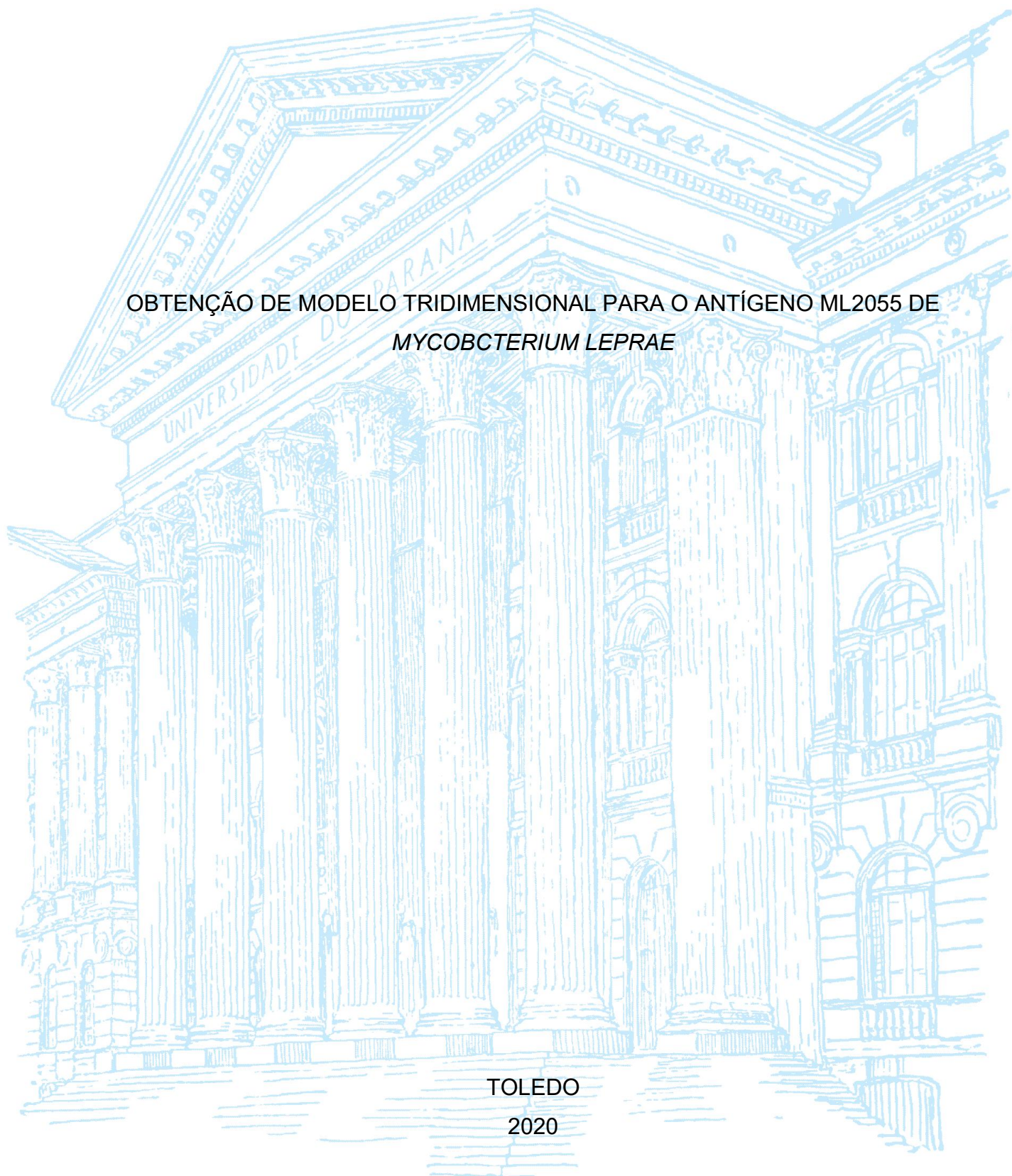


UNIVERSIDADE FEDERAL DO PARANÁ

BIANCA LUIZA MELO DE ASSIS

OBTENÇÃO DE MODELO TRIDIMENSIONAL PARA O ANTÍGENO ML2055 DE  
*MYCOBACTERIUM LEPRAE*

TOLEDO  
2020



BIANCA LUIZA MELO DE ASSIS

OBTENÇÃO DE MODELO TRIDIMENSIONAL PARA O ANTÍGENO ML2055 DE  
*MYCOBACTERIUM LEPRAE*

Trabalho de curso apresentado ao curso de medicina da Universidade Federal do Paraná-Campus Toledo, como requisito parcial de obtenção do título de Bacharel em medicina.

Orientadora: Profa. Dra. Kádima Nayara Teixeira

TOLEDO  
2020



Dedico este trabalho aos ousados, àqueles que ousam sonhar, que ousam acreditar que o mundo pode ser melhor, diferente. Dedico a todos, que de alguma forma, contribuem para que a vida, a ciência e a sociedade evoluam sempre. Assim, dedico este trabalho a cada um de vocês.

## **AGRADECIMENTOS**

Agradeço, primeiramente, a Deus, pois me trouxe a casa que jamais imaginei ser a minha, a inalcançável Universidade Federal do Paraná. Obrigada! Agradeço a cada professor que teve paciência e maestria ao me ensinar, a vocês toda a minha gratidão e admiração, não teria conseguido sem a participação de vocês. A minha família e amigos, sou grata por sempre acreditarem no meu sonho, principalmente quando eu não achei que seria possível. Vocês me motivam todos os dias. Agradeço, especialmente, minha orientadora Prof<sup>a</sup>. Dra. Kádima Nayara Teixeira, por primeiro ter me aceito como aluna de iniciação científica e por sugerir este trabalho. A todos vocês, o meu mais sincero, muito obrigada!

“O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo. Mesmo não atingindo o alvo, quem busca e vence obstáculos, no mínimo fará coisas admiráveis.” (ALENCAR, JOSÉ)

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>9</b>
<b>2 ARTIGO</b>	<b>11</b>
<b>3 CONSIDERAÇÕES FINAIS</b>	<b>22</b>

## 1 INTRODUÇÃO

A Hanseníase é uma doença causada pelo *Mycobacterium leprae* cujos relatos datam dos tempos bíblicos. Ao longo da história possuiu momentos em que esteve em alta e outros flertou com o esquecimento. Entretanto, trata-se de uma doença negligenciada que se perpetua em situações de pobreza.

A patologia é caracterizada por ter acometimento cutâneo e nervoso, mais comumente, entretanto, pode acometer o trato respiratório superior, ossos e olhos (ANTUNES, 2019). A lesão característica é uma área que pode variar de eritematosa a hipopigmentada, geralmente cursando com perda da sensibilidade térmica e tátil. Existe, hoje, cinco espectros da doença, entretanto, os de maior relevância nesse caso são a forma multibacilar (lepromatosa) caracterizada por mais de cinco lesões de pele e acometimento de mais de um tronco nervoso e a forma paucibacilar (tuberculoide) caracterizada por menos de cinco lesões cutâneas e no máximo um tronco nervoso acometido.

Diante desses espectros de doença nenhum teste diagnóstico laboratorial é suficiente para estabelecer o diagnóstico da doença. Tem-se disponível hoje a intradermorreação de Mitsuda (não sendo exclusivo para doença) e a baciloscopia realizada a partir de esfregaço de lesões de pele. Ambos são insuficientes para diagnosticarem os pacientes com a forma branda por estes apresentarem baixa contagem de bacilos.

Dessa forma, o desenvolvimento de teste sorológico sensível para as diferentes apresentações da doença é de suma importância. Tal feito pode começar a ser realizado a partir da modelagem de peptídeos imunogênicos identificados por Santana em 2017. O problema a ser resolvido é que tais peptídeos não possuem sua conformação e estrutura tridimensional conhecidas.

De maneira geral, o intuito do presente trabalho é obter um modelo de estrutura tridimensional para uma proteína de *Mycobacterium leprae* com potencial imunogênico para desenvolvimento de teste sorodiagnóstico capaz de detectar a forma paucibacilar da hanseníase.

Mais especificamente, os objetivos são modelar por meio do método molecular o antígeno ML2055 que faz parte de uma proteína imunogênica; construir um modelo tridimensional para a proteína antigênica por meio de ferramentas de



bioinformática; localizar os epítomos imunodominantes dentro da estrutura do antígeno ML2055 e caracterizar bioquimicamente os epítomos imunodominantes.

