

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE CIÊNCIAS FARMACÊUTICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FARMÁCIA (FISIOPATOLOGIA E
TOXICOLOGIA)
ÁREA DE TOXICOLOGIA

THABATA CORAZZA NAVARRO VINHA

Aplicação de ferramentas de quimioinformática na comparação da estrutura tridimensional da dihidrofolato redutase expressa em diferentes organismos

São Paulo
2023

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE CIÊNCIAS FARMACÊUTICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FARMÁCIA (FISIOPATOLOGIA E
TOXICOLOGIA)
ÁREA DE TOXICOLOGIA

THABATA CORAZZA NAVARRO VINHA

**Aplicação de ferramentas de quimioinformática na comparação da
estrutura tridimensional da dihidrofolato redutase expressa em
diferentes organismos**

Versão Original

Dissertação para obtenção do título de mestre na Faculdade de Ciências Farmacêuticas, Programa de Pós-graduação em Farmácia (Fisiopatologia e Toxicologia)

Área de concentração: Toxicologia

Orientador: Prof. Dr. Gustavo Henrique Goulart Trossini

São Paulo
2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:

Marlene Aparecida Vieira - CRB - 8/5562

V766a	Vinha, Thabata Aplicação de ferramentas de quimioinformática na comparação da estrutura tridimensional da dihidrofolato redutase expressa em diferentes organismos / Thabata Vinha. - São Paulo, 2023. 93 p. Dissertação (mestrado) - Faculdade de Ciências Farmacêuticas da Universidade de São Paulo. Departamento de Farmácia. Orientador: Trossini, Gustavo 1. Quimioinformática. 2. Dihidrofolato Redutase. 3. Planejamento de Fármacos. I. T. II. Trossini, Gustavo, orientador.
-------	---

Thabata Corazza Navarro Vinha

Aplicação de ferramentas de quimioinformática na comparação da estrutura tridimensional da dihidrofolato redutase expressa em diferentes organismos; dissertação para obtenção do título de mestre na Faculdade de Ciências Farmacêuticas, Programa de Pós-Graduação em Farmácia (Fisiopatologia e Toxicologia); área de concentração Toxicologia.

Comissão Julgadora

Prof. Dr. Gustavo Henrique Goulart Trossini
Orientador / Presidente

1o. Examinador

2o. Examinador

3o. Examinador

4o. Examinador

São Paulo, ____ de _____ de 2024.

RESUMO

VINHA, T. C. N. Aplicação de ferramentas de quimioinformática na comparação da estrutura tridimensional da dihidrofolato redutase expressa em diferentes organismos. 2023. 94f. Dissertação (Mestrado) - Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, 2023.

A quimioinformática, definida como o emprego de técnicas informáticas na solução de problemas da química, evolui em conjunto com o desenvolvimento de ferramentas computacionais e é de grande relevância para o planejamento racional de fármacos ao otimizar etapas do desenvolvimento de novas moléculas e economizar recursos e tempo. Dentre as técnicas disponíveis destacam-se o planejamento de fármacos baseado na estrutura e no ligante, que quando combinadas auxiliam na identificação e otimização de moléculas ativas frente a alvos farmacológicos. A Dihidrofolato Redutase (DHFR) é uma importante enzima da via dos folatos que catalisa a redução do dihidrofolato em tetrahydrofolato, utilizando NADPH como cofator, reação essencial para a replicação celular, visto que este ciclo resulta na síntese de precursores das bases nitrogenadas que compõem o DNA, conseqüentemente, inibidores de DHFR são utilizados no tratamento de infecções bacterianas e alguns tipos de câncer. *Trypanosoma cruzi*, protozoário causador da doença de Chagas, é um dos organismos que expressam a DHFR, além do próprio *Homo sapiens*. Analisaram-se ligantes conhecidos e as estruturas da proteína expressa pelos dois organismos, visando identificar pontos de divergência que possam ser explorados no planejamento de moléculas seletivas para o tratamento da doença de Chagas. Os 6 modelos cristalográficos de *T. cruzi* e 2 de *H. sapiens* foram obtidos do banco de dados de proteínas (PDB) após aplicação de filtros de qualidade. Foram analisadas as sequências de aminoácidos dos modelos, com o uso do Cluster Ômega, sua estrutura tridimensional com os programas Pymol e Chimera X, além da análise das cavidades proteicas com o CavityPlus, que também gerou os farmacóforos de ambos alvos. A análise de estrutura primária identificou mutações em três aminoácidos nos cristais do parasita, que podem ser explicados por diferentes caminhos evolutivos de grupos segregados, embora nenhuma mutação observada esteja em regiões de sítio ativo. A análise dos modelos permitiu que fossem identificados os 25 aminoácidos que estão a menos de 5 Å de distância dos ligantes de *T. cruzi*, sendo 5 aminoácidos responsáveis por interações de hidrogênio com pelo menos um dos ligantes analisados. Destes, 18 se repetem na proteína humana ou são substituídos por outro aminoácido que mantém a mesma interação. Quanto às diferenças observadas, destacam-se a asparagina 44 substituída por uma prolina na proteína humana e a prolina 92, substituída por uma lisina. A análise de cavidades identificou três cavidades em cada proteína, embora somente as cavidades correspondentes ao sítio ativo sejam *druggables*. A cavidade da proteína humana é maior e mais alongada, além de apresentar o aspecto de um túnel, enquanto a cavidade da proteína parasita é mais aberta, tal abertura permite que ligantes com o anel benzeno meta substituídos explorem uma região existente na cavidade de *T. cruzi* que é fechada na humana. O farmacóforo de ambas proteínas foi identificado, apresentando diferenças no tamanho e angulação que também podem ser explorados no planejamento de fármacos seletivos.

Palavras-chave: Quimioinformática, Dihidrofolato Redutase, Planejamento de Fármacos

ABSTRACT

VINHA, T. C. N. Application of cheminformatics tools for comparing dihydrofolate reductase's tridimensional structure expressed by different organisms. 2023. 94f. Dissertação (Mestrado) - Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, 2023.

Cheminformatic, defined as the use of informatic techniques to solve chemical problems, has evolved together with new computational tools and it is quite important for rational drug designing, by optimizing different steps on the development pipeline of new molecules, saving resources and time. From all the available tools, structure and ligand based drug design shall be highlighted, when combined, they support the identification and optimization of active molecules from pharmaceutical targets. Dihydrofolate reductase (DHFR) is an important enzyme of the folate pathway that catalyzes the reduction of dihydrofolate to tetrahydrofolate, by using NADPH as cofactor. This reaction is essential for cell replication, as this pathway results in the synthesis of nucleobases that build the DNA. That's the reason why DHFR inhibitors are used for treating bacterial infections and some types of cancer. *Trypanosoma cruzi*, a protozoa that causes Chagas disease, is one of the organisms that express DHFR, besides *Homo sapiens* itself. This work analyzed known ligands and the structure of the protein expressed by both organisms, aiming to identify divergence points that could be explored for designing selective drugs for Chagas disease treatment. The 6 protein's crystallographic models from *T. cruzi* and 2 from *H. sapiens* were obtained from protein data bank (PDB) after the application of quality filters. The amino acid sequence of each model was analyzed by Clustal Omega, its tridimensional structure by Pymol and Chimera X and the cavity analysis by CavityPlus, that also generated the pharmacophore from both targets. The primary structure analysis identified mutations on three amino acids on the parasite crystal, which may be explained by different evolutive paths from segregated groups, although none of the observed mutations are on the active site region. The model's analysis allowed the identification of 24 amino acids that are closer than 5 Å from the *T. cruzi* ligands, 5 of them responsible for hydrogen interactions on at least one of the ligands analyzed. 18 of them are repeated on the human protein or are replaced by another amino acid that preserves the same interaction. As by the differences observed that shall be highlighted, asparagine 44 is replaced by a proline on the human protein, and proline 92 by a lysin. The cavity analysis identified three cavities on each protein, although only the cavities of the active site are druggables. The human protein cavity is bigger and longer, besides it looks like it's a tunnel, when the parasite protein is open, that opening allows ligands with benzene ring meta substituted to explore the existing regions of the *T. cruzi* protein that is closed on the human protein. Lastly, the pharmacophore from both proteins was identified, it shows differences on size and angulation that also could be explored in the designing of selective drugs.

Palavras-chave: Cheminformatics, Dihydrofolate Reductase, Drug Design

LISTA DE ABREVIATURAS E SIGLAS

1CY - 1-(4-clorofenil)-6,6-dimetil-1,6-dihidro-1,3,5-triazina-2,4-diamina
1QZ - 6-etil-5-((3S)-3-[3-(isoquinolin-5-il)-5-metoxifenil]but-1-in-1-il)pirimidina-2,4-diamina
1R0 - 6-etil-5-[3-[3-(isoquinolin-5-il)-5-metoxifenil]prop-1-in-1-il]pirimidina-2,4-diamina
2CY - 5-[3-(3-fluorofenoxi)propoxi]quinazolina-2,4-diamina

ACT - íon acetato
ALA / A - Alanina
ARG / R - Arginina
ASN / N - Asparagina
ASP / D - Aspartato
C - cisteína
C50 - 1-[3-(2,3-diclorofenoxi)propoxi]-6,6-dimetil-1,6-dihidro-1,3,5-triazina-2,4-diamina
CD - íon cádmio
CL - íon cloreto

DHF - dihidrofolato

DHFR - dihidrofolato redutase

DNA - *deoxyribonucleic acid* (ácido desoxirribonucleico)

DQ1- eEtil-4-(5-[[2,4-diaminoquinazolin-6-il]metil]amino)-2-metoxifenoxi)butanoato
DR-TS - dihidrofolato redutase timidilato sintase

dTMP - monofosfato de desoxitimidina

DU - 2'-desoxiuridina-5'-monofosfato (vide dUMP)

dUMP / UMP - monofosfato de desoxiuridina

EDO - 1,2-etanodiol
EOH - etanol
FOL - ácido fólico
GLN / Q - glutamina
GLU / E - glutamato

Gly / G - glicina
GOL - glicerol
H - Histidina
ILE / I - Isoleucina
K - íon potássio

LBDD - *Ligand Based Drug Desing* (Planejamento de Fármacos Baseado no Ligante)
LEU / L - leucina

LG3 - pirimidina-2,4-diamina

LYS / K - lisina

MET / M - metionina

MG - íon magnésio

MTX - metotrexato

NAP / NADPH / NADP+ - *nicotinamide adenine dinucleotide phosphate* (fosfato de dinucleótido de nicotinamida e adenina)

NPX - ácido (2R)-2-(6-metoxinaftalen-2-il)pronaóico

PDB - *Protein Data Bank* (Banco de Dados de Proteínas)

pH - potencial hidrogeniônico

PHE / F - fenilalanina

PO4 - íon fosfato

PRO / P - prolina

RMN - Ressonância Magnética Nuclear

RNA - *ribonucleic acid* (ácido ribonucleico)

SBDD - *Structure Based Drug Design* (Planejamento de Fármacos Baseado na Estrutura)

SER / S - serina

SHMT - serina hidroximetil transferase

SO4 - íon sulfato

THF - tetraidrofolato

THG - (6S)-5,6,7,8-tetraidrofolato

THR / T - treonina

TMQ - trimetrexato

TRP / W - triptofano

TS - timidilato sintase

TYR / Y - tirosina

VAL / V - valina

VG9 - 6-[[2,5-dietoxifenil]amino]metil]-5-metilpirido[2,3-D]pirimidina-2,4-diamina

SUMÁRIO

1.	INTRODUÇÃO	10
1.1.	QUIMIOINFORMÁTICA NO PLANEJAMENTO DE FÁRMACOS.....	10
1.1.1.	Estratégias Quimioinformáticas	11
1.1.2.	Protein Data Bank	12
1.2.	DIHIDROFOLATO REDUTASE.....	14
1.2.1.	Ligantes conhecidos da Dihidrofolato Redutase	15
1.3.	TRYPANOSOMA CRUZI E A DOENÇA DE CHAGAS.....	17
2.	OBJETIVO E JUSTIFICATIVA	20
3.1.	OBTENÇÃO DOS MODELOS CRISTALOGRAFICOS.....	21
3.2.	ANÁLISE DA SEQUÊNCIA DE AMINOÁCIDOS.....	22
3.3.	ANÁLISE DA ESTRUTURA PROTEICA.....	23
3.4.	ANÁLISE DE CAVIDADES	23
4.	RESULTADOS E DISCUSSÃO	25
4.1.	MODELOS CRISTALOGRAFICOS DE T. CRUZI.....	25
4.2.	COMPARAÇÃO DAS SEQUÊNCIAS PRIMÁRIAS DA DHFR DE T. CRUZI..	28
4.3.	COMPARAÇÃO DAS ESTRUTURAS TRIDIMENSIONAIS DA DHFR DE T. CRUZI.....	28
4.4.	ANÁLISE DOS LIGANTES DISPONÍVEIS.....	31
4.5.	ANÁLISE DOS POSSÍVEIS PONTOS DE INTERAÇÃO.....	33
4.6.	MODELOS CRISTALOGRAFICOS DE H. SAPIENS.....	46
4.7.	ANÁLISE COMPARATIVA DA SEQUÊNCIA PRIMÁRIA.....	50
4.8.	ANÁLISE DOS POSSÍVEIS PONTOS DE INTERAÇÃO DA PROTEÍNA HUMANA.....	52
4.9.	COMPARAÇÃO DOS POSSÍVEIS PONTOS DE INTERAÇÃO <i>T. cruzi</i> X <i>H. sapiens</i>	56
4.10.	ANÁLISE DE CAVIDADES DA DIHIDROFOLATO REDUTASE DE <i>T. cruzi</i>	62
4.11.	ANÁLISE DE CAVIDADES DA DIHIDROFOLATO REDUTASE DE <i>H. sapiens</i>	67
5.	CONCLUSÃO	73
	REFERÊNCIAS	75
	APÊNDICES	79

1. INTRODUÇÃO

1.1. QUIMIOINFORMÁTICA NO PLANEJAMENTO DE FÁRMACOS

O surgimento do computador eletrônico na década de 40 revolucionou diversos aspectos da sociedade e da ciência ao redor do mundo, e as ciências químicas não seriam exceção. Com o avanço das tecnologias de hardware e software, além da criação de grandes bancos de dados, as duas áreas do conhecimento se encontraram e evoluíram de forma conjunta, dando origem à “quimioinformática”. O termo, cunhado no fim dos anos 90, pode ser explicado como a área do conhecimento que se vale da aplicação de ferramentas e metodologias informáticas e estatísticas na análise e investigação de desafios da química (MARTINEZ-MAYORGA, *et al.*, 2020).

Ao longo do tempo, a quimioinformática vem sendo amplamente utilizadas, principalmente no ramo farmacêutico para o planejamento racional de novos fármacos, onde as técnicas de quimioinformática podem ser aplicadas em diferentes etapas deste processo, como na identificação de alvos potenciais, na identificação de moléculas com potencial farmacológico, na otimização de tais moléculas e na análise de ensaios pré-clínicos (ALVES, *et al.*, 2017; MARTINEZ-MAYORGA, *et al.*, 2020).

Com o avanço da aplicação de técnicas quimioinformáticas as ferramentas disponíveis também foram se desenvolvendo, otimizando e especificando suas aplicações, tornando a velocidade das análises e a confiabilidade dos dados gerados cada vez maiores. Um exemplo de aplicação destas técnicas é o uso de programas de aprendizado de máquina para encontrar padrões que relacionam as estruturas química de moléculas com seu comportamento biológico; ou ainda para filtrar grandes bancos de dados em busca de compostos que possuam as características desejadas. Cabe ressaltar, no entanto, que em ambos os casos exemplificados, como em demais áreas de aplicação, a quimioinformática apresenta uma relação de dependência com outras áreas da ciência experimental para que tenha um ponto de partida na construção de suas predições (ALVES, *et al.*, 2017; MARTINEZ-MAYORGA, *et al.*, 2020).

Em comparação com as tradicionais técnicas de bancada, as computacionais

são mais baratas, dependem de menos recursos, tem uma logística mais simplificada, sofrem com menos variáveis externas e apresentam resultados mais rapidamente. Por estas vantagens, elas são empregadas nas etapas iniciais do desenvolvimento de um novo fármaco, com o objetivo de selecionar e otimizar possíveis *hits*, direcionando de forma racional e aumentando as chances de sucesso das etapas subseqüentes de ensaios químicos e biológicos tradicionais (CHEN; KIRCHMAIR, 2020).

1.1.1. Estratégias Quimioinformáticas

Existem diversas estratégias que se enquadram dentro do campo da quimioinformática, dentre elas, destacam-se o Planejamento de Fármacos Baseado na Estrutura (SBDD, em inglês, *Structure Based Drug Design*) e o Planejamento de Fármacos Baseado no Ligante (LBDD, *Ligand Based Drug Design*, em inglês) (ALVES, *et al.*, 2017; MARTINEZ-MAYORGA, *et al.*, 2020).

Estratégias de SBDD partem de informações sobre a estrutura tridimensional de um alvo farmacológico, que ao ser analisado permite a identificação de pontos de interação com ligantes, bem como as características de tais pontos. Essas informações podem, então, ser utilizadas na busca por moléculas compatíveis, ou seja, que tenham potencial para serem ligantes do alvo selecionado (BATOO; AHMAD; CHOI, 2019).

Por sua vez, LBDD parte de uma abordagem onde o ligante é utilizado como base de estudo, sem que sejam necessárias informações do alvo. Nestes estudos, informações sobre a atividade da molécula, ou das moléculas, ativa são correlacionadas com suas características estruturais e físico-químicas, que por sua vez são utilizadas como base para o desenho e busca de novas moléculas (FERREIRA; ANDRICOPULO, 2018).

Para que seja possível a aplicação de técnicas de SBDD ou LBDD, como o próprio nome sugere, é preciso possuir informações prévias quanto à estrutura tridimensional da proteína alvo de interesse ou de ligantes conhecidos. Além deste pré-requisito, os principais desafios das técnicas de SBDD estão na flexibilidade das moléculas estudadas: a maioria dos programas empregados baseiam-se na construção de um ambiente rígido ou pouco flexível, principalmente no caso da macromolécula, quando no ambiente biológico macromoléculas e ligantes

apresentam um alto grau de flexibilidade. Tal diferença resulta em limitações nas aplicações e resultados observados com as técnicas de SBDD e LBDD (GUIDO; ANDRICOPULO; OLIVA, 2010; HASSAN BAIG, *et al.*, 2016).

Apesar da dependência de informações pré-existentes sobre alvos e ligantes, estratégias de SBDD e LBDD tem sido cada vez mais relevante no planejamento de fármacos, principalmente no contexto de doenças negligenciadas, onde o baixo custo relativo destas técnicas contorna a falta de investimento em pesquisas neste campo (FERREIRA; MORAES; ANDRICOPULO, 2022).

1.1.2. *Protein Data Bank*

Considerando o pré-requisito do SBDD supracitado, o *Protein Data Bank* (PDB) é um importante repositório de estruturas tridimensionais de proteínas. Desenvolvido em 1971 e que cresceu de forma acelerada nos anos seguintes, resultando em números significativos de estruturas depositadas (BERMAN, *et al.* 2002; BERNSTEIN, *et al.*, 1977).

Existem três metodologias comumente empregadas na determinação da estrutura tridimensional de proteínas: cristalografia de raio-X, espectroscopia de ressonância magnética nuclear (RMN) e espectroscopia eletrônica tridimensional (ZHAO, 2019).

A cristalografia de raio-X é uma técnica biofísica que depende da obtenção de um cristal da proteína de interesse para determinação da estrutura. Esse cristal é então, bombardeado com raios-X e a leitura dos raios difratados permite gerar um mapa de densidade eletrônica da proteína que, conseqüentemente, indica a posição de cada átomo da macromolécula em questão. Os maiores desafios dessa metodologia estão na própria obtenção do cristal, visto que realizar o processo de cristalização de proteínas sem que elas percam suas características estruturais é delicado e dispendioso, levando algumas proteínas a cristalizar-se em conformações diferentes das observadas naturalmente e, além disso, o cristal confere uma rigidez às moléculas que pode ser incompatível com proteínas mais flexíveis e deve ser considerado no uso das informações cristalográficas em experimentos e análises (PARKER, 2003).

Por outro lado, a espectroscopia de RMN permite que a estrutura da proteína

seja determinada em solução, o que supera dois desafios da cristalografia de raio X. Por essa técnica, proteínas que não podem ou que apresentam dificuldade em ser cristalizadas podem ser avaliadas, sendo que ela permite um maior grau de mobilidade ao sistema estudado. Assim, as proteínas são submetidas a um forte campo magnético e então irradiadas com ondas de rádio, a alteração da frequência dessas ondas permite identificar quais átomos estão próximos e, conseqüentemente, determinar a estrutura tridimensional da proteína. Infelizmente, devido à complexidade dos sinais gerados e a sobreposição de sinais, essa metodologia é normalmente aplicada para estudo de proteínas menores (BAX; CLORE, 2019)

Por fim, na microscopia eletrônica tridimensional amostras da proteína são resfriadas a temperaturas criogênicas, formando finas camadas de gelo amorfo que são então bombardeadas com elétrons. A difração e refração dos elétrons permite gerar uma imagem bidimensional da proteína, várias imagens, de diferentes cristais, fornecem informações da proteína em variadas posições, que são então combinadas criando um modelo tridimensional da proteína. A técnica, desenvolvida nos anos 70, evoluiu em conjunto com o desenvolvimento das ferramentas de captura de imagem, permitindo que hoje sejam obtidas imagens de alta resolução, que devido a sua relevância científica agraciaram seus desenvolvedores com o prêmio Nobel de química em 2017 (BAKER, 2018; SHEN, 2018).

A evolução das diferentes técnicas para determinação da estrutura tridimensional de macromoléculas, que permitiram a obtenção destas de forma mais fácil e precisa, bem como o incentivo à cooperação internacional entre cientistas na busca por compartilhar resultados e informações, contribuiu para o crescimento e sucesso do PDB. Hoje, além de informações de proteínas de diversos organismos, o PDB possui informações de diferentes moléculas e ligantes e também de estruturas de DNA e RNA (BERMAN, *et al.*, 2002; RCSB PDB, 2020).

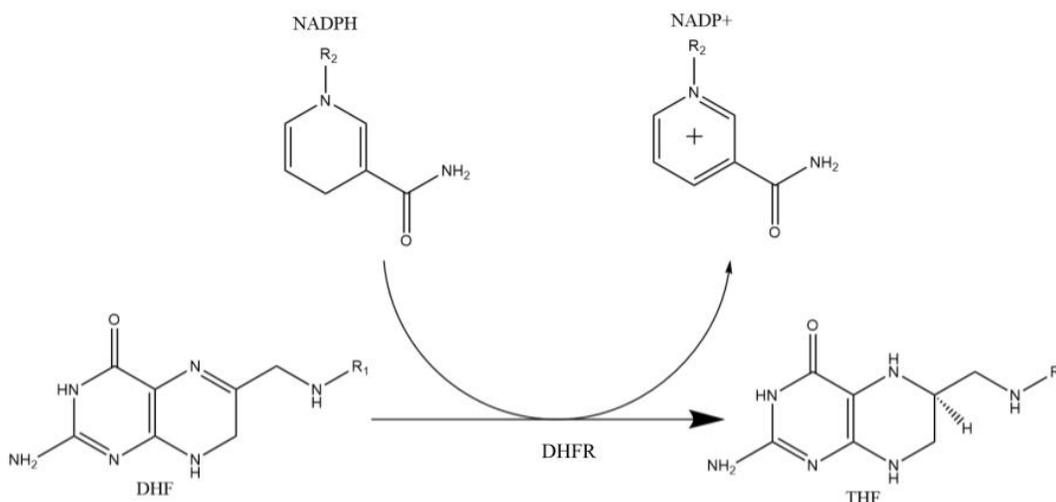
Do ponto de vista da aplicação de cristais para estratégias de SBDD, a forma como a estrutura foi obtida não apresenta grandes impactos, desde que a boa definição do cristal seja assegurada, ou seja, a resolução do cristal deve ser a menor possível (HASSAN BAIG, 2016).

