



**UNIVERSIDADE FEDERAL DE GOIÁS
CAMPUS JATAÍ
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS
À SAÚDE**

Raisa Melo Lima

Busca de peptídeos com potencial modulador da enzima
malato sintase, a partir de estudos de interação proteína-
proteína

JATAÍ, 2016

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1 **1. Identificação do material bibliográfico:** ☒ **Dissertação** ☐ **Tese**

1 **2. Identificação da Tese ou Dissertação**
2

Nome completo do autor: Raisa Melo Lima

Título do trabalho: Busca de peptídeos com potencial modulador da enzima malato sintase, a partir de estudos de interação proteína-proteína

3. Informações de acesso ao documento:

Concorda com a liberação total do documento ☒ SIM ☐ NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Raisa melo lima
Assinatura do (a) autor (a)

Data: 03 / 11 / 2016

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

Raisa Melo Lima

Busca de peptídeos com potencial modulador da enzima
malato sintase, a partir de estudos de interação proteína-
proteína

Dissertação apresentada como
parte dos requisitos para obtenção
do título de Mestre, no Programa
de Pós-graduação em Ciências
Aplicadas à Saúde da
Universidade Federal de Goiás

Orientador: Prof. Dr. Roosevelt
Alves da Silva

JATAÍ, 2016

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Lima, Raísa Melo

Busca de peptídeos com potencial modulador da enzima malato
sintase, a partir de estudos de interação proteína-proteína [manuscrito] /
Raísa Melo Lima. - 2016.

XC, 90 f.: il.

Orientador: Prof. Dr. Roosevelt Alves da Silva.

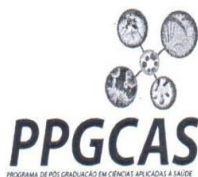
Dissertação (Mestrado) - Universidade Federal de Goiás, Unidade
Acadêmica Especial de Ciências Exatas e Tecnológicas, Programa
de Pós-Graduação em Ciências Aplicadas à Saúde, Jataí, 2016.

Bibliografia. Anexos. Apêndice.

Inclui abreviaturas, lista de figuras, lista de tabelas.

1. Malato sintase. 2. Paracoccidioides spp.. 3. Ancoragem
molecular. 4. Dinâmica molecular. 5. Desenho de peptídeos. I. da Silva,
Roosevelt Alves, orient. II. Título.

CDU 577.32



SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DE GOIÁS
CAMPUS JATAÍ

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS APLICADAS À SAÚDE

Rod. BR 364, km 192, n. 3800, Setor Industrial
75.801.615 – Jataí/GO- Brasil

ATA DE DEFESA PÚBLICA DE DISSERTAÇÃO DE MESTRADO

Ao terceiro dia do mês de Outubro do ano de dois mil e dezesseis às 14 horas no auditório Unidade Jatobá, Regional Jataí da universidade Federal de Goiás, realizou-se a sessão pública de defesa de dissertação, intitulada **“Busca de peptídeos com potencial modulador da enzima Malato sintase, a partir de estudos de interação proteína-proteína”**. De autoria do Candidato **Raísa Melo Lima** aluno do Programa de Pós-Graduação em Ciências Aplicadas a Saúde, em nível de Mestrado. A Comissão Examinadora esteve constituída pelos professores: Prof. Dr. Roosevelt Alves da Silva, Presidente. Prof^ª. Dr^ª. Maristela Pereira e Prof^ª. Dr^ª. Mirian Machado Mendes.

Concluídos os trabalhos de apresentação e arguição, os membros da banca consideraram o projeto de dissertação:

(X) aprovado

() reprovado

Banca Examinadora:


Prof. Dr. Roosevelt Alves da Silva. UFG/Jataí


Prof. Dr.ª. Maristela Pereira. UFG/Goiânia


Prof. Dr.ª. Mirian Machado Mendes. UFG/Jataí

Foi concedido um prazo máximo de 30 dias, a contar da data de defesa o candidato a efetuar as correções sugeridas pela Comissão Examinadora e apresentar o trabalho em redação definitiva, sob pena de não expedição do Diploma. E, para constar, foi lavrada a presente ata, que vai assinada pelos membros da Comissão.

AGRADECIMENTOS

À minha família, em especial meus pais Raimundo e Maria Virgínia, que sempre me proporcionaram boa educação, e estiveram ao meu lado em todos os momentos.

Às minhas filhas, Júlia e Sophia, minha afilhada-irmã Ana Clara, e meu irmão Renan, que são minhas razões de querer algo melhor, foi por elas todo o sacrifício desprendido, e será por elas que seguirei os meus objetivos.

Ao meu marido, Amaro Vinícius, pelo companheirismo e paciência durante essa fase tão turbulenta em nossas vidas.

À minha sogra Fátima, meus cunhados Hugo e Franciny, pelo carinho e paciência, principalmente nos últimos momentos, que foram de tanta correria.

Ao meu orientador, Dr.^o Roosevelt, pelos ensinamentos, e por me acolher como sua aluna, mesmo com as minhas dificuldades, sendo sempre atento e prestativo. Muito bom ter feito parte do seu time! Isso não é uma despedida! Serei sempre a primeira dama do NCBios, certo?!

Aos meus amigos e colegas do NCBios, Uessiley, Ricardo e Fausto, e às novatas Mirla, Diva, Gabriela, e em especial a Lu Sussuchi, que se tornou uma grande amiga de todas as horas, de estudo, de conselhos, de desesperos, de comilanças e bebedeiras, KKK. Você é demais Sushi!!

À professora Dr.^a Maristela Pereira pela sua parceria e oportunidades oferecidas a nós alunos.

Às agências de fomento pelo financiamento desta pesquisa.

À todos que fizeram parte dessa história de alguma maneira, meu muito obrigada.

E a Deus, claro, pois tudo só é possível com a permissão Dele.

LISTA DE ABREVIATURAS

Acetil-CoA	Acetil coenzima A
AIDS	<i>Acquired immunodeficiency syndrome</i>
CASP	<i>Critical Assessment of Protein Structure</i>
cDNA	DNA complementar
DM	Dinâmica Molecular
ELISA	<i>Enzyme Linked Immuno Sorbent Assay</i>
FDA	<i>Food and Drug Administration</i>
FFT	Transformadas de Fourier
HIV	<i>Human Immunodeficiency vírus</i>
ID	Ensaio de Imunodifusão
IPP	Interface Proteína-proteína
PI	Proteínas Ligantes
MAPT	Proteína associada a microtúbulo
MEC	Matriz Extracelular
MLS	Malato sintase
MS	Espectrometria de massas
OMS	Organização Mundial da Saúde
pb	Pares de bases
<i>PbMLS</i>	Malato sintase de <i>Paracoccidioides brasiliensis</i>
PCM	Paracoccidioidomicose
PCR	Reação em Cadeia da polimerase
PDB	<i>Protein Databank</i>
PPI	Interação proteína-proteína
RMN	Ressonância Magnética Nuclear
RMSD	<i>Root mean square deviation</i>
SNC	Sistema nervoso central
3D	Tridimensional

LISTA DE TABELAS

Tabela 1. Comparativo dos preditores de <i>hotspots</i>	20
Tabela 2. Aspectos gerais das proteínas envolvidas no estudo.....	35
Tabela 3. Combinações realizadas no <i>molecular docking</i>	40
Tabela 4. Regiões da MLS observadas com maior frequência nas interações proteína-proteína	45
Tabela 5. Peptídeos selecionados para interação com o modo 1 da MLS, e comparação das suas estruturas com os modelados <i>ab initio</i>	47
Tabela 6. Peptídeos selecionados para interação com o modo 2 da MLS, e comparação das suas estruturas com os modelados <i>ab initio</i>	47
Tabela 7. Peptídeos selecionados para interação com o modo 3 da MLS, e comparação das suas estruturas com os modelados <i>ab initio</i>	48

LISTA DE FIGURAS

Figura 1. Amostra de tecido com leveduras de <i>Paracoccidioides</i> spp.	3
Figura 2. Mapa epidemiológico.	5
Figura 3. Ciclo biológico de <i>Paracoccidioides</i> spp.	6
Figura 4. Tecido epitelial.	8
Figura 5. Ciclo do Glioxilato, ciclo do ácido cítrico e gliconeogênese.	10
Figura 6. Princípio do sistema duplo-híbrido em levedura.	12
Figura 7. Esquema do ensaio de pull-down <i>in vitro</i>	13
Figura 8. Etapas essenciais no desenvolvimento de um fármaco.	14
Figura 9. Emprego de peptídeos na terapêutica.	21
Figura 10. Rede de interações de MLS.	30
Figura 11. Rede de interações de malato desidrogenase.	30
Figura 12. Rede de interações de Enolase.	31
Figura 13. Rede de interações de Triosefosfato isomerase.	31
Figura 14. Rede de interações de Ubiquitina.	32
Figura 15. Rede de interações de GPD.	32
Figura 16. Rede de interações das proteínas envolvidas no estudo.	33
Figura 17. Gráfico de clusters da MLS.	36
Figura 18. Gráfico de RMSD e estrutura tridimensional da MLS.	36
Figura 19. Diagrama de Ramachandran da Malato sintase antes e depois da DM.	37
Figura 20. Alinhamento dos modelos da MLS.	38
Figura 21. Modelos 3D dos complexos Malato sintase e Enolase.	41
Figura 22. Representação tridimensional do complexo MLS e Malato Desidrogenase.	41
Figura 23. Representação tridimensional do complexo MLS e Gama actina.	42
Figura 24. Representação tridimensional do complexo MLS e Gliceraldeído-3-fosfato desidrogenase.	42
Figura 25. Representação tridimensional do complexo MLS e Enolase.	43
Figura 26. Representação tridimensional do complexo MLS e TFI.	43
Figura 27. Representação tridimensional do complexo MLS e Ubiquitina.	44
Figura 28. Regiões frequentes de interação da MLS.	45
Figura 29. Gráfico das regiões frequentes de interação da MLS.	46
Figura 30. Peptídeo 3 ancorado à superfície da MLS caracterizada quanto a sua hidrofobicidade.	50
Figura 31. Peptídeo 3 ancorado à superfície da MLS, evidenciando suas interações intermoleculares.	51
Figura 32. Peptídeo 4 em contato com a superfície da MLS, caracterizada quanto a sua hidrofobicidade.	52
Figura 33. Peptídeo 4 ancorado à superfície da MLS, evidenciando suas interações intermoleculares.	52
Figura 34. Peptídeo 7 ancorado à superfície hidrofóbica da MLS.	53
Figura 35. Peptídeo 7 ancorado à superfície da MLS, evidenciando suas interações intermoleculares.	54

Figura 36. Peptídeo 10 ancorado a superfície da MLS evidenciando as interações intermoleculares.	55
Figura 37. Peptídeo 14 ancorado à superfície da MLS, evidenciando a superfície hidrofóbica de contato.	55
Figura 38. Peptídeo 14 ancorado à superfície da MLS	56

RESUMO

A paracoccidioidomicose (PCM) é uma micose sistêmica endêmica no Brasil, onde são registrados cerca de 80% dos casos mundiais, e possui como agente etiológico o fungo *Paracoccidioides* spp. A Malato sintase (MLS) é uma importante enzima relacionada ao metabolismo fúngico, uma vez que é essencial no ciclo do glioxilato, uma via metabólica importante de produção de glicose para parede celular, sendo exclusiva de micro-organismos e plantas. Sua ausência em humanos a torna um alvo interessante de estudo, principalmente, no desenho racional de fármacos. A partir de recentes estudos *in vitro*, várias proteínas que interagem com a *PbMLS* (receptor) foram classificadas, porém os modos de interação e as regiões chaves envolvidas nas interfaces proteína-proteína (IPP's), não foram ainda descritas. Neste trabalho, 6 (seis) proteínas ligantes (PL) foram selecionadas para verificar suas interações com *PbMLS*. As estruturas tridimensionais dessas proteínas, bem como de *PbMLS*, foram preditas por homologia, usando o servidor I-TASSER. Simulações de dinâmica molecular (DM) foram realizadas pelo programa GROMACS, e os modos das conformações mais representativas de cada proteína foram determinados baseando-se nas análises de agrupamentos a partir das trajetórias geradas por DM. Simulações de ancoragem molecular com GRAMM-X foram então realizadas entre os modos conformacionais de *PbMLS* contra os obtidos de PL's, resultando num total de 36 complexos. Baseado na frequência maior de alguns pequenos fragmentos de proteínas, observados nas IPP's, 57 peptídeos de tamanhos entre 5 e 20 resíduos de aminoácidos, foram inicialmente selecionados a partir de 5 regiões da *PbMLS* consideradas mais frequentes na interação proteína-proteína. Simulações com FlexPepDock foram realizadas para otimizar as coordenadas atômicas dos peptídeos complexados com MLS e, concomitantemente, simulações com PepFOLD foram realizadas para avaliar a estabilidade de cada peptídeo em solução. Com base nos mais baixos scores de energia dos peptídeos ligados a MLS bem como na estabilidade de suas estruturas não ligadas em solução, 5 peptídeos foram selecionados como promissores ligantes ao modo 1 de MLS. A estabilidade e os padrões de interações destes peptídeos são avaliados em detalhes.

Palavras-chaves: desenho de peptídeos, dinâmica molecular, interação proteína-proteína, malato sintase, *molecular docking* e *Paracoccidioides* spp.

ABSTRACT

Paracoccidioidomycosis (PCM) is a systemic mycosis endemic in Brazil, where are recorded about 80% of cases worldwide, and has *Paracoccidioides* sp. as the etiologic agent. Malate synthase (MLS) is an important enzyme related to the fungal metabolism, once it is essential in the glyoxylate cycle, a secondary metabolic pathway of the citric acid cycle exclusive to microorganisms and plants. Its absence in humans makes this enzyme an interesting subject to study, mainly in rational drug design. From recent in vitro studies, several interacting proteins of MLS (receiver) were classified, but the modes of interaction and key regions involved in protein-protein interfaces (PPIs) have not yet been described. In this work, six (6) binding proteins (BPs) were selected to describe the IP's of MLS. Their tridimensional structures, as well as MLS, were predicted by homology modeling using I-TASSER server, and subsequent molecular dynamics simulations (MD). The most common conformational modes of each protein were obtained by cluster analysis of the trajectories generated by MD. Molecular docking simulations using Gramm-X were then performed for the conformational modes of MLS against the BP's, resulting in a total of 36 complexes. Based on the higher frequency of some small fragments of proteins observed in the IPP's, 57 peptides with sizes between 5 and 20 residues, were initially selected from 5 regions of MLS which are considered more frequent in protein-protein interaction. FlexPepDock simulations were performed to optimize the atomic coordinates of the peptide complexed with MLS, and concomitantly, PepFOLD simulations were performed to evaluate the stability of each peptide in solution. Based on the lower energy score of peptides linked to MLS, and the stability of their structures in solution (MLS-free), 6 peptides were selected as promising ligands to MLS mode 1. The stability and patterns of interactions of these peptides were evaluated in detail.

Keywords: malate synthase, molecular docking, molecular dynamics, *Paracoccidioides* sp, peptide design and protein-protein interaction.

SUMÁRIO

AGRADECIMENTOS	i
LISTA DE ABREVIATURAS.....	vii
LISTA DE TABELAS	viii
LISTA DE FIGURAS	ix
RESUMO.....	xi
ABSTRACT.....	xii
SUMÁRIO	1
1. INTRODUÇÃO	3
1.1. A paracoccidioidomicose (PCM)	3
1.1.1. Epidemiologia	4
1.1.2. Patogênese.....	5
1.1.3. Aspectos clínicos	6
1.1.4. Imunopatologia da PCM	7
1.1.5. Diagnóstico	8
1.2. Importância da <i>PbMLS</i>	9
1.2.1. Rede de interações proteicas	11
1.3. Aplicação da bioinformática na busca de alvos para drogas	13
1.4. Metodologias computacionais	15
1.4.1. Modelagem molecular de proteínas.....	15
1.4.2. Dinâmica molecular	16
1.4.3. <i>Molecular docking</i>	17
1.4.4. Peptidologia	18
2. JUSTIFICATIVA	22
3. OBJETIVOS	23
3.1. Objetivo geral	23
3.2. Objetivos específicos	23
4. PROCEDIMENTOS EXPERIMENTAIS	24
4.1. Obtenção das estruturas 3D	24
4.2. Simulações de Dinâmica Molecular.....	25
4.3. Ancoragem molecular.....	26

4.4. Design e ancoragem dos peptídeos	26
5. RESULTADOS E DISCUSSÕES	28
5.1. Modelos estruturais das proteínas	28
5.1.1. Malato sintase (<i>PbMLS</i>).....	28
5.2. Rede de interações de proteínas	28
5.3. Dinâmica Molecular e Qualidade dos modelos	33
5.3.1. Análise dos modos da MLS	33
5.4. Complexos formados e interface de interação.....	39
5.5. Análises das regiões mais interagentes de MLS	44
5.6. Design de peptídeos	46
5.6.1. Aspectos gerais dos peptídeos	46
5.7. Avaliação das interações proteína-peptídeo	49
5.7.1. Peptídeos promissores ancorados ao Modo 1 da MLS	49
6. CONSIDERAÇÕES FINAIS	57
6.1. Conclusões	57
6.2. Perspectivas	57
7. REFERÊNCIAS	59
ANEXO 1 – Alguns peptídeos usados na terapêutica atualmente	68
APÊNDICE 1.....	69
1. Características gerais das proteínas ligantes de MLS	69
1.1. Enolase (EC 4.2.1.11).....	69
1.2. Ubiquitina (PAAG_07080).....	70
1.3. Triosefosfato isomerase (PAAG_02585 EC- 5.3.1.1)	72
1.4. Gamma-actina	74
1.5. Malato desidrogenase.....	75
1.6. Gliceraldeído-3-fosfato desidrogenase (EC- 1.2.1.12).....	76

1. INTRODUÇÃO

1.1. A paracoccidioidomicose (PCM)

A PCM é uma micose sistêmica causada pelo fungo *Paracoccidioides* spp., sendo endêmica e importante na América Latina (BRUMMER et al., 1993). Estes fungos são termo-dimórficos e foram descritos primeiramente pelo médico Adolpho Lutz em 1908, que isolou estes microorganismos a partir de pacientes que apresentavam lesões na mucosa oral (LUTZ, 1908; STURME et al., 2011). De acordo com Prado et al. (2009), a PCM foi responsável por mais de 50% dos casos de micose sistêmica no Brasil, entre 1996 e 2006.

Os fungos *Paracoccidioides* spp. podem se apresentar de duas formas, dependendo da temperatura do ambiente onde estão. Vivendo de forma sapróbia no solo (temperatura abaixo de 25° C) encontram-se como micélio, esta é composta por hifas finas e septadas com poucos clamidósporos e vários núcleos. Nos tecidos humanos (temperatura acima de 35°C) assumem a forma de levedura (BAGAGLI et al, 2006), apresentando uma estrutura multinuclear com muitos brotamentos (FRANCO et al, 1987; STURME et al, 2011), como mostrado na Figura 1.

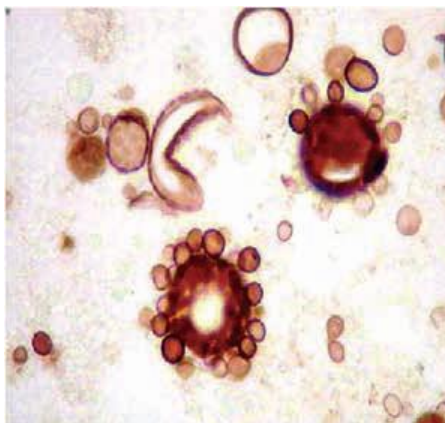


Figura 1. Amostra de tecido com leveduras de *Paracoccidioides* spp. Observar os brotamentos em aspecto de roda de leme. Fonte: GÓES et al., 2014.

Até o ano de 2006 o gênero *Paracoccidioides* incluía apenas uma espécie, *P. brasiliensis*. Porém, estudos mais recentes de Matute et al. (2006) avaliaram polimorfismos genéticos, e observaram a existência de 3 espécies filogeneticamente distintas. Em 2009, estudos de Teixeira et al, identificaram uma nova espécie, distinta das 3 anteriores, endêmica na região centro-oeste do Brasil, sendo classificada como Pb01. Esta nova espécie vem sendo

denominada para *Paracoccidioides lutzii*, desde meados de 2009 nos trabalhos de Teixeira et al., em homenagem ao micologista Adolpho Lutz (BOCCA, et al 2013).

1.1.1. Epidemiologia

A PCM é uma doença fúngica que geralmente é adquirida em países latino-americanos, com maior prevalência na América do Sul. A PCM tem um grande impacto social e médico nas áreas endêmicas, não apenas pelo número de casos, mas também pela cronicidade da doença, com longos períodos de tratamento, que geralmente limita o paciente a realizar atividades rotineiras, visto que pode desencadear insuficiência respiratória a partir de uma fibrose pulmonar semelhante à tuberculose (MARTINEZ,2015). A abordagem da epidemiologia da doença é complicada, embora em alguns estados brasileiros a PCM integre o rol de doenças de notificação compulsória, isso ainda não é uma regra. de difícil diagnóstico laboratorial, principalmente na fase precoce da doença (MARTINEZ, 2010).

De acordo com a OMS, o Brasil é o país com maior incidência de PCM no mundo, sendo considerada endêmica nas regiões Sudeste e Centro-Oeste, como se pode observar na Figura 2 (MARQUES, 2013). A letalidade da PCM varia de 2 a 23%, podendo chegar a 30% dos casos associados com AIDS. Cerca de 80% da incidência ocorre no Brasil e os outros 20% na Venezuela, Colômbia e Argentina. Por se tratar de uma doença de clima subtropical, a região Nordeste possui baixa endemia da doença. Outro fator que pode diminuir a estatística desta região, pode ser o fato da maior dificuldade de notificação e diagnóstico específico nos serviços de saúde da região (MATOS et al., 2012; VIEIRA et al., 2014).

A PCM acomete mais trabalhadores rurais, do sexo masculino, provavelmente devido ao maior contato destes com o solo e a vegetação, que são habitats naturais destes fungos (MARQUES et al., 1983; RAMOS & SARAIVA, 2008). Outro motivo provável para a maior incidência em homens, seria a presença de receptores de estrógeno no citoplasma do fungo, sendo que este hormônio seria capaz de bloquear a transformação de micélio para levedura infectante (ARISTIZABAL et al., 2002).

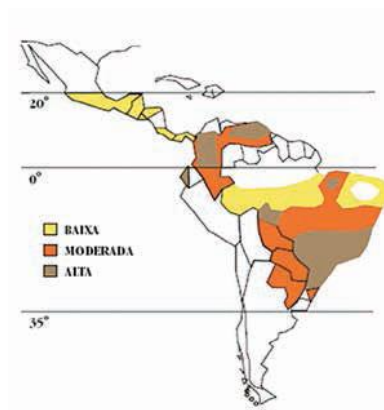


Figura 2. Mapa epidemiológico. Distribuição geográfica da PCM de acordo com o Ministério da Saúde. Fonte: GÓES et al., 2014.

1.1.2. Patogênese

Através da inalação de propágulos de micélio ou conídeos presente no ambiente, principalmente no solo, o hospedeiro adquire o fungo, que no parênquima pulmonar se diferencia em levedura. A partir daí, o fungo inicia sua reprodução assexuada, podendo se disseminar para vários órgãos via linfática e hematogênica, causando a micose sistêmica (BRUMMER et al, 1993; BAGAGLI et al., 2006). Em alguns casos o fungo pode se instalar via traumática na pele e mucosas (CASTRO et al., 1975). Alguns estudos, como o de Zanetti et al. (2011), apontam que o tatu pode ser um transmissor do fungo, como observado na Figura 3, quando utilizado como alimento pelo ser humano.

Para o sucesso da infecção, fungo *Paracoccidioides sp.* tem habilidade em aderir a componentes da matriz extracelular (MEC) através de moléculas adesinas. Algumas destas já foram descritas, como a glicoproteína de 43 kDa (gp43), a gliceraldeído-3-fosfato desidrogenase, a triosefosfato isomerase, a malato sintase e a enolase (BERNARD & MENDES-GIANNINI, 2009; ZAMBUZZI-CARVALHO et al., 2009).

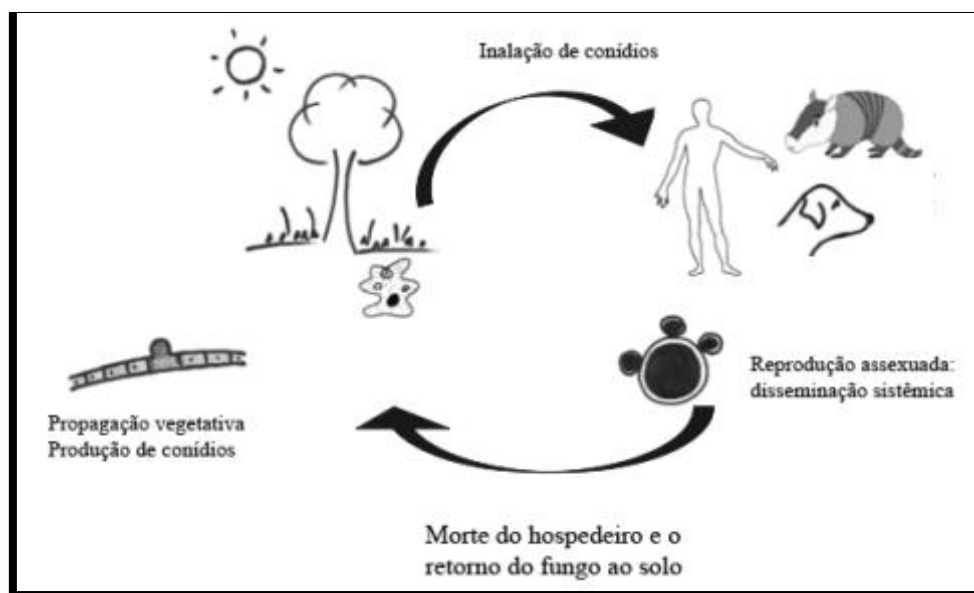


Figura 3. Ciclo biológico de *Paracoccidioides* spp. Os indivíduos ou mamíferos como tatus e cães inalam propágulos do fungo presentes no solo, que chegam ao sistema pulmonar e diferenciam em leveduras. Essas formas se multiplicam por brotamento no interior das células do hospedeiro, até que estas se rompem e as leveduras estão livres para se instalar em outras células. O fungo retorna ao solo com a morte do hospedeiro, transformando-se em micélio. Adaptado de (ELETRÔNICAS et al., 2014).

1.1.3. Aspectos clínicos

A doença tem duas apresentações clínicas importantes: aguda/subaguda e crônica. A forma crônica granulomatosa é a mais comum (cerca de 90% dos casos) e afeta indivíduos entre 30 e 60 anos, tendo um grande período de latência (BOCCA et al., 2013). Na maior parte dos casos, inicia-se nos pulmões e quando progride é de forma lenta e geralmente para pele e mucosas (MARQUES et al., 2007).

