

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE CIÊNCIAS FARMACÊUTICAS  
Curso de Graduação em Farmácia-Bioquímica

**Construção de modelos tridimensionais de Sirtuína 2 de espécies  
de *Leishmania* por meio de modelagem comparativa**

**André Berndt Penteado**

Trabalho de Conclusão do Curso de  
Farmácia-Bioquímica da Faculdade de  
Ciências Farmacêuticas da  
Universidade de São Paulo.

Orientador(a):  
Prof. Dr. Gustavo Henrique Goulart  
Trossini

São Paulo

2020

## SUMÁRIO

	Pág.
Lista de Abreviaturas .....	1
RESUMO .....	2
1. INTRODUÇÃO .....	3
2. OBJETIVOS .....	9
3. MATERIAIS E MÉTODOS .....	9
4. RESULTADOS .....	11
5. DISCUSSÃO .....	24
6. CONCLUSÃO .....	29
7. BIBLIOGRAFIA .....	30

## LISTA DE ABREVIATURAS

DTN	Doenças Tropicais Negligenciadas
hSIRT	<i>Human Sirtuin</i>
I-TASSER	<i>Iterative Threading Assembly Refinement</i>
LiSir2rp1	<i>Leishmania infantum Sirtuin 2 related protein 1</i>
NCBI	<i>National Center for Biotechnology Information</i>
OMS	Organização Mundial da Saúde
PDB	<i>Protein Data Bank</i>
QMEAN	<i>Quantitative Model Energy Analysis</i>
RMSD	<i>Root Mean Square Deviation</i>
SAVES	<i>Structure Analysis and Verification server</i>
SBDD	<i>Structure Based Drug Design</i>
SBVS	<i>Structure Based Virtual Screening</i>
Sirtuin	<i>Silent Information Regulator (Sirtuína)</i>
Sir2	Sirtuína 2
Sir2rp1	<i>Sirtuin 2 related protein 1</i>
Sir2rp2	<i>Sirtuin 2 related protein 2</i>
Sir2rp3	<i>Sirtuin 2 related protein 3</i>

## RESUMO

Penteado, A. B. **Construção de modelos tridimensionais de Sirtuína 2 de espécies de Leishmania por meio de modelagem comparativa.** 2020. no. f. Trabalho de Conclusão de Curso de Farmácia-Bioquímica – Faculdade de Ciências Farmacêuticas – Universidade de São Paulo, São Paulo, 2020.

Palavras-chave: Modelagem Comparativa, Dinâmica Molecular, Sirtuína 2, Leishmania.

As leishmanioses são doenças negligenciadas que afetam uma grande parte da população mundial, principalmente em países em desenvolvimento. Os fármacos disponíveis para essas doenças são ineficientes e apresentam graves efeitos adversos. A sirtuína 2 é uma enzima essencial para a sobrevivência dos parasitas do gênero *Leishmania*. Assim sendo, se apresenta como um alvo na busca de novos fármacos contra essas doenças. O planejamento de fármacos baseado na estrutura do receptor exige o conhecimento da estrutura tridimensional da proteína alvo. Neste sentido, a elucidação das estruturas tridimensionais de sirtuínas 2 se apresenta como uma excelente ferramenta para a aplicação desta estratégia na busca de quimioterápicos contra *Leishmania*. Como apenas uma estrutura 3D de Sir2 de *Leishmania* é conhecida, a aplicação da modelagem comparativa é uma alternativa para o entendimento dos pontos de interação e seletividade destas hidrolases parasitárias. Os estudos envolvidos neste projeto nos fornecerão embasamento teórico para a busca de candidatos a fármacos leishmanicidas mais seguros e eficazes.

## 1. INTRODUÇÃO

As doenças tropicais negligenciadas (DTN) caracterizam-se por um grupo de diversas patologias que, apesar de provocarem enfermidades de grande relevância médica, social e econômica, não recebem a devida atenção e investimento de governos, indústrias farmacêuticas e grupos de pesquisa públicos e privados. Estas acometem, majoritariamente, as populações mais carentes de países em desenvolvimento como da África, Ásia e Américas Central e do Sul (OMSa, 2020).

Atualmente, a Organização Mundial da Saúde (OMS) reconhece 20 doenças como sendo DTNs, as quais representam cerca de 11,4% das doenças mundiais, afetando mais de um bilhão de pessoas em 149 países e causam de 500 mil a 1 milhão de óbitos anualmente. Algumas delas são tidas como extremamente negligenciadas, pois acometem milhões de pessoas no mundo e recebem quase nenhum investimento em inovação, como, por exemplo, a Leishmaniose Visceral, que possui alta prevalência no Brasil (OMSa, 2020).

As Leishmanioses são endemias infecciosas, consideradas doenças tropicais negligenciadas, causadas por mais de 20 espécies de protozoários do gênero *Leishmania*. Os parasitas, agentes causadores destas doenças, são transmitidos majoritariamente por insetos hematófagos vetores da família *Phlebotominae*, pertencentes ao gênero *Phlebotomus* e conhecidos como “sand flies”, no velho mundo, e ao gênero *Lutzomyia*, nas Américas, sendo popularmente conhecidos como mosquito-palha, birigui, cangalha ou tatuquira (CLABORN, 2010; OMSb, 2020).

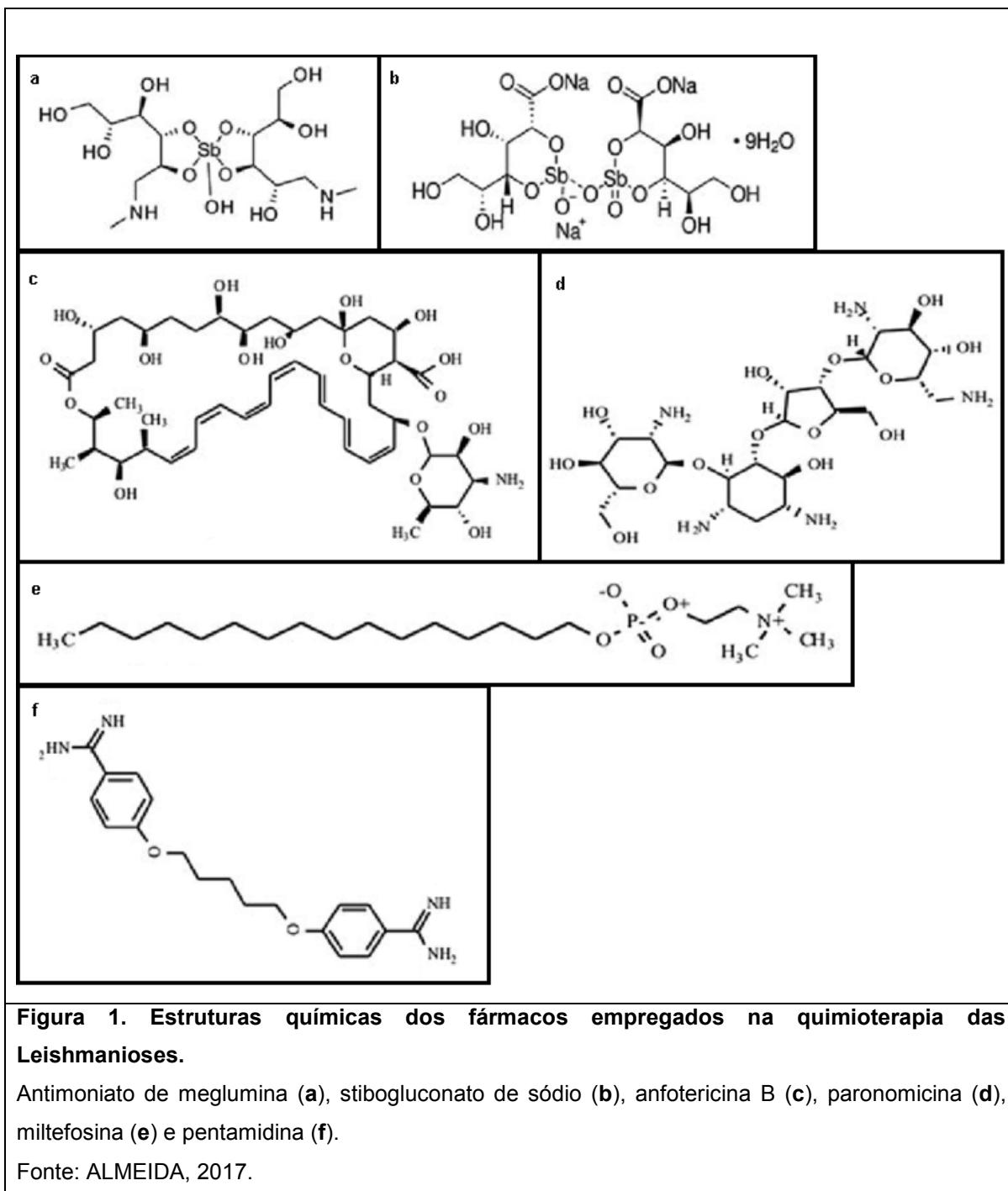
As infecções por *Leishmania* se manifestam em três formas principais, sendo elas: a cutânea, a visceral, chamada kala-azar e a mucocutânea. Todas as três formas da doença provocam sequelas irreparáveis e podem levar a morbidades desabilitantes, com impactos sócio-econômicos gravíssimos. A forma visceral é a mais grave, podendo levar a caquexia profunda, anemia, hipergamaglobulinemia, processos hemorrágicos e infecções secundárias, devido à indução de trombocitopenia e leucopenia pelo parasita. Caso não tratada, sua

taxa de mortalidade pode chegar a 100% em 2 anos. (MURRAY, 2005; OMSb, 2020; OMSc, 2020).

Afetando 88 países, apresenta 1,6 milhões de novos casos e provoca 65 mil mortes anualmente. Segundo dados epidemiológicos, estima-se que 12 milhões de pessoas estejam infectadas e que aproximadamente 350 milhões de pessoas encontram-se sob risco de infecção. No brasil existem 97 milhões de pessoas sob risco de infecção, sendo a incidência de 13 em 100 mil e 12690 novos casos foram relatados em 2016, sendo 94% da forma cutânea (GEORGIADOU, 2015; OMSd, 2020).

Para o tratamento das Leishmanioses apenas poucos agentes quimioterapêuticos estão disponíveis. Os mais empregados são os compostos de antimônio pentavalentes, como o antimoníato de meglumina e o stibogluconato de sódio, desenvolvidos na década de 1920, e a anfotericina B, cuja eficácia frente a parasitas deste gênero fora notada nos anos 50. Ambos são de uso parenteral e induzem efeitos adversos graves. Além disso, as taxas de falha do tratamento são altas, sendo relatados casos de resistência aos fármacos e consequentemente levando ao aumento das dosagens necessárias ao longo dos anos. Há algumas alternativas ao tratamento clássico, porém também apresentam efeitos adversos graves e eficácia limitada, como é o caso da miltefosina, de administração oral, e da paronomicina, que demonstra eficácia inferior aos compostos de antimônio pentavalente, porém tem se mostrado útil quando empregada em associação aos mesmos, possibilitando redução do tempo de tratamento. Há também a pentamidina, cujos efeitos adversos e baixa eficácia a levam a não ser empregada frequentemente, salvo em casos de resistência a primeira linha de tratamento. O tratamento também é caro e inacessível para a maior parte da população afetada. Até o momento, não existem vacinas eficazes para Leishmaniose, sendo a primeira linha de tratamento os compostos de antimônio pentavelente, administrados por via parenteral. O segundo agente mais empregado é a anfotericina B, que apesar de mais eficaz, apresenta toxicidade bastante elevada e as suas formulações mais modernas, que tentam mitigar tais efeitos colaterais, como preparações lipossômicas e PEGuiladas, apresentam custo proibitivo para

esta população. A combinação de dois agentes de mecanismos de ação distintos é normalmente empregada, visando aumentar a eficácia, reduzir o tempo de tratamento e a incidência de resistência aos agentes quimioterápicos (MURRAY, 2005; ULIANA, 2018; OMSe, 2020).