1.2. DIHIDROFOLATO REDUTASE

A Dihidrofolato Redutase (DHFR) é uma importante enzima da via de biossíntese dos folatos, essencial para a síntese de bases nitrogenadas purina e pirimidina utilizadas na construção do DNA e RNA, estando presente em diferentes organismos, como bactérias, protozoários e mamíferos (IRVINE, 2020).

A DHFR é responsável por catalisar a redução do dihidrofolato (DHF) em tetrahydrofolato (THF) utilizando NADPH como cofator, conforme ilustrado na Figura 1. A Dihidrofolato Redutase satura a ligação entre o Nitrogênio (5) e o Carbono (6) do anel pterina do DHF, adicionando dois hidrogênios à molécula, que se torna THF. Para que tal reação ocorra, a DHFR também oxida a molécula de NADPH, promovendo a insaturação do anel e transformando-a em NADP⁺ (REDDISH, *et al.*, 2016).

Figura 1 - Reação de redução do dihidrofolato (DHF) em tetrahydrofolato (THF) catalisada pela enzima dihidrofolato redutase (DHFR) com o conversão do cofator NADPH em NADP⁺

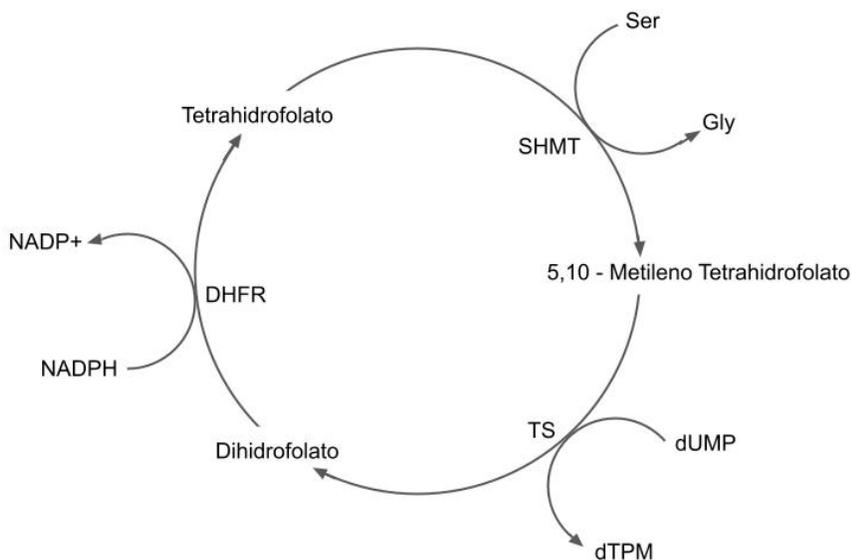


Fonte: Schnell, Dyson e Wright (2004)

Seguindo na via dos folatos (Figura 2), o THF é convertido em 5,10-metileno tetrahydrofolato pela serina hidroximetil transferase (SHMT) que por sua vez é reciclado em DHF pela timidilato sintase (TS) - nesse passo da via ocorre a conversão de monofosfato de desoxiuridina (dUMP) em monofosfato de desoxitimidina (dTMP), que é essencial para a síntese dos monômeros de DNA (CLARK; PAZDERNIK; MCGEHEE, 2019).

Sendo assim, a DHFR é uma proteína com características relevantes para ser estudada como um alvo de modulação farmacológica, além de ser um alvo promissor para diferentes doenças - podendo ser explorada no tratamento de doenças como malária, tuberculose e tripanossomíases, além de infecções bacterianas e fúngicas e no tratamento de doenças autoimunes e câncer (SHARMA; CHAUHAN, 2012).

Figura 2 - Trecho da via de síntese dos folatos - A dihidrofolato redutase (DHFR) converte dihidrofolato em tetrahydrofolato com a conversão de NADPH em NADP+, o tetrahydrofolato é então convertido em 5,10-metileno tetrahydrofolato pela serina hidroximetil transferase (SHMT), consumindo uma serina (Ser) e gerando uma glicina (Gly), a reciclagem de 5,10-metileno tetrahydrofolato em dihidrofolato é catalisada pela timidilato sintase (TS) ao mesmo tempo em que converte monofosfato de desoxiuridina (dUMP) em monofosfato de desoxitimidina (dTMP) - sendo este um precursor das bases nitrogenadas do DNA.

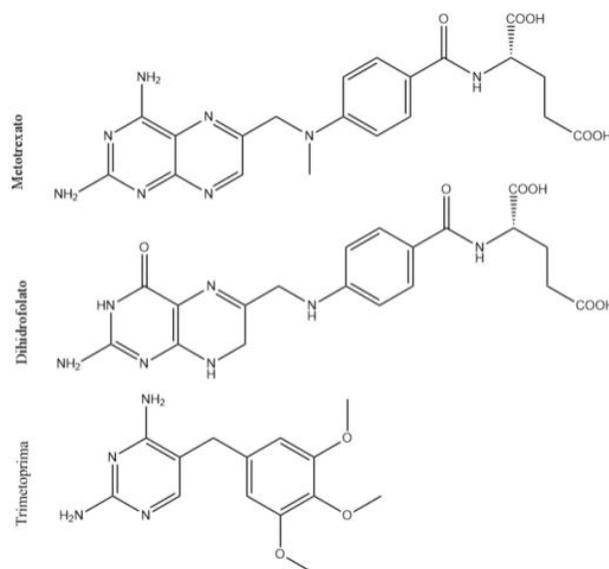


Fonte: Anderson (2005).

1.2.1. Ligantes conhecidos da Dihidrofolato Redutase

Visto a importância da dihidrofolato redutase na regulação da via dos folatos, já existem fármacos disponíveis no mercado que atuam por essa via, como o metotrexato e a trimetoprima (Figura 3) (CLARK; PAZDERNIK; MCGEHEE, 2019).

Figura 3 - Moléculas do metotrexato, dihidrofolato e trimetoprima, é possível observar a semelhança estrutural das três moléculas pela presença do anel heterocíclico e seus substituintes mais próximos, o que assegura a capacidade de acoplamento das três moléculas no sítio ativo da DHFR.



Fonte: Clark, Pazdernik e Mcgehee (2019)

O metotrexato é um antimetabólito da via dos folatos. Sua estrutura foi planejada para ser similar ao DHF (Figura 3), deste modo ele tem a capacidade de acoplar-se ao sítio ativo da DHFR, porém é incapaz de gerar os substratos necessários para a continuação dessa via, inibindo, assim, a proteína e, conseqüentemente, a replicação celular. Seu uso se dá no tratamento de alguns tipos de câncer, inibindo a replicação de células tumorais (CLARK; PAZDERNIK; MCGEHEE, 2019).

De forma similar, a trimetoprima também é um fármaco disponível no mercado que possui como alvo a DHFR. Embora sua estrutura tenha menor semelhança com o substrato natural da enzima em relação ao metotrexato, é possível observar a presença do anel pirimidina meta substituído, além disso, o substituinte metil benzeno garante que a estrutura tridimensional do fármaco seja similar o suficiente ao DHF para garantir o acoplamento da trimetoprima no sítio ativo da DHFR e, assim, conferir sua ação como antimetabólito. Uma vantagem da trimetoprima é possuir maior afinidade pela DHFR bacteriana, logo, ela é empregada como antibacteriano (CLARK; PAZDERNIK; MCGEHEE, 2019).

1.3. TRYPANOSOMA CRUZI E A DOENÇA DE CHAGAS

T. cruzi é o patógeno causador da Doença de Chagas, uma doença tropical negligenciada, que, de acordo com a Organização Mundial da Saúde, afeta cerca de 6 a 7 milhões de pessoas ao redor do mundo, principalmente na América Latina (OMS, 2022).

A forma de contágio mais reportada é por contato com o vetor da doença, o inseto conhecido como Barbeiro. Ao se alimentar do sangue de uma pessoa, as fezes contaminadas do vetor são depositadas no local e o parasita em sua forma infectante, tripomastigota, é capaz de invadir a epiderme por alguma via de acesso existente e, conseqüentemente, atingir a circulação sanguínea do hospedeiro. O parasita então infecta diferentes células do hospedeiro e diferentes tecidos, onde se multiplicam na forma de amastigotas. Após diferenciar-se novamente em tripomastigotas, ocorre a lise celular que libera o parasita novamente na corrente sanguínea, que por sua vez passam a infectar novas células e atingir novos tecidos, principalmente tecidos musculares. De forma a complementar o ciclo, quando um inseto barbeiro se alimenta do sangue de uma pessoa infectada, este se infecta com a forma tripomastigota do parasita, que se diferencia em epimastigota, no trato digestivo do vetor o parasita se multiplica e se diferencia novamente em tripomastigota, reiniciando o ciclo da doença (MORETTI; MORTARA; SCHENKMAN, 2020; ZUMA; BARRIAS; SOUZA, 2021).

No momento inicial do ciclo de infecção ocorre a fase aguda da doença, que pode ser assintomática, ou causar febre baixa no hospedeiro, na minoria dos casos sintomáticos a fase aguda da doença de chagas pode resultar em miocardite ou meningoencefalite. Na maioria dos casos, no entanto, a fase aguda evolui para a fase crônica da doença, onde o parasita infecta principalmente os tecidos musculares, sem causar sintomas claros no hospedeiro e ficando indetectável nos exames clínicos de anticorpos contra *T. cruzi*. No decorrer dos anos, a doença de Chagas na sua forma crônica pode resultar em sintomas cardíacos, digestivos ou uma combinação destes, os sintomas mais comuns são o aumento dos tecidos atingidos, como coração esôfago e cólon, que levam a um mal funcionamento destes órgãos e podem levar à morte prematura do paciente (RASSI JR; RASSI; MARIN-NETO, 2010).

Formas alternativas de contaminação também estão descritas na literatura, atualmente a forma mais importante de transmissão é pela via oral, através do consumo de alimentos não processados contaminados com o parasita na sua forma tripomastigota. O parasita é capaz de infiltrar-se na corrente sanguínea do hospedeiro pela mucosa bucal e do estômago, dando sequência ao ciclo previamente descrito. Essa forma de contágio é responsável pelos surtos de doença de chagas aguda observados em áreas endêmicas. Outras formas menos comuns de transmissão são por transfusão sanguínea e transmissão congênita (SUZUKI, *et al.*, 2023; RUEDA, *et al.*, 2014).

Atualmente só existem dois fármacos disponíveis para o tratamento da Doença de Chagas: benzonidazol e nifurtimox. Porém nenhuma das duas opções apresentam eficácia satisfatória e a toxicidade dos fármacos faz com que os pacientes desistam do tratamento por conta dos efeitos adversos observados, como anorexia, perda de apetite e irritação gastrointestinal. Além disso, já foram identificadas cepas resistentes a estes fármacos (SALES JR, *et al.*, 2017).

Tanto o nifurtimox quanto o benzonidazol são pró-fármacos nitroheterocíclicos, que são catalisados pela nitroreductase tipo I do parasita. No caso do benzonidazol, o grupo nitroimidazol é reduzido em hidroxilamina que sofre transformações não enzimáticas até a formação de glioxal. Já o nifurtimox sofre redução no grupo nitrofurano, formando também a hidroxilamina, que novamente é transformada, resultando em uma molécula com um grupo nitrila saturado. Ambas moléculas resultantes são capazes de interagir com macromoléculas protéicas e de DNA/RNA, causando os efeitos tóxicos e mutagênicos observados no parasita. No caso do benzonidazol, há também a formação do íon nitrênio, capaz de reagir com o DNA e danificá-lo (MULLER KRATZ, *et al.*, 2018; PATTERSON; WYLLIE, 2014).

A busca por alternativas para o tratamento da doença de Chagas levou a identificação de novos alvos promissores e desenvolvimento de novas moléculas. Inibidores da via de biossíntese do ergosterol chegaram a etapa de ensaios clínicos - o ergosterol é um esteroide essencial para a viabilidade celular do parasita e inibir sua produção, além de causar a falta do esteroide essencial, leva ao acúmulo de intermediários tóxicos da via biossintética. Outros alvos estudados para o planejamento de inibidores são a cruzipaina (ou cruzaina) - essencial para a diferenciação do parasita e sua capacidade de infectar novas células - e a farnesil

pirofosfato sintase - ao inibi-la os esteróis essenciais são reduzidos, prejudicando a viabilidade celular - no entanto nenhum fármaco com tais alvos chegou a fase de ensaio clínico por enquanto (SALES JR, *et al.*, 2017).

2. OBJETIVO E JUSTIFICATIVA

Considerando a relevância da doença de Chagas, seu impacto na saúde e qualidade de vida de populações carentes da América Latina e outras regiões do hemisfério sul e a falta de alternativas farmacológicas que sejam eficazes e seguras, faz-se necessário investir na pesquisa de novos alvos promissores para possíveis planejamentos futuros de fármacos.

Neste contexto, a dihidrofolato redutase apresenta-se como um destes possíveis alvos promissores, uma vez que sua função bioquímica está diretamente relacionada à multiplicação celular do parasita. No entanto, a dihidrofolato redutase é uma proteína também expressa pelo ser humano, desta forma faz-se necessário planejar fármacos que sejam seletivos para a enzima de *T. cruzi*, de forma a evitar a ocorrência de efeitos adversos graves que impactam o tratamento. Assim, avaliar comparativamente as estruturas da proteína humana e da proteína do parasita permite que sejam identificados e abordados pontos de divergências entre ambas, pontos tais que possam ser explorados no planejamento de moléculas que tenham maior afinidade por uma estrutura em detrimento da outra, ou seja, que sejam seletivos para a dihidrofolato redutase parasita.

Logo, os estudos realizados nesta dissertação propõem a comparação da estrutura tridimensional da dihidrofolato redutase de *Trypanosoma cruzi* com sua homóloga humana, no intuito de identificar diferenças estruturais que propiciem o planejamento de fármacos seletivos e corroborem a escolha da dihidrofolato redutase como potencial alvo farmacológico para a doença de Chagas.

3. MATERIAIS E MÉTODOS

3.1. OBTENÇÃO DOS MODELOS CRISTALOGRAFICOS

Os modelos cristalográficos das proteínas foram obtidos a partir do Banco de Dados de Proteínas (PDB - *Protein Data Bank*, em inglês). Foram utilizados como critérios de busca o nome da proteína dihidrofolato redutase em inglês - *dihydrofolate reductase*; além do nome científico dos organismos de interesse - *Trypanosoma cruzi* e *Homo sapiens*.

Considerando a importância de selecionar modelos cristalográficos de alta qualidade para que os resultados sejam confiáveis, foram utilizados três parâmetros como critérios de escolha dos modelos: a resolução, a diferença entre *R-free* e *R-Work/R-Value* e a posição dos outliers de Ramachandran, visto que todos os modelos disponíveis foram obtidos via difração de Raio-X, utilizou-se os mesmos critérios para todos os modelos (KLEYWEGT; JONES, 1997; EVANSA; MURSHUDOV, 2013).

R-Value, ou *R-Work*, são utilizados para determinar a qualidade do refinamento do modelo cristalográfico gerado: após determinar o modelo atômico do cristal, simula-se os dados de difração que seriam gerados pelo modelo determinado, essa difração é comparada aos dados experimentais da difração original, gerada pelo cristal de proteína, e o *R-Value* é calculado, sendo que quanto mais próximo a 0, maior a similaridade entre os perfis de difração, ou seja, mais próximo aos dados experimentais e mais confiável é o modelo cristalográfico determinado (KLEYWEGT; JONES, 1997).

Antes de iniciar o refinamento do modelo atômico parte dos dados experimentais são separados e não são utilizados. Após a construção do modelo esses dados são utilizados para calcular o *R-Free*: mesmo indicador do *R-Value*, porém com dados que não foram utilizados nos cálculos iniciais. Quanto mais próximo o *R-Free* do *R-Value*, menos enviesada foi o refinamento e, portanto, mais confiável é o modelo gerado (KLEYWEGT; JONES, 1997).

O gráfico de Ramachandran calcula quais os possíveis ângulos ϕ e Ψ entre as cadeias laterais de aminoácidos e o *backbone* da proteína, considerando a rotação e a angulação das ligações peptídicas e a cadeia lateral dos aminoácidos vizinhos, uma vez que impedimentos estéricos diminuem as possibilidades de

movimentação das cadeias laterais. Os *outliers* de Ramachandran são aminoácidos que, na determinação do modelo cristalográfico, apresentam angulações não permitidas pelo gráfico de Ramachandran, indicando possíveis falhas no posicionamento dos átomos da cadeia lateral do aminoácido em questão. Embora *outliers* sejam comuns em modelos cristalográficos, é importante que a quantidade seja pequena em relação ao tamanho da proteína e que as regiões afetadas por *outliers* não impactam nas análises realizadas (GOPALAKRISHNAN *et al.*, 2007; RAMACHANDRAN; RAMAKRISHNAN; SASISEKHARAN, 1963).

3.2. ANÁLISE DA SEQUÊNCIA DE AMINOÁCIDOS

Para a análise comparativa da sequência primária de cada um dos modelos cristalográficos selecionados empregou-se as sequências de aminoácidos no formato FASTA na plataforma online Clustal Omega (SIEVERS, *et al.*, 2011).

A ferramenta utilizada, além de realizar o alinhamento das sequências, avalia o tipo de variação com base no *score* da matrix de Gonnet PAM 250, que determina o número de mutações da sequência genética necessárias para que a mudança em um aminoácido da proteína aconteça com base na comparação de proteínas homólogas conhecidas. Este indicador é importante para o alinhamento de sequências proteicas, uma vez que permite que mutações sejam levadas em consideração ao executar tal alinhamento (SIEVERS, 2011; STRASSER, 2009).

O programa utilizado faz a seguinte classificação dos *scores* das trocas de aminoácidos: se for maior que 0,5 é considerada uma substituição fortemente conservadora, ou seja, a alteração ocorre entre aminoácidos com propriedades muito similares. Se o *score* da substituição for entre 0 e 0,5 a mudança conserva poucas características, uma vez que os aminoácidos são levemente similares. Acima de 0,5 é considerada uma mutação que não conserva similaridade (SIEVERS, 2011).

O resultado é representado visualmente no alinhamento, onde um asterisco indica que o mesmo aminoácido ocupa a posição nas duas sequências, dois pontos indicam que a substituição conserva fortemente as características do aminoácido e um ponto indica que a substituição é pouco conservadora.

Alinharam-se as sequências da dihidrofolato redutase de *T. cruzi* de todos os modelos cristalográficos disponíveis no PDB para que fossem identificadas possíveis

mutações. Da mesma forma, os modelos cristalográficos da DHFR de *H. sapiens* foram alinhados entre si e com outros dois modelos representantes da proteína do parasita. Para o segundo caso, como a proteína de *T. cruzi* é uma proteína bifuncional, ou seja, possui também a subunidade timidilato sintase, somente a sequência da porção correspondente a dihidrofolato redutase foi considerada no alinhamento.

3.3. ANÁLISE DA ESTRUTURA PROTEICA

A estrutura tridimensional dos modelos cristalográficos foi analisada com dois programas diferentes: PyMol e Chimera, escolhidos pela sua utilização simples e interface amigável e por serem programas acessíveis de uso livre para pesquisa. Ambos foram utilizados para a visualização das estruturas, além da visualização e análise dos aminoácidos divergentes entre os modelos e identificação e análise dos pontos de interação com o ligante.

O programa PyMol foi utilizado para a visualização da proteína e das moléculas ligantes, além de ser utilizado para geração das imagens utilizadas neste trabalho. Dentre as ferramentas disponíveis no programa, foi utilizada a *measurement wizard* para determinação da distância atômica indicadas nas figuras (SCHRÖDINGER; DELANO, 2020).

Já o programa Chimera X, por sua vez, foi utilizado para determinar os pontos de interação entre os ligantes e a proteína, tanto para *H. sapiens* quanto para *T. cruzi*, além da geração de algumas das imagens aqui representadas.

Para determinar os pontos de interação utilizou-se a ferramenta *contacts* do programa, foi determinada a distância máxima de 5 Å entre qualquer átomo do ligante e um átomo da macromolécula ou do co-fator. Essa distância foi escolhida pois todas as interações intermoleculares, como interações íon-dipolo, dipolo-dipolo e Van der Waals dependem de distâncias menores para ocorrer, desta forma, todas as possíveis interações estariam contempladas na análise. Além de apresentar como resultado os pontos de interação e suas respectivas distâncias, o programa também indica quais interações podem ser contatos não favoráveis, devido a distâncias muito pequenas entre os átomos, o que indicaria possíveis erros nos modelos analisados (GODDARD, 2018; KOHLHOFF, *et al.* 2009).

Para determinar os pontos de interação de hidrogênio, utilizou-se a ferramenta *H-bonds*, essa ferramenta considera critérios como natureza dos átomos envolvidos na interação (nitrogênio, oxigênio ou enxofre) e o tipo de hibridização, distância e angulação interatômica para determinar os possíveis pontos de interação de hidrogênio entre duas moléculas selecionadas (GODDARD, 2018).

As ferramentas de alinhamento foram utilizadas em ambos os programas para alinhar os diferentes modelos selecionados.

3.4. ANÁLISE DE CAVIDADES

O programa CavityPlus é uma plataforma *online* que permite identificar bolsões acessíveis de uma estrutura proteica. A identificação de bolsões é essencial no planejamento racional de fármacos, uma vez que interações entre ligante-alvo ocorrem justamente em cavidades da proteína que permitem o acoplamento de tais ligantes. A identificação de bolsões e suas características permite determinar quais as características necessárias para um acoplamento favorável, além de permitir a identificação de possíveis sítios de interação alostérica, isto é, pontos de interação entre a proteína e o ligante, que não sejam o sítio ativo, mas que ainda assim permitam a modulação da atividade proteica (XU, *et al.*, 2018).

Foram utilizadas três ferramentas da plataforma: *Cavity*, que permite a identificação de bolsões, *CavPharmer* que identifica as características de cada bolsão para o planejamento de farmacóforos, como a hidrofobicidade e a presença de doadores ou aceptores de interações de hidrogênio, e o *CorrSite* que identifica sítios alostéricos na proteína (XU, *et al.*, 2018).

Realizou-se as análises com um modelo de *T. cruzi* (5T7O) e um modelo de *H. sapiens* (4M6K), uma vez que as estruturas proteicas de cada modelo analisado nos itens anteriores já haviam sido comparadas e não foram identificadas divergências relevantes entre eles. Optou-se por tais modelos como representantes devido aos seus ligantes serem os mais próximos entre si e em relação ao substrato natural da proteína, diminuindo assim a influência do ligante na conformação estrutural do sítio ativo.