A forma aguda/subaguda é menos frequente, porém é rápida e mais grave, acometendo jovens e crianças (BOCCA et al, 2013; BARROZO et al, 2009). Desenvolve-se a partir de uma lesão primária pulmonar que progride e se dissemina via linfática e hematogênica para órgãos como baço, fígado, medula óssea, ossos e linfonodos. Neste caso, observam-se altos níveis de anticorpos específicos e uma depressão da imunidade celular, sendo a forma mais ligada à letalidade da doença (FRANCO et al, 1987).

Quando apenas um órgão ou sistema é acometido, tem-se o estado unifocal, o tipo multifocal está relacionado com o envolvimento de mais de um órgão, sendo os sítios mais acometidos os pulmões, mucosa e pele. A doença pode seguir para um período de remissão, ausente de sintomas e sinais clínicos,

porém os granulomas podem apresentar fungos viáveis, o que justificaria o reaparecimento da doença em alguns casos de deficiência imune. (SHIKANAI-YASUDA *et al*, 2006).

O estabelecimento da doença, sua gravidade e disseminação dependem de vários fatores, como a virulência da cepa, de fatores ambientais e condições do hospedeiro, como estado nutricional, e imunológico (BENARD, 2008; MENDES-GIANNINI *et al*, 2008). No caso de pacientes imunodeprimidos, como pacientes com AIDS, em tratamento quimioterápico e tuberculose, muitas vezes apresentam formas mais graves da doença (MINISTÉRIO DA SAÚDE, 2012).

1.1.4. Imunopatologia da PCM

O fungo *Paracoccidioides* spp. produz antígenos que interagem com o sistema imune provocando uma resposta imunológica complexa (BERNARD, 2008). Este patógeno apresenta uma complexa estrutura antigênica, sendo que o principal antígeno relacionado à patogenicidade é uma glicoproteína estrutural da parede do fungo (gp43). Tal antígeno está associado ao escape ao sistema imunológico, sendo este o principal fator de virulência por evadir os mecanismos de defesa do hospedeiro, possibilitando sua instalação nos tecidos. Para que isso ocorra, a glicoproteína gp43, componente da superfície celular que é liberada pelo fungo em sua fase exponencial de crescimento, possui ação lítica (proteinase) sobre proteínas de matriz extracelular, como colágeno, elastina e fibronectina, o que favorece esta instalação (VICENTINI *et al.*, 1994; MENDES-GIANINI *et al.*, 1990, 2008).

O contato inicial gera uma inflamação ao redor do fungo, com a formação de edema, congestão vascular e predomínio de polimorfonucleados, como podemos observar na Figura 4. Posteriormente, surgem as células mononucleares e linfócitos que começam a estimulação de uma resposta específica, com formação de granuloma com células gigantes multinucleadas, células epitelióides e necrose caseosa (MOREIRA, 2008; PALMEIRO, CHERUBINI, YURGEL, 2005).

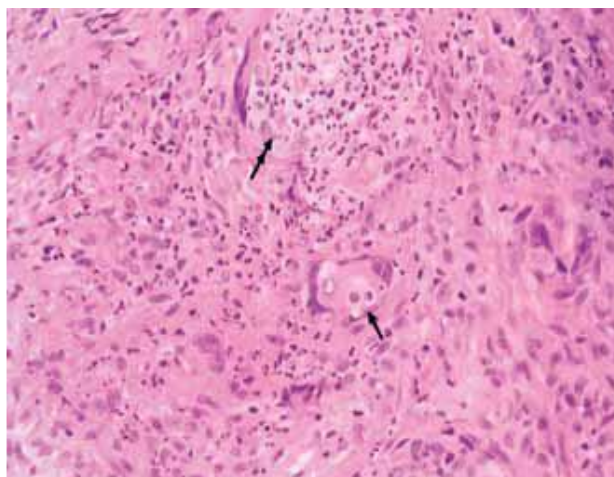


Figura 4. Tecido epitelial. Pode-se observar reação inflamatória crônica granulomatosa, com células gigantes, macrófagos, com fungos em seu interior. Coloração por HE, aumento de 200x. Fonte: FORTES et al., 2011.

1.1.5. Diagnóstico

O diagnóstico da PCM baseia-se na identificação direta do fungo, exame histopatológico, ou cultura do fungo, sendo estes considerados “padrão-ouro” para essa doença. O exame micológico direto é realizado a partir de amostras das lesões que se suspeitam estarem contaminadas, podendo ser realizado com material a fresco, ou após cultivo (MARTINEZ, 2010). A secreção ganglionar geralmente é rica em células do *Paracoccidioides* spp., e a biopsia do linfonodo é especialmente importante no diagnóstico da PCM em crianças e adolescentes, devido a linfadenomegalia observada nesta faixa etária (PEREIRA et al, 2004).

O diagnóstico sorológico também é muito útil em pacientes com suspeita de PCM, visto que estes produzem anticorpos específicos que podem ser detectados por ELISA, Imunodifusão dupla (ID) e *Western blotting*. A ID é o teste mais utilizado entre estes, por apresentar sensibilidade de cerca de 90% e especificidade de 99 a 100%. Podem ocorrer reações cruzadas com outros fungos, além disso, pacientes imunocomprometidos podem apresentar baixos níveis de anticorpos circulantes, o que dificulta a detecção pelo teste (SHIKANAI-YASUDA et al, 2006; BERNARD & MENDES-GIANNINI, 2009).

A PCR (Reação em Cadeia da Polimerase) e outras técnicas moleculares, são essenciais para identificar o genótipo circulante nos pacientes, ou em uma determinada região de estudo. Apesar de se restringirem ao âmbito da pesquisa, o diagnóstico molecular tem demonstrado sucesso na detecção do fungo, quando as outras técnicas são insatisfatórias (MARQUES, 2013).

De acordo com Ambrósio et al (2014), todos os métodos diagnósticos possuem limitações, seja por baixa sensibilidade ou especificidade, seja por reações cruzadas com outras doenças fúngicas e bacterianas, ou até mesmo por falta de padronização ou custo elevado. Desta maneira, é importante a reunião de evidências clínicas com diferentes métodos laboratoriais para obtenção de um diagnóstico seguro (AMBROSIO et al, 2014).

1.2. Importância da *PbMLS*

Dentre as vias metabólicas relacionadas aos micro-organismos patogênicos destaca-se o ciclo do glioxilato. Este ciclo permite a utilização de acetil-CoA oriunda da beta-oxidação de lipídeos, para sintetizar compostos que serão usados na síntese de carboidratos (gliconeogênese), uma vez que macrófagos, células que fagocitam o fungo, são deficientes em carboidratos e ricos em ácidos graxos (COZZONE, 1998).

As enzimas exclusivas do ciclo do glioxilato, isocitrato liase e malato sintase, tem como objetivo a conversão de duas moléculas de acetil-coA em succinato e malato. Da clivagem de isocitrato, pela enzima isocitrato liase, temos succinato e glioxalato, este é convertido em malato, pela enzima MLS, que depois será convertido em oxaloacetato, que pela gliconeogênese será convertido em glicose (Figura 5) (LENHINGER et al, 2000; LORENZ & FINK, 2001; ANDRADE, 2006).

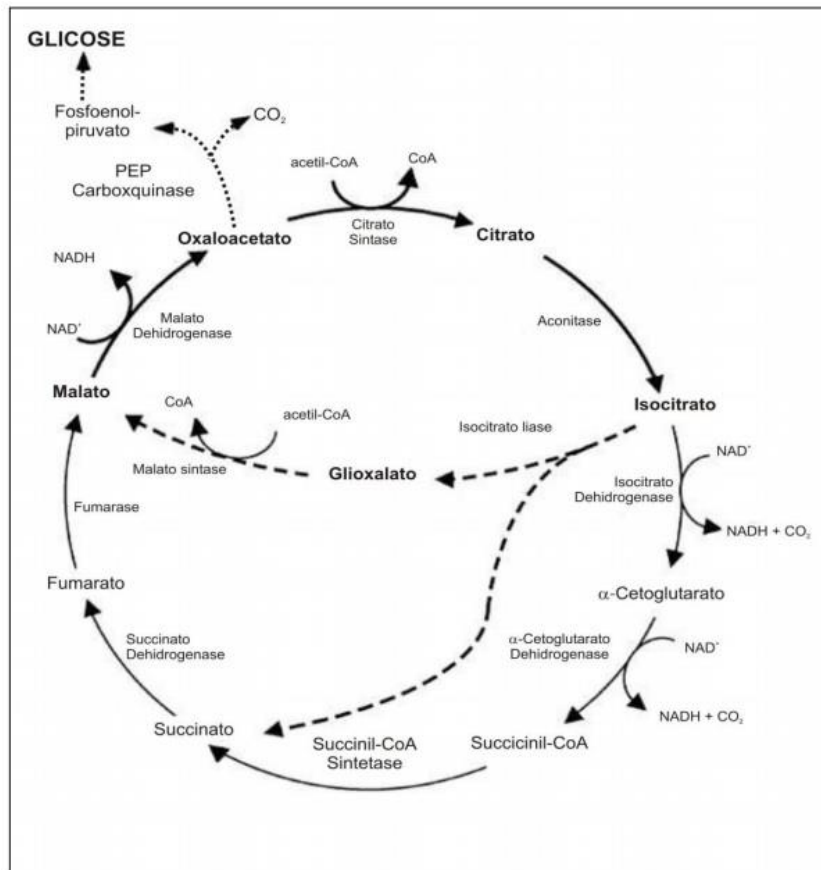


Figura 5. Ciclo do Glioxilato, ciclo do ácido cítrico e gliconeogênese. Neste esquema estão evidenciados os principais passos enzimáticos no ciclo do ácido cítrico (linhas finas), o qual é comum para todos os organismos. Os passos no ciclo do glioxilato que é um ciclo comum para micro-organismos e plantas, estão representados pelas linhas tracejadas. Já os passos compartilhados pelos dois ciclos estão representados por linhas grossas, e a reação inicial de gliconeogênese por linhas pontilhadas. Adaptado de Lorenz e Fink (2002).

O ciclo do glioxilato geralmente ocorre nos peroxissomos, onde a enzima MLS está localizada, principalmente, de acordo com todos os estudos recentes em organismos eucariotos (NETO et al., 2009). A reação de condensação catalisada pela MLS é altamente específica e necessita de Mg^{2+} como cofator (NOGALES et al., 2004).

O gene *PbMLS* do *Paracoccidioides spp.*, foi clonado por Zaambuzzi-Carvalho et al. (2009). O cDNA possui 1882 pb (pares de bases) e a sequência traduzida da proteína possui 539 resíduos de aminoácidos e massa molecular de 60 kDa. A proteína faz parte da família das MLS, apresentando resíduos característicos de atividade catalítica e localização peroxissomal e glioxissomal (ZAMBUZZI-CARVALHO et al., 2009).

As proteínas reguladoras do ciclo do glioxilato (malato sintase e isocitrato liase) são consideradas promissores alvos de fármacos, uma vez que esta via é

essencial para crescimento do patógeno e não está presente nos animais, o que poderia minimizar o risco de reações adversas (LORENZ & FINK, 2001; SELTRENNIKOFF & NAKATA, 2003).

Estudos demonstraram que ocorre, *in vitro*, super expressão dos genes que codificam as duas proteínas: malato sintase e a isocitrato liase em condições de escassez de carboidrato. Provavelmente, estes genes se comportam como fatores de virulência para favorecer a sobrevivência do fungo neste ambiente. Estudos de Derengowski (2006), confirmaram um aumento significativo na expressão desses genes quando cultivados em meio ausente em glicose e rico em acetato, simulando o ambiente do interior da célula fagocitária (FINLAY & FALCOW, 1997; TAVARES et al., 2005).

Acredita-se que a proteína *PbMLS*, além de desempenhar função enzimática no ciclo do glioxilato, possua várias outras funções. De acordo com Silva Neto et al. (2013), a proteína também possui a função de adesina, estando presente na superfície do fungo, sendo ligante de fibronectina e colágeno I e IV e participando no processo de adesão. Outro estudo demonstrou que ela interage com proteínas de diferentes categorias (transportadoras, metabólicas, transdutoras de sinais), o que sugere outros papéis na célula fúngica (OLIVEIRA et al., 2013).

1.2.1. Rede de interações proteicas

A maioria dos processos biológicos essenciais tem o envolvimento de interações entre duas ou mais proteínas; comunicação celular, expressão gênica e resposta imunológica, são alguns exemplos. A partir dessas interações podemos inferir a função biológica da proteína, conhecendo a rede de proteínas que ela interage, além de facilitar a descoberta de novos alvos para fármacos (SKRABANEK et al., 2008; AUERBACH et al, 2002).

De acordo com Oliveira et al (2013), no sistema de duplo híbrido, ocorre a fusão de duas proteínas de interesse a domínios complementares de transcrição, é uma metodologia usada para verificar a interação ou não destas proteínas. Uma vez havendo a interação ocorre a expressão dos genes responsivos a este fator de transcrição. Uma desvantagem dessa técnica é que como os domínios BD e AD (genes que codificam proteínas conhecidas) são ligados à porção N-terminal das proteínas de interesse, eventuais interações

entre elas nessa região podem ser prejudicadas. Na Figura 6, temos o esquema representativo desta técnica.

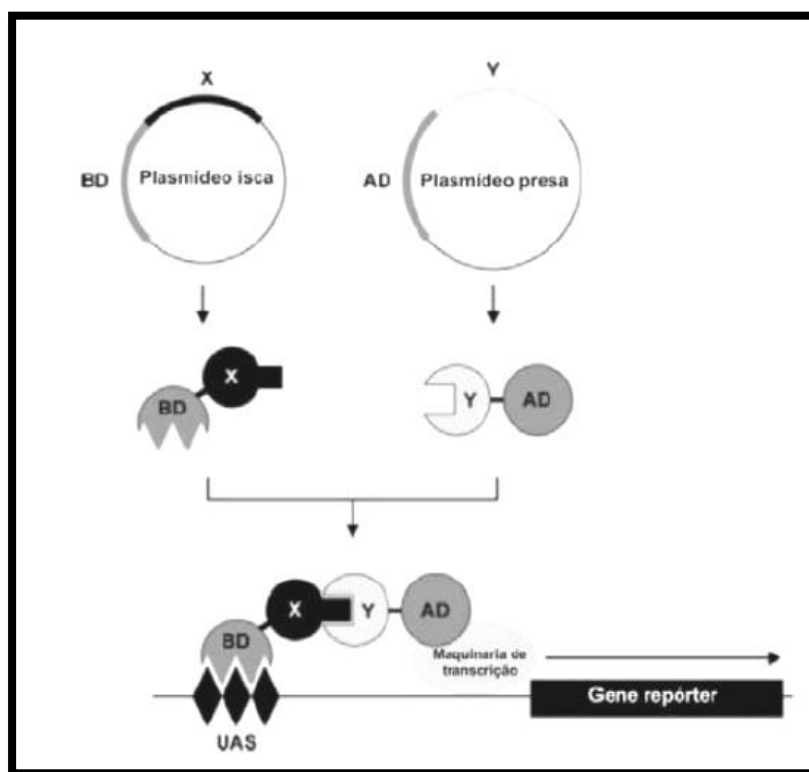


Figura 6. Princípio do sistema duplo-híbrido em levedura. Após a construção dos híbridos da isca, os plasmídeos são inseridos em uma mesma linhagem de levedura. Somente se as proteínas X e Y interagirem fisicamente, os domínios BD e AD ficarão próximos, reconstituindo o fator de transcrição ativo funcionalmente. Fonte: de Oliveira et al., 2013.

Além dessa metodologia apresentada, técnicas de purificação por afinidade podem trazer bons resultados na detecção de interação proteína-proteína (PPI). O ensaio *in vitro pull down*, consiste na fusão da proteína de interesse (*PbMLS*) à uma glutationa, geralmente a GST em uma resina. Desta forma, as proteínas que se deseja analisar, “iscas”, são incubadas com essa resina ligada à proteína “presa” (*PbMLS*). Os complexos são analisados por eletroforese em gel de poliacrilamida SDS-PAGE, e as proteínas podem ser identificadas por espectrometria de massas (MS), como esquematizado na Figura 7 (CAUSIER et al., 2004; CHO et al., 2004).

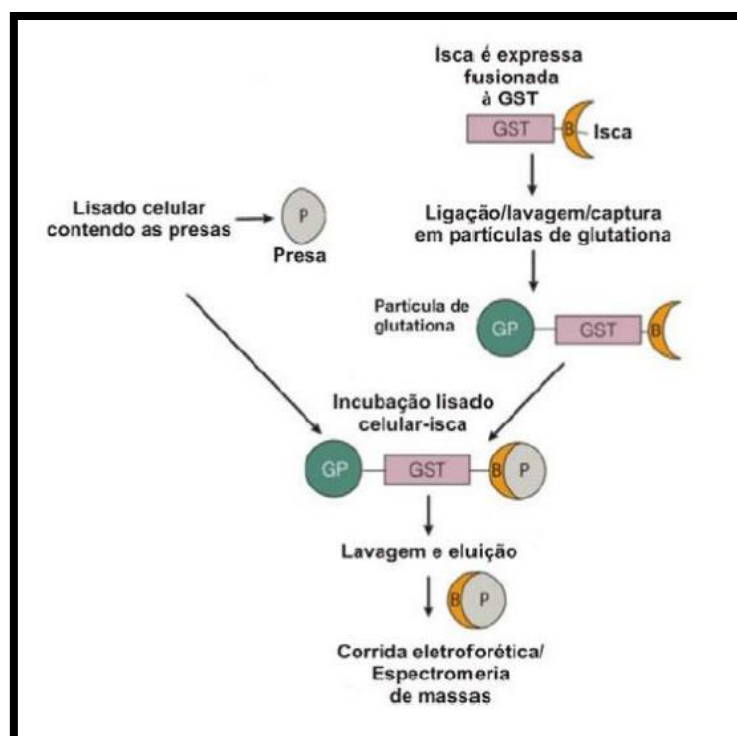


Figura 7. Esquema do ensaio de pull-down *in vitro*. A proteína isca é expressa fusionada a uma proteína adjuvante (GST) no processo de purificação por afinidade. O extrato celular (presa) é incubado com a proteína de fusão purificada. O complexo proteico é submetido à SDS-PAGE e cada proteína é identificada por MS. Fonte: de Oliveira et al., 2013.

A partir das técnicas de *pull down* e duplo-híbrido nas duas fases do fungo, Oliveira et al (2013) identificaram possíveis proteínas que interagem *in vitro* com a *PbMLS*. Tais informações são importantes para descrever os mecanismos e a função destas interações, assim como a descrição dos principais pontos de interação de MLS para uma possível prospecção de moléculas inibidoras de *PbMLS* e suas interações.

1.3. Aplicação da bioinformática na busca de alvos para fármacos

A bioinformática surge como uma ferramenta que tem a finalidade de diminuir os custos com as pesquisas, direcionando com maior rapidez um caminho a seguir. Essa ferramenta é aplicável na pesquisa inicial, de sequenciamento genético de um organismo, bem como, na predição de estrutura de uma molécula proveniente de sua sequência traduzida, e também em fases avançadas, como na aplicação de metodologias para o design de fármacos. Na figura 8, temos uma síntese destas etapas importantes para o desenho racional de fármacos (SILVA, 2012).

Existem dois princípios fundamentais para a validação de um possível alvo terapêutico para descobertas de novos fármacos: 1) o alvo deve possuir propriedades drogáveis e, 2) ao ter a sua atividade modulada por pequenas moléculas, apresentar um resultado terapêutico (FISHMAN e PORTER, 2005; RUSS e LAMPEL, 2005).

Uma maneira de se prever se um alvo (proteína) é drogável, ou seja, responsivo a uma modulação, se baseia na família à qual ele pertence. Se esta família for conhecidamente alvo de drogas, esta provavelmente será, embora não seja regra absoluta (HOPKINS e GROOM, 2002; OWENS, 2007). Outra maneira é baseada na análise da superfície do alvo, ou seja, a partir da identificação de possíveis sítios de interação com ligantes e das características físico-químicas dessa superfície (CHENG et al., 2007). Na Figura 8 temos um esquema representativo das principais fases no desenvolvimento de um fármaco.

Já para o estudo de novos fármacos, a regra de 5 de Lipinski foi e ainda é útil, e parte do princípio de que uma molécula para apresentar uma boa biodisponibilidade via oral deve possuir algumas características físico-químicas específicas (LIPINSKI et al., 1997). Porém, apenas metade dos medicamentos aprovados pela FDA (*Food and Drugs Administration*) se encaixam nos parâmetros da regra dos 5, uma vez que esta não é válida para moléculas derivadas de extratos naturais, e nem moléculas biológicas (anticorpos, e outras macromoléculas), se fazendo necessária a adaptação de alguns parâmetros a fim de atualizar esse “postulado” (ZHANG e WILKINSON, 2007, WALTERS, 2012).

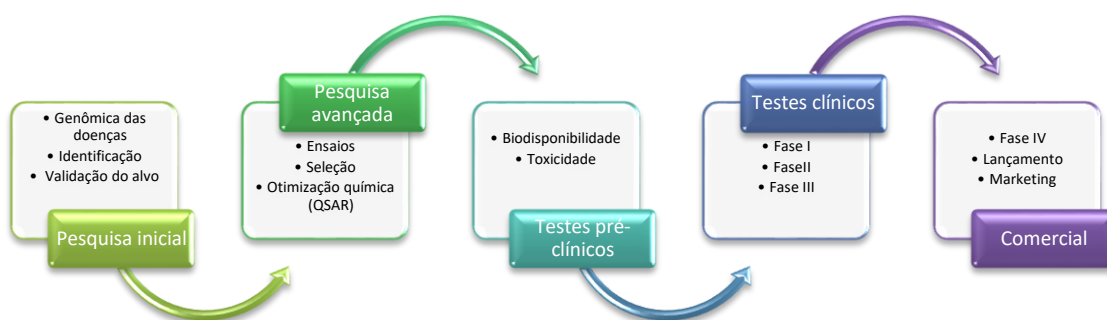


Figura 8. Etapas essenciais no desenvolvimento de um fármaco. Adaptado de (SILVA, 2012).

1.4. Metodologias computacionais

1.4.1. Modelagem molecular de proteínas

Nesta era pós-Genômica um dos maiores desafios é a atribuição de função aos produtos gênicos (proteínas). Como a quantidade de sequências proteicas conhecidas é muito maior que as estruturas de proteínas depositadas em banco de dados, instituições de todo o mundo vem se esforçando para resolver o maior número de estruturas (HEINEMANN, 2000; YANG et al., 2015).

A obtenção de amostras para cristalografia de raios X, ressonância magnética nuclear (RMN) e difração de elétrons é complicada. De acordo com Maggio e Ramnarayan (2001), a cada 20 proteínas, apenas uma produz resultados satisfatórios para estudo com RMN. Assim, as abordagens computacionais de predição de estrutura e função de proteínas vêm ganhando espaço e resultando em estudos bastante relevantes (NORWELL e MACHALEK, 2000).

A metodologia da modelagem molecular *ab initio*, também conhecida como modelagem por primeiros princípios, é baseada na estrutura primária da proteína. Este método utiliza-se das propriedades físico-químicas de cada aminoácido a fim de construir funções de energias que são minimizadas por algoritmos representando as conformações prováveis da proteína (PETSKO & RINGE, 2004). Este método não é totalmente preciso, pois os processos de enovelamento e empacotamento proteicos ainda não são perfeitamente conhecidos. Porém, alguns estudos demonstram bons resultados para pequenas proteínas, de até 100 aminoácidos (SANTOS FILHO; BICCA DE ALENCASTRO, 2003).

A modelagem por homologia, ou modelagem comparativa é considerada a ferramenta mais bem-sucedida, e mais utilizada para predição de estruturas de proteínas (SANTOS FILHO, 2000). Neste caso a estrutura 3D de outra proteína de referência é usada como molde, sendo que esta já possui dados experimentais armazenados em banco de dados (SALI & BLUNDELL, 1993; DUTRA et al., 2014).

Na modelagem por homologia alguns critérios observados no processo de evolução biológica são considerados: 1º) quando há homologia entre sequências de aminoácidos, pode-se inferir que há semelhança estrutural e

funcional; 2º) quando as proteínas são homólogas elas apresentam as regiões internas conservadas; 3º) as maiores diferenças entre proteínas homólogas estão em regiões mais externas, geralmente alças; 4º) proteínas membros de uma mesma família provavelmente podem ser resolvidas por homologia (SANTOS FILHO; BICCA DE ALENCASTRO, 2003).

A metodologia básica da modelagem por homologia se resume em quatro passos principais: identificação e seleção do molde; alinhamento das sequências; construção das coordenadas do modelo e validação (YANG et al., 2015). Pode-se destacar, dentro deste contexto, o servidor I-TASSER que apresenta boa reputação frente às suas participações no CASP7, CASP8, CASP9 e CASP10. O CASP é uma competição dos servidores de predição de estruturas de proteínas do mundo todo, que ocorre a cada 2 anos, e objetiva avaliar a qualidade destes métodos (ZHANG, 2011; YANG et al., 2015).

1.4.2. Dinâmica molecular (DM)

Simulações de dinâmica molecular são aplicadas em estudo de biomoléculas a fim de se estudar propriedades físicas das moléculas – flexibilidade, torções e estabilização - envolvidas em um sistema que simula a realidade biológica. Esse processo é importante para o relaxamento da estrutura tridimensional da molécula e para alcançar uma conformação com melhor resolução. Através da DM é possível também descrever os modos normais de vibração das proteínas e entender quais são as principais conformações envolvidas em torno da estrutura nativa predita. São justamente estas conformações que muitas vezes fornecem informações a cerca da flexibilidade de segmentos da proteína, podendo indicar importantes mecanismos de interação a partir da estabilidade ou instabilidade destas regiões. Todas estas informações são importantes para selecionar as principais conformações e identificar os pontos de interação mais prováveis para a ligação com outras moléculas, sejam elas outras proteínas ou ligantes (PRONK et al., 2013).

Fundamentada nos princípios de mecânica clássica, a simulação de dinâmica molecular fornece informações do comportamento atômico do sistema em função do tempo. Por meio de cálculos estatísticos, algumas propriedades podem ser analisadas, como temperatura, pressão, volume, entropia e energia livre, por exemplo (PRONK et al., 2013; CSERMEY et al., 2013).

Baseada nos estudos Newtonianos, na DM as moléculas são formadas por átomos que se interagem por meio de vários potenciais, os quais, juntos, definem um particular campo de força. É a partir das derivações destes potenciais que as coordenadas atômicas do sistema são atualizadas para cada intervalo de tempo. Os termos de energias envolvidos num campo de força para DM levam em consideração ângulos de ligação, comprimentos de ligação, forças de van der Waals e eletrostáticas. Alguns servidores web oferecem passos de minimização de energia e refinamento automatizados, além dos cálculos de energia livre que fornecem estimativas para uma melhor interação (ZHANG et al., 2011).

1.4.3. *Molecular docking*

Ancoragem molecular é uma metodologia bastante utilizada para prever os melhores complexos de proteína-proteína e proteína-ligante (COLEMAN et al., 2013). A compreensão dessas interações bem como as características da interface de ligação são essências no estudo de interatoma. Muitos complexos proteína-proteína vem sendo caracterizados por cristalografia de raios X e os dados armazenados no PDB (*protein data bank*), porém esse número ainda é muito baixo comparado à importância e frequência desse tipo de interação. Dessa maneira, metodologias de *docking* vêm sendo uma saída para ajudar a elucidar estes complexos (PIERCE, HOURAI e WENG, 2011).

