrópico. A figura 4.5 mostra os parâmetros anisotrópicos medidos experimentalmente.

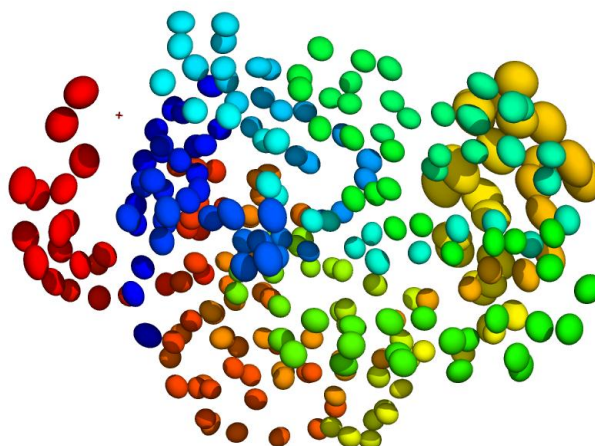


Figura 4.5 - Representação dos parâmetros de deslocamento anisotrópicos para cada carbono α como um elipsoide. Os valores mostrados são referentes ao da estrutura de um monômero da GlcN6PD na conformação T, resolvida por cristalografia com resolução de $1,9\text{\AA}$ de código pdb 1fsf. Figura feita com o programa Pymol, colorida na forma *rainbow*, em que o C terminal está em vermelho e o N terminal em azul.

Já a figura 4.6 mostra os calculados para cada cadeia como uma média do deslocamento atômico ao longo da trajetória.

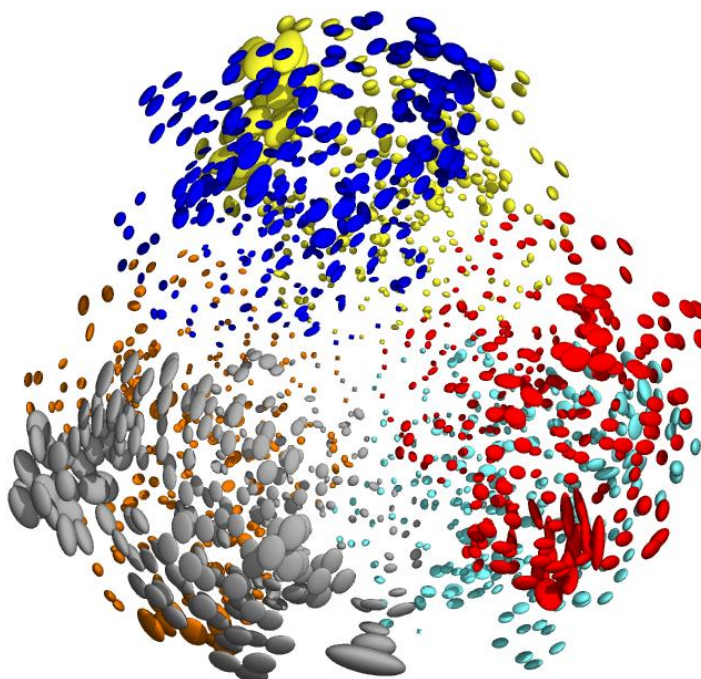


Figura 4.6 - Parâmetros de deslocamento anisotrópicos representados em elipsoides para cada carbono α . Eles foram calculados como uma média do deslocamento durante a trajetória da dinâmica 3. Figura desenhada no programa Pymol. Em azul está identificada a cadeia A, em vermelho a cadeia B, em cinza C, em laranja a D, em amarelo a E e em ciano a cadeia F.

Comparando-se as figuras 4.5 e 4.6, pode-se dizer que a medida cristalográfica possui elipsoides que se aproximam mais a esferas que os calculados para a dinâmica, e na dinâmica percebe-se uma direção mais acentuada para o movimento. Essa direção é justamente a de rotação da região externa. Mais uma vez é evidente a rigidez da região interna, mais ressaltada na dinâmica que no cristal. Também fica clara a movimentação da tampa do sítio ativo, com uma direção ligeiramente diferente do restante da região externa, com uma inclinação um pouco voltada para o centro do hexâmero.

Para quantificar a rotação que a região externa faz durante a dinâmica foi calculado o deslocamento angular do centro de massa dela em relação ao eixo x. A figura 4.7 mostra 6 gráficos desse deslocamento, um para cada cadeia, durante a trajetória.

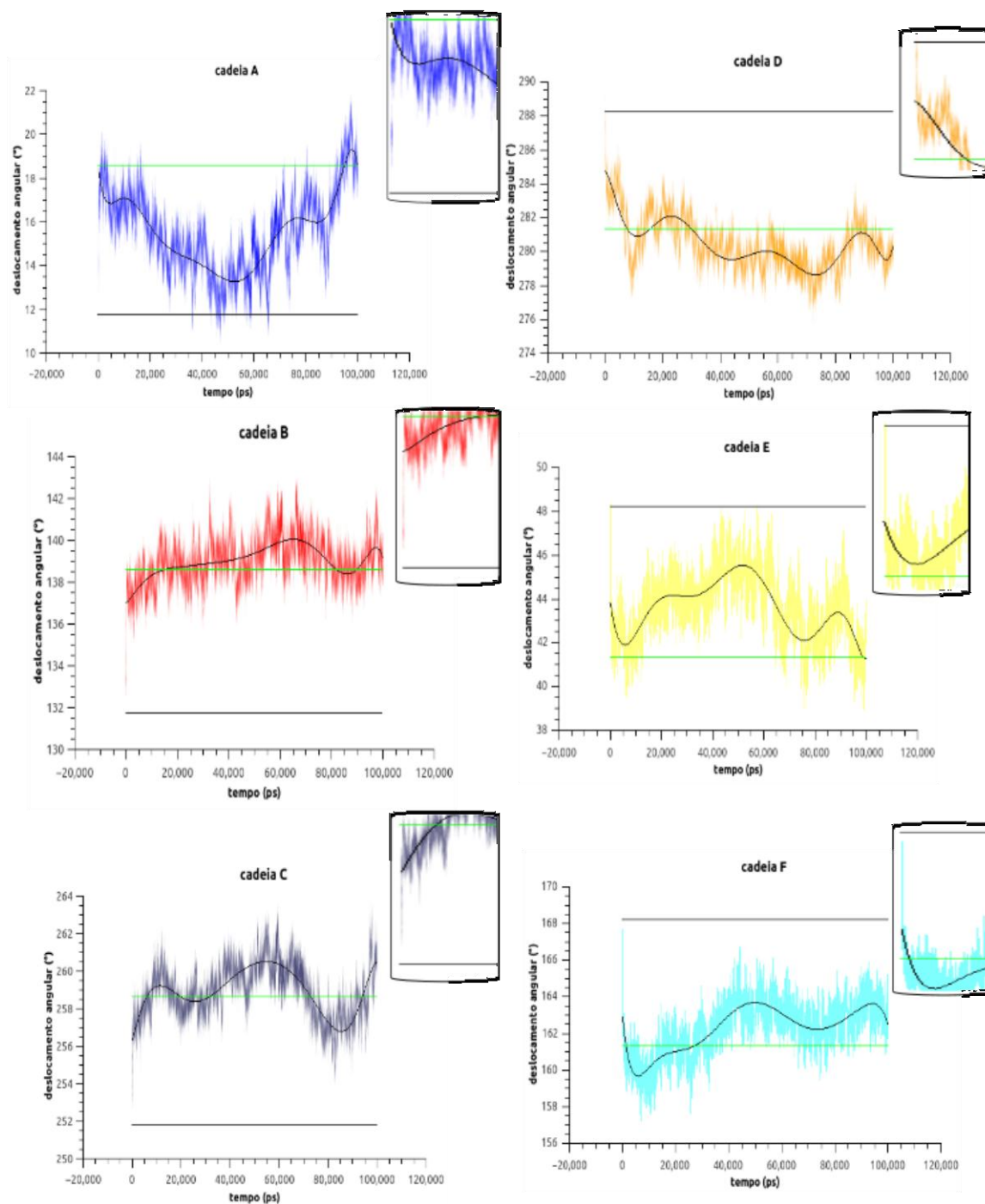


Figura 4.7 - Seis gráficos, um para cada cadeia do hexâmero, que mostram o deslocamento angular do centro de massa de região externa. Dados calculados para a trajetória da dinâmica 3. As cadeias A, B e C formam um trîmero de simetria 3, assim como as cadeias D, E e F. Em detalhe, no canto superior direito de cada gráfico está uma figura aumentada do início da trajetória, ratificando que a dinâmica se iniciou na conformação T, e após breve período de equilíbrio ainda se encontrava nela. As linhas horizontais verde e preta representam o ângulo calculado para cada cadeia nas conformações R e T respectivamente. A curva preta é uma função polinomial de grau 9 ajustada para os pontos do gráfico.

O deslocamento angular mostrado na figura 4.7 é o ângulo formado pelo centro de massa e o eixo x, representado em função do tempo e comparado com os calculados para as estruturas resolvidas por cristalografia dos conformémeros T e R, representados pelas linhas horizontais em preto e verde respectivamente. Vê-se que para todas as cadeias o centro de massa realmente passa por um deslocamento nesse sentido, na direção do movimento de transição alostérica. No entanto, esse deslocamento é bastante diferente para cada cadeia. Por exemplo, a cadeia A mostra deslocar-se entre os dois estados e parece alcançar a amplitude do deslocamento necessária para a transição. Enquanto que a cadeia B parece ter sua conformação estabilizada próxima ao estado R, e não possui uma amplitude de deslocamento muito alta. Isso sugere uma falta de simetria na movimentação.

Os gráficos também mostram que os centros de massa não se limitam a uma rotação restrita pelo ângulo entre a conformação T e a R, ou seja, os centros de massa não oscilam somente no espaço delimitado entre as linhas verde e preta, mas se movimentam um pouco além dela. Isso é bastante coerente com o conceito de que, em cristalografia, a estrutura resolvida é uma média num ensemble. Enquanto na dinâmica seria possível alcançar conformações com características mais extremas que a média, nesse caso, passar do ângulo encontrado para o centro de massa.

Ainda vê-se pela figura 4.7, em todos os quadros em detalhe do início da dinâmica, que todos os monômeros partem de uma conformação próxima da T e rapidamente mudam para a R. Isso mostra não só que a curta equilibração serviu seu propósito de detectar esta mudança durante a dinâmica, mas também que é uma mudança independente dos parâmetros da equilibração, visto que ocorreu também rapidamente para a dinâmica.

Ademais, ainda que cada protômero em particular mostre movimentar-se na direção da transição alostérica, era previsto que um movimento concertado possibilitasse a formação do sítio alostérico capaz de prender um ligante. Uma análise para isso é a distância entre resíduos desse sítio. Especificamente entre o N terminal de uma cadeia e a arginina 158 e a lisina 160 de outra. A figura 4.8 mostra 6 gráficos para cada sítio alostérico do hexâmero.

Da figura 4.8, vê-se que, a exceção do final da dinâmica com o N terminal da cadeia A, a arginina e a lisina seguem o mesmo padrão de movimentação, envolvendo no mínimo a estrutura secundária e não somente uma movimentação local de resíduos individuais, o que

sugere uma movimentação global da estrutura. Também foi uma exceção da dinâmica com o

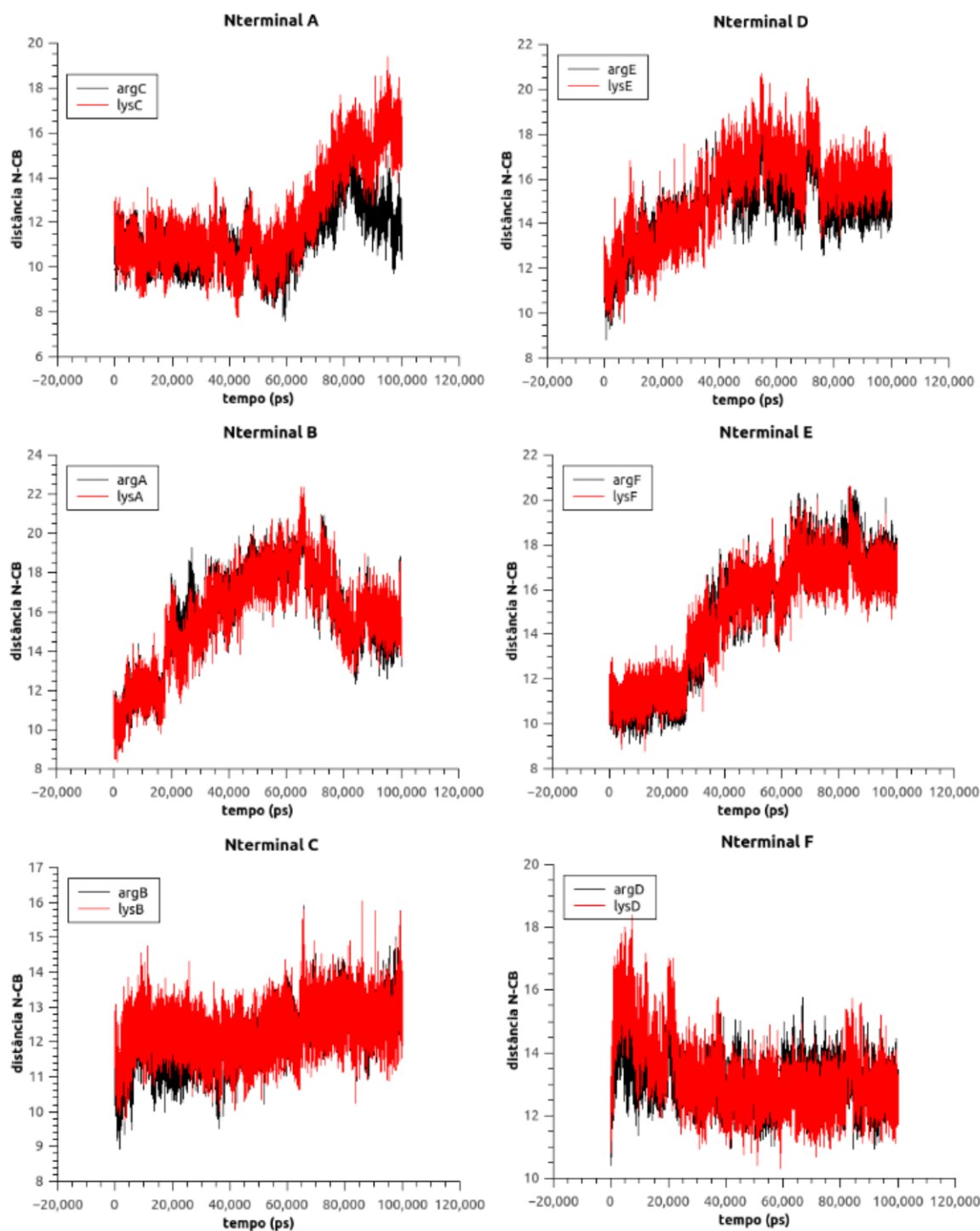


Figura 4.8 - Seis gráficos, um para cada sítio alostérico do hexâmero, que mostram o comprimento da ligação entre o nitrogênio da metionina N terminal e o carbono β da arginina 158 em vermelho, e entre o mesmo nitrogênio da metionina e o carbono β da lisina 160 em preto. Como os sítios alostéricos são formados entre os protômeros, a metionina se encontra em uma cadeia, identificada no título de cada gráfico, enquanto as arginina e lisina pertencem à outra cadeia identificada na legenda de cada gráfico. Cálculo feito para a dinâmica 3.

N terminal da cadeia A, quando comparada com as outras, ter realmente alcançado as distâncias compatíveis com o conformero R. Talvez por tratar-se de um movimento que necessite a colaboração de duas cadeias, na maioria do tempo, para todos os sítios, a conformação mais estável parece ser a do conformero T, em que os resíduos estão mais distantes. Ainda assim, a movimentação da arginina e da lisina parece ser a mais determinante na formação do sítio, enquanto o N terminal parece ser menos flexível. Isso é mais evidente para a dinâmica do N terminal da cadeia B, gráfico cujo padrão se assemelha ao do deslocamento de centro de massa da cadeia A, mesma cadeia da arginina e da lisina do sítio alostérico em questão. Também os gráficos dos N terminais das cadeias C e F se assemelham a movimentação do centro de massa das cadeias B e D. Cadeias cujas arginina e lisina se mantêm mais próximas de um estado R, quando comparadas com os outros gráficos, durante toda a dinâmica.

Também é ressaltado por essa análise que esses três resíduos de fato se distanciam muito mais que o previsto para a conformação T, e alcançam distâncias entre si de mais de 18 Å em três dos sítios alostéricos.

Mesmo trazendo indícios de movimentos globais na dinâmica, esse tipo de análise ainda carrega muito ruído e infere sobre uma característica muito local, que é o sítio alostérico. Uma análise que pretende filtrar ruídos e de fato caracterizar mudanças globais no decorrer da trajetória é a de tensores inércia. Ela foi feita para a região externa de cada protômero durante a dinâmica, e seus resultados estão a seguir organizados para cada cadeia de forma separada. Cada conjunto de análises é formado por uma figura dos eixos principais do tensor de inércia e seu respectivo protômero; um gráfico da variação de amplitude (autovalor) de cada autovetor do tensor; e três pares de gráficos da projeção dos autovetores num plano perpendicular à média dos autovetores ao longo do tempo, projeção que representa a variação na direção dos autovetores. Essa variação dos autovetores, tanto na direção quanto na amplitude, quando projetada no plano é capaz de separar grupos que teriam conformações semelhantes, caracterizando distintas populações visitadas durante a dinâmica.