Tais fatos se justificam, pois, a doença continua sendo um problema de saúde pública no Brasil. Uma vez que, é uma condição tratável cujo número de casos diminui anualmente. Entretanto, ao tratar-se de indivíduos menores de 15 anos essa perspectiva não se mantém. Reforçando, assim, que há ineficácia ou no tratamento ou no rastreio e diagnóstico do restante da população.

O estabelecimento da infecção é inteiramente dependente do status imunológico de cada hospedeiro. Portanto, a resposta ao microrganismo é individual e a existência de pessoas que não desenvolvem a doença clínica não reflete a população como um todo. Segundo dados do ministério da saúde até 2018, os pacientes mais diagnosticados são homens com mais de 60 anos, demonstrando mais uma vez que os jovens são os doentes assintomáticos ou pouco sintomáticos disseminadores da condição. Daí a necessidade de diagnosticar e tratar esses pacientes.

## 2 ARTIGO

O presente estudo foi realizado em formato de artigo conforme deliberação da Comissão de Trabalho de Curso da UFPR – Campus Toledo conforme as normas da Revista Medicina (Ribeirão Preto). Entretanto, por se tratar de uma vertente de um projeto maior, os dados obtidos neste trabalho, juntamente com dados de outros pesquisadores, serão usados para elaboração de um artigo que será submetido para a revista Journal of Biomédical Informatics. O artigo encontra-se em fase de preparação, cuja responsável é a Prof<sup>a</sup>. Juliana Ferreira de Moura da UFPR- Setor de Biológicas.

## Introdução

A hanseníase é uma doença granulomatosa crônica causada pelo *Mycobacterium leprae*, uma bactéria intracelular ácido-álcool resistente (BAAR) que infecta a pele e os nervos periféricos (ZHU et al., 2006); além desses locais mais comuns, pode afetar o trato respiratório superior, ossos e olhos (ANTUNES, 2019). Tal doença tem seus primeiros relatos datando de tempos bíblicos e permanece como um problema de saúde pública em países em desenvolvimento como o Brasil.

Existem duas principais apresentações clínicas da doença. A forma paucibacilar é caracterizada por um número de lesões inferior a cinco ou que acomete apenas um tronco nervoso. Essa apresentação desencadeia uma resposta imunológica celular do tipo Th1 promovendo controle da infecção, resultando em baixo índice baciloscópico, sendo assim insensível aos métodos diagnósticos sorológicos disponíveis atualmente. A segunda forma de apresentação é a forma multibacilar ou lepromatosa caracterizada por múltiplas lesões e acometimento de diferentes troncos nervosos. Tal apresentação é mediada por uma resposta celular do tipo Th2 e possui alto índice baciloscópico, sendo diagnosticada pela detecção de anticorpos anti-glicolípido fenólico I (PGL-I).

Diante da dificuldade de se obter um teste sorológico sensível para ambas as formas clínicas da doença pesquisas têm sido desenvolvidas na tentativa de sanar essa questão. Santana, em 2017, mapeou epítomos com potencial imunológico satisfatório a partir de antígenos identificados por Spencer e colaboradores em 2012.

Santana identificou sete proteínas imunogênicas com potencial para serem utilizadas em testes sorológicos. Dentre essas proteínas, a ML2055 apresenta o maior potencial imunogênico. Todavia, sua estrutura tridimensional ainda não foi resolvida, logo o objetivo deste trabalho foi obter um modelo de estrutura tridimensional para essa proteína e assim identificar seus epítomos imunodominantes com potencial para o desenvolvimento de um teste sorológico sensível.

## Materiais e Métodos

### Modelagem molecular da proteína ML2055 de *M. leprae*

#### Proteína alvo: ML2055

A sequência primária de ML2055 mapeada por SANTANA (2017) corresponde a uma proteína com 287 resíduos de aminoácidos identificada no banco de dados do NCBI (*National Center for Biotechnology Information*) como:

gi|499211152|ref|WP\_010908692.1| *alanine and proline-rich secreted protein* Apa [*Mycobacterium leprae*]

#### Análise e seleção do modelo (*template*) para modelagem por homologia

A sequência primária de ML2055 foi submetida a busca por moldes para predição de estrutura tridimensional por meio de alinhamento global utilizando a ferramenta *BLAST* (*Basic Local Alignment Search Tool*) fornecida pelo *Protein data bank* (PDB) e pela mesma ferramenta disponível no Swiss-model (GUEX & PEITSCH, 1997).

Procurou-se por homologia no banco de dados do PDB proteínas já estruturalmente resolvidas de forma a encontrar moldes (*templates*) com sequência

primária com tamanho compatível à proteína alvo (cobertura de sequência), bem como com uma alta identidade de sequência primária e alta resolução.

Preconizou-se ainda o gênero de microrganismo ao qual a sequência do PDB pertencia, bem como o método de resolução de estrutura. As sequências primárias dos moldes selecionados foram alinhadas individualmente com a proteína alvo, ML2055, utilizando o programa ClustalX2 para uma análise mais detalhada de identidade e cobertura de sequência.

### **Modelagem por predição de enovelamento (*Threading*)**

A modelagem molecular por *Threading* foi realizada por meio de ferramentas disponíveis no servidor I-Tasser (ZHANG, 2008). A modelagem foi realizada de modo automatizado, sendo utilizadas as configurações de parâmetros padrão (“default”).

### **Modelagem *ab initio***

O método em questão não depende do conhecimento prévio da estrutura da proteína analisada ou de moldes. Assim, não há uso de base de dados de estruturas experimentais. Essa técnica pode ser descrita como *de novo modelling* (BRADLEY; MISURA; BAKER, 2005), que se baseia no comportamento físico, formação de campos de força e como esses fatores interagem entre si e com o solvente. Para isso utilizou-se o servidor Robetta (<http://robetta.bakerlab.org>), fornecido pela plataforma ROSETTA, cujos dados são avaliados por meio do CAMEO (<https://www.cameo3d.org>) (HAAS, 2019).

### **Análise e refinamento do modelo tridimensional de ML2055**

A qualidade do modelo foi analisada pelos programas da plataforma MolProbity (DAVIS et al., 2007), pelos valores de GMQE (*Global Model Quality Estimation*) e QMEAN Z-score (BENKERT et al., 2009). A compatibilidade entre a estrutura tridimensional e sua sequência primária foi analisada utilizando o programa Verify3D (LUTHY et al., 1997).

### **Análise da porção amino-terminal de ML2055**

A porção amino-terminal da proteína ML2055, correspondente aos 100 primeiros resíduos de aminoácidos, foi analisada quanto às estruturas secundárias possíveis de serem encontradas e presença de padrões de assinatura de sequência e/ou estrutura. Os programas ou plataformas utilizadas são mostrados na Tabela 1.

**Tabela 1.** Programas e plataformas para busca de padrões e/ou assinaturas de sequência.

<b>Plataformas/Programas</b>	<b>Sítio</b>
JPred 4	<a href="http://www.compbio.dundee.ac.uk/jpred/">http://www.compbio.dundee.ac.uk/jpred/</a>
GOR4	<a href="https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html">https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html</a>
PSORT	<a href="https://www.psort.org/psortb/">https://www.psort.org/psortb/</a>
PSIPRED	<a href="http://bioinf.cs.ucl.ac.uk/psipred/">http://bioinf.cs.ucl.ac.uk/psipred/</a>

### Previsão de epítomos lineares para células B na proteína ML2055

A previsão de epítomos da proteína ML2055 indica a existência de sequências lineares que podem ser reconhecidas por receptores de linfócitos B, ou seja, por anticorpos. Logo, tais sequências são consideradas sequências imunogênicas. Os programas utilizados buscaram por sequências de 20 aminoácidos de comprimentos. Os dados obtidos foram confrontados com aqueles obtidos experimentalmente por Santana (2017). Foram utilizados três programas/plataformas, listados na Tabela 2.