Existem 3 pontos importantes para a realização da ancoragem molecular: 1) escolha da representação do sistema; 2) busca de espaço conformacional; e 3) ranking de potenciais soluções. Sobre o ponto de vista de flexibilidade, os procedimentos de *docking* podem ser classificados como de corpo rígido, semi-flexíveis e flexíveis. No primeiro caso, as duas proteínas são tratadas como dois corpos sólidos e rígidos (modelo simplista), nas metodologias semi-flexíveis, uma das moléculas, geralmente a menor, é considerada flexível. Já em métodos flexíveis, ambas as moléculas são tratadas como flexíveis, com flexibilidade limitada (HALPERIN et al., 2002).

Existem servidores que são capazes de refinar os complexos por um rearranjo de cadeia lateral e pela otimização de corpo rígido. A flexibilidade da cadeia lateral é modelada por rotâmeros, a posição relativa dos complexos é refinada por minimização de *Monte Carlo* e classificados pelo *score* de ligação.

Esta pontuação inclui contatos atômicos, força de *van der Waals* e eletrostática, e energia livre de ligação (MASHIACH et al., 2008).

Um dos métodos mais importantes de ancoragem molecular é o AutoDock, um algoritmo genético que utiliza de um modelo físico aproximado para avaliar as possíveis conformações proteína-ligante. Baseado nos princípios de evolução lamarckiana, o genótipo dos indivíduos é substituído pela melhor conformação obtida a cada passo. Nele, o ligante é flexível o que garante alteração da sua conformação ao longo da simulação, porém a proteína é rígida (TAVARES; MESMOUDI; TALBI, 2009). Chang et al. (2010), realizaram um estudo comparativo entre o Autodock e o Autodock Vina para o rastreio virtual e seleção de compostos inibidores de protease do vírus HIV. Eles observaram que, o AutoDock 4 e o AutoDock Vina foram capazes de classificar os melhores compostos, porém o Vina executa mais rapidamente e classifica de forma mais precisa moléculas maiores (CHANG et al., 2010).

Ancoragem proteína-proteína exige algoritmos de encaixe de corpos rígidos, devido ao tamanho das moléculas e as possibilidades de conformações e graus de liberdade. Apesar desse tipo de *docking* não ser totalmente fiel ao estado nativo das proteínas, ele se faz necessário neste contexto, e tem demonstrados bons resultados (HALPERIN et al., 2002). Um servidor que podemos destacar que se encaixa nestes termos é o GRAMM-X (TOVCHIGRECHKO; VAKSER, 2006).

Um trabalho realizado recentemente, buscou-se resíduos importantes na interação de proteínas que participam de vias de sinalização para o desenvolvimento do câncer de pulmão. Nele, os pesquisadores utilizaram o servidor I-TASSER para a modelagem por homologia e o GRAMM-X para a ancoragem molecular (NAQSH E ZAHRA; KHATTAK; MIR, 2013).

Jayapalan e Natarajan avaliaram as interações de duas enzimas com a proteína associada a microtúbulos (MAPT) que quando hiperfosforilada resulta na doença de Alzheimer. Para isto, utilizaram o servidor GRAMM-X para a etapa de ancoragem molecular (JAYAPALAN; NATARAJAN, 2013).

1.4.4. Peptidologia

A síntese de peptídeos de até 15 resíduos tem emergido como uma alternativa para interações proteicas, uma vez que as combinações de 20

aminoácidos primários garantem uma variedade de moléculas e funções, e a síntese de peptídeos deste tamanho é relativamente fácil e barata. Todavia, peptídeos muito pequenos normalmente não apresentam estabilidade, pois podem se ligar em várias regiões e modificar sua estrutura após a ligação na proteína alvo (alta flexibilidade). Para tentar minimizar esses problemas, alguns autores utilizam de informações de complexos proteína-proteína para selecionar possíveis arcabouços para o local de interação (HONG ENRIQUEZ et al., 2012).

As ferramentas de bioinformática e de modelagem molecular, podem ser muito úteis na projeção de peptídeos candidatos a uma função farmacológica, uma vez que permitem identificar peptídeos com alta afinidade a uma proteína alvo específica (LOPEZ-MARTINEZ et al., 2013). Para a predição de peptídeos *in silico*, são necessários dados da interação proteína-proteína, de ontologia gênica (GO), cálculos físico-químicos, dados de interação, e informações de locais funcionais e resíduos “chave” (*hotspots*) (SHAMMEER et al., 2010).

A escolha de uma sequência peptídica específica com afinidade a uma proteína alvo é uma tarefa complexa, partindo-se da ideia de que temos milhares de combinações com os 20 aminoácidos primários existentes. Dessa maneira, uma metodologia seletiva, rápida e eficiente se faz essencial. Pode-se destacar 3 passos importantes no sucesso de uma interação peptídeo-proteína: 1) analisar a região da superfície da proteína, onde o peptídeo poderá se ligar; 2) confirmar essa interação; e 3) otimizar essa interação de forma a deixá-la mais estável e energeticamente viável (UNAL, GURSOY e ERMAN, 2009).

Nos últimos anos vários métodos computacionais foram desenvolvidos para prever a interface e prováveis *hotspots*, alguns baseados em simulações de DM, enquanto outros são baseados em funções de energia. O servidor KFC2 (ZHU; MITCHELL, 2011) identifica os aminoácidos envolvidos na interação, bem como os resíduos “chaves”.

KFC2 é um método baseado em abordagem de aprendizado de máquinas (SVM), tais como KFC e Minerva, baseados em características geométricas e bioquímicas. Além disso, o ambiente estrutural no qual está localizado determinado resíduo é considerado, comparando-o com *hotspots* experimentalmente já comprovados (ZHU e MITCHELL, 2011).

A nova metodologia de KFC2 apresenta uma precisão de predição e velocidade de cálculos superiores, principalmente comparando-se com outros

métodos. Características relacionadas à área de superfície ao solvente e propriedades dos resíduos vizinhos, foram também adicionadas à nova metodologia KFC2 (CHO; KIM; LEE, 2009; ZHU e MITCHELL, 2011; FALLIS, 2013). Thakur et al. (2014) utilizaram o servidor KFC2 para identificar os resíduos chaves da interação de duas proteínas responsáveis pela manutenção da homeostase de ferro; ferroportina e hepcidina.

Na Tabela 1, estão alguns resultados de uma pesquisa feita dentre os preditores de *hotspots*. Observa-se na Tabela 1.A, que KFC2a apresentou o maior número de *hotspots* verdadeiros positivos (51) e menor quantidade de falsos negativos comparando-se com os outros programas. Já para KFC2b, temos menor quantidade de falsos positivos comparando com KFC2a. Na Tabela 1.B, observa-se que KFC2a apresentou a maior sensibilidade e acurácia dentre os métodos, e KFC2b apresentou uma boa especificidade e acurácia.

A

		Predicted ^b	
	Actual ^a	Hot spot	Other
KFC2a	Hot spot	51 (TP)	14 (FN)
	Other	15 (FP)	52 (TN)
KFC2b	Hot spot	36 (TP)	29 (FN)
	Other	9 (FP)	58 (TN)
Robetta	Hot spot	33 (TP)	32 (FN)
	Other	8 (FP)	59 (TN)
FOLDEF	Hot spot	20 (TP)	45 (FN)
	Other	5 (FP)	62 (TN)
KFC	Hot spot	36 (TP)	29 (FN)
	Other	8 (FP)	59 (TN)
MINERVA2	Hot spot	38 (TP)	27 (FN)
	Other	3 (FP)	64 (TN)
HotPoint	Hot spot	34 (TP)	31 (FN)
	Other	9 (FP)	58 (TN)

B

	Sensitivity	Specificity	Accuracy
KFC2a	0.78	0.78	0.78
KFC2b	0.55	0.87	0.71
MINERVA2 ^a	0.58	0.96	0.77
HotPoint	0.52	0.87	0.70
KFC ^a	0.55	0.88	0.72
Robetta	0.51	0.88	0.70
FOLDEF	0.31	0.93	0.62

Tabela 1. Comparativo dos preditores de *hotspots*. Na tabela A, estão listados os servidores analisados para predição de interface e a quantidade de *hotspots* preditos por cada servidor. Na tabela B, temos os valores de sensibilidade, especificidade e acurácia destes métodos. TP: verdadeiros positivos, FP: falsos positivos, FN: falsos negativos e TN: verdadeiros negativos. FONTE: ZHU e MITCHELL, 2011.

Tais abordagens sobre as informações de PPI são essenciais na busca de peptídeos que possam inibir tais interações. Para isto, a predição de peptídeos pode ser baseada em estrutura ou por metodologias *ab initio/de novo*. O primeiro, considera a estrutura tridimensional da sequência peptídica de escolha a partir de uma região da “proteína molde”. Já o segundo, baseia-se apenas na sequência peptídica, utilizando-se de uma abordagem sem uma estrutura tridimensional inicial (SHAMMEER et al., 2010). O software PEP-FOLD utiliza uma abordagem *de novo* para prever estruturas peptídicas 3D, a partir da

sequência de aminoácidos. Sua metodologia é baseada no alfabeto estrutural e utiliza um modelo Markoviano (MAUPETIT, DERREUMAUX e TUFFERY, 2009).

A aplicação de peptídeos na terapia de doenças nos fornece várias vantagens; alta bioatividade, especificidade e afinidade de ligação ao seu alvo, e, além disso, toxicidade relativamente baixa. Apesar da investigação de peptídeos, ocorra desde os anos 40, apenas nas últimas décadas estes realmente foram empregados na terapia, como mostrado na Figura 9. Hoje, podemos encontrar vários peptídeos empregados como fármacos antiinflamatórios, antitumorais, antimicrobianos, sinalizadores do SNC e imunomoduladores (REICHERT, 2010; CRAIK et al., 2013).

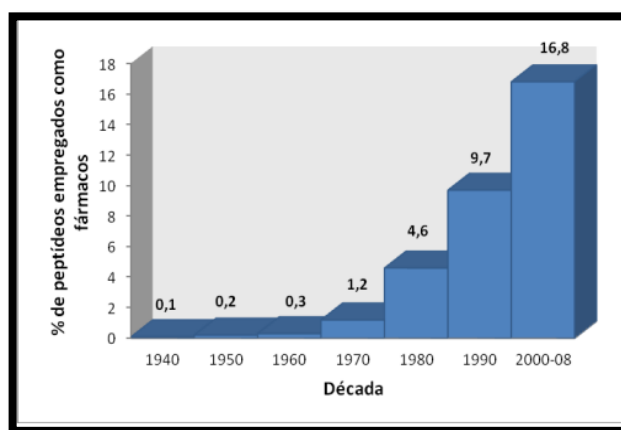


Figura 9. Emprego de peptídeos na terapêutica. Dados do crescimento do emprego de peptídeos como fármacos a partir da década de 80 (FONTE: REICHERT, 2010).

No anexo 1, temos uma lista de alguns peptídeos utilizados atualmente pela indústria farmacêutica na terapêutica. Finalmente, os estudos de peptídeos como moléculas promissoras atuando como fármacos, despertam atenção de muitos grupos de pesquisadores devido a suas vantagens, já apresentadas. Dessa forma, pretendemos com este estudo apresentar subsídios interessantes para a síntese de peptídeos e possíveis testes *in vitro*.

2. JUSTIFICATIVA

A PCM é uma doença negligenciada, de alta letalidade (podendo chegar a 30%) principalmente em segmentos sociais específicos, como os trabalhadores rurais, representando um grave problema de saúde pública. No Brasil, é responsável por 80% dos novos casos de micoses sistêmicas, sendo o país com maior incidência do mundo. Desta maneira, estudos relacionados à biologia do fungo *Paracoccidioides* spp., bem como seus mecanismos de adaptação são essenciais para o entendimento melhor da doença.

Neste contexto, a enzima *PbMLS*, foco principal desta pesquisa, possui várias funções tanto no metabolismo energético quanto no processo de adesão do fungo à célula hospedeira. A *PbMLS* pode ser considerada um fator de virulência do *Paracoccidioides* spp., por favorecer a instalação da doença e pelo fato de ter sua expressão aumentada em condições de escassez de carboidrato *in vitro*. Estudar essa proteína, portanto, é interessante para descobrir possíveis novas funções e localizações, visto que esta pode ser um possível alvo de ligantes inibidores.

O desenho de peptídeos tem sido cada vez mais explorado por se tratar de moléculas relativamente mais específicas e menos tóxicas, comparando-se com ligantes não peptídicos. Buscas *in silico* de prováveis alvos de fármacos tornou-se uma etapa importante para otimização de tempo, refinamento de dados e direcionamento de pesquisas. Com isso, o estudo das interações de *PbMLS* com outras proteínas de *Paracoccidioides* spp. e identificação dos resíduos chave dessas interações podem nos direcionar ao desenho de pequenas moléculas peptídicas com afinidade à *PbMLS*.

3. OBJETIVOS

3.1. Objetivo geral

Identificar os principais pontos de interação da proteína *PbMLS* de *Paracoccidioides* spp. e propor peptídeos capazes de modular a atuação desta enzima.

3.2. Objetivos específicos

- Determinar as estruturas tridimensionais das proteínas que interagem com *PbMLS* e analisar a rede de interações de cada uma delas;
- A partir dos experimentos de ancoragem molecular (*docking*) identificar os principais resíduos envolvidos e suas contribuições para a interação;
- Propor peptídeos que poderão ter uma afinidade pela *PbMLS* em experimentos de DM ,
- Apontar os peptídeos mais promissores à interação com MLS para posterior síntese e testes *in vitro* e *in vitro*.

4. PROCEDIMENTOS EXPERIMENTAIS

4.1. Obtenção das estruturas 3D

Do total de 45 proteínas interagentes com a MLS, *in vitro*, apontadas por de Oliveira et al. (2013), foram selecionadas 6 para estudo *in silico* de interação, são elas: Enolase, Gama actina, Gliceraldeído - 3 fosfato desidrogenase (GPD), Malato desidrogenase, Triosefosfato isomerase e Ubiquitina. Essas proteínas foram selecionadas por diversas justificativas; 1ª: No estudo de Oliveira et al. (2013) as proteínas enolase, triosefosfato isomerase e gama actina interagiram com PbMLS por *Far Western- blot*. 2ª: A proteína GPD é possivelmente regulador alostérico da via glicolítica, o que a torna interessante alvo de estudo com interações proteína-proteína. 3ª: A interação entre ubiquitina e PbMLS se tem registrada no software de interações OSPREY (<https://osprey.thebiogrid.org/>). 4ª: Já a proteína malato desidrogenase apresenta interação com a PbMLS de acordo com o servidor STRING (OLIVEIRA et al., 2013).

Como as estruturas 3D das proteínas ainda não foram resolvidas experimentalmente, por cristalografia ou Ressonância Magnética Nuclear (RMN), propomos modelos iniciais baseados em estruturas já conhecidas (homólogas) a partir do PDB (*Protein DataBank*). Todos esses modelos iniciais foram preditos pelo servidor I-TASSER, que baseia-se na sequência de aminoácidos da proteína, realizando uma predição da estrutura 3D e suas possíveis funções (YANG et al., 2015). Para a análise da qualidade dos modelos gerados, utilizamos o servidor MolProbity do qual extraímos os gráficos de Ramachandran, o *score* geral da molécula (*score molProbity*) e dos átomos sobrepostos (*clashscore*), sendo que quanto mais próximo de 0 esses *scores*, melhor é a estrutura predita. Com este servidor foi possível comparar a estrutura da proteína antes e depois da DM (CHEN et al., 2010).

O servidor STRING foi utilizado para prever as interações já conhecidas ou possíveis das proteínas desse estudo, sejam elas associações físicas (diretas) ou funcionais (indiretas). Nele, os dados são provenientes de diversas fontes: conhecimentos experimentais extraídos de bancos de dados, estatísticas de interações proteína-proteína descritas na literatura, predições *de novo* por

algoritmos que utilizam de informações genômicas (co-expressão) e relações de ontologia gênica entre elas. Além disso, é possível adicionar várias proteínas como entrada e visualizar a rede de interações disponíveis para todas elas (SZKLARCZYK et al, 2014).

O servidor *2struct* foi utilizado para a análise do perfil de estrutura secundária das moléculas, a qual realiza a partir de comparações com várias informações de estruturas secundárias conhecidas (KLOSE & JANES, 2010).

4.2. Simulações de Dinâmica Molecular

Cada proteína foi submetida a simulação com o software GROMACS 4.5.5, usando o campo de força AMBER (ff99SB-ILDN), a fim de conseguir uma estabilidade estrutural do modelo, na presença de solvente (BERENDSEN et al, 1995; LINDAHL et al, 2001; PRONK et al, 2013).

A primeira etapa tem a função de remover os contatos conflitantes, a qual corresponde a uma rápida minimização de energia, sendo considerada completa apenas quando a tolerância de 1.000 kJ/mol não foi excedida. Para equilíbrio das variáveis termodinâmicas do sistema, este foi sujeito a uma simulação de 100 ps de NVT e NPT, tendo a proteína restringida em suas posições. Na fase NVT, onde ocorre variação na pressão do sistema, a temperatura foi ajustada pelo termostato em 300 K (26,85°C), e as velocidades calculadas a partir de equações de Maxwell. Já na simulação NPT, onde o volume é permitido variar, a pressão foi mantida pelo barostato Parrinello-Rahman (HUTTER, 2012; IANNUZZI et al, 2003).

Após estes passos as proteínas foram submetidas à simulação de 100 ns, temperatura de 300 K, pressão de 1 atm e intervalo de tempo de 2 fs, sem restrição da conformação. Apenas para a *PbMLS*, nossa proteína alvo do estudo, foi utilizado um tempo de simulação de 200 ns para garantir maior estabilidade (PRONK et al, 2013).

A trajetória foi analisada por meio da Raiz do Desvio Quadrático Médio (RMSD) em relação à estrutura inicial, pelo algoritmo *gromos*, como descrito por Daura et al. (2009). O programa *g_cluster* (pacote GROMACS) foi utilizado para determinar as conformações que mais estiveram presentes ao longo da simulação, com *cut-off* de 0,3 nm para distinguir os conjuntos conformacionais com base no perfil RMSD. Com a análise de cluster e RMSD, observamos o

perfil de cada proteína ao longo da simulação, e assim, foram selecionados de 1 a 3 modos para cada uma delas para os procedimentos de *docking* molecular.

4.3. Ancoragem molecular

O servidor de ancoragem proteína-proteína GRAMM-X (Tovchigrechek e Vakser, 2006) foi utilizado para a determinar as melhores conformações entre os complexos *PbMLS*-proteína ligante. Na etapa seguinte, um refinamento da interação e, conseqüentemente de energia, foi realizado pelo servidor FireDock (Mashiach et al., 2008).

Com os complexos formados, a etapa seguinte foi identificar os aminoácidos envolvidos na interação. O servidor KFC2 (Zhu e Mitchell, 2011) reconhece todos os resíduos da interface de interação, tanto da cadeia A (*PbMLS*) quanto da cadeia B (proteínas ligantes).

Uma vez conhecendo os resíduos da *PbMLS* mais frequentemente envolvidos nas interações, foi utilizado o servidor CoCoMAPS (Vangone et al., 2011). Este fornece quais os resíduos das proteínas ligantes que interagem com as regiões de escolha da *MLS*, bem como classifica-os como *hotspots* ou não. Desta maneira, com as informações de resíduos interagentes, e com o auxílio de programas de visualização molecular (VMD, Pymol e Chimera), pudemos propor alguns peptídeos promissores, que pudessem apresentar complementaridade físico-química às regiões selecionadas como as mais frequentes envolvidas nas interações entre a *PbMLS* e as outras proteínas.

4.4. Desenho e ancoragem dos peptídeos

O desenho de peptídeos foi realizado a partir das análises das regiões de interesse da *MLS*, selecionadas a partir da alta frequência em interações com outras proteínas. Dessa forma, foram selecionados peptídeos a partir do arcabouço tridimensional das proteínas ligantes. Os arquivos *pdb* (arquivos de texto com as coordenadas atômicas) dos complexos proteína-proteína foram recortados a fim de se ter apenas a proteína receptora (*PbMLS*) e a sequência delimitada ao peptídeo.

Tais complexos, foram submetidos ao servidor FlexPepDock (LONDON et al., 2011), para o refinamento e modelagem dos mesmos (proteína-peptídeo).

Neste servidor é realizado um refinamento da interação em torno da região inicial, isto é, a partir do peptídeo no local de ligação, permitindo flexibilidade ao ligante tanto das cadeias laterais como da cadeia principal.

Paralelamente, realizamos a predição da estrutura tridimensional dos peptídeos em solvente (modelagem *de novo*), a partir da sequência do peptídeo selecionado na etapa anterior. Para isto, utilizamos o PepFOLD (MAUPETIT, DERREUMAUX, TUFFÉRY, 2009).

Os peptídeos que foram refinados no FlexPepDock, foram alinhados ao melhor peptídeo (modo conformacional) *rankeado* pelo PepFOLD e tiveram o valor de RMSD calculados. Desta maneira, o alinhamento foi feito para cada peptídeo proposto inicialmente, com a finalidade de escolher o de menor RMSD, levando em consideração o menor valor de *score* de energia (mais próximo ao modelado no solvente)

Os peptídeos de menor *score* de energia foram selecionados para análises de interações hidrofóbicas, eletrostáticas e interações de hidrogênio. As interações desfavoráveis também foram avaliadas, e aqueles peptídeos com muitas destas foram excluídos das etapas seguintes. Tais análises foram feitas através do programa Discovery Studios.

5. RESULTADOS E DISCUSSÃO

5.1. Modelos estruturais das proteínas

5.1.1. Malato sintase (*PbMLS*)

O principal modelo de *PbMLS* definido pelo servidor I-TASSER, foi baseado na estrutura da Malato sintase de *Escherichia coli*, resolvida por difração de raio X e depositado no PDB (ID:3CV1A). Esta estrutura apresentou 41% de identidade e 98% de cobertura com *PbMLS*, de acordo com alinhamento do servidor BLAST. Através da análise no servidor *2Struct*, observamos que mais de 50% do modelo da MLS corresponde a alças.

Estas análises foram realizadas para todas as proteínas, como mostrado na Tabela 2. Nesta tabela temos informações da estrutura inicial, o percentual de estrutura secundária, scores de qualidade da estrutura tanto antes da DM quanto depois, e a quantidade de modos conformacionais selecionados para o *molecular docking*. Informações gerais, de cada uma das proteínas do estudo, bem como suas funções estão inseridos no Apêndice 1.

5.2. Rede de interações de proteínas

Apresentamos nesta subseção os resultados obtidos pelo servidor STRING-DB, no qual foram selecionadas interações com alta confiança (score de 0,9) e interações baseadas em banco de dados e experimentos. Temos as redes de interações de cada uma das proteínas separadamente (Figuras 10 a 15) e ao final o mapa de interação provenientes dessa metodologia (Figura 16), para todas as proteínas juntas. Dessa forma, buscamos padrões que possam caracterizar a rede de proteínas da *PbMLS*.

Na Figura 10, observam-se interações proteína-proteína que estão depositadas em banco de dados e oriundas de dados experimentais. Destaca-se a MLS (círculo vermelho central) e as proteínas ICL e malato desidrogenase, que pertencem ao ciclo do glioxilato.

Na Figura 11, temos a Malato desidrogenase no centro (círculo vermelho), as linhas cinza representam possíveis interações provenientes de bancos de dados. Já as linhas azuis indicam interações confirmadas por experimentos.

Na Figura 12, a Proteína enolase (círculo vermelho) realiza várias interações descritas em banco de dados, mas apenas as interações com a FBA1 e PAAG 06104 são comprovadas experimentalmente.

Na Figura 13, temos a rede da triosefosfato isomerase observam-se duas proteínas ilustradas com círculos maiores, que indicam a presença de estruturas 3D de proteínas homólogas a estas (alta identidade). As linhas cinzas e azuis indicam o mesmo que as figuras anteriores. Na rede de interações da ubiquitina (Figura 14), observa-se proteínas de vários grupos funcionais, e portanto não apenas interações diretas (linhas em azul). As linhas rosas indicam interação com modificação pós traducional, as pretas indicam reações químicas e as linhas roxas indicam catálise. As pontas arredondadas nas linhas indicam que o efeito de uma proteína sobre a outra ainda não é conhecido.

Na Figura 15, temos a rede de interações da GPD, onde destacam-se algumas interações comprovadas experimentalmente (linhas azuis) que inclui a TPI1, mesmo que indiretamente. Por fim, na Figura 16, temos a rede das interações preditas pelo servidor STRING para as proteínas envolvidas no estudo (exceto a gama actina que corresponde à sequência de *Mus musculus*). As linhas em cinza representam as interações possíveis de acordo com análise de banco de dados, já as linhas em azul indicam interações com comprovação experimental. A ubiquitina apresenta a maior rede de interação, e juntamente com a GPD apresentam estruturas altamente homólogas já cristalizadas. A MLS não apresentou interações significativas por esta metodologia

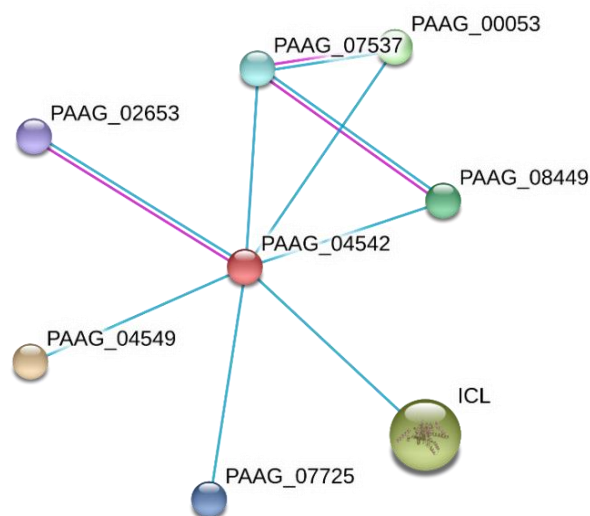


Figura 10. Rede de interações de *PbMLS*. As linhas azuis indicam interações oriundas de banco de dados, e as linhas rosas indicam interações confirmadas por experimentos. Legenda: PAAG_04542-*PbMLS*; 02653: CoA acetato ligase; 04549: metil-isocitrato liase; 07725: hidroxioxidase peroxissomal; 07537: enzima málica; 00053: malato desidrogenase (domínio beta); 08449: malato desidrogenase (domínio Ldh); ICL: isocitrato liase.