Para todas as análises a entrada foi o código PDB do modelo cristalográfico, selecionou-se sempre a subunidade A para análise e para as análises que fosse

necessário indicar o ligante foi escolhido o ligante ativo, ou seja, o tetraidrofolato ou o folato.

4. RESULTADOS E DISCUSSÃO

4.1. MODELOS CRISTALOGRAFICOS DE *T. CRUZI*

Ao realizar uma busca no PDB por resultados que contenham as palavras “*Dihydrofolate Reductase*” ou “DHFR” e cujo organismo contenha “*Trypanosoma cruzi*” foram obtidos 10 resultados.

A DHFR de *T. cruzi* é uma proteína homodimérica bifuncional, isto é, ela possui duas porções catalíticas, uma responsável pela ação de dihidrofolato redutase e outra responsável pela ação de timidilato sintase, que catalisa o produto tetrahydrofolato em timidilato. Por se tratar de uma proteína única com duas subunidades catalíticas, todos os modelos cristalográficos obtidos possuem ambas subunidades representadas (BELTRAN-HORTELANO, 2022).

Os resultados foram filtrados com base nos seguintes parâmetros: resolução < 3,0 Å; R-Free - R-Work < 0,23 e posição dos *outliers* de Ramachandran não coincidente com sítio ativo ou pontos de interação com ligantes ou co-fator.

Foram descartados todos os resultados que apresentassem resolução > 3,0 (1 resultado), R-free - R-Work > 0,23 (0 resultados), no caso dos *Outliers* de Ramachandran (4 resultados), os modelos foram descartados caso a posição dos *outliers* coincidissem com o sítio ativo ou com alguma posição de interação com ligantes ou co-fator, conforme listado na Tabela 01. Os cristais descartados estão sublinhados, assim como o parâmetro responsável por tal descarte.

Tabela 01 - Modelos cristalográficos de DHFR disponíveis no PDB para *T. cruzi*.

PDB ID	Nº de cadeias	Peso Molecular (kDa)	Resolução (Å)	Diferença R-Free R-Work	Outliers (backbone / side chain)	Ligantes (ID - descrição)
3HBB	4	243,51	3,00	0,038	31 / 106	NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida TMQ - trimetrexato SO4 - íon sulfato EDO - 1,2-etanodiol
3CLB	4	240,83	3,00	0,038	28 / 105	NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida TMQ - trimetrexato SO4 - íon sulfato EDO - 1,2-etanodiol
3KJS	4	241,87	2,5	0,034	10 / 28	DQ1 - etil-4-(5-[(2,4-diaminoquinazolin-6-il)metil]amino)-2-metoxifenoxi)butanoato NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida SO4 - íon sulfato EDO - 1,2-etanodiol
2H2Q	2	119,91	2,4	0,035	3 / 39	NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida DU - 2'-desoxiuridina-5'-monofosfato
<u>3IRM</u>	4	237,95	2,1	0,043	<u>27 / 92</u>	1CY - 1-(4-clorofenil)-6,6-dimetil-1,6-dihidro-1,3,5-triazina-2,4-diamina PO4 - íon fosfato ACT - íon acetato
3INV	2	121,14	2,37	0,054	3 / 49	C50 - 1-[3-(2,3-diclorofenoxi)propoxi]-6,6-dimetil-1,6-dihidro-1,3,5-triazina-2,4-diamina NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida UMP - 2'-desoxiuridina-5'-monofosfato GOL - glicerol

(continua)

PDB ID	N° de cadeias	Peso Molecular (kDa)	Resolução (Å)	Diferença R-Free R-Work	Outliers (<i>backbone / side chain</i>)	Ligantes (ID - descrição)
(continuação)						
5T7O	4	244,62	1,8	0,043	0 / 15	NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida THG - (6S)-5,6,7,8-tetrahydrofolate UMP - 2'-desoxiuridina-5'-monofosfato EDO - 1,2-etanodiol
<u>3IRO</u>	4	241,16	2,8	0,071	<u>57 / 93</u>	2CY - 5-[3-(3-fluorofenoxi)propoxi]quinazolina-2,4-diamina NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida PO4 - íon fosfato ACT - íon acetato GOL - glicerol
<u>3IRN</u>	4	240,08	2,6	0,069	<u>48 / 101</u>	1CY 1-(4-clorofenil)-6,6-dimetil-1,6-dihidro-1,3,5-triazina-2,4-diamina NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida PO4 - íon fosfato
<u>3CL9</u>	1	60,83	<u>3,3</u>	0,043	<u>11 / 46</u>	SO4 - íon sulfato CL - íon cloreto NAP / NADP - fosfato de dinucleotídeo de adenina e nicotinamida MTX - metotrexato UMP - 2'-desoxiuridina-5'-monofosfato EDO - 1,2-etanodiol

(conclusão)

4.2. COMPARAÇÃO DAS SEQUÊNCIAS PRIMÁRIAS DA DHFR DE *T. CRUZI*

Realizou-se um alinhamento de sequências de aminoácidos com os 10 modelos cristalográficos de *T. cruzi* disponíveis no PDB (listados na Tabela 1), inclusive aqueles descartados por critérios de qualidade. Esperava-se assim avaliar se algum modelo com mutação relevante tivesse sido desconsiderado dos estudos subsequentes. Os alinhamentos resultantes encontram-se no apêndice A.

Como resultado, observou-se que das 10 sequências disponíveis, 5 possuem uma histidina na posição 32, uma leucina na posição 55 e uma arginina na posição 137, enquanto as outras 5 sequências apresentam, respectivamente, uma arginina, uma valina e uma glutamina. As substituições mantêm as características físico-químicas dos aminoácidos, exceto na substituição da arginina, um aminoácido básico, pela glutamina, um aminoácido polar, de toda forma, torna-se necessário uma análise da estrutura tridimensional da proteína para avaliar o impacto de tais substituições.

A dihidrofolato redutase timidilato sintase (DHFR-TS) de *T. cruzi* é uma proteína que apresenta polimorfismos em sua sequência genética. Segundo Machado e Ayala (2002), as variações genéticas encontradas em cepas do parasita permitem organizá-las em quatro clados filogenéticos distintos, o que indica a possibilidade do parasita ter evoluído de forma isolada e diferenciado-se em quatro regiões diferentes. Ainda que as variações tenham ocorrido naturalmente, a maioria dos polimorfismos observados são silenciosos, ou seja, não resultam na modificação da sequência de aminoácidos da proteína e, ainda assim, aqueles que levam à modificações na sequência primária são em regiões que não afetam os sítios ativos da proteína nem sua funcionalidade (Machado, Ayala, 2002).

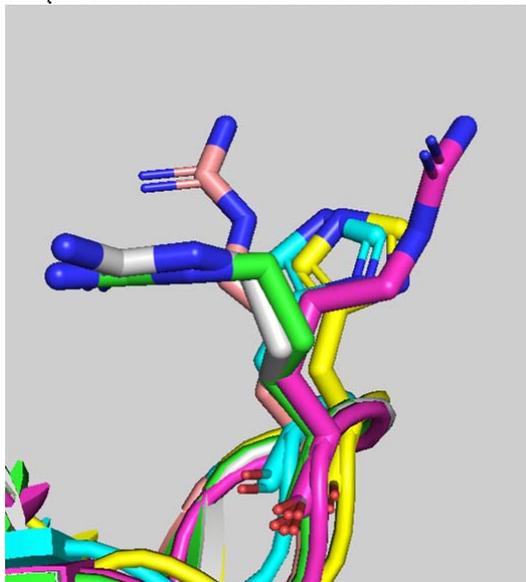
4.3. COMPARAÇÃO DAS ESTRUTURAS TRIDIMENSIONAIS DA DHFR DE *T. CRUZI*

Primeiramente, foi realizada a comparação das diferentes estruturas de DHFR frente a mudança de aminoácidos observada na análise da sequência das estruturas. Para esta análise utilizou-se o programa Pymol como ferramenta de visualização do modelo cristalográfico e foi possível observar que as três trocas de aminoácidos identificadas, H32R, L55V e R137Q, estão em posições periféricas da

proteína e não coincidem com o sítio ativo, nem com pontos de interação com cofatores.

A alteração H32R, indicada na Figura 4, encontra-se em uma alça da proteína, as cadeias laterais dos dois aminoácidos, histidina e arginina, apresentam cadeias de carbono formada por rotâmeros, o que confere flexibilidade ao resíduo e permite que a cadeia lateral ocupe diferentes posições, devido aos confôrmeros possíveis das ligações C-C, ainda que a substituição da cadeia linear da arginina pelo anel aromático da histidina aumente a rigidez da molécula, a geometria estrutural de ambas permanece similar, além disso, como dito anteriormente, as características físico-químicas são mantidas, uma vez que a composição das cadeias laterais são similares, formadas por ligações simples e duplas com a presença de átomos de nitrogênio, o que resulta em resíduos positivamente carregados em pH neutro.

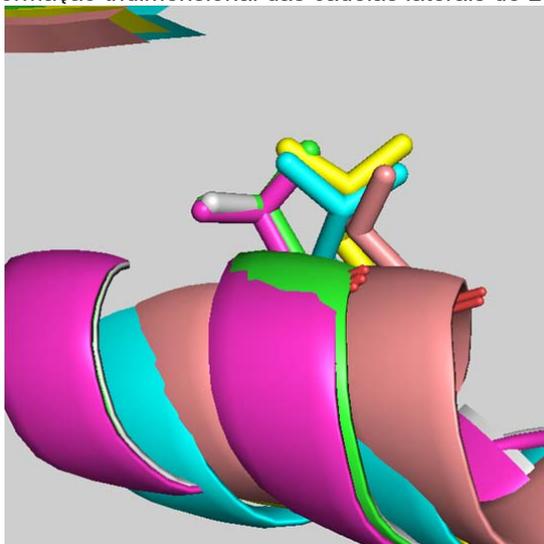
Figura 4 - Conformação tridimensional das cadeias laterais de Histidina e Alanina



Legenda: branco - 3CLB; verde - 3HBB; magenta - 3KJS; rosa - 2H2Q; ciano - 3INV; amarelo - 5T7O.

Já a substituição L55V, indicada na Figura 5, encontra-se em uma das alfa-hélices da proteína, sem alterar a estrutura terciária da macromolécula, uma vez que a troca de aminoácido apenas acrescenta um carbono na cadeia lateral. Embora a leucina seja um pouco maior que a valina, ambas são compostas apenas por carbonos, possuindo as mesmas características físico-químicas, e possuem a mesma conformação estrutural.

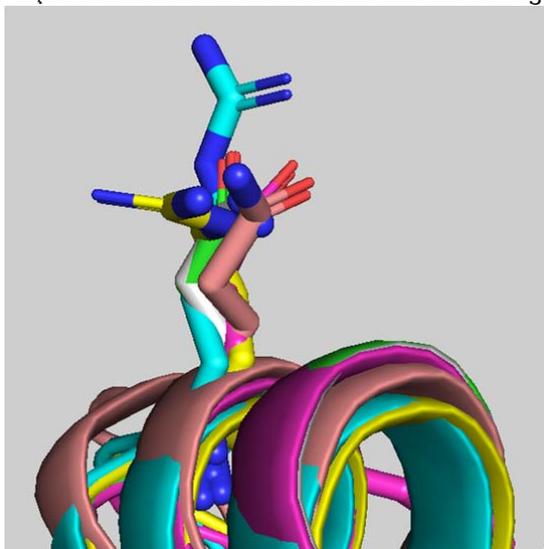
Figura 5 - Conformação tridimensional das cadeias laterais de Leucina e Valina.



Legenda: branco - 3CLB; verde - 3HBB; magenta - 3KJS; rosa - 2H2Q; ciano - 3INV; amarelo - 5T7O.

Por fim, na modificação R137Q, indicada na Figura 6, também em uma região de alfa-hélice, a cadeia linear de três carbonos da arginina é substituída por uma cadeia linear de quatro carbonos, além disso, a guanidina é substituída por uma amida no resíduo de glutamina. Essa mudança, embora impacte na eletronegatividade da cadeia lateral, além do seu volume, não resultou em alterações conformacionais da proteína como um todo que pudessem ser observadas na análise dos cristais.

Figura 6 - Conformação tridimensional das cadeias laterais de Arginina e Glutamina.



Legenda: branco - 3CLB; verde - 3HBB; magenta - 3KJS; rosa - 2H2Q; ciano - 3INV; amarelo - 5T7O.

O resultado observado neste trabalho reflete o registrado na literatura: as três variações de aminoácidos observadas entre os modelos de *T. cruzi* não estão em

regiões relativas à atividade enzimática e são substituições entre aminoácidos de características físico-químicas similares, que também não impactam a estrutura tridimensional da proteína (MACHADO; AYALA, 2002).

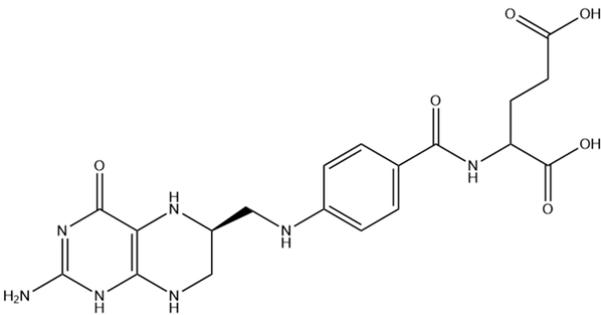
Sendo assim, as diferenças observadas, embora devam ser pontuadas, não interferem nas análises realizadas posteriormente, uma vez que o foco de estudo deste trabalho foi a estrutura tridimensional da proteína, mais especificamente a região do sítio ativo e os pontos de interação com ligantes e, como discutido, tal região permanece inalterada em ambas mutações da proteína de *T. cruzi* analisadas.

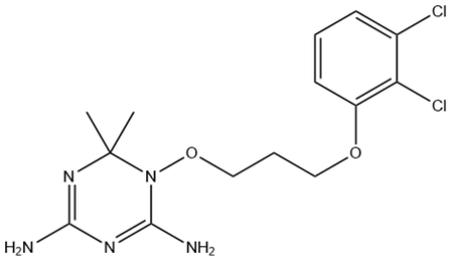
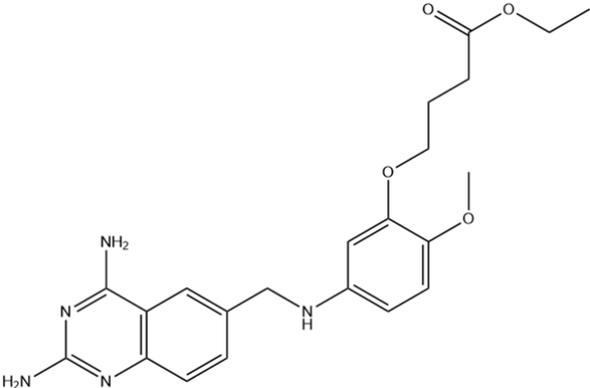
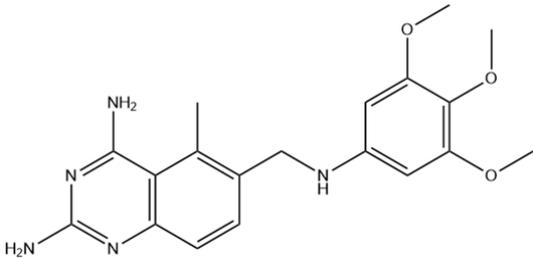
4.4. ANÁLISE DOS LIGANTES DISPONÍVEIS

Os modelos cristalográficos selecionados para este trabalho apresentavam diferentes ligantes co-cristalizados. Desta forma, optou-se por avaliar a conformação de cada um destes ligantes, de forma comparativa, no acoplamento com a proteína.

Dentre os ligantes selecionados, o tetrahydrofolato é o resultado da catálise natural da DHFR, o trimetrexato é um inibidor competitivo disponível no mercado e C50 e DQ1 são inibidores competitivos não comercializados.

Tabela 02 - Estrutura Molecular dos Ligantes de Dihidrofolato Redutase

Ligante	Estrutura Molecular
Tetrahydrofolato (Fonte: PubChem CID 5460413)	 <p style="text-align: right;">(continua)</p>

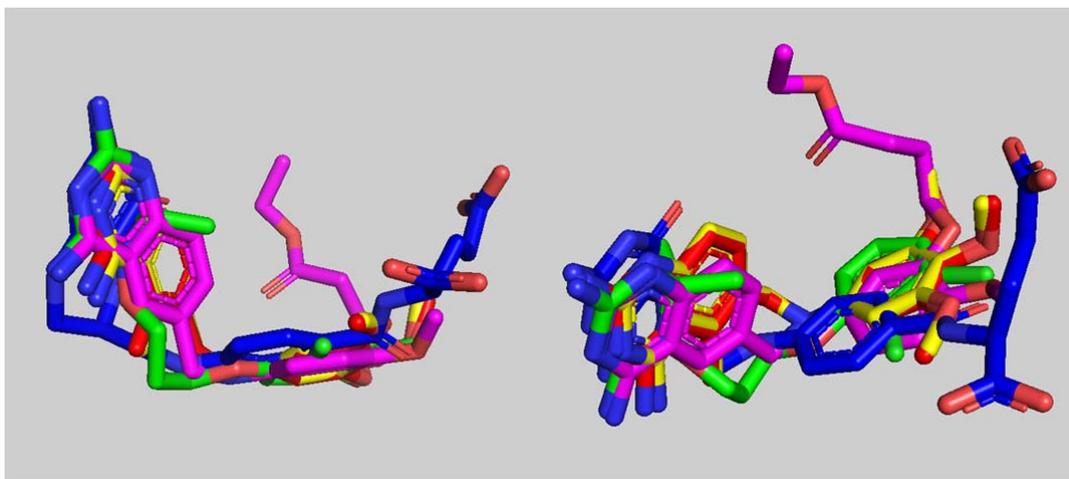
Ligante	Estrutura Molecular
(continuação)	
C50 (Fonte: PubChem CID 46829298)	
DQ1 (Fonte PubChem CID 16038397)	
Trimetrexato (Fonte PubChem CID 59961053)	
	(conclusão)

Para realizar a análise comparativa dos ligantes co-cristalizados nos diferentes modelos de DHFR, foi realizado o alinhamento das estruturas proteicas no PyMol. Após o alinhamento, os ligantes foram extraídos mantendo-se a orientação original do cristal, permitindo observar-se somente os ligantes e seu posicionamento espacial. É possível observar que todos eles apresentam geometrias similares e os grupos eletronegativos, normalmente formados pelos anéis aromáticos, encontram-se sempre nas mesmas posições, como pode ser observado nas Figuras 7 e 8, que mostram cada um dos ligantes em rotações diferentes para melhor observação em 2D de uma estrutura tridimensional.

A maior variabilidade está nos substituintes do anel benzeno, que possuem diferentes características e tamanhos, indicando, provavelmente, tratar-se de uma região de maior espaço e flexibilidade da proteína, permitindo o interação com

cadeias longas de carbono substituídas por átomos de oxigênio, o que confere maior eletronegatividade à molécula. Principalmente quando observado as cadeias laterais do tetrahidrofolato e DQ1, que possuem grande flexibilidade e ocupam espaços em direções opostas. Considerando que o tetrahidrofolato é o produto da reação catalítica da enzima, o fato do DQ1 explorar uma região oposta pode indicar a presença de um bolsão cuja interação com ligantes resulte em atividade inibidora.

Figuras 7 e 8 - Ligantes dos modelos cristalográficos.



Legenda: amarelo e vermelho - trimetrexato (3HBB e 3CLB); rosa - DQ1 (3KJS); azul - tetrahidrofolato (5T7O); verde - C50 (3INV). Imagens geradas pelo Pymol.

O modelo cristalográfico 2H2Q (representado em ciano nas figuras 4 a 6) apresentava como ligante a desoxiuridina, acoplada no sítio ativo da porção timidilato sintase da proteína e, portanto, não aparece nesta análise, uma vez que todos os outros ligantes - e o foco deste trabalho - estão na porção dihidrofolato redutase.

4.5. ANÁLISE DOS POSSÍVEIS PONTOS DE INTERAÇÃO

O estudo dos pontos de interação da proteína-alvo, dihidrofolato redutase, com os diferentes ligantes, permitiu que fossem feitas algumas análises relativas ao sítio ativo da proteína. Como os ligantes são inibidores competitivos, essas interações podem ser exploradas para entender melhor o mecanismo de ação ou, mesmo, planejar novas moléculas que atuam de maneira semelhante.

Com essa análise, é possível identificar quais aminoácidos são capazes de realizar interações intermoleculares com as moléculas acopladas no sítio ativo:

região da proteína onde a reação catalítica ocorre e, para que tal reação seja possível, é preciso que o substrato interaja com os aminoácidos que compõem a região do sítio ativo, de forma a acoplar-se a ele. O acoplamento do ligante na proteína é favorecido quanto maiores e mais fortes forem as interações intermoleculares, ou seja, maior poderá ser a afinidade do ligante pela proteína.

Com o uso do programa Chimera X, foram identificados os resíduos de aminoácidos que estivessem em um raio de pelo menos 5 Å da molécula acoplada. Essa distância foi escolhida pois todas as interações intermoleculares quimicamente possíveis dependem de distâncias iguais ou menores a 5 Å para ocorrerem, sendo que quanto menor a distância, mais forte a interação. (GODDARD, *et al.*, 2018; KOHLHOFF, *et al.* 2009).

Para esta análise o modelo 3HBB foi desconsiderado, uma vez que ele possui o mesmo ligante do modelo 3CLB (trimetrexato), porém somente 2 subunidades proteicas, ao invés das 4 subunidades apresentadas por 3CLB.

Antes da determinação dos pontos de interação, todos os modelos restantes (3CLB, 3KJS, 3INV e 5T7O) foram alinhados, tendo como base o modelo cristalográfico 3CLB. Como os modelos apresentavam pelo menos duas regiões de sítio ativo, foi escolhida a região de melhor sobreposição dos modelos cristalográficos para análise dos resultados.

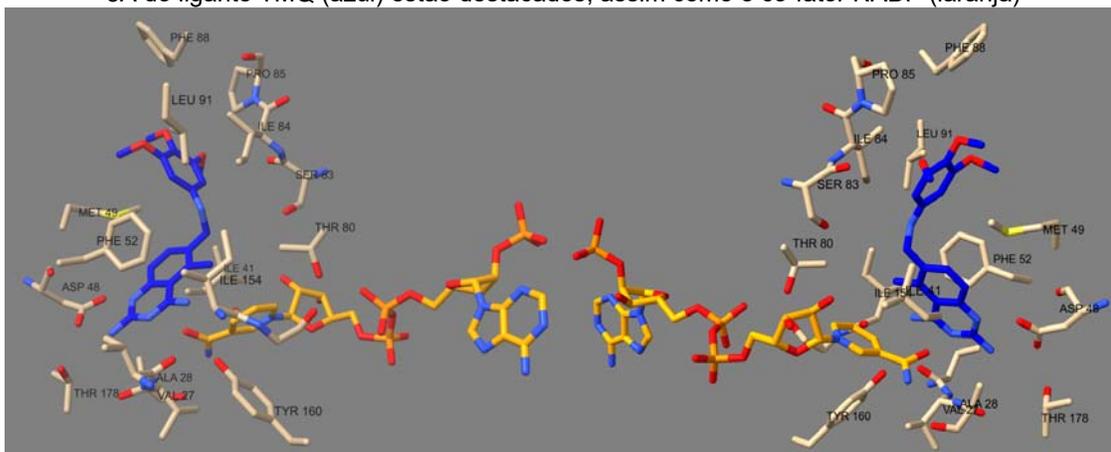
Os aminoácidos que apresentavam pelo menos um átomo dentro do raio de 5Å do ligante estão destacados nas Figuras 9 a 16, cada par de figuras representa o mesmo modelo e ligantes, de forma espelhada, ou seja, o modelo foi rotacionado 180° para que fosse possível visualizar todos os resíduos identificados.