**Tabela 2.** Programas e plataformas para busca de epítomos lineares.

Plataformas/Programas	Sítio
Bepipred Linear Epitope Prediction 2.0	<a href="http://tools.iedb.org/bcell/">http://tools.iedb.org/bcell/</a>
SVMTrip	<a href="http://sysbio.unl.edu/SVMTriP/prediction.php">http://sysbio.unl.edu/SVMTriP/prediction.php</a>
ABCpred	<a href="http://crdd.osdd.net/raghava/abcpred/ABC_submission.htm">http://crdd.osdd.net/raghava/abcpred/ABC_submission.htm</a>

### Análise dos peptídeos imunogênicos experimentais de ML2055

Os epítomos experimentalmente testados e relatados por Santana (2017) foram confrontados quanto à sua imunogenicidade. A análise foi feita por meio de uma busca no banco de dados de epítomos - *Immune Epitope Database* (IEDB), o qual reúne dados experimentais sobre epítomos reconhecidos por anticorpos e células T. A busca foi feita utilizando a sequência primária dos epítomos de ML2055, sendo selecionadas as opções para epítomos lineares com até 90% de identidade para bactérias do gênero *Mycobacterium*, reconhecidos pelo organismo humano.

Banco de dados	Sítio
<i>Immune Epitope Database</i>	<a href="http://www.iedb.org/">http://www.iedb.org/</a>

### Resultados e Discussão

A modelagem por homologia para a proteína ML2055 foi a primeira a ser realizada. O BLAST contra o PDB foi realizado e obteve-se como *template* sugerido a estrutura 5ZX9, a proteína secretada Apa rica em alanina e prolina de *Mycobacterium tuberculosis*, um homo-hexâmero. O BLAST do Swiss-model também foi utilizado sugerindo como molde a mesma estrutura. O método de resolução da estrutura 5ZX9 é a difração de raios-X com resolução de 1,55 Å. Além dessas variáveis, foram consideradas a identidade de sequência entre o *template* e a proteína alvo - 141/189 aa (75%), assim como a cobertura - 81 a 269 aa (64%) entre as sequências primárias.

Posteriormente, alinhou-se a sequência primária do molde sugerido com a sequência da proteína alvo por meio do ClustalX2 para analisar a identidade ao longo da sequência proteica como mostra a Figura 1. Como o número de aminoácidos entre as sequências são diferentes, não é levado apenas em consideração o número de resíduos idênticos para se calcular a identidade, mas

também as substituições por resíduos quimicamente e/ou fisicamente semelhantes, presença e tamanho *gaps* (regiões não cobertas pelo molde).

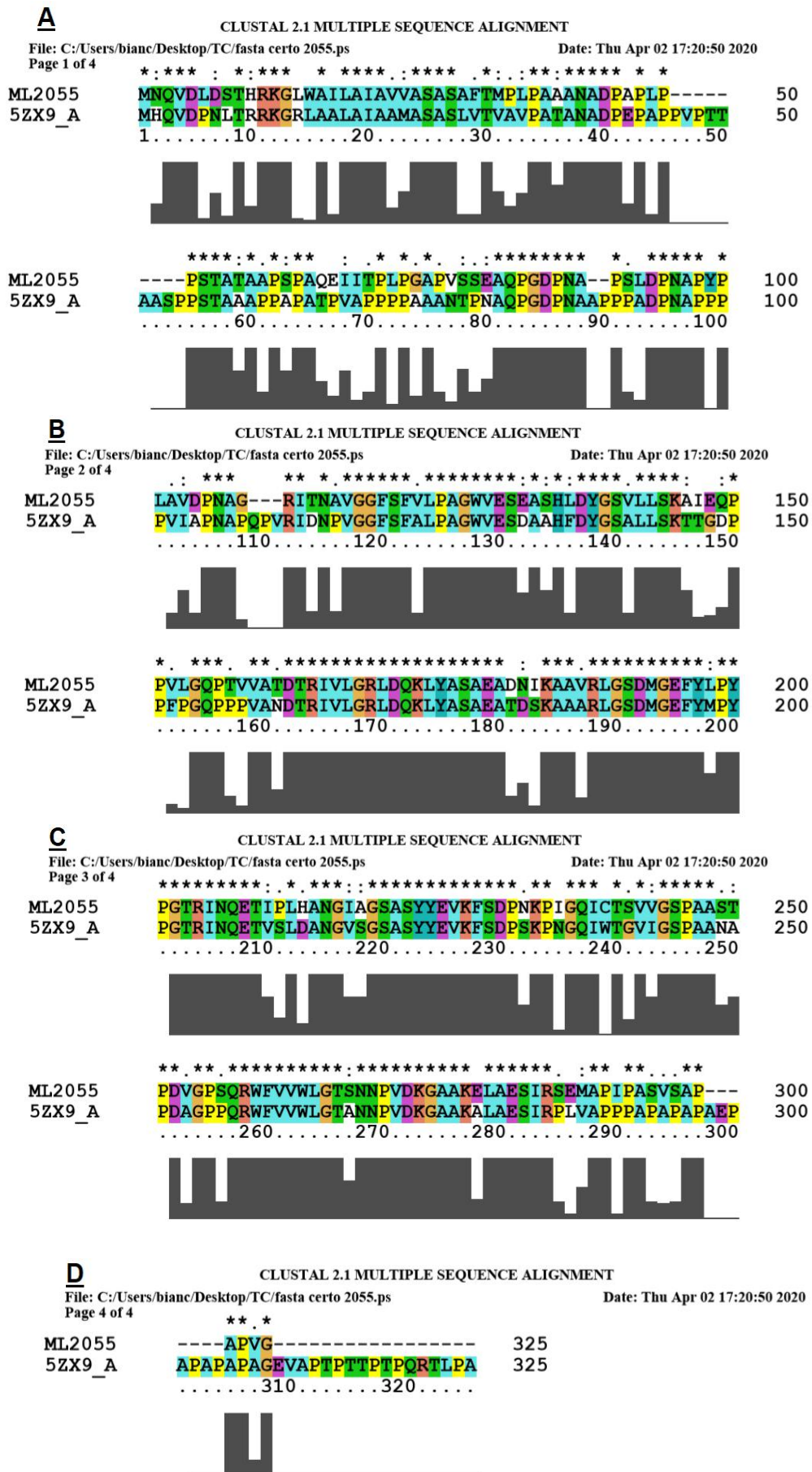
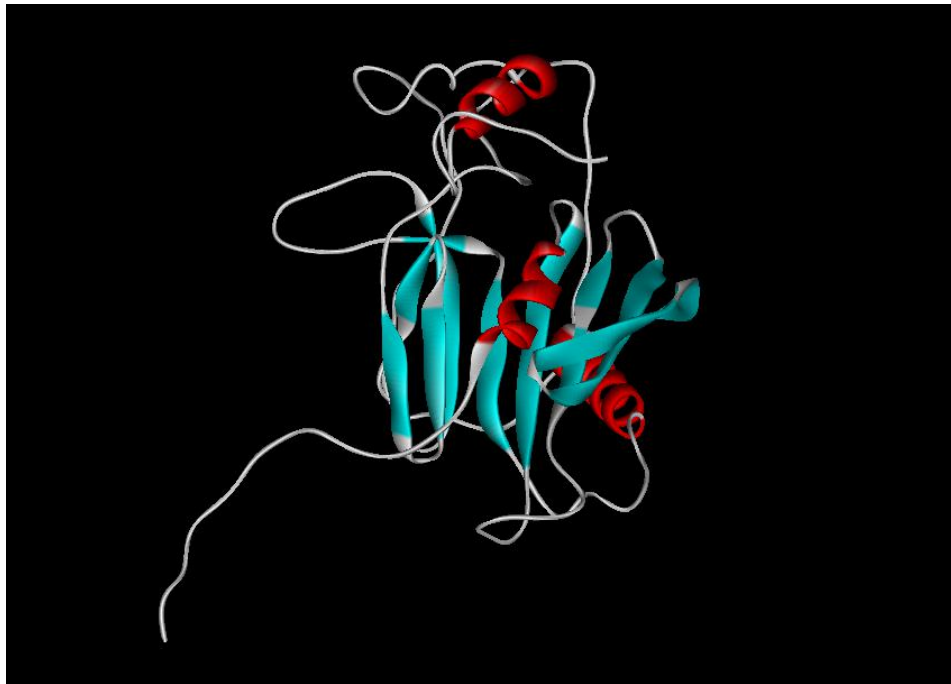