Protômero A

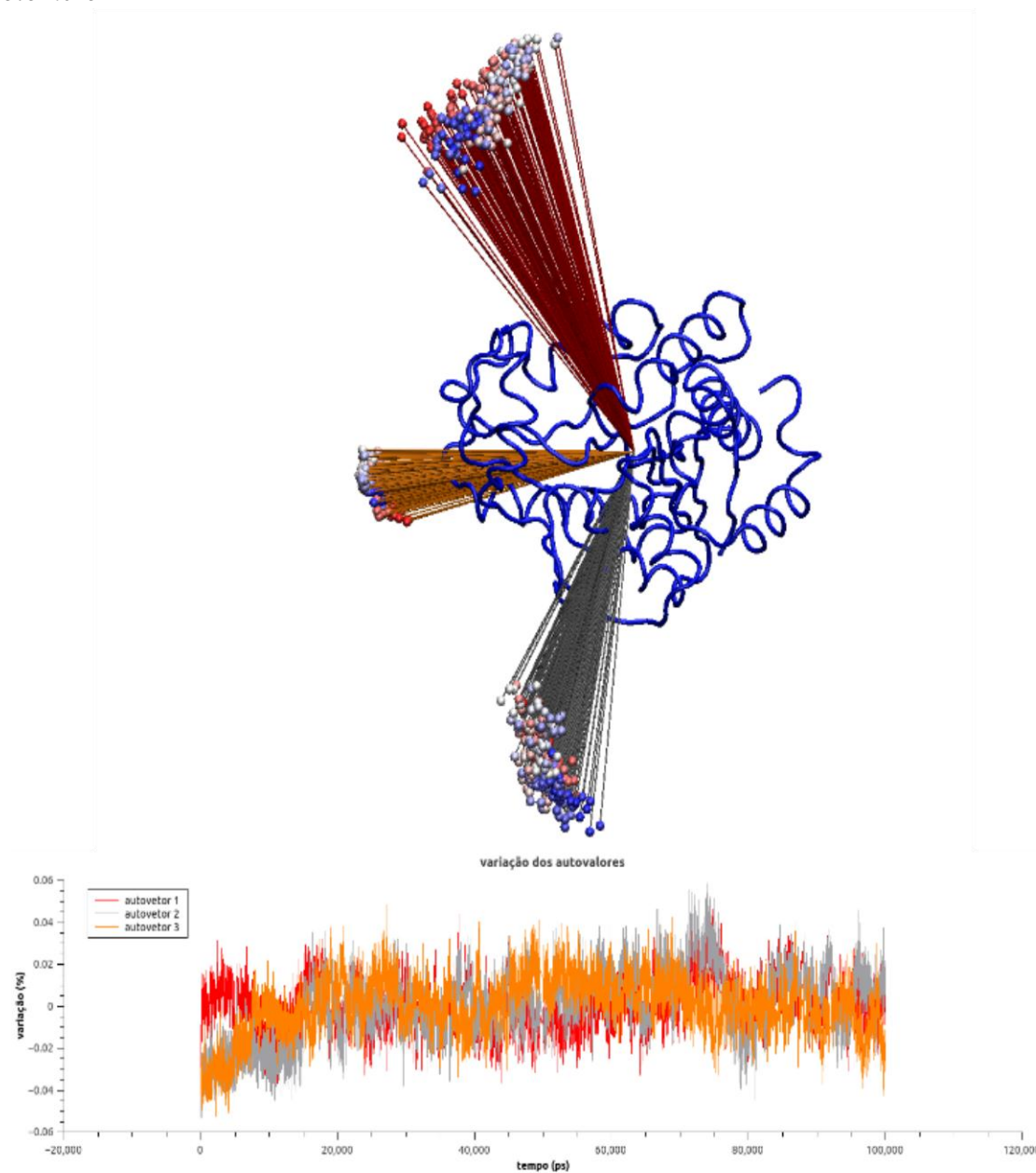


Figura 4.9 - acima, um desenho (feito com o programa VMD) do protômero A completo, visto com o eixo z perpendicular ao plano da figura, acompanhado dos eixos principais do tensor de inércia calculados para a região externa dessa cadeia. Cada conjunto de três eixos representa um tensor médio para cada 500ps da trajetória, calculada para a dinâmica 3. Todos eles têm a origem centrada no centro de massa da região externa do respectivo protômero na conformação T. Cada conjunto possui três pequenas esferas, cada uma na ponta de um vetor. Elas estão coloridas de acordo com o tempo na trajetória que representam, tempos iniciais estão em vermelho, medianos em branco e passos finais estão coloridos em azul. O gráfico de baixo mostra a variação em porcentagem dos autovalores respectivos a seus autovetores. As cores estão de acordo com a figura acima, o autovetor desenhado como um eixo vermelho tem a variação de sua amplitude (seu autovalor), em função do tempo, graficada na cor vermelha, e assim sucessivamente. Dez mil pontos compõem a linha graficada, sendo cada um a média de cada 10ps na trajetória calculada.

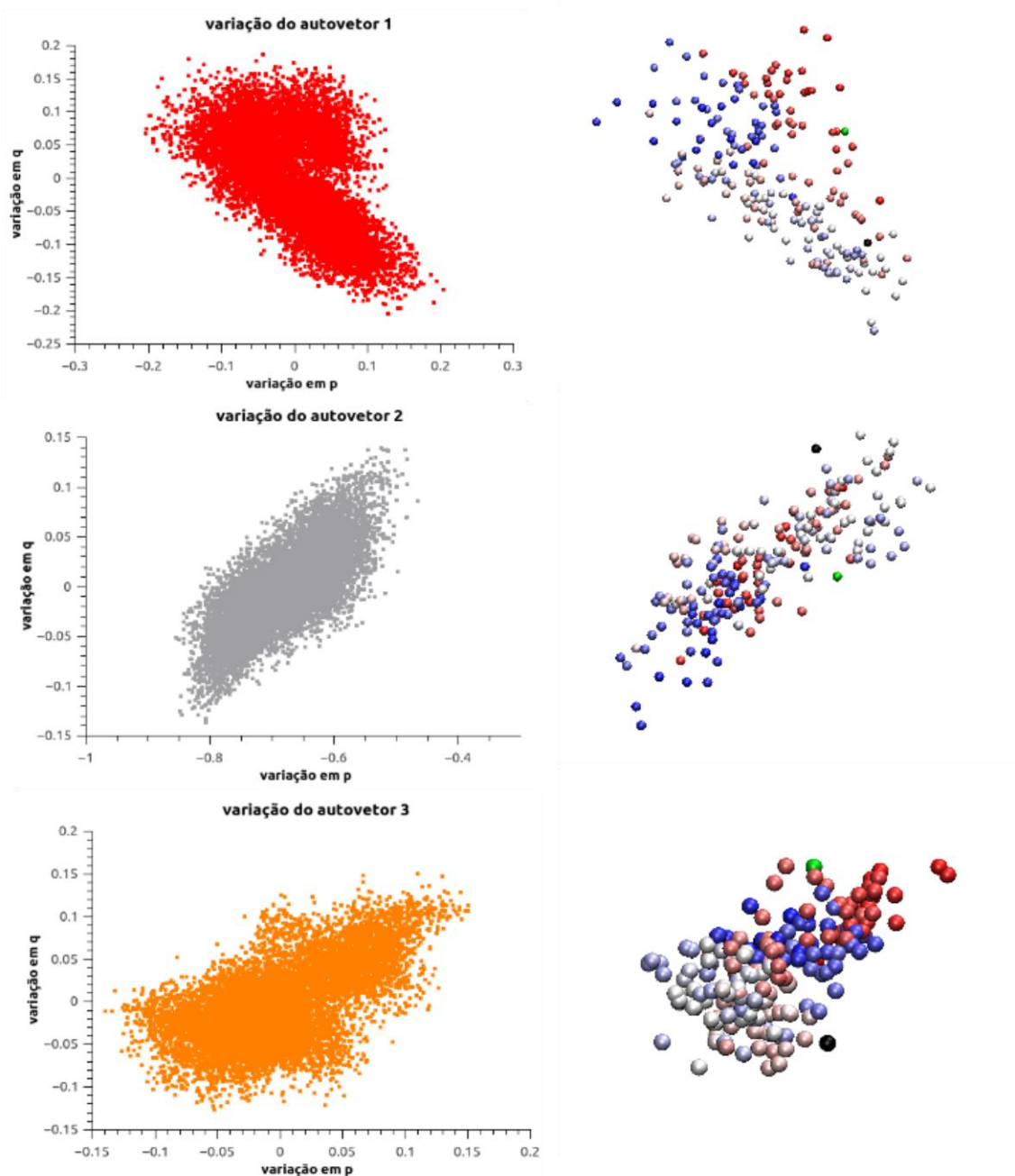


Figura 4.10 - Os três gráficos à esquerda correspondem à variação na direção dos três eixos do tensor de inércia, projetados cada qual a um plano perpendicular à média de cada eixo, sendo estes planos definidos por dois novos versores, p e q , diferentes para cada eixo. As cores dos pontos nos gráficos correspondem às cores dos eixos desenhados na figura anterior. Dez mil pontos compõem o gráfico, sendo cada um a média de cada 10ps na trajetória calculada. As figuras a direita são imagens feitas no programa VMD, mostrando somente as pequenas esferas das pontas dos eixos do tensor de inércia mostradas na figura anterior. Elas estão coloridas de acordo com sua posição no tempo, sendo cada uma uma média de cada 500ps da trajetória, tempos iniciais estão em vermelho, intermediários em branco e finais em azul. As duas esferas, preta e verde, que acompanham esta figura representam a direção correspondente deste eixo para as conformações T e R, respectivamente, com respeito ao protômero A.

Protômero B

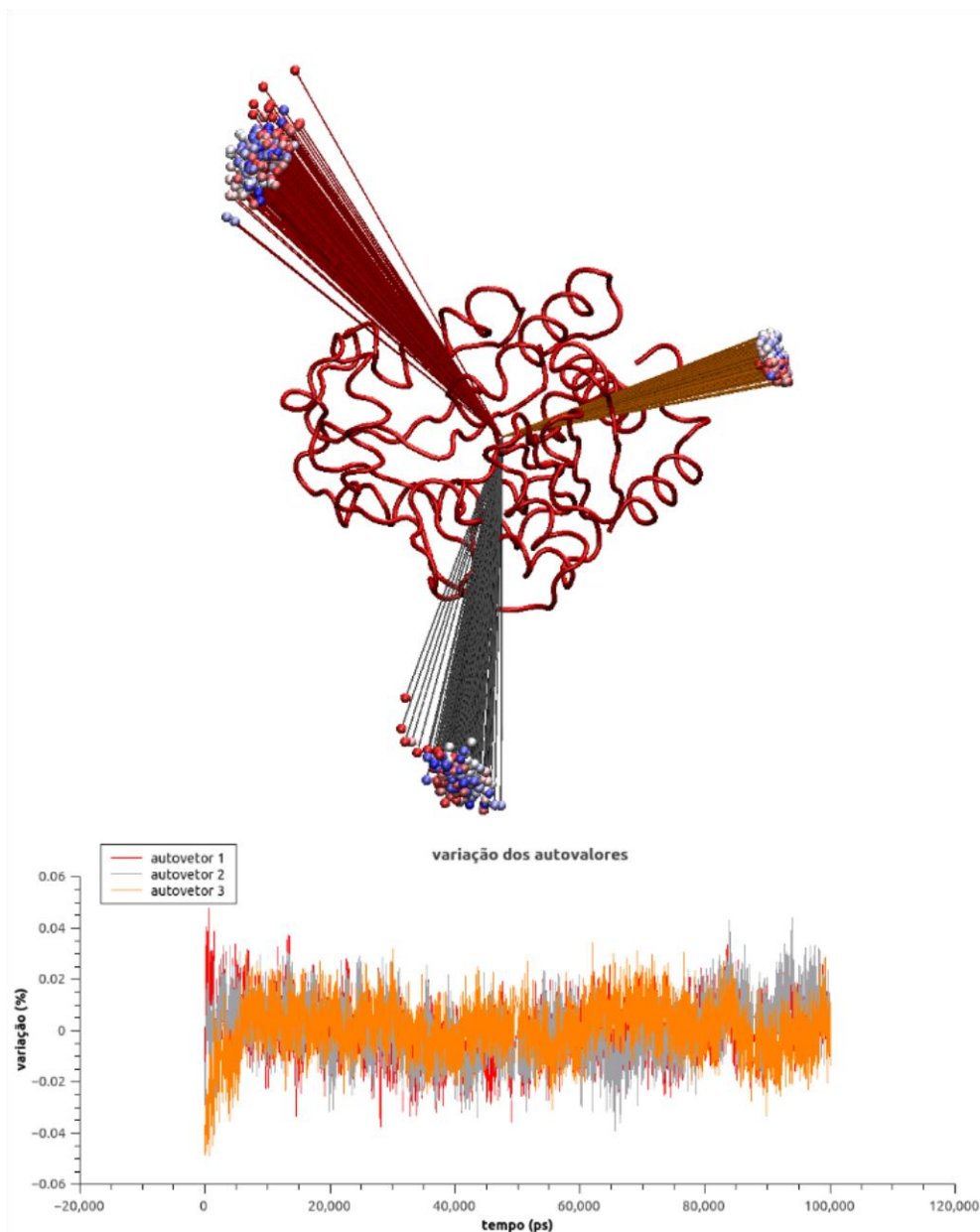


Figura 4.11 - acima, um desenho (feito com o programa VMD) do protômero B completo, visto com o eixo z perpendicular ao plano da figura, acompanhado dos eixos principais do tensor de inércia calculados para a região externa dessa cadeia. Cada conjunto de três eixos representa um tensor médio para cada 500ps da trajetória, calculada para a dinâmica 3. Todos eles têm a origem centrada no centro de massa da região externa do respectivo protômero na conformação T. Cada conjunto possui três pequenas esferas, cada uma na ponta de um vetor. Elas estão coloridas de acordo com o tempo na trajetória que representam. Tempos iniciais estão em vermelho, medianos em branco e passos finais estão coloridos em azul. O gráfico de baixo mostra a variação em porcentagem dos autovalores respectivos a seus autovetores. As cores estão de acordo com a figura acima, o autovetor desenhado como um eixo vermelho tem a variação de sua amplitude (seu autovalor), em função do tempo, graficada na cor vermelha, e assim sucessivamente. Dez mil pontos compõem a linha graficada, sendo cada um a média de cada 10ps na trajetória calculada.

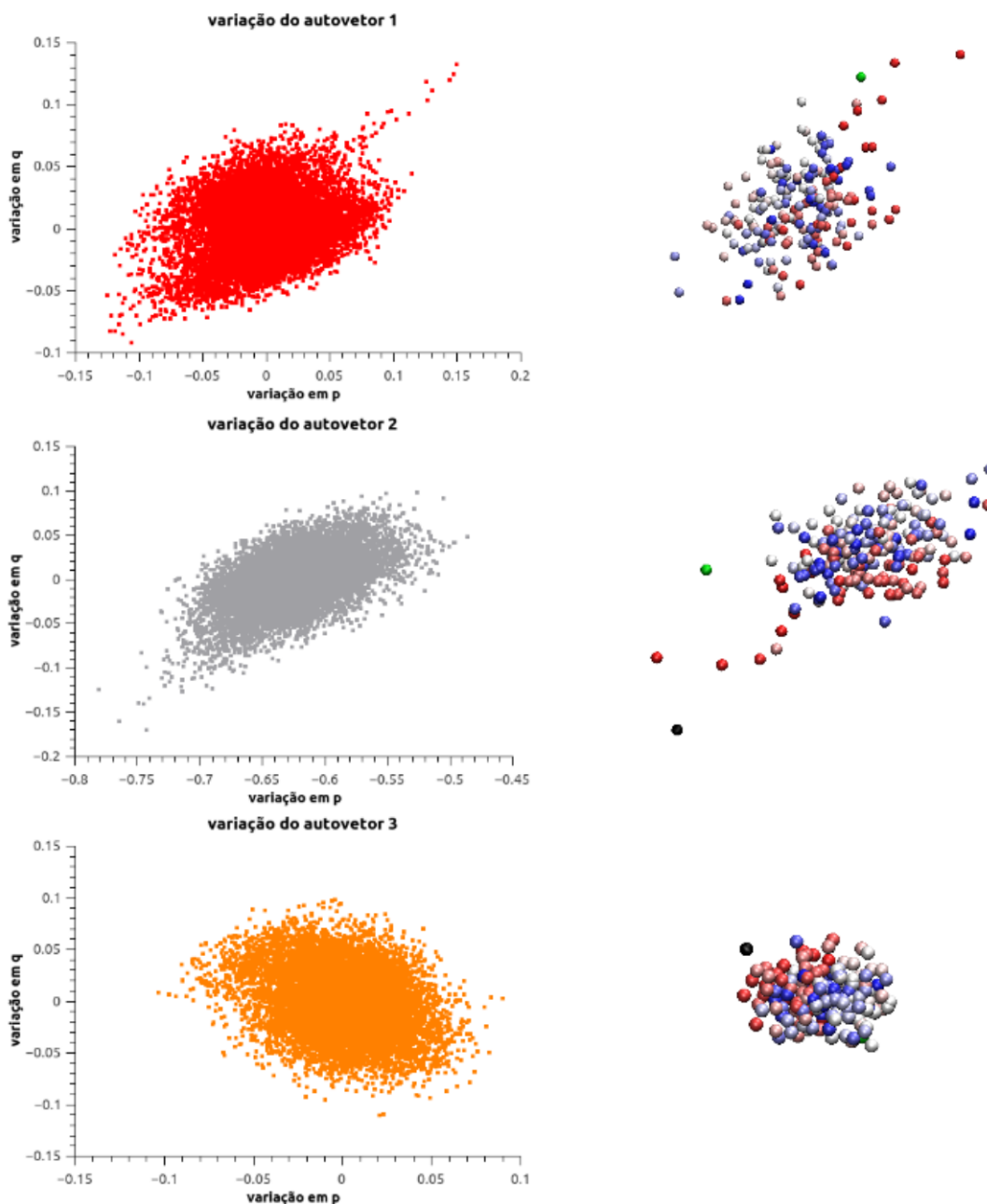


Figura 4.12 - Os três gráficos à esquerda correspondem à variação na direção dos três eixos do tensor de inércia, projetados cada qual a um plano perpendicular à média de cada eixo, sendo estes planos definidos por dois novos versores, p e q , diferentes para cada eixo. As cores dos pontos nos gráficos correspondem às cores dos eixos desenhados na figura anterior. Dez mil pontos compõem o gráfico, sendo cada um a média de cada 10ps na trajetória calculada. As figuras a direita são imagens feitas no programa VMD, mostrando somente as pequenas esferas das pontas dos eixos do tensor de inércia mostradas na figura anterior. Elas estão coloridas de acordo com sua posição no tempo, sendo cada uma uma média de cada 500ps da trajetória. Tempos iniciais estão em vermelho, intermediários em branco e finais em azul. As duas esferas, preta e verde, que acompanham esta figura representam a direção correspondente deste eixo para as conformações T e R, respectivamente, com respeito ao protômero B.