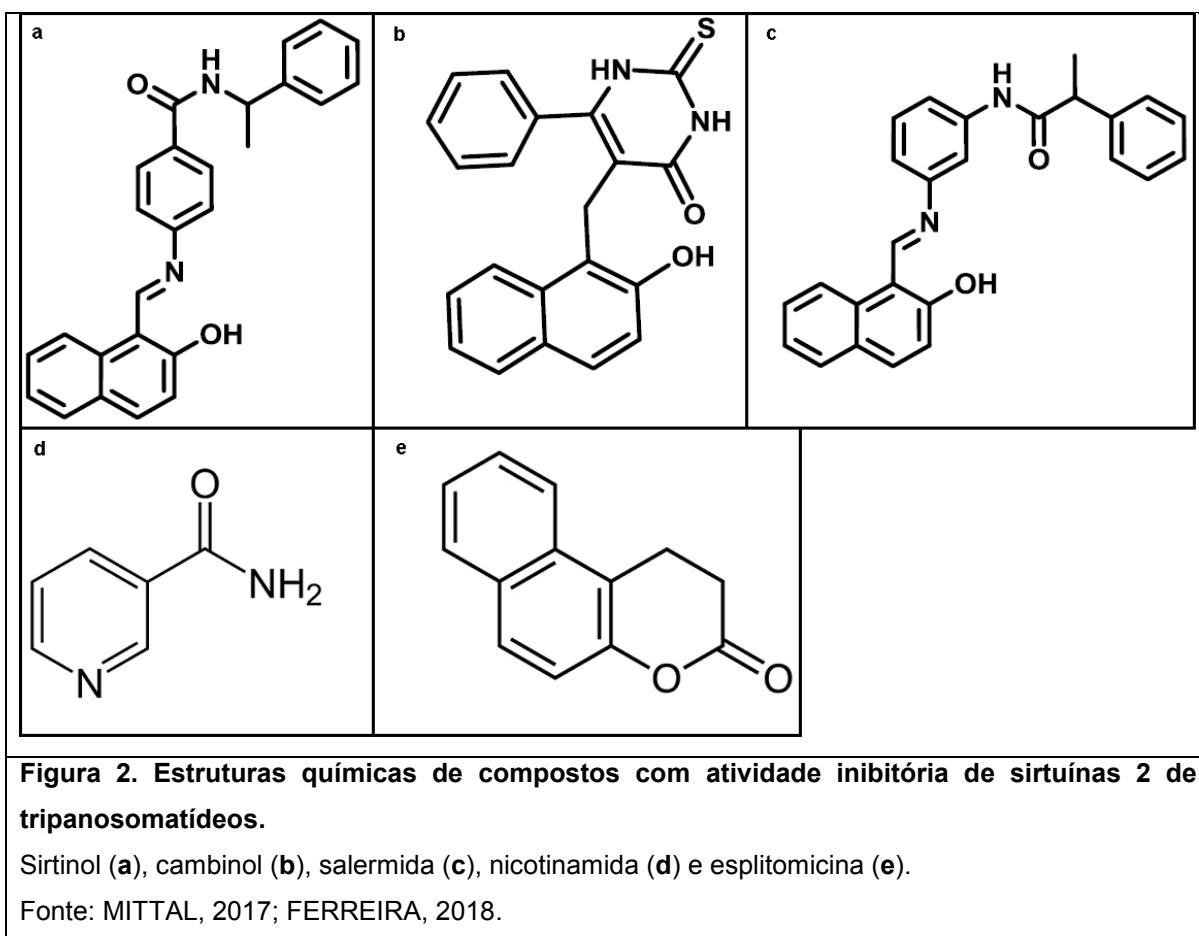


Neste contexto, a busca por novos agentes seguros e eficazes para combater protozoários do gênero *Leishmania*, com um perfil de tratamento mais simples e com um custo mais acessível, possibilitando um maior acesso e adesão da população ao mesmo é premente (FERREIRA, 2018; ULIANA, 2018).

A *Silent Information Regulator (Sirtuin)* é uma família de enzimas que apresentam alto nível de conservação estrutural e funcional entre todos os reinos da vida. São caracterizadas por sua dependência no cofator NAD<sup>+</sup> para a catálise de reações de deacetilação de resíduos de lisina de histonas e outros substratos peptídicos. As Sirtuínas são expressas por uma grande variedade de seres vivos, incluindo seres Humanos e espécies do gênero *Leishmania*. Em Humanos, diversas isoformas já foram caracterizadas (hSIRT1-7), enquanto que em Trypanosomatídeos foram descritos três grupos homólogos da Sirtuína 2 humana (hSIRT2): *Sirtuin 2 related protein 1* (Sir2rp1), *Sirtuin 2 related protein 2* (Sir2rp2) e *Sirtuin 2 related protein 3* (Sir2rp3). Enzimas desta família estão envolvidas em diversos processos fisiológicos celulares, incluindo a regulação da expressão de vários genes e a resposta a danos no DNA (MONIOT, 2013; RONIN, 2018).

Em espécies do gênero *Leishmania*, as Sirtuínas 2 se mostram essenciais para a sobrevivência do parasita, levando a morte do mesmo quando inibidas por algum agente farmacológico ou quando ausentes em estudos de *knockout*. Diante disto, se apresentam como alvo bioquímico na busca e desenvolvimento de novos fármacos contra estas parasitoses (VERGNES, 2005; MITTAL, 2017).

Além do mais, diferenças estruturais entre os homólogos Humanos e de espécies de *Leishmania* podem ser exploradas, na busca de compostos seletivos, o que pode resultar em fármacos mais seguros e com menor toxicidade. Alguns compostos com atividade inibitória de sirtuínas 2 de tripanosomatídeos já foram reportados, como o sirtinol, o cambinol, a salermida, a nicotinamida e a esplitomicina (MITTAL, 2017; FERREIRA, 2018).



A busca de novos fármacos pela abordagem de planejamento baseada na estrutura do receptor (do inglês *Structure Based Drug Design*, SBDD) se fundamenta no conhecimento estrutural detalhado da proteína alvo a nível tridimensional. Analisando a composição de sítios ativos e alostéricos de enzimas ou receptores em termos de quais resíduos de aminoácidos os compõem e o posicionamento espacial dos mesmos, torna-se possível desenhar de maneira racional compostos que apresentem características complementares a estes sítios, isto é, grupamentos químicos funcionais cujas propriedades eletrônicas e estéricas possibilitem interações favoráveis com tais resíduos de aminoácidos (ANDERSON, 2003).

A metodologia de triagem virtual baseada no alvo (do inglês *Structure Based Virtual Screening*) também requer conhecimento estrutural detalhado do alvo. Empregando algorítimos computacionais, modelos tridimensionais de

compostos ou fragmentos de compostos provenientes de algum banco de dados são virtualmente posicionados em alguma região eleita da estrutura alvo, como um possível sítio ativo de uma enzima. Os compostos são então avaliados e ranqueados de acordo com sua complementariedade com o alvo bioquímico. As interações estéricas e elétroestáticas são as principais informações para avaliar a afinidade do ligante por seu receptor. Deste modo, a otimização de um protótipo pode ser continuada à partir dos compostos que apresentam interações mais favoráveis (GLAAB, 2016).

Assim, a busca se torna muito mais eficiente, pois à partir de uma biblioteca contendo centenas, ou até milhares de compostos, é possível eleger algumas poucas dezenas de compostos promissores para a realização de ensaios biológicos, além da possibilidade de desenhar alguns compostos que provavelmente, segundo a estrutura do alvo, irão apresentar afinidade por este, o que propicia uma grande redução de custos e de tempo nesta primeira etapa do desenvolvimento (ANDERSON, 2003).

Porém, para que seja possível prever e avaliar a interação de compostos com o alvo escolhido, se faz necessário conhecimento estrutural extensivo do mesmo, a estrutura do sítio ativo, quais resíduos de aminoácidos estão envolvidos no reconhecimento do ligante e na catálise da reação, no caso de uma enzima, e o posicionamento destes no espaço. Desta forma, este método requer um modelo estrutural da proteína alvo, que pode ser obtido pela técnica de cristalografia de raios X, ressonância magnética nuclear, ou modelagem comparativa (FERREIRA, 2018).

A técnica computacional de modelagem comparativa, ou modelagem por homologia, se baseia no emprego de propriedades tridimensionais já elucidadas experimentalmente, provenientes de uma proteína usada como molde, na construção de modelos tridimensionais de outras proteínas à partir de suas sequências de aminoácidos, desde que as estruturas molde e modelada apresentem alto grau de identidade entre suas sequências primárias (VERLI, 2014; FERREIRA, 2018).

Considerando-se, que apenas uma Sirtuína 2 do gênero *Leishmania* teve sua estrutura elucidada após obtenção de dados de cristalografia de raios X até o momento (código PDB 5OL0), o presente trabalho visou a aplicação da metodologia de modelagem comparativa na construção de modelos tridimensionais de enzimas da família das Sirtuínas 2 de outras espécies do gênero. Essa abordagem tende a ser importante para contribuir no entendimento de suas estruturas tridimensionais, seu funcionamento bioquímico, seus sítios catalíticos e de ligação a substratos e cofatores, ao estudo das diferenças entre estas e possíveis pontos de interação para candidatos a novos fármacos (FERREIRA, 2018; RONIN, 2018).

## **2. OBJETIVO(S)**

Construção e validação de modelos tridimensionais de Sirtuínas 2 *related protein 1* (Sir2rp1) de *L. major*, *L. mexicana* e *L. braziliensis*, empregando-se metodologias e ferramentas de modelagem comparativa.

## **3. MATERIAIS E MÉTODOS**

### **3.1 Aquisição das sequências**

Inicialmente, todas as estruturas primárias de Sirtuínas 2 *related protein 1* (Sir2rp1) de *L. major*, *L. mexicana*, *L. braziliensis* e *L. Infantum* depositadas na plataforma NCBI para posterior análise e comparação (VERLI, 2014; FERREIRA, 2018).

### **3.2 Predição de localização intracelular**

As sequências foram analisadas pela ferramenta Phobius do *Stockholm Bioinformatics Centre* a fim de predizer sua localização subcelular e presença de peptídeo sinal (KÄLL, 2004).

### **3.3 Alinhamento entre as sequências**

Foi realizada a comparação, o alinhamento e a obtenção do grau de identidade das sequências de Sir2rp1 por meio da ferramenta *T-Coffee*, do *Centre for Genomic Regulation of Barcelona* (NOTREDAME, 2000; TOMMASO, 2011).

### 3.4 Construção dos modelos

Empregando o servidor *I-Tasser*, do *Zhang Lab* da *University of Michigan*, foram construídos os modelos tridimensionais de Sir2rp1 de *L. major*, *L. mexicana* e *L. braziliensis*, por meio da técnica de Predição de Enovelamento, ou *Threading*, tendo como molde a estrutura tridimensional de Sir2rp1 de *L. infantum* extraída da plataforma PDB (5OL0 cadeia A). As sequências escolhidas para serem modeladas foram, respectivamente, sp|Q25337.2, CBZ27844.1 e CBZ14728.1 (ZHANG, 2008; ROY, 2010; YANG, 2015).

### 3.5 Verificação de Estereoquímica dos Modelos

Os modelos gerados foram avaliados por meio de gráficos de Ramachandran (diagramas de ângulos phi e psi), distorção de grupos planares, desvio de ângulos ômega e análise de qualidade local dos resíduos de aminoácidos nos modelos, empregando para tal as ferramentas PROCHECK e WHATCHECK do servidor SAVES, do *Molecular Biology Institute* da *University of California* e QMEAN, do *Center for Molecular Life Sciences* da *University of Basel*, respectivamente (LASKOWSKI, 1993; HOOFT, 1996; LASKOWSKI, 1996; BENKERT, 2011).