Figuras complementares estão no apêndice B, onde estão também representados os pontos de cada interação identificada e suas respectivas distâncias.

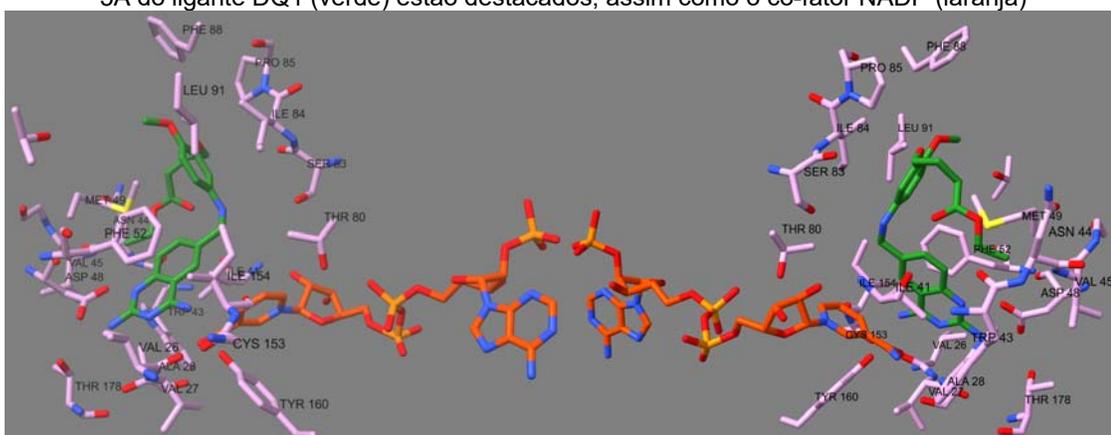
Além disso, todas as possíveis interações identificadas foram listadas na tabela 3, de forma a complementar a visualização dos resultados.

O modelo 3CLB, cujo ligante é o trimetrexato, apresentou 16 pontos de interação entre a proteína e o ligante, representados nas Figuras 9 e 10. O modelo 3KJS (figuras 11 e 12) apresentou 19 pontos de interação entre o DQ1 e a macromolécula. Para o modelo 3INV foram identificadas 18 interações entre C50 e DHFR, ilustradas nas Figuras 13 e 14. Por fim, o modelo do ligante THF, 5T7O, apresentou 21 possíveis interações (Figuras 15 e 16.)

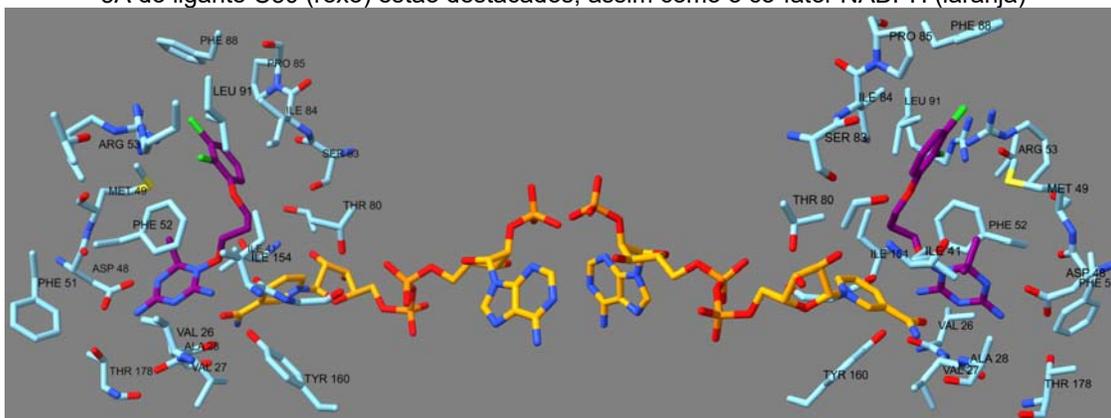
Figuras 9 e 10 - Modelo cristalográfico 3CLB, os aminoácidos com pontos de interação a menos de 5Å do ligante TMQ (azul) estão destacados, assim como o co-fator NADP (laranja)



Figuras 11 e 12 - Modelo cristalográfico 3KJS, os aminoácidos com pontos de interação a menos de 5Å do ligante DQ1 (verde) estão destacados, assim como o co-fator NADP (laranja)



Figuras 13 e 14 - Modelo cristalográfico 3INV, os aminoácidos com pontos de interação a menos de 5Å do ligante C50 (roxo) estão destacados, assim como o co-fator NADPH (laranja)



Figuras 15 e 16 - Modelo cristalográfico 5T7O, os aminoácidos com pontos de interação a menos de 5Å do ligante THG (magenta) estão destacados, assim como o co-fator NADP (laranja)

Figura 17 - Modelo cristalográfico 3CLB, as possíveis interações de hidrogênio com o ligante TMQ (azul) estão destacadas, assim como o co-fator NADP (laranja)

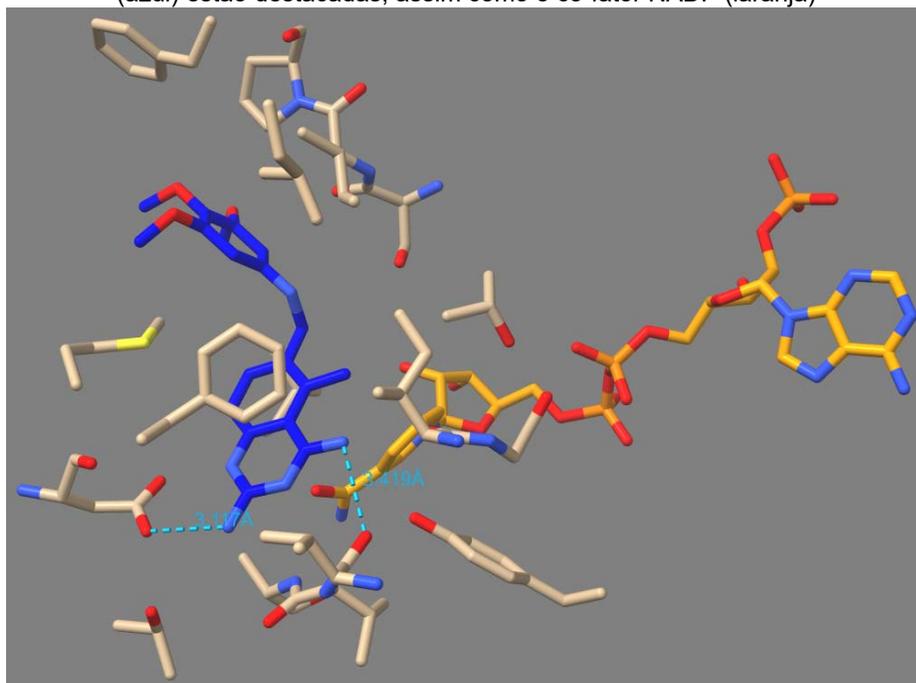


Figura 18 - Modelo cristalográfico 3KJS, as possíveis interações de hidrogênio com o ligante DQ1 (verde) estão destacadas, assim como o co-fator NADP (laranja). O aminoácido PHE52 foi ocultado para melhor visualização.

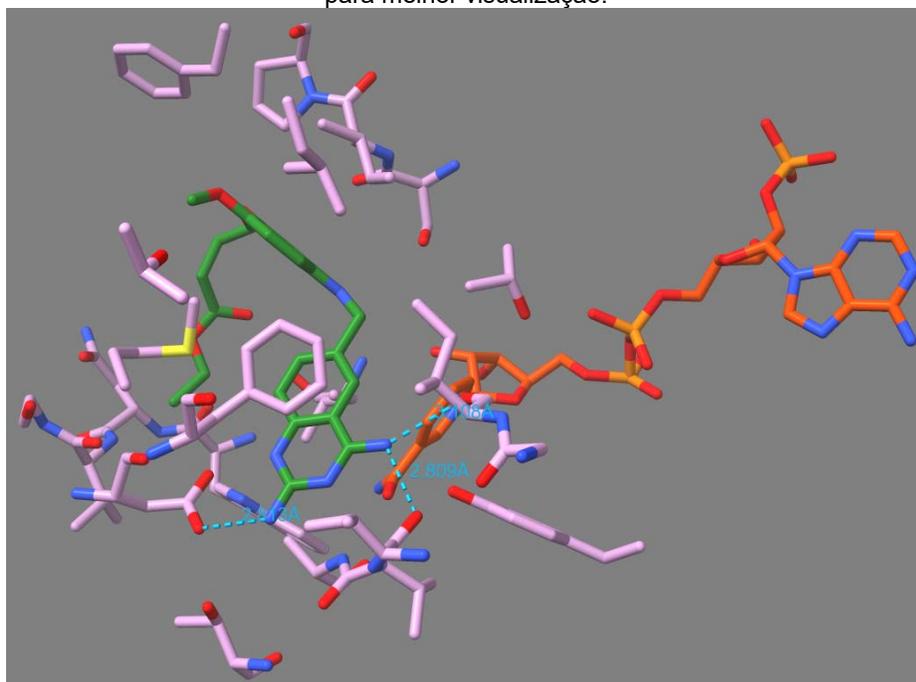


Tabela 03 - Lista de aminoácidos com possíveis interações para cada um dos modelos avaliados.

Aminoácido	Modelo cristalográfico (ligante)			
	3CLB (TMQ)	3KJS (DQ1)	3INV (C50)	5T7O (THG)
VAL26	H	H	H	H
VAL27	X	X	X	X
ALA28	X	X	X	X
ILE41	X	X	X	
TRP43		X		
ASN44		X		
VAL45		X		
ASP48	H	H	H	H
MET49	X	X	X	X
LYS50				X
PHE51			X	X
PHE52	X	X	X	X
ARG53			X	H
THR56				X
THR80	X	X	X	X
SER83	X	X	X	X
ILE84	X	X	X	X
PRO85	X	X	X	X
PHE88	X	X	X	X
LEU91	X	X	X	X
PRO92				X
ARG94				H
ILE154	X	H	H	X
TYR160	X	X	X	X
THR178	X	X	X	X

Legenda: os pontos de interação estão destacados com um X, os pontos de interação de hidrogênio destacados com um H.

O modelo cristalográfico 5T7O foi o que apresentou o maior número de possíveis interações entre o ligante THG e a proteína. Embora o número de interações por si só não possa ser correlacionado diretamente a uma maior afinidade do ligante pelo alvo, o tetrahydrofolato é o produto da catálise da dihydrofolato, sendo muito similar estruturalmente ao dihydrofolato: a diferença entre o ligante analisado e o substrato natural da enzima é somente a troca de uma insaturação por um hidrogênio.

Como a afinidade do dihydrofolato pela dihydrofolato redutase é produto da evolução da proteína ao longo do tempo para otimizar sua atividade, é de se esperar que existam mais pontos de interação entre ligante-receptor e, devido à similaridade estrutural, os mesmos pontos de interação presentes para o dihydrofolato estão disponíveis, em sua maioria, para o tetrahydrofolato.

Os modelos 3CLB (ligante TMQ), 3KJS (DQ1) apresentam interações de hidrogênio nos nitrogênios substituintes do anel pirimidina com os resíduos de aminoácidos VAL26 e ASP48, representados nas figuras 17 e 18 e detalhados nas figuras 21 e 22, respectivamente.

Figura 21 - Pontos de interação do ligante TMQ (azul) com VAL26 e ASP48

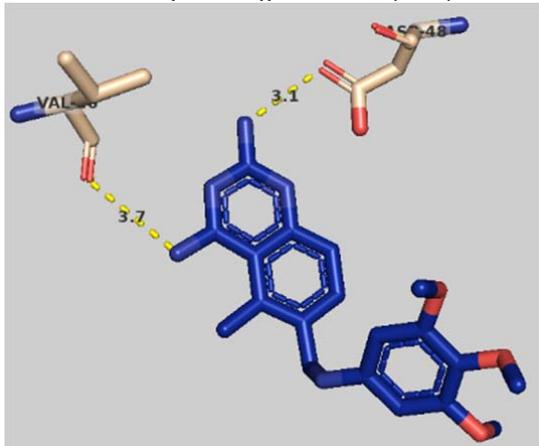
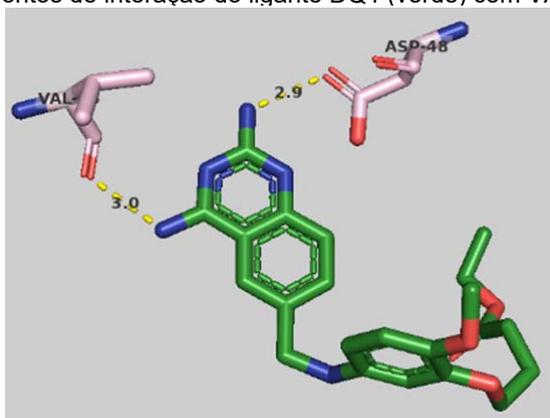
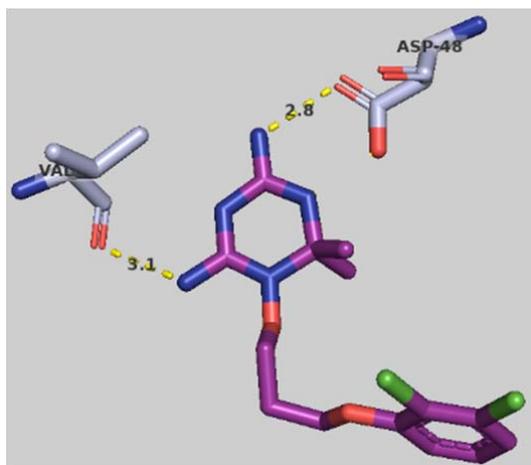


Figura 22 - Pontos de interação do ligante DQ1 (verde) com VAL26 e ASP48



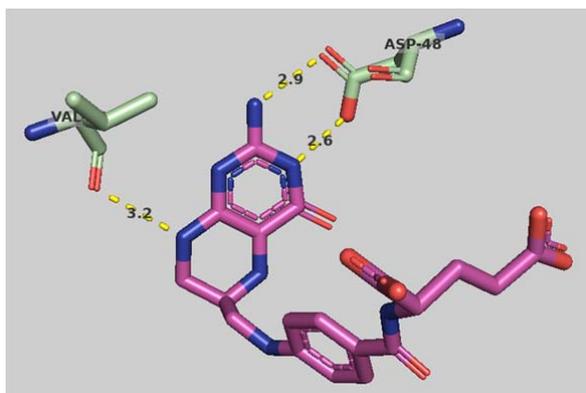
O modelo 3INV (C50), apresentou o mesmo ponto de interações (Figuras 19 e 23), também com os nitrogênios substituintes do anel, no entanto, ao invés de possuir uma pirimidina fundida ao anel benzeno, a pirimidina possui um carbono substituído por um nitrogênio e não há o anel benzeno.

Figura 23 - Pontos de interação do ligante C50 (roxo) com VAL26 e ASP48



No caso do modelo 5T7O, o ligante THG realiza as interações com o nitrogênio substituinte do anel, e com os nitrogênios dos próprios anéis pirimidina fundidos, conforme Figuras 20 e 24.

Figura 24 - Pontos de interação do ligante THG (magenta) com VAL26 e ASP48



O ácido aspártico 48 pode interagir tanto como doador de ligação hidrogênio, na porção álcool, quanto como receptor na porção cetona, uma vez que existe ressonância entre os elétrons do ácido carboxílico e a posição do hidrogênio não é fixa. O mesmo ocorre com os nitrogênios do substituinte e do anel do ligante, que podem atuar de ambas formas, dependendo da posição do hidrogênio.

Já a interação com a valina, sugerida pelo programa, ocorre pelo oxigênio do *backbone* do aminoácido, que atuaria como átomo receptor de hidrogênio, enquanto os nitrogênios dos heteroátomos dos ligantes atuam como doadores da interação de hidrogênio.

O mesmo ocorre para a interação de hidrogênio entre o aminoácido ILE154 com os ligantes DQ1 (3KJS) e C50 (3INV), representados nas figuras 25 e 26, conforme sugerido, a interação ocorreria entre a cetona da isoleucina, como grupo receptor, e o nitrogênio do heteroátomo dos ligantes, como átomos doadores.

As interações entre ligantes e o *backbone* da proteína não costumam ocorrer, uma vez que a região do *backbone* é, normalmente, estericamente impedida pelos resíduos dos aminoácidos próximos, sendo assim, apesar da indicação do programa, essas interações foram desconsideradas, uma vez que a probabilidade de realmente ocorrer é baixa.

Figura 25 - Pontos de interação do ligante DQ1 (verde) com ILE154

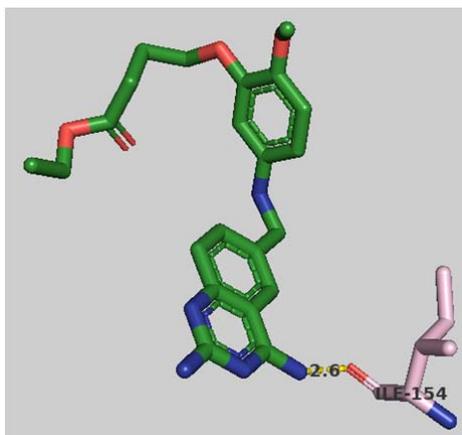
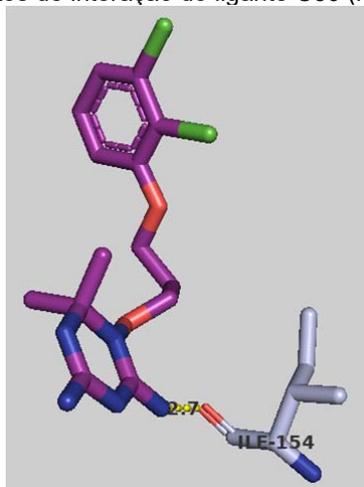


Figura 26 - Pontos de interação do ligante C50 (roxo) com ILE154



As interações entre os anéis nos mesmos pontos, permite também identificar que, apesar da substituição de átomos do anel, ou da sua simplificação, no caso do DQ1, a conformação espacial é mantida, reproduzindo os mesmos pontos de interação, ou seja, o anel pirimidina fundido ao anel benzeno não é essencial para afinidade do ligante pelo alvo, embora a estrutura dos substituintes nitrogenados deva ser preservada.

As argininas nas posições 53 e 94 fazem interações de hidrogênio somente com o tetrahydrofolato (PDB 5T7O), uma vez que esse ligante é o único que possui átomos capazes de realizar interações de hidrogênio nesta região (Figura 27). O grupo substituinte do anel benzeno possui uma amida e dois ácidos carboxílicos, sendo estes responsáveis por realizar as interações de hidrogênio com as cetonas das argininas, gerando um efeito de ressonância tanto no ligante quanto na proteína. Esse grupo substituinte do benzeno é importante justamente para explorar uma

região da proteína que não é explorada pelos outros ligantes, o único ligante que apresenta uma cadeia carbônica capaz de explorar tal região é o DQ1 (PDB 3INV), embora ele possui um éster ao invés do ácido carboxílico, que não permitiria o efeito de ressonância, ainda que pudesse ocorrer a interação de hidrogênio com o oxigênio da cetona e o grupo guanidínico da ARG53, no entanto, ao analisar a estrutura tridimensional é possível observar que esta cadeia lateral do ligante ocupa um espaço oposto à posição das argininas (Figura 28).

Figura 27 - Pontos de interação do ligante THG (magenta) com ARG53 e ARG94

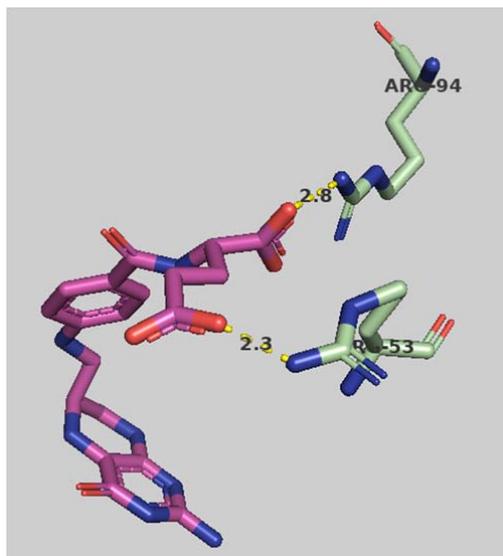
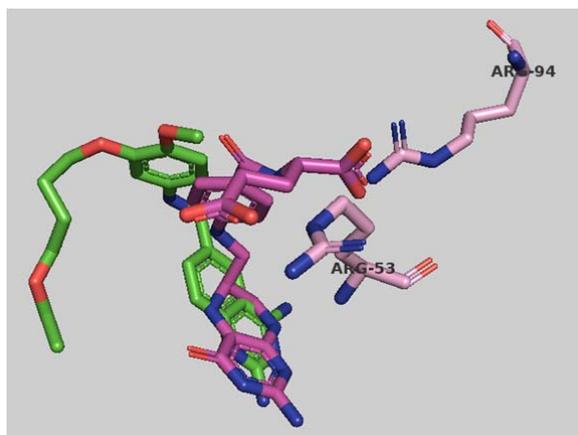


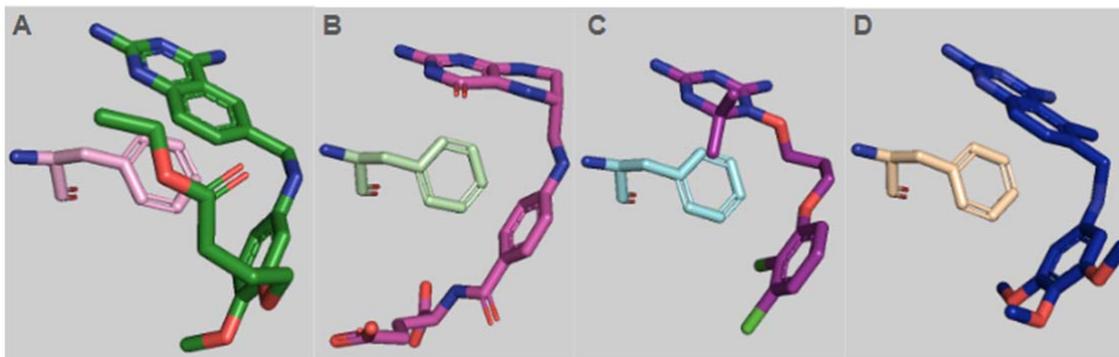
Figura 28 - Comparação das posições dos ligantes THG (magenta) e DQ1 (verde) em relação aos aminoácidos ARG53 e ARG94



Além dos pontos destacados, faz-se importante ressaltar o papel dos anéis nas interações com a fenilalanina 52 da proteína: todos os ligantes possuem dois grupos cíclicos, um anel benzeno e um heterocíclico insaturado, formado por dois anéis fundidos para a maioria dos ligantes analisados, exceto para o C50. Conforme

representado na figura 29, os ligantes apresentam uma torção que confere uma conformação côncava às moléculas, sendo que o anel da fenilalanina se encaixa justamente no centro desta formação, entre os dois grupos cíclicos.