Protômero C

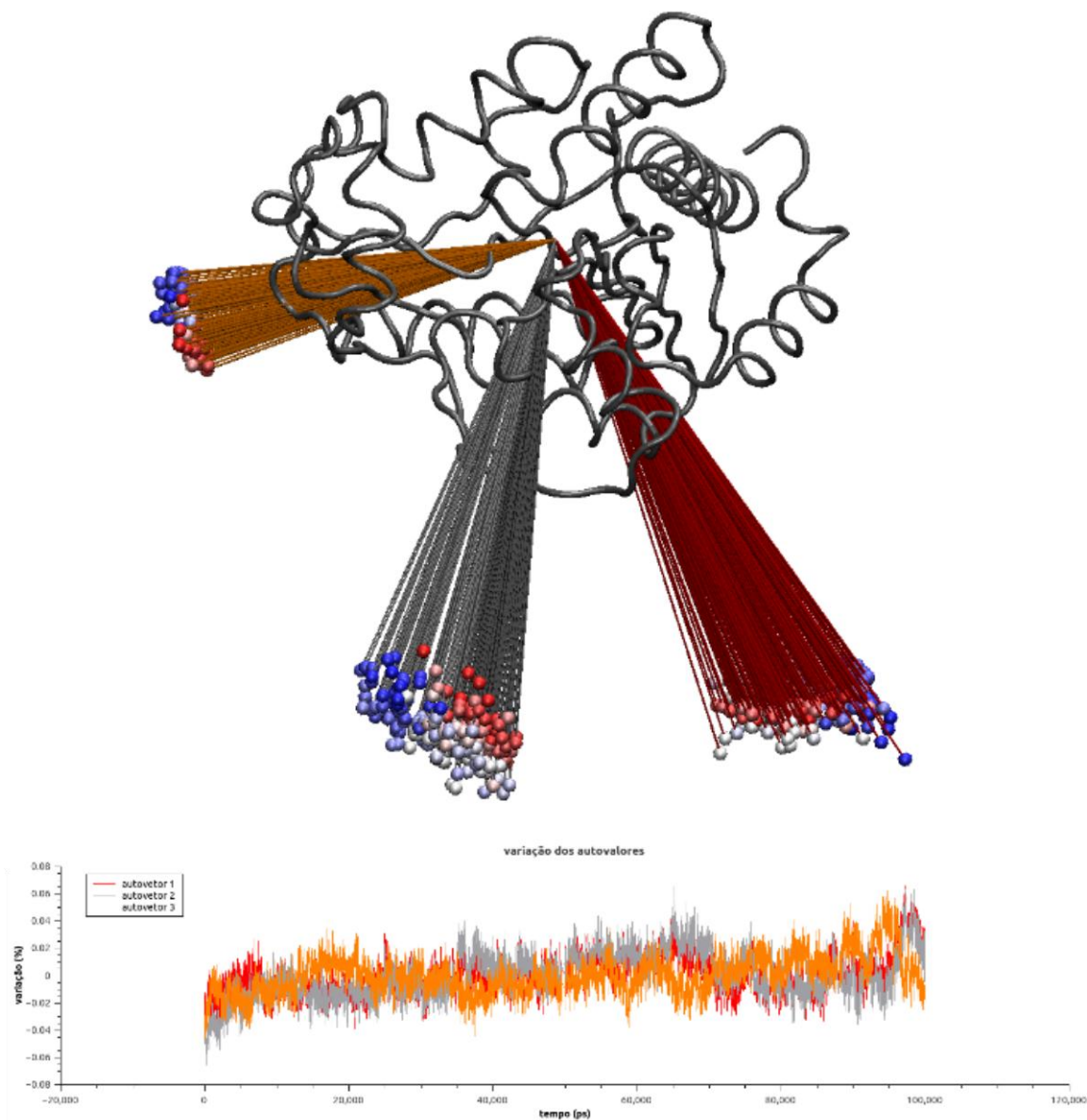


Figura 4.13 - acima, um desenho (feito com o programa VMD) do protômero C completo, visto com o eixo z perpendicular ao plano da figura, acompanhado dos eixos principais do tensor de inércia calculados para a região externa dessa cadeia. Cada conjunto de três eixos representa um tensor médio para cada 500ps da trajetória, calculada para a dinâmica 3. Todos eles têm a origem centrada no centro de massa da região externa do respectivo protômero na conformação T. Cada conjunto possui três pequenas esferas, cada uma na ponta de um vetor. Elas estão coloridas de acordo com o tempo na trajetória que representam. Tempos iniciais estão em vermelho, medianos em branco e passos finais estão coloridos em azul. O gráfico de baixo mostra a variação em porcentagem dos autovalores respectivos a seus autovetores. As cores estão de acordo com a figura acima, o autovetor desenhado como um eixo vermelho tem a variação de sua amplitude (seu autovalor), em função do tempo, graficada na cor vermelha, e assim sucessivamente. Dez mil pontos compõem a linha graficada, sendo cada um a média de cada 10ps na trajetória calculada.

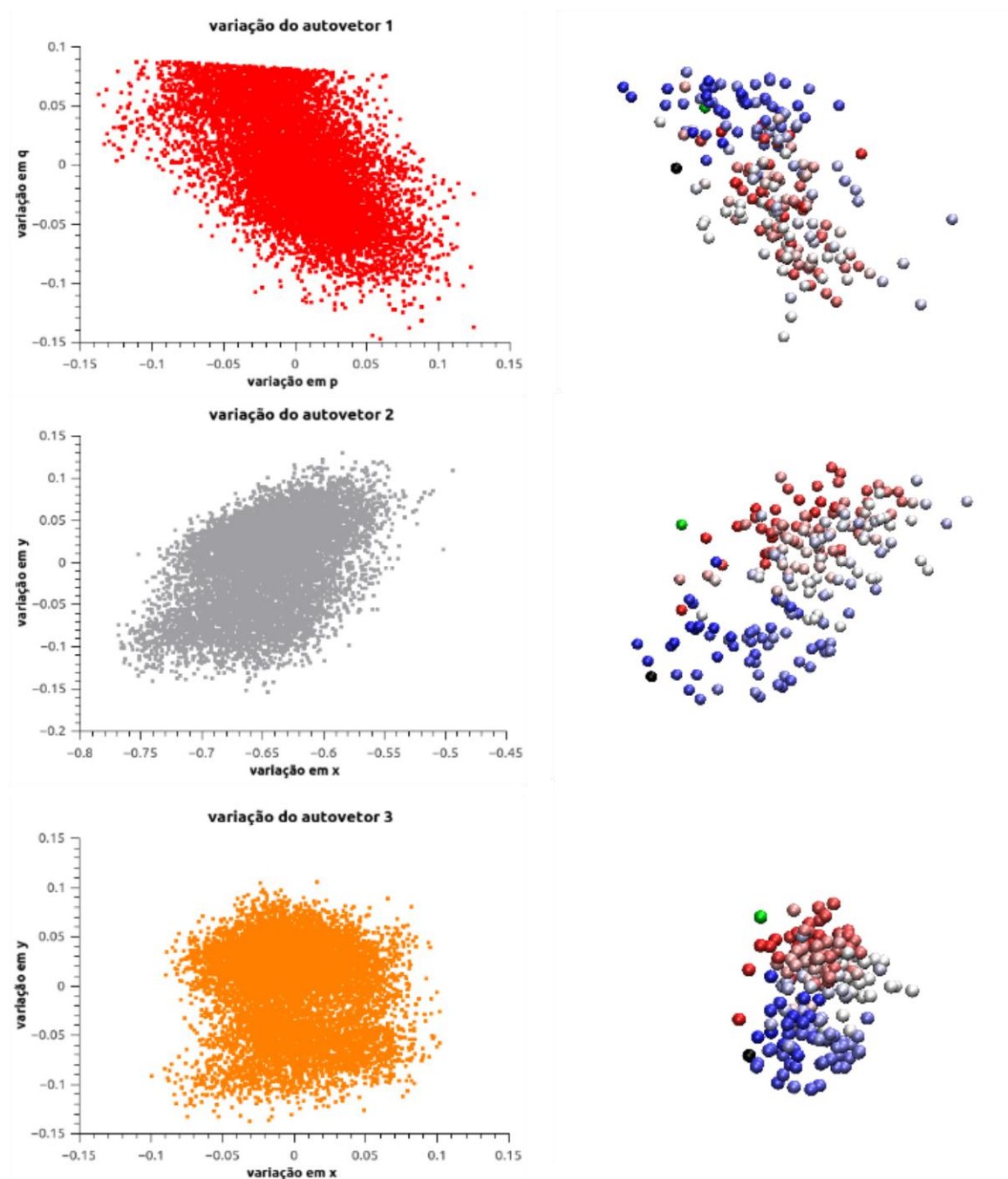


Figura 4.14 - Os três gráficos à esquerda correspondem à variação na direção dos três eixos do tensor de inércia, projetados cada qual a um plano perpendicular à média de cada eixo, sendo estes planos definidos por dois novos versores, p e q , diferentes para cada eixo. As cores dos pontos nos gráficos correspondem às cores dos eixos desenhados na figura anterior. Dez mil pontos compõem o gráfico, sendo cada um a média de cada 10ps na trajetória calculada. As figuras a direita são imagens feitas no programa VMD, mostrando somente as pequenas esferas das pontas dos eixos do tensor de inércia mostradas na figura anterior. Elas estão coloridas de acordo com sua posição no tempo, sendo cada uma uma média de cada 500ps da trajetória. Tempos iniciais estão em vermelho, intermediários em branco e finais em azul. As duas esferas, preta e verde, que acompanham esta figura representam a direção correspondente deste eixo para as conformações T e R, respectivamente, com respeito ao prótômero C.

Protômero D

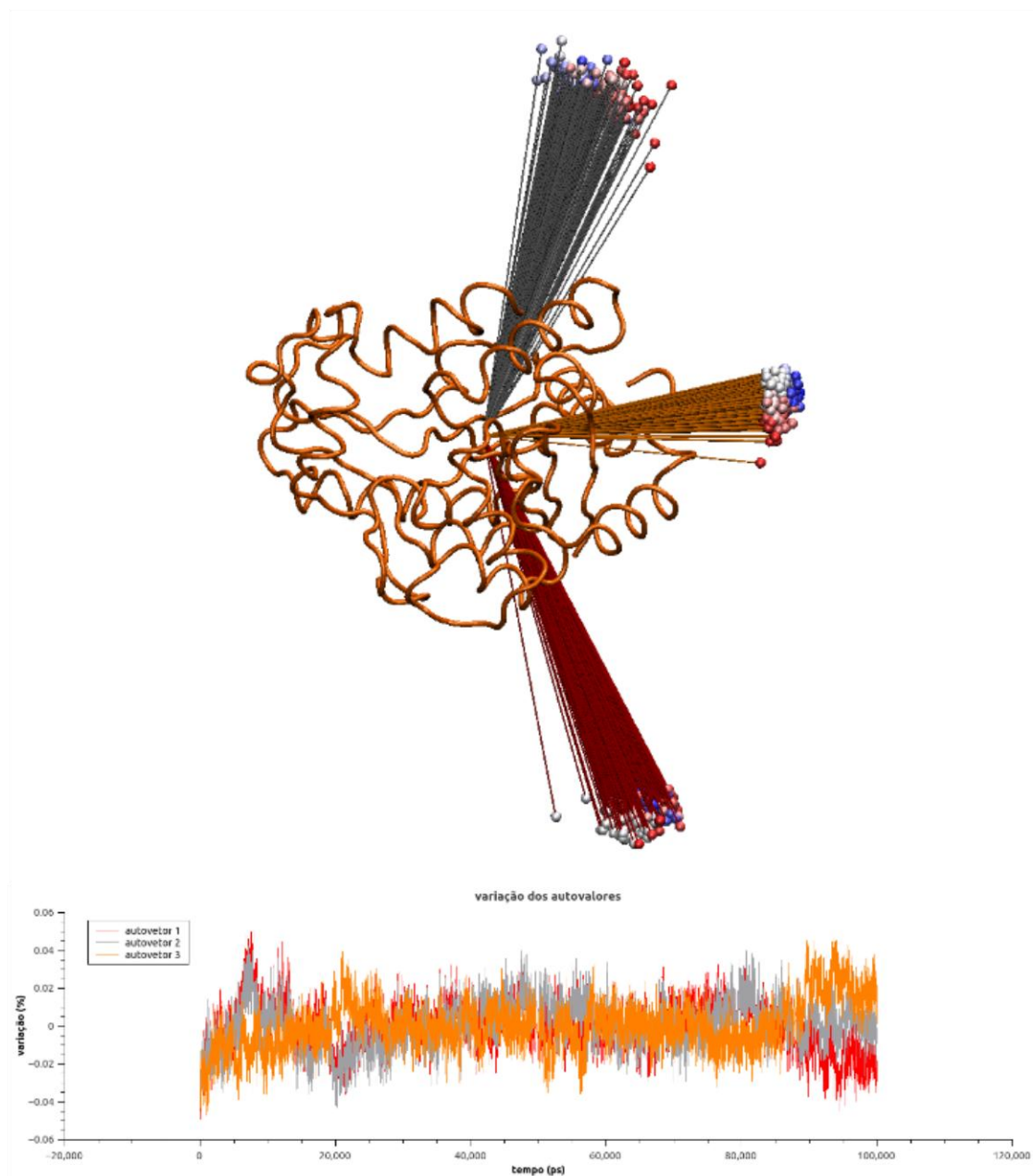


Figura 4.15 - acima, um desenho (feito com o programa VMD) do protômero D completo, visto com o eixo z perpendicular ao plano da figura, acompanhado dos eixos principais do tensor de inércia calculados para a região externa dessa cadeia. Cada conjunto de três eixos representa um tensor médio para cada 500ps da trajetória, calculada para a dinâmica 3. Todos eles têm a origem centrada no centro de massa da região externa do respectivo protômero na conformação T. Cada conjunto possui três pequenas esferas, cada uma na ponta de um vetor. Elas estão coloridas de acordo com o tempo na trajetória que representam. Tempos iniciais estão em vermelho, medianos em branco e passos finais estão coloridos em azul. O gráfico de baixo mostra a variação em porcentagem dos autovalores respectivos a seus autovetores. As cores estão de acordo com a figura acima, o autovetor desenhado como um eixo vermelho tem a variação de sua amplitude (seu autovalor), em função do tempo, graficada na cor vermelha, e assim sucessivamente. Dez mil pontos compõem a linha graficada, sendo cada um a média de cada 10ps na trajetória calculada.

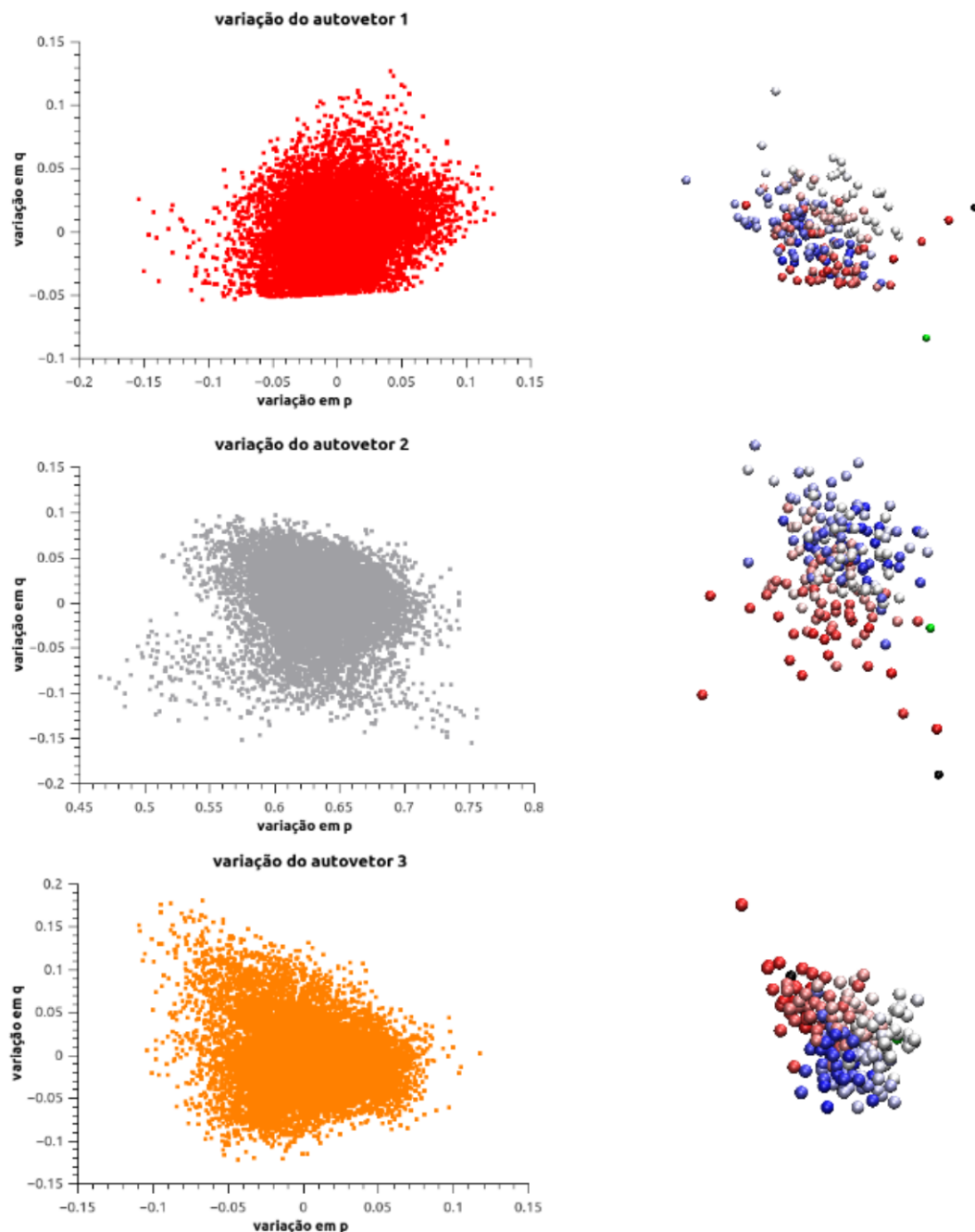


Figura 4.16 - Os três gráficos à esquerda correspondem à variação na direção dos três eixos do tensor de inércia, projetados cada qual a um plano perpendicular à média de cada eixo, sendo estes planos definidos por dois novos versores, p e q , diferentes para cada eixo. As cores dos pontos nos gráficos correspondem às cores dos eixos desenhados na figura anterior. Dez mil pontos compõem o gráfico, sendo cada um a média de cada 10ps na trajetória calculada. As figuras a direita são imagens feitas no programa VMD, mostrando somente as pequenas esferas das pontas dos eixos do tensor de inércia mostradas na figura anterior. Elas estão coloridas de acordo com sua posição no tempo, sendo cada uma a média de cada 500ps da trajetória. Tempos iniciais estão em vermelho, intermediários em branco e finais em azul. As duas esferas, preta e verde, que acompanham esta figura representam a direção correspondente deste eixo para as conformações T e R, respectivamente, com respeito ao protômero D.