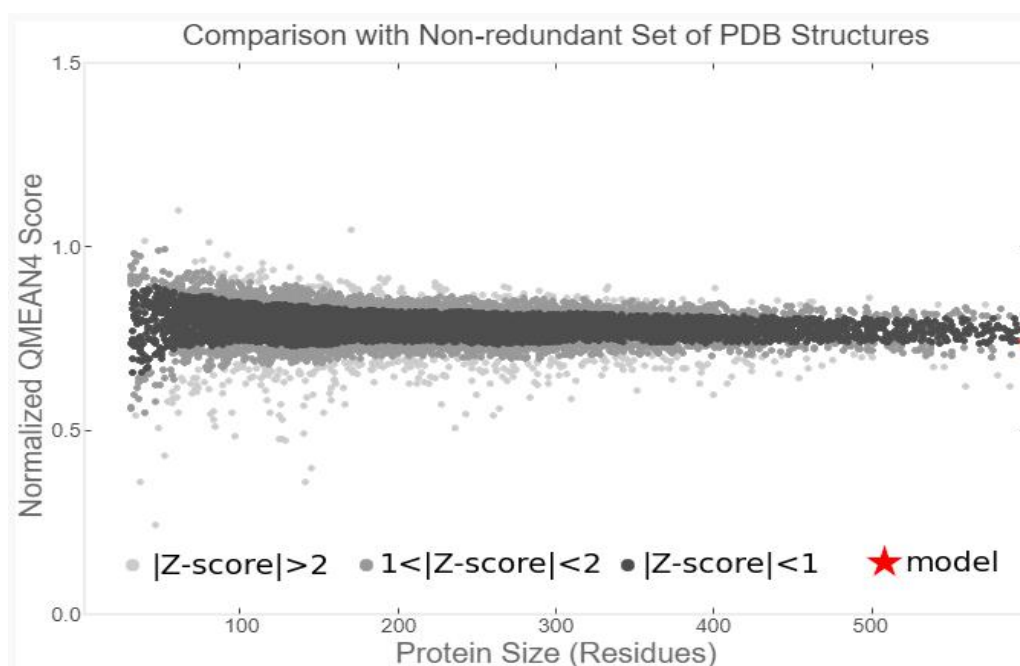
Figura 1. Alinhamento de seqüência primária entre as proteínas ML2055 e 5ZX9.

Quando um molde satisfatório é sugerido e há mais de 25-30% de identidade a técnica de modelagem molecular por homologia pode ser admitida.

O modelo obtido e seus gráficos de *QMEAN Z-score* e Ramachandran são mostrados nas Figuras 2, 3 e 4. O valor de *Q-mean Z-score* foi de -0,64. Esses valores apresentados pelo Swiss-model permitem uma avaliação preliminar da confiabilidade do modelo fornecido. Valores de *QMEAN Z-score* em torno de zero indicam uma boa concordância entre a estrutura do modelo obtido e estruturas experimentais de tamanho semelhante, quando esse valor é igual ou menor que -4,0 indica que o modelo obtido possui baixa qualidade.

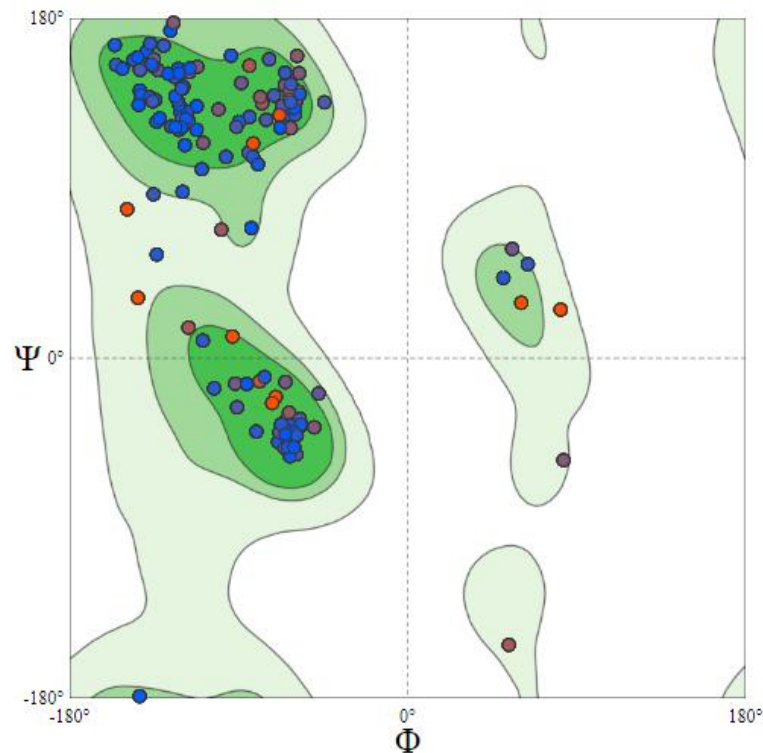


**Figura 2.** Modelo 3D do monômero da proteína ML2055. Em vermelho:  $\alpha$ -hélices; em azul: fitas  $\beta$ ; em branco: *loops*.





**Figura 3.** Gráfico de QMEAN Z-score. A estrela vermelha indica a posição do modelo dentre as estruturas do PDB.

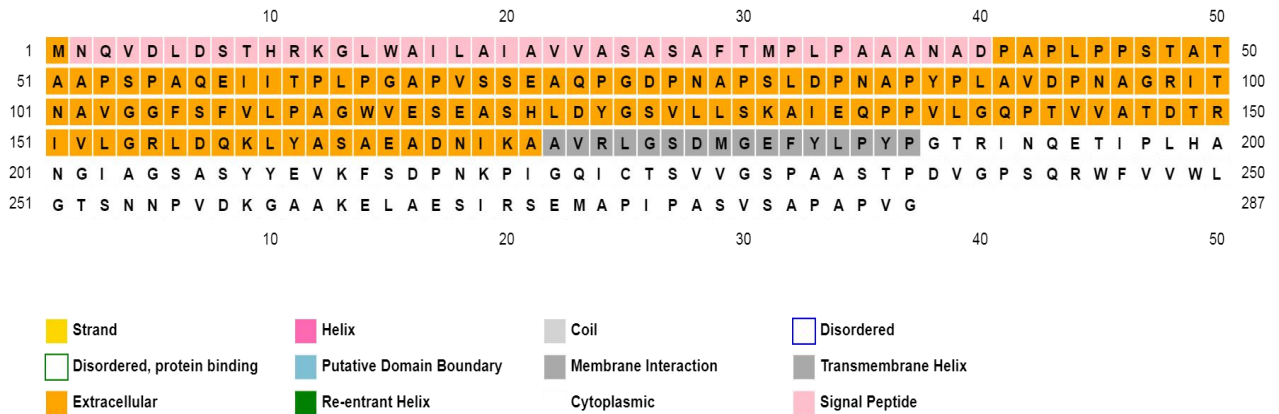


**Figura 4.** Gráfico de Ramachandran. Em verde escuro: regiões muito favoráveis; verde intermediário: regiões favoráveis; verde claro: regiões menos favoráveis; branca: regiões desfavoráveis para os ângulos diédros  $\phi$  e  $\psi$ .

Ao se observar apenas os resultados dos gráficos, valores fornecidos pelo QMEAN Z-score e a porcentagem de identidade pelo alinhamento, o modelo obtido parece ser satisfatório. Entretanto, uma análise mais detalhada é necessária para avaliar outras variáveis pertinentes à confiabilidade do modelo. A ferramenta computacional utilizada para se alcançar esse objetivo foi a plataforma MolProbity.