Figura 29 - Posição dos ligantes em relação à PHE52



Legenda: A - modelo 3KJS com ligante DQ1 (verde); B- modelo 5T7O com ligante THG (magenta); C - modelo 3INV com ligante C50 (roxo); D - modelo 3CLB com ligante TMQ (azul)

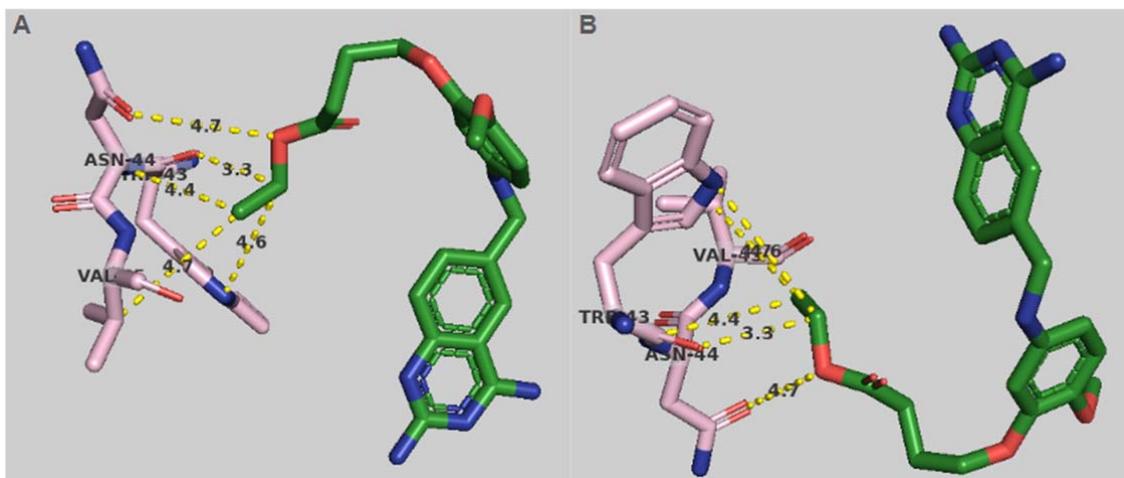
Essa interação ocorre entre todos os modelos analisados, independente na simplificação do anel, no caso do C50 (Figura 29 C) indicando ser possível simplificar a molécula ligante, desde que sejam mantidos dois anéis insaturados conectados por cadeia linear que permita a ocorrência da torção observada e preserve a distância entre os grupos.

É possível observar que existe uma torção entre os ângulos, a fenilalanina não encaixa de forma paralela ao anel benzeno ou ao anel heterociclo de cada modelo, formando uma interação *pi-stacking*, no entanto, por se tratarem de anéis insaturados com ressonância de elétrons, as nuvens eletrônicas da fenilalanina, do heterociclo insaturado e do benzeno são capazes de interagir entre si, estabilizando o ligante nesta conformação observada e aumentando a interação entre as duas moléculas.

O ligante DQ1 apresenta uma característica diferente dos demais ligantes a cadeia lateral do anel benzeno com 8 átomos lineares e um ácido carboxílico, onde todos os outros ligante não apresentam substituintes na mesma posição. Ao analisar os pontos de interação do DQ1, ele está próximo a resíduos de três aminoácidos que não são observados nas análises das demais moléculas: TRP43, ASN44 e VAL45.

As distâncias entre o DQ1, modelo 3KJS, e os resíduos de tais aminoácidos estão representadas na figura 30, a figura A e B correspondem a mesma representação em posições diferentes para melhor visualização.

Figura 30 - Posição do ligante DQ1 em relação aos aminoácidos TRP43, ASN44 e VAL45



Os aminoácidos triptofano e asparagina possuem resíduos volumosos com heteroátomos, além de possuírem características opostas: enquanto o triptofano é um aminoácido apolar com um resíduo cíclico, insaturado, com ressonância de elétrons, enquanto a asparagina é um aminoácido polar neutro formado por uma carboxamida. As variadas possibilidades de interações com tais grupos químicos tornam interessante esta região da proteína, que é pouco explorada pelos demais ligantes.

Embora nem todo aminoácido presente no sítio ativo seja diretamente responsável por interações intermoleculares com o substrato, conhecer os aminoácidos que são capazes de tais interações se faz importante para o planejamento de novos ligantes, ou inibidores, justamente por direcionar tal planejamento: sabendo-se que existe um ponto possível de interação na macromolécula alvo é possível planejar uma molécula que explore tal ponto, aumentando sua afinidade pela proteína.

Ou seja, ao desenvolver um ligante para a dihidrofolato redutase, a molécula precisa ter geometria e composição química compatíveis com os pontos de interação identificados.

4.6. MODELOS CRISTALOGRÁFICOS DE *H. SAPIENS*

A busca pelos modelos cristalográficos da dihidrofolato redutase humana seguiu os mesmos critérios da sua homóloga parasita: utilizou-se como chave de pesquisa as palavras “dihydrofolate reductase” ou “DHFR” e cujo organismo de origem fosse “Homo sapiens”, o que obteve 87 resultados.

Para a escolha dos modelos de *H. sapiens*, além dos critérios de qualidade de cristais considerados e descritos anteriormente, foram levados em consideração os ligantes co-cristalizados, priorizando-se os modelos que apresentassem os mesmos ligantes dos modelos selecionados para *T. cruzi*, o que infelizmente não foi possível uma vez que nenhum dos 87 resultados apresentava como ligantes o tetrahydrofolato, o trimetrexato, o C50 ou o DQ1.

Optou-se então por realizar um novo filtro, considerando os modelos com um ligante que apresentasse ação inibidora conhecida e que pudesse ser comparado aos ligantes trimetrexato, C50 e DQ1 - ou seja, o metotrexato (MTX) -, e um ligante sem ação inibidora que pudesse ser comparado ao tetrahydrofolato, neste caso o folato (FOL). Cabe ressaltar que foram priorizados os modelos que fossem co-cristalizados somente com o cofator NADP, sem outros possíveis ligantes.

Como os resultados de modelos cristalográficos disponíveis foram todos obtidos a partir de difração de raio-X, mantiveram-se os mesmos critérios de qualidade anteriores, ou seja, resolução > 3,0, R-free - R-Work > 0,23 e, no caso dos Outliers de Ramachandran, os modelos foram descartados caso a posição dos outliers coincidissem com o sítio ativo ou com alguma posição de interação com ligantes ou co-fator. Caso mais de um modelo atendesse aos requisitos, seria escolhido o de melhor resolução.

Os resultados da pesquisa encontram-se na Tabela 04. Os modelos de cristais descartados estão sublinhados, assim como o parâmetro responsável por tal descarte. Os modelos cristalográficos escolhidos para a sequência deste trabalho foram 1U72 e 4M6K.

Tabela 04 - Modelos cristalográficos de DHFR disponíveis no PDB para *H. sapiens* após escolha dos ligantes.

PDB ID	N° de cadeias	Peso Molecular (kDa)	Resolução (Å)	Diferença R-Free R-Work	Outliers (backbone / side chain)	Ligantes (ID - descrição)
<u>1DLS</u>	1	22,6	2,3	<u>Dados não informados</u>	1 / 7	NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida MTX - metotrexato
<u>3EIG</u>	1	22,7	1,7	0,04855	0 / 3	<u>MTX - metotrexato</u> <u>SO4 - íon sulfato</u> <u>CD - íon cádmio</u>
1U72	1	22,55	1,9	0,054	0 / 25	NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida MTX - metotrexato
<u>1DHF</u>	2	43,58	2,3	Dados não informados	<u>3 / 59</u>	<u>FOL - ácido fólico</u>
2W3M	2	45,34	1,6	0,041	0 / 2	NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida FOL - ácido fólico
4M6K	1	22,76	1,396	0,0393	0 / 3	FOL - ácido fólico NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida GOL - glicerol
7ESE	1	24,68	1,85	0,445	0 / 5	NAP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida FOL - ácido fólico
<u>2W3B</u>	2	46,15	1,27	0,032	1 / 4	NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida <u>VG9 - 6-[[[2,5-Dietoxifenil]amino]metil]-5-metilpirido[2,3-D]pirimidina-2,4-diamina</u> FOL - ácido fólico K - íon potássio

(continua)

PDB ID	N° de cadeias	Peso Molecular (kDa)	Resolução (Å)	Diferença R-Free R-Work	Outliers (backbone / side chain)	Ligantes (ID - descrição)
(continuação) <u>1DRF</u>	1	21,98	2,0	Dados não informados	4 / 37	<u>SO4 - íon sulfato</u> <u>FOL - ácido fólico</u>
<u>6VCJ</u>	4	92,25	2,34	0,0466	0 / 2	LG3 - pirimidina-2,4-diamina NPX - ácido (2R)-2-(6-metoxinaftalen-2-il)pronaóico NAP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida FOL - ácido fólico
<u>4KEB</u>	2	45,05	1,45	0,0388	0 / 0	NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida <u>1QZ - 6-etil-5-[(3S)-3-[3-(isoquinolin-5-il)-5-metoxifenil]but-1-in-1-il]pirimidina-2,4-diamina</u> FOL - ácido fólico
<u>4KFJ</u>	2	45,28	1,76	0,0489	0 / 8	NDP / NADPH - fosfato de dinucleotídeo de adenina e nicotinamida MG - íon magnésio CL - íon cloreto EOH - etanol FOL - ácido fólico <u>1R0 - 6-etil-5-[(3S)-3-[3-(isoquinolin-5-il)-5-metoxifenil]prop-1-in-1-il]pirimidina-2,4-diamina</u>

(conclusão)

4.7. ANÁLISE COMPARATIVA DA SEQUÊNCIA PRIMÁRIA

Assim como para os modelos de *T. cruzi*, foi feito um alinhamento das sequências de aminoácidos, com a ferramenta Clustal Omega. Onde não foram observadas alterações nos aminoácidos das duas sequências, conforme indicado na Figura 31.

Diferente da proteína parasita, que apresenta mutações naturais na dihidrofolato redutase, a proteína humana é conservada, sendo idêntica entre as duas sequências. A diferença de uma metionina no início da sequência, presente no modelo 4M6K e não no modelo 1U72, pode ser explicada como uma falha na expressão ou no sequenciamento da proteína, não representando uma mutação ou alteração da sequência proteica. Por se tratar do aminoácido inicial e não impactar nas análises realizadas, a metionina será desconsiderada e a posição dos aminoácidos descrita neste trabalho considerará a contagem a partir da valina como aminoácido na posição 1.

Figura 31 - Comparação das sequências de aminoácidos dos modelos cristalográficos 4M6K e 1U71 obtida pelo Clustal Omega

```

1U72      -VGSLNCIVAVSQNMGIGKNGDLPWPLRNEFRYFQRMTTSSVEGKQNLVIMGKKTWFS  59
4M6K      MVGSLNCIVAVSQNMGIGKNGDLPWPLRNEFRYFQRMTTSSVEGKQNLVIMGKKTWFS  60
          *****

1U72      IPEKNRPLKGRINLVLSRELKEPPQGAHFLSRSLDDALKLTEQPELANKVDMMVIVGGSS  119
4M6K      IPEKNRPLKGRINLVLSRELKEPPQGAHFLSRSLDDALKLTEQPELANKVDMMVIVGGSS  120
          *****

1U72      VYKEAMNHPGHLKLVTRIMQDFESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKYKF  179
4M6K      VYKEAMNHPGHLKLVTRIMQDFESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKYKF  180
          *****

1U72      EVYEKND      186
4M6K      EVYEKND      187
          *****

```

Adicionalmente, foi realizado o alinhamento da sequência de DHFR humana e de *T. cruzi*, para isso utilizou-se novamente o Clustal Omega. Como a proteína do parasita apresenta dois domínios, considerou-se somente a sequência do domínio correspondente à dihidrofolato redutase, excluindo-se a porção referente à timidilato sintase, ou seja, manteve-se somente os 231 aminoácidos iniciais da sequência.

Optou-se por realizar o alinhamento com dois modelos de *T. cruzi* de forma que as mutações discutidas anteriormente fossem consideradas nesta análise

comparativa, portanto, o modelo 3CLB foi utilizado no alinhamento além do modelo 5T7O.

Figura 32 - Comparação das sequências de aminoácidos dos modelos cristalográficos de *H. sapiens* e *T. cruzi* obtida pelo Clustal Omega

3CLB Chains	MSLFKIRMPETVAEGTRLALRAFSLVVAVDERGGIGDGRSIPWNVPE -DMKFFRDVTTKL	59
5T7O Chains	MSLFKIRMPETVAEGTRLALRAFSLVVAVDEHGGIGDGRSIPWNVPE -DMKFFRDLTTKL	59
4M6K Chain	-----MVGSLNCIVAVSQNMGIKNGDLPWPPLRNEFRYFQRMTTT-	41
1U72 Chain	-----VGSLNCIVAVSQNMGIKNGDLPWPPLRNEFRYFQRMTTT-	40
	: :. :.***. . .***. . :** . :. :. :.***.	
3CLB Chains	RGKNVKPSPAKRNAVVMGRKTWDSIPPKFRPLPGRLNVVLSSTLTTQHLLDGLPDEEKRN	119
5T7O Chains	RGKNVKPSPAKRNAVVMGRKTWDSIPPKFRPLPGRLNVVLSSTLTTQHLLDGLPDEEKRN	119
4M6K Chain	-----SSVEGKQNLVIMGKKTWFSIPEKNRPLKGRINLVLSRELKEPPQ-----	85
1U72 Chain	-----SSVEGKQNLVIMGKKTWFSIPEKNRPLKGRINLVLSRELKEPPQ-----	84
	. . :.* * :.* :.* * * * * * * * * * * * * * * * * .	
3CLB Chains	LHADSIVAVNGGLEQALQLLASPNYTPSIETVYICIGGGSVYAEALRPPCVHLLQAIYRRT	179
5T7O Chains	LHADSIVAVNGGLEQALRLLASPNYTPSIETVYICIGGGSVYAEALRPPCVHLLQAIYRRT	179
4M6K Chain	----GAHFLSRSLDDALKLTEQPELANKVDMMWIVGGSSVYKEAMNHPGHLKLFV---TR	138
1U72 Chain	----GAHFLSRSLDDALKLTEQPELANKVDMMWIVGGSSVYKEAMNHPGHLKLFV---TR	137
	. :. . :.* :.* :.* * * * * * * * * * * * * * * * * .	
3CLB Chains	IRASESSCSVFFR-----VPE-----SGTEAAAGIEWQRETISEELTSANGNETKY	225
5T7O Chains	IRASESSCSVFFR-----VPE-----SGTEAAAGIEWQRETISEELTSANGNETKY	225
4M6K Chain	IMQDF-ESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKYKFEVYEKND-----	187
1U72 Chain	IMQDF-ESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKYKFEVYEKND-----	186
	* * .	
3CLB Chains	YFEKLI 231	
5T7O Chains	YFEKLI 231	
4M6K Chain	----- 187	
1U72 Chain	----- 186	

A primeira divergência entre as proteínas encontra-se no fato da dihidrofolato redutase de *T. cruzi* ser uma proteína bifuncional, apresentando ação de dihidrofolato redutase e de timidilato sintase, enquanto a enzima humana exerce somente a primeira função. Para que as duas subunidades catalíticas sejam acopladas de forma funcional a DHFR-TS possui um peptídeo conectando a porção N-terminal DHFR com a porção C-terminal TS (BELTRAN-HORTELANO, 2022).

Enquanto a porção timidilato sintase é preservada entre as duas espécies e não foi analisada neste estudo, a porção dihidrofolato redutase apresenta maior variabilidade entre os organismos ou seja, existem diferenças entre a enzima de diferentes organismos que podem ser exploradas na busca por seletividade, objetivo deste estudo (BELTRAN-HORTELANO, 2022).

As diferenças analisadas estão principalmente no tamanho das proteínas, enquanto a proteína do parasita apresenta 231 aminoácidos, a proteína humana apresenta somente 186, ou seja, existem aminoácidos na proteína de *T. cruzi* que

não possuem correspondentes na estrutura humana para pareamento. Quanto às regiões pareadas, é possível observar que existe uma certa preservação dos aminoácidos, 58 (31% da sequência humana) dos aminoácidos são idênticos entre as sequências.

Como análise complementar cabe avaliar quais dos aminoácidos divergentes encontram-se próximos a pontos de interação com o ligante, de forma que possam ser explorados no favorecimento de interações com a proteína do parasita.

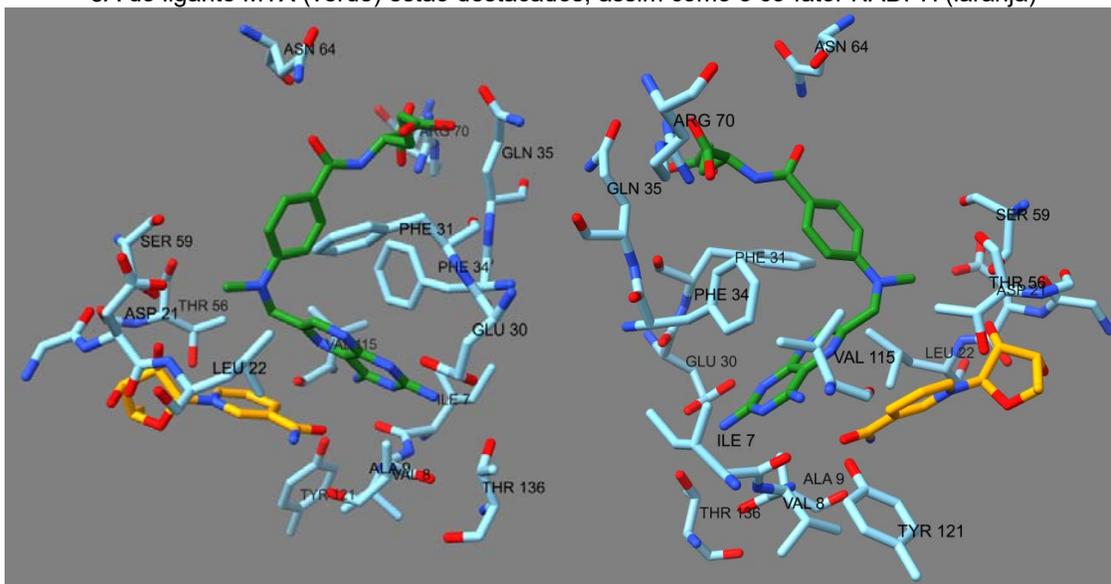
4.8. ANÁLISE DOS POSSÍVEIS PONTOS DE INTERAÇÃO DA PROTEÍNA HUMANA

Assim como para a proteína de *T. cruzi*, os pontos de interação entre os ligantes e os modelos cristalográficos da proteína humana foram determinados utilizando-se o programa Chimera, considerando como pontos de interação átomos da proteína e do ligante que estivessem dentro de um raio de 5,0 Å.

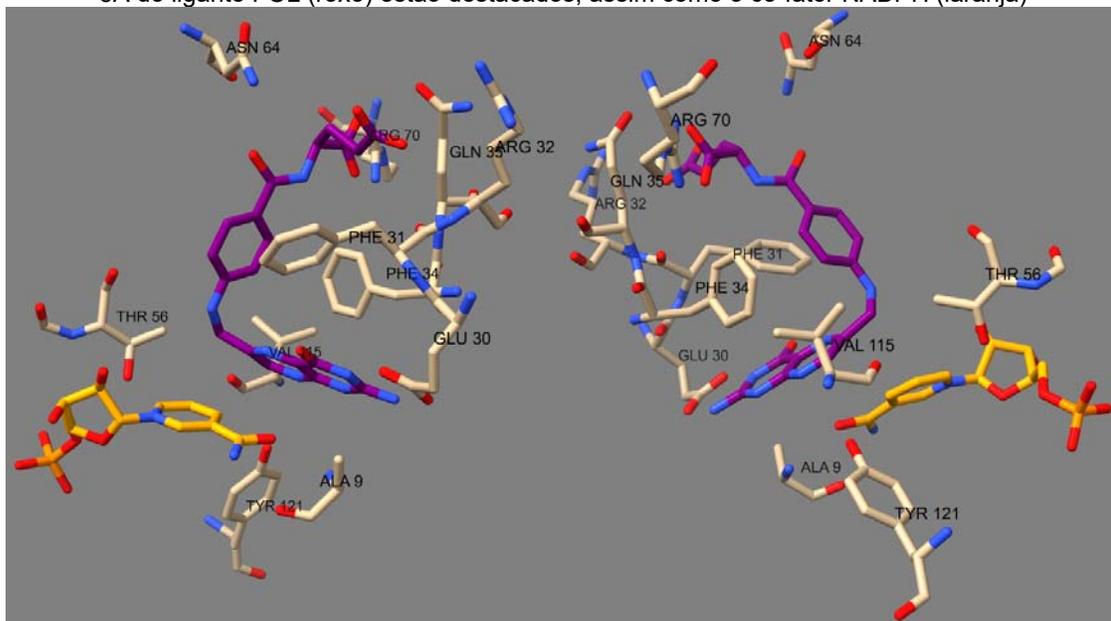
As Figuras 33 e 34 representam o modelo 1U72, do ligante MTX, e as Figuras 34 e 35 representam o modelo 4M6K, ligante FOL. As imagens são espelhos uma da outra, representando a rotação de 180° do modelo, de forma que permita a observação de pontos diferentes de interação. Imagens detalhando os pontos de interação identificados e suas respectivas distâncias estão representados no apêndice C.

A Tabela 5 lista todas as possíveis interações identificadas.

Figuras 33 e 34 - Modelo cristalográfico 1U72, os aminoácidos com pontos de interação a menos de 5Å do ligante MTX (verde) estão destacados, assim como o co-fator NADPH (laranja)



Figuras 35 e 36 - Modelo cristalográfico 4M6K, os aminoácidos com pontos de interação a menos de 5Å do ligante FOL (roxo) estão destacados, assim como o co-fator NADPH (laranja)



O programa também foi utilizado para identificação de pontos de interação de hidrogênio, representados nas Figuras 37 e 38, para os dois modelos cristalográficos selecionados.