Protômero E

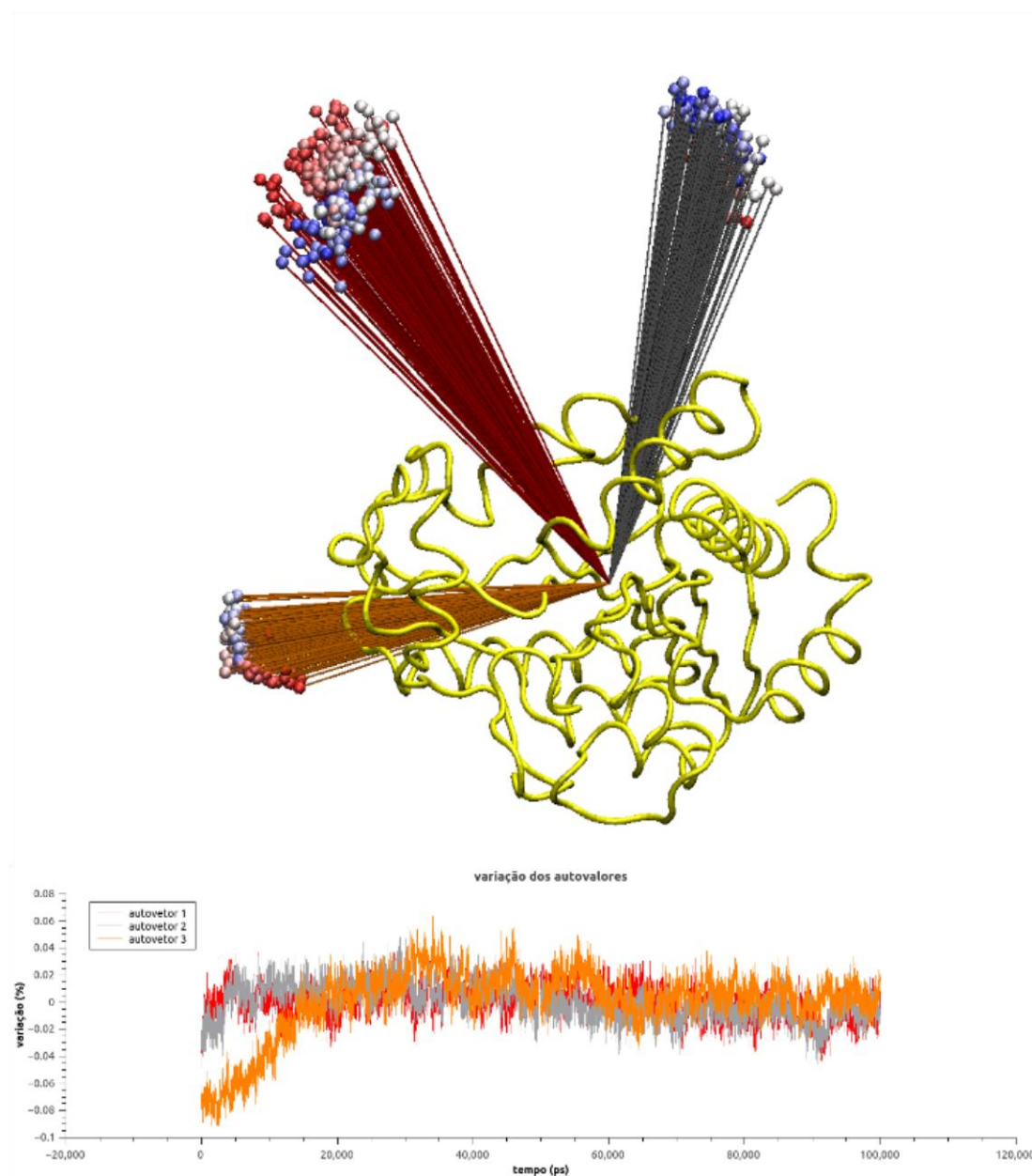


Figura 4.17 - acima, um desenho (feito com o programa VMD) do protômero E completo, visto com o eixo z perpendicular ao plano da figura, acompanhado dos eixos principais do tensor de inércia calculados para a região externa dessa cadeia. Cada conjunto de três eixos representa um tensor médio para cada 500ps da trajetória, calculada para a dinâmica 3. Todos eles têm a origem centrada no centro de massa da região externa do respectivo protômero na conformação T. Cada conjunto possui três pequenas esferas, cada uma na ponta de um vetor. Elas estão coloridas de acordo com o tempo na trajetória que representam. Tempos iniciais estão em vermelho, medianos em branco e passos finais estão coloridos em azul. O gráfico de baixo mostra a variação em porcentagem dos autovalores respectivos a seus autovetores. As cores estão de acordo com a figura acima, o autovetor desenhado como um eixo vermelho tem a variação de sua amplitude (seu autovalor), em função do tempo, graficada na cor vermelha, e assim sucessivamente. Dez mil pontos compõem a linha graficada, sendo cada um a média de cada 10ps na trajetória calculada.

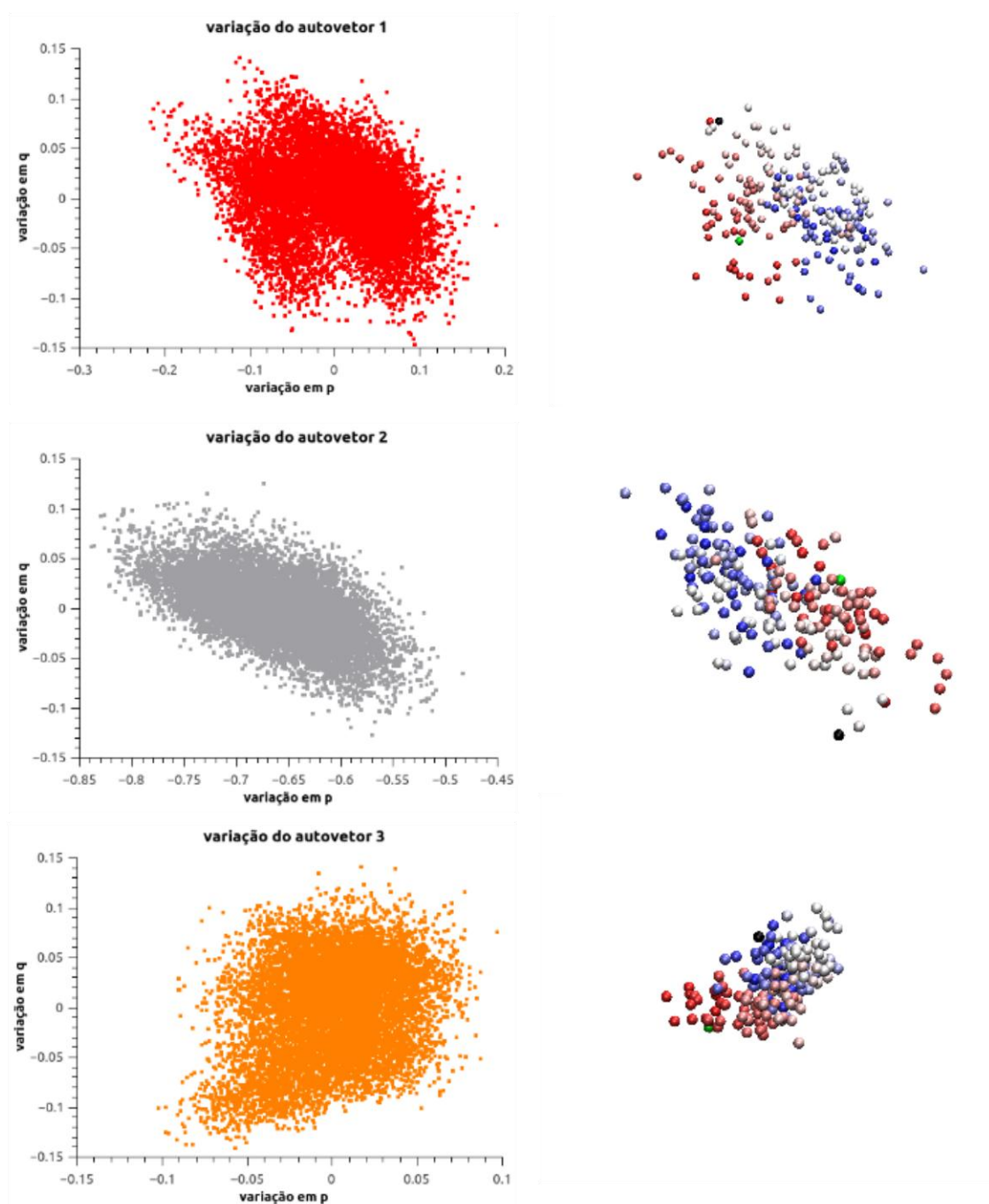


Figura 4.18 - Os três gráficos à esquerda correspondem à variação na direção dos três eixos do tensor de inércia, projetados cada qual a um plano perpendicular à média de cada eixo, sendo estes planos definidos por dois novos versores, p e q , diferentes para cada eixo. As cores dos pontos nos gráficos correspondem às cores dos eixos desenhados na figura anterior. Dez mil pontos compõem o gráfico, sendo cada um a média de cada 10ps na trajetória calculada. As figuras a direita são imagens feitas no programa VMD, mostrando somente as pequenas esferas das pontas dos eixos do tensor de inércia mostradas na figura anterior. Elas estão coloridas de acordo com sua posição no tempo, sendo cada uma uma média de cada 500ps da trajetória. Tempos iniciais estão em vermelho, intermediários em branco e finais em azul. As duas esferas, preta e verde, que acompanham esta figura representam a direção correspondente deste eixo para as conformações T e R, respectivamente, com respeito ao protômero E.

Protômero F

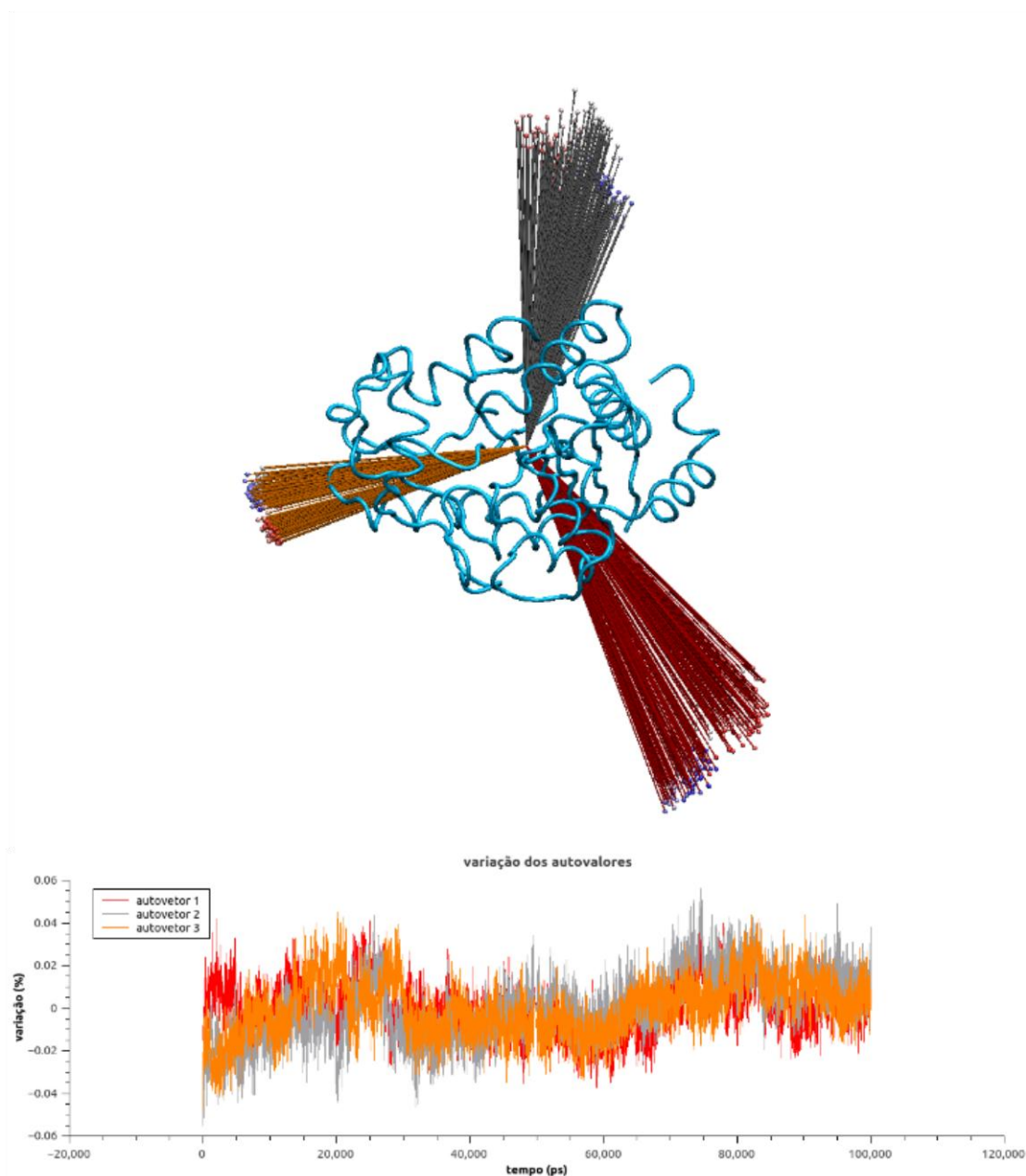


Figura 4.19 - acima, um desenho (feito com o programa VMD) do protômero F completo, visto com o eixo z perpendicular ao plano da figura, acompanhado dos eixos principais do tensor de inércia calculados para a região externa dessa cadeia. Cada conjunto de três eixos representa um tensor médio para cada 500ps da trajetória, calculada para a dinâmica 3. Todos eles têm a origem centrada no centro de massa da região externa do respectivo protômero na conformação T. Cada conjunto possui três pequenas esferas, cada uma na ponta de um vetor. Elas estão coloridas de acordo com o tempo na trajetória que representam. Tempos iniciais estão em vermelho, medianos em branco e passos finais estão coloridos em azul. O gráfico de baixo mostra a variação em porcentagem dos autovalores respectivos a seus autovetores. As cores estão de acordo com a figura acima, o autovetor desenhado como um eixo vermelho tem a variação de sua amplitude (seu autovalor), em função do tempo, graficada na cor vermelha, e assim sucessivamente. Dez mil pontos compõem a linha graficada, sendo cada um a média de cada 10ps na trajetória calculada.

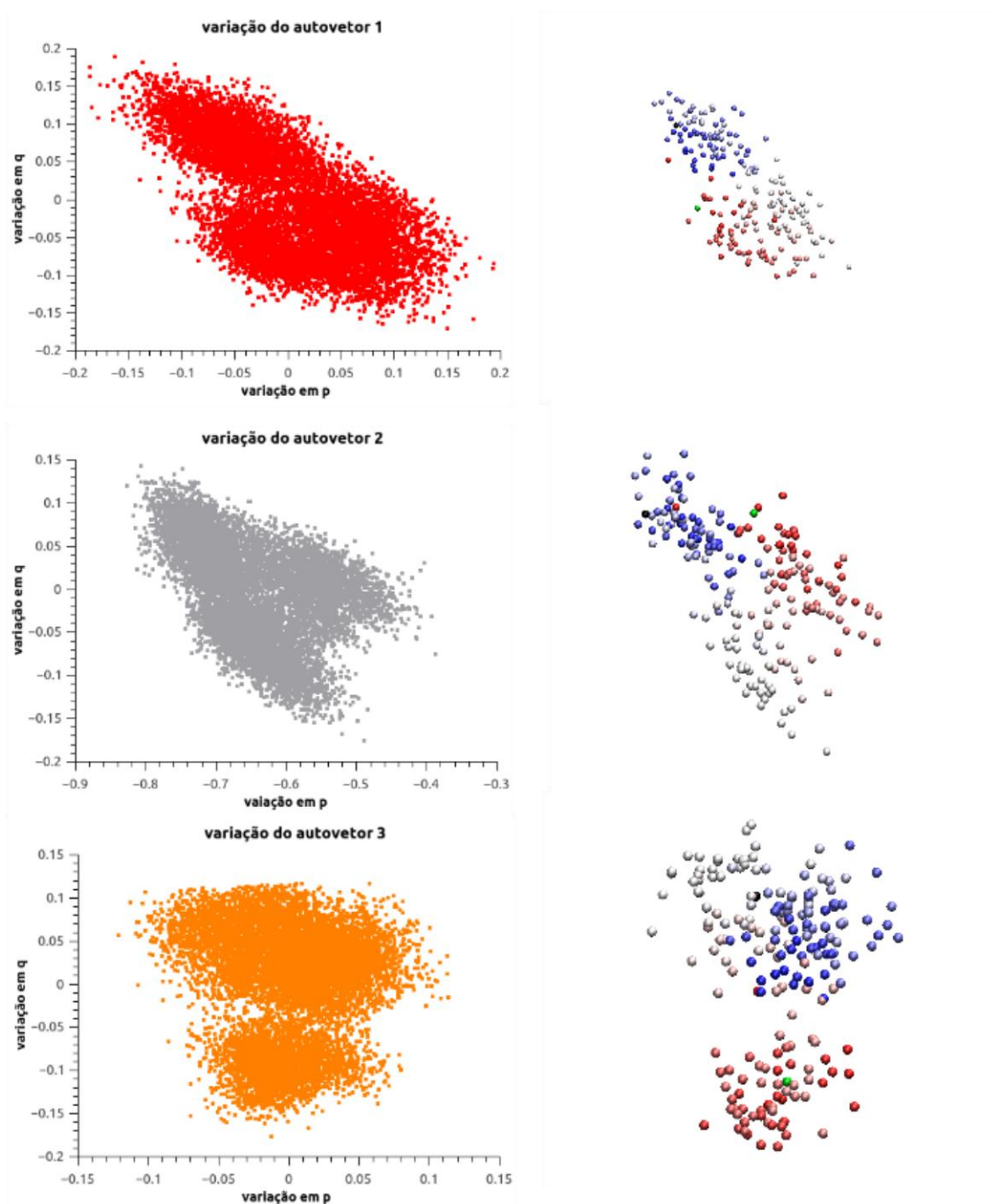


Figura 4.20 - Os três gráficos à esquerda correspondem à variação na direção dos três eixos do tensor de inércia, projetados cada qual a um plano perpendicular à média de cada eixo, sendo estes planos definidos por dois novos versores, p e q , diferentes para cada eixo. As cores dos pontos nos gráficos correspondem às cores dos eixos desenhados na figura anterior. Dez mil pontos compõem o gráfico, sendo cada um a média de cada 10ps na trajetória calculada. As figuras a direita são imagens feitas no programa VMD, mostrando somente as pequenas esferas das pontas dos eixos do tensor de inércia mostradas na figura anterior. Elas estão coloridas de acordo com sua posição no tempo, sendo cada uma uma média de cada 500ps da trajetória. Tempos iniciais estão em vermelho, intermediários em branco e finais em azul. As duas esferas, preta e verde, que acompanham esta figura representam a direção correspondente deste eixo para as conformações T e R, respectivamente, com respeito ao protômero F.

Para todos os protômeros, o desenho dos tensores de inércia no tempo, como conjunto de vetores ortogonais que mudam em direção e amplitude, conforme o corpo roda ou deforma, mostram que de fato a região escolhida, a região externa se comporta como um corpo quase rígido, pois se mantêm muito constantes nessas duas características, mantendo os autovalores com variações muito pequenas. Em termos de amplitude, eles variam no máximo 10% (situação encontrada para o protômero E como mostrado na figura 4.17), enquanto, na figura 3.10, o tensor calculado para o monômero inteiro tem amplitudes que variam em 12%. Ainda assim, as pequenas diferenças são importantes e facilmente visualizadas nos gráficos de variação. Quanto à direção dos autovetores, para todos os protômeros o terceiro autovetor, de menor valor é sempre na direção da rotação da transição alostérica, logo a região é mais alongada nessa direção, tendo os átomos não muito espalhados com relação a esse vetor. Os outros dois maiores autovetores, estão no plano perpendicular ao primeiro, mas sem evidente predominância pelo eixo z, sempre perpendicular à direção de rotação, e de simetria três no hexâmero.