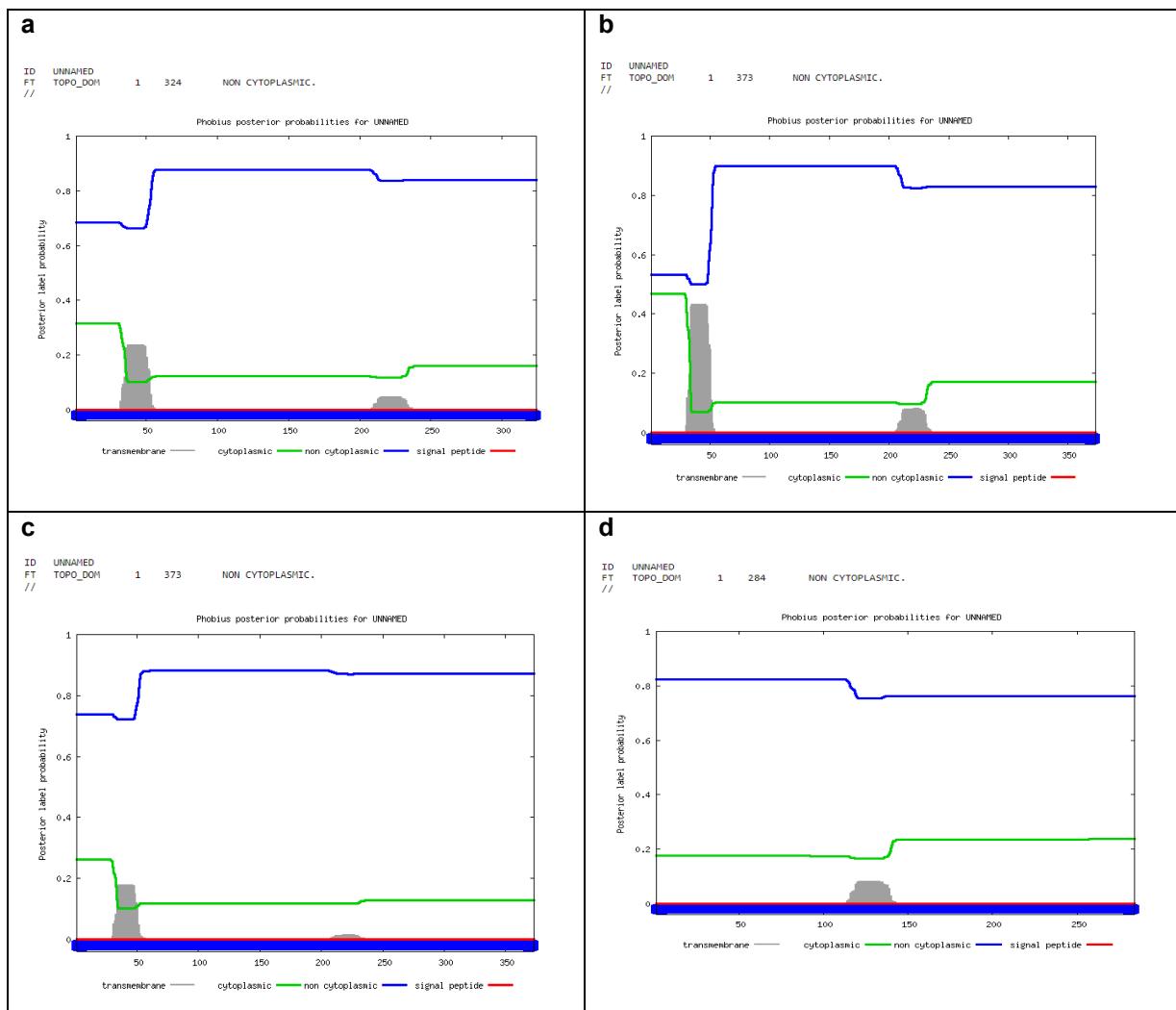
Estes dados foram correlacionados com resultados da ferramenta preditora de regiões de desordem, Remark465 do servidor DisEMBL, do *European Molecular Biology Laboratory* (LINDING, 2003).

Os modelos e as propriedades de seus resíduos foram visualizados com o auxílio dos programas PyMol do *Schrödinger Team* (*The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.*) e do BIOVIA *Discovery Studio Visualizer* (BIOVIA, D. S. 2015. *Discovery studio modeling environment. San Diego, Dassault Systèmes, Release, 4*).

## 4. RESULTADOS

### 4.1 Phobius - predição da localização intracelular e presença de peptídeo sinal

As sequências foram analisadas pela ferramenta Phobius do *Stockholm Bioinformatics Centre* afim de predizer sua localização intracelular e presença de peptídeo sinal. Observa-se que o perfil de localização intracelular predito pela ferramenta segue um padrão semelhante para todas as Sir2rp1 (Figuras 3a-d), e a ausência de sequência sinalizadora nestas enzimas. (KÄLL, 2004).



**Figura 3. Predição de localização intracelular e presença de peptídeo sinal em sequências de Sirtuínas 2 do gênero Leishmania.**

Em verde, probabilidade de localização citoplasmática ao longo da cadeia peptídica. Em azul, probabilidade de localização não citoplasmática ao longo da cadeia peptídica. Em cinza, probabilidade de regiões apresentarem localização transmembrânica. Em vermelho, probabilidade de uma região apresentar a função de peptídeo sinal.

(a) Sir2rp1 de *L. infantum* (PDB 5OL0 cadeia A), (b) Sir2rp1 de *L. major* (sp|Q25337.2), (c) Sir2rp1 de *L. mexicana* (CBZ27844.1) e (d) Sir2rp1 de *L. braziliensis* (CBZ14728.1).

Gráficos e análises obtidos através da ferramenta Phobius.

#### 4.2 *T-Coffee* - Alinhamento das sequências

Foi realizada a comparação, alinhamento e obtenção do grau de identidade das sequências de Sir2rp1 por meio da ferramenta *T-Coffee*, do *Centre for Genomic Regulation* of Barcelona. Estas apresentam alto grau de identidade entre si (de 90 a 99%, Figura 4). Nota-se também que a sequência de Sir2rp1 de *L. braziliensis* é mais curta do que as demais (aproximadamente 100 resíduos de aminoácidos a menos no início da sequência). (NOTREDAME, 2000; TOMMASO, 2011; LIMA, 2016).

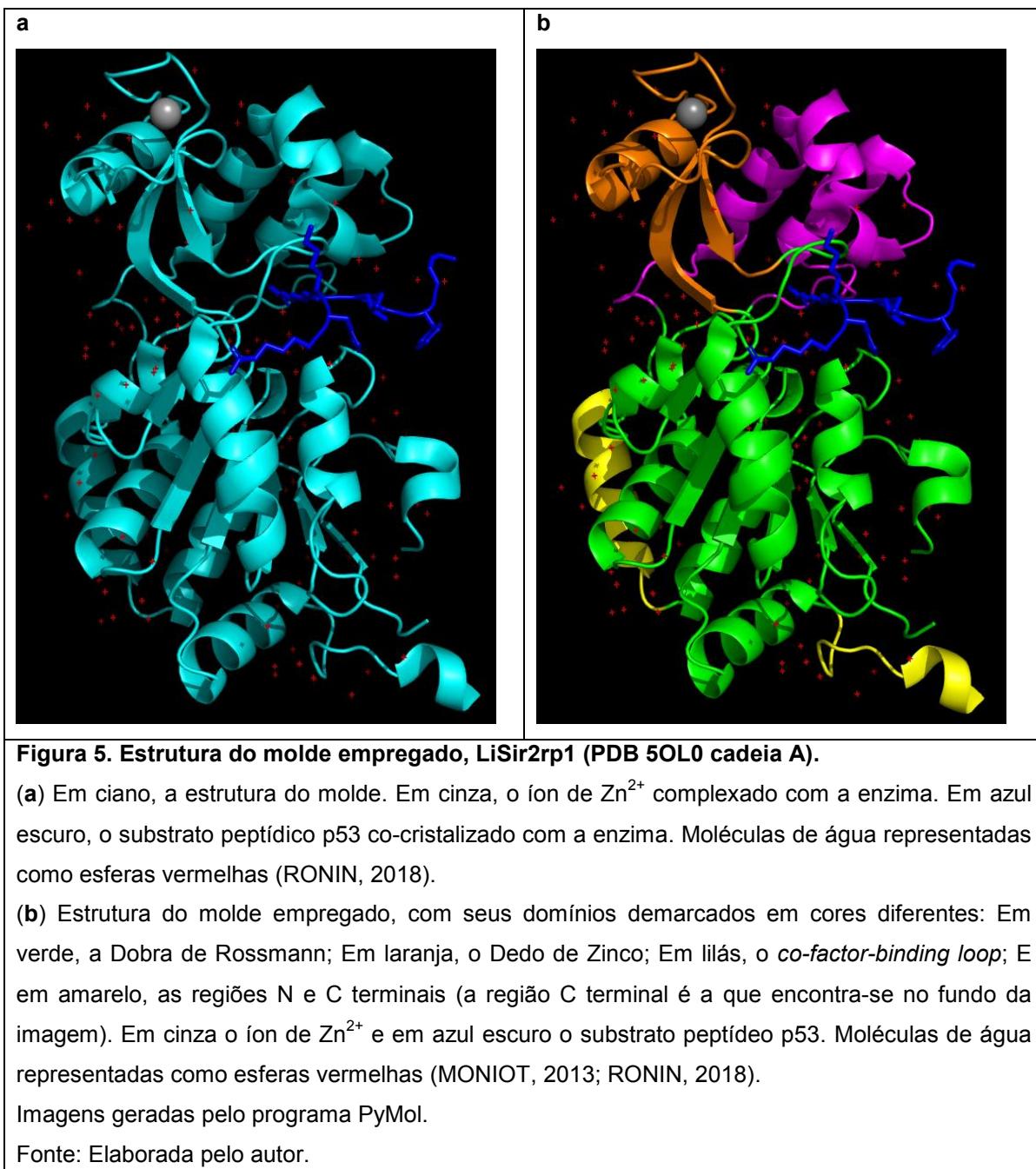


### 4.3 I-Tasser - Modelos gerados

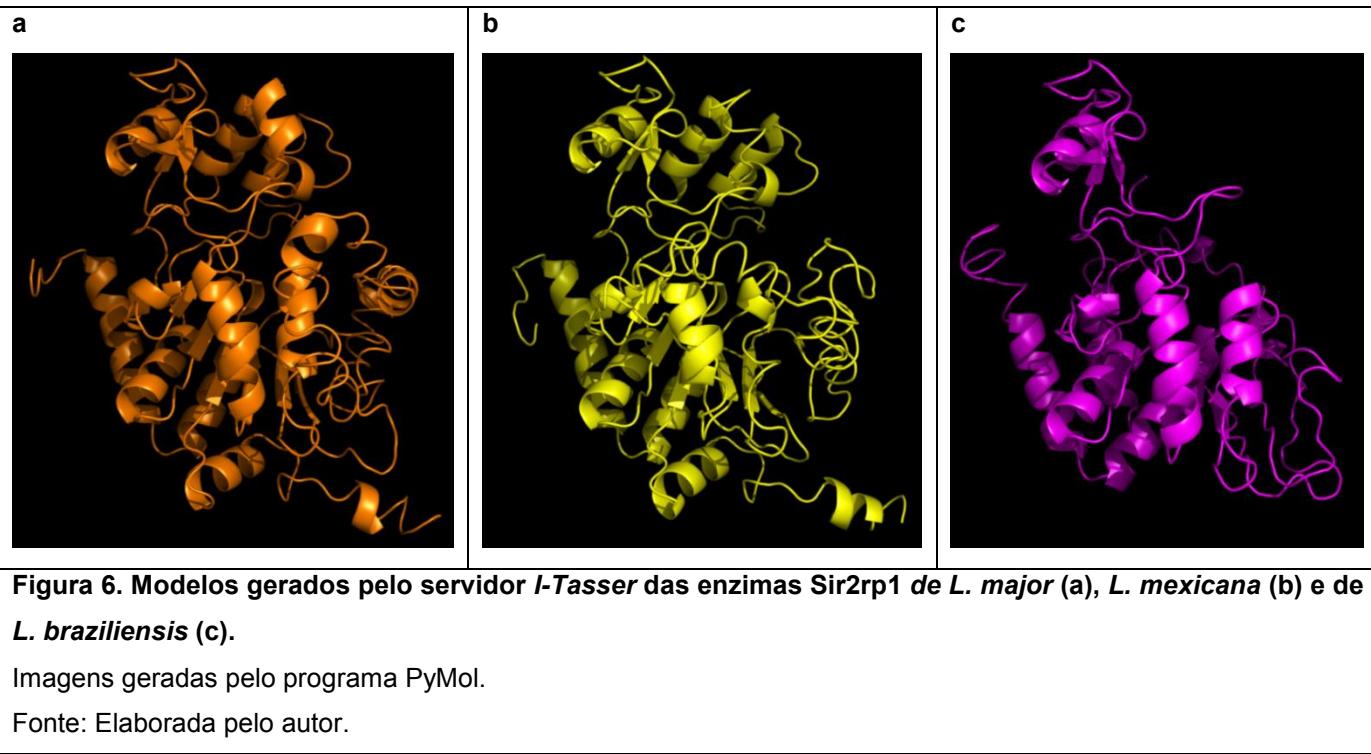
Empregando o servidor *I-Tasser*, do *Zhang Lab* da *University of Michigan*, foram construídos os modelos tridimensionais de Sir2rp1 de *L. major*, *L. mexicana* e *L. braziliensis*, por meio da técnica de predição de enovelamento, ou *Threading*, tendo como molde a estrutura tridimensional de Sir2rp1 de *L. infantum* extraída do banco de dados *Protein Data Bank* (5OL0 cadeia A). As sequências escolhidas para serem modeladas foram, respectivamente, spJQ25337.2, CBZ27844.1 e

CBZ14728.1, e seus valores de Erro Quadrático Médio (*Root Mean Square Deviation* ou RMSD) com relação ao molde foram de 0,428, 0,228 e 1,583 Angstroms, respectivamente. (ZHANG, 2008; ROY, 2010; YANG, 2015).