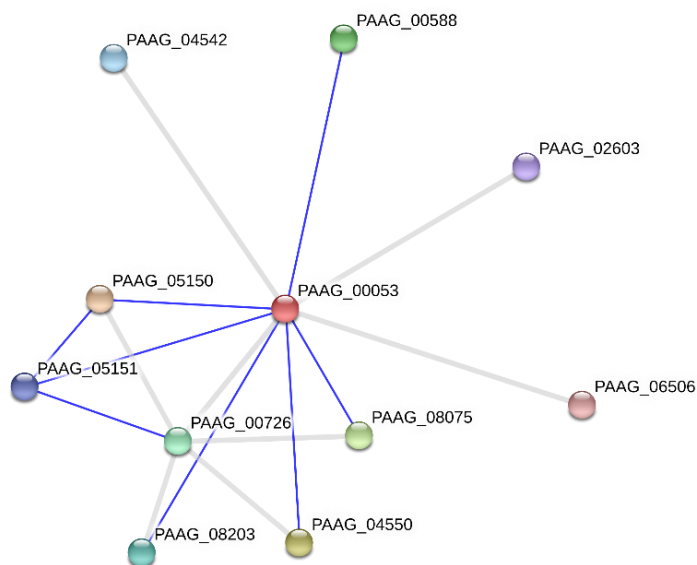


Figura 11. Rede de interações de malato desidrogenase. Legenda: PAAG_04550: 2-metilcitrato sintase; 08075: citrato sintase; 00588: fumarato hidratase; 00726: piruvato carboxylase ; 08203: phosphoenolpyruvate carboxykinase; 04542: malate synthase ; 05151: ATP-citrate-lyase; 02603: aspartate aminotransferase ; 06506:aspartate aminotransferase .

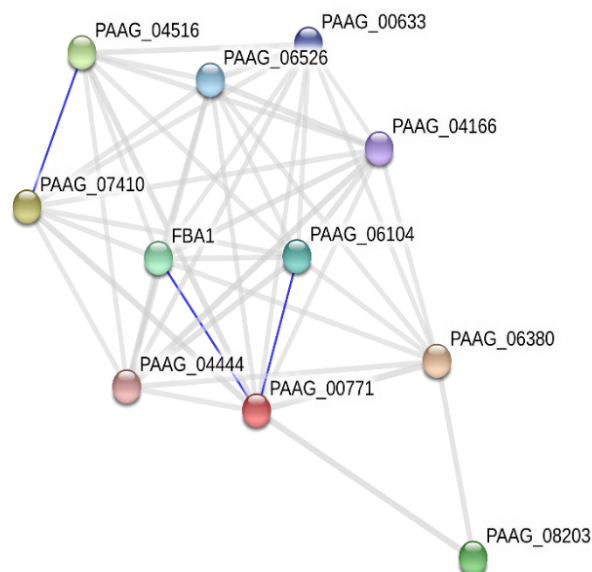


Figura 12. Rede de interações de Enolase. Legenda: PAAG_07410: 2,3-bisphosphoglycerate-independent phosphoglycerate mutase; 04516: phosphoglycerate mutase family protein ; 08203: phosphoenolpyruvate carboxykinase ; FBA1: fructose-bisphosphate aldolase ; 06104: fructose-bisphosphate aldolase ; 06526: glucose-6-phosphate isomerase; 00633: glucose-6-phosphate 1-dehydrogenase; 04166: transaldolase; 04444: transketolase.

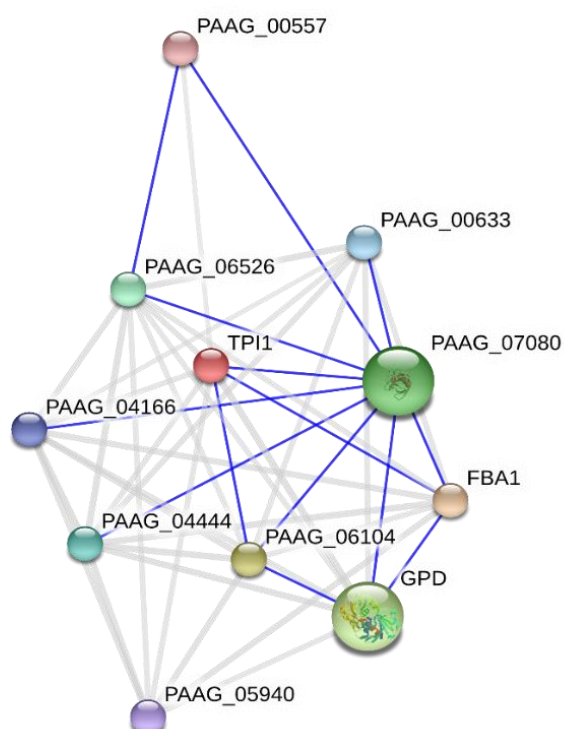


Figura 13. Rede de interações de Triosefosfato isomerase. Legenda: FBA1: fructose-bisphosphate aldolase; PAAG_06104 fructose-bisphosphate aldolase; GPD: glyceraldehyde-3-phosphate dehydrogenase; 07080 ubiquitin; 06526: glucose-6-phosphate isomerase; 04444 transketolase ; 00633 glucose-6-phosphate 1-dehydrogenase; 04166 transaldolase ; 05940 deoxyribose-phosphate aldolase; 00557 mannose-6-phosphate isomerase.

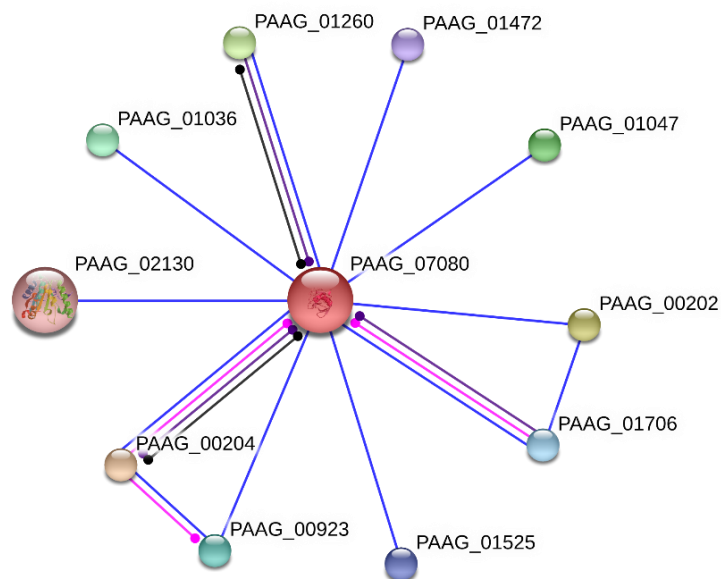


Figura 14. Rede de interações de Ubiquitina. Legenda: PAAG_00204: ubiquitin-conjugating enzyme ; 00202: ubiquitin carboxyl-terminal hydrolase ; 01260: vacuolar protein sorting-associated protein ; 01047: DUF814 domain-containing protein ; 01036:negative regulator of cdc42 ; 00923: proliferating cell nuclear antigen ; 01706: 26S proteasome regulatory subunit RPN10 ; 01525: fatty acid synthase subunit alpha reductase ; 01472: ubiquitin-conjugating enzyme ; 02130: heat shock protein HSP98.

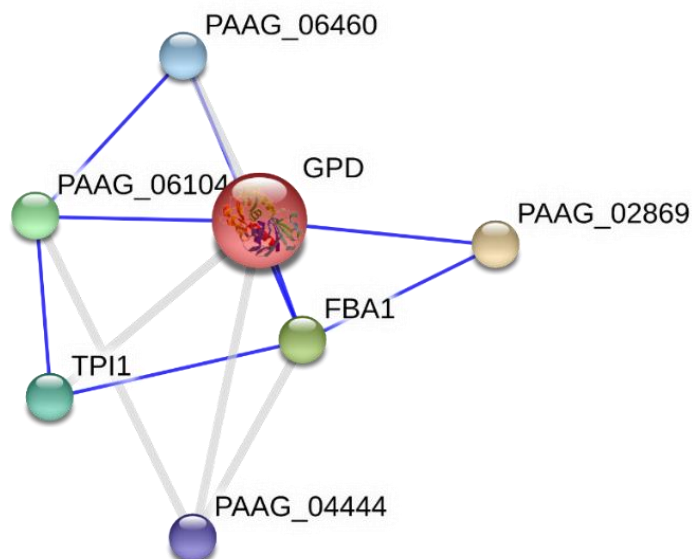


Figura 15. Rede de interações de GPD. Legenda: 02869: phosphoglycerate kinase ; FBA1: fructose-bisphosphate aldolase ; 06104: fructose-bisphosphate aldolase; TPI1: triosephosphate isomerase; 06460 triosephosphate isomerase; 04444: transketolase.

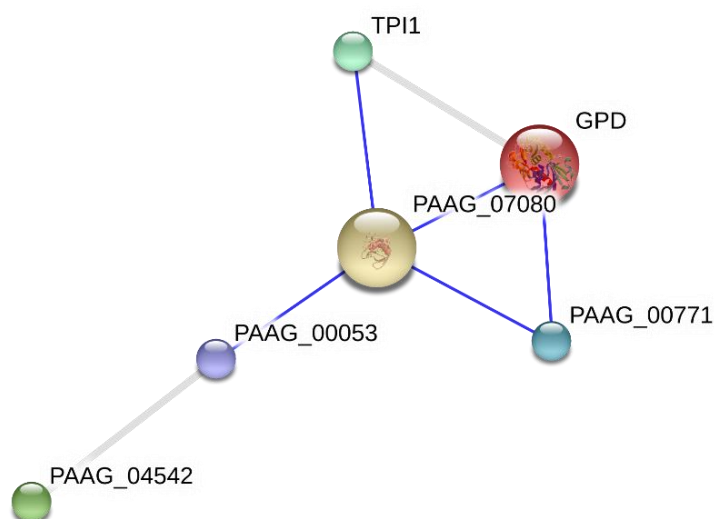


Figura 16. Rede de interações das proteínas envolvidas no estudo. Legenda: GPD: glyceraldehyde-3-phosphate dehydrogenase; 07080: ubiquitin; 04542: malate synthase; TPI1 : triosephosphate isomerase; 00771: enolase; 00053: malate dehydrogenase.

5.3. Dinâmica Molecular e Qualidade dos modelos

5.3.1. Análise dos modos da MLS

Recentemente, Costa et al. (2015) realizou a modelagem de *PbMLS* numa simulação de DM de 200 ns. A partir das trajetórias destas simulações, foi realizadas análises de agrupamento, onde observaram-se 14 *clusters* (grupos) conformacionais, porém os que se mantiveram por maior tempo foram os *clusters* 1 e 2 (Figura 17).

O *cluster* 1 ocorre justamente a partir do momento de maior estabilidade estrutural, aproximadamente 80 ns, como mostrado no gráfico de RMSD (Figura 18). O *cluster* 2 foi o segundo que permaneceu por mais tempo durante a simulação. Porém, foram selecionados os *clusters* 1, 2 e 6 para o estudo de interações com as outras proteínas, uma vez que o *cluster* 6 surgiu nos momentos finais da DM, que poderia representar um modelo conformacional estável alcançado. Os 3 modelos selecionados e o modelo inicial (anterior à DM) estão alinhados na Figura 20.

Tabela 2. Aspectos gerais das proteínas envolvidas no estudo

Proteínas		<i>PbMLS</i>	Enolase	Gama-actina	TPI	Ubiquitina	GAPDH	Malato desidrogenase
PDB template -I-Tasser		3CV1A	1P43	3U4L_A	1YPI_A	4KSL_A	1U8F_O	2DFD_A
Identidade		41%	61%	99%	100%	91%	71%	56%
Cobertura		98%	98%	100%	99%	99%	97%	94%
Nº resíduos		539	404	375	249	156	388	340
Estrutura secundária **	% hélice	38,4	36,4	39,5	34,1	13,6	26,6	39,1
	% fita beta	10,8	17,6	18,1	12,2	15,6	15,1	8,2
	% outros	50,8	46	42,4	53,7	70,8	58,3	52,6
Antes da DM *	MolProbity score	3,74 (5%)	2,48 (48%)	2,13(68%)	2,32 (58%)	3,14 (18%)	2,78 (32%)	3,12 (19%)
	Clashscore	149 (0%)	4,81 (94%)	2,4 (99%)	3,1 (98%)	9,45 (74%)	6,44 (89%)	12,21 (61%)
	Ramachandran	90%	96%	96,50%	96,90%	90%	90%	88,50%
Após a DM *	MolProbity score	1,86 (86%)	1,50 (95%)	1,12 (99%)	1,36 (98%)	1,17 (99%)	1,62 (92%)	1,27 (99%)
	Clashscore	0 (100%)	0 (100%)	0(100%)	0 (100%)	0 (100%)	0(100%)	0 (100%)
	Ramachandran	94%	99,30%	98,90%	98,20%	98,70%	97,30%	95%
Clusters selecionados		3 (14)	1 (1)	1 (1)	3 (18)	3 (19)	2 (8)	2 (7)

*Análises do servidor MolProbity, relacionadas à qualidade do modelo gerado antes e depois das simulações de Dinâmica Molecular. **Análises realizadas pelo servidor 2Struct, relacionadas à estrutura secundária.

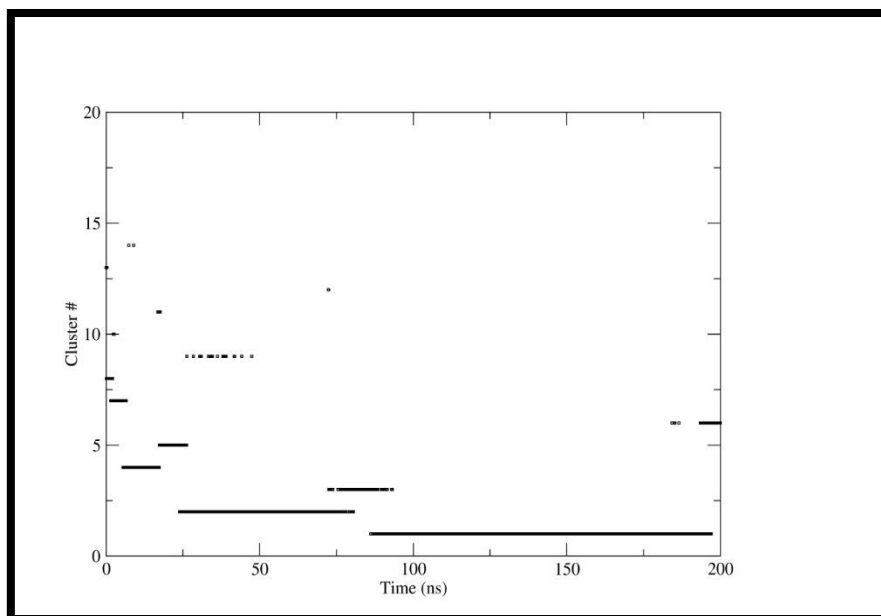


Figura 17. Gráfico de clusters da MLS. Observam-se 14 conjuntos de conformações ao longo da trajetória de DM. Cut-off: 0,3 angstrom. Fonte: Costa et al., 2015.

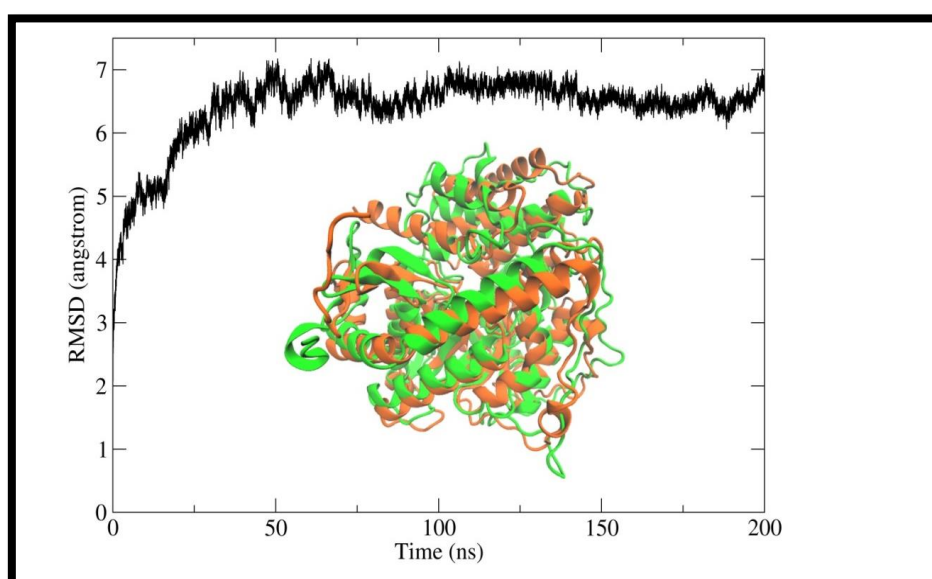


Figura 18. Gráfico de RMSD e estrutura tridimensional da MLS. Evolução do RMSD (raiz do desvio quadrático médio) ao longo de 200 ns de simulação de Dinâmica Molecular. As estruturas sobrepostas representam o modelo gerado por homologia (laranja) e o modelo final (*model 1*) da DM (verde). Fonte: Costa et al., 2015.

No Molprobability score da estrutura total observamos uma melhora de 3,74 para 1,86, e o *clashscore* que antes era 149 chegou a 0, após a DM, demonstrando a redução acentuada do score que indica a quantidade de átomos estruturalmente sobrepostos no modelo principal da Malato sintase. Na Figura

19, temos o mapa de Ramachandran do modelo gerado pelo I-TASSER, em A, 94% dos resíduos encontram-se em regiões permitidas, com relação aos ângulos *phi* e *psi*. Já na Figura 19.B, há uma notória diminuição da quantidade de resíduos em regiões proibidas (de 6 % para 2%). O comportamento das demais proteínas durante a DM, e as análises de *cluster* e RMSD, estão inseridas no Apêndice 1.

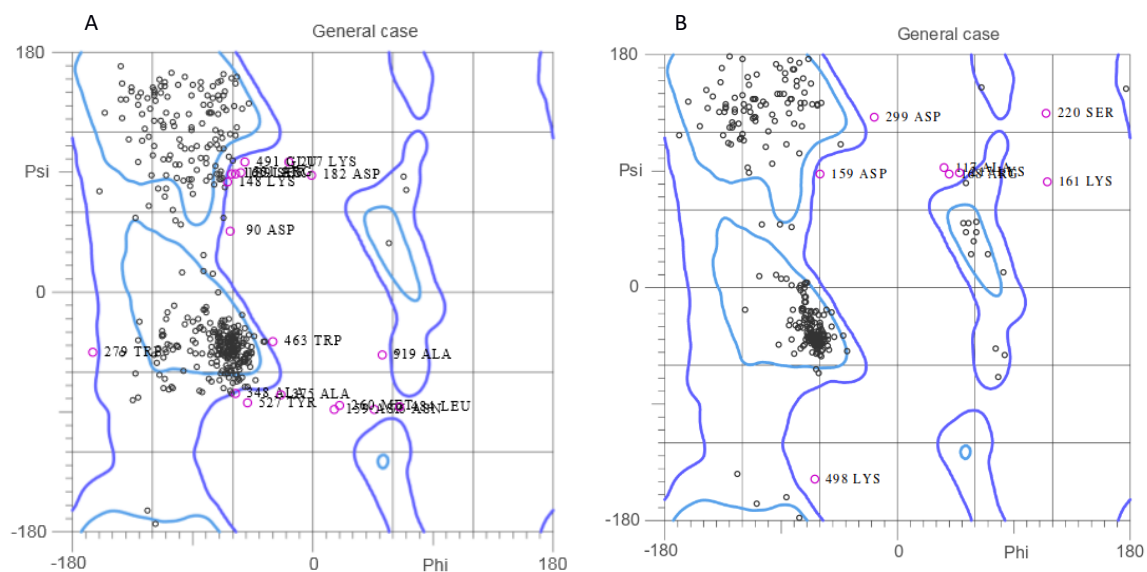


Figura 19. Diagrama de Ramachandran da Malato sintase antes e depois da DM. Em A temos o diagrama da MLS antes da DM (modelo gerado pelo I-TASSER), com 94% dos resíduos da proteína em regiões favoráveis (dentro dos espaços delimitados pela linha azul escuro). Em B, temos 97% dos resíduos em regiões permitidas, sendo possível notar um menor número de resíduos fora das regiões delimitadas pela linha azul escuro.

De uma maneira geral, observamos uma melhora significativa do modelo inicial, com a utilização da DM, como pode ser observado na Tabela 2 .Uma vez que no Mapa de Ramachandran (Figura 19) os resíduos que possuíam átomos conflituosos, ou seja, átomos em posições angulares não favoráveis à estrutura da molécula, diminuiriam nitidamente. Os 31 resíduos que estavam em uma angulação não permitida , após a DM eram apenas 8. Além disso, todos os *scores* de avaliação da qualidade chegaram a valores de 86.º e 100.º de percentil, o que corresponde a valores de proteínas de alta resolução.

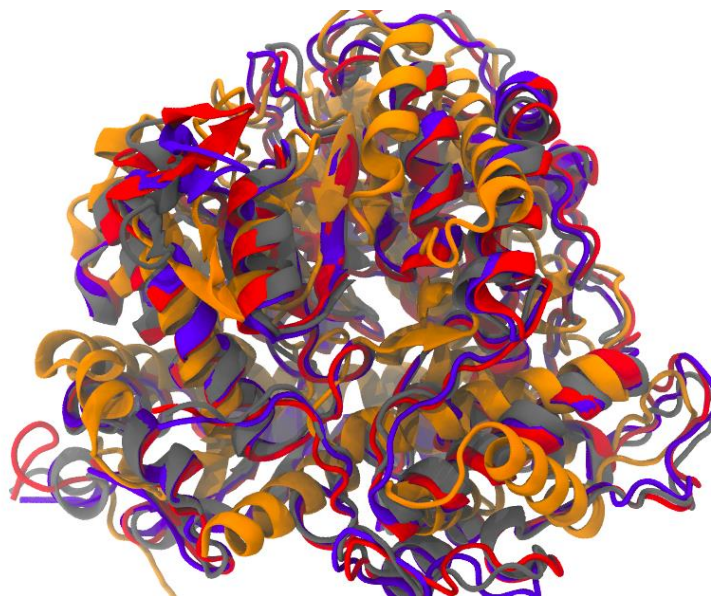


Figura 20. Alinhamento dos modelos da MLS. Em azul o modelo predito pelo servido I-Tasser (antes da DM), em vermelho temos o modo 1 selecionado após a DM (cluster mais frequente durante a simulação de DM. Em cinza, temos o modo 2, e em laranja o modo 3 após a DM.

5.3.2. Características gerais das proteínas ligantes à *PbMLS*

Observa-se que todas as proteínas ligantes à *PbMLS* apresentaram alta cobertura (acima de 94%) com relação ao modelo homólogo principal utilizado pelo I-TASSER. Isso significa que, o modelo estrutural usado como molde para modelagem da proteína de interesse, cobre 94% da sequência total. A identidade corresponde a porcentagem de resíduos de aminoácidos iguais em posições iguais, entre a sequência de interesse, e a proteína estruturalmente resolvida, utilizada como molde principal. Vale destacar que, as proteínas Gama-actina e TPI (Triosefosfato Isomerase) apresentaram identidade e cobertura de quase 100%. que corresponde a praticamente a mesma sequência de aminoácidos.

Com relação à estrutura secundária, a proteína Ubiquitina apresenta mais de 70% de alças em sua estrutura. Sabe-se que esse tipo de região é altamente flexível, comparado a regiões de alfa-hélice e fita-beta. A Ubiquitina é uma molécula muito flexível, o que provavelmente dificultou a mesma atingir um estado de maior estabilidade, confirmado pelos números de *clusters* (grupos) obtidos durante a simulação de DM. Essa relação entre a porcentagem de alças da molécula e o número de clusters durante a simulação, também pode ser observada para as proteínas *PbMLS*, TPI, GPDH e MD.

Com relação a qualidade dos modelos, todos tiveram um *score* melhorado após a aplicação da DM. Esse fato está relacionado à sobreposição de átomos, Ramachandran e ângulos da molécula em posições favoráveis. Na tabela 2, observa-se que após a simulação de DM, as proteínas atingiram um *clashscore* de 0, com percentil de 100. Isso demonstra a importância da aplicação da DM para buscar modelos de maior qualidade e estabilidade estrutural.

Para cada proteína foram selecionados uma quantidade específica de modos (*clusters*) obtidos na DM, para prosseguir com os estudos de *docking*. No caso da *PbMLS*, nossa proteína alvo, foram extraídos 3 modos que foram considerados representativos, já para Enolase, por exemplo, foi extraído 1 modo conformacional devido a sua estabilidade durante a DM. Essa metodologia foi realizada para todas as outras proteínas, sendo selecionado no mínimo 1 e no máximo 3 modos conformacionais. A ideia principal foi cruzar os modos da *PbMLS* com todos os outros, que resultou em 36 complexos proteína-proteína.

5.4. Complexos formados e interface de interação

A partir dos modelos obtidos na DM, os principais modos conformacionais de cada proteína foram submetidas ao *molecular docking* com os 3 modos da *PbMLS* no GRAMM-X. Na Tabela 3, são apresentados os complexos formados e a nomenclatura para cada um deles.

Para observarmos os resíduos da interação entre uma proteína e outra, utilizamos o servidor KFC-2 que fornece a região de interface e os resíduos *hotspots*. As interações dos complexos entre os modos 1 estão exemplificadas nas Figuras 22 a 27 .

Tabela 3. Combinações realizadas no *molecular docking*.

Modos conformacionais das proteínas		MLS		
		Modo 1	Modo 2	Modo 3
Enolase	Modo 1	1.1	2.1	3.1
	Modo 2	1.2	2.2	3.2
GA	Modo 1	1.1	2.1	3.1
	Modo 2	1.2	2.2	3.2
TFI	Modo 1	1.1	2.1	3.1
	Modo 2	1.2	2.2	3.2
	Modo3	1.3	2.3	3.3
GFD	Modo 1	1.1	2.1	3.1
	Modo 2	1.2	2.2	3.2
Ubiquitina	Modo 1	1.1	2.1	3.1
	Modo 2	1.2	2.2	3.2
	Modo 3	1.3	2.3	3.3
MD	Modo 1	1.1	2.1	3.1
	Modo 2	1.2	2.2	3.2

Legenda: GA- Gama actina; TFI-Triosefosfato Isomerase; GFI –Gliceraldeído-3-fosfato desidrogenase; MD-Malato desidrogenase.