Quanto à variação dos autovalores, de forma geral, os autovetores 1 e 2 têm o mesmo padrão, diminuem e aumentam de forma conjunta, o que sugere que apenas um deles, ou uma média entre os dois seria suficiente para descrever as diferenças do corpo nessa região. Também o autovetor 3 parece, na maioria das vezes ter exatamente o padrão oposto aos outros dois. De forma que quando o corpo se estica na direção 3 (direção do autovetor 3) os átomos também se aproximam desse eixo, enquanto quando o corpo se encolhe nessa direção os átomos também se espalham mais em relação ao autovetor 3. Duas exceções seriam para os protômeros B e F, sendo que para o B os autovalores parecem simplesmente não variar muito, enquanto para o F eles aparentam variar, mas em colaboração, ou o corpo se expande em todas as direções ou se encolhem em todas elas, como num movimento de respiração.

Analisar a variação na direção dos autovetores exige observar os gráficos em pares, os gráficos da direita, que possuem tantos pontos quanto conformações que constituem a trajetória, permitem identificar se a estrutura varia entre populações diferentes, que teriam orientações diferentes na conformação, ou se variam em torno de uma só conformação. Já os gráficos da esquerda, que possuem um ponto médio para cada 500 conformações que constituem a trajetória, permitem identificar se existe um caminho lógico para a mudança conformacional, ou se apenas oscila aleatoriamente em torno de uma posição média. Para o

protômero F, por exemplo, ficou evidente a existência de populações distintas, separáveis em quase três grupos, que são visitados durante a dinâmica em tempos diferentes. Enquanto que para o protômero B, dificilmente se distinguem populações e os pontos coloridos que determinam o tempo na trajetória estão bastante misturados.

Mas algum padrão em relação ao caminho percorrido, em termos de mudança conformacional, é mantido para todos os protômeros. Todos eles se iniciam perto da conformação T (possuem um ponto vermelho perto de um preto) e rapidamente se encaminham para próximo da conformação R (maioria dos pontos vermelhos perto do verde). Depois passam para outro grupo de pontos, outra população, onde a maioria é branca e por fim, se mudam para outra população onde a maioria é azul. A conformação T, o ponto preto, para alguns protômeros está próximo do grupo branco, como para o protômero A, indicando que neste ponto da trajetória, a estrutura já voltava a sua conformação inicial e os pontos azuis voltam a se aproximar do grupo vermelho, indicando que está se fechando um ciclo no caminho de volta a uma conformação já visitada. Em outros casos, o ponto preto, está próximo da população azul, indicando que apenas no final da trajetória a estrutura volta a sua conformação inicial. Em dois casos particulares, para os protômeros B e D o ponto preto está bastante distante dos pontos azuis e dos brancos, mostrando que a estrutura se distancia dessa conformação e em nenhum instante volta à ela.

Enfim, essas análises permitem dizer que, partindo-se da conformação T é possível que a proteína passe espontaneamente por uma mudança conformacional tal que se assemelhe ao estado R. No entanto, essa afirmação é somente válida para cada protômero. Todos eles passam pela mesma transição, mas não simultaneamente e a simetria da proteína não se mantém.

Entretanto, é possível que esse movimento observado seja uma combinação de outros movimentos harmônicos simétricos ou antissimétricos. Em que um descreveria a movimentação da tampa do sítio ativo e o outro poderia descrever a transição alostérica, por exemplo. Para estudar essa possibilidade foi feita uma Análise de Modos Normais. Para melhor validar os resultados desse método, foi feita uma primeira minimização de energia e depois uma continuação desta para se ter certeza de que uma estrutura de mínima energia tinha sido alcançada. Elas serão denominadas 1 e 2, respectivamente. As energias e forças calculadas para elas estão na tabela 4.2.

Tabela 4.2 - Energia potencial, força máxima encontrada para um átomo e força média encontrada para o *ensemble* de átomos no último instante da minimização, que termina com o método l-bfgs, para ambas as minimizações efetuadas.

	Energia potencial (kJ/mol)	F máxima (kJ mol⁻¹ nm⁻¹)	F norm (kJ mol⁻¹ nm⁻¹)
Minimização 1	-3.29862945849849e+05	3.27712774277860e-04	6.47989330441580e-06
Minimização 2	-3.29862945849851e+05	9.06613555569008e-07	1.04801322290570e-07

Ao fim das duas minimizações, o centro de massa da região externa foi calculado e ele permaneceu como na conformação T, mostrando que uma minimização prolongada não é capaz de provocar a transição, mas que ela ocorre assim que a proteína é colocada em banho térmico, como mostrado em resultados anteriores.

Com as duas estruturas alcançadas foram feitas Análises de Modos Normais. Os dez primeiros autovalores constam na tabela 4.3.

Tabela 4.3 - Autovalores (rad² ps⁻²) de ambas as minimizações para os dez primeiros modos normais encontrados, calculados pelo programa Gromacs usando a rotina g_nmeig.

Modo	Autovalores da minimização 1	Autovalores da minimização 2
1	-9.96007e-06	-8.5222e-06
2	-2.28683e-06	-2.43485e-06
3	1.16969e-08	6.42829e-08
4	1.30508e-07	8.96649e-07
5	2.68439e-06	2.37648e-06
6	8.39438e-06	6.66186e-06
7	0.0215904	0.0216275
8	0.04536	0.0453754
9	0.0576717	0.0576747
10	0.0641275	0.0641424

Sendo os autovalores encontrados para ambas muito semelhantes, é possível dizer que minimizar mais a estrutura não resultaria em modos diferentes. Também como os seis primeiros autovalores foram muito próximos de zero quando comparados com o restante, é correto afirmar que de fato a estrutura estava num mínimo de energia. Daqui para frente serão mostrados somente resultados da segunda minimização e como os seis primeiros autovalores não representam um modo da proteína, mas somente rotação e translação,

também eles não serão estudados. A tabela 4.4 mostra para os dez primeiros modos (a partir do sétimo) seus autovalores, e seus respectivos número de onda, autofrequência e período calculados.

Tabela 4.4 - Autovalores e números de onda calculados pelo programa Gromacs com a rotina `g_nmeig`, autofrequência e período calculados manualmente correspondentes para os dez primeiros modos normais encontrados (excetuando-se os seis primeiros que correspondem a modos de rotação e translação do sistema), para a minimização 2.

Modo	Autovalores (rad² ps⁻²)	Número de onda (cm⁻¹)	Autofrequência (GHz)	Período (ns)
7	0.0216275	0.780733	3.727725486	0.268260097
8	0.0453754	1.13086	5.399458768	0.185203748
9	0.0576747	1.27495	6.087437841	0.164272725
10	0.0641424	1.34453	6.419657869	0.155771541
11	0.0788333	1.49058	7.116995252	0.140508735
12	0.0972719	1.65574	7.905576164	0.126492994
13	0.106379	1.73152	8.267399012	0.120957026
14	0.11741	1.81908	8.685467216	0.115134854
15	0.131692	1.92654	9.198550922	0.108712775
16	0.142559	2.00445	9.57054377	0.104487271

Ainda que as frequências encontradas sejam muito altas para um movimento global – frequências altas são características de movimentos locais – o primeiro modo normal de fato descreve uma rotação que se assemelha à da transição alostérica, como mostra a figura 4.21.

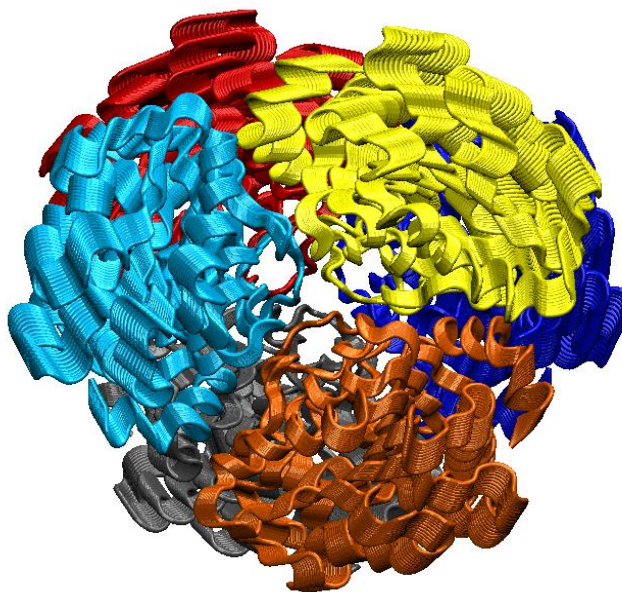


Figura 4.21 - Representação do movimento do primeiro modo normal (modo 7) encontrado. Todos os quadros da trajetória calculada para esse estão superpostos para passar uma noção do movimento. A trajetória foi calculada usando a rotina `g_nmtraj` com um parâmetro de temperatura de 1000K. Os protômeros estão coloridos conforme as cores padrões usadas nesta dissertação, a cadeia A está em azul, a B em vermelho, a C em cinza, a D em laranja, a E em amarelo e a F está em ciano.

Na figura 4.21 o movimento aparenta ser simétrico, acontecendo da mesma forma para todos os protômeros. A figura 4.22 foca apenas em uma cadeia para detalhar o movimento.

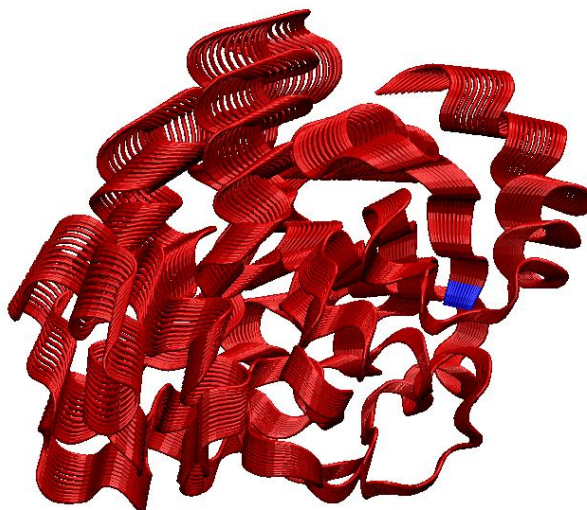


Figura 4.22 - Representação do movimento encontrado para o primeiro modo normal (modo 7) para a cadeia B em destaque. Em azul está ressaltado a metionina N terminal. Todos os quadros da trajetória foram superpostos para passar a noção de movimento. A trajetória foi calculada usando a rotina `g_nmtraj` com um parâmetro de temperatura de 1000K.

A figura 4.22 mostra claramente que a região externa passa por uma rotação enquanto que a região interna permanece praticamente fixa. No entanto, o N terminal (em azul) mostra uma movimentação um pouco maior que a esperada. Para melhor caracterizar o movimento e comparar com a transição ideal calculada inicialmente foram calculados os parâmetros de deslocamento atômico para este primeiro modo normal, mostrado no gráfico 4.8.

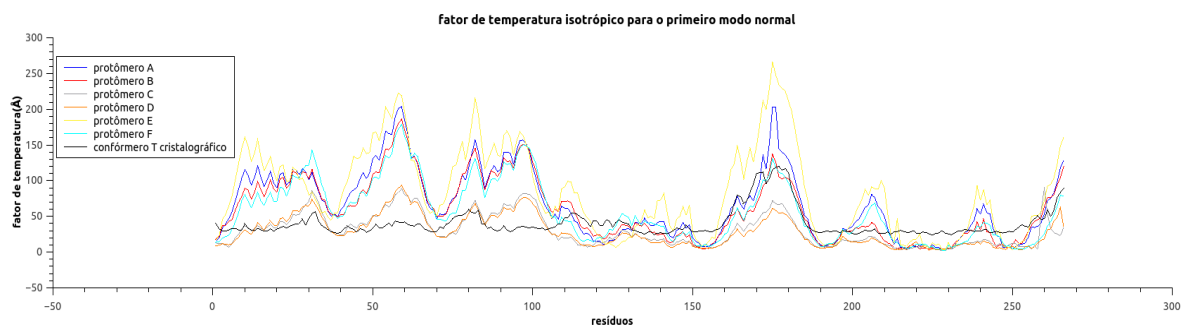


Gráfico 4.8 - Parâmetros de deslocamento atômico isotrópicos, ou fator de temperatura isotrópico, calculados para cada Carbono α . As seis linhas coloridas representam cada protômero na trajetória calculada para o primeiro modo normal usando o programa Gromacs e a rotina `g_nmtraj` com um parâmetro de temperatura igual a 1000K. A linha preta representa o mesmo parâmetro encontrado por cristalografia para a estrutura de código 1fsf, um monômero na conformação T.

O gráfico 4.8 mostra que o movimento não é tão simétrico quanto aparentava. Ainda que todas as cadeias apresentem o mesmo padrão, que também é o padrão do dado experimental cristalográfico, elas possuem amplitudes bastante variadas.

Comparando-se com a transição ideal, vê-se que as amplitudes da tampa do sítio ativo e do loop do resíduo 80 são proporcionais. Mas para este movimento aparecem dois picos por volta do resíduo 205 e 240 que não são tão evidentes nem na transição calculada nem no dado cristalográfico. Talvez o movimento encontrado se assemelhe ao de uma rotação por completo do monômero. Esse resultado já foi encontrado antes por redes elásticas e acreditava-se que um método mais robusto, que considera todos os átomos e seus devidos potenciais pudesse contornar esse resultado. No entanto, a análise de modos normais não parece descrever um movimento muito diferente e também resulta em frequências altas demais para um movimento global. Talvez o campo de forças escolhido não tenha sido adequado, e algum potencial essencial para este movimento não esteja sendo considerado.

O último experimento realizado para caracterizar a dinâmica da proteína foi a Análise de Componentes Principais. Procurando saber se a trajetória poderia ser descrita como uma

soma de movimentos harmônicos, e se seria o de maior contribuição justamente o que descreve a transição alostérica. O gráfico 4.9 indica a porcentagem da contribuição dos modos encontrados.

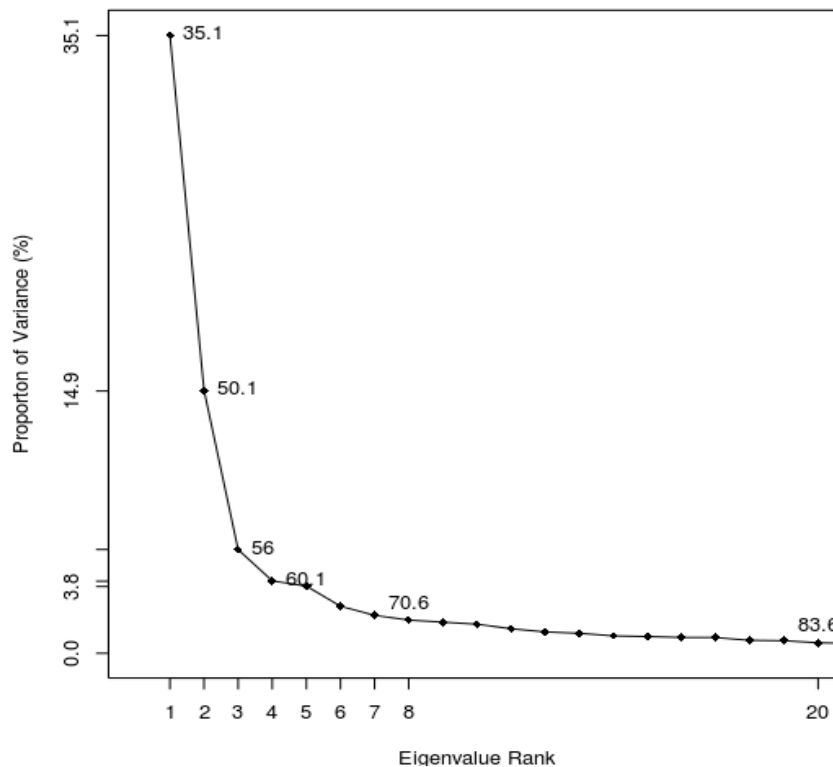


Gráfico 4.9 - Resultados oriundos do uso do programa bio3d para cálculos de Análise de Componentes Principais. O gráfico mostra a contribuição das primeiras 20 componentes na descrição da trajetória, sendo atribuído a cada uma quanto do movimento na trajetória pode ser descrito como uma variação na direção determinada pelo respectivo autovetor.

O gráfico 4.9, no entanto mostra que nem somando 20 modos é possível descrever a trajetória, e a soma de todos os modos indica não convergir para 100% da trajetória descrita. Esse resultado é característico de um movimento anarmônico, mas também pode ser consequência de uma dinâmica muito curta. De qualquer forma, os três primeiros modos já são bastante representativos. A flutuação em Å por resíduo para estes está na figura 4.23.