#### 4.3.1 Estrutura do molde empregado



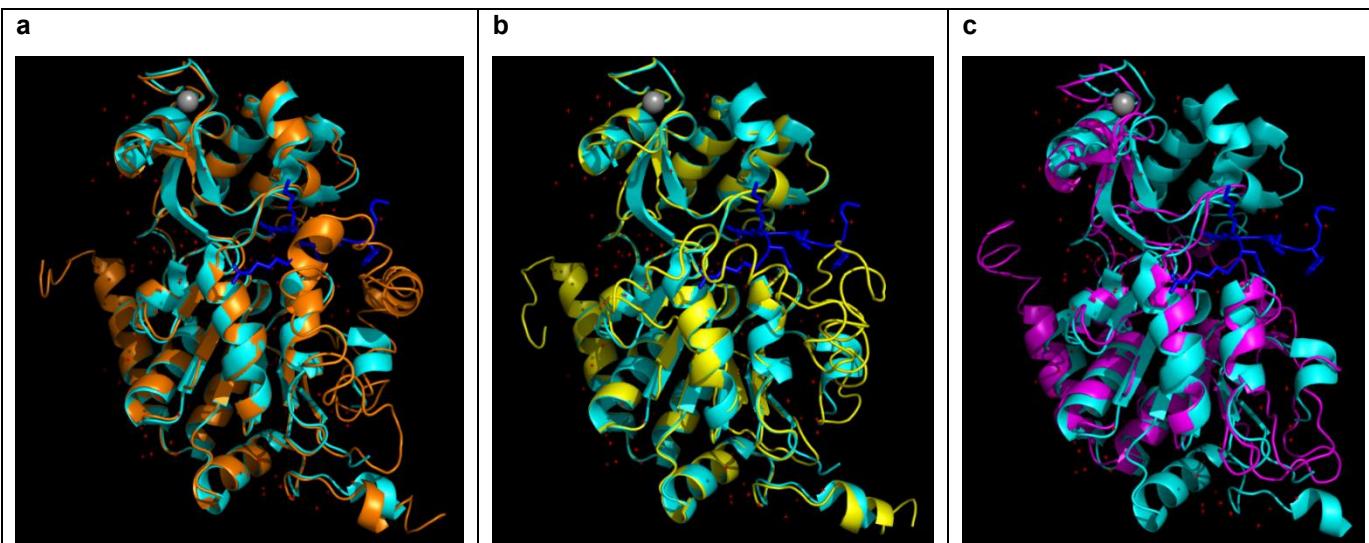
#### 4.3.2 Estrutura dos modelos gerados



#### 4.3.3 Sobreposição e alinhamento dos modelos com o molde

No software PyMOL, os modelos obtidos foram sobrepostos à estrutura cristalizada da LiSir2rp1 para cálculo de RMSD em relação ao molde estrutural.

O RMSD, ou erro quadrático médio, é um valor médio do desvio no posicionamento espacial dos átomos de uma determinada estrutura, ou entre os átomos de duas ou mais estruturas tridimensionais alinhadas e sobrepostas. Esta medida é muito empregada nos campos de cristalografia e modelagem, pois denota, no caso da cristalografia, o desvio médio de cada átomo de uma determinada estrutura, ou sua resolução, e para a modelagem, o quanto próximas duas estruturas tridimensionais são uma da outra. O valor arbitrário de RMSD < 2 Angstroms é comumente empregado como sendo indicativo de uma estrutura cristalográfica resolvida com alta resolução, ou como um alinhamento tridimensional favorável entre duas estruturas (COLE, 2005; HAWKINS, 2017).



**Figura 7. Alinhamento e sobreposição dos modelos gerados com o molde usado (em ciano).**

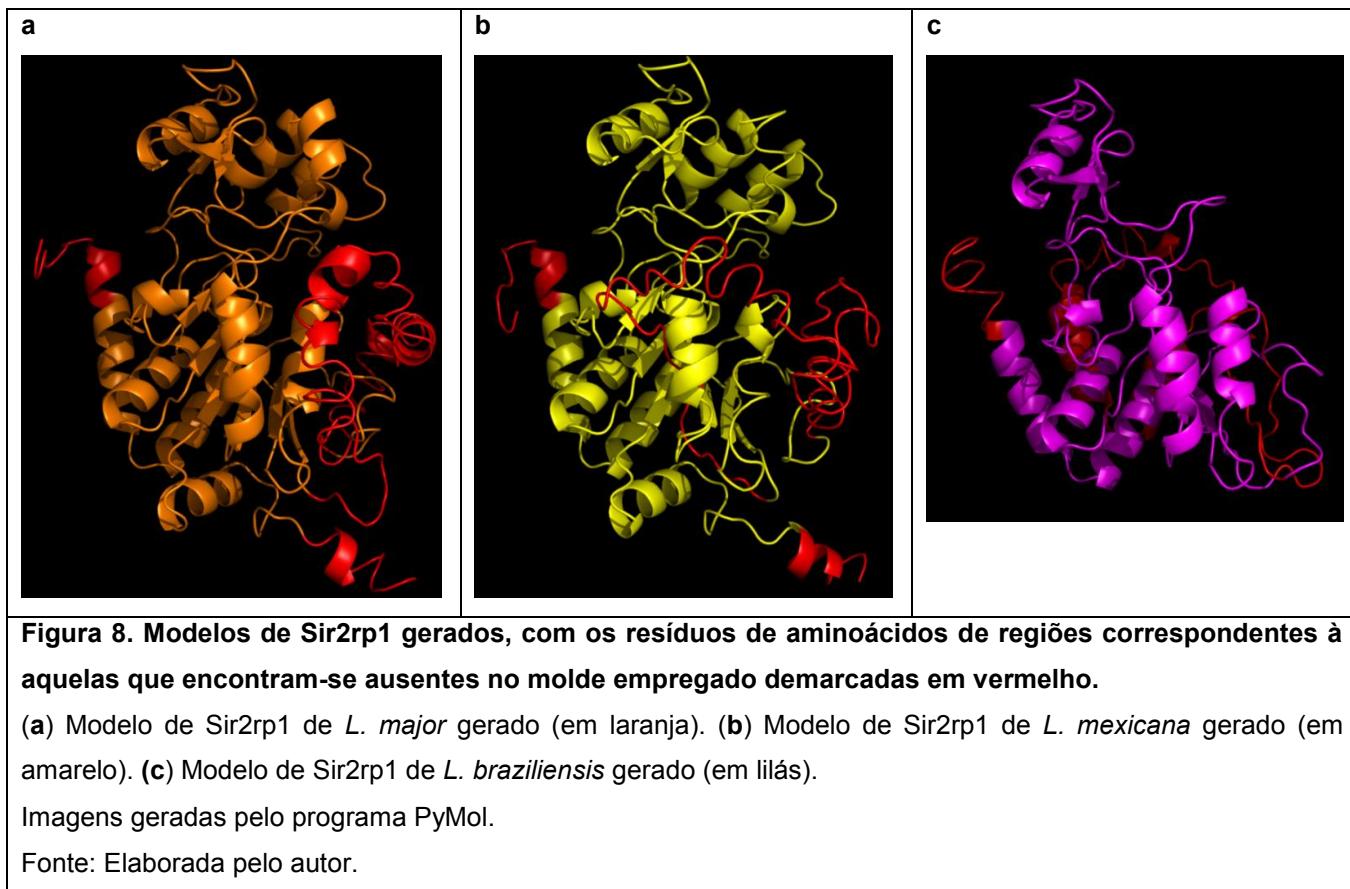
(a) Modelo de Sir2rp1 de *L. major* gerado (em laranja). RMSD = 0,428 Å. (b) Modelo de Sir2rp1 de *L. mexicana* gerado (em amarelo). RMSD = 0,228 Å. (c) Modelo de Sir2rp1 de *L. brasiliensis* gerado (em lilás). RMSD = 1,583 Å.

Alinhamentos tridimensionais e imagens obtidos através do programa PyMol.

Fonte: Elaborada pelo autor.

#### 4.3.4 Regiões dos modelos que encontram-se ausentes no molde

O molde apresenta as extremidades das porções N e C terminas e uma porção que ocorre no meio da sequência da enzima ausentes, em decorrência de um tratamento proteolítico breve que fora necessário para a obtenção de seu cristal, devido a natureza desordenada e altamente flexível destas regiões, o que impede a cristalização da proteína quando de sua presença (RONIN, 2018).



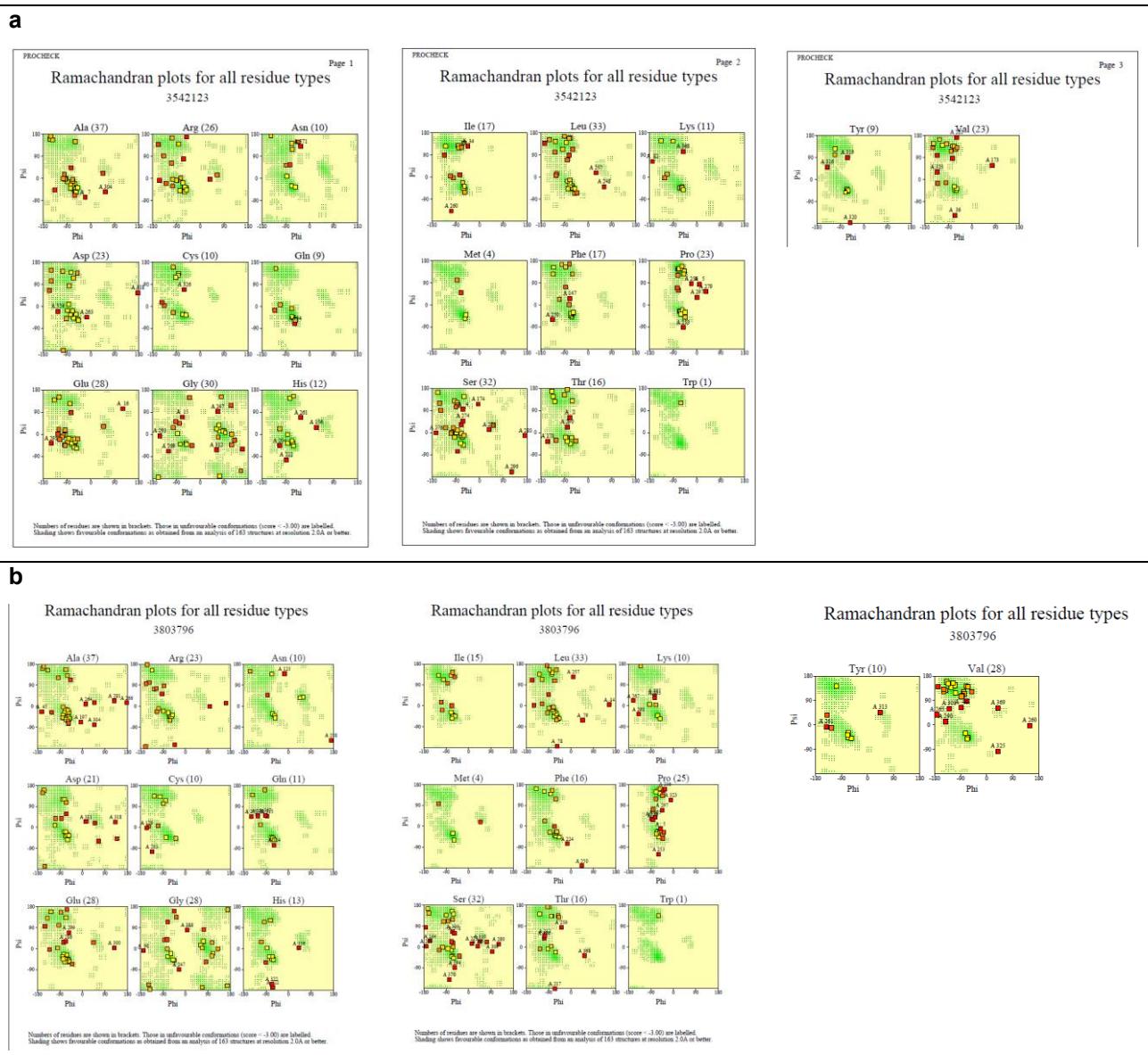
#### 4.4 SAVES, PROCHECK, WHATCHECK, QMEAN e DisEMBL – Verificação de Estereoquímica dos Modelos