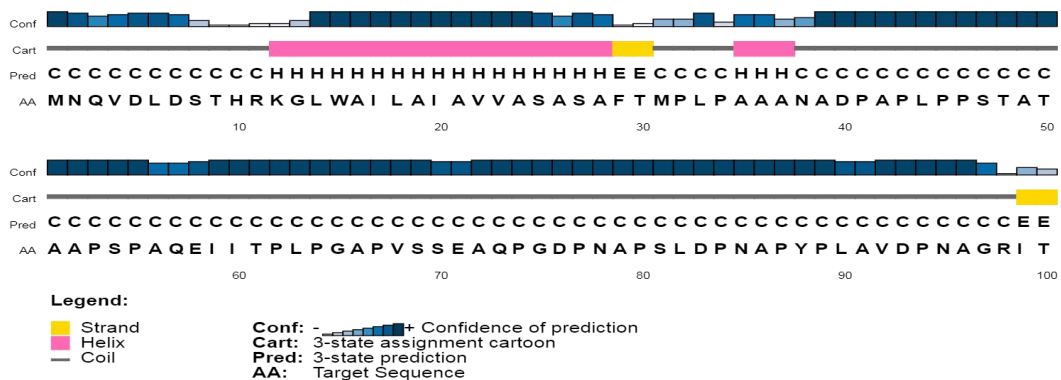
Uma análise de visualização de sequência primária pelo programa ViewerLite indicou que o modelo fornecido pela técnica de modelagem molecular por homologia apresentava um problema, uma vez que a sua porção amino-terminal não estava presente na estrutura. Os 94 primeiros resíduos (incluindo a metionina inicial) não estavam presentes no modelo devido à ausência de cobertura de sequência.

De acordo com o molde homólogo presente no PDB, ML2055 seria uma proteína secretada, logo ela deve apresentar na porção amino-terminal um peptídeo sinal para exportação para o meio extracelular. Os resultados dos quatro programas/plataformas usados para localização de peptídeo sinal indicaram que ML2055 realmente é uma proteína secretada e o peptídeo sinal, provavelmente correspondiam aos primeiros 40 resíduos. A Figura 5 mostra o resultado de um dos programas.



**Figura 5.** Resultado obtido pelo programa PsiPred. O peptídeo sinal, de acordo com a legenda, é mostrado em rosa.

Uma vez que o peptídeo sinal é excisado da proteína funcional, sua retirada resolveria em parte o problema do modelo 3D obtido para ML2055, entretanto ainda faltaria a estrutura de 54 resíduos. Os mesmos quatro programas/plataformas usados para análise de peptídeo sinal foram usados para predição de estrutura secundária, como mostrado na Figura 6.



**Figura 6.** Resultado de predição de estrutura secundária pelo programa PsiPred para os 100 primeiros resíduos de aminoácidos da proteína ML2055.

A Figura 6 mostra que para a porção que se deseja obter dados de estrutura (resíduos 41 a 95) há uma alta confiabilidade para a presença de *coils*. Em modelagem *coil (loop)* se refere a regiões de estrutura secundária sem padrão de repetição definido, diferente da alfa-hélice e folha-beta.

Apesar dessa confiabilidade, o modelo obtido foi descartado devido à impossibilidade de acréscimo desses 54 resíduos à estrutura final.

Como essa primeira tentativa de modelagem molecular não gerou um resultado satisfatório prosseguiu-se com a modelagem utilizando a técnica de *Threading* por meio da plataforma I-Tasser. O I-Tasser constrói cinco possíveis modelos de estrutura baseados em dez estruturas resolvidas de proteínas disponíveis no PDB e fornece um valor de *C-score* para cada modelo. O *C-score* é

um índice de confiabilidade que varia de -5 a 2, e serve para estimar a qualidade dos modelos previstos pela plataforma. Dos cinco modelos sugeridos o modelo 1 (Figura 7), com melhor valor de *C-score*, foi analisado e submetido a validação pelo MolProbity (Quadro 1).

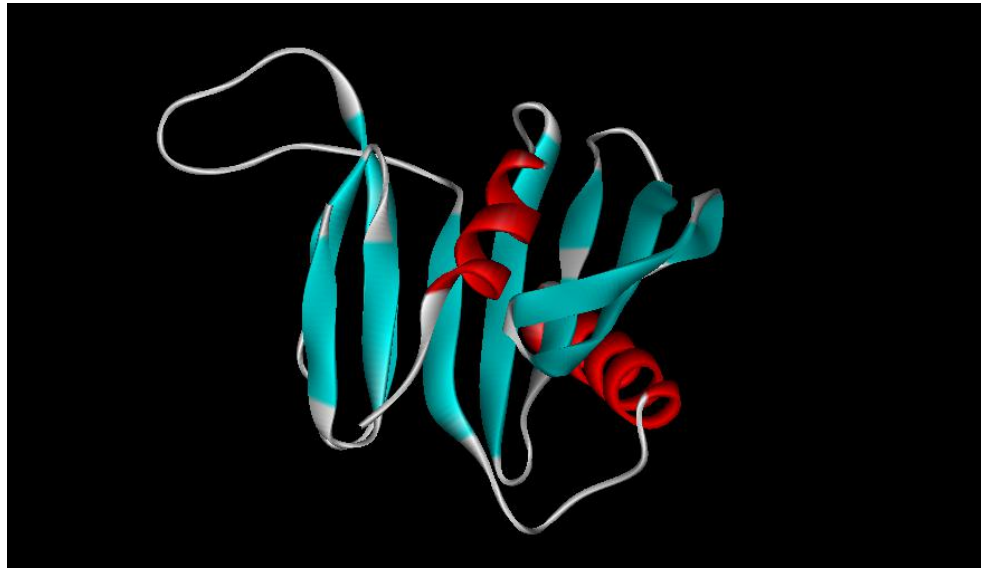


**Figura 7.** Modelo tridimensional da proteína ML2055 obtida pelo I-Tasser. Em vermelho:  $\alpha$ -hélices; em azul: fitas  $\beta$ ; em branco: *loops*.

O modelo tridimensional gerado pela plataforma I-Tasser apresentava a porção amino-terminal faltante no modelo obtido pelo Swiss-model. Entretanto, ele foi invalidado pelos parâmetros do programa MolProbity, que apontou várias inconsistências de posicionamento de resíduos, sugerindo uma má qualidade do modelo.

O método, por fim, utilizado para a construção da estrutura foi a técnica *Ab initio*, utilizando-se o programa Robetta. O molde obtido foi uma estrutura monohexamérica, da qual foi isolada sua cadeia alfa por meio do ViewerLite. A porção amino-terminal faltante no modelo sugerido pelo Swiss-model foi modelada. Este modelo final (Figura 8) foi submetido a duas plataformas de validação a plataforma Molprobity, já usada anteriormente, e ao Verify3D (Quadro 1).

O posicionamento de todos os resíduos da cadeia primária no modelo tridimensional obedeceu às regras físico-químicas e permaneceram em regiões favoráveis ou permitidas para os ângulos diedros  $\phi$  e  $\psi$ . Além disso, provou-se por meio do Verify3D que há compatibilidade entre a estrutura tridimensional sugerida e a sequência primária da proteína ML2055. Assim, o modelo é considerado satisfatório pelos três diferentes métodos de validação, o resultado, entretanto, é visualmente muito semelhante aos demais - já descartados - quando se avalia as regiões de *loops*, fitas  $\beta$  e  $\alpha$ -hélices.



**Figura 8.** Modelo final sugerido pelo Robetta. Em vermelho:  $\alpha$ -hélices; em azul: fitas  $\beta$ ; em branco: *loops*.

Técnica de modelagem	Programas/ Plataformas	Molprobit		Referência	Verify 3D	Referência	Ramachandram		Referência
<i>Ab initio</i>	Robetta	<i>Poor rotamers</i>	0%	< 0,3%	81,53%	80% de resíduos com score =< 0,2	Resíduos em regiões favoráveis	99,30%	98%
		<i>Favoured rotamers</i>	100,00%	> 98%			Resíduos em regiões permitidas	100%	> 99,8%
		<i>Bad bonds</i>	0%	0%			Resíduos em regiões favoráveis	57,89%	98%
		<i>Bad angles</i>	0,14%	< 0,1%			Resíduos em regiões permitidas	76,10%	> 99,8%
<i>Theading</i>	I-Tasser	<i>Poor rotamers</i>	17,04%	< 0,3%	82,93%	80% de resíduos com score =< 0,2	Resíduos em regiões favoráveis	99,30%	98%
		<i>Favoured rotamers</i>	65,02%	> 98%			Resíduos em regiões permitidas	100%	> 99,8%
		<i>Bad bonds</i>	0,61%	0%			Resíduos em regiões favoráveis	57,89%	98%
		<i>Bad angles</i>	5,32%	< 0,1%			Resíduos em regiões permitidas	76,10%	> 99,8%

**Quadro 1.** Resultado da validação dos modelos de estrutura gerados pelo I-Tasser e pelo Robetta.

O resultado da previsão dos epítomos lineares potencialmente reconhecidos por células B, ou seja, anticorpos são mostrados na Figura 9.