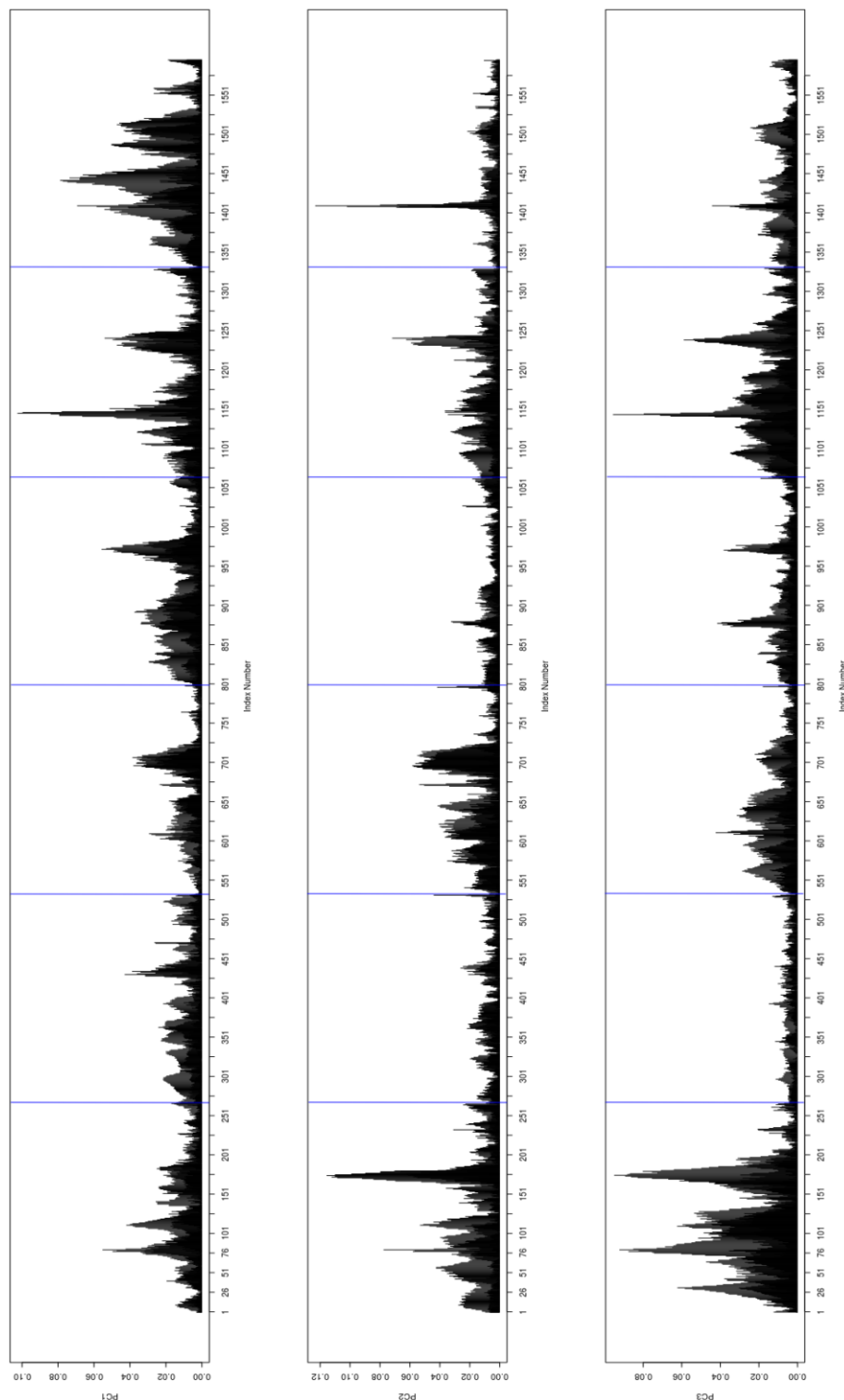


Figura 4.23 - Resultados oriundos do uso do programa bio3d para cálculos de Análise de Componentes Principais. Flutuações médias em Å calculadas para cada resíduo do hexâmero completo, considerando somente carbonos α , para a trajetória da dinâmica 3. As linhas azuis verticais indicam o início e o fim de cada protômero. Da esquerda para a direita estão mostradas as cadeias A, B, C, D, E e F. De cima para baixo estão mostrados os gráficos para as três primeiras componentes principais em ordem crescente. Cada um representando um padrão de movimentação na direção do respectivo autovetor. Tal padrão é semelhante e comparável aos parâmetros de deslocamento atômico isotrópico.

Considerando apenas a cadeia A, vê-se que o primeiro modo não apresenta alta amplitude para a tampa do sítio ativo, sendo o loop do resíduo 80 a flutuação que mais se destaca, além de toda a região externa ter uma flutuação proporcional e a região interna ainda se manter bastante baixa. Já no segundo modo a tampa do sítio ativo parece ser muito mais influente não só para a cadeia A como também para a E e a F. O terceiro modo parece apenas contribuir para o movimento aumentando a amplitude do que já era alta, particularmente para os protômeros A e E. Essa análise parece indicar que o primeiro modo descreve a transição alostérica enquanto o segundo descreve a movimentação da tampa do sítio ativo, como imaginado, mas somente para as cadeias A e E. A cadeia F apresenta um comportamento bastante diferente no modo 2, em que apenas o loop do resíduo 80 parece se movimentar. As outras cadeias parecem apenas seguir o padrão encontrado na cristalografia. É possível perceber ainda, que cada protômero possui amplitudes distintas em um mesmo modo, o que ressalta a falta de simetria do movimento do hexâmero. Algum padrão quanto às amplitudes relativas a cada cadeia pode ser observado para os modos 2 e 3. As cadeias B, D e F possuem amplitudes médias claramente menores que as outras cadeias nestes dois modos ou componentes principais, mas nenhuma relação aparente é encontrada entre eles, a não ser o fato de serem todos pertencentes a dímeros diferentes.

Uma análise muito semelhante -em termos de padrão de gráfico- e complementar a esta é de parâmetros de deslocamento atômico isotrópico. Ela se encontra no gráfico 4.10, calculada apenas para a primeira componente principal, ou primeiro modo.

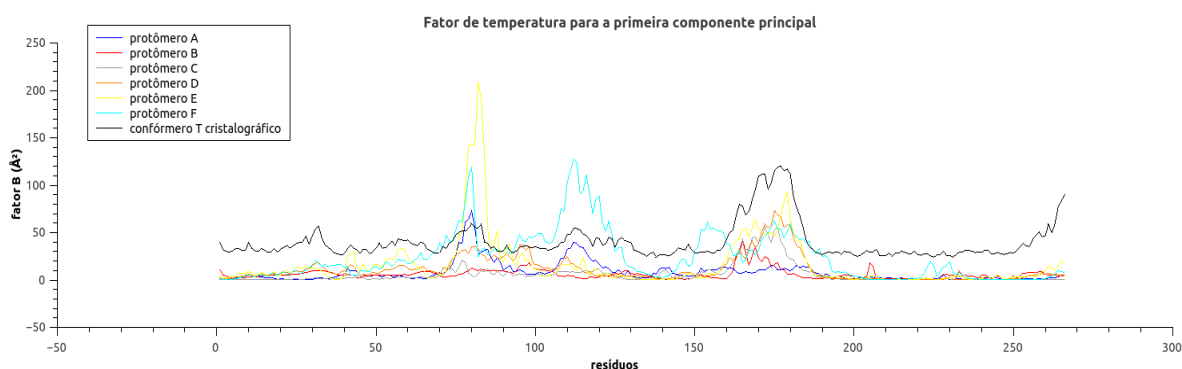


Gráfico 4.10 - Parâmetros de deslocamento atômico isotrópicos, ou fator de temperatura isotrópico, calculados para cada Carbono α . As seis linhas coloridas representam cada protômero na trajetória calculada para o primeiro modo, ou primeira componente principal, encontrado por PCA. Análise feita com o uso do programa bio3d. A linha preta representa o mesmo parâmetro encontrado por cristalografia para a estrutura de código 1fsf, um monômero na conformação T.

Mais uma vez o parâmetro de deslocamento mostra para as cadeias A, E e F o loop do resíduo 80 mais flexível que a tampa do sítio ativo, enquanto outras cadeias apenas seguem o padrão cristalográfico. É interessante notar que as amplitudes para este movimento são relativamente mais baixas que para os outros movimentos avaliados. Talvez isto seja devido ao fato de a primeira componente principal conseguir contribuir com apenas 35% para a descrição do movimento completo da dinâmica, como mostra o gráfico 4.9.

Como análises anteriores mostraram que nem todos os protômeros de fato chegam a completar um período no movimento de transição, também foi realizada uma Análise de Componentes Principais para cada protômero. Os gráficos de porcentagem de contribuição estão na figura 4.24.

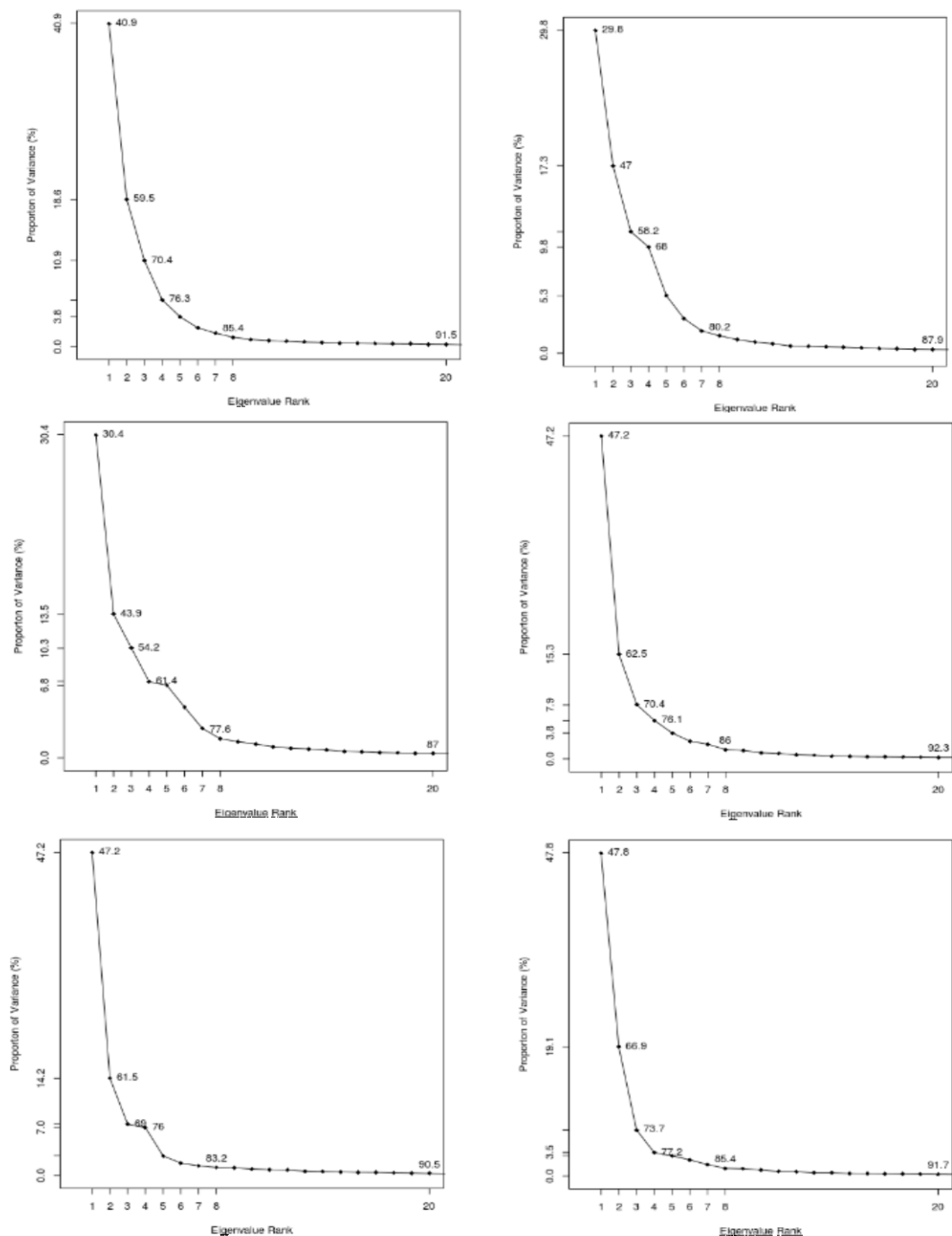


Figura 4.24 - Resultados oriundos do uso do programa bio3d para cálculos de Análise de Componentes Principais. Os gráficos mostram a contribuição das primeiras 20 componentes na descrição da trajetória, sendo atribuído a cada uma quanto do movimento na trajetória pode ser descrito como uma variação na direção determinada pelo respectivo autovetor. O movimento de cada cadeia foi analisado separadamente, resultando em seis gráficos, um para cada monômero. Na coluna da esquerda, de cima para baixo estão os dados para as cadeias A, B e C, nesta ordem, enquanto na coluna da direita estão os dados das cadeias D, E e F, de cima para baixo.

A figura 4.24 mostra que a maioria dos protômeros é mais bem descrita

separadamente, ou seja, a soma de movimentos harmônicos - a soma dos movimentos em diferentes eixos ortogonais-, descreve bem a trajetória. A exceção são as cadeias B e D, cujos primeiros modos mal representam um terço da trajetória e a soma não indica convergir para próximo de 100%. Já os protômeros A, E e F são muito bem descritos para este método. A primeira componente principal de cada cadeia tem seu movimento representado na figura 4.25.

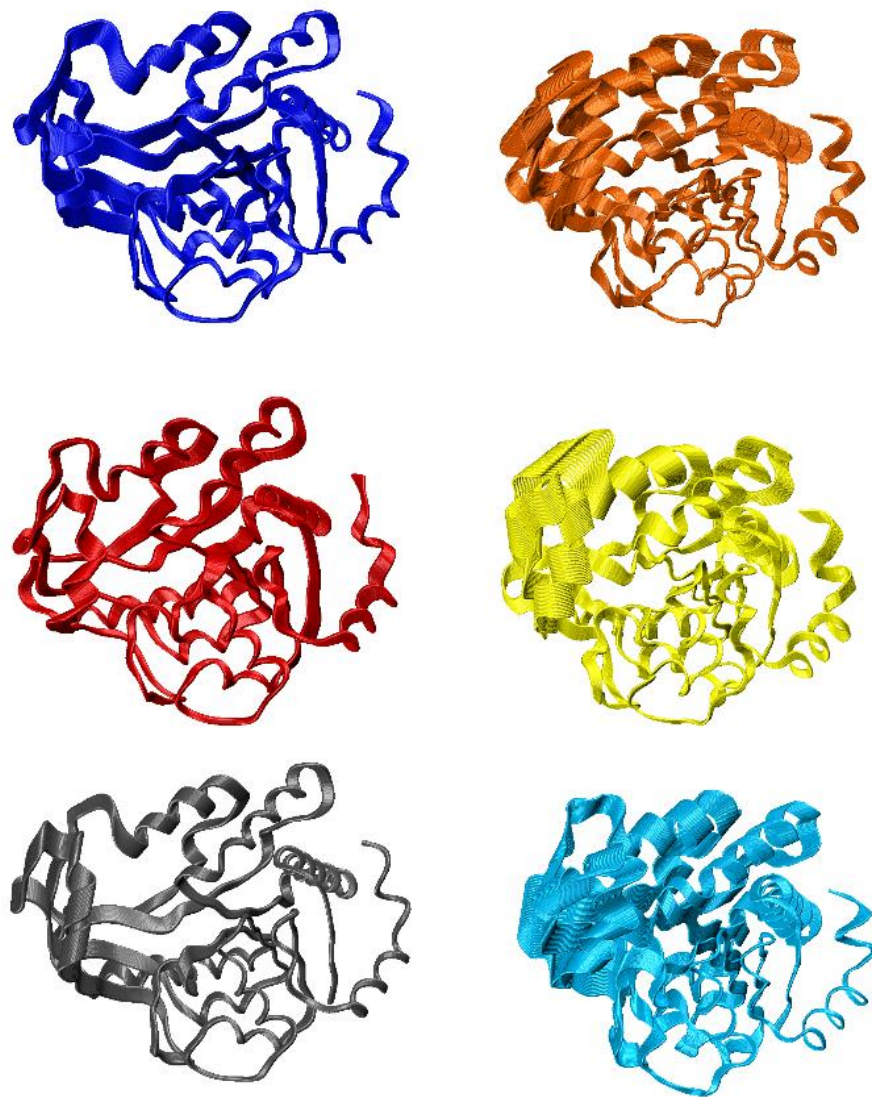


Figura 4.25 - Representação do movimento da primeira componente principal encontrada, quando considerado cada monômero individualmente. Todos os quadros da trajetória calculada para esse estão superpostos para passar uma noção de dinâmica. A trajetória foi calculada usando o programa bio3d. Os protômeros estão coloridos conforme as cores padrões usadas nesta dissertação, a cadeia A está em azul, a B em vermelho, a C em cinza, a D em laranja, a E em amarelo e a F está em ciano.

A figura 4.25 mostra que os protômeros A, E e F possuem uma flexibilidade maior no lado esquerdo da estrutura se comparado com o restante da cadeia. O protômero C também, porém com amplitudes menores para toda a estrutura. Esses protômeros são também aqueles que apresentam alta porcentagem de correspondência da trajetória com a primeira componente principal (acima de 40% - como mostra a figura 4.24).

Todos parecem mostrar maior flexibilidade na região externa e fazendo um movimento como uma rotação. Os protômeros A, B e C não aparentam uma mudança conformacional muito grande, sendo bastante estáveis em torno de uma conformação, o que condiz com o gráfico para a primeira componente principal mostrado na figura 4.23.

Por fim, numa tentativa de analisar a relação entre os movimentos encontrados por Análise de Modos Normais e análise de Componentes Principais, calculou-se o vetor direção de movimento para cada átomo nas coordenadas x, y e z, compondo um vetor 3N para cada um dos modos normais ou componentes principais. Para cada par destes vetores foi feito um produto escalar, pretendendo encontrar se as direções do movimento eram coincidentes. No entanto, como ambas as análises encontram 3N autovetores, a probabilidade de que se encontrem direções coincidentes é de $(3N)^{-2}$. E de fato o produto escalar entre a primeira componente principal e qualquer um dos cinco primeiros modos normais (modos de 7 a 11) resultava num escalar da ordem de 10^{-2} , um número bastante próximo de zero para considerar as direções como ortogonais. Entretanto, produtos escalares dos mesmos modos normais com a segunda componente principal resultaram da ordem de 10^{-3} . Isso leva a reconsiderar o primeiro resultado, mas como tal análise não é usual, é difícil especular quão ortogonais estão os vetores analisados, e a conclusão mais segura é de que os movimentos em questão refletem direções bastante diferentes.

Como nenhuma das análises acima mostrou uma flexibilidade com conservação da simetria, em que aparecem só conformações simétricas, duas dinâmicas alternativas foram realizadas. A primeira (dinâmica 2) iniciou, após minimização e equilibração, com velocidades atômicas calculadas para manterem a simetria do grupo espacial. A segunda (dinâmica F) iniciou com um fosfato numa posição ótima de um dos sítios alostéricos, de tal forma que se mantivesse preso por ligações de hidrogênio assim como no estado R da proteína.

A dinâmica 2 mostrou-se bastante estável em energia e pressão conforme os gráficos 4.11 e 4.12.

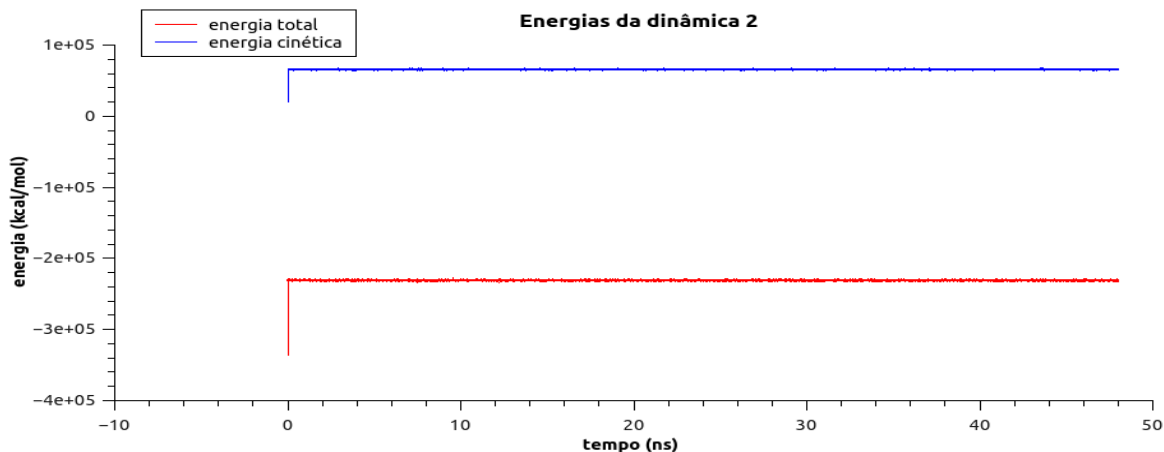


Gráfico 4.11 - Energias total (em vermelho) e cinética (em azul) do sistema, proteína em solvente, durante a dinâmica 2, calculadas pelo programa NAMD.