Figura 37 - Modelo cristalográfico 1U72, as possíveis interações de hidrogênio com o ligante MTX (verde) estão destacadas, assim como o co-fator NAP (laranja)

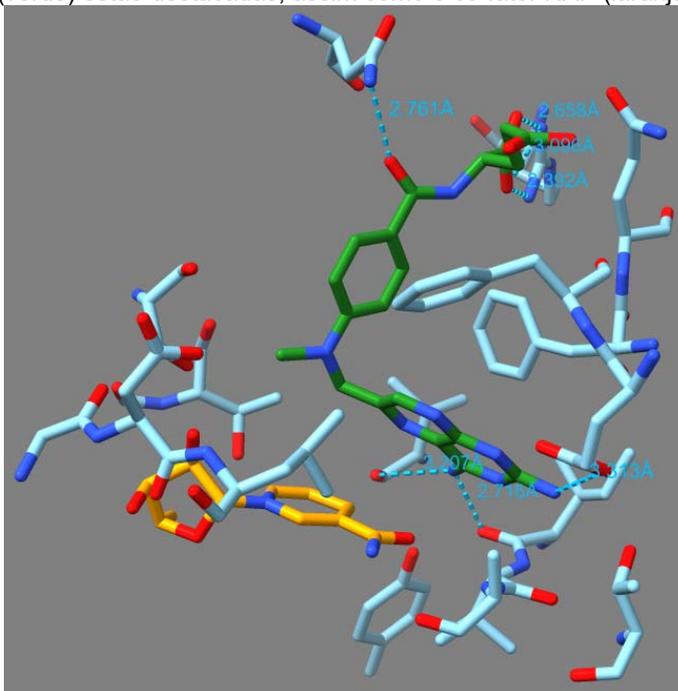


Figura 38 - Modelo cristalográfico 4M6K, as possíveis interações de hidrogênio com o ligante FOL (roxo) estão destacadas, assim como o co-fator NAP (laranja)

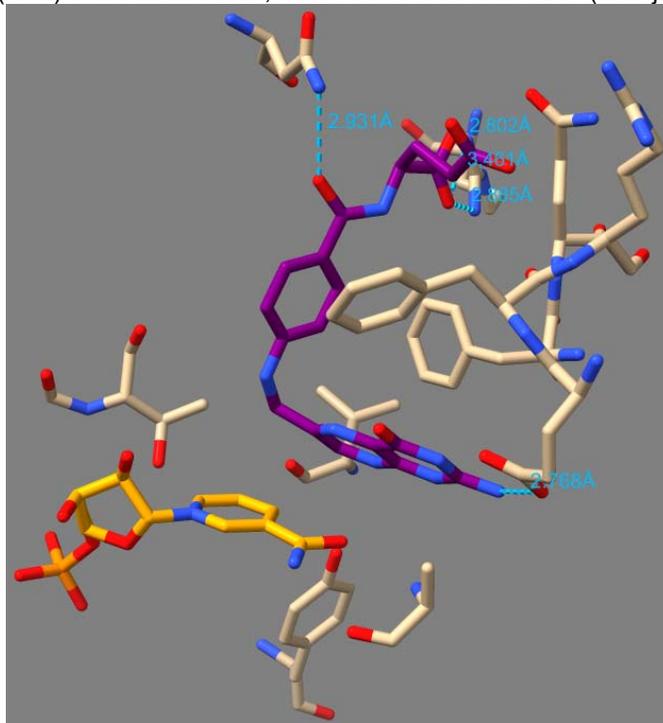


Tabela 05 - Lista de aminoácidos com possíveis interações para cada um dos modelos avaliados.

Aminoácido	Modelo cristalográfico (ligante)	
	1U72 (MTX)	4M6K (FOL)
ILE7		H
VAL8		X
ALA9	X	X
ASP21		X
LEU22		X
GLU30	H	H
PHE31	X	X
ARG32	X	
PHE34	X	X
GLN35	X	X
THR56	X	X
SER59		X
ASN64	H	H
ARG70	H	H
VAL115	X	H
TYR121	X	X
THR136		X

Legenda: os pontos de interação estão destacados com um X, os pontos de interação de hidrogênio destacados com um H.

Assim como ocorreu para os modelos de parasita avaliados, o co-fator NADPH foi identificado em todas as análises realizadas, novamente, não devido a uma possível interação do ligante com o NADPH que contribua para a afinidade do ligante pela proteína, mas devido a proximidade deste ao sítio ativo, menor que os 5 Å determinados como parâmetro de análise.

Além disso, também de forma similar ao observado anteriormente, o folato apresentou mais possíveis pontos de interação com a proteína do que o metotrexato, o que pode ser explicado pela similaridade estrutural e química do

folato com o dihidrofolato, substrato natural da DHFR, onde a diferença entre as duas moléculas é somente uma insaturação e um hidrogênio.

Considerando o foco deste trabalho o planejamento de fármacos voltados para tratamento de doença de Chagas que sejam seletivos, os possíveis pontos de interação da DHRF humana com os ligantes não serão avaliados diretamente, como foi realizado em item anterior para a proteína do parasita, tal avaliação será realizada de forma comparativa com os resultados anteriores.

4.9. COMPARAÇÃO DOS POSSÍVEIS PONTOS DE INTERAÇÃO *T. cruzi* X *H. sapiens*

De forma a facilitar a discussão dos resultados e padronizar as representações utilizadas, a comparação entre os modelos será feita entre o modelo 4M6K da proteína humana com o modelo 5T7O da proteína parasita. Os modelos foram selecionados devido aos seus ligantes, folato e tetrahydrofolato, respectivamente, serem os mais próximos estruturalmente do substrato natural da proteína. Caso algum outro modelo apresente pontos a serem destacados, tais fatores serão descritos.

Para comparar os pontos de interação discutidos anteriormente com os pontos identificados na interação da proteína humana com os dois ligantes analisados de forma mais visual, recorreu-se ao alinhamento das sequências primárias das proteínas. Essa comparação proporciona o entendimento de diferenças a serem exploradas na busca de seletividade dos ligantes para a enzima do parasita em relação à humana (FERREIRA; MORAES; ANDRICOPULO, 2022).

A figura 39 ilustra o alinhamento entre as duas sequências de proteína humana selecionadas e duas sequências da enzima de *T.cruzi*, equivalente ao apresentado anteriormente na figura 32, porém destacando-se os aminoácidos que possuem possíveis interações com o ligante: em azul, os resíduos de aminoácidos que participam de interações com o ligante, estando na mesma posição da sequência primária tanto na proteína do parasita quanto na proteína humana, e em verde aqueles cuja interação é especificamente uma interação de hidrogênio. Por fim, em vermelho estão destacados os aminoácidos que realizam interações com o ligante na proteína do parasita mas não na proteína humana e em amarelo os aminoácidos que realizam interação de hidrogênio em um organismo mas não no outro.

Figura 39 - Alinhamento das sequências primárias de dihidrofolato redutase de *T. cruzi* (3CLB e 5T70)

H. sapiens (4M6K e 1U72)

3CLB Chains	MSLFKIRMPETVAEGRTRLALRAFSLVAVVDERGGIGDGRSIPWNVPE-DMKFFRDVTTKL	59
5T70 Chains	MSLFKIRMPETVAEGRTRLALRAFSLVAVVDEHGGIGDGRSIPWNVPE-DMKFFRDVTTKL	59
4M6K Chain	-----MVGSLNCIVAVSQNMGIGKNGDLPWPPLRNEFRYFQRM TTT-	41
1U72 Chain	-----VGS LNCIVAVSQNMGIGKNGDLPWPPLRNEFRYFQRM TTT-	40
	: : : : * * * : : * * * . . . * * . . : : : * : * * .	
3CLB Chains	RGKNVKPSAKRNAVVMGRKTWDSIPPFRPLPGRNLNVLSSTLTQHLLDGLPDEEKRN	119
5T70 Chains	RGKNVKPSAKRNAVVMGRKTWDSIPPFRPLPGRNLNVLSSTLTQHLLDGLPDEEKRN	119
4M6K Chain	-----SSVEGKQNLVIMGKKTWFSIPEKNRPLKGRINLVLSRELKEPPQ-----	85
1U72 Chain	-----SSVEGKQNLVIMGKKTWFSIPEKNRPLKGRINLVLSRELKEPPQ-----	84
	. . * : * * : * * : *	
3CLB Chains	LHADSIVAVNGGLEQALQLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTT	179
5T70 Chains	LHADSIVAVNGGLEQALRL L ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTT	179
4M6K Chain	----GAHFLSRSLDDALKLTEQPELANKVDMVWIVGGSSVYKEAMNHPGHLKLFV---TR	138
1U72 Chain	----GAHFLSRSLDDALKLTEQPELANKVDMVWIVGGSSVYKEAMNHPGHLKLFV---TR	137
	. . . * : * * : *	
3CLB Chains	IRASESSCSVFFR-----VPE-----SGTEAAAGIEWQRETISEELTSANGNETKY	225
5T70 Chains	IRASESSCSVFFR-----VPE-----SGTEAAAGIEWQRETISEELTSANGNETKY	225
4M6K Chain	IMQDF-ESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKYKFEVYEKND-----	187
1U72 Chain	IMQDF-ESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKYKFEVYEKND-----	186
	* *	
3CLB Chains	YFEKLI 231	
5T70 Chains	YFEKLI 231	
4M6K Chain	----- 187	
1U72 Chain	----- 186	

Legenda: os aminoácidos que realizam interações de hidrogênio em ambos (em verde) ou apenas em um organismo (amarelo) ou que realizam interações em ambos (azul) ou apenas na proteína parasita (vermelho)

Embora a sequência primária seja uma representação linear de uma estrutura tridimensional, é possível observar que a maioria dos aminoácidos responsáveis por interações com o ligante estão na mesma posição das sequências de cada organismo, e mesmo quando o aminoácido em si é diferente, a capacidade de tal interação é mantida, indicando que o sítio ativo é um sítio bem conservado, inclusive entre diferentes espécies.

Considerando o objetivo do presente trabalho, se faz necessário analisar as diferenças entre as proteínas estudadas. Dentre os nove aminoácidos que podem realizar interações com a proteína do parasita mas não na humana, cinco são o mesmo aminoácido em ambas sequências: triptofano na posição 43 (24 da proteína humana), treonina 56 (38), isoleucina 84 (60), prolina 85 (61), leucina 91 (68), ou seja, embora não tenha sido observada a possível interação da proteína humana com o ligante nesses pontos, eles não apresentam diferenças reais que possam ser exploradas no planejamento de fármacos seletivos. A ausência de interações,

apesar da similaridade estrutural, pode ser explicada pela diferença dos ligantes, isto é, a enzima apresentava pontos favoráveis a interação, porém o ligante não apresentava grupos químicos nas posições corretas para promover tal interação, ou ainda, a não observação da interação na análise pode ser resultado do uso de ferramentas rígidas na observação de fenômenos flexíveis.

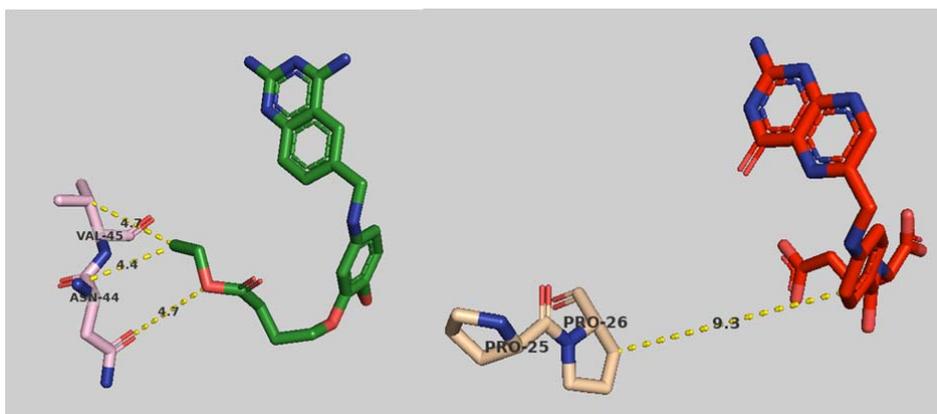
Proteínas são macromoléculas que, embora possuam uma conformação estrutural definida, apresentam grande flexibilidade, principalmente a nível molecular e atômico. A observação de modelos cristalográficos omite a flexibilidade dos complexos proteína-ligante, tornando rígida e estática tal relação. Desta forma é preciso considerar que a posição ocupada por um aminoácido em uma análise pode sofrer variações em ambiente biológico.

Os quatro aminoácidos que apresentam uma diferenciação estrutural são asparagina 44 e valina 45, ambas substituídas por prolina (24 e 25), fenilalanina (51), substituída por tirosina (33) e prolina (92) substituída por lisina (68).

A substituição da asparagina por uma prolina remove os resíduos de aminoácido importantes para a interação com a cadeia lateral do DQ1, destacada na análise do ligante anteriormente, e os substitui por um aminoácido de cadeia curta que não costuma realizar interações com o ligante devido à sua proximidade do *backbone* da proteína e do impedimento estérico gerado pelos resíduos dos aminoácidos vizinhos. Além disso, devido ao tamanho da cadeia lateral, o resíduo fica distante do ligante para uma possível interação, conforme indicado na Figura 41.

Um ligante de cadeia lateral longa, em posição similar ao DQ1, seria capaz de interagir com a asparagina e o triptofano 43 nesta região da enzima parasita, mas não com a prolina da proteína humana. Conforme discutido anteriormente, o triptofano e a asparagina são aminoácidos de características opostas, considerando a substituição da asparagina na DHFR humana, o ideal seria que o ligante possuísse características que favorecessem interações com a asparagina, aumentando a afinidade pela proteína do parasita e desfavorecendo a interação com a proteína humana.

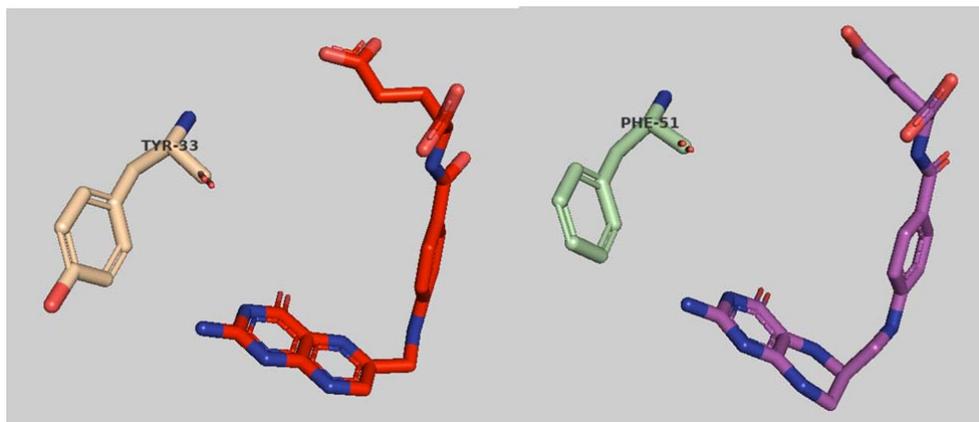
Figuras 40 e 41 - Comparação da alteração da ASN44 e VAL45 por PRO25 e PRO26



Legenda: magenta - modelo 3KJS com ligante DQ1 (verde); bege - modelo 4M6K com ligante FOL (vermelho)

A fenilalanina substituída por uma tirosina substitui o anel benzeno por um fenol, o que aumenta a eletronegatividade do substituinte, favorecendo interações de hidrogênio ou dipolo-dipolo, por exemplo, o que seria favorável para aumentar a afinidade pela proteína humana, o que não é o objetivo deste trabalho. No entanto, um ligante que possuísse um grupo eletronegativo em tal posição poderia causar o efeito contrário, e criar uma repulsão entre ligante e resíduo.

Figuras 42 e 43 - Comparação da alteração da TYR33 por PHE51

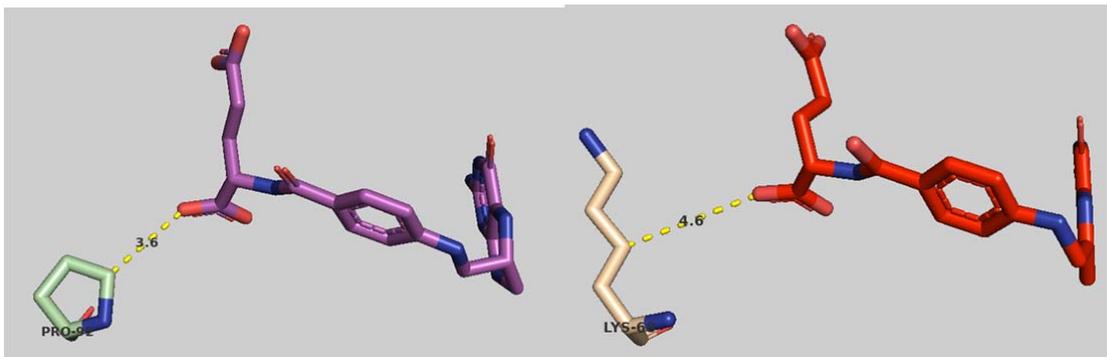


Legenda: verde - modelo 5T7O com ligante THG (roxo); bege - modelo 4M6K com ligante FOL (vermelho)

De forma similar, a substituição da prolina por uma lisina favorece a ocorrência de interações do ligante com a proteína humana, por aumentar a cadeia lateral, conferindo novos átomos de carbono e, adicionalmente, incluir um heteroátomo doador de ligação de hidrogênio. Porém, este ponto de alteração pode

ser explorado de outra forma: ao adicionar um substituinte no ligante que aumente o volume da molécula nessa posição, é possível explorar o volume da lisina, desfavorecendo o acoplamento do ligante na proteína humana.

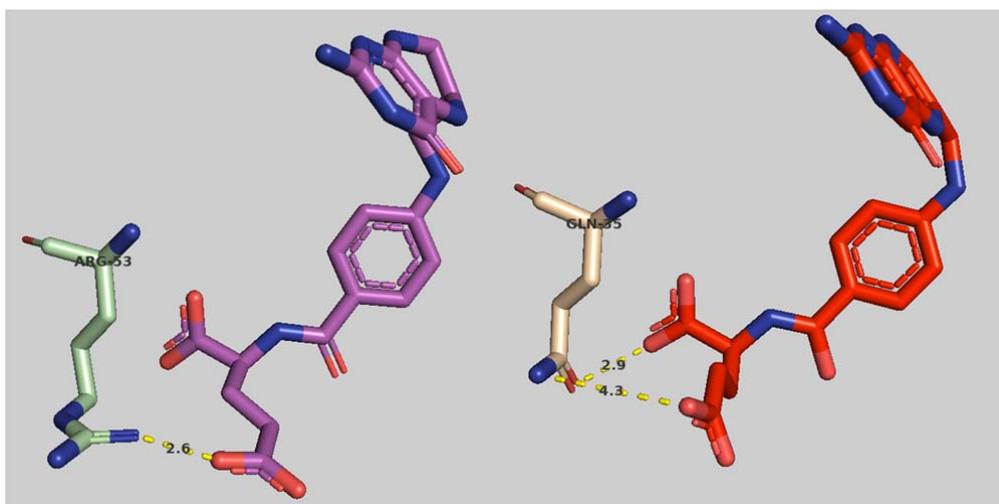
Figuras 44 e 45 - Comparação da alteração da PRO92 por LYS68



Legenda: verde - modelo 5T7O com ligante THG (roxo); bege - modelo 4M6K com ligante FOL (vermelho)

Por fim, resta avaliar os pontos que apresentaram divergências entre os modelos relacionados às interações de hidrogênio. Na posição 53 a proteína de *T. cruzi* possui um resíduo de arginina que é substituído por uma glutamina (posição 35 da proteína humana). A arginina possui nitrogênios doadores de ligação de hidrogênio, enquanto a glutamina possui uma amida, que pode ser tanto doadora (pelo nitrogênio) quanto aceptora de elétrons (pelo oxigênio). Este ponto torna-se portanto, difícil de ser explorado visando a seletividade de um ligante, embora o planejamento de um ligante com átomos aceptores de ligação de hidrogênio na posição seja favorável.

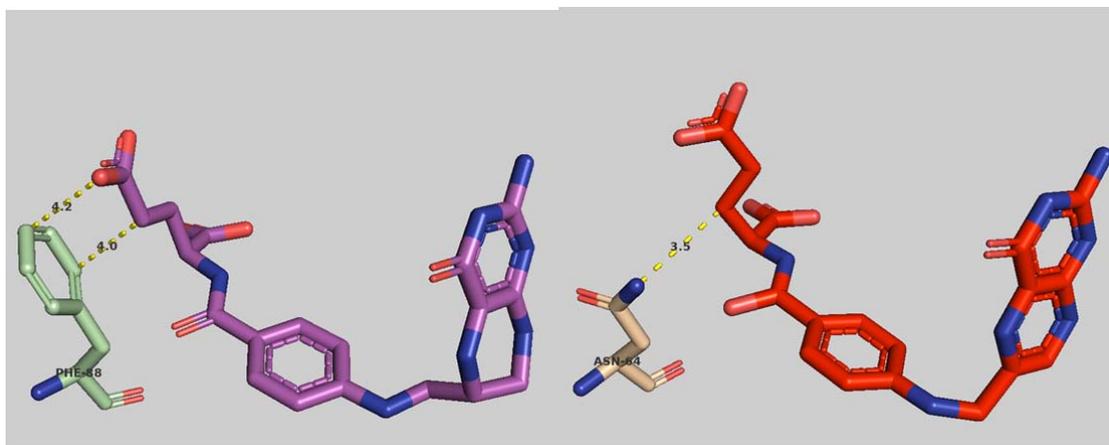
Figuras 46 e 47 - Comparação da alteração da ARG53 por GLN35



Legenda: verde - modelo 5T7O com ligante THG (roxo); bege - modelo 4M6K com ligante FOL (vermelho)

A fenilalanina (88) é substituída por uma asparagina (64), tal substituição resultou em uma interação de hidrogênio na proteína de *H. sapiens*. Por ser um ponto favorável à interação do ligante com a proteína humana, é preciso desfavorecer tal interação, planejando um ligante que não consiga realizar interações de hidrogênio nesse ponto, por exemplo, incluindo doadores de ligação de hidrogênio ou carbonos nesta posição.

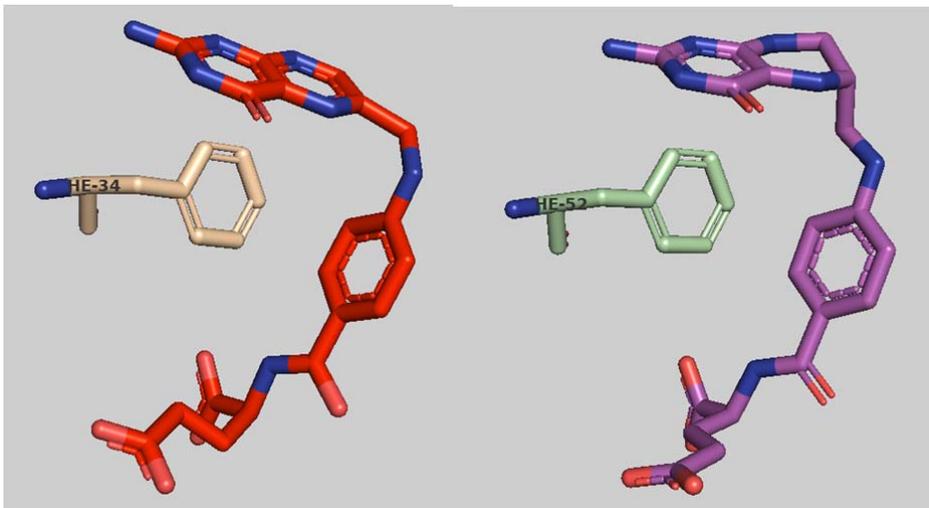
Figuras 48 e 49 - Comparação da alteração da PHE88 por ASN64



Legenda: verde - modelo 5T7O com ligante THG (roxo); bege - modelo 4M6K com ligante FOL (vermelho)

Por fim, o resíduo de fenilalanina 52 é mantido em ambas proteínas (posição 34 da proteína humana), ou seja, o fenômeno de acoplamento discutido para a proteína de *T. cruzi* também é observado no caso de *H. sapiens*.