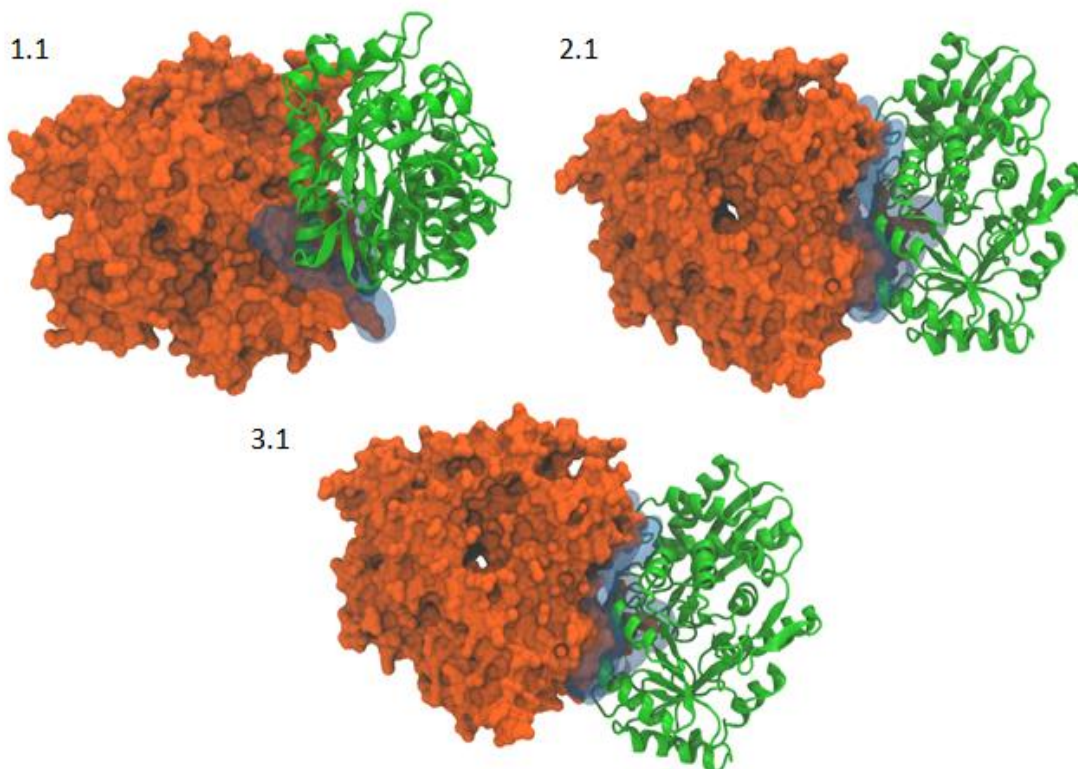


Figura 21. Modelos 3D dos complexos *PbMLS* e enolase. Em laranja temos a *PbMLS* em seus diferentes modelos provenientes da DM e em verde temos o modelo 1 da Enolase. A região em azul representa a região de interface da interação. Na face frontal da *PbMLS*, em todas as imagens, está representado a cavidade do sítio ativo desta enzima.

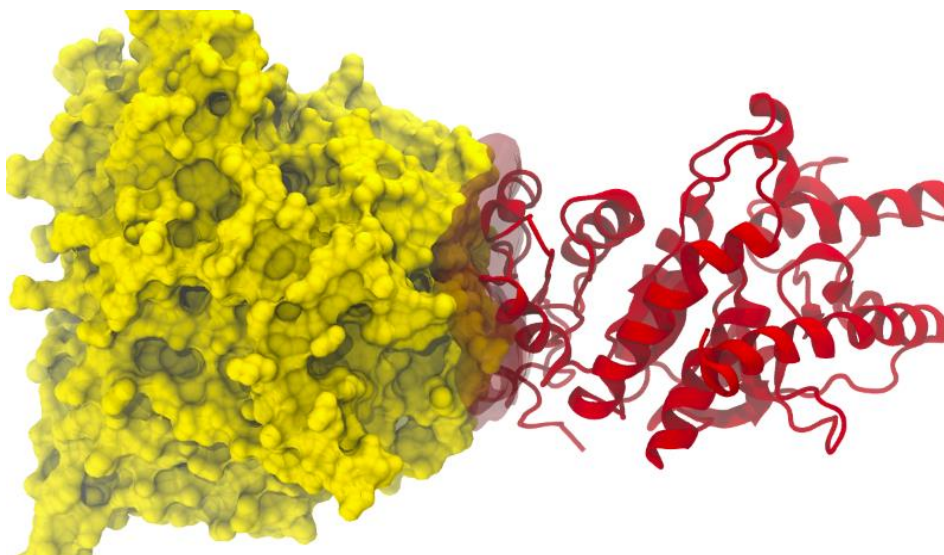


Figura 22. Representação tridimensional do complexo *PbMLS* e malato desidrogenase. Em amarelo temos a *MLS*, em vermelho formato *newcartoon*, temos o modo 1 da malato desidrogenase, e a região sombreada corresponde à interface de interação entre elas, predita pelo servidor KFC-2.

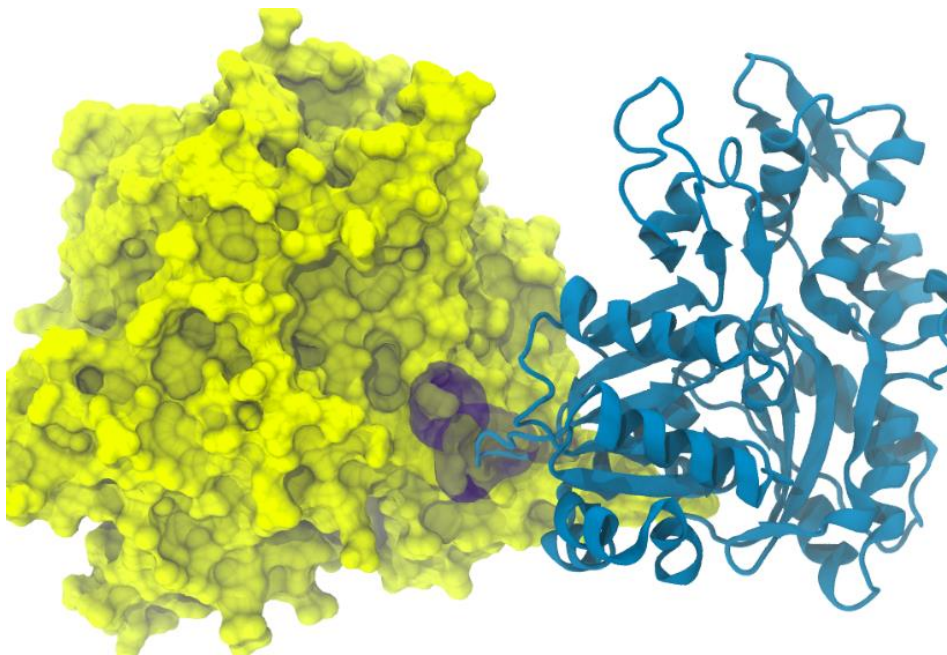


Figura 23. Representação tridimensional do complexo *PbMLS* e gama actina. Em amarelo temos a *PbMLS*, em azul (*newcartoon*), temos gama actina, e a região sombreada corresponde à interface de interação entre elas, predita pelo servidor KFC-2.

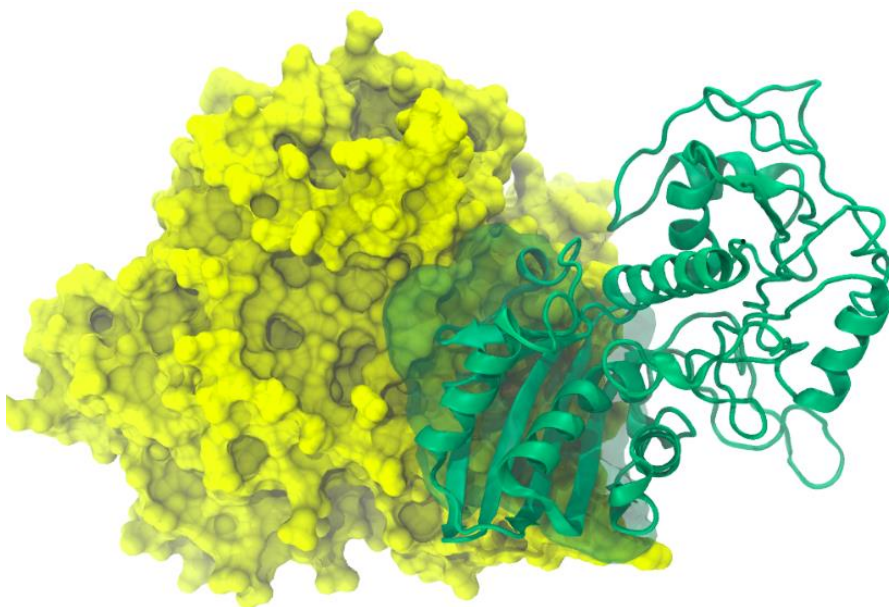


Figura 24. Representação tridimensional do complexo *MLS* e Gliceraldeído-3-fosfato desidrogenase. Em amarelo temos a *MLS*, em verde (*newcartoon*), temos a GAPDH, e a região sombreada corresponde à interface de interação entre elas, predita pelo servidor KFC-2

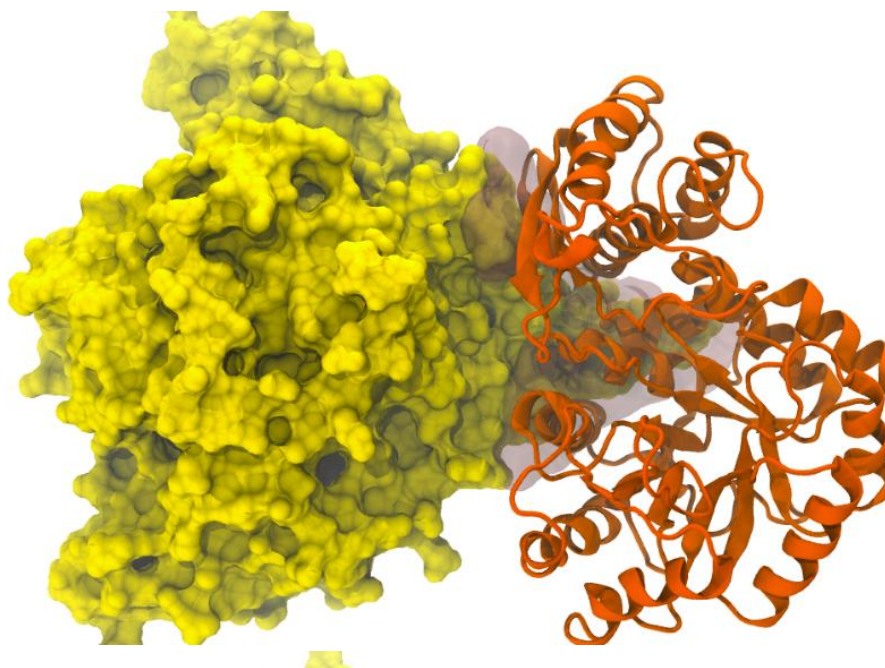


Figura 25. Representação tridimensional do complexo *PbMLS* e enolase. Em amarelo temos a *PbMLS*, em laranja (*newcartoon*), temos a enolase, e a região sombreada corresponde à interface de interação entre elas, predita pelo servidor KFC-2.

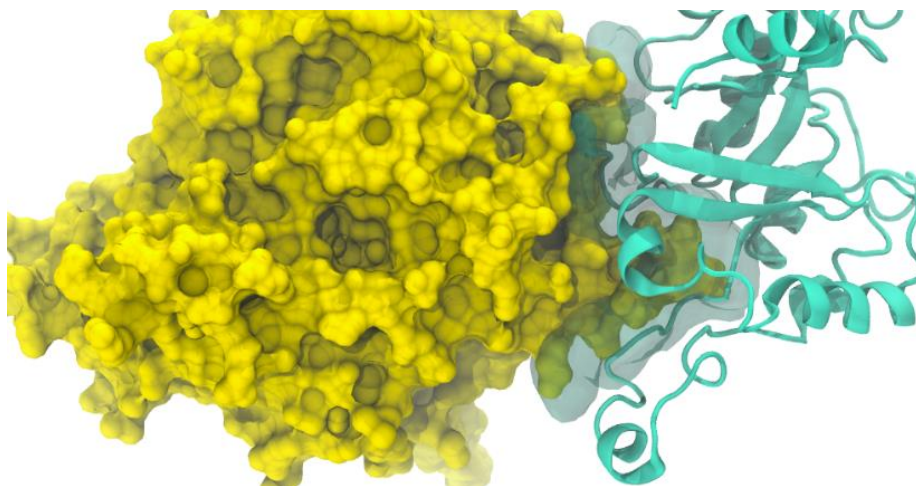


Figura 26. Representação tridimensional do complexo *PbMLS* e TFI. Em amarelo temos a MLS, em azul claro (*newcartoon*), temos a TPI, e a região sombreada corresponde à interface de interação entre elas, predita pelo servidor KFC2.

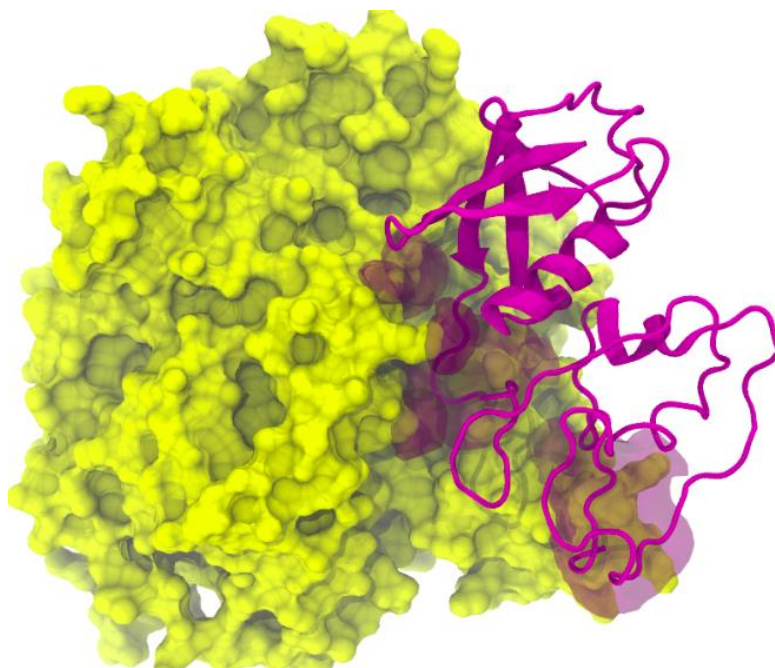


Figura 27. Representação tridimensional do complexo *PbMLS* e Ubiquitina. Em amarelo temos a *PbMLS*, em roxo (*newcartoon*), temos a Ubiquitina, e a região sombreada corresponde à interface de interação entre elas, predita pelo servidor KFC-2.

5.5. Análises das regiões mais interagentes de MLS

Entre todos os complexos analisados, 5 regiões definem as principais IPP's observadas em no mínimo 15 complexos do total de 36. A região mais frequente (R4), esteve presente em 22 complexos. Tais regiões estão registradas na Tabela 4. Na Figura 28 temos a representação molecular e na Figura 29 a representação gráfica dessas regiões da *PbMLS*.

Com o auxílio do servidor CoCoMAPS, foi possível identificar os resíduos das proteínas ligantes que interagiram nestas 5 regiões definidas como as mais frequentemente envolvidas nas interações.

Tabela 4. Regiões da MLS observadas com maior frequência nas interações proteína-proteína

	RESÍDUOS ENVOLVIDOS	Quantidade de complexos
R1	ALA85, PRO86, GLY87 e TYR88	22
R2	ARG160 e LYS161	23
R3	VAL205, GLY208, ALA209, GLY210, PRO 211 e TYR212	22
R4	ASP236, PRO241, ARG242, GLY243, THR244, ARG246	26
R5	PHE393 e VAL394	20



Figura 28. Regiões frequentes de interação da *PbMLS*. Malato sintase evidenciando as regiões mais frequentes de interação proteína-proteína. R1 (amarelo), R2 (azul), R3 (roxo), (verde) e R5 (rosa).

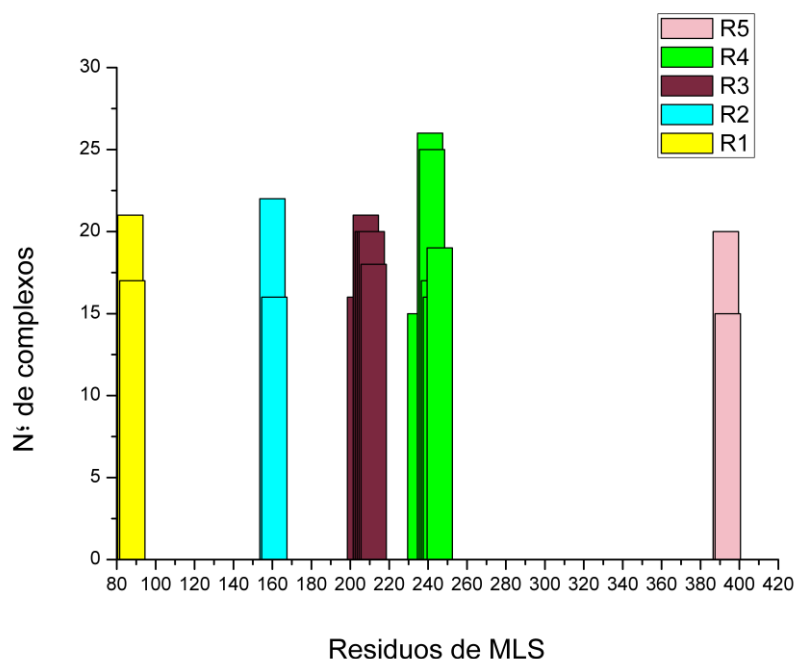


Figura 29. Gráfico das regiões frequentes de interação da MLS. As barras correspondem aos resíduos da IPP, que estão plotados em função das regiões determinadas como importantes da MLS e a quantidade de complexos que apresentam tais regiões envolvidas nas IPP's. Podemos observar que a região 4 (R4) apresentou a maior frequência, presentes em 26 IPP's.

5.6. Design de peptídeos

5.6.1. Aspectos gerais dos peptídeos

Uma vez conhecendo tais regiões da proteína alvo (MLS) pudemos desenhar e propor alguns peptídeos a partir das proteínas ligantes, para interagirem com esses resíduos da MLS. Os peptídeos propostos, foram montados a partir da estrutura tridimensional do arcabouço das interações proteína-proteína. Paralelamente, os peptídeos foram modelados no solvente (PepFOLD), a partir da sequência dos selecionados da forma já citada.

Todos estes foram submetidos ao FlexPepDock. Nas Tabelas 5, 6 e 7 temos a comparação das estruturas dos peptídeos modelados no solvente (PepFOLD) com os provenientes do arcabouço proteína-proteína e com os de menor energia oriundo do FlexPepDock, para os 3 modos da MLS.

Tabela 5. Peptídeos selecionados para interação com o modo 1 da MLS, e comparação das suas estruturas com os modelados *ab initio*.

Peptídeos	Sequência Peptídica	RMSD – PepFOLD X Arcabouço	RMSD - PepFOLD X FlexPepDock	Energia *
Pep 1	VPTANV	1,459	0,872	-220,045
Pep 2	LVSTD L	0,450	0,317	-222,366
Pep 3	LKGIL	0,675	0,58	-229,044
Pep 4	LDLIAYIAK	0,479	0,223	-223,944
Pep 5	MAPPEYPL	0,439	0,243	-222,256
Pep 6	ASRFISQVKGTD	1,323	0,676	-234,059
Pep 7	TLSMAM	0,314	0,122	-209,539
Pep 8	AASEGE	0,375	0,208	-217,988
Pep 9	EGELY	0,198	0,158	-222,618
Pep 10	MAPPE	0,45	0,07	-221,193
Pep 11	TSEDA	0,684	0,811	-217,654
Pep 12	MAGAG	0,252	0,422	-213,12
Pep 13	VPSYAP	1,319	1,324	-217,43
Pep 14	LESVPSYAPVVF	0,491	0,634	-229,522

*Score de energia do laboratório Rosetta

Tabela 6. Peptídeos selecionados para interação com o modo 2 da MLS, e comparação das suas estruturas com os modelados *ab initio*.

Peptídeos	Sequência Peptídica	RMSD – PepFOLD X Arcabouço	RMSD-Pep-FOLD	Energia
Pep 1	LDFEQ	0,298	0,354	-250,07
Pep 2	FQPSFLGMQ	1,042	0,723	-258,362
Pep 3	RMQKEITAL	0,462	0,277	-256,662
Pep 4	AGIAL	0,378	0,28	-247,076
Pep 5	DQIKA	0,158	0,123	-247,598
Pep 6	VSAPQ	0,565	0,65	-252,171
Pep 7	LLMKLSPRV	0,713	0,35	-253,319
Pep 8	PIRIKK	0,540	0,48	-247,998
Pep 9	GTPNK	0,417	0,407	-246,214
Pep 10	GQHEA	1,156	0,354	-246,143
Pep 12	PSSTGAAKA	0,275	0,191	-255,944
Pep 13	LDFEQEMATAASSSSLE	0,858	0,708	-273,249
Pep 14	DALVSTD	1,247	0,953	-253,134
Pep 15	ADLSHI	0,860	0,315	-250,471
Pep 18	ARGWHL D	1,378	0,684	-250,682
Pep 23	AEDDWEAGSY Y	0,986	0,432	-255,07
Pep 25	AVGKV	1,010	0,432	-255,294
Pep 30	MAMRVPTANVSV	0,891	1,289	-254,563
Pep 36	ECGAGV	1,365	0,645	-248,041

Tabela 7. Peptídeos selecionados para interação com o modo 3 da MLS, e comparação das suas estruturas com os modelados *ab initio*.

Peptídeos	Sequência Peptídica	RMSD – PepFOLD X Arcabouço	RMSD-Pep-FOLD	Energia
Pep 1	ELPDGQV	0,488	0,472	-247,468
Pep 2	SYTATQKTV	1,389	0,983	-246,183
Pep 3	INTNSTV	1,015	0,129	-237,100
Pep 4	VPSYAPVV	1,207	0,297	-246,575
Pep 5	KESTLH	0,242	0,170	-231,214
Pep 6	MKILTERGYSF	0,694	0,512	-251,009
Pep 7	IRGGPGVAA	0,653	0,610	-246,805
Pep 8	ASTGQHE	1,191	0,440	-234,273
Pep 9	SPSCLKD	0,484	0,267	-234,812
Pep 10	ELPPDGQV	0,486	0,679	-242,495
Pep 11	RTAAQNI	1,233	0,831	-244,755
Pep 12	AASEGELK	0,310	0,244	-236,933
Pep 13	ADLSHINT	0,745	0,659	-241,804
Pep 14	FAEDDW	0,680	0,5	-224,714
Pep 17	KTLTGKT	0,685	0,858	-241,951
Pep 18	ESSDTIDN	1,375	1,076	-243,300
Pep 19	SLLMKL	0,460	0,211	-247,131
Pep 20	AEDDWEAGSY	0,721	0,444	-252,886
Pep 21	QNNQRQ	1,111	0,764	-238,285
Pep 22	AAVSVH	0,587	0,184	-234,678
Pep 23	TITLE	0,828	0,795	-236,789
Pep 24	QSKRGI	0,390	0,324	-238,579
Pep 25	SEDALV	0,539	0,572	-233,761
Pep 26	LESVPSYAPVVF	1,462	1,058	-248,023

5.7. Avaliação das interações proteína-peptídeo

Para seleção dos melhores peptídeos para cada modelo da MLS, foram analisados o *score* de energia livre de interação, o menor RMSD comparado ao peptídeo correspondente modelado no solvente (PepFOLD), e, as características dos contatos (interações eletrostáticas, de hidrogênio e hidrofóbicas). Tais aspectos de contato foram analisados com o auxílio do programa DiscoveryStudios. Inicialmente, podemos destacar 6 peptídeos para cada modo da *PbMLS* que apresentaram os menores *scores* de energia*, e menores RMSDs comparado ao PepFOLD.

As análises de interação permitiram a identificação de interações favoráveis e desfavoráveis. Dentre as favoráveis, destacamos as hidrofóbicas classificadas de acordo com o grupo funcional envolvido na interação com grupos alquil, pi-alquil e pi-hidrofóbica. A interação pi envolve anel aromático, sendo pi-alquil quando este interage com o grupo alquila (C-H). As interações hidrofóbicas são essenciais para o reconhecimento intermolecular, uma vez que estão envolvidas nas etapas iniciais de ligação a outras moléculas. Dessa forma, uma superfície hidrofóbica otimiza o processo de interação (GARDE, 2015).

Além disso, temos as interações de hidrogênio clássicas que envolvem C-H e não clássicas onde outros átomos, diferentes de C, interagem com o H. Com relação a interação eletrostática, temos as interações carga-pi, a qual envolve um átomo carregado com um anel aromático, e as do tipo pi, onde o anel aromático interage com o H de outro resíduo (HUDSON et al, 2015).

Uma outra interação observada ocorre entre o anel aromático e o enxofre, sendo classificada como pi-sulfur. A respeito das interações não favoráveis, foram observados choques estéricos, repulsão de carga e repulsão de doadores e aceptores de H.

5.7.1. Peptídeos promissores ancorados ao Modo 1 da *PbMLS*

Para o modo 1 da *PbMLS*, o peptídeo 3 apresenta 5 resíduos, sendo as duas extremidades apolares, e carga total líquida de +1 (apenas a LYS como resíduo carregado positivamente). Na Figura 30, temos a superfície de contato com a *PbMLS* e a sua caracterização quanto à hidrofobicidade.

A respeito das interações hidrofóbicas, podemos destacar 4 interações do tipo Alquil, na qual temos interação de resíduos hidrofóbicos e seus grupos funcionais alquil, e uma interação do tipo *pi*-alquil na qual o grupo alquil de um resíduo hidrofóbico interage com um anel de outro resíduo hidrofóbico. No primeiro caso, temos a LEU 270 como principal resíduo interagente, pois interage com os resíduos ILE273 (4,39 Å), LEU274 (4,98 Å), que são do próprio peptídeo, e PRO211 (4,9 Å) da MLS. Além disso, a ILE273 realiza interação alquil com a VAL89 (5,04 Å) da MLS. No segundo caso, temos o grupo alquil da ILE273 interagindo com o fenol da TYR113 da MLS.

Na Figura 31, temos as principais interações destacadas entre os resíduos do peptídeo com os resíduos da MLS. Destacamos 3 interações de hidrogênio, sendo que a LEU270 realiza duas interações deste tipo, uma com o resíduo ILE273 e outra com o resíduo LEU274 do próprio peptídeo, essas com cerca de 2 Å de comprimento. Observamos uma interação de atração de cargas (eletrostática) entre os resíduos ARG246 (MLS) e o resíduo LEU274 (peptídeo) de 4,11 Å.

Tais interações são favoráveis entre o peptídeo e a proteína, destacamos as interações hidrofóbicas estiveram bastante presentes neste caso, e são essenciais para o contato inicial do peptídeo na região desejada da proteína, mesmo em solvente.

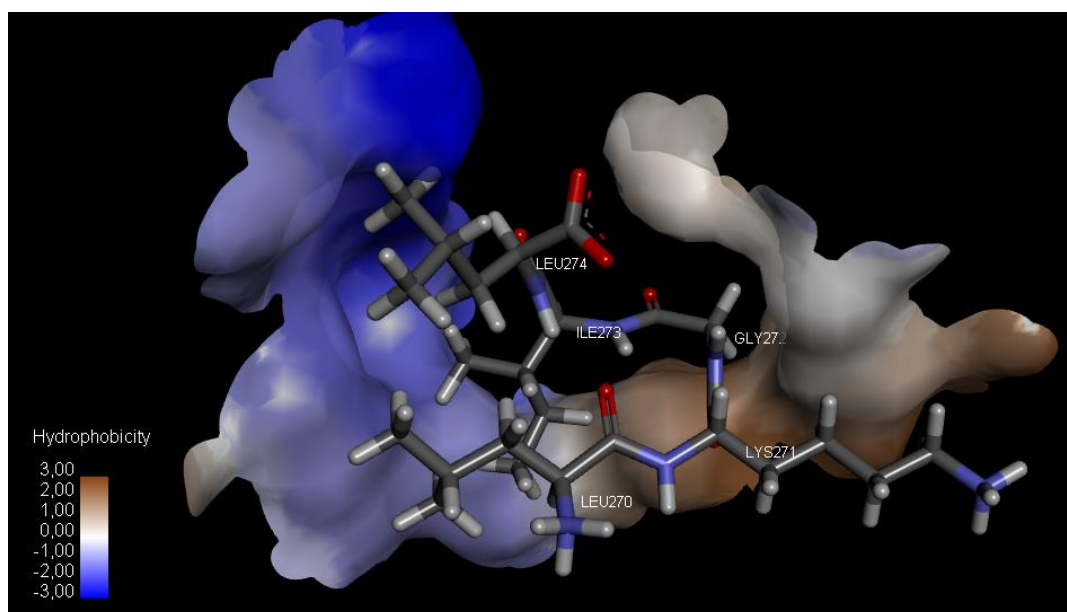


Figura 30. Peptídeo 3 ancorado à superfície da MLS caracterizada quanto a sua hidrofobicidade.