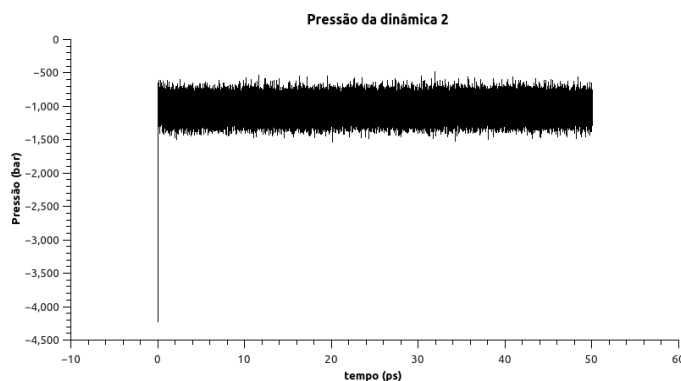


Gráfico 4.12 - Pressão calculada para o sistema, proteína em solvente, durante a dinâmica 2, calculada pelo programa NAMD.

Uma análise feita para esta dinâmica, a fim de verificar se se manteve a simetria inicialmente imposta pelas velocidades foi a de deslocamento angular dos centros de massa da região externa. A figura 4.26 mostra os gráficos referentes para cada cadeia.

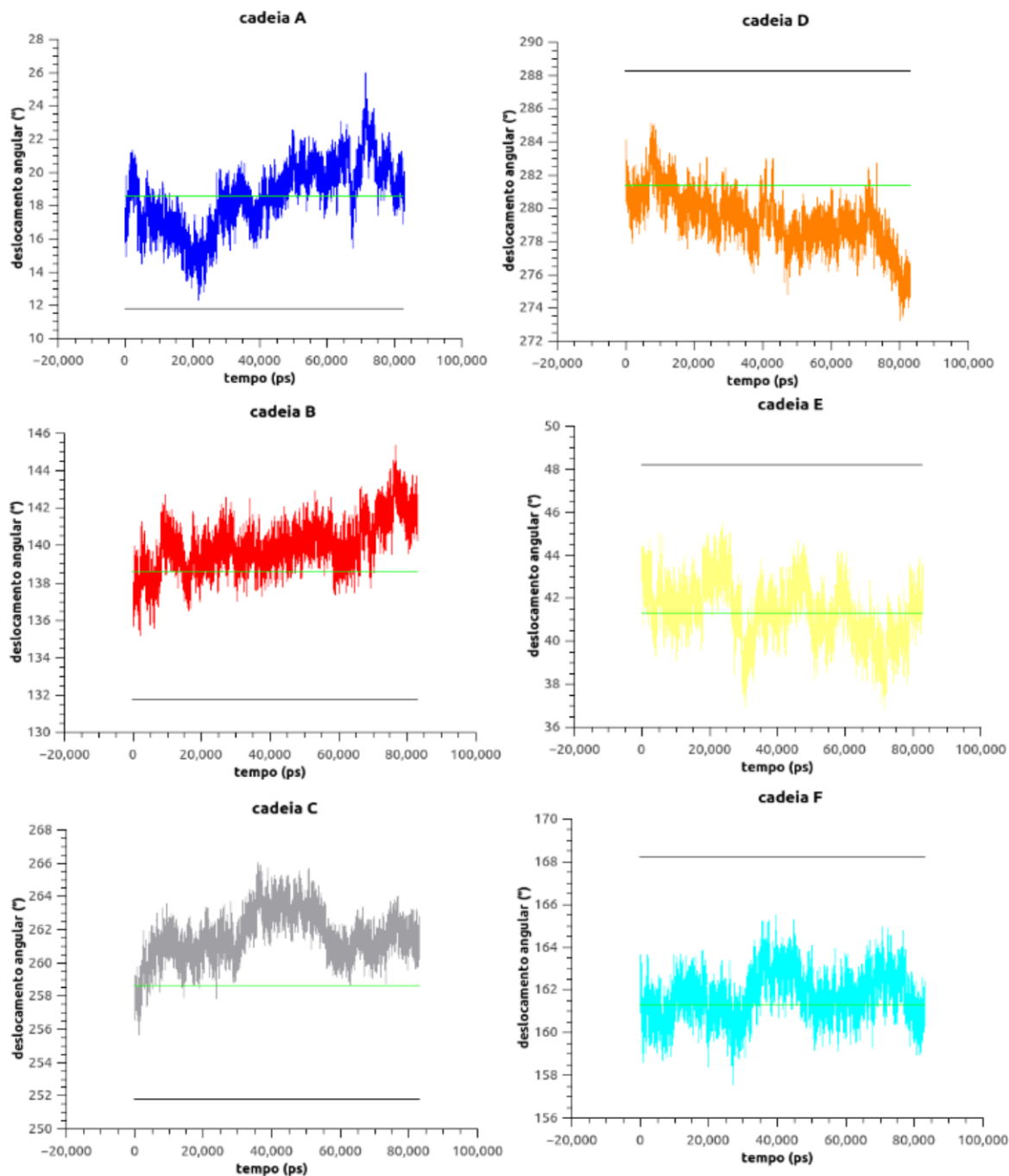


Figura 4.26 - Seis gráficos, um para cada cadeia do hexâmero, que mostram o deslocamento angular do centro de massa. Dados calculados para a trajetória da dinâmica 2. As cadeias A, B e C formam um trímero de simetria 3, assim como as cadeias D, E e F. Cálculo feito para a trajetória da dinâmica 2. As linhas horizontais verde e preta representam o ângulo calculado para cada cadeia nas conformações R e T respectivamente.

Ainda que tenham um início semelhante, com o centro de massa tendendo a se posicionar como no confórmero R, a simetria é logo perdida.

A dinâmica F também se mostrou bastante estável quanto à energia. A análise calculada para esta foi a de distâncias entre resíduos do sítio alostérico. A figura 4.27 mostra os resultados para cada sítio.

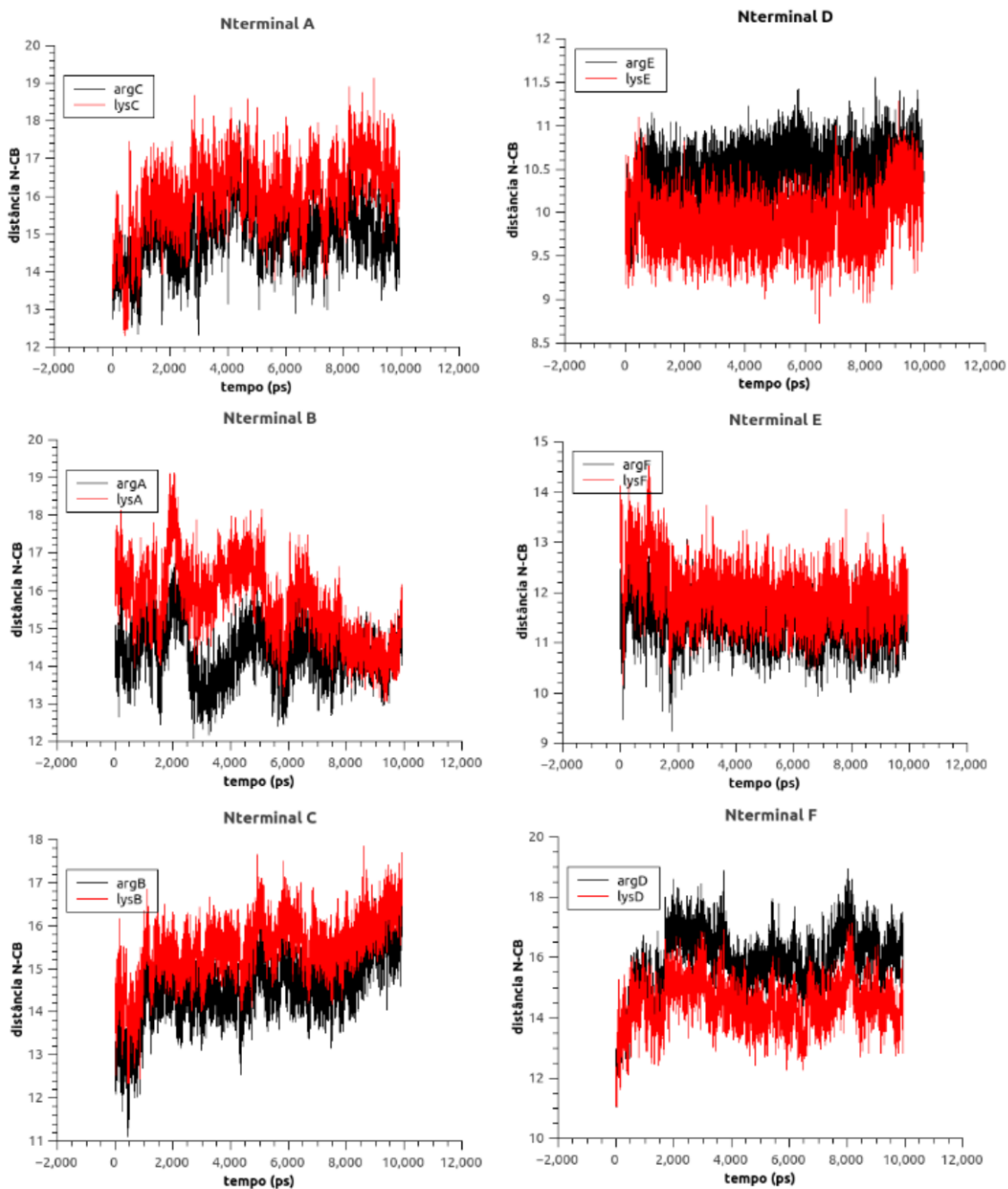


Figura 4.27 - Seis gráficos, um para cada sítio alostérico do hexâmero, que mostram o comprimento da ligação entre o nitrogênio da metionina N terminal e o carbono β da arginina 158 em vermelho, e entre o mesmo nitrogênio da metionina e o carbono β da lisina 160 em preto. Como os sítios alostéricos são formados entre os protômeros, a metionina se encontra em uma cadeia, identificada no título de cada gráfico, enquanto as arginina e lisina pertencem à outra cadeia identificada na legenda de cada gráfico. Cálculo feito para a dinâmica F.

Na figura 4.27, o gráfico referente ao sítio em que o fosfato se encontra (entre as cadeias D e E) mostra que ele permaneceu preso durante toda a dinâmica, estabilizando a conformação como no estado R. Isso indica que a partir do conformero T é possível que os protômeros alcancem a conformação R. Também o sítio alostérico formado pelo N terminal da cadeia E se mostrou mais estável e próximo da conformação R que o restante. Mas não é possível dizer que apenas um fosfato ligado desencadeou a mudança conformacional da transição alostérica de todos os outros protômeros.

5 Mais Discussões e Conclusões

Este trabalho procurou estudar a transição alostérica de uma proteína em particular, a Glucosamina 6-fosfato Desaminase, e para isso algumas propostas de métodos computacionais de análise foram levantadas. Ainda que algumas dificuldades tenham surgido, que impossibilitaram análises conclusivas, dado o tamanho do sistema em estudo, ou mesmo a natureza desconhecida da forma da transição, a maioria das análises usadas mostrou ser eficiente em algum aspecto.

Ademais, por mais que transições alostéricas sejam mudanças conformacionais bastante específicas para cada sistema em particular, as análises aqui propostas são abrangentes e passíveis de serem estendidas e aplicadas a outras proteínas.

A começar pelo algoritmo de alinhamento que encontra regiões de core rígido na proteína – como descrito na seção 3.4.3. Alinhamentos locais já é uma preocupação em estudos de estrutura proteica, mas eles dependem de um conhecimento prévio e alguma análise detalhada de parâmetros de flexibilidade atômica nem sempre disponíveis. Mesmo não sendo uma ideia original, já encontrada em programas como o bio3d, ou o antigo O, ela é imprescindível para a maioria das análises conseguintes aqui propostas, visto que são muito dependentes de mudanças conformacionais de algumas regiões em particular. Este alinhamento também possibilita uma primeira descrição da transição como somente a rotação de um corpo, como proposto por Monod, *et al.*. E, por fim, encontra regiões que se comportam como um corpo elástico, ou quase rígido, que podem ser analisadas por tensores de inércia.

A análise por tensores de inércia, proposta aqui, mostrou-se positiva em alguns pontos. Em primeiro lugar, foi capaz de distinguir um conjunto de átomos que compõe um corpo quase rígido de outro que não compõe, como mostrado na metodologia, em que um monômero inteiro teve variações nos seus autovalores de até 12%, enquanto apenas a região externa variava somente 6%. Em segundo lugar, ainda que pequenas, as variações dos autovalores permitiram diferenciar movimentos de respiração, como o encontrado para o monômero F na figura 19 dos resultados, de outros que mantêm um volume constante do corpo, como um movimento de alongamento e encolhimento acompanhado de um de afinamento e alargamento, como o encontrado para os outros monômeros, ou mesmo para a

hélice alfa usada de exemplo na metodologia.

A variação dos autovetores também possibilitou análises frutíferas. A projeção deles no plano, acompanhada da visualização espacial permitiu encontrar grupos distintos no espaço amostral das conformações, o que pressupõe a existência de mínimos locais na energia potencial, que caracterizam alguns estados conformacionais metaestáveis. E por fim, mas não menos importante, foi possível também acompanhar em que instante essas populações características de mínimos locais foram visitadas, e se se trata de um movimento cíclico, em que a proteína volta a assumir uma estrutura global similar a outras já visitadas.

Outros algoritmos escritos para a composição deste trabalho, como o cálculo de parâmetros de deslocamento atômico, o de distâncias entre resíduos chaves para a transição ou o de deslocamento do centro de massa de corpos quase rígidos, já são mais simples, mas compõem um conjunto de técnicas, que como num protocolo experimental, carrega muita informação, talvez até de forma mais clara que os tensores de inércia, e possibilita a comparação com experimentos *in vitro*, como a cristalografia, além de permitirem avaliar se a transição alostérica realmente aconteceu - como encontrado pelo menos individualmente para alguns monômeros, conforme a figura 4.7, que uma vez próximos da conformação R, o sítio alostérico formado é capaz de fixar um ligante e manter-se nesta conformação enquanto ligado ao ativador, como mostra a figura 4.27.

Além destas análises, o uso de técnicas já bem estabelecidas como a Análise de Modos Normais e a Análise de Componentes Principais permitiu discutir sobre a harmonicidade do movimento, sua coletividade e sua simetria.

Em suma, a metodologia apresentada aqui para o estudo de transições alostéricas mostrou-se construtiva. Aponta dificuldades para o estudo do alosterismo que não estavam claras, permitindo algumas propostas de soluções, um caminho de quê parâmetros são necessários para descrever uma transição, além de fornecer as ferramentas para encontrá-los.

Já o experimento *in silico* realizado, as longas Dinâmicas Moleculares (simples ou com ligante), não apenas forneceu os dados para as análises citadas acima, como levantou questões sobre espaço amostral conformacional, podendo compará-lo com o obtido por cristalografia, e sua relação com a temperatura; além de levantar aspectos importantes para o estudo do alosterismo em particular, como a influência da periodicidade do sistema e a

simetria usada (a de uma caixa de água, diferente da simetria do cristal).

A água pode ser um fator importante na movimentação da proteína, fornecendo-lhe energia, participando de movimentos de vibração e conferindo entropia ao sistema. É possível que sua periodicidade e a simetria da caixa de solvatação estejam influenciando para retirar a simetria que se esperava para a oscilação da proteína, um hexâmero de simetria 32.

Um problema encontrado para a realização destas dinâmicas foi o tamanho do sistema. Não só no que concerne tempo de máquina, que aumenta conforme o número de átomos, mas também no que diz respeito aos parâmetros utilizados. Em especial, parâmetros de busca por vizinho mais próximo como a distância mínima para pares de átomos que devem interagir por Coulomb ou Van der Waals. Um sistema grande necessita uma distância mínima maior que o usual, o que leva a uma região do potencial, principalmente o de Coulomb, mais distante do mínimo que se aproxima a uma parábola, e isso pode conferir alguns termos de anarmonicidade ao sistema. Também o uso do método da soma de Ewald para interações de longa distância torna-se mais difícil, visto que o raio de corte da esfera de Ewald é proporcional ao tamanho da caixa do sistema.

Quanto ao espaço de conformações que a Dinâmica Molecular proporciona, ele pode ser comparado ao existente nos cristais analisando-se os parâmetros de deslocamento anisotrópicos, considerando os elipsoides proporcionais ao número de conformações. A começar, um cristal possui mais conformações que as calculadas para uma trajetória de dinâmica (mesmo tomando como tamanho de passo para guardar uma conformação um tempo comparável ao das vibrações percebidas num cristal). Por exemplo, o cristal de Glucosamina 6-fosfato Desaminase possui lados de tamanho próximo a 0,1mm e uma cela cristalográfica de lados por volta de 120Å, o que pressupõe a existência 10^{12} unidades de proteína para compor os dados cristalográficos, muito mais que o alcançável por simulação computacional. No entanto, a temperatura mais alta em que é simulada a proteína numa dinâmica pressupõe maior flexibilidade, ou seja, maior amplitude nos parâmetros de deslocamento atômico. De fato, comparando-se as figuras 5 e 6 da seção de resultados, vê-se que alguns elipsoides calculados para a Dinâmica possuem um eixo sobressalente, enquanto os da cristalografia se assemelham mais a esferas. Assim, pode-se concluir que a dinâmica possibilita uma população de conformações de forma mais direcional que a cristalografia, ou seja, que divergem preponderantemente em poucas direções do movimento total.

Provavelmente, se o tempo pudesse ser bastante prolongado outras direções de movimento seriam encontradas. Uma análise viável, que aproximaria mais estes resultados, seria calcular a média dos deslocamentos anisotrópicos considerando os seis protômeros. É possível até que a proteína de fato apresente deslocamentos assimétricos no cristal (mesmo com posições médias simétricas) para cada cadeia, mas sua simetria implícita permite apenas encontrar a média deles.