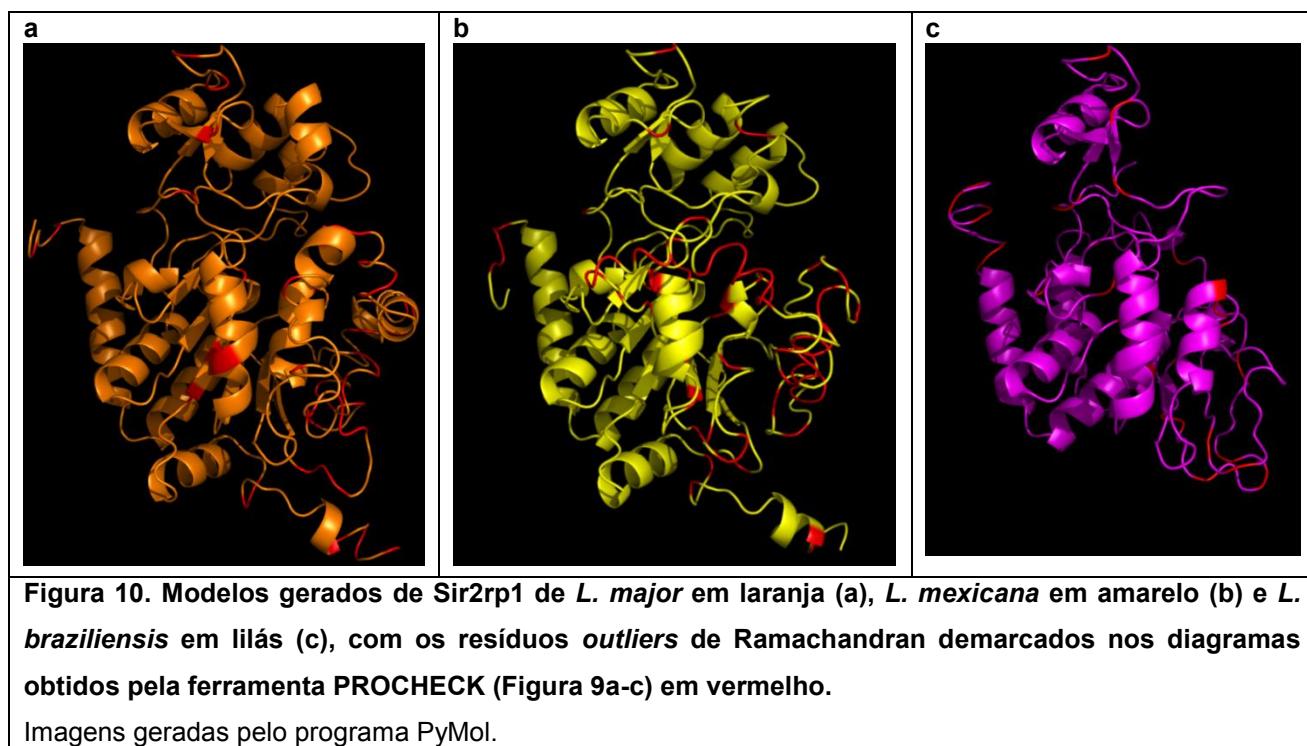
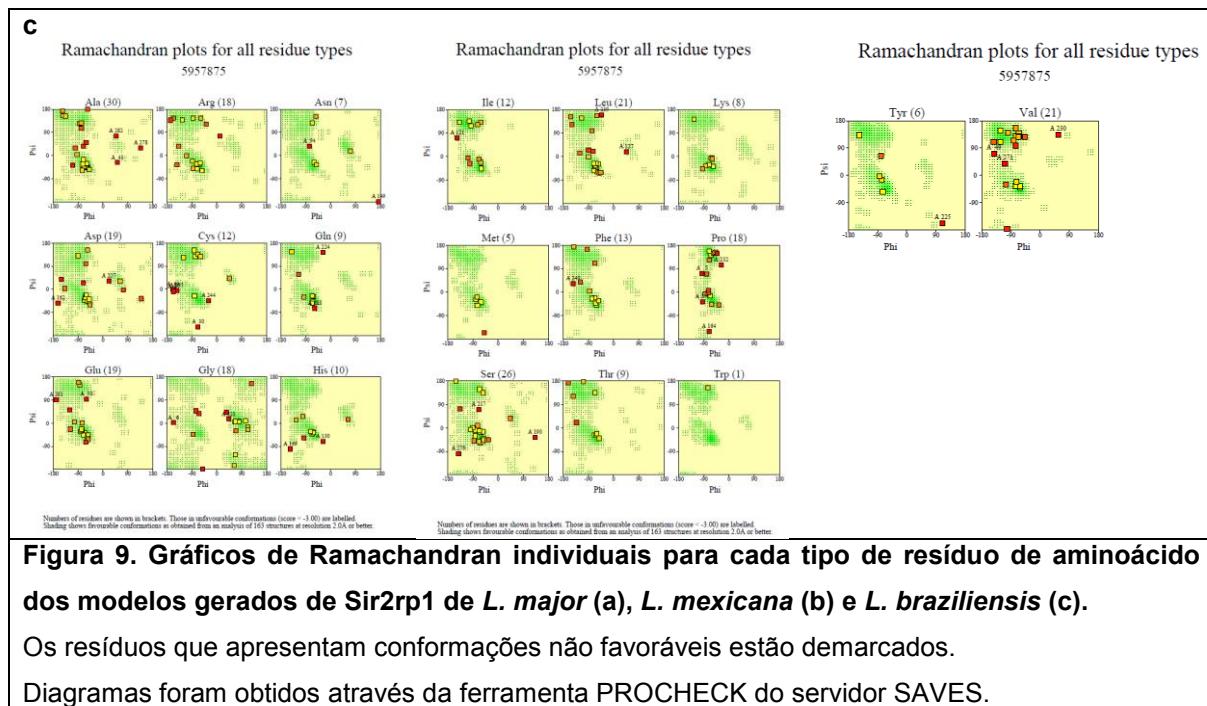
Os modelos gerados foram avaliados por meio de gráficos de Ramachandran (diagramas de ângulos phi e psi), distorção de grupos planares, desvio de ângulos ômega e análise de qualidade local dos resíduos de aminoácidos dos modelos. Estes dados também foram correlacionados com resultados da ferramenta preditora de regiões de desordem, Remark465.

##### 4.4.1 Gráficos de Ramachandran

Os gráficos de Ramachandran são diagramas dos valores de ângulos diédricos da cadeia principal dos resíduos de aminoácidos que compõem uma determinada proteína. Os ângulos phi representam a torsão da ligação química entre o grupo amino e o carbono alfa, enquanto que os ângulos psi representam a torção da ligação entre o carbono alfa e a carbonila do resíduo de aminoácido em questão (RICHARDSON, 1981).

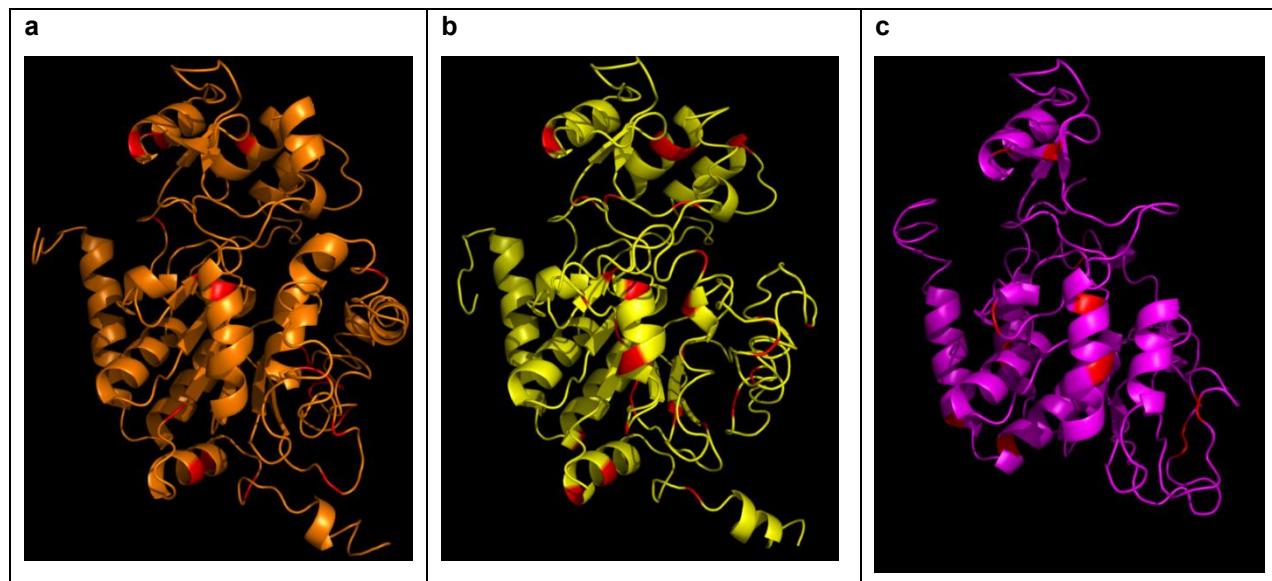
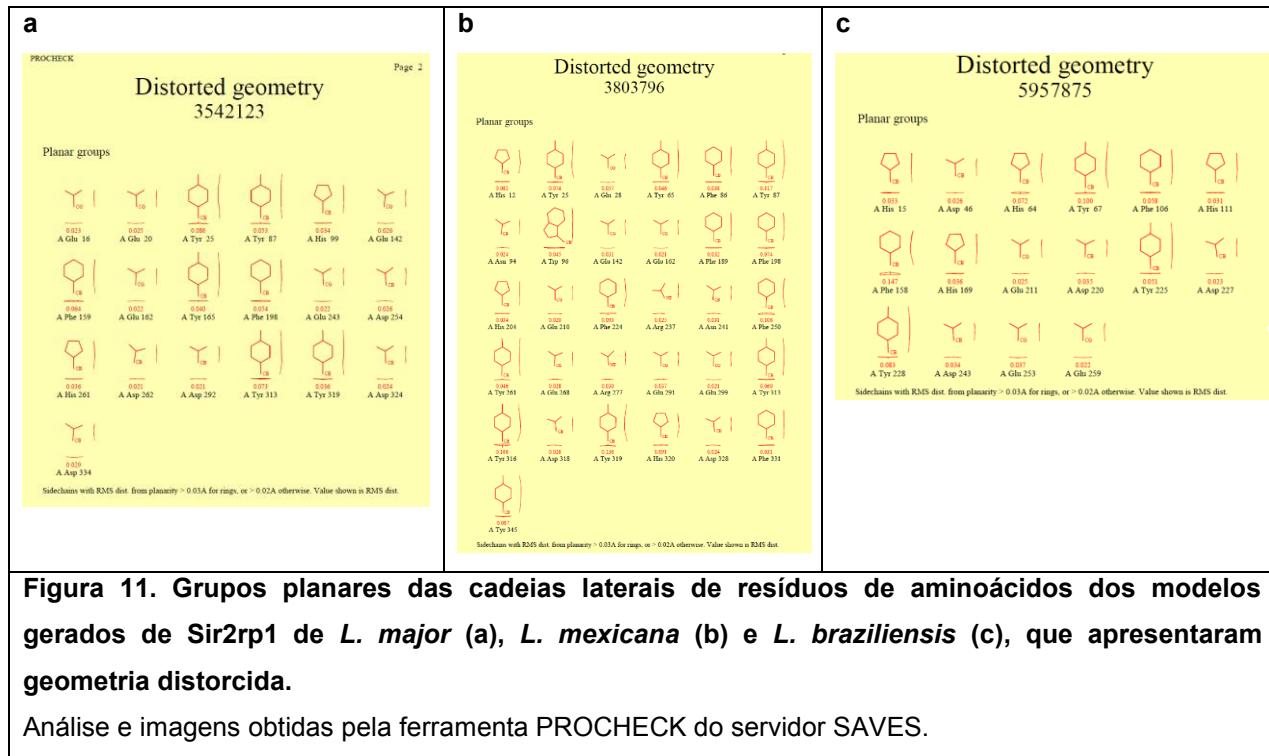
Os resíduos de aminoácidos considerados *outliers* de Ramachandran são aqueles cujos ângulos phi e/ou psi apresentam valores incompatíveis com as configurações de estruturas secundárias mais comuns para aquele tipo de aminoácido, sendo elas folhas-betas ou alfa-hélices (RICHARDSON, 1981).