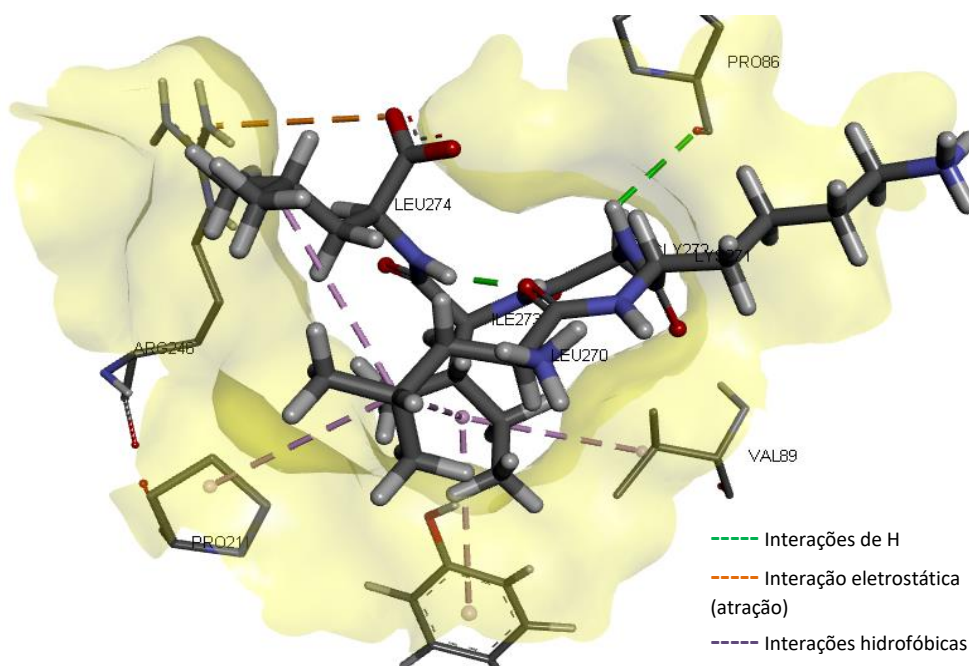


Figura 31. Peptídeo 3 ancorado à superfície da MLS, evidenciando suas interações intermoleculares.

O peptídeo 4 apresenta 9 resíduos, sendo a maioria de caráter apolar, e carga total líquida nula, uma vez que temos um resíduo carregado positivamente (LYS) e outro negativamente (ASP). De acordo com a Figura 32, observamos uma região hidrofóbica na superfície da MLS que não interage diretamente com resíduos hidrofóbicos do peptídeo.

Em especial pode-se destacar o resíduo 241 da MLS realizando 3 interações com o peptídeo: duas interações do tipo Carbono-Hidrogênio (interação de H não clássica) com a TYR330 e a LEU325 do peptídeo, a uma distância inferior a $3\text{\AA}$, e a terceira interação corresponde a uma interação hidrofóbica do tipo *pi*-alquil com a TYR330. O resíduo ARG242 realiza duas interações hidrofóbicas com ALA e LEU do peptídeo. Dentro do próprio peptídeo, podemos destacar uma interação eletrostática de atração da LEU325 com o ASP326, além de outras interações de hidrogênio clássicas ou não, como observamos na Figura 33.

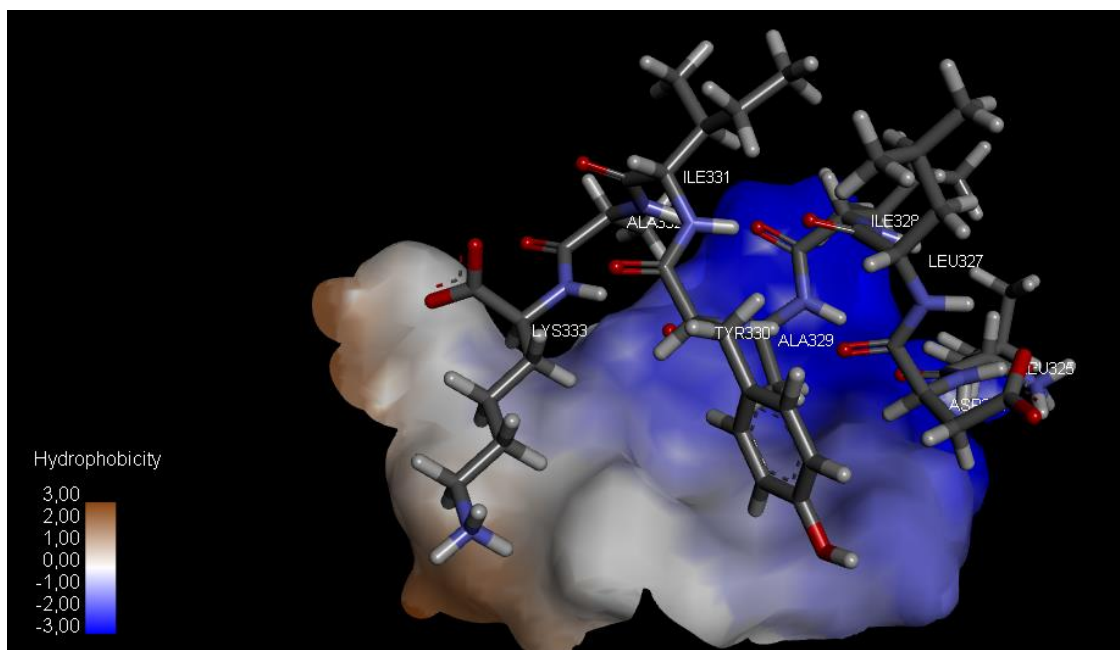


Figura 32. Peptídeo 4 em contato com a superfície da MLS, caracterizada quanto a sua hidrofobicidade.

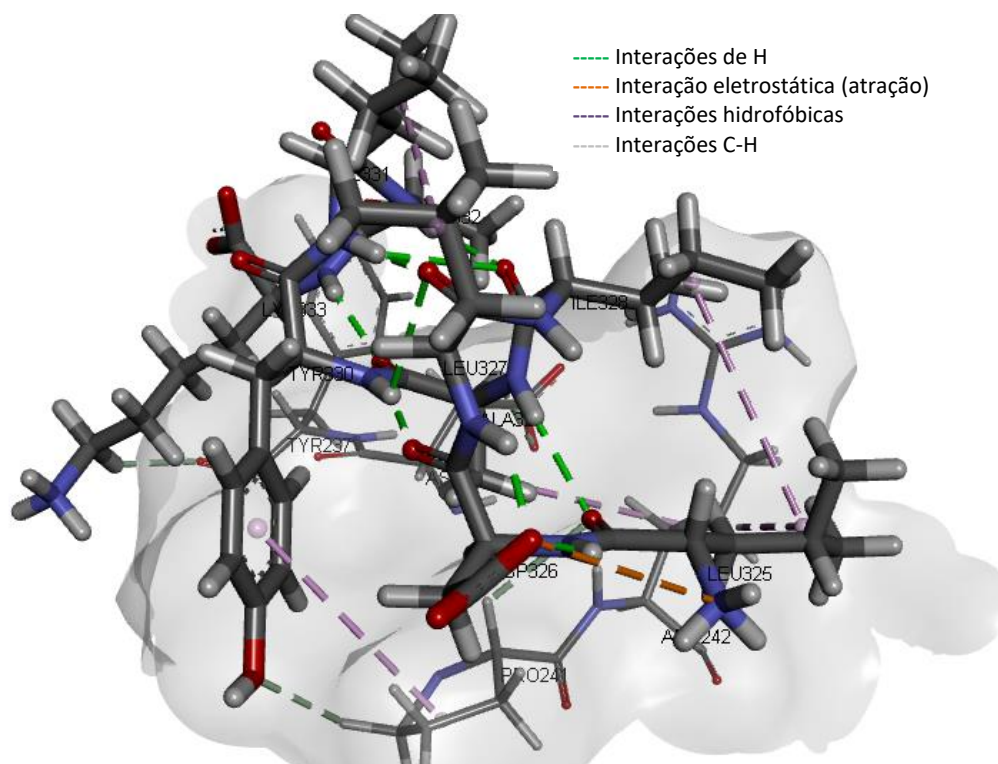


Figura 33. Peptídeo 4 ancorado à superfície da MLS, evidenciando suas interações intermoleculares.

Outro peptídeo que podemos destacar é o de número 7, que apresentou estrutura próxima ao seu correspondente modelado no solvente. Nele, temos 6 resíduos dos quais 4 possuem caráter apolar (MET, ALA e LEU), o que é complementar à estrutura da MLS como observado na Figura 34.

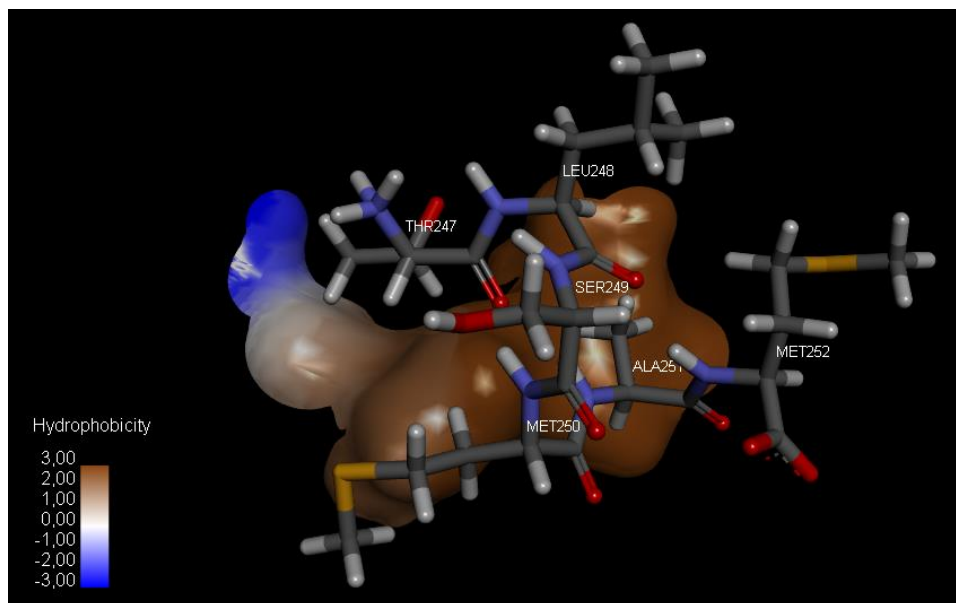


Figura 34. Peptídeo 7 ancorado à superfície hidrofóbica da MLS. Observa-se a complementaridade apolar entre o peptídeo e a proteína neste caso.

O peptídeo 7 realiza além das interações hidrofóbicas do tipo alquil e *pi*-alquil (entre proteína e peptídeo) e das interações de hidrogênio clássicas (entre resíduos do peptídeo), como observado para os peptídeo anteriores, nota-se também uma interação do tipo *pi*-sulfur. Tal interação ocorre entre o enxofre (S) a MET250 com o anel aromático da PHE393 da MLS, destacadas na Figura 35.

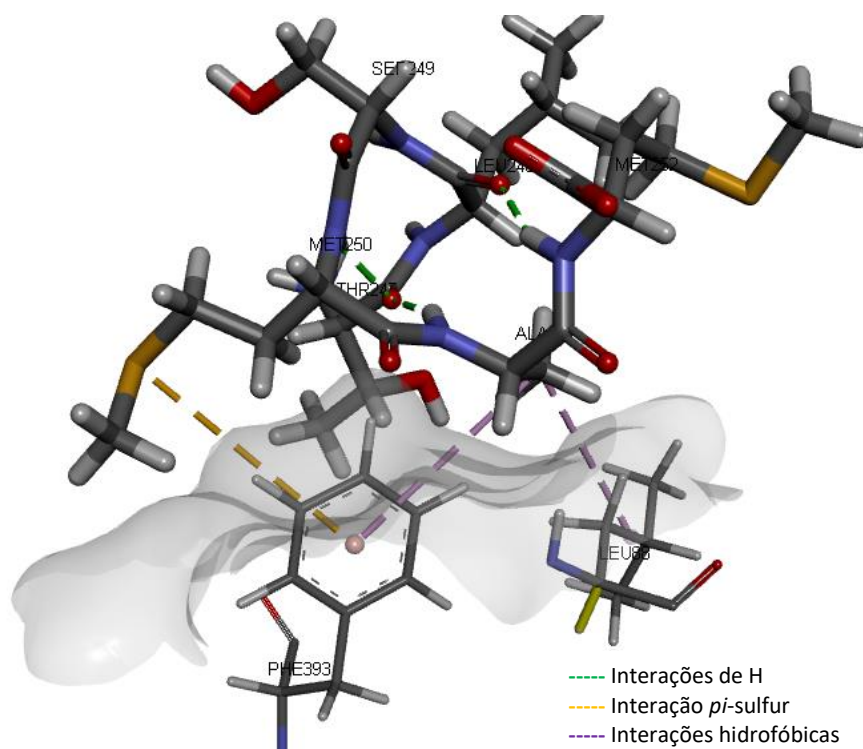


Figura 35. Peptídeo 7 ancorado à superfície da MLS, evidenciando suas interações intermoleculares.

O peptídeo 10 possui 5 resíduos, sendo em sua maioria de caráter apolar, e carga total líquida -1. A respeito de suas interações podemos destacar várias interações de caráter hidrofóbico, na Figura 36, com todos os resíduos apolares do peptídeo 10. Observam-se duas interações do tipo *pi*-alquil com o resíduo TYR113 da MLS, uma interação de hidrogênio entre o enxofre (S) da MET1 com a ARG246 da MLS. Além disso, pode-se observar 3 interações eletrostáticas, sendo 2 pontes salinas e 1 interação com cargas atrativas.

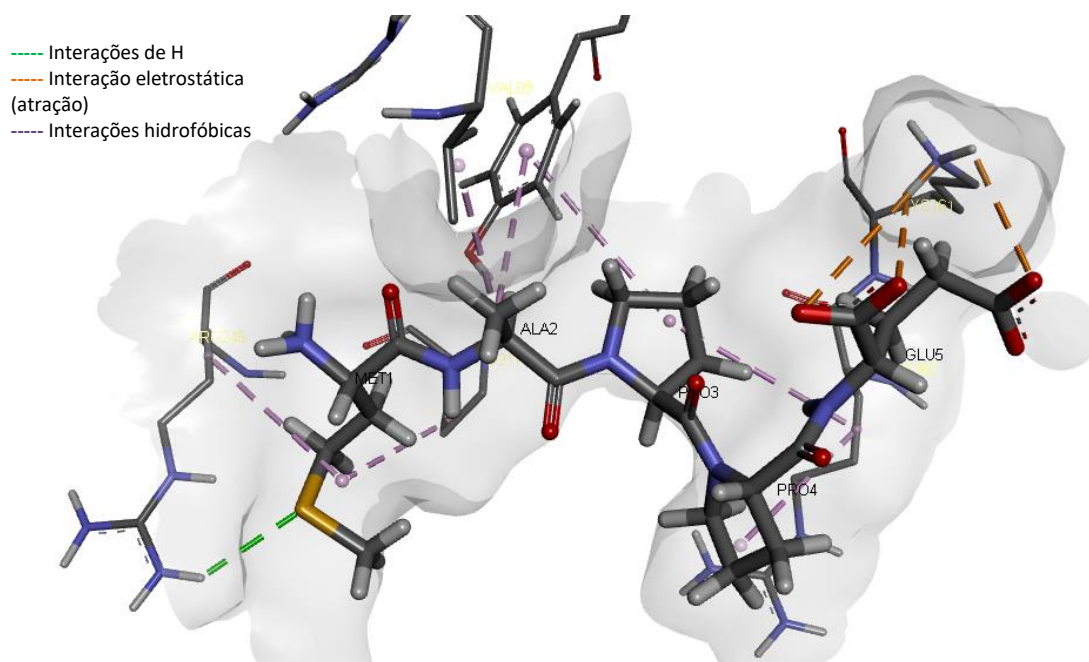


Figura 36. Peptídeo 10 ancorado a superfície da MLS evidenciando as interações intermoleculares.

O peptídeo 14 é interessante pois apresentou baixo *score* de energia, e corresponde a uma estrutura com 12 aminoácidos com carga total líquida -1. A superfície de contato da MLS apresenta uma complementaridade hidrofóbica interessante com o peptídeo, como pode-se observar na Figura 37.

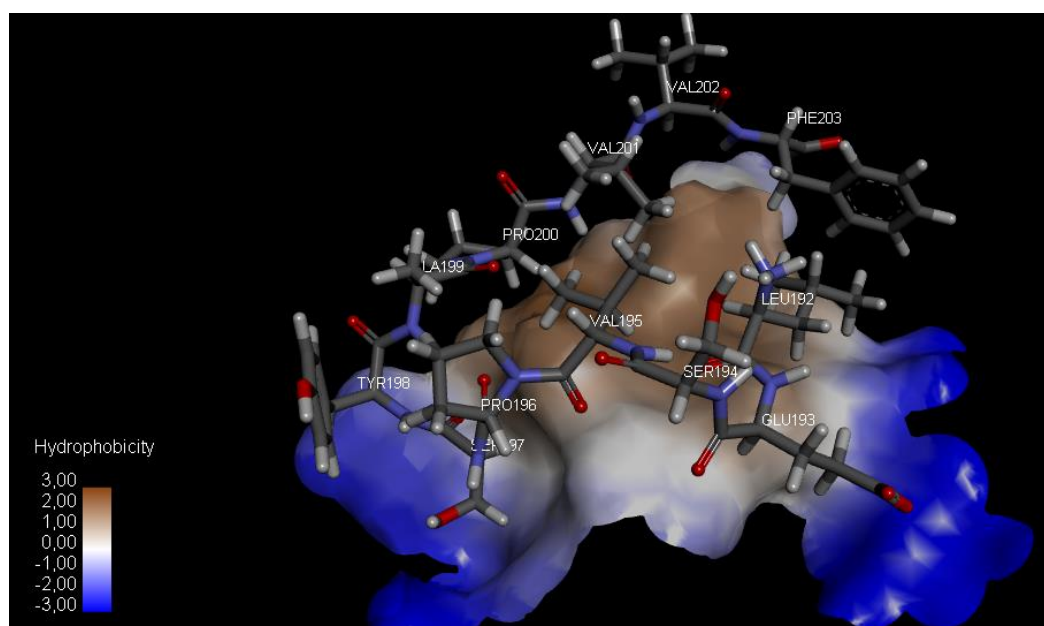


Figura 37. Peptídeo 14 ancorado à superfície da MLS, evidenciando a superfície hidrofóbica de contato.

Na Figura 38, observamos com maior detalhe as interações envolvidas. Temos várias interações do tipo hidrofóbica, o que confirma a complementaridade observada na Figura 37. Várias estruturas de anéis aromáticos estão envolvidas na interação, tanto no peptídeo quanto na proteína, o que permite a realização de 3 interações do tipo *pi*. Dessas, uma delas envolve os resíduos LEU192 e PHE193 e é caracterizada como uma interação *pi-sigma*, dentro do próprio peptídeo. Além disso, temos algumas interações de hidrogênio dentro do próprio peptídeo e algumas entre proteína e peptídeo.

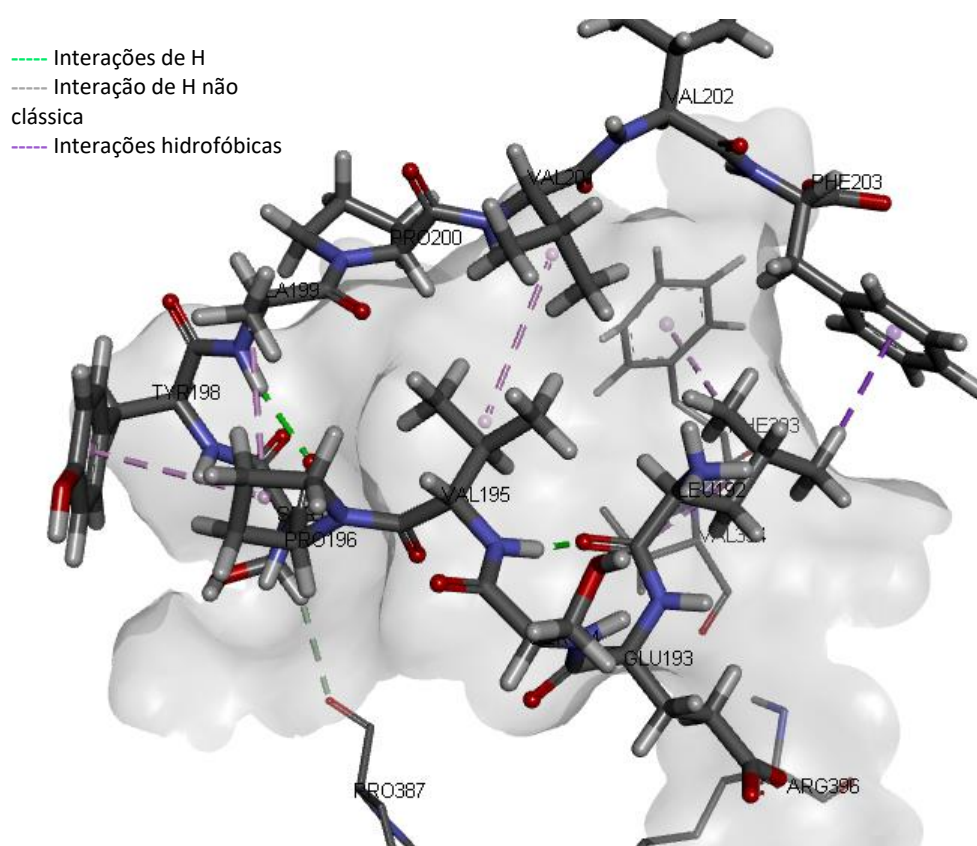


Figura 38. Peptídeo 14 ancorado à superfície da MLS, evidenciando as interações intermoleculares. Nota-se grande complementaridade hidrofóbica, com muitas interações do tipo *pi*.

6. CONSIDERAÇÕES FINAIS

6.1. Conclusões

As conclusões deste trabalho são promissoras, porém parciais em torno do objetivo final. A respeito dos objetivos propostos inicialmente, todos foram cumpridos, como por exemplo, a modelagem das proteínas, ancoragem proteína-proteína e determinação de regiões mais frequentes envolvidas nas interações. Além disso, foram propostos 5 peptídeos promissores para interação com a *PbMLS*.

As análises realizadas, juntamente com os peptídeos para os Modos 2 e 3 da *PbMLS*, que serão realizados em breve, servirão como um afunilamento para a etapa de mutações aleatórias, onde iremos observar a qualidade dessas interações e possivelmente encontrar outros peptídeos com boa estabilidade. Outra etapa importante será a DM com os complexos proteína-peptídeo, a fim de verificar quão estável são essas interações e encontrar modos conformacionais mais promissores.

Os 5 peptídeos pré-selecionados nos resultados, com o modo 1 da *MLS*, apresentaram interações interessantes do ponto de vista físico-químico, as quais servirão como passo inicial de análise dos mesmos. As análises de afinidade do complexo proteína-peptídeo são iniciais, visto que para inferirmos moléculas para testes *in vitro*, são necessárias análises mais detalhadas com DM dos complexos e montagem de peptídeos com possibilidade de síntese.

6.2. Perspectivas

As interações proteína-proteína são essenciais para a homeostase do sistema, uma vez que as formações de complexos são importantes em várias vias metabólicas, na interação parasito-hospedeiro e no processo de resposta celular. Assim, o estudo das interações *in silico* de algumas proteínas importantes para a sobrevivência do fungo *Paracoccidioides* sp. nos permitiu observar os aminoácidos mais envolvidos nas interações.

Conhecendo os resíduos mais importantes da *PbMLS* pudemos propor peptídeos que possam apresentar afinidade às regiões. As análises das interações, do score de energia e da flexibilidade desses peptídeos foram

importantes para selecionar os mais promissores para validação. Esta, será realizada por meio de mutações aleatórias e pontuais e dinâmica molecular dos complexos proteína-peptídeo, que serão realizadas em breve.

A partir desses resultados poderão ser apresentados complexos interessantes para testes *in vitro* e *in vivo*, com a intenção de modulação por interação com a PbMLS. Esta modulação pode ser a partir da alteração da conformação funcional da proteína, ou até mesmo do bloqueio de interações essenciais com outras proteínas.

Sabe-se que a proteína no sistema vivo, pode variar sua conformação de acordo com a situação em que está envolvida, porém os estudos *in silico* servem como uma ferramenta inicial de estudo com a finalidade de direcionar a pesquisas e diminuir assim gastos financeiros e tempo.

7. REFERÊNCIAS

- ANDRADE, R. V. Análise do Transcriptoma e da Expressão Diferencial de Genes de Micélio e Levedura de *Paracoccidioides brasiliensis*. p. 136, 2006.
- ARISTIZABAL B.H., CLEMONS K.V., COCK A.M., RESTREPO A., STEVENS D.A. Experimental *Paracoccidioides brasiliensis* infection in mice: influence of the hormonal status of the host on tissue responses. **Medical Mycology**, v.40, p.169-78, 2002.
- AUERBACH, D.; THAMINY, S.; HOTTIGER, M.O.; STAGLJAR, I. The post-genomic era of interactive proteomics: facts and perspectives. **Proteomics** v. 2, n. 6, p. 611-623, 2002.
- BAGAGLI E., BOSCO S.M.G., THEODORO R.C., FRANCO M. Phylogenetic and evolutionary aspects of *Paracoccidioides brasiliensis* reveal a long coexistence with animal hosts that explain several biological features of the pathogen. **Infection, Genetics and Evolution** , v. 6, p. 344-351, 2006.
- BARROZO L.V., MENDES R.P., MARQUES S.A., BENARD G., SILVA M.E.S., BAGAGLI E. Climate and acute/subacute paracoccidioidomycosis in a hyper-endemic area in Brazil. **International Journal of Epidemiology** v.38, p. 1642-1649, 2009.
- BENARD G. An overview of the immunopathology of human paracoccidioidomycosis. **Mycopathologia**.v. 165, p. 209-21, 2008.
- Benard G., Mendes-Giannini, M.J.S. Paracoccidioidomycosis. In: Feigin RD, Cherry JD, Demmler-Harrison G, Kaplan SL, editors. **Textbook of pediatric infectious diseases**. 6^a ed. Philadelphia: Elsevier Science. p.2762-76, 2009.
- BOCCA A.L., AMARAL A.C., TEIXEIRA M.M., SATO P.K., SATO P., SHIKANAI-YASUDA M.A., et al. Paracoccidioidomycosis: eco-epidemiology, taxonomy and clinical and therapeutic issues. **Future Microbiol.** v. 8, p. 1177-91, 2013.
- BRUMMER E., CASTANEDA E., RESTREPO A. Paracoccidioidomycosis: an update. **Clinical Microbiology Reviews**. v.6, p. 89-117, 1993.
- CARNEIRO, L. C. et al. *Paracoccidioides brasiliensis* presents two different cDNAs encoding homologues of the fructose 1,6-biphosphate aldolase: Protein isolation, cloning of the cDNAs and genes, structural, phylogenetic, and expression analysis. **Fungal Genetics and Biology**, v. 42, n. 1, p. 51–60, 2005.
- CASTRO R.M., CUCE L.C., NETTO C.F. Paracoccidioidomycosis. Accidental inoculation "in anima nobile." Report of a case. **Medicina Cutânea Ibero-Latino-Americana**. v. 3, 289-92, 1975.