Por último, no que diz respeito à transição alostérica da Glucosamina 6-fosfato Desaminase em particular, a primeira conclusão é de que não foi percebido um movimento simétrico durante as dinâmicas, apesar da simetria 32 da proteína, e do predito por Monod (de que a transição alostérica aconteceria de forma simétrica para todos os protômeros). No entanto, foi detectado sim que a proteína se move de tal forma a se aproximar das duas conformações, R e T, por diferentes análises – pelo deslocamento de centro de massa, pela distância de ligações no sítio alostérico e por análise de tensor de inércia. Essa mudança de conformação foi percebida em dois instantes. No decorrer da dinâmica, para alguns protômeros separada e lentamente, e no início dela para todos os protômeros coletiva e rapidamente.

A mudança lenta é claramente percebida para as cadeias A e E, como mostra a figura 4.7, e em tempos similares. De fato o deslocamento angular de centro de massa mostrou um padrão bastante semelhante para estas duas cadeias, enquanto os outros protômeros pareceram movimentar-se de forma independente. A cadeia F ainda tende à conformação T de volta, mas as outras permanecem com alguma estabilidade na forma R. Imagina-se que para tempos maiores de simulação todas as cadeias retornariam à forma T, mesmo que de forma assimétrica, momento em que, se feita uma Análise de Componentes Principais, todas as cadeias apresentariam gráficos de contribuição das componentes como o encontrado para as cadeias A ou E (na figura 4.24), cuja soma de componentes converge para mais 90% da trajetória, e isso deveria levar a uma convergência também, da mesma análise, para o hexâmero inteiro. Essa perspectiva está baseada na ideia de que a não convergência da soma dos modos de PCA pode ser proveniente não de uma anarmonicidade do sistema, mas da sua falta de simetria, o que requer trajetórias mais longas para que todos os monômeros mostrem um movimento oscilatório entre as duas conformações. Isso significa que a transição

alostérica possui um período, na média para os protômeros, maior que 100ns, ou uma frequência menor que 10MHz – a esperada frequência baixa para movimentos globais.

Já a mudança rápida e inicial para R é simétrica para todas as cadeias, mas acontece em aproximadamente 500ps, o que corresponderia a uma frequência de 2GHz, muito rápida para um movimento global. Essa rápida transição aconteceu para as três dinâmicas simuladas, seja durante a equilibração ou não, e é sempre seguida da transição mais lenta descrita acima. Ainda que seja uma frequência inesperada, ela é coerente com o calculado por Análise de Modos Normais para o modo de mais baixa frequência (que corresponde ao movimento da transição alostérica) e mantém a simetria predita por Monod.

Assim, as duas formas distintas de transição observadas mantêm, cada uma, uma das características esperadas para a transição alostérica, em uma a simetria, e na outra a frequência.

Portanto, é preciso encontrar como é, realmente, a função energia potencial que coordena a transição alostérica. Questão essa já levantada em trabalhos anteriores (62-63).

Uma análise da energia potencial, baseada no gráfico 4.6, mostra que durante a transição rápida e simétrica ocorre uma queda na energia de aproximadamente 8,7Mcal/mol, até 500ps quando alcança uma média de -271,8Mcal. Ela continua caindo até uma média de -272,2Mcal/mol, o que constitui uma diferença de 0,4Mcal/mol. Assim, admitindo que até os 500ps ainda se trata de uma transição simétrica, conforme a estrutura se afasta dessa simetria, também a energia potencial cai. Já a transição identificada como de baixa frequência e assimétrica ocorre concomitantemente a uma média na energia potencial de -272,4Mcal/mol, que oscila numa frequência de 7ns, fazendo essa média ter uma variação de 100kcal/mol durante a dinâmica, mas ainda em conjunto com rápidas oscilações de amplitude de 1Mcal/mol.

Comparando os dois tipos de transição observados, a primeira é então caracterizada por uma queda exponencial da energia, enquanto a segunda transição, por uma oscilação. Desde o ponto em que a transição deixa de ser simétrica (em 500ps, quando se pode considerar a última estrutura como simétrica), até um ponto em que a oscilação na média da energia se estabelece (e também as conformações já se caracterizam pela assimetria entre os protômeros), a diferença na energia é de, aproximadamente, 0,6Mcal/mol (como comentado acima). No entanto, essa diferença é menor que a faixa de energias acessíveis durante a

dinâmica, de amplitude 1Mcal/mol. Isso sugere que uma estrutura simétrica ainda é acessível, mesmo durante uma transição assimétrica. Assim, a hipótese de Monod permanece viável, na medida em que os estados simétricos são acessíveis, mesmo que passando por conformações não simétricas.

Complementando, essas diferenças energéticas encontradas, tanto entre as estruturas simétricas e assimétricas (0,6Mcal/mol), quanto para a variação de energia ao longo de uma dinâmica de mudanças conformacionais assimétricas (1Mcal/mol), são menores que NkT (da ordem de 50Mcal/mol), sugerindo que a energia fornecida pelo banho térmico é mais que suficiente para promover a transição entre os estados alostéricos.

Uma forma imaginada de se determinar com maior precisão a frequência – e energia associada - dos movimentos em questão, seja na dinâmica ou nos modos normais calculados, seria por uma análise de Fourier, capaz de decompor um movimento em oscilações de frequência bem determinada. Outra maneira de procurar determinar qual a forma da energia potencial seria tentar distinguir as diferentes populações conformacionais visitadas numa dinâmica.

Entretanto, este trabalho já pôde dar, ainda que de forma simplificada, uma ideia do conjunto de conformações visitadas pela análise de tensores de inércia. A figura 4.20 mostra claramente para a cadeia F a existência de três populações conformacionais, e isso pressupõe a existência de três mínimos locais na forma do potencial.

Essas análises apontam para um padrão na energia potencial que se assemelha mais ao formato de um poço de potencial com mínimos locais, que o potencial de barreira com dois estados conformacionais bem definidos.

Por fim, para a GlcN6PD, este trabalho ainda questiona a influência do ativador alostérico na transição. As dinâmicas simples provaram a espontânea formação do sítio alostérico, e a dinâmica com ligante mostrou que um fosfato ligado é capaz de estabilizar o sítio em que se encontra (e mais um vizinho) no estado R, mas, até o presente momento, sua ligação não mostrou influenciar na convergência dos outros protômeros para o conformero R (conforme a figura 4.8).

Concluindo, este trabalho apresentou uma eficiente metodologia para descrever a

transição alostérica, sendo merecedora de um maior desenvolvimento. E também contribuiu com algumas perspectivas sobre a frequência e a simetria que caracterizam este movimento.

6 Perspectivas

Uma continuação deste trabalho englobaria os seguintes tópicos:

- Prolongar as dinâmicas iniciadas e analisá-las como proposto nesta dissertação, para avaliar se a transição alostérica ocorre para todos os protômeros, e se, então, o movimento seria bem representado por uma soma de oscilações harmônicas espaciais calculada por PCA;
- Prolongar também a dinâmica com o fosfato ligado para analisar se existe a convergência dos outros sítios alostéricos para a conformação R;
- Tentar reescrever a trajetória somente com alguns modos encontrados por PCA, e então buscar as reais amplitudes e frequências com que as conformações indicadas por PCA são visitadas;
- Comparar autovetores de distintas análises (um de ANM e outro de PCA, por exemplo) e avaliar quão correspondentes eles são;
- Aprimorar os algoritmos escritos durante a realização deste projeto de forma a torná-los abrangentes o suficiente para aplicar a metodologia sugerida nesta dissertação a outras proteínas alostéricas;
- Desenvolver análises de Fourier para encontrar as frequências de oscilações que compõem o movimento simulado por dinâmica molecular;
- Estender as análises desenvolvidas neste trabalho a outras proteínas alostéricas, e preferencialmente menores, para ampliar as conclusões sobre o alosterismo;

REFERÊNCIAS

1 MONOD, J.; WYMAN, J.; CHANGEUX, J. P. On the nature of allosteric transitions: a plausible model. **Journal of Molecular Biology**, v. 12, p. 88-118, May 1965.

2 VILLALI, J.; KERN, D. Choreographing an enzyme's dance. **Current Opinion in Chemical Biology**, v. 14, n. 5, p. 636-43, Oct 2010.

3 HORJALES, E. et al. The allosteric transition of glucosamine-6-phosphate deaminase: the structure of. **Structure**, v. 7, n. 5, p. 527-37, May 1999.

4 PERUTZ, M. F. **Mechanisms of cooperativity and allosteric regulation in proteins**. Great Britain: Press Syndicate of the University of Cambridge, 1990. 103p.

5 LEHNINGER, A.; NELSON, D.; COX, M. **Principles of biochemistry**. New York: W. H. Freeman, 2008.

6 CUI, Q.; KARPLUS, M. Allostery and cooperativity revisited. **Protein Science**, v. 17, n. 8, p. 1295-307, Aug. 2008.

7 KOSHLAND, D. E.; NÉMETHY, G.; FILMER, D. Comparison of experimental binding data and theoretical models in proteins containing subunits. **Biochemistry**, v. 5, n. 1, p. 365-385, 1966.

8 GOLDSTEIN, H.; POOLE, C. **Classical mechanics**. San Francisco: Addison Wesley, 2001.

9 CHASLES, M. Note sur les proprietes generales du systeme de deux corps semblables entr'eux et places d'une maniere quelconque dans l'espace; et sur le deplacement fini ou infiniment petit d'un corps solide libre. **Bulletin des Sciences Mathematiques, Astronomiques, Physiques et Chimiques**, v. 14, p. 321-326, 1830.

10 OKAZAKI, K.-I.; TAKADA, S. Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms. **Proceedings of the National Academy of Sciences**, v. 105, n. 32, p. 11182-11187, 2008.

11 FRAUENFELDER, H.; SLIGAR, S. G.; WOLYNES, P. G. The energy landscapes and

motions of proteins. **Science**, v. 254, p. 1598-1603, 1991.

12 FERSHT, A. **Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding**. New York: W. H. Freeman, 1998.

13 HAMMES-SCHIFFER, S.; BENKOVIC, S. J. Relating protein motion to catalysis. **Biochemistry**, v. 75, p. 519-541, 2006.

14 HOOD, D. A. Invited Review: contractile activity-induced mitochondrial biogenesis in skeletal. **Journal of Applied Physiology**, v. 90, n. 3, p. 1137-57, Mar. 2001.

15 BAHAR, I.; CHENNUBHOTLA, C.; TOBI, D. Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. **Current Opinion in Structural Biology**, v. 17, n. 6, p. 633-640, 2007.

16 KERN, D.; ZUIDERWEG, E. R. P. The role of dynamics in allosteric regulation. **Current Opinion in Structural Biology**, v. 13, n. 6, p. 748-757, 2003.

17 CALCAGNO, M. et al. Purification, molecular and kinetic properties of glucosamine-6-phosphate isomerase (deaminase) from *Escherichia coli*. **Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology**, v. 787, n. 2, p. 165-173, 1984.

18 ÁLVAREZ-AÑORVE, L. I.; CALCAGNO, M. L.; PLUMBRIDGE, J. Why does *Escherichia coli* grow more slowly on Glucosamine than on N-Acetylglucosamine? Effects of enzyme levels and allosteric activation of GlcN6P Deaminase (NagB) on growth rates. **Journal of Bacteriology**, v. 187, n. 9, p. 2974-2982, 2005.

19 MONTERO-MORÁN, G. M. et al. On the multiple functional roles of the active site histidine in catalysis and allosteric regulation of *Escherichia coli* Glucosamine 6-phosphate Deaminase. **Biochemistry**, v. 40, n. 34, p.10187-10196, 2001

20 ALTAMIRANO, M. M. et al. Asymmetric allosteric activation of *Escherichia coli* glucosamine-6-phosphate deaminase produced by replacements of Tyr 121. **Biochemistry**, v. 34, n. 18, p. 6074-6082, 1995.

21 ALTAMIRANO, M. M. et al. Secondary structure of *Escherichia coli* glucosamine-6-phosphate deaminase from amino acid sequence and circular dichroism spectroscopy. **Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology**, v.

1076, n. 2, p. 266-272, 1991.

22 RUDIÑO-PIÑERA, E. et al. Structural flexibility, an essential component of the allosteric activation in *Escherichia coli* glucosamine-6-phosphate deaminase. **Acta Crystallographica Section D**, v. 58, n. 1, p. 10-20, 2002.

23 OLIVA, G. et al. Structure and catalytic mechanism of glucosamine 6-phosphate deaminase from *Escherichia coli* at 2.1 Å resolution. **Structure**, v. 3, n. 12, p. 1323-1332, 1995.

24 HOLBROOK, S. R. et al. Local mobility of nucleic acids as determined from crystallographic data: II. Z-form DNA. **Journal of Molecular Biology**, v. 187, n. 3, p. 429-440, 1986.

25 STEC, B.; ZHOU, R.; TEETER, M. M. Full-matrix refinement of the protein crambin at 0.83 Å and 130 K. **Acta Crystallographica Section D**, v. 51, n. 5, p. 663-681, 1995.

26 RHODES, G. **Crystallography made crystal clear:** guide for users of macromolecular models (complementary science). 3rd.ed. London: Academic Press, 2006.

27 BUSTOS-JAIMES, I. et al. On the role of the conformational flexibility of the active-site lid on the allosteric kinetics of Glucosamine-6-phosphate Deaminase. **Journal of Molecular Biology**, v. 319, n. 1, p. 183-189, 2002.

28 WEBER, G. Energetics of ligand binding to proteins. In: ANFINSEN, C. B.; EDSALL, J. T.; RICHARDS, F. M. (Ed.). **Advances in Protein Chemistry**. New York: Academic Press, 1975. v.29, p.1-83.

29 REIF, F. **Fundamentals of statistical and thermal physics**. New York: McGraw-Hill, 1965.

30 LEVITT, M. Molecular dynamics of native protein: I. computer simulation of trajectories. **Journal of Molecular Biology**, v. 168, n. 3, p. 595-617, 1983.

31 BERENDSEN, H. **Simulating the physical world:** hierarchical modeling from quantum mechanics to fluid dynamics. Cambridge: Cambridge University Press, 2007.

32 ANTONIOU, D.; SCHWARTZ, S. D. Low-frequency collective motions in proteins. **Journal of Theoretical and Computational Chemistry**, v. 02, n. 02, p. 163-169, 2003.

33 HAYWARD, S.; GROOT, B. L. Normal modes and essential dynamics. In: KUKOL, A.(Ed.). **Molecular modeling of proteins**. Totowa,NJ: Humana Press,2008. p.89-106. (Methods in molecular biology, v.443).

34 CASE, D. A. Normal mode analysis of protein dynamics. **Current Opinion in Structural Biology**, v. 4, n. 2, p. 285-290, 1994.

35 AMADEI, A.; LINSSSEN, A. B. M.; BERENDSEN, H. J. C. Essential dynamics of proteins. **Proteins: Structure, Function, and Bioinformatics**, v. 17, n. 4, p. 412-425, 1993.

36 PARAK, F.; FRAUENFELDER, H. Protein dynamics. **Physica A: Statistical Mechanics and its Applications**, v. 201, n. 1-3, p. 332-345, 1993.

37 LEVITT, M. et al. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. **Computer Physics Communications**, v. 91, n. 1-3, p. 215-231, 1995.

38 HESS, B. et al. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. **Journal of Chemical Theory and Computation**, v. 4, n. 3, p. 435-447, 2008.

39 VAN DER SPOEL, D. et al. **Gromacs user manual versão 4.5.4**. Disponível em <<http://www.gromacs.org/>> Acesso em: 7 set. 2012.

40 ARFKEN, G. **Mathematical methods for physicists**. Orlando, FL: Academic Press, 1985. p.428-436.

41 SHEWCHUK, J. R. **An introduction to the conjugate gradient method without the agonizing pain**. Carnegie Mellon University. 1994 Disponível em <<http://www.cs.cmu.edu/~jrs/jrspapers.html>> Acesso em: 28 nov. 2012

42 HESTENES, M. R.; STIEFEL, E. Methods of conjugate gradients for solving linear systems. **Journal of Research of the National Bureau of Standards**, v. 49, n. 6, p. 28, 1952.

- 43 PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1781-1802, 2005.
- 44 BERMAN, H. M. et al. The protein data bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235-42, Jan. 2000.
- 45 EMSLEY, P. et al. Features and development of Coot. **Acta Crystallographica Section D**, v. 66, n. 4, p. 486-501, 2010.
- 46 MARK, P.; NILSSON, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. **The Journal of Physical Chemistry A**, v. 105, n. 43, p. 9954-9960, 2001.
- 47 HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33-38, 1996.
- 48 BROOKS, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, v. 4, n. 2, p. 187-217, 1983.
- 49 MARTYNA, G. J.; TOBIAS, D. J.; KLEIN, M. L. Constant pressure molecular dynamics algorithms. **The Journal of Chemical Physics**, v. 101, n. 5, p. 4177-4189, 1994.
- 50 FELLER, S. E. et al. Constant pressure molecular dynamics simulation: the Langevin piston method. **The Journal of Chemical Physics**, v. 103, n. 11, p. 4613-4621, 1995.
- 51 PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. **Journal of Computational Chemistry**, v. 26 p. 1781-1802, 2005.
- 52 BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **The Journal of Chemical Physics**, v. 126, n. 1, p. 014101-7, 2007.
- 53 PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: a new molecular dynamics method. **Journal of Applied Physics**, v. 52, n. 12, p. 7182-7190, 1981.
- 54 GRANT, B. J. et al. Bio3d: an R package for the comparative analysis of protein

structures. **Bioinformatics**, v. 22, n. 21, p. 2695-2696, 2006.

55 HAYWARD, S.; KITAO, A.; GO, N. Harmonic and anharmonic aspects in the dynamics of BPTI: a normal mode analysis and principal component analysis. **Protein Science**, v.3 n.6 p. 936-943, 2008.

56 MARTINEZ, L. **LM-Mdanalysis**. Disponível em: < <http://lm-mdanalysis.googlecode.com> >. Acesso em: 13 out. 2011.

57 PRESS, W. et al. **Numerical recipes in Fortran: the art of scientific computing**. Cambridge: Cambridge University Press, 1992.

58 KEARSLEY, S. On the orthogonal transformation used for structural comparisons. **Acta Crystallographica Section A**, v. 45, n. 2, p. 208-210, 1989.

59 LARA-GONZÁLEZ, S. et al. On the role of the N-terminal group in the allosteric function of glucosamine-6-phosphate deaminase from *Escherichia coli*. **Journal of Molecular Biology**, v. 301, n. 1, p. 219-227, 2000.

60 IN K S C A P E : Draw Freely. Disponível em: < <http://inkscape.org/> >. Acesso em: 08 jan. 2013

61 QTIPLLOT. Disponível em <<http://softproindependent.com/qtiplot.html>>. Acesso em: 08 jan. 2013

62 WOLYNES, P. G. Symmetry and the energy landscapes of biomolecules. **Proceedings of National Academy of Science U S A**, v. 93, n. 25, p. 14249-55, 1996.

63 Symmetry and function of biological systems at the macromolecular level. In: ENGSTRÖM, A. e STRANDBERG, B. **Nobel symposium no. 11**. Sweden.: John Wiley and Sons, 1969. p.436.