ML2055		100 resíduos/linha
Programa	Epít./score <sup>1</sup>	MNQVDLSDTHRKGLWAILAIIVVASASAFMTPLPAAANADPAPLPPSTATAAPSPAQEIIITPLPGA PVSSEAQPGDPNAPS LDPNAPYPLAVDPNAGRIT
ABCPred	4 (> 0,83)	PSPAQEIIITPLPGA PVSSEAQPGDPNAPS LDPNAPYPLAVDPNAGRIT
SVMTriP	0	
Bepipred 2	0	

ML2055		100 resíduos/linha
Programa	Epít./score <sup>1</sup>	NAVGGFSFVLPAGWVVESEASHLDYGSVLLSKAIEQPPVLGQPTVVATDTRIVLGRDLQKLYASAEADNIKAAVRLGSDMGFEFYPYGPTRINQETIPLHA
ABCPred	3 (> 0,83)	NAVGG FVLPAGWVVESEASHLDYGSV MGEFYLPGYGPTRINQETIPLHA
SVMTriP	2 (> 0,83)	ESEASHLDYGSVLLSKAIEQ GSDMGFEFYPYGPTRINQET
Bepipred 2	3 (> 0,50)	WVVESEASHLDY 11* LSKAIEQPPVLGQPTVVAT 19* YASAEADNIKA AVR 14*

ML2055		87 resíduos/linha
Programa	Epít./score <sup>1</sup>	NGIAGSASYEVKFSDPNPKPIGQICTSVVGSAASTPDVGPQRWFVWVLGTSNNPVDKGA AKELAESIRSEM APIPASVSAPAPVG
ABCPred	3 (> 0,83)	NGIAG GPSQRWFVWVLGTSNNPVDKGA AKELAESIRSEM APIPASVSAPAPVG
SVMTriP	0	
Bepipred 2	2 (> 0,50)	AASTPDVGPS 10* RSEM APIPASVSAP 14*

1. Número de epítotos encontrados pelo programa e seu respectivo score. O score corresponde à probabilidade do peptídeo ser imunogênico para resposta por LB.

\*Número de resíduos para o Bepipred 2.0; para os demais programas o padrão são 20 resíduos/epítoto.

Amarelo: peptídeo imunogênico (experimental).

Resíduos em branco nas regiões em verde: compartilhados por 2 ou mais epítotos.

**Figura 9.** Resultados obtidos a partir dos programas ABCPred, SVMTriP e Bepipred 2.0 para os epítotos potencialmente imunogênicos. O epítoto PVSSEAQPGDPNAPS tem seus resíduos como parte de diferentes epítotos; os epítotos AQP GDPNAAPPPADPNAPPP e PPPAAANTPNAQPGDPNAAP da proteína Apa rica em alanina de *M. tuberculosis* são imunogênicos de acordo com IEDB.

O epítoto selecionado experimentalmente por Santana (PVSSEAQPGDPNAPS) foi identificado como parte de resíduos compartilhados por diferentes epítotos. Corroborando a hipótese de que estes sejam realmente reconhecidos pelo sistema imune, dois epítotos a partir de uma proteína secretada Apa rica em alanina e prolina (AQP GDPNAAPPPADPNAPPP e PPPAAANTPNAQPGDPNAAP) de *Mycobacterium tuberculosis* são identificados pelo *Immune Epitope Database* como imunogênicos e conservam parte dos resíduos identificados por Santana e pelos programas utilizados neste artigo.

## Conclusão

O modelo obtido nesse trabalho mostra que mesmo que se possua alta identidade de sequência existem outros parâmetros a serem obedecidos para que o modelo escolhido seja satisfatório, revelando assim, que a análise superficial e apenas visual pode levar a escolha de modelos tridimensionais errôneos que não respeitam as regras físico-químicas de interação entre os aminoácidos da cadeia e seu comportamento *in vivo*. Além disso, por mais que não seja o ideal e esperado, regiões de *coil* são permitidas dentro da molécula desde que os mesmos parâmetros já citados sejam obedecidos.

Com o modelo obtido será possível localizar espacialmente os peptídeos experimentalmente imunogênicos e tentar estabelecer uma relação com seu potencial de imunogenicidade, sua localização na proteína e suas características químicas. Dessa forma, há a possibilidade de se desenvolver ensaios clínicos para obtenção de um teste sorológico eficaz para diagnóstico das diferentes formas da doença, bem como acessível as diferentes realidades do país.

### **3 CONSIDERAÇÕES FINAIS**

A utilização dos dados gerados através da pesquisa em estudos posteriores para o desenvolvimento de testes sorodiagnósticos, aparentemente é plausível. A grande limitação do trabalho foi resolver a estrutura da ML2055 apenas por métodos de bioinformática por falta de equipamentos, quantidade substancial da proteína a ser cristalizada, bem como a análise dos dados que seriam obtidos a partir da difração de raio-x, como é realizado no Laboratório Nacional de Luz Sincontron com sede em Campinas.

Além disso, outra dificuldade está no fato de que há uma quantidade limitada de informações sobre as proteínas identificadas por Santana, uma vez que se trata de uma descoberta recente.

## REFERÊNCIAS

- ANTUNES, S.; JARDIM, M. R.; VITAL, R. T.; et al. Fibrosis: a distinguishing feature in the pathology of neural leprosy. **Memorias do Instituto Oswaldo Cruz**, v. 114, n. 6, p. 1–8, 2019.
- BENKERT, P.; KÜNZLI, M.; SCHWEDE, T. QMEAN server for protein model quality estimation. **Nucleic Acids Research**, v. 37, n. SUPPL. 2, p. 510–514, 2009.
- BLOOM, B. R.; GROSSET, J. H.; LECHAT, M. F.; et al. Who Expert Committee on Leprosy. **WHO Technical Report Series**, , n. 768, p. 1–61, 2012.
- BRADLEY, P.; MISURA, K. M. S.; BAKER, D. Biochemistry: Toward high-resolution de novo structure prediction for small proteins. **Science**, v. 309, n. 5742, p. 1868–1871, 2005.
- BRASIL. Guia prático sobre a hanseníase. Secretaria de Vigilância em Saúde Departamento de Vigilância e Doenças Transmissíveis. **Anais...** . p.68, 2017. Disponível em: <[http://bvsmms.saude.gov.br/bvs/publicacoes/guia\\_pratico\\_hanseniose.%0Apdf](http://bvsmms.saude.gov.br/bvs/publicacoes/guia_pratico_hanseniose.%0Apdf)> .
- DAVIS, I. W.; LEAVER-FAY, A.; CHEN, V. B.; et al. MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 375–383, 2007.
- DEGANG, Y.; NAKAMURA, K.; AKAMA, T.; et al. Leprosy as a model of immunity. **Future Microbiology**, v. 9, n. 1, p. 43–54, 2014.
- DUTHIE, M. S.; BALAGON, M. F.; MAGHANOY, A.; et al. Rapid quantitative serological test for detection of infection with *Mycobacterium leprae*, the causative agent of leprosy. **Journal of Clinical Microbiology**, v. 52, n. 2, p. 613–619, 2014.
- GUEx, N.; PEITSCH, M. C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. **Electrophoresis**, v. 18, n. 15, p. 2714–2723, 1997.
- HAAS, J.; GUMIENNY, R.; BARBATO, A.; et al. Introducing “best single template” models as reference baseline for the Continuous Automated Model Evaluation (CAMEO). **Proteins: Structure, Function and Bioinformatics**, v. 87, n. 12, p. 1378–1387, 2019.
- LASTÓRIA, J. C.; DE ABREU, M. A. M. M. Leprosy: A review of laboratory and therapeutic aspects - Part 2. **Anais Brasileiros de Dermatologia**, v. 89, n. 3, p. 389–401, 2014.
- LEGENDRE, D. P.; MUZNY, C. A.; SWIATLO, E. Hansen’s disease (Leprosy): Current and future pharmacotherapy and treatment of disease-related immunologic reactions. **Pharmacotherapy**, v. 32, n. 1, p. 27–37, 2012.
- LUTHY, R.; BOWEI, J.; EINSENBURG, D. Verify3D: Assessment of protein models with three-dimensional profiles. **Methods in enzymology**, v. 277, p. 396–404, 1997.
- MOYLE, P. M.; TOTH, I. Modern Subunit Vaccines: Development, Components, and Research Opportunities. **ChemMedChem**, v. 8, n. 3, p. 360–376, 2013.