#### 4.4.2 Distorção de grupos planares

Por meio da ferramenta PROCHECK, foram identificados os resíduos dos modelos gerados cujas cadeias laterais que deveriam apresentar geometria planar foram distorcidas e não se encaixam mais em um único plano.

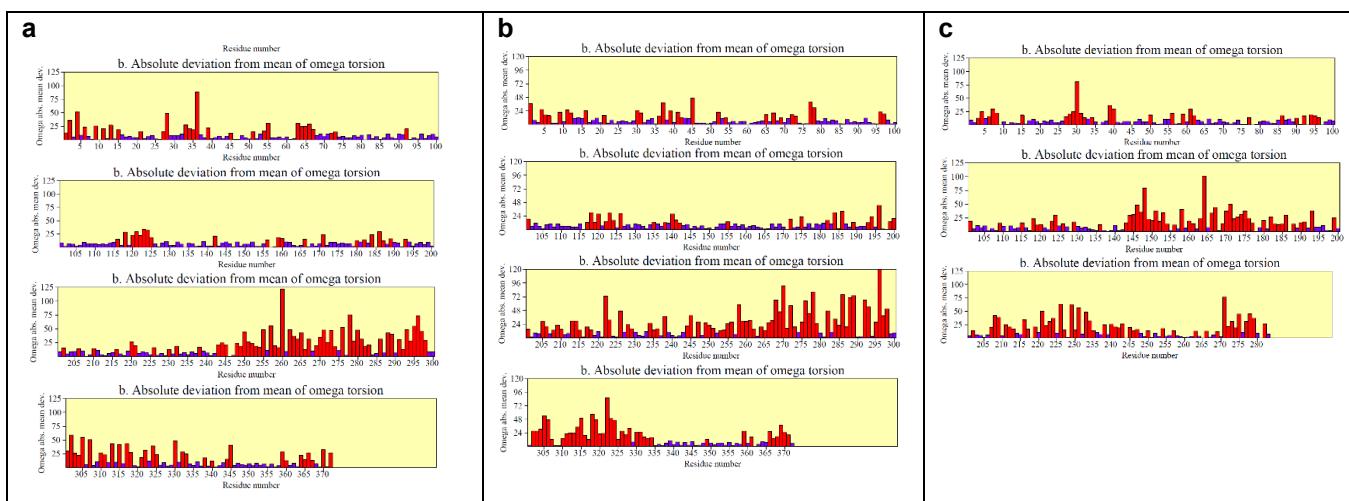


**Figura 12.** Modelos gerados de Sir2rp1 de *L. major* em laranja (a), *L. mexicana* em amarelo (b) e *L. braziliensis* em lilás (c), com os resíduos que apresentaram geometria planar distorcida demarcados pela ferramenta PROCHECK (Figura 11a-c) em vermelho.

Imagens geradas pelo programa PyMol.

#### 4.4.3 Desvio de torsão de ângulos Ômega

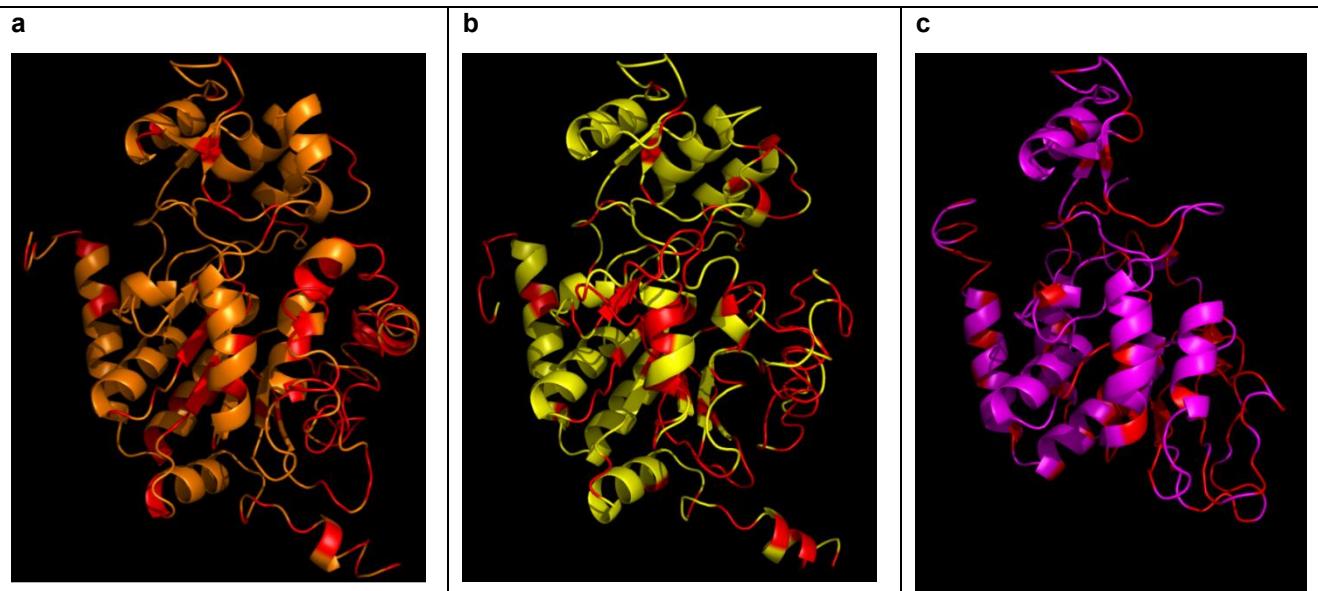
A ferramenta PROCHECK nos indicou os resíduos dos modelos gerados cujos valores de ângulos ômega, isto é, a torção da ligação peptídica, formada entre a carbonila de um resíduo de aminoácido e o grupo amino do próximo resíduo, apresentaram um desvio maior que 2 desvios padrão da média para aquele tipo de aminoácido (RICHARDSON, 1981).



**Figura 13.** Desvio absoluto médio de ângulos ômega dos resíduos de aminoácidos dos modelos gerados de Sir2rp1 de *L. major* (a), *L. mexicana* (b) e *L. braziliensis* (c).

Aqueles que aparecem em vermelho apresentaram um desvio maior que 2 desvios padrão do ideal para aquele tipo de aminoácido.

Análise e imagens obtidas através da ferramenta PROCHECK do servidor SAVES.

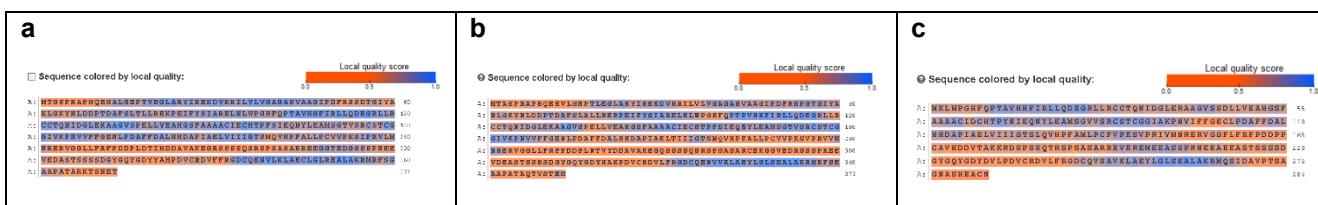


**Figura 14.** Modelos gerados de Sir2rp1 de *L. major* em laranja (a), *L. mexicana* em amarelo (b) e *L. braziliensis* em lilás (c), com os resíduos que apresentaram grande desvio de torsão de ângulos ômega demarcados pela ferramenta PROCHECK (Figura 13a-c) em vermelho.

Imagens geradas pelo programa PyMol.

#### 4.4.4 Análise de qualidade local dos resíduos de aminoácidos

Utilizou-se o QMEAN para avaliar os resíduos dos modelos gerados e classifica-los de acordo com sua pontuação de qualidade local.

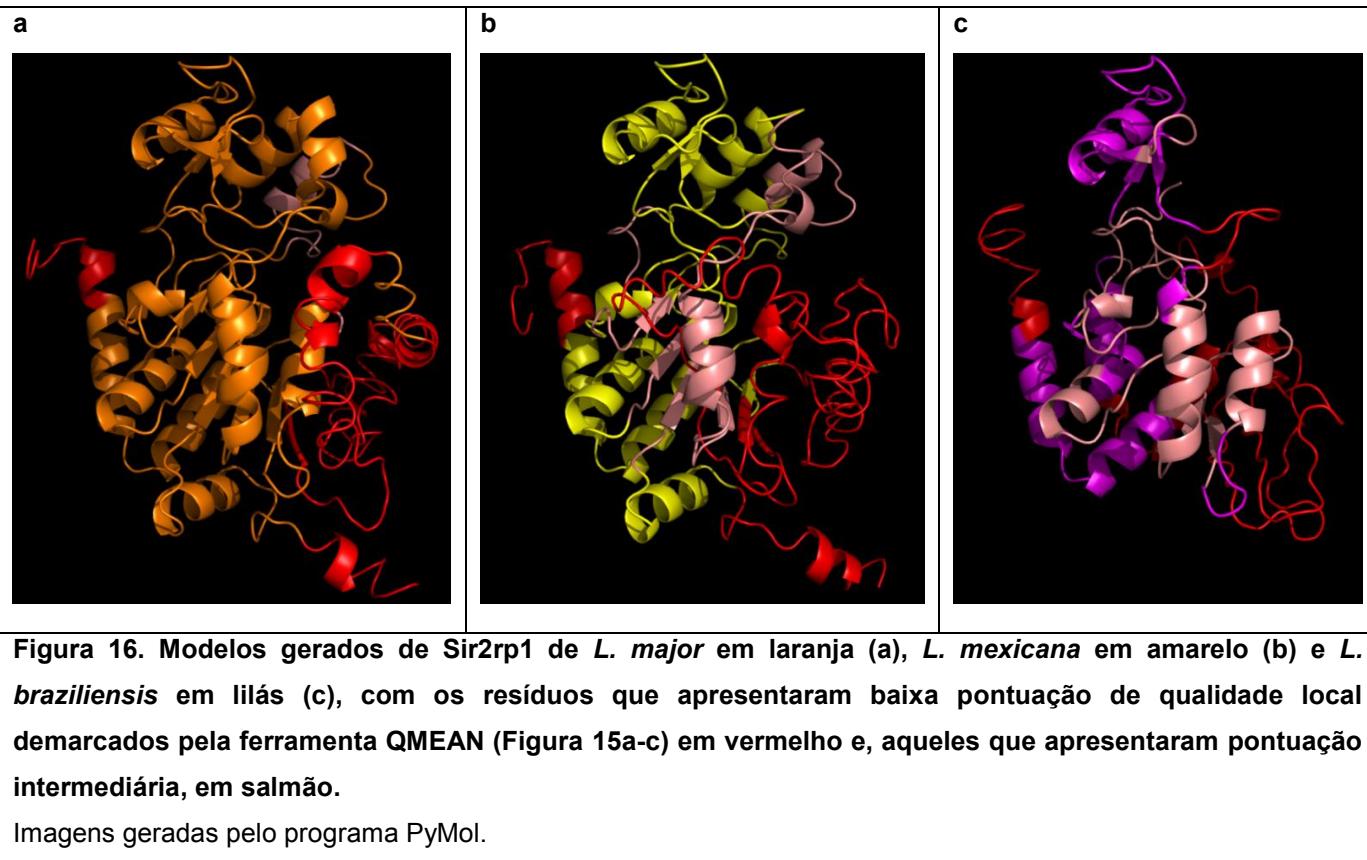


**Figura 15.** Sequências dos modelos de Sir2rp1 de *L. major* (a), *L. mexicana* (b) e *L. braziliensis* (c) gerados com os resíduos de aminoácidos demarcados de acordo com sua pontuação de qualidade local.

O gradiente de coloração varia de ruim (laranja) a bom (azul).