Figuras 50 e 51 - Comparação da posição da PHE52 com a PHE34



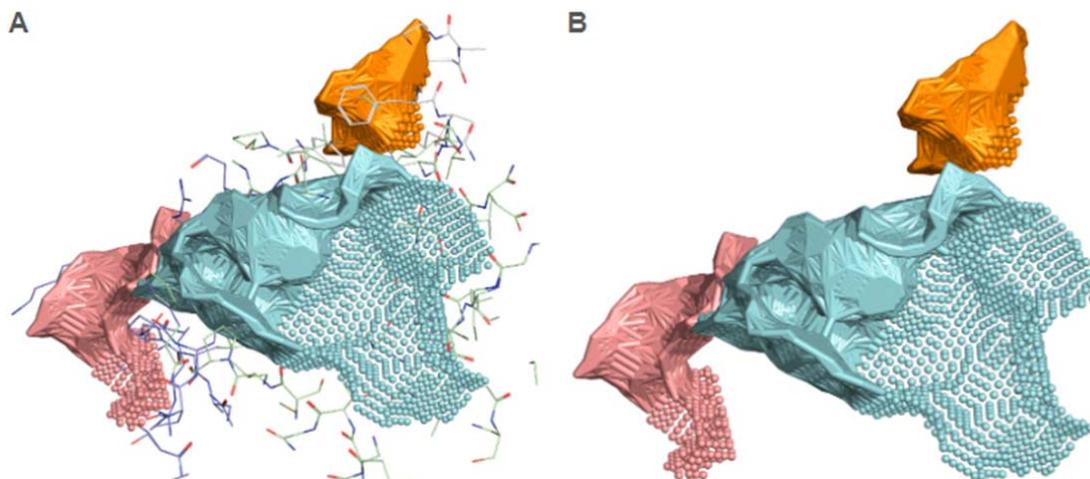
Legenda: verde - modelo 5T7O com ligante THG (roxo); bege - modelo 4M6K com ligante FOL (vermelho)

4.10. ANÁLISE DE CAVIDADES DA DIHIDROFOLATO REDUTASE DE *T. cruzi*

A análise de cavidades da proteína dihidrofolato redutase de *T. cruzi* realizada pela plataforma CavityPlus identificou três cavidades, representadas na Figura 52, com as seguintes características:

Tabela 06 - Características das cavidades da DHFR de *T. cruzi*

Cavidade	pKd máximo previsto	pKd mínimo previsto	DrugScore	Druggability
1	11.36	7.00	2332.00	Forte
2	6.31	4.78	-1256.00	Fraca
3	5.76	4.59	-1469.00	Fraca

Figura 52 - Cavidades da DHFR *T. cruzi* identificadas pelo CavityPlus.

Legenda: A - representação da superfície das cavidades com os aminoácidos que as formam; B - representação da superfície das cavidades. Azul - 1; Laranja - 2; rosa - 3

“*Druggability*” é o termo utilizado pelo programa para representar a capacidade de uma região de permitir o acoplamento de ligantes que modulem a atividade enzimática. Das três regiões identificadas somente a cavidade 1 possui a característica de ser “*druggable*”, bolsão que corresponde, justamente, ao sítio ativo da proteína.

O sítio ativo da dihidrofolato redutase do parasita possui uma superfície de 753,75 Å² e volume de 1180,88 Å³, composta por 60 aminoácidos, listados na tabela abaixo, onde os aminoácidos que haviam sido identificados nas análises dos ligantes foram destacados em negrito. Todos se referem a cadeia A da proteína pois foi a cadeia selecionada para a análise.

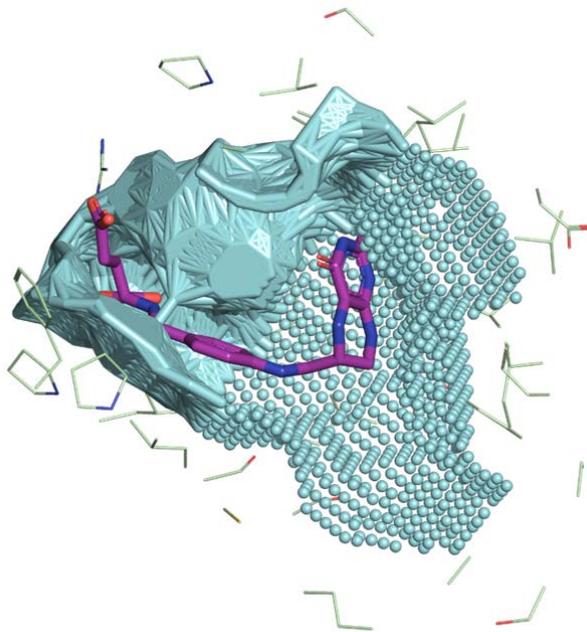
Tabela 07 - Aminoácidos que compõe o sítio ativo da DHFR de *T. cruzi*

Aminoácido
VAL-189-A, VAL-27-A , PHE-88-A , MET-49-A , GLY-93-A, VAL-74-A, VAL-159-A, THR-57-A, THR-178-A , ARG-94-A , LEU-91-A , LYS-79-A, ASP-54-A, VAL-26-A , GLY-33-A, PHE-51-A , ASP-48-A , ALA-28-A , ASP-82-A, GLY-155-A, GLU-31-A, GLY-157-A, ILE-154-A , THR-56-A , PRO-90-A, ILE-180-A, ILE-35-A, TRP-81-A, ALA-182-A, THR-80-A , VAL-29-A, CYS-153-A, SER-158-A, GLY-77-A, SER-83-A , ASN-219-A, ASP-30-A, GLY-34-A, GLY-36-A, LYS-50-A , ARG-181-A, ARG-177-A, VAL-45-A , ARG-78-A, THR-179-A, PRO-46-A, MET-76-A, TYR-160-A , TYR-176-A, GLY-156-A, PRO-85-A , PHE-52-A , LYS-87-A, LEU-55-A, THR-223-A, ARG-53-A , ARG-89-A, ASN-44-A , PRO-92-A, ILE-84-A

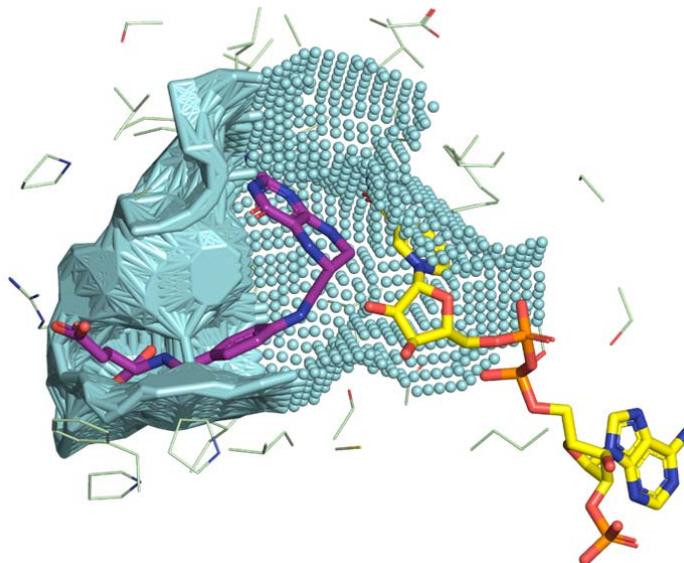
Ao acoplar o ligante no bolsão identificado é possível avaliar o acoplamento entre a proteína e o ligante, conforme ilustrado na Figura 53. A cadeia lateral do anel

benzeno acopla-se na porção mais estreita do bolsão, enquanto o anel bicíclico ocupa a região de maior abertura do bolsão. Embora essa região do bolso seja aberta, até para permitir o acesso do solvente para chegada do substrato, é possível observar que os anéis fundidos ocupam toda a região, o que explica o fato de nenhum ligante avaliado ter substituintes com cadeias maiores de um átomo.

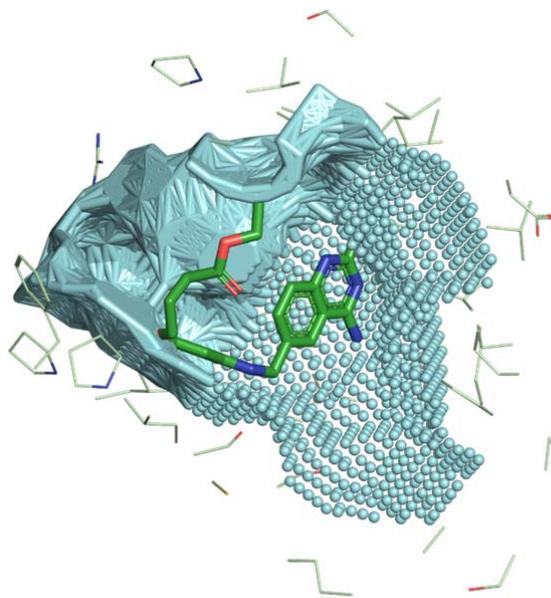
Figura 53 - Cavidades da DHFR *T. cruzi* acoplada ao THG (roxo)



Próximo a esta região ocorre o acoplamento do co-fator NADPH, representado na figura 54, onde é possível observar que somente a nicotinamida ocupa a região do bolsão. A presença do co-fator na mesma região justifica, também, a impossibilidade de grupos volumosos nesta porção do ligante.

Figura 54 - Cavidades da DHFR *T. cruzi* acoplada ao THG (roxo) e ao NADPH (amarelo)

Dentre os ligantes observados anteriormente, o DQ1 apresentou como singularidade possuir uma cadeia substituinte do anel benzênico que apresentasse geometria oposta aos demais ligantes. Ao realizar o acoplamento desta molécula no mesmo bolsão é possível observar que a cadeia linear de carbonos, ao estar na posição meta explora uma região diferente do bolsão, ao invés de se acomodar na mesma posição da cadeia lateral do THG, substituindo a posição para do anel benzênico.

Figura 55 - Cavidades da DHFR *T. cruzi* acoplada ao DQ1 (verde)

Foi também realizada uma busca por sítios alostéricos, pela ferramenta CorrSite, no entanto não foram identificados possíveis sítios alostéricos na proteína.

A análise do CavPharmer identificou que o farmacóforo, representado na Figura 56, para interação no bolsão do sítio ativo, deveria possuir três centros doadores de ligação de hidrogênio (azul) e três aceptores (vermelho e rosa), dois centros eletrostáticos negativos (amarelo) e oito centros hidrofóbicos (azul claro), além de nenhum centro eletrostático positivo, além de ser mantida a estrutura cíclica insaturada. A partir do ciclo, a distância para os centros doadores de interação de hidrogênio é de 6.7 e 6.5 Å, para os aceptores, em orientação oposta, é de 5.8 Å para os dois primeiros e mais 2.5 Å para o terceiro. Portanto, o tamanho das cadeias laterais deve condizer com estas distâncias e, no caso dos centros doadores, a inserção de grupos cíclicos garante que os dois centros sejam explorados, respeitando-se as distâncias, sem que sejam necessários dois substituintes no anel, como observado nos ligantes que apresentam um segundo centro cíclico nesta região, com nitrogênios como heteroátomos e grupos substituintes. Por fim, a angulação entre estes pontos deve ser de 105,3° e 141,1°, condizente com a torção da molécula sobre a fenilalanina 52 discutida anteriormente.

Figura 56 - Farmacóforo acoplado na cavidade da DHFR

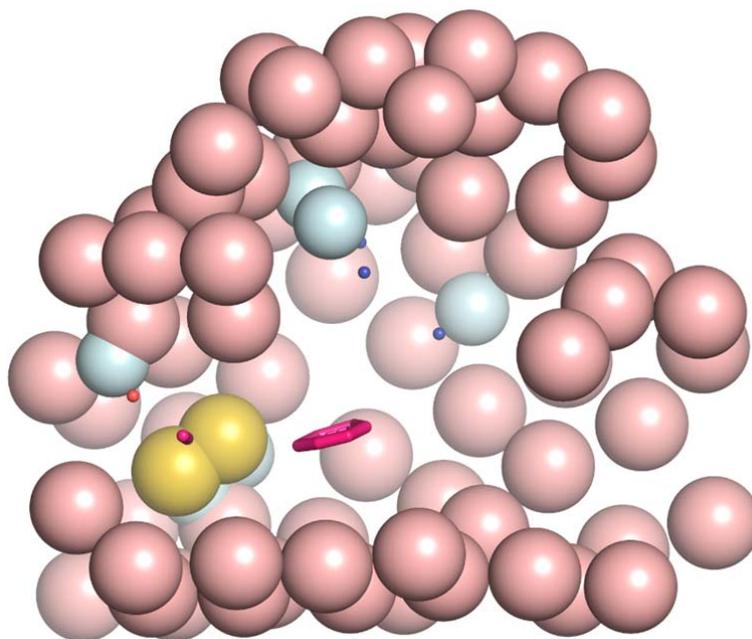


Figura 57 - Representação do farmacóforo identificado

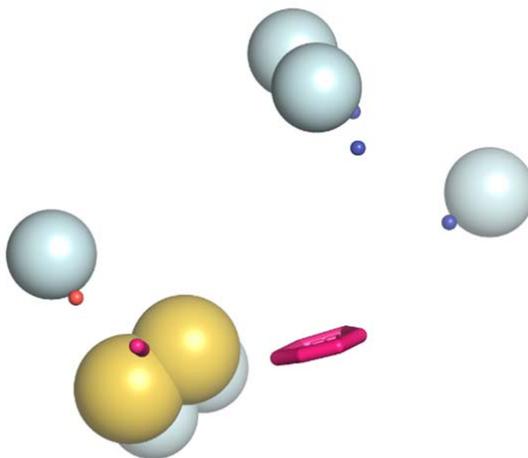
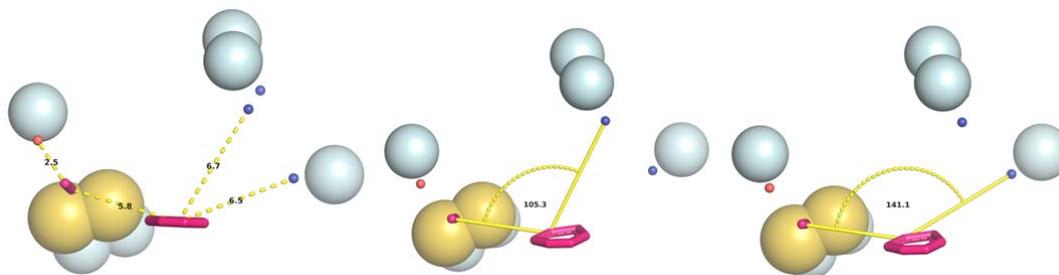


Figura 58 - Representação do farmacóforo identificado e suas medidas de distância e angulação



4.11. ANÁLISE DE CAVIDADES DA DIHIDROFOLATO REDUTASE DE *H. sapiens*

Assim como realizado para a proteína do parasita, executou-se uma avaliação das cavidades da dihidrofolato redutase humana, que identificou também três bolsões, sendo o maior deles correspondente ao sítio ativo da proteína. Assim como observado para o parasita, o maior bolsão, e o único “*druggable*”, é o próprio sítio ativo da proteína.

Tabela 08 - Características das cavidades da DHFR de *H. sapiens*

Cavidade	pKd máximo previsto	pKd mínimo previsto	<i>DrugScore</i>	<i>Druggability</i>
1	11,40	6,53	1572,00	Forte
2	6,41	4,82	-1256,00	Fraca
3	6,33	4,79	-1268,00	Fraca

A área superficial medida para a cavidade do sítio ativo foi de de 1112,00 Å² e um volume de 1447,75 Å³, composta pelos aminoácidos listados a seguir,

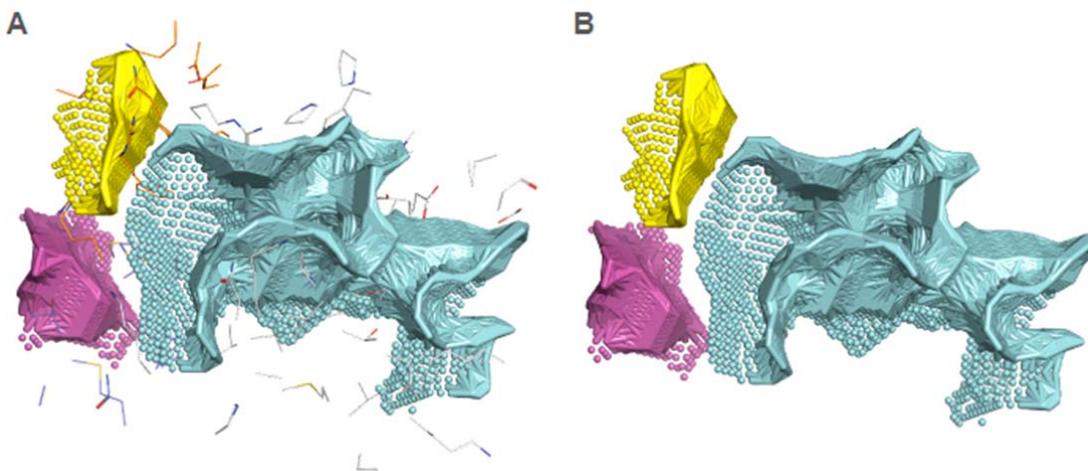
novamente, os aminoácidos que foram analisados nas etapas anteriores deste trabalho estão destacados em negrito.

Tabela 09 - Aminoácidos que compõe o sítio ativo da DHFR de *H. sapiens*

Aminoácido
GLU-62-A, PRO-61-A, VAL-115-A , GLY-17-A, ARG-32-A , PRO-66-A, ARG-36-A, GLY-69-A, VAL-10-A, ARG-70-A , TRP-57-A, ASN-19-A, ASN-29-A, PHE-58-A, THR-146-A, THR-39-A, PRO-23-A, THR-56-A , ASN-64-A , LEU-67-A, ILE-7-A , ASP-145-A, GLY-15-A, ILE-60-A, GLY-53-A, LYS-18-A, PHE-147-A, LYS-63-A, ARG-65-A, PHE-34-A , PRO-25-A, TYR-33-A, GLN-35-A, ASP-21-A, LYS-68-A, TYR-121-A, GLY-20-A, THR-38-A, PHE-134-A, LEU-22-A , TRP-24-A, THR-136-A , GLY-117-A, LYS-54-A, LEU-27-A, LYS-55-A, LEU-75-A, ALA-9-A , ILE-16-A, PRO-26-A, SER-59-A , GLY-116-A, SER-119-A, VAL-8-A , GLU-30-A , SER-118-A, VAL-120-A, ILE-114-A, PHE-31-A , MET-52-A, ARG-28-A.

Dentre as características dos bolsões correspondentes ao sítio ativo, ambos possuem valores de pKd máximo e mínimo similares, além de serem, obviamente, “druggables”. Quanto aos outros dois bolsões, considerados não “druggables”, isto é, não podem contribuir para o acoplamento de moduladores da atividade enzimática, um deles, cavidade 3, encontra-se em posição similar a cavidade 3 da proteína de *T. cruzi*, além de apresentar formato próximo. Já a cavidade 2 encontra-se em posição diametralmente oposta ao observado na proteína do parasita, além de possuir formato único, indicando tratarem-se de bolsões distintos.

Figura 59 - Cavidades da DHFR *H. sapiens* identificadas pelo CavityPlus.

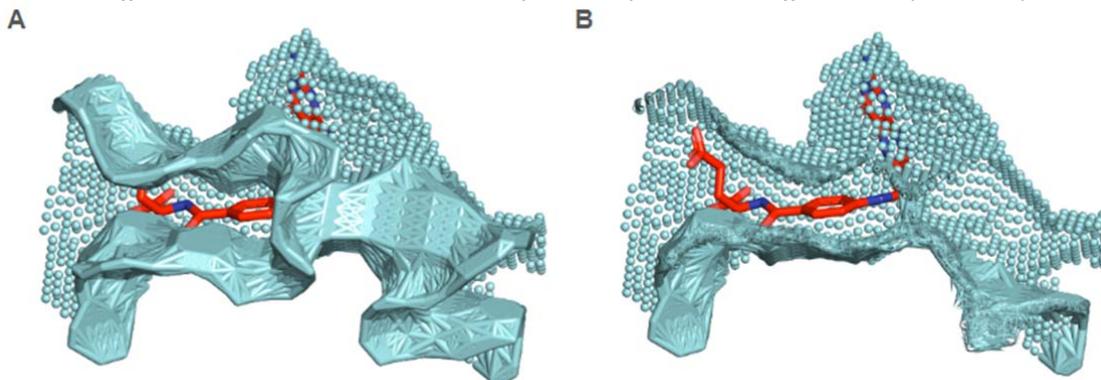


Legenda: A - representação da superfície das cavidades com os aminoácidos que as formam; B - representação da superfície das cavidades. Azul - 1; Amarelo - 2; Rosa - 3.

A cavidade correspondente ao sítio ativo da proteína humana é maior que o sítio ativo da proteína do parasita, tanto em área quanto em superfície. Na Figura 60 é possível observar que o bolsão da proteína humana apresenta maior comprimento,

enquanto é mais estreita nos pontos de acoplamento. Para facilitar a visualização, na Figura A a cavidade é mostrada completa e na Figura B parte da superfície do bolsão foi retirada para que fosse possível visualizar o posicionamento do ligante folato.

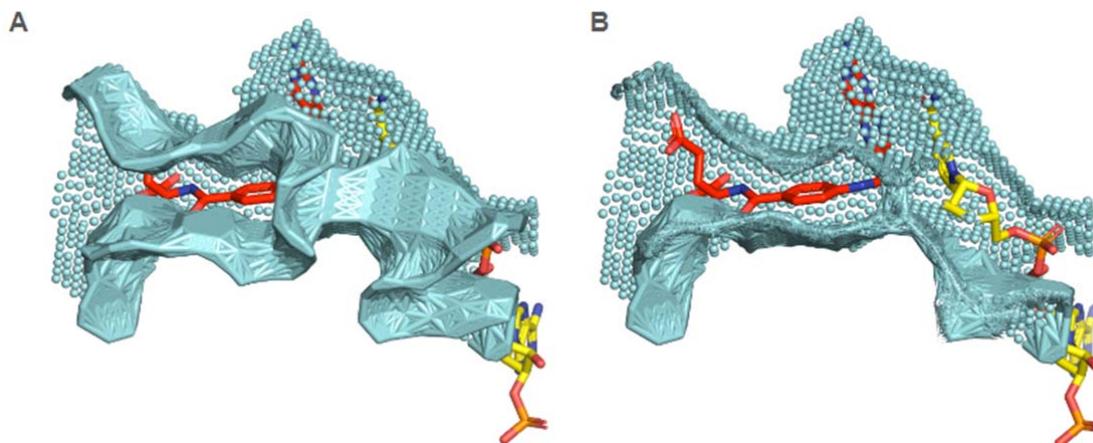
Figura 60 - Cavidades da DHFR *H. sapiens* acoplada com o ligante FOL (vermelho)



A região ocupada pelo substituinte do grupo benzeno é menos estreita na proteína humana do que na proteína do parasita, no entanto, como a cavidade humana forma um túnel semi-aberto, cadeias na posição meta, como no caso do ligante C50, provavelmente sofreriam um impedimento estérico que dificultaria seu encaixe no sítio ativo da proteína.

Na Figura 61 acoplou-se, além do folato, o NADPH no sítio ativo, novamente, na Figura 61 B parte da superfície do bolsão foi retirada para permitir a visualização das moléculas.

É possível observar que a região ocupada pelo biciclo é mais fechada na proteína humana, em comparação com a proteína de *T. cruzi*, além disso, o NADPH é melhor englobado pelo bolsão na proteína humana do que observado anteriormente, em que somente o grupo adenosina acoplava-se à região. Possivelmente ao substituir o heterociclo por grupos pouco maiores, ou substituir os anéis por anéis com mais membros, pode contribuir para que ocorra um impedimento estérico no acoplamento do ligante à proteína humana, desfavorecendo a afinidade da molécula pela proteína.