NATH, I.; SAINI, C.; VALLURI, V. L. Immunology of leprosy and diagnostic challenges. **Clinics in Dermatology**, v. 33, n. 1, p. 90–98, 2015. Elsevier Inc. Disponível em: <<http://dx.doi.org/10.1016/j.clindermatol.2014.07.005>>. .

PARREN, PAUL W. H. I.; POIGNARD, PASCAL; DITZEL; HENRIK J.; WILLIAMSON, R. ANTHONY; BURTON, D. R. Antibodies in Human Infectious Disease. **Journal of Visual Languages & Computing**, v. 11, n. 3, p. 55, 2559. Disponível em: <[https://www.m-culture.go.th/mculture\\_th/download/king9/Glossary\\_about\\_HM\\_King\\_Bhumibol\\_Adul\\_yadej's\\_Funeral.pdf](https://www.m-culture.go.th/mculture_th/download/king9/Glossary_about_HM_King_Bhumibol_Adul_yadej's_Funeral.pdf)>. .

PELLEQUER, JEAN-LUC; WESTHOF, ERIC; VAN REGENMORTEL, M. H. V. Correlation between the location of antigenic sites and the prediction of turns in proteins. **Immunology Letters**, v. 36, p. 83–100, 1993.

PEREIRA, HELENA LÚCIA ALVES; RIBEIRO, SANDRA LÚCIA EUZÉBIO; CICONELLI, ROZANA MESQUITA; FERNANDES, A. DA R. C. Imaging Methods Evaluation in Osteoarticular and Peripheral Nerves Involvement in Leprosy. .

REIBEL, F.; CAMBAU, E.; AUBRY, A. Update on the epidemiology, diagnosis, and treatment of leprosy. **Medecine et Maladies Infectieuses**, v. 45, n. 9, p. 383–393, 2015. Elsevier Masson SAS. Disponível em: <<http://dx.doi.org/10.1016/j.medmal.2015.09.002>>. .

RODRIGUES, L. C.; LOCKWOOD, D. N. J. Leprosy now: Epidemiology, progress, challenges, and research gaps. **The Lancet Infectious Diseases**, v. 11, n. 6, p. 464–470, 2011. Elsevier Ltd. Disponível em: <[http://dx.doi.org/10.1016/S1473-3099\(11\)70006-8](http://dx.doi.org/10.1016/S1473-3099(11)70006-8)>. .

ŠALI, A.; BLUNDELL, T. L. 1-s2.0-S0022283683716268-main.pdf. **Journal of Molecular Biology**, v. 234, p. 779–815, 1993.

SANTANA, J. F. DE. **Mapeamento de epítomos imunodominantes de antígenos de Mycobacterium leprae: caracterização in vitro, in vivo e in silico.**, 2017. UNIVERSIDADE FEDERAL DO PARANÁ.

SILVA, E. A.; IYER, A.; URA, S.; et al. Utility of measuring serum levels of anti-PGL-I antibody, neopterin and C-reactive protein in monitoring leprosy patients during multi-drug treatment and reactions. **Tropical Medicine and International Health**, v. 12, n. 12, p. 1450–1458, 2007.

SMITH, W. CAIRNS S.; SMITH, CHRISTINE M.; CREE, IAN A.; JADHAV, RUPRENDRA S.; MACDONALD, MURDO; EDWARD, VIJAY K.; OSKAM, LINDA; BEERS, STELLA VAN; KLATSER, P. An Approach to Understanding the Transmission of Mycobacterium leprae Using Molecular and Immunological Methods: Results from the MILEP2 Study. **International Journal of Leprosy**, v. 72, p. 269–277, 2004.

SMITH, C. S.; NOORDEEN, S. K.; RICHARDUS, J. H.; et al. A strategy to halt leprosy transmission. **The Lancet Infectious Diseases**, v. 14, n. 2, p. 96–98, 2014.

SPENCER, J. S.; DUTHIE, M. S.; GELUK, A.; et al. Identification of serological biomarkers of infection, disease progression and treatment efficacy for leprosy. **Memorias do Instituto Oswaldo Cruz**, v. 107, n. SUPPL.1, p. 79–89, 2012.



TECNOLOGIA, D. DE C. E. Neglected diseases: the strategies of the Brazilian Ministry of Health. Neglected diseases: the strategies of the Brazilian Ministry of Health. **Anais...** . p.1–3, 2010. Rev Saúde Pública.

VIRMOND, M.; GRZYBOWSKI, A. Leprosy-An intriguing disease. **Clinics in Dermatology**, v. 33, n. 1, p. 1–2, 2015. Elsevier Inc. Disponível em: <<http://dx.doi.org/10.1016/j.clindermatol.2014.10.002>>. .

VISSA, V. D.; BRENNAN, P. J. The genome of *Mycobacterium leprae*: A minimal mycobacterial gene set. **Genome Biology**, v. 2, n. 8, p. 1–8, 2001.

WHITE, C.; FRANCO-PAREDES, C. Leprosy in the 21st century. **Clinical Microbiology Reviews**, v. 28, n. 1, p. 80–94, 2015.

WILSON, P. C.; ANDREWS, S. F. Tools to therapeutically harness the human antibody response. **Nature Reviews Immunology**, v. 12, n. 10, p. 709–719, 2012. Nature Publishing Group.

ZHANG, Y. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, v. 9, p. 1–8, 2008.

ZHU, L.; ZHANG, Y.; TEH, J. S.; et al. Characterization of mRNA interferases from *Mycobacterium tuberculosis*. **Journal of Biological Chemistry**, v. 281, n. 27, p. 18638–18643, 2006.

## APÊNDICE 1 – ANÁLISE E REFINAMENTO DO MODELO 3D

Modelagem molecular		Análise e refinamento do modelo 3D								
		Programas/Plataformas								
Técnica de modelagem	Programas/ Plataformas	Molprobit		Referência	Verify 3D	Referência	Ramachandram		Referência	
ML2055	Comparativa/homologia	SwissModel	<i>Bad rotamers</i>	0%	< 0,3%	91,57%	80% de resíduos com score =< 0,2	Resíduos em regiões favoráveis	95,50%	98%
			<i>Bad rotamers</i>	98,60%	> 98%			Resíduos em regiões permitidas	100%	> 99,8%
			<i>Bad bonds</i>	0%	0%			Resíduos em regiões favoráveis	99,30%	98%
			<i>Bad angles</i>	0,27%	< 0,1%			Resíduos em regiões permitidas	100%	> 99,8%
	<i>Ab initio</i>	Robetta	<i>Bad rotamers</i>	0%	< 0,3%	81,53%	80% de resíduos com score =< 0,2	Resíduos em regiões favoráveis	99,30%	98%
			<i>Bad rotamers</i>	100,00%	> 98%			Resíduos em regiões permitidas	100%	> 99,8%
			<i>Bad bonds</i>	0%	0%					
			<i>Bad angles</i>	0,14%	< 0,1%					

### ANEXO 1 – SEQUÊNCIAS PRIMÁRIAS

>gi|499211152|ref|WP\_010908692.1| alanine and proline-rich secreted protein Apa  
[Mycobacterium leprae]

MNQVDLDSTHRKGLWAILAIAVVASASAFSTMPLPAAANADPAPLPPSTATAAPSPAQ  
EIITPLPGAPVSSEAQPGDPNAPSLDPNAPYPLAVDPNAGRITNAVGGFSFVLPAGW  
VESEASHLDYGSVLLSKAIEQPPVLGQPTVVATDTRIVLGRLDQKLYASAEADNIKAA  
VRLGSDMGEFYLPYPGTRINQETIPLHANGIAGSASYEYVKFSDPNKPIGQICTSVVG  
SPAASTPDVGPSQRWFVWVLGTSSNNPVDKGAAKELAESIRSEMAPIPASVSAPAPV  
G

>5ZX9\_1|Chain A|Alanine and proline-rich secreted protein Apa|Mycobacterium  
tuberculosis (strain ATCC 25618 / H37Rv) (83332)

MHQVDPNLTRRKGRLLAALIAAMASASLVTVAVPATANADPEPAPPVPTTAASPPS  
TAAAPPAPATPVAPPPAAANTPNAQPGDPNAAPPPADPNAPPPVIAPNAPQPVR  
IDNPVGGFSFALPAGWVESDAAHFDYGSALLSKTTGDPPFPGQPPPVANDTRIVLG  
RLDQKLYASAEATDSKAAARLGSDMGEFYMPYPGTRINQETVSLDANGVSGSASY  
YEVKFSDPSKPNGQIWTGVIGSPAANAPDAGPPQRWFVWVLGTANNPVDKGAACA  
LAESIRPLVAPPPAPAPAPAEPAPAPAPAGEVAPTPTTPQRTLPA

## ANEXO 2 – REGRAS DA REVISTA

### SEÇÕES DO MANUSCRITO

Apresentamos abaixo os requisitos gerais para todos os tipos de manuscritos a serem submetidos à revista Medicina (Ribeirão Preto).