CAUSIER, B. Studying the interactome with the yeast two-hybrid system and mass spectrometry. **Mass Spectrometry Reviews**. v. 23, n. 5, p350-367, 2004.

CHANG, M. W. et al. Virtual screening for HIV protease inhibitors: A comparison of AutoDock 4 and Vina. **PLoS ONE**, v. 5, n. 8, p. 1–9, 2010.

CHENG, A. C.; COLEMAN, R. G.; SMYTH, K. T. et al. Structure-based maximal affinity model predicts small-molecule druggability. **Nature biotechnology**, v. 25, n. 1, p. 71-5, 2007.

CHO, K. II.; KIM, D.; LEE, D. A feature-based approach to modeling protein-protein interaction hot spots. **Nucleic Acids Research**, v. 37, n. 8, p. 2672–2687, 2009.

CHO, S.; PARK, S.G.; LEE, D.H.; PARK, B.C. (). Protein-protein interaction networks: from interactions to networks. **Journal of Biochemistry and Molecular Biology**. 37(1):45-52, 2004.

COLEMAN R.G., CARCHIA M., STERLING T., IRWIN J.J., SHOICHET B.K. Ligand Pose and Orientational Sampling in Molecular Docking. **PLoS One** v.8 n.10, 2013.

CRAIK D.J., FAIRLIE D.P., LIRAS S., PRICE D. The future of peptide-based drugs. **Chemical Biology & Drug Design**. p. 136-47, 2013.

COSTA, F. G. et al. Alkaloids as inhibitors of malate synthase from *Paracoccidioides* spp.: Receptor-ligand interaction-based virtual screening and molecular docking studies, antifungal activity, and the adhesion process. **Antimicrobial Agents and Chemotherapy**, v. 59, n. 9, p. 5581–5594, 2015.

COZZONE, A.J. Regulation of acetate metabolism by protein phosphorylation in enteric bacteria. **Annual Review of Microbiology**. 52:127-164, 1998.

DAURA, X. et al. Peptide folding: when simulation meets experiment. **Angewandte Chemie International Edition**, v. 38, p. 236–240, 1999.

DA SILVA NETO, B. R. et al. The malate synthase of *Paracoccidioides brasiliensis* is a linked surface protein that behaves as an anchorless adhesin. **BMC Microbiology**, v. 9, p. 272, 2009. ISSN 1471-2180. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20034376> >.

DE OLIVEIRA, K. M. et al. Intermolecular interactions of the malate synthase of *Paracoccidioides* spp. **BMC Microbiology**, v. 13, p. 107, 2013. ISSN 1471-2180. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23672539> >.

DERENGOWSKI L.S 2006. Análise da expressão gênica de potenciais fatores de virulência do fungo *Paracoccidioides brasiliensis* submetido a diferentes condições experimentais. Dissertação de mestrado. **Universidade de Brasília**. Instituto de Ciências Biológicas.

DUTRA, L.M.F. et al. Modelagem molecular de proteínas: o caso de uma glucoronosiltransferase (GumK) de *Gluconacetobacter diazotrophicus* PAL5. **Journal of Biology & Farmacy and Agricultural Management**. v. 10, p. 56–67, 2014.

FALLIS, A. **Journal of Chemical Information and Modeling**, v. 53, n. 9, p. 1689–1699, 2013.

LORENZ, M.C., FINK, G.R. The glyoxalate cycle is required for fungal virulence. **Nature**. 412: 83-86, 2001.

FINLAY, B.B. and FALKOW, S. Common themes in microbial pathogenicity revisited. **Microbiol. Mol. Biol. Rev.** 61:136-169, 1997.

FISHMAN, M. C. & PORTER, J. A. Pharmaceuticals: a new grammar for drug discovery. **Nature**, v. 437, n. 7058, p. 491-3, 2005.

FORTES, M. R. P. et al. Imunologia da paracoccidioidomicose. **Anais Brasileiros de Dermatologia**, v. 86, p. 516–524, 2011.

FRANCO M., MONTENEGRO M.R., MENDES R.P., MARQUES S.A., DILLON N.L., MOTA N.G. Paracoccidioidomycosis: a recently proposed classification of its clinical forms. **Rev Soc Bras Med Trop**.v. 20, p. 129-32, 1987.

Guex, N. and Peitsch, M.C. (1997). **and the Swiss-PdbViewer: An environment for comparative protein modeling.** *Electrophoresis* **18**, 2714-2723.

GÓES, A. M. et al. Paracoccidioidomycosis disease (Lutz-Splendore-Almeida): etiology, epidemiology, and pathogenesis. **Revista Médica de Minas Gerais**, v. 24, n. 1, p. 61–66, 2014.

HALPERIN, I. et al. Principles of docking: An overview of search algorithms and a guide to scoring functions. **Proteins: Structure, Function and Genetics**, v. 47, n. 4, p. 409–443, 2002.

HEINEMANN, U. Structural genomics in Europe: Slow start, strong finish? **Nat. Struct. Biol.** v. 7, p. 940–942, 2000.

HETENYI C., VAN DER SPOEL D. Efficient docking of peptides to proteins without prior knowledge of the binding site. **Protein Science**. v. 11, p. 1729–1737, 2002.

HOPKINS, A. L. & GROOM, C. R. The druggable genome. **Nature reviews. Drug discovery** v.1, n. 9, p. 727-30, 2002.

HONG ENRIQUEZ, R. P. et al. Designing short peptides with high affinity for organic molecules: A combined docking, molecular dynamics, and Monte Carlo approach. **Journal of Chemical Theory and Computation**, v. 8, n. 3, p. 1121–1128, 2012.

HUMPHREY, W., DALKE, A. AND SCHULTEN, K. VMD - Visual Molecular Dynamics. **J. Molec. Graphics**. v. 14, p. 33-38, 1996.

HUTTER, J. Car-Parrinello molecular dynamics. **Computational Molecular Science**, v. 2, n. February, p. 604–612, 2012.

IANNUZZI, M.; LAIO, A.; PARRINELLO, M. Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics. **Physical Review Letters**, v. 90, 2003.

JAYAPALAN, S.; NATARAJAN, J. The role of CDK5 and GSK3B kinases in hyperphosphorylation of microtubule associated protein tau (MAPT) in Alzheimer's disease. **Bioinformation**, v. 9, n. 20, p. 1023–30, 2013.

LEHNINGER, A.L., NELSON, D.L., COX, M.M. **Principios de Bioquímica**. Editora Savier. Ed. 2ª, 2000. São Paulo, Brasil.

LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, v. 23, n. 1-3, p. 3-25, 1997.

LONDON, N. et al. Rosetta FlexPepDock web server - High resolution modeling of peptide-protein interactions. **Nucleic Acids Research**, v. 39, n. SUPPL. 2, p. 249–253, 2011.

LOPEZ-MARTÍNEZ, R. et al. Inhibition of Influenza A Virus Infection In Vitro by Peptides Designed In Silico. **PLoS ONE**, v. 8, n. 10, p. 1–12, 2013.

LOVELL, I.W. DAVIS, W.B. ARENDALL III, P.I.W. DE BAKKER, J.M. WORD, M.G. PRISANT, J.S. RICHARDSON and D.C. RICHARDSON. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. **Proteins: Structure, Function & Genetics**. v.50, p. 437-450, 2002.

LORENZ, M.C. e FINK, G.R. The glyoxylate cycle is required for fungal virulence. **Nature**. v. 412, p. 83-86, 2001.

LORENZ, M.C. E FINK, G.R. Life and Death in a Macrophage: Role of the Glyoxylate Cycle in Virulence. **Eukaryotic cell**. v. 1, p. 657-662, 2002.

LUTZ A. Uma micose pseudococcídica localizada na boca e observada no Brasil: contribuição ao conhecimento das hiphoblastomicoses americanas. **Brasil Med**, v. 22, p. 121-124, 1908.

MAGGIO, E. T.; RAMNARAYAN, K. Recent developments in computational proteomics. **Drug Discovery Today**, v. 6, n. 19, p. 996–1004, 2001.

MATUTE D.R., MCEWEN J.G., PUCCIA R., MONTES B.A., SAN-BLAS G., BAGAGLI E., RAUSCHER J.T., RESTREPO A., MORAIS F., NIÑO-VEGA G., TAYLOR J.W. Cryptic speciation and recombination in the fungus *Paracoccidioides brasiliensis* as revealed by gene genealogies. **Mol. Bio. Evol.** v. 23, n.1, p. 65-73, 2006.

MARQUES DA SILVA SH, RODRIGUES AM, DE HOOG S, SILVEIRA-GOMES F, CAMARGO ZP. Occurrence of *Paracoccidioides lutzii* in the Amazon region: description of two cases. **Am J Trop Med Hyg.** v. 87, p.710-4, 2012.

MARQUES S.A., FRANCO M.F., MENDES R.P., SILVA N.C., BACCILI C., CURCELLI E.D., et al. Epidemiologic aspects of paracoccidioidomycosis in the endemic area of Botucatu (São Paulo - Brazil). **Rev Inst Med Trop Sao Paulo.** v. 25, p. 87-92, 1983.

MARQUES S.A., CORTEZ D.B., LASTORIA J.C., CAMARGO R.M.P., MARQUES M.E.A. Paracoccidioidomycose: Frequência, morfologia e patogênese de lesões tegumentares. **An Bras Dermatol.** v. 82, p. 411-7, 2007.

MARQUES S.A. Paracoccidioidomycosis: epidemiological, clinical, diagnostic and treatment up-dating. **An Bras Dermatol.**; v. 88, n. 5, p. 700-11, 2013.

MARTINEZ R. Paracoccidioidomycosis: The dimension of the problem of a neglected disease. **Rev Soc Bras Med Trop.** v. 43, p. 480, 2010.

MASHIACH, D. SCHNEIDMAN-DUHOVNY, N. ANDRUSIER, R. NUSSINOV, H. J. WOLFSON. FireDock: a web server for fast interaction refinement in molecular docking. **Nucleic Acids Res**, 2008.

MAUPETIT, J.; DERREUMAUX, P.; TUFFERY, P. PEP-FOLD: An online resource for de novo peptide structure prediction. **Nucleic Acids Research**, v. 37, n. SUPPL. 2, p. 498–503, 2009.

MATOS W.B., SANTOS G.M.C., SILVA U.E.B., GONÇALVES E.G.R., SILVA A.R. Paracoccidioidomycosis in the state of Maranhão, Brazil: geographical and clinical aspects. **Rev Soc Bras Med Trop.** v. 45, p. 385-9, 2012;

MENDES-GIANNINI MJ, MORAES RA, RICCI TA. Proteolytic activity of the 43,000 molecular weight antigen secreted by *Paracoccidioides brasiliensis*. **Rev Inst Med.Trop** Sao Paulo. v. 32, p. 384-5, 1990.

MENDES-GIANNINI M.J., MONTEIRO DA SILVA J.L., DE FÁTIMA DA SILVA J, DONOFRIO F.C., MIRANDA E.T., ANDREOTTI P.F., et al. Interactions of *Paracoccidioides brasiliensis* with host cells: recent advances. **Mycopathologia.** v. 165, p. 237-48, 2008.

MINISTÉRIO DA SAÚDE (Brasil). **Proposta de vigilância epidemiológica da paracoccidioidomycose.** Brasília, 2012. 11p.

MOREIRA A.P.V. Paracoccidioidomycose: histórico, etiologia, epidemiologia, patogênese, formas clínicas, diagnóstico laboratorial e antígenos. **Biol Epidemiol Paulista**.v. 5, n. 51, 2008.

NOGALES, J., GUIJO, M.I, QUESADA, A., MERCHAN, F. . Functional analysis of the malate synthase from *Clamydomonas reinhardtii*. **Planta**. v.219, p. 325-331, 2004.

NAQSH E ZAHRA, S.; KHATTAK, N. A.; MIR, A. Comparative modeling and docking studies of p16ink4/cyclin D1/Rb pathway genes in lung cancer revealed functionally interactive residue of RB1 and its functional partner E2F1. **Theoretical biology & medical modelling**, v. 10, n. 1, p. 1, 2013.

NORWELL, J.C. AND MACHALEK, A.Z. Structural genomics programs at the US National Institute of General Medical Sciences. **NAT. STRUCT. BIOL.** v. 7, n. 931, 2000.

OWENS, J. Determining druggability. **Nature Reviews Drug Discovery**, v. 6, n. 3, p. 187-187, 2007.

PALMEIRO M, CHERUBINI K, YURGEL LS. Paracoccidioidomycose: revisão da literatura. **Scientia Med**. v. 15, n. 4, p. 234-7, 2005.

PEREIRA R.M., BUCARETCHI F., BARISON E.D.E.M., HESSEL G., TRESOLDI A.T. Paracoccidioidomycosis in children: clinical presentation, follow-up and outcome. **Rev Inst Med Trop Sao Paulo**. v. 46, n. 3, p. 127-31, 2004.

PETSALAKI E., STARK A., GARCIA-URDIALES E., RUSSELL R.B. Accurate Prediction of Peptide Binding Sites on Protein Surfaces. **Plos Computational Biology**. v. 5, 2009.

PETSKO, G.A.; RINGE, D. Protein Structure and Function. **New Science Press**. p.195, 2004.

PIERCE B.G., HOURAI Y., WENG Z. .Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. **PLoS ONE** v. 6, n. 9, 2011.

PRADO, M. *et al*. Mortality due to systemic mycoses as a primary cause of death or in association with AIDS in Brazil: A review from 1996 to 2006. **Memorias do Instituto Oswaldo Cruz**, v. 104, n. May, p. 513–521, 2009.

PRONK S., PÁLL S., SCHULZ R. *et al*. 2013. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. **Bioinformatics** v. 29, n. 7, p. 845-854, 2013.

RAMOS E.S.M., SARAIVA L.D.O. Paracoccidioidomycosis. **Dermatol Clin**.v. 26, p. 257-69, 2008.

RUSS, A. P.; LAMPEL, S. The druggable genome: an update. **Drug discovery today**, v. 10, n. 23-24, p. 1607-10, 2005.

ŠALI, A.; BLUNDELL, T.L. Comparative protein modeling by satisfaction of spatial restraints. **Journal of Molecular Biology**, v.234, p.779-815, 1993.

SANTOS FILHO, O. A.; BICCA DE ALENCASTRO, R. Modelagem de proteínas por homologia. **Química Nova**, v. 26, n. 2, p. 253–259, 2003.

SELITRENNIKOFF, C.P. E NAKATA, M. New cell wall for antifungal drugs. **Current Opinion In Investigational Drugs**. v. 4, p. 200-205, 2003.

SHAMEER et al. PeptideMine - A webserver for the design of peptides for protein-peptide binding studies derived from protein-protein interactomes. **BMC Bioinformatics**. v. 11, 2010.

SHIKANAI-YASUDA M.A., TELLES FILHO F., MENDES R.P., COLOMBO A.L., MORETTI M.L. Guidelines in paracoccidioidomycosis. **Rev Soc Bras Med Trop**. v. 9, p. 297-310, 2006.

SKRABANEK, L.; SAINI, H.K.; BADER, G.D.; ENRIGTH, A.J. Computational prediction of protein-protein interactions. **Mol Biotechnol**. v. 38, n.1, p.1-17, 2008.

STURME M.H.J., PUCCIA R., GOLDMAN G.H., RODRIGUES F. Molecular biology of the dimorphic fungi *Paracoccidioides spp*. **Fungal Biology Reviews**, v. 25, p. 89-97, 2011.

SZKLARCZYK D, FRANCESCHINI A, WYDER S, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. **Nucleic Acids Research**. 2015;43(Database issue):D447-D452. doi:10.1093/nar/gku1003.

TAVARES, A.H., SILVA, S.S., BERNARDES, V.V., MARANHÃO, A.Q., KYAW, C.M., POÇAS- FONSECA, M. and SILVA-PEREIRA, I. Virulence insights from the *Paracoccidioides brasiliensis* transcriptome. **Genetics and Molecular Research**. v. 4, n. 2, p. 358-371, 2005.

TAVARES, J.; MESMOUDI, S.; TALBI, E.-G. On the efficiency of local search methods for the molecular docking problem. **Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics**, p. 104–115, 2009.

TEIXEIRA M.M., THEODORO R.C., CARVALHO M.J.A., FERNANDES L., PAES H.C., HAHN R.C., MENDOZA L., BAGAGLI E., SAN-BLAS G., FELIPE M.S. Phylogenetic analysis reveals a high level of speciation in the *Paracoccidioides* genus. **Molecular Phylogenetics and Evolution** v. 52, p. 273-283, 2009.

THÉVENET, P. et al. PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. **Nucleic Acids Research**, v. 40, n. W1, p. 288–293, 2012.

THAKUR P.K, PRAKASH A., KHAN P., FLEMING R. E., WAHEED A., AHMAD F., HASSAN I. Identification of Interfacial Residues Involved in Hepcidin-Ferroportin Interaction. **Letters in Drug Design & Discovery**. v.11, p. 363-374, 2014.

TOVCHIGRECHKO A., VAKSER I.A. GRAMM-X public web server for protein-protein docking. **Nucleic Acids Res**. v. 34, p. 310-314, 2006.

TROTT, O.; OLSON, A. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. 455–461, 2010.

UNAL EB, GURSOY A, ERMAN B. VitAL: Viterbi Algorithm for de novo Peptide Design. **PLoS ONE**.v. 5, n. 6, 2010.

VICENTINI A.P., GESZTESI J.L., FRANCO M.F., DE SOUZA W., DE MORAES J.Z., TRAVASSOS L.R., et al. Binding of *Paracoccidioides brasiliensis* to laminin through surface glycoprotein gp43 leads to enhancement of fungal pathogenesis. **Infect Immun**.v. 62, p. 1465-9, 1994.

VIEIRA G. DE D., ALVES T. DE C., LIMA S.M., CAMARGO L.M., SOUZA C.M. Paracoccidioidomycosis in a western Brazilian Amazon State: clinical-epidemiologic profile and spatial distribution of the disease. **Rev Soc Bras Med Trop**.v. 47, p. 63-8, 2014.

WALTERS, W. P. Going further than Lipinski's rule in drug design. **Expert opinion on drug discovery**, v. 7, n. 2, p. 99-107, 2012.

YANG, J. et al. The I-TASSER Suite: Protein structure and function prediction. **Nature Methods**, v. 12, n. 1, p. 7–8, 2015.

ZAMBUZZI-CARVALHO PF, CRUZ AH, SANTOS-SILVA LK, GOES AM, SOARES CMA, PEREIRA M. The malate synthase of *Paracoccidioides brasiliensis* Pb01 is required in the glyoxylate cycle and in the allantoin degradation pathway. **Med Mycol**. v. 47, p. 734–744, 2009.

ZANETTI, E. M.; ROSSI, P.; MORAES, P. C. Carne de tatu como potencial transmissor da paracoccidioidomicose. In: Congresso Brasileiro de Estomatologia, XIX, 2011, Recife. **International Journal of Dentistry**. Recife, v. 10, n. 2, p. 25-163, 2011.

ZHANG, M.Q. ; WILKINSON, B. Drug discovery beyond the “rule-of-five”. **Current opinion in biotechnology**, v. 18, n. 6, p. 478-88, 2007.

ZHANG, J. et al. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. **Structure**. v. 19, p. 1784–1795, 2011.

ZHU X., and MITCHELL J.C. KFC2: A knowledge-based hot spot prediction method based on interface solvation, atomic density and plasticity features. **Proteins**. v. 79, n. 9, p. 2671-2683, 2011.

ANEXO 1 – Alguns peptídeos usados na terapêutica atualmente

Nome	Ação farmacológica	Emprego terapêutico	Via de adm.
Adalimumab (Humira®)	Anticorpo monoclonal inibidor do fator de necrose tumoral (TNF)	Doenças auto-imunes	Subcutânea
Infliximabe (Remicade®)	Anticorpo monoclonal inibidor de TNF α	Doenças auto-imunes	Intravenosa
Rituximab (Rituxan®)	Anticorpo citolítico para proteína CD20	Linfoma não Hodgkin's, granulomatose de Wegener (WG), poliangite microscópica (MPA) e leucemia linfocítica crônica de células B.	Intravenosa
Bevacizumabe (Avastin®)	Anticorpo monoclonal que neutraliza a atividade biológica do fator de crescimento do endotélio vascular humano (VEGF)	Antineoplásico	Intravenosa
Trastuzumabe (Herceptin®)	Anticorpo monoclonal que atinge o domínio extracelular da proteína do receptor-2 do fator de crescimento epidérmico humano (HER2)	Câncer de mama metastático	Intravenosa
Insulina (Lantus®)	Equipotente à insulina humana	Diabetes I e II	Subcutânea
Pegfilgrastim (Neulasta®)	Estimulador de glóbulos brancos	Mielossupressão	Subcutânea
Epoetina alfa (Épogen®)	Glicoproteína estimuladora de eritropoiese	Anemia	Intravenosa, Subcutânea
Interferon beta-1a (Avonex®)	Indutor de proteínas específicas, as quais ainda não foram claramente definidas	Esclerose múltipla	Intramuscular
Acetato de glatirâmer (Copaxone®)	Alteração de funções imunológicas por mecanismos ainda não esclarecidos	Esclerose múltipla	Subcutânea
Goserelina (Zoladex®)	Supressão do sistema pituitário-gonadal	Câncer de mama e próstata, endometriose	Subcutânea
Acetato de Octreotida (Sandostatin®)	Derivado sintético da somatostatina, inibe a liberação do hormônio de crescimento, do glucagon e da insulina	Acromegalia e câncer	Intramuscular
Enfuvirtida (Fuzeon®)	Inibidor do rearranjo estrutural da gp41 do vírus HIV, bloqueando a entrada do vírus na célula	Infecção por HIV	Subcutânea

APÊNDICE 1

1. Características gerais das proteínas ligantes de MLS

1.1. Enolase (EC 4.2.1.11)

Pertencente à família das liases, a enolase de *Paracoccidioides sp.* possui 404 resíduos de aminoácidos e entre 40 e 50kDa. Está envolvida no processo de degradação do RNA, fazendo parte de um complexo degradassomo composto pela RNA helicase e a PNPase (software KEGG). De acordo com o banco de dados NCBI, a enolase é uma enzima homodimérica que catalisa a desidratação de 2- fosfo-D-glicerato em fosfoenolpiruvato, participando das vias glicolítica e gliconeogênica, com participação de um co-fator, geralmente Mg^{2+} .

Estudos recentes de Donofrio e colaboradores (2009) e Nogueira et al., 2010, mostraram que na presença da enolase recombinante, o fungo aumentou sua capacidade de associação com células epiteliais e fagocíticas. Por imunofluorescência, foi detectada no citoplasma e na superfície da célula fúngica, e acredita-se que ela se ligue a fibronectina da MEC. Além disso, a enolase foi relacionada a conversão de plasminogênio em plasmina, sendo essencial para a ligação de plasminogênio na superfície fúngica.

O modelo representativo de código PDB 1P43 de *S. cerevisiae*, apresentou *query cover* de 98%, identidade de 61%. De acordo com servidor MolProbity, 96% dos resíduos se encontram em regiões favoráveis antes da DM. Além disso, temos um *score* da estrutura de 2.48 , com um *Clashscore* de 4.81 antes da DM. Após a DM, a porcentagem de resíduos em regiões favoráveis foi aumentada para 99,3% , e os *scores* também foram melhorados (*score* 1.50 e *clashscore* de 0).

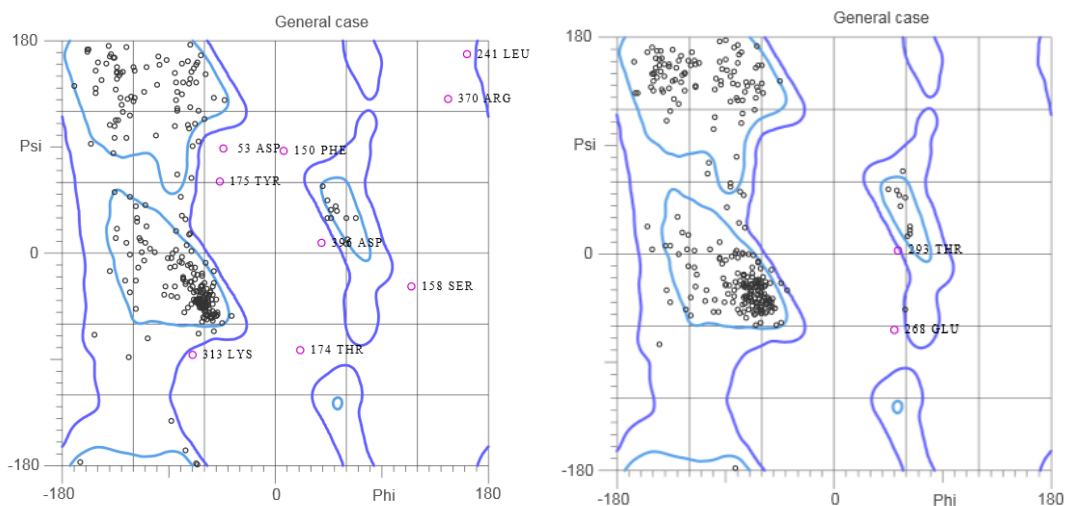


Figura 1-A. Diagrama de Ramachandran da Enolase. Em A, temos o diagrama antes da simulação de DM, com 96% dos resíduos em regiões favoráveis. E, em B temos a análise após a DM, com 99,3% dos resíduos em regiões favoráveis, evidenciando uma melhora na qualidade da estrutura desta proteína com a DM.