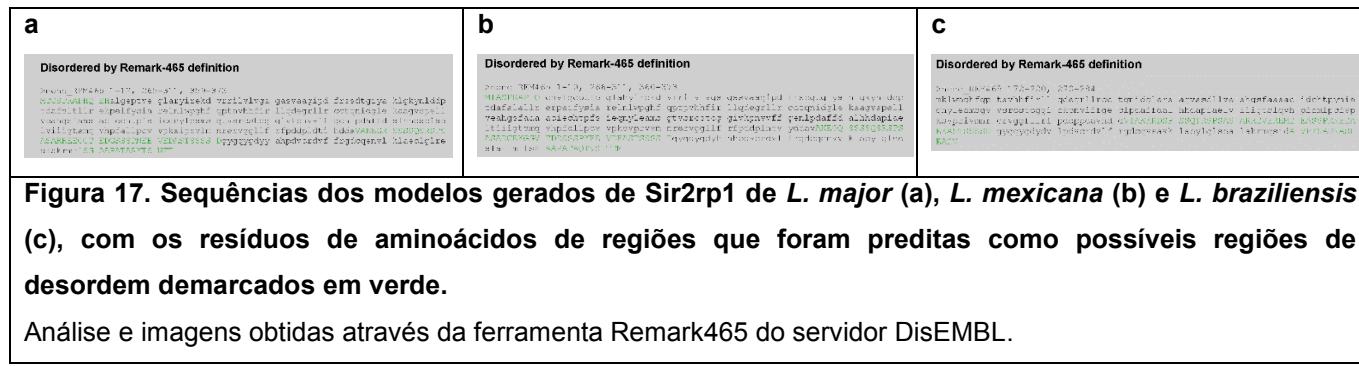
Seus valores de pontuação global (QMEAN4 Value) foram de: -6.69, -10.93 e -7.37, respectivamente.

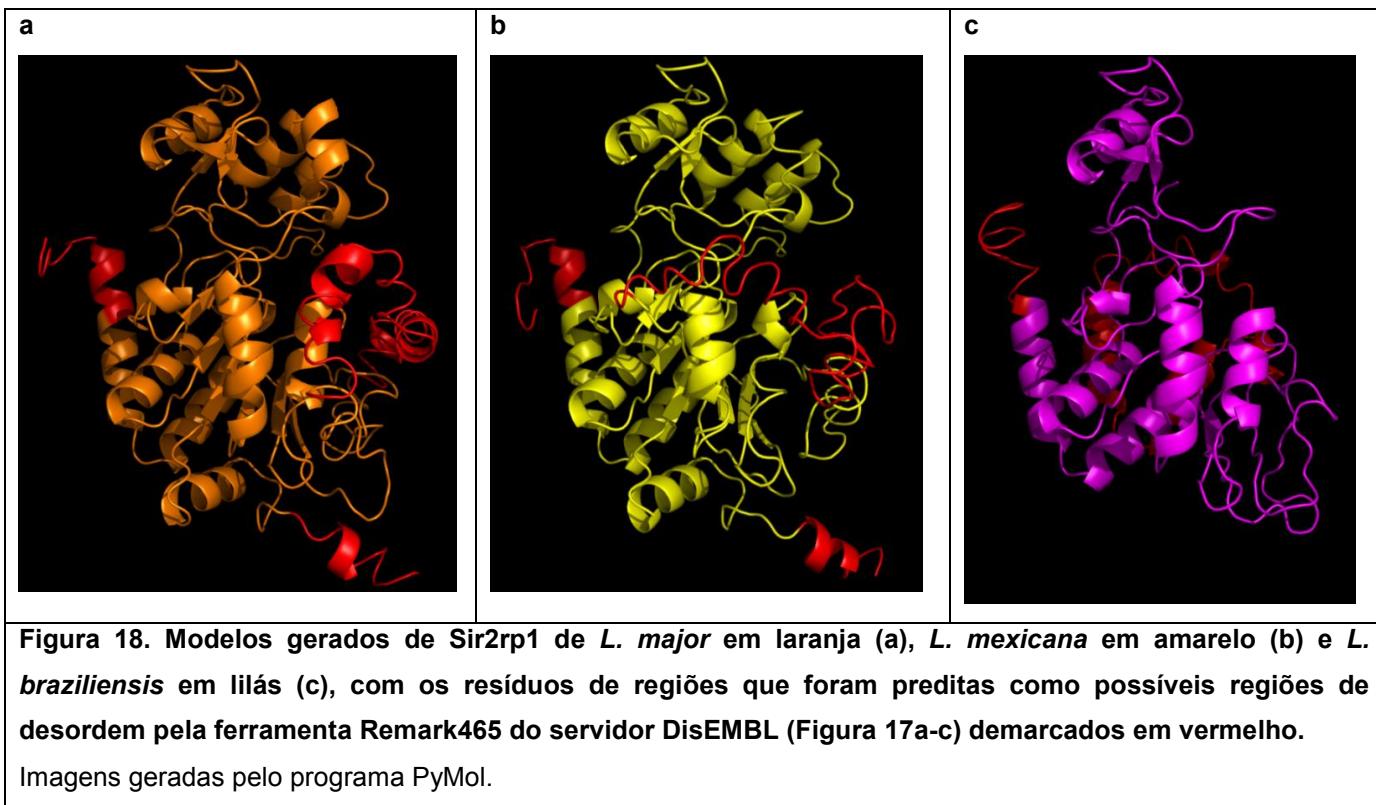
Análise e imagens obtidas através da ferramenta QMEAN.



#### 4.4.5 Predição de regiões estruturais de desordem à partir das sequências

A predição dos resíduos com alta propensão de formar regiões desordenadas nos modelos construídos foi feita pela ferramenta DisEMBL.





## 5. DISCUSSÃO

### 5.1 Escolha das sequências

A primeira etapa para a realização do trabalho de modelagem comparativa de proteínas é a escolha da sequência de aminoácidos a ser modelada e do molde mais adequado para a mesma. Tendo em vista que o servidor *I-Tasser* realiza predição do enovelamento com base em um molde (*Template Based*), deve-se empregar idealmente para tal finalidade uma sequência que apresente grau de identidade de sua estrutura primária com as sequências a serem modeladas de no mínimo 40%, porém estruturas com até 20% de identidade também podem ser empregadas, especialmente no caso do emprego de múltiplos moldes para diferentes regiões das estruturas modeladas (VERLI, 2014; FERREIRA, 2018).

Neste caso, considerando que até então apenas uma enzima da família das Sirtuínas 2 do gênero *Leishmania* fora cristalizada e teve a sua estrutura terciária elucidada por meio da técnica de cristalografia por difração de raios X (código

PDB 5OL0), o presente trabalho empregou esta como molde, na construção de modelos de Sirtuínas 2 *related protein 1* (Sir2rp1) homólogas de outras espécies do mesmo gênero (*L. major*, *L. mexicana* e *L. braziliensis*), devido a sua proximidade filogenética e identidade entre suas estruturas primárias. Foram então levantadas todas as sequências primárias de enzimas da família das Sirtuínas 2 do gênero *Leishmania* que foram depositadas na plataforma NCBI, para posterior análise e seleção daquelas a serem modeladas (LIMA, 2016; RONIN, 2018; FERREIRA, 2018).

### **5.1.2 Predição de localização intracelular e presença de peptídeo sinal**

As sequências foram analisadas pela ferramenta Phobius, a fim de predizer sua localização intracelular e presença de peptídeo sinal. Desta maneira, foi possível distinguir as enzimas homólogas do molde escolhido, LiSir2rp1, daquelas pertencentes à família das Sirtuínas 2, porém mais distantes, como aquelas dos grupos Sir2rp2 e Sir2rp3, devido às notáveis diferenças estruturais e de localização intracelular preditas pela ferramenta. Observa-se que o perfil de localização intracelular predito pela ferramenta segue um padrão semelhante para todas as Sir2rp1 (Figura 3), e a ausência de sequência sinalizadora nestas enzimas. Este resultado é condizente com a localização citoplasmática das mesmas e também reflete a semelhança entre as suas estruturas primárias (KÄLL, 2004; RONIN, 2018).

### **5.1.3 Alinhamento das sequências**

Dentre as sequências de Sir2rp1 de *L. major*, *L. mexicana* e *L. braziliensis* escolhidas, respectivamente sp|Q25337.2, CBZ27844.1 e CBZ14728.1, realizou-se um alinhamento com a sequência do molde de *L. infantum*, por meio da ferramenta *T-Coffee*. Observa-se na figura 4 o grau de identidade de 99% entre a sequência molde e as sequências de *L. major* e *L. mexicana* e de 90% para a sequência de *L. braziliensis*. Nota-se também que a sequência de Sir2rp1 de *L. braziliensis* é mais curta do que as demais (aproximadamente 100 resíduos de aminoácidos a menos), e apresenta a porção correspondente à região N terminal e ao *co-factor-binding loop* (Figura 5b) ausentes de sua estrutura. Tal discrepância se deve, provavelmente, a um erro na interpretação dos segmentos do gene que

correspondem a íntrons e éxons, no trabalho de sequenciamento genético do qual essa sequência de aminoácidos provém. (NOTREDAME, 2000; ROGERS, 2011; TOMMASO, 2011; LIMA, 2016).

### **5.2 Modelagem**

A construção dos modelos tridimensionais foi realizada no servidor *I-Tasser*, a partir das sequências previamente selecionadas. Não foi indicado qual molde deveria ser utilizado, e dessa forma a plataforma realizou um processo preliminar automatizado de alinhamento e busca por candidatos a molde, resultando na escolha da *Sirtuin 2 related protein 1* de *L. infantum* como molde (Figura 5), o mesmo que havia sido eleito como o mais adequado no início deste trabalho. O grupo responsável pelo desenvolvimento desta ferramenta optou por esta abordagem tendo em vista que nem sempre a sequência proveniente de uma espécie de maior proximidade filogenética à aquela que está sendo estudada é a mais adequada para ser empregue como molde. Muitas vezes, existe uma sequência de alguma espécie mais distante que apresenta um alinhamento mais favorável com a sequência a ser modelada. Além disso, este servidor tem a capacidade de empregar múltiplos moldes, identificando o mais adequado para cada região individual da sequência a ser modelada. Desta maneira, esta metodologia visa aprimorar a qualidade dos modelos gerados, evitando enviesar a modelagem pela escolha de um molde que em uma primeira análise aparenta ser o mais adequado, devido à proximidade filogenética entre as espécies das quais as sequências provém, mas que na prática apresenta um alinhamento menos favorável do que outra sequência que não fora considerada pelo usuário da ferramenta (ZHANG, 2008; ROY, 2010; VERLI, 2014; YANG, 2015).

### **5.3 Verificação de Estereoquímica dos Modelos**

Os modelos de Sir2rp1 gerados (Figura 6) tiveram a sua estereoquímica avaliada através de gráficos de Ramachandran (Figuras 9 e 10), distorção de grupos planares (Figuras 11 e 12), desvio de ângulos ômega (Figuras 13 e 14) e análise de qualidade local dos resíduos de aminoácidos dos mesmos (Figuras 15 e 16), além de obtenção de seu RMSD com relação ao molde empregado através de alinhamento tridimensional entre as suas estruturas (Figura 7). Estes

resultados também foram correlacionados com a análise de suas estruturas primárias pela ferramenta preditora de regiões de desordem, Remark465 (Figuras 17 e 18) (LASKOWSKI, 1993; HOOFT, 1996; LASKOWSKI, 1996; LINDING, 2003; BENKERT, 2011).

Constatou-se que diversos resíduos de aminoácidos dos modelos gerados apresentam anomalias estruturais (Figuras de 9 a 16), sendo estas: ângulos phi e psi incompatíveis com as estruturas secundárias das quais fazem parte, grupos planares das cadeias laterais que apresentaram sua geometria distorcida de um único plano e ângulos ômega com desvio maior que 2 desvios padrão da média para o tipo de aminoácido em questão, além da pontuação de qualidade local baixa para alguns resíduos. Tais anomalias podem ser atribuídas a alguns fatores, tais como: a estrutura do molde ter sido elucidada por cristalografia de raios X, a partir de um cristal da enzima, que dificilmente apresentaria uma estrutura idêntica àquela que assume em condições fisiológicas, considerando o quão distintas estas são daquelas empregadas na cristalização de proteínas. A modelagem comparativa ter sido realizada em um meio virtual, sem a presença de moléculas de solvente e outros solutos. O fato de que a maior parte destes resíduos de aminoácidos que apresentaram anomalias estruturais encontram-se nas regiões N e C terminais e em uma região específica que aparece no meio da sequência das enzimas (Figura 8). Pois, não só estas três regiões são preditas pela ferramenta Remark465 como sendo prováveis regiões de desordem (Figuras 17 e 18), mas também porque estas estão ausentes no molde empregado, devido ao tratamento proteolítico breve que foi necessário para a obtenção do cristal, justamente por conta da impossibilidade de cristalizar a enzima íntegra. A ausência destas regiões no molde prejudica a capacidade preditiva do algorítimo modelador (VERLI, 2014; MONIOT, 2013; FERREIRA, 2018; RONIN, 2018).