#### 1. Folha de rosto / Página inicial

Incluir os itens abaixo:

a. Título do artigo no idioma da submissão e em língua inglesa – sucintos, chamativos e representativos do conteúdo do manuscrito (não há um limite rígido para o tamanho do título);

b. Título abreviado/resumido no idioma da submissão (limitado a até 50 caracteres incluindo letras e espaços);

c. Nome completo dos autores, separados por vírgula, na ordem em que devem aparecer na versão final, com indicação de afiliação por meio de número sobrescrito;

d. Afiliação (Instituição, cidade, estado, país) dos autores, precedidos de números sobrescritos correspondentes;

e. Titulação (título máximo de pós-graduação, atividade profissional e/ou acadêmica relevante) e URL completa do registro no ORCID iD;

f. Indicação sobre as contribuições específicas de cada autor para o trabalho submetido, inserindo as iniciais dos autores envolvidos em cada uma das tarefas listadas;

g. Informar as fontes de apoio ou financiamento que tenham contribuído para o desenvolvimento do trabalho;

h. Na seção de agradecimentos incluir os “não autores” Informando abaixo do seu(s) nome(s) uma breve descrição das contribuições específicas à pesquisa.

Observação: Será considerado para indexação o vínculo profissional mais forte para autores com mais de uma afiliação institucional.

## **2. Resumos**

Os autores precisam garantir que o resumo represente com precisão o conteúdo do texto.

Artigos originais, revisões sistemáticas e metanálises requerem resumos estruturados (objetivos, métodos, resultados e conclusões); ensaios clínicos incluirão os itens que o grupo CONSORT identifica como essencial. Relatos de caso requerem resumos “não-estruturados”, que descrevem de forma sucinta e objetiva o artigo, sua importância clínica, desfechos e conclusão sumária em um único parágrafo.

Recomendamos o limite de até 400 palavras no resumo.

## **3. Palavras-chave / Descritores**

Incluir, após o resumo e abstract, entre 3 e 5 palavras-chave e keywords separadas por ponto e vírgula. Os descritores deverão ser obtidos, obrigatoriamente, no vocabulário DeCS e/ou MeSH.

## **4. Texto (padronização)**

O texto de um manuscrito só pode ser aceito como um arquivo do Microsoft Word criado com o MS Word como um documento ".doc" ".docx" ou ".rtf".

No arquivo da versão a ser submetida use de preferência a fonte Arial, tamanho 12 com espaçamento simples; inclua o título, resumo e palavras-chave (no idioma da submissão e em língua inglesa), texto (estruturado conforme a categoria do manuscrito contendo as tabelas e figuras) e referências.

Não identifique no corpo do texto informações dos autores e instituição à qual pertencem de forma a assegurar uma revisão cega.

Observar os quantitativos de palavras recomendadas do texto principal no quadro de resumo.

## **5. Tabelas**

As tabelas capturam informações de maneira concisa e as exibem eficientemente e não deverão conter dados previamente informados no texto.

Sugerimos limitar o número máximo de tabelas conforme o quadro de resumo.

São inseridas no texto principal, numerando-as sequencialmente, juntamente com seus títulos e enviadas no Word (.doc) ou Excel (.xls), não como uma imagem.

O título da tabela deve ser claro, explicativo e deve ser colocado acima da mesma, no canto superior esquerdo, logo após a palavra “Tabela” acompanhada de sua numeração (Tabela 1, Tabela 2, etc).

Os marcadores (\*, †, ‡, §, //, ¶, #, \*\*, ††, etc) são indicados no rodapé da tabela.

Linhas verticais e diagonais não devem ser usadas em tabelas; em vez disso, recuo e espaço vertical ou horizontal devem ser usados para agrupar dados.

## **6. Figuras (fotografia, gráfico, imagem entre outros)**

As Figuras incluem ilustrações, imagens, esquemas ou qualquer outro elemento gráfico que não seja uma tabela e devem ser numeradas de forma sequencial com chamadas no texto.

Sugerimos limitar o número máximo de figuras conforme o quadro de resumo.

O título da figura deve fornecer explicação de maneira concisa de forma que o leitor compreenda do que se trata sem necessidade de se remeter ao texto. As legendas acrescentam informações sobre aspectos das figuras que necessitem detalhamento.

Observe que não deve haver título na parte superior da figura.

É de responsabilidade do(s) autor(es) obter permissão do detentor dos direitos autorais para reproduzir figuras (ou tabelas) que foram publicadas anteriormente em outros lugares.

Se forem usadas fotografias de pessoas, os sujeitos não devem ser identificáveis ou suas fotografias devem estar acompanhadas por consentimento escrito.

Utilizar figuras em alta resolução (mínimo de 300 dpi), nos formatos JPG, GIF ou TIFF. Caso sejam enviadas figuras com resolução inadequada, os autores serão convidados a substituí-la por outra de melhor qualidade.

## 7. Abreviações e nomenclaturas

Quando se mencionar pela primeira vez um termo a ser abreviado, deve-se descrevê-lo inteiramente e a seguir escrever a abreviação apropriada entre parênteses. Não use abreviações no título e limite seu uso no resumo e no texto.

Encorajamos o uso de unidades SI: s por segundo; min por minuto; h por hora; L por litro; m por metro; nomes de espécies (por exemplo, *Homo sapiens*), genes, mutações, genótipos e alelos devem estar em itálico.

As medidas de comprimento, altura, peso e volume devem ser relatadas em unidades métricas (metro, quilograma ou litro) ou em seus múltiplos decimais.

## 8. Referências

Os autores devem fornecer referências diretas de fontes originais de pesquisas sempre que possível. Listas menores de trabalhos “originais-chave” muitas vezes servem tão bem quanto listas mais exaustivas de referências.

As citações ao longo do texto são feitas por números em sobrescrito antes da pontuação, sem espaços:

De acordo com Moura<sup>15</sup>, o índice de vacinação...

...de forma significativa<sup>1</sup>. (detalhe importante: observar que o ponto é inserido após a citação).

...foi descrita considerando alérgenos inalados<sup>22</sup>, na ocasião... (a vírgula deverá ser inserida após a citação).

Se duas ou mais referências forem citadas em sequência, apenas a primeira e a última devem ser digitadas, separadas por um traço (exemplo:<sup>7-11</sup>).

As referências ficam organizadas de acordo com a ordem em que são citadas no texto e padronizadas no estilo Vancouver.

Sugerimos limitar o número máximo de referências conforme o quadro de resumo.

#### Quadro de resumo

Tipo artigo	de N° autores	Título de (n° max. de Caracteres)	Resumo (n° máximo de palavras)	Texto principal (n° máximo de palavras)	N° de tabelas e figuras	de Referências
Artigo original	8	Sem limite	400	7000	8	40
Artigo de revisão	6	Sem limite	400	7000	8	80
Relato de caso	6	Sem limite	400	3000	6	20
Demais categorias	6	Sem limite	400	7000	10	50

Observação: estes valores indicados no quadro servem como referências para a redação dos artigos. Caso haja necessidade, os autores devem justificar em carta aos editores, quaisquer mudanças desses parâmetros.

OBS: Essas são as regras atuais de submissão disponíveis na revista para conferência e não correspondem ao modelo aplicado no artigo, uma vez que este foi finalizado em julho de 2020.