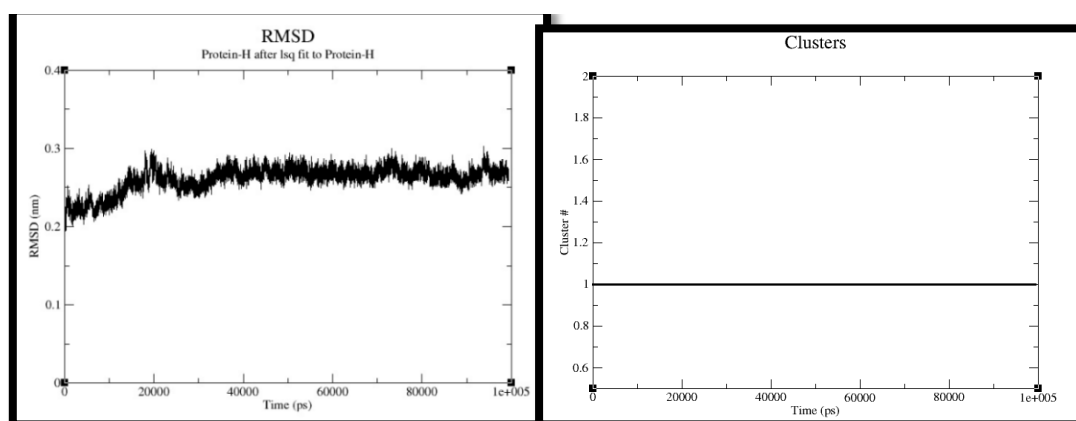


Figura 2-A. Gráficos de RMSD e clusters da Enolase. Observa-se que a proteína antigui “estabilidade” estrutural a partir dos 40 ns, aproximadamente. De acordo com o algoritmo g_cluster do pacote GROMACS, a proteína Enolase apresentou apenas 1 conjunto de conformações estruturais (cluster) muito próximas, durante a simulação de DM de 100 ns, que foi o modelo selecionado para as etapas *molecular docking*.

1.2. Ubiquitina (PAAG_07080)

A ubiquitina é uma importante proteína para degradação seletiva, pois “marca” proteínas que deverão ser destruídas pelo complexo proteossoma, em resposta a um estresse celular. O modelo utilizado para a modelagem 3D foi o de ID 4KSL_D (156 resíduos) de *Homo sapiens*, com *e-value* de 6e-100, identidade de 96% e *query cover* de 99%.

O *Molprobability score* teve uma melhora de 3,14 (antes da DM) para 1,17 (após a DM) e o *clashscore* de 9,45 (antes da DM) para 0 (após a DM). A seguir, temos os mapas de Ramachandran evidenciando as regiões que estão localizados os resíduos.

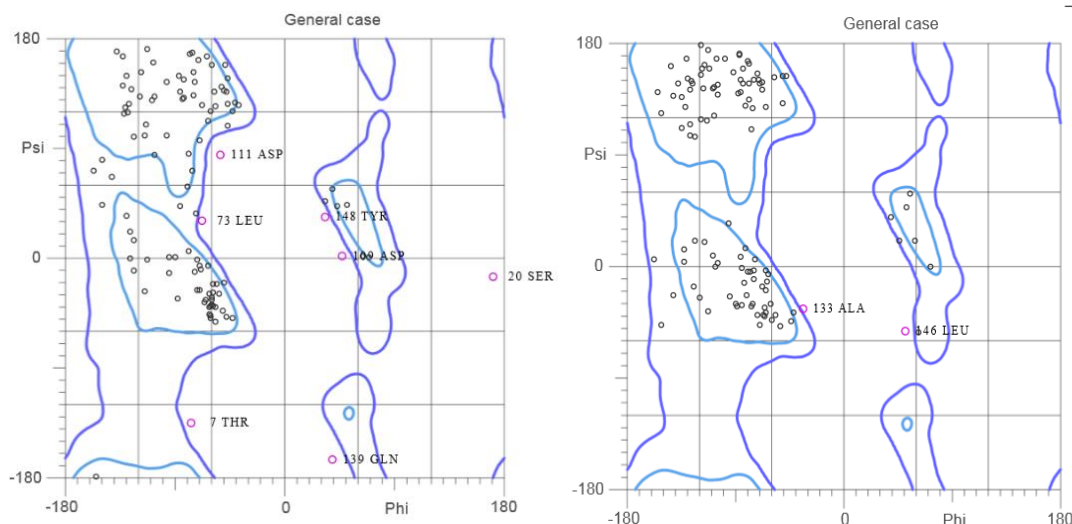


Figura 3-A. Diagramas de Ramachandran da Ubiquitina. Em A temos o mapa de localização dos resíduos, antes da DM 90% dos resíduos se localizavam em regiões favoráveis. Já em B (após a DM) 98,7% dos resíduos estão em regiões favoráveis.

Ubiquitina, ao passar pela simulação de DM, apresentou 16 modelos conformacionais, evidenciados na figura 3- A. Pode-se observar que o modelo 1 se manteve por mais tempo, porém nos momentos finais da simulação os modelos 2 e 8 também estiveram presentes, semelhante ao ocorrido com a Malato sintase.

No gráfico de RMSD (Figura 4-A) podemos observar uma maior estabilidade no desvio quadrático médio a partir de 25 ns aproximadamente, com uma pequena flutuação nos momentos finais, coincidindo com os dados de clusters. Deste modo, para a realização das próximas etapas, os modelos 1, 2 e 8 foram selecionados.

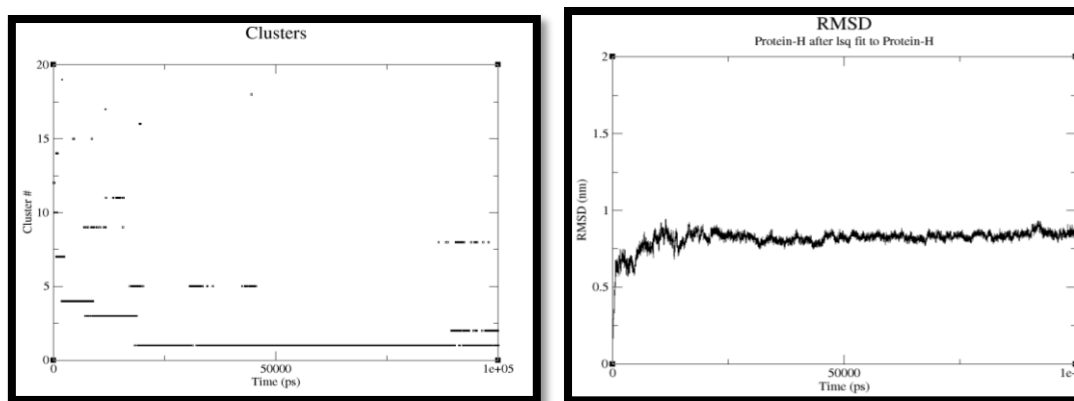


Figura 4-A. Gráficos de Clusters e RMSD da Ubiquitina. O cluster 1 se manteve bastante estável durante a DM, porém foram selecionados outros 2 clusters para as etapas de ancoragem, que estiveram presentes ao final da simulação, a fim de aumentar a variabilidade de complexos.

1.3. Triosefosfato isomerase (PAAG_02585 EC- 5.3.1.1)

É uma enzima homodimérica glicolítica, com 249 resíduos, presente em bactérias e eucariotos, que catalisa a síntese de gliceraldeído-3 fosfato a partir de glicerona fosfato, na via oxidativa das pentoses fosfato. Por ser uma reação muito eficiente, não necessita de co-fatores ou íons metálicos (NCBI; UniprotKB).

Além disso, já foi descrito por Pereira e colaboradores (2007) que esta é preferencialmente expressa na fase leveduriforme do *Pb01*, talvez pela sua função de adesina se ligando a fibronectina e laminina (importante no estabelecimento do patógeno). O modelo utilizado para a modelagem foi o de ID: 1YPI_A com query cover 99%, e-value de 0 e identidade 100%, do organismo *S. cerevisiae*.

Os valores de *Molprobit score* passaram de 2,32 para 1,36 e de *clashscore* de 3,1 para 0, após o procedimento de DM. Os resíduos que estavam em regiões favoráveis representavam 96,9% antes da DM. Após a simulação de DM essa porcentagem aumentou para 98,2%, como pode-se observar nos diagramas da Figura 5-A.

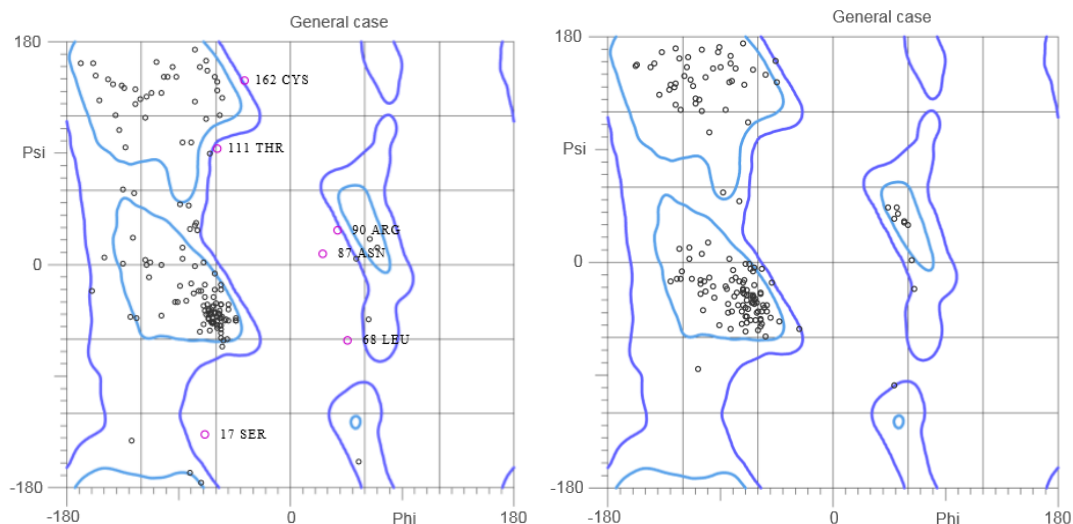


Figura 5-A. Ramachandran da triosefosfato isomerase. Em A, temos o diagrama antes da DM, e em B após a DM.

A proteína triosefosfato isomerase apresentou vários modelos estruturais ao longo da simulação de DM (18 clusters), alcançando maior estabilidade estrutural a partir de 50 ns, aproximadamente, como mostrado no gráfico de RMSD. Portanto, para as etapas seguintes foram selecionados os modelos 1 e 3, por serem os modelos de maior frequência ao longo da trajetória, principalmente a partir de 50 ns, como mostrado na figura 6-A.

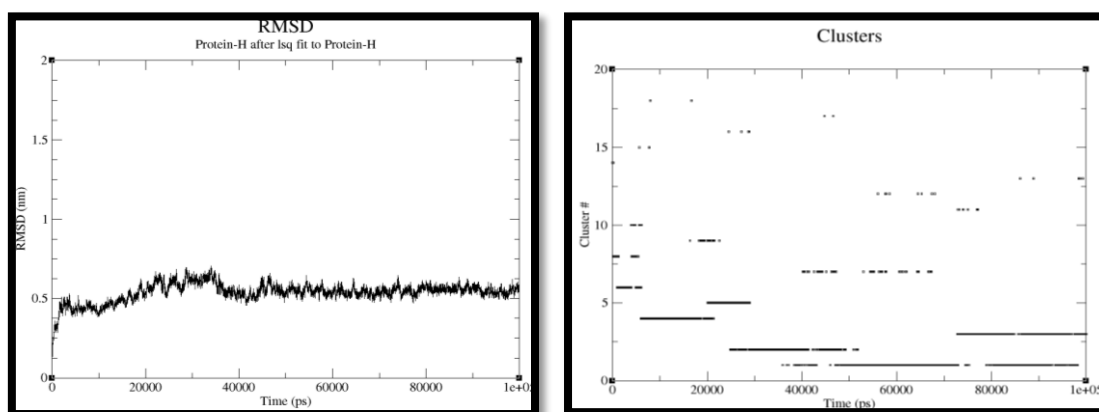


Figura 6-A. Gráficos de RMSD e Clusters da Triosefosfato Isomerase. De acordo com o gráfico de RMSD a proteína TFI alcançou certa estabilidade apenas em 50 ns aproximadamente, e, portanto, foram selecionados os clusters 1 e 3 para as etapas de ancoragem.

1.4. Gamma-actina

Gamma actina é uma proteína com várias funções estruturais e de motilidade, sendo expressa em todas as células eucarióticas. Está envolvida no processo de adesão plaquetária, na organização dos sarcômeros que são unidades repetidas de miofibrilas nas células musculares, e constituição do citoesqueleto (UniProtKB – ID:P63260).

A modelagem por homologia foi realizada utilizando como modelo principal a proteína cristalizada gamma-actina de *Bos taurus* (PDB ID: 3U4L_A) com 375 resíduos, que apresentou cobertura de 100%, identidade de 99%, e-value 0. O *molprobability score* observado foi de 2,13 para 1,12, e o *clashscore* de 2,9 para 0, antes e após a DM, respectivamente. Na figura 7-A, temos os diagramas de ramachandran que nos mostra a qualidade do modelo gerado.

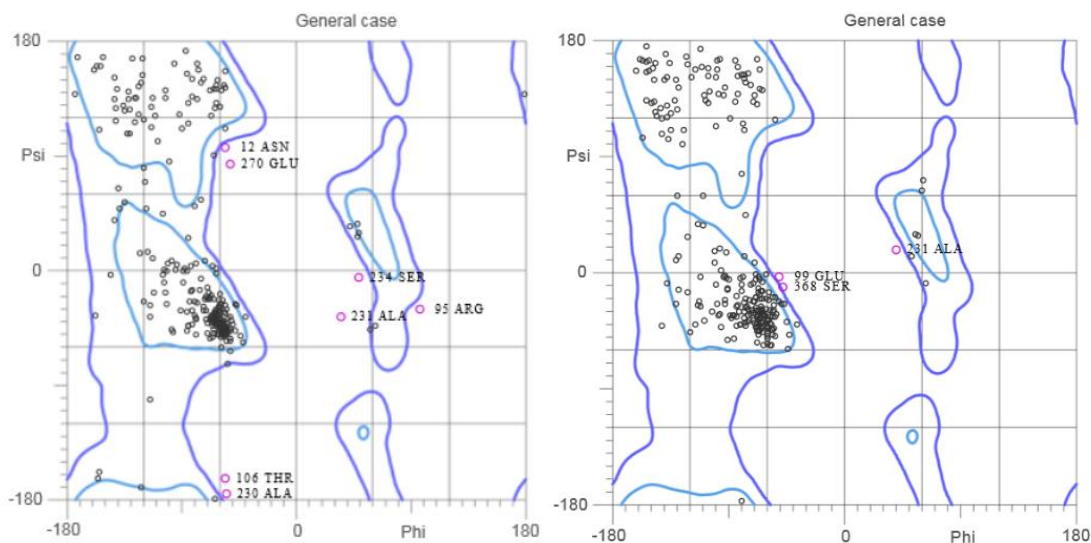


Figura 7-A. Diagramas de Ramachandran de Gama actina. Em A (antes DM), 96,5% dos resíduos em regiões favoráveis, e em B (após a DM), 98,9%.

De acordo com as análises pós DM, podemos inferir que esta proteína apresentou apenas um conjunto de conformações aproximadas, para um cut-off de 0,3 angstrom, o que coincide com a estrutura bastante estável no gráfico de RMSD, mostrados na Figura 8-A.

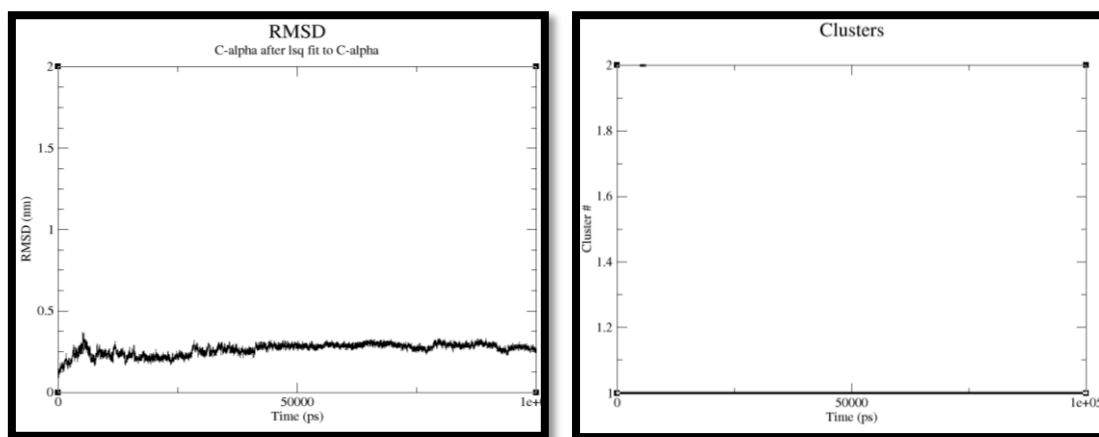


Figura 8-A. Gráficos de RMSD e Cluster da Gama actina. Semelhante ao observado para a proteína Enolase, a Gama actina apresentou apenas um cluster ao longo da DM, confirmado pela alta estabilidade no gráfico de RMSD.

1.5. Malato desidrogenase

Esta enzima com 340 resíduos, está envolvida com o ciclo do ácido cítrico e do glicoxalato, estando presente em algumas bactérias e eucariotos, sendo que nestes últimos podem existir com várias isoformas (citoplasmáticas, mitocondriais e peroxissomais). No ciclo do ácido cítrico, esta enzima dependente de NAD, facilita a conversão de malato em oxaloacetato, e reabastece os níveis de oxaloacetato por carboxilação redutiva de piruvato (NCBI).

O modelo principal utilizado para a modelagem foi o 2DFD_A (estrutura cristalizada da malato desidrogenase de humano), com query cover de 94%, e-value $3e-115$ e identidade 56%. De acordo com dados do banco PDB, o sítio de ligação à NAD corresponde à uma cavidade composta pelos resíduos: GLY17, ILE18, ASP39, PRO81, VAL84 e MET233 da cadeia A da proteína malato desidrogenase.

Molprobit score antes da DM estava em 3,12 e atingiu 1,27 após a simulação de DM. Já o *clashscore* melhorou de 12,21 para 0. Na Figura 9-A, temos os mapas de Ramachandran para a os modelos antes e após a DM, e na Figura 10-A estão os gráficos de RMSD e Clusters provenientes da DM.

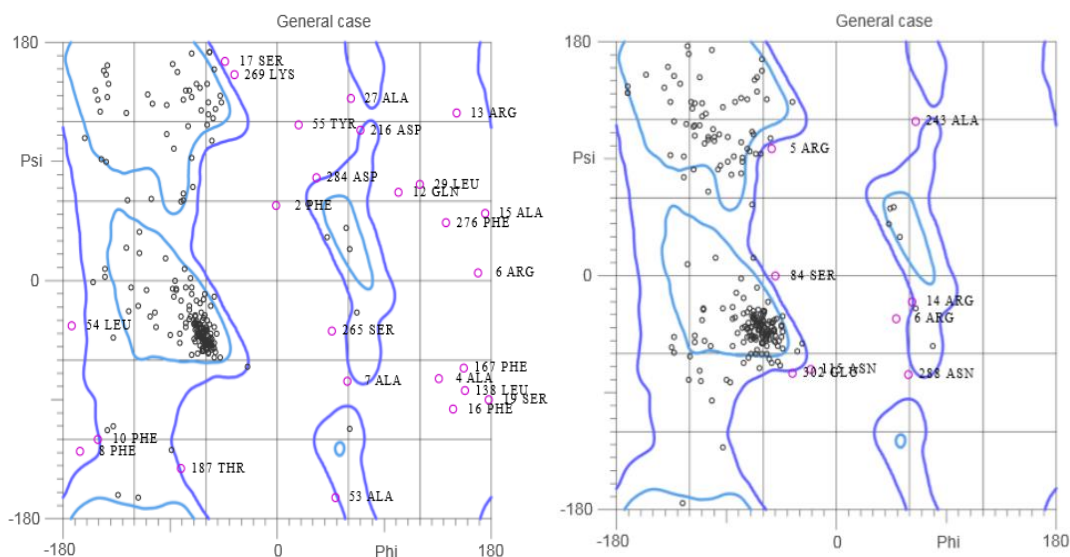


Figura 9-A Diagrama de Ramachandran da Malato desidrogenase. Em A, temos 88,5% dos resíduos em regiões favoráveis (antes da DM) e em B, temos 95% (após a DM).

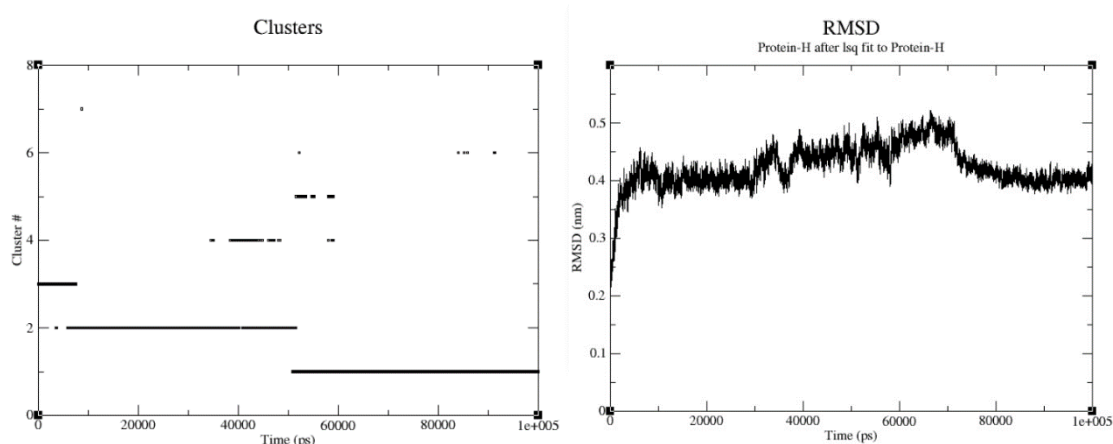


Figura 10-A. Gráficos de cluster e RMSD da Malato desidrogenase. Notam-se dois grandes grupos conformacionais ao longo da DM, e estes 2 foram selecionados para as etapas de ancoragem molecular.

1.6. Gliceraldeído-3-fosfato desidrogenase (EC- 1.2.1.12)

Enzima com 388 resíduos de aminoácidos, da família das oxidoredutases com 75kDa, responsável principalmente pela síntese piruvato a partir de D-gliceraldeído-3-fosfato da via glicolítica, dependente de NAD. O modelo cristalizado utilizado PDB (1U8F_O) derivado de placenta humana, apresentou cobertura de 97%, e-value de $3e-171$ e identidade de 71% de acordo com o software BLAST.

De acordo com Barbosa et al. (2006), é encontrada em grande quantidade na parede celular de *Paracoccidioides sp.*, provavelmente envolvida nos estágios iniciais da infecção (ligando a laminina, fibronectina e colágeno I). Dessa forma, acredita-se que diminuição ou inibição prejudique a infecção e adesão aos pneumócitos.

O *molprobability score* estava em 2,78 antes da DM, e após esse score diminuiu para 1,62. Já o *clashscore* que estava em 6,44 antes da DM, atingiu o valor de 0 após a DM, demonstrando a melhora na qualidade geral da estrutura. Na Figura 11-A, temos os diagramas de Ramachandran, antes e após a simulação de DM. Já na Figura 12-A, temos as análises de DM, com os gráficos de RMSD e clusters.

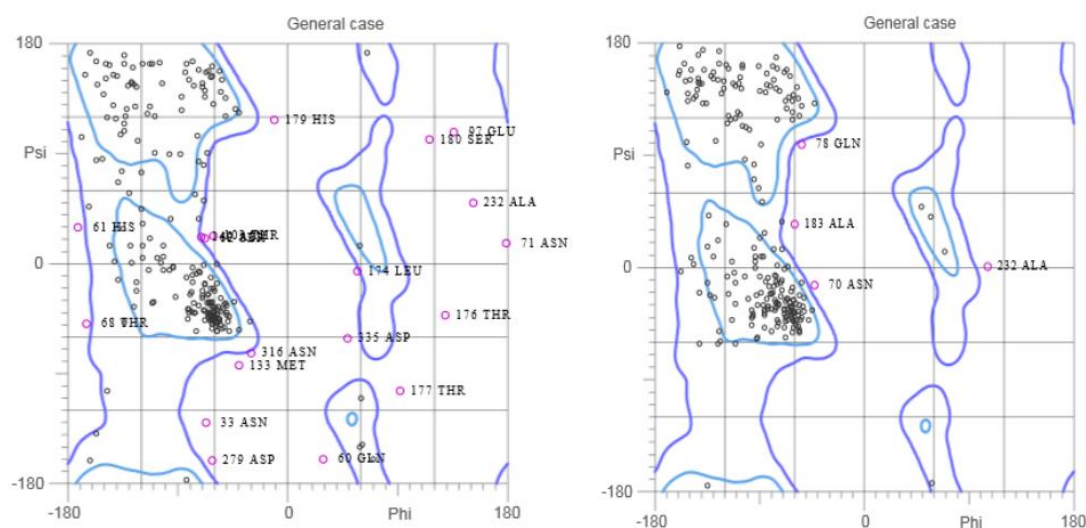


Figura 11-A. Diagramas de Ramachandran da Gliceraldeído-3-fosfato desidrogenase. Em A, temos 90% dos resíduos em regiões favoráveis (antes da DM). Em B, temos 97,3% dos resíduos em regiões favoráveis (após a DM).

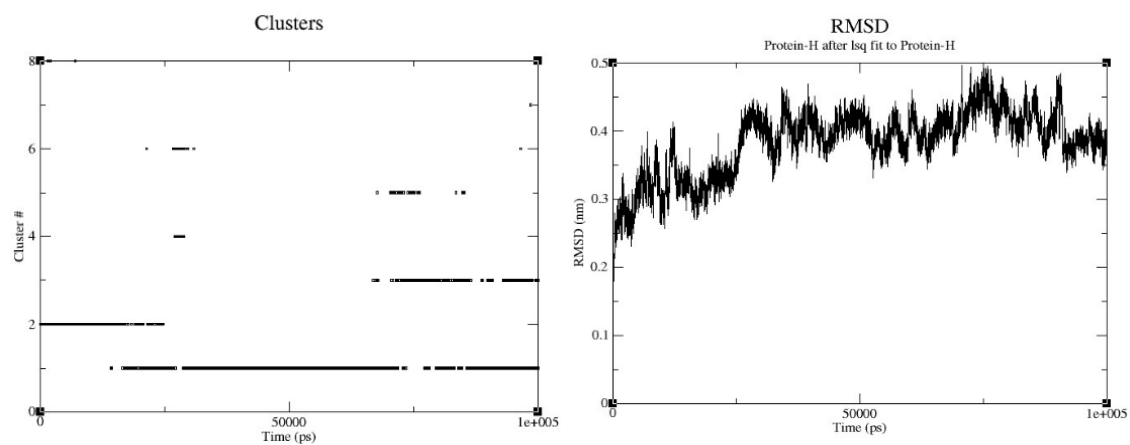


Figura 12-A. Gráficos de Cluster e RMSD da Gliceraldeído-3 fosfato desidrogenase. A estrutura tridimensional da proteína GPDH apresentou bastante instabilidade como observado no gráfico de RMSD, e os dois clusters mais representativos (1 e 3) foram selecionados para as etapas de ancoragem.