Além do mais, as constatações de que: estas regiões (Figura 8) são muito mais suscetíveis à hidrólise por proteases do que outras presentes na enzima, os dois segmentos resultantes da proteólise permanecem unidos após sua clivagem, a estrutura resultante desta hidrólise apresenta atividade catalítica inalterada, e a inabilidade de se obter um cristal quando da presença destas regiões, todas

corroboram o conceito de que estas são, de fato, regiões de baixa rigidez estrutural, além de não comporem os domínios funcionais da enzima (RONIN, 2018).

Estas anomalias nas cadeias dos resíduos de aminoácidos dos modelos, principalmente desvios de Ramachandran e de ângulos ômega, também tendem a ocorrer em regiões de alinhamento com motivos estruturais bastante conservados filogeneticamente e de grande importância para o desempenho das funções enzimáticas, no molde. Pois, os algorítmos modeladores apresentam neste caso um viés no posicionamento dos resíduos destas regiões, visando manter as suas estruturas secundárias e terciária dentro do esperado para aquele domínio que compõem, como é o caso do Dedo de Zinco e da Dobra de Rossmann (Figura 5b), ambos domínios funcionais essenciais para a atividade enzimática das sirtuínas e, também, bastante conservados ao longo de todos os clados filogenéticos que as expressam (MONIOT, 2013; VERLI, 2014; FERREIRA, 2018).

A análise dos valores de RMSD calculados à partir da sobreposição e alinhamento tridimensional entre os modelos obtidos e o molde estrutural empregado (Figura 7) indicam o quanto próximas são suas estruturas espacialmente. Os valores obtidos para os modelos de Sir2rp1 de *L. major* e *L. mexicana*, respectivamente 0,428 e 0,228 Angstroms, refletem sua proximidade filogenética e o alto grau de identidade de 99% entre suas estruturas primárias (Figura 4a-b), e também o alto nível de conservação entre suas estruturas secundárias e terciárias, o que implica que estas proteínas apresentem atividade enzimática similar, como deacetilases de peptídeos dependentes de NAD<sup>+</sup> e Zn<sup>2+</sup> como cofatores, devido à presença em comum dos domínios Dedo de Zinco e da Dobra de Rossmann, ambas estruturas altamente conservadas (Figura 5b). Já o valor maior de RMSD calculado para o modelo obtido de Sir2rp1 de *L. braziliensis* (1,583 Angstroms) reflete a identidade ligeiramente menor entre sua estrutura primária e a do molde, de 90% (Figura 4c), o que se deve tanto à ausência da porção de aproximadamente 100 aminoácidos que abrange a região N-terminal e o *co-factor-binding-loop* (Figura 5b) de sua sequência obtida do NCBI como a sua

distância filogenética um pouco maior com relação ao molde (COLE, 2005; MONIOT, 2013; HAWKINS, 2017; RONIN, 2018).

Desta forma outras técnicas complementares, como a Dinâmica Molecular (DM), se fazem necessárias para refinar o enovelamento dos modelos e a configuração de seus resíduos de aminoácidos, visando a obtenção de modelos fidedignos que sejam adequados a aplicação em futuros estudos de planejamento de fármaco baseado na estrutura do alvo (SBDD) e de triagem virtual baseada na estrutura do alvo (SBVS) (VERLI, 2014; FERREIRA, 2018).

## 6. CONCLUSÃO

Diante dos resultados obtidos, o presente trabalho se mostra uma contribuição para o conhecimento estrutural das Sirtuínas 2 *related protein 1* (Sir2rp1) de espécies do gênero *Leishmania*, possibilitando que futuros estudos empreguem os modelos construídos para o desenvolvimento de fármacos visando o tratamento das Leishmanioses, uma doença negligenciada de alto impacto sócio-econômico que é a Leishmaniose, causada por estes parasitas.

No mais, as ferramentas empregadas no desenvolvimento deste trabalho, tanto para a análise e triagem de sequências peptídicas, como para a modelagem de suas estruturas secundárias e terciárias, tanto quanto para a avaliação dos modelos gerados, contribui para o conhecimento de modelagem virtual, necessário no desenvolvimento de modelos de outros alvos terapêuticos em potencial.

A qualidade dos resultados obtidos neste trabalho, também, nos da indicações que temos modelos de Sir2rp1 adequados a aplicação em estudos de planejamento de fármaco baseado na estrutura do alvo (SBDD) e de triagem virtual baseada na estrutura do alvo (SBVS), visando a obtenção de novos candidatos à fármacos com maior seletividade, eficácia e segurança frente aos parasitas do gênero *Leishmania*.

## 7. BIBLIOGRAFIA

BIOVIA *Discovery Studio Visualizer* (BIOVIA, D. S. 2015. *Discovery studio modeling environment. San Diego, Dassault Systèmes, Release, 4*)

ALMEIDA, L.; FUJIMURA, A.T.; FONSECA-SANTOS, B.; IMAMURA, K.B.; MICHELS, P.; CHORILLI, M.; GRAMINHA, M. Nanotechnological Strategies for Treatment of Leishmaniasis—A Review. **Journal of Biomedical Nanotechnology**, v. 13(2), p. 1-17, 2017.

ANDERSON, A.C. The Process of Structure-Based Drug Design. **Chemistry & Biology**, v. 10, p. 787-797, 2003.

BENKERT, P.; BIASINI, M.; SCHWEDE, T. Toward the estimation of the absolute quality of individual protein structure models. **Bioinformatics**, v. 27, p. 343-350, 2011.

CLABORN, M.D. The Biology and Control of Leishmaniasis Vectors. **Journal of Global Infectious Diseases**, v. 2 (2), p. 127-134, 2010.

GLAAB, E. Building a virtual ligand screening pipeline using free software: a survey. **Briefings in Bioinformatics**, v. 17(2), p. 352-366, 2016.

COLE, J.C.; MURRAY, C.W.; NISSINK, J.W.M.; TAYLOR, R.D.; TAYLOR, R. Comparing Protein–Ligand Docking Programs Is Difficult. **PROTEINS: Structure, Function, and Bioinformatics**, v. 60, p. 325-332, 2005.

FERREIRA, G.M. Busca por inibidores seletivos de Sirtuín 2 de *T. cruzi* empregando técnicas de planejamento de fármacos baseado na estrutura do receptor. **Tese de Doutorado - Faculdade de Ciências Farmacêuticas – Universidade de São Paulo**, São Paulo, 2018.

GEORGIADOU, P.S.; MAKARITSIS, P.K.; DALEKOS, N.G. Leishmaniasis revisited: Current aspects on epidemiology, diagnosis and treatment. **Journal of Translational Internal Medicine**, v. 3, p. 43-50, 2015.

HAWKINS, P.C.D. Conformation Generation: The State of the Art. **Journal of Chemical Information and Modeling**, v. 57, p. 1747-1756, 2017.

HOOFT, R.W.W.; VRIEND, G.; SANDER, C.; ABOLA, E.E. Errors in protein structures. **Nature**, v. 381, p. 272-272, 1996.

KÄLL, L.; KROGH, A.; SONNHAMMER, E.L.L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. **Journal of Molecular Biology**, v. 338(5), p. 1027-1036, 2004.

LASKOWSKI, R.A.; MACARTHUR, M.W.; MOSS, D.S.; THORNTON, J.M. PROCHECK - a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, p. 283-291, 1993.

LASKOWSKI, R.A.; RULLMANNN, J.A.; MACARTHUR, M.W.; KAPTEIN, R.; THORNTON, J.M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. **Journal of Biomolecular NMR**, v. 8, p. 477-486, 1996.

LIMA, N.R.; FAHEEM, M.; BARBOSA, A.R.G.J.; POLÊTO, D.M.; VERLI, H.; MELO, L.F.; RESENDE, O.R. Homology modeling and molecular dynamics provide structural insights into *tospovirus* nucleoprotein. **BMC Bioinformatics**, v. 17, s. 18, p. 489, 2016.

LINDING, R.; JENSEN, L.J.; DIELLA, F.; BORK, P.; GIBSON, T.J.; RUSSELL, R.B. Protein disorder prediction: implications for structural proteomics. **Structure**, v. 11, p. 4, 2003.

MITTAL, N.; MUTHUSWAMI, R.; MADHUBALA, R. The mitochondrial SIR2 related protein 2 (SIR2RP2) impacts *Leishmania donovani* growth and infectivity. **PLOS Neglected Tropical Diseases**, v. 11, p. 5, 2017.

MONIOT, S.; SCHUTKOWSKI, M.; STEEGBORN, C. Crystal structure analysis of human Sirt2 and its ADP-ribose complex. **Journal of Structural Biology**, v. 182, p. 136 – 143, 2013.

MURRAY, H.W.; BERMAN, J.D.; DAVIES, C.R.; SARAVIA, N.G. Advances in leishmaniasis. **The Lancet**, v. 366, i. 9496, p. 1561-1577, 2005.

NOTREDAME, C.; HIGGINS, D.G.; HERINGA, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. **Journal of Molecular Biology**, v. 302(1), p. 205-217, 2000.

OMS<sub>a</sub> [https://www.who.int/neglected\\_diseases/diseases/en/](https://www.who.int/neglected_diseases/diseases/en/) (acesso em 01/09/2020)

OMS<sub>b</sub> [https://www.who.int/gho/neglected\\_diseases/leishmaniasis/en/](https://www.who.int/gho/neglected_diseases/leishmaniasis/en/) (acesso em 01/09/2020)

OMS<sub>c</sub> <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis> (acesso em 01/09/2020)

OMS<sub>d</sub> [https://www.who.int/leishmaniasis/burden/Leishmaniasis\\_Brazil/en/](https://www.who.int/leishmaniasis/burden/Leishmaniasis_Brazil/en/) (acesso em 01/09/2020)

OMS<sub>e</sub> <https://www.who.int/leishmaniasis/research/en/> (acesso em 01/09/2020)

RICHARDSON, J.S. Anatomy and Taxonomy of Protein Structures. **Advances in Protein Chemistry**, v. 34, p. 167-339, 1981.

ROGERS, M.B.; HILLEY, J.D.; DICKENS, N.J.; WILKES, J.; BATES, P.A.; DEPLEDGE, D.P.; HARRIS, D.; HER, Y.; HERZYK, P.; IMAMURA, H.; OTTO, T.D.; SANDERS, M.; SEEGER, K.; DUJARDIN, J.C.; BERRIMAN, M.; SMITH, D.F.; HERTZ-FOWLER, C.; MOTTRAM, J.C. Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. **Genome research**, v. 21(12), p. 2129–2142, 2011.

RONIN, C.; COSTA, M.D.; TAVARES, J.; FARIA, J.; CIESIELSKI, F.; CIAPETTI, P.; SMITH, K.T.; MACDOUGALL, J.; CORDEIRO-DA-SILVA, A.; PEMBERTON, K.L. The crystal structure of the *Leishmania infantum* Silent Information Regulator 2 related protein 1: Implications to protein function and drug design. **PLOS ONE**, 2018.

ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: a unified platform for automated protein structure and function prediction. **Nature Protocols**, v. 5, p. 725-738, 2010.

The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

TOMMASO, P.D.; MORETTI, S.; XENARIOS, I.; OROBITG, M.; MONTANYOLA, A.; CHANG, J.M.; TALY, J.F.; NOTREDAME, C. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. **Nucleic Acid Research**, v. 39, p. 13-17, 2011.

ULIANA, R.B.S.; TRINCONI, T.C.; COELHO, C.A. Chemotherapy of leishmaniasis: present challenges. **Parasitology**, v. 145, p. 464-480, 2018.

VERGNES, B.; SERENO, D.; TAVARES, J.; CORDEIRO-DA-SILVA, A.; VANHILLE, L.; MADJIDIAN-SERENO, N.; DEPOIX, D.; MONTE-ALEGRE, A.; OUAISI, A. Targeted disruption of cytosolic SIR2 deacetylase discloses its essential role in *Leishmania* survival and proliferation. **GENE**, v. 363, p. 85-96, 2005.

VERLI, H. **Bioinformática: da Biologia à Flexibilidade Molecular**. 1.ed. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2014.

YANG, J.; YAN, R.; ROY, A.; XU, D.; POISSON, J.; ZHANG, Y. The I-TASSER Suite: Protein structure and function prediction. **Nature Methods**, v. 12, p. 7-8, 2015.

ZHANG, Y. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, v. 9, p. 40, 2008.



---

Data e assinatura do aluno(a)



---

Data e assinatura do orientador(a)