Figura 61 - Cavidades da DHFR *H. sapiens* acoplada com FOL (vermelho) e NADPH (amarelo)

Realizou-se também a identificação de sítios alostéricos com a plataforma CorrSite, no entanto, novamente, não foram identificados sítios alostéricos.

Por fim, realizou-se a determinação do farmacóforo com o CavPharmer, como resultado, identificou-se que o farmacóforo, representado nas Figuras 62 e 63, para interação no bolsão do sítio ativo, deveria possuir dois centros doadores de interação de hidrogênio (azul), três centros aceptores (vermelho), seis centros hidrofóbicos (azul claro) e nenhum centro eletrostático, tanto positivo quanto negativo. O farmacóforo também possui um centro cíclico insaturado, correspondente ao anel benzênico, e as distâncias dos centros doadores e aceptores estão representadas na Figura 64, sendo de 8,6 Å para o centro doador e 7,2 Å para o centro aceptor e a angulação entre os centros e o benzênio é de 101,6°.

Figura 62 - Farmacóforo acoplado na cavidade da DHFR

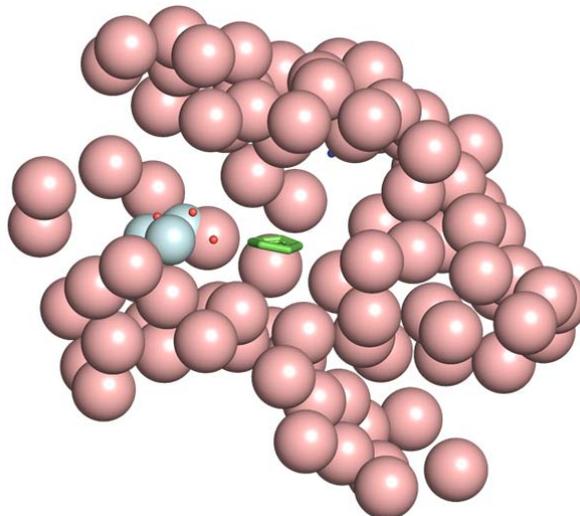


Figura 63 - Representação do farmacóforo identificado

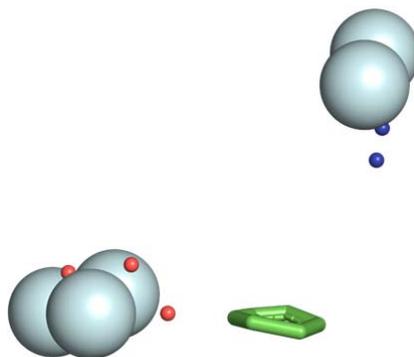
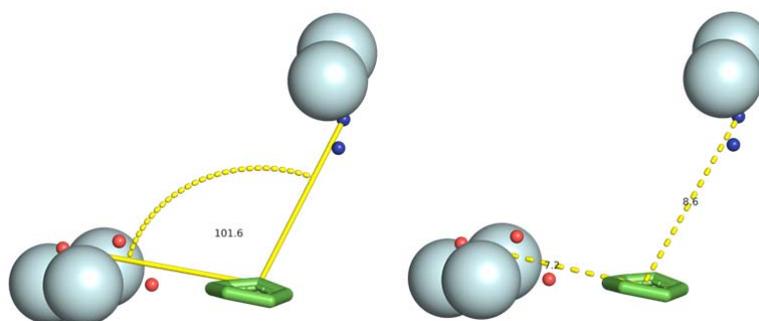


Figura 64 - Representação do farmacóforo identificado e suas medidas de distância e angulação



A distância do anel para os centros doadores de ligação de hidrogênio na proteína humana é maior que a distância na proteína parasita, de 6,7-6,5 Å para 8,6 Å, permitindo que os substituintes do benzeno tenham cadeias menores de forma a interagir com os centros da proteína parasita mas não sejam suficientes para interagir com a proteína humana.

Quanto à distância para os centros doadores de interação de hidrogênio, ambas são próximas, 7,2 Å para a humana e 5,8 + 2,5 Å para a DHFR parasita. A presença de um centro de interação mais próximo ao benzeno, a 2,5 Å, torna-se um ponto interessante a ser explorado, uma vez que não existe correspondente no modelo humano. Além disso, considerando o observado pela molécula C50, que explora uma região diferente do bolsão, e onde existe um impedimento estérico na proteína humana, indica ser possível incluir dois substituintes no anel benzeno, um para substituinte a a 2,5 Å para interação com o centro doador de interação de hidrogênio, e outro na posição meta, para exploração desta outra região da cavidade.

A angulação dos dois centros com o anel é de $101,6^\circ$ para a proteína humana e $105,3^\circ$ a $141,1^\circ$ para o parasita. A maior variabilidade de angulação do parasita se deve pela distância entre os dois centros doadores, que é maior entre eles, enquanto os centros doadores na cavidade da proteína humana estão mais próximos entre si. Além disso, esses valores são resultado da abertura do bolsão da proteína parasita, que é maior em relação à proteína humana. Moléculas com ângulos mais abertos podem ser interessantes para promover maior afinidade pela proteína do parasita, em relação à proteína humana.

5. CONCLUSÃO

Como resultado deste trabalho identificou-se fatores que são comuns aos ligantes de dihidrofolato redutase e, portanto, devem estar presentes em qualquer ligante planejado para este alvo: a molécula precisa possuir dois grupos cíclicos, o primeiro deles um heterocíclo insaturado de seis átomos, com presença de átomos de nitrogênio, podendo ou não estar fundido a outro anel de seis átomos. Assim, os grupos substituintes, principalmente aminas, aumentam os pontos de interação deste anel e, obviamente, é essencial que ele não possua a amida participante da reação de oxido redução da enzima.

O segundo anel, um benzeno, deve ser conectado ao grupo heterocíclico por quatro átomos, permitindo assim que a molécula atinja a conformação côncava necessária para acoplar-se ao sítio ativo e interagir com a fenilalanina. A presença de grupos substituintes neste anel, principalmente cadeias longas de seis carbonos na posição para, o que aumenta os pontos de interação com a proteína.

As diferenças entre as duas proteínas são muito sutis no sítio ativo, indicando que é um sítio bem conservado, mesmo considerando espécies bastantes distantes evolutivamente, entretanto, as poucas diferenças observadas podem ser exploradas no planejamento de inibidores seletivos.

O anel benzeno também pode receber substituintes por um grupo volumosos, na posição meta, de forma a explorar a prolina 92, e os centros doadores de elétrons, estes centros devem ser menores, principalmente se o ligante apresentar substituinte na posição meta, para assegurar a interação com o centro próximo, a 2,5 Å. e de forma a dificultar a interação com o centro doador da proteína humana, a 7,2 Å.

O centro aceptor de ligação de hidrogênio possui pouca variabilidade, principalmente por ser onde efetivamente ocorre a catálise, no entanto a proteína do parasita apresenta os dois centros em posições mais próximas ao benzeno, a 6,5 e a 6,7 Å do anel, enquanto os centros aceptores da proteína humana estão a 8,6 Å, ou seja, substituintes de cadeias menores podem explorar essa região de interação na proteína parasita e não na proteína humana.

Essa mesma região possui uma certa abertura na enzima de *T. cruzi*, enquanto é fechada na correspondente humana, o que torna necessário a

investigação de anéis e seus conectores com grupos substituintes maiores, para que seja observado se é possível explorar essa abertura da proteína do parasita de forma a garantir seletividade.

Por fim, a asparagina 44 substituída por uma prolina na proteína humana, permite interações com o grupo substituinte do anel benzeno, na posição meta. Além disso, substituintes que explorem essa região apresentam uma conformação diferenciada, que ocupam um espaço diferente do espaço ocupado pelo substrato na proteína, além de ser uma região que é estericamente impedida no caso da proteína humana, podendo ser um ponto importante para a afinidade do ligante.

Embora sejam poucos os pontos destacados, ao tratar-se de uma proteína altamente preservada entre espécies, pequenas mudanças podem refletir de forma significativa na interação alvo-ligante, ainda mais tratando-se de uma interação de hidrogênio, com forte energia de interação, e um impedimento estérico, que desfavorece o acoplamento da molécula ao sítio ativo.

Para validar tais informações, faz-se necessário estudos mais aprofundados como, por exemplo, o desenho de tais ligantes, ou o filtro de bancos de moléculas por aquelas que apresentem as características destacadas, e a simulação do acoplamento destas moléculas com as duas proteínas, tanto de *T. cruzi* quanto de *H. sapiens*, para medir a energia do sistema acoplado e inferir se há, realmente, uma diferença significativa na afinidade das moléculas pelo sítio ativo de cada proteína.

REFERÊNCIAS

- ALVES, V. et al. QUIMIOINFORMÁTICA: UMA INTRODUÇÃO. **Química Nova**, 2017.
- ANDERSON, A. Targeting DHFR in parasitic protozoa. **Drug Discovery Today**, v. 10, n. 2, p. 121–128, 15 jan. 2005.
- BAKER, M. Cryo-electron microscopy shapes up. **Nature**, v. 561, n. 7724, p. 565–567, 25 set. 2018.
- BATOOL, M.; AHMAD, B.; CHOI, S. A Structure-Based Drug Discovery Paradigm. **International Journal of Molecular Sciences**, v. 20, n. 11, p. 2783, 6 jun. 2019.
- BAX, A.; CLORE, G. M. Protein NMR: Boundless opportunities. **Journal of Magnetic Resonance**, v. 306, p. 187–191, set. 2019.
- BELTRAN-HORTELANO, I. et al. Examination of multiple Trypanosoma cruzi targets in a new drug discovery approach for Chagas disease. **Bioorganic & Medicinal Chemistry**, v. 58, p. 116577, mar. 2022.
- BERMAN, H. M. et al. The Protein Data Bank. **Acta Crystallographica Section D Biological Crystallography**, v. 58, n. 6, p. 899–907, 1 jun. 2002.
- BERNSTEIN, F. C. et al. The protein data bank: A computer-based archival file for macromolecular structures. **Journal of Molecular Biology**, v. 112, n. 3, p. 535–542, maio 1977.
- CHEN, Y.; KIRCHMAIR, J. Cheminformatics in Natural Product-based Drug Discovery. **Molecular Informatics**, v. 39, n. 12, 6 dez. 2020.
- CLARK, D. P.; PAZDERNIK, N. J.; MCGEHEE, M. R. Cell Division and DNA Replication. In: **Molecular Biology**. [s.l.] Elsevier, 2019. p. 296–331.
- EVANS, P. R.; MURSHUDOV, G. N. How good are my data and what is the resolution? **Acta Crystallographica Section D Biological Crystallography**, v. 69, n. 7, p. 1204–1214, 1 jul. 2013.
- FERREIRA, L. L. G.; ANDRICOPULO, A. D. Chemoinformatics Strategies for Leishmaniasis Drug Discovery. **Frontiers in Pharmacology**, v. 9, 1 nov. 2018.
- FERREIRA, L. L. G.; DE MORAES, J.; ANDRICOPULO, A. D. Approaches to advance drug discovery for neglected tropical diseases. **Drug Discovery Today**, v. 27, n. 8, p. 2278–2287, ago. 2022.

GODDARD, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. **Protein Science**, v. 27, n. 1, p. 14–25, 6 jan. 2018.

K. GOPALAKRISHNAN et al. Ramachandran Plot on The Web (2.0). **Protein & Peptide Letters**, v. 14, n. 7, p. 669–671, 1 jul. 2007.

GUIDO, R. V. C.; ANDRICOPULO, A. D.; OLIVA, G. Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. **Estudos Avançados**, v. 24, n. 70, p. 81–98, 2010.

HASSAN BAIG, M. et al. Computer Aided Drug Design: Success and Limitations. **Current Pharmaceutical Design**, v. 22, n. 5, p. 572–581, 26 jan. 2016.

HUBBARD, R. E.; KAMRAN HAIDER, M. Hydrogen Bonds in Proteins: Role and Strength. In: **Encyclopedia of Life Sciences**. [s.l.] Wiley, 2010.

IRVINE, M. W. Agents acting on pyrimidine metabolism. In: **Antimalarial Agents**. [s.l.] Elsevier, 2020. p. 133–185.

KLEYWEGT, G. J.; JONES, T. A., [11] Model building and refinement practice. In: [s.l.: s.n.]. p. 208–230.

KOHLHOFF, K. J. et al. Fast and Accurate Predictions of Protein NMR Chemical Shifts from Interatomic Distances. **Journal of the American Chemical Society**, v. 131, n. 39, p. 13894–13895, 7 out. 2009.

MACHADO, C. A.; AYALA, F. J. Sequence variation in the dihydrofolate reductase-thymidylate synthase (DHFR-TS) and trypanothione reductase (TR) genes of *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 121, n. 1, p. 33–47, abr. 2002.

MARTINEZ-MAYORGA, K. et al. The impact of chemoinformatics on drug discovery in the pharmaceutical industry. **Expert Opinion on Drug Discovery**, v. 15, n. 3, p. 293–306, 3 mar. 2020.

MORETTI, N. S.; MORTARA, R. A.; SCHENKMAN, S. *Trypanosoma cruzi*. **Trends in Parasitology**, v. 36, n. 4, p. 404–405, abr. 2020.

MÜLLER KRATZ, J. et al. Clinical and pharmacological profile of benznidazole for treatment of Chagas disease. **Expert Review of Clinical Pharmacology**, v. 11, n. 10, p. 943–957, 3 out. 2018.

PARKER, M. W. Protein Structure from X-Ray Diffraction. **Journal of Biological Physics**, v. 29, n. 4, p. 341–362, 2003.

PATTERSON, S.; WYLLIE, S. Nitro drugs for the treatment of trypanosomatid diseases: past, present, and future prospects. **Trends in Parasitology**, v. 30, n. 6, p. 289–298, jun. 2014.

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, n. 1, p. 95–99, jul. 1963.

RASSI, A.; RASSI, A.; MARIN-NETO, J. A. Chagas disease. **The Lancet**, v. 375, n. 9723, p. 1388–1402, abr. 2010.

RCSB PDB. **About RCSB PDB: Enabling Breakthroughs in Scientific and Biomedical Research and Education.** Disponível em: <<https://www.rcsb.org/pages/about-us/index>>.

REDDISH, M. J. et al. Ligand-Dependent Conformational Dynamics of Dihydrofolate Reductase. **Biochemistry**, v. 55, n. 10, p. 1485–1493, 15 mar. 2016.

RUEDA, K. et al. Transmissão oral de *Trypanosoma cruzi*: un nuevo escenario epidemiológico de la enfermedad de Chagas en Colombia y otros países suramericanos. **Biomédica**, v. 34, n. 4, 9 jun. 2014.

SALES JUNIOR, P. A. et al. Experimental and Clinical Treatment of Chagas Disease: A Review. **The American Journal of Tropical Medicine and Hygiene**, v. 97, n. 5, p. 1289–1303, 8 nov. 2017.

SCHNELL, J. R.; DYSON, H. J.; WRIGHT, P. E. Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase. **Annual Review of Biophysics and Biomolecular Structure**, v. 33, n. 1, p. 119–140, 9 jun. 2004.

SHARMA, M.; CHAUHAN, P. M. Dihydrofolate reductase as a therapeutic target for infectious diseases: opportunities and challenges. **Future Medicinal Chemistry**, v. 4, n. 10, p. 1335–1365, jun. 2012.

SHEN, P. S. The 2017 Nobel Prize in Chemistry: cryo-EM comes of age. **Analytical and Bioanalytical Chemistry**, v. 410, n. 8, p. 2053–2057, 9 mar. 2018.

SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, v. 7, n. 1, 11 jan. 2011.

SCHRÖDINGER, L.; DELANO, W. **PyMol**, 2020. Disponível em <<http://www.pymol.org/pymol>>, último acesso em dez. 2023.

STRASSER, B. J. Collecting, Comparing, and Computing Sequences: The Making of Margaret O. Dayhoff's Atlas of Protein Sequence and Structure, 1954–1965. **Journal of the History of Biology**, v. 43, n. 4, p. 623–660, 24 dez. 2010.

SUZUKI, A. F. et al. Oral Infection and Survival of *Trypanosoma cruzi* in Sugarcane Juice Conditioned at Different Temperatures. **Acta Parasitologica**, 22 nov. 2023.

XU, Y. et al. CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. **Nucleic Acids Research**, v. 46, n. W1, p. W374–W379, 2 jul. 2018.

ZHAO, F. et al. A guide to sample delivery systems for serial crystallography. **The FEBS Journal**, v. 286, n. 22, p. 4402–4417, 6 nov. 2019.

ZUMA, A. A.; DOS SANTOS BARRIAS, E.; DE SOUZA, W. Basic Biology of *Trypanosoma cruzi*. **Current Pharmaceutical Design**, v. 27, n. 14, p. 1671–1732, abr. 2021.

APÊNDICES

APÊNDICE A - Comparação das sequências de aminoácidos dos modelos cristalográficos de *T. cruzi*. As posições em que há diferenças estão indicadas por “:”

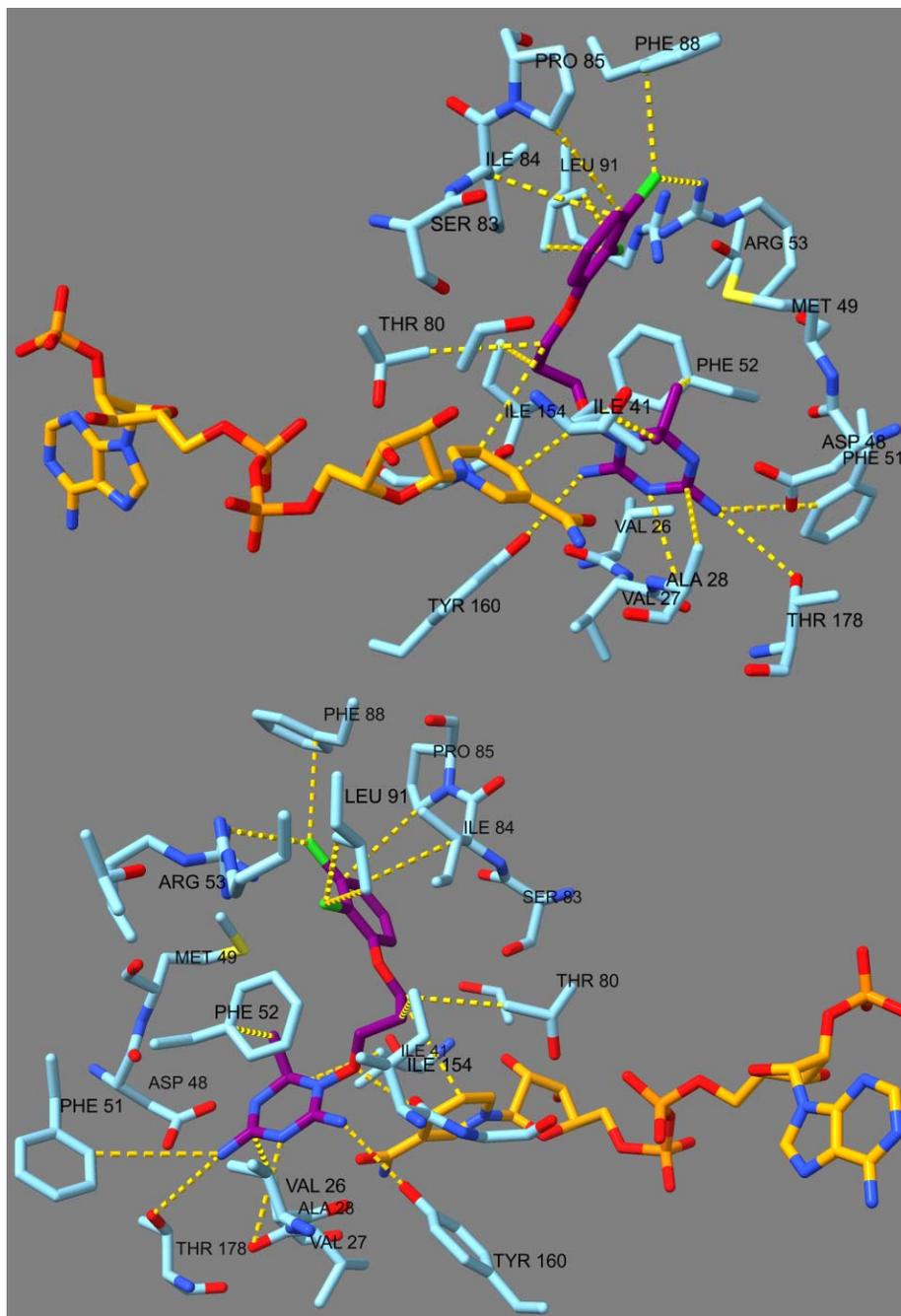
3HBB	MSLFKIRMPETVAEGRALRALRAFSLVAVDERGGIGDGRSIPWNPEDMKFFRDVTTKLR	60
3CLB	MSLFKIRMPETVAEGRALRALRAFSLVAVDERGGIGDGRSIPWNPEDMKFFRDVTTKLR	60
3CL9	MSLFKIRMPETVAEGRALRALRAFSLVAVDERGGIGDGRSIPWNPEDMKFFRDVTTKLR	60
3KJS	MSLFKIRMPETVAEGRALRALRAFSLVAVDERGGIGDGRSIPWNPEDMKFFRDVTTKLR	60
2H2Q	MSLFKIRMPETVAEGRALRALRAFSLVAVDERGGIGDGRSIPWNPEDMKFFRDVTTKLR	60
3IRN	MSLFKIRMPETVAEGRALRALRAFSLVAVDEHGGIGDGRSIPWNPEDMKFFRDLTTKLR	60
3IRM	MSLFKIRMPETVAEGRALRALRAFSLVAVDEHGGIGDGRSIPWNPEDMKFFRDLTTKLR	60
3IRO	MSLFKIRMPETVAEGRALRALRAFSLVAVDEHGGIGDGRSIPWNPEDMKFFRDLTTKLR	60
3INV	MSLFKIRMPETVAEGRALRALRAFSLVAVDEHGGIGDGRSIPWNPEDMKFFRDLTTKLR	60
5T70	MSLFKIRMPETVAEGRALRALRAFSLVAVDEHGGIGDGRSIPWNPEDMKFFRDLTTKLR	60
	*****.*****.*****	
3HBB	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3CLB	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3CL9	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3KJS	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
2H2Q	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3IRN	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3IRM	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3IRO	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
3INV	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
5T70	GKNVKPSAKRNAVVMGRKTWDSIPPKFRPLPGRLNVLSSTLTQHLLDGLPDEEKRL	120
	*****.*****.*****	
3HBB	HADSI VAVNGGLEQALQLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3CLB	HADSI VAVNGGLEQALQLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3CL9	HADSI VAVNGGLEQALQLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3KJS	HADSI VAVNGGLEQALQLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
2H2Q	HADSI VAVNGGLEQALQLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3IRN	HADSI VAVNGGLEQALRLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3IRM	HADSI VAVNGGLEQALRLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3IRO	HADSI VAVNGGLEQALRLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
3INV	HADSI VAVNGGLEQALRLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
5T70	HADSI VAVNGGLEQALRLL ASPNYTPSIETVYCI GGGSVYAEALRPPCVHLLQAIYRTTI	180
	*****.*****.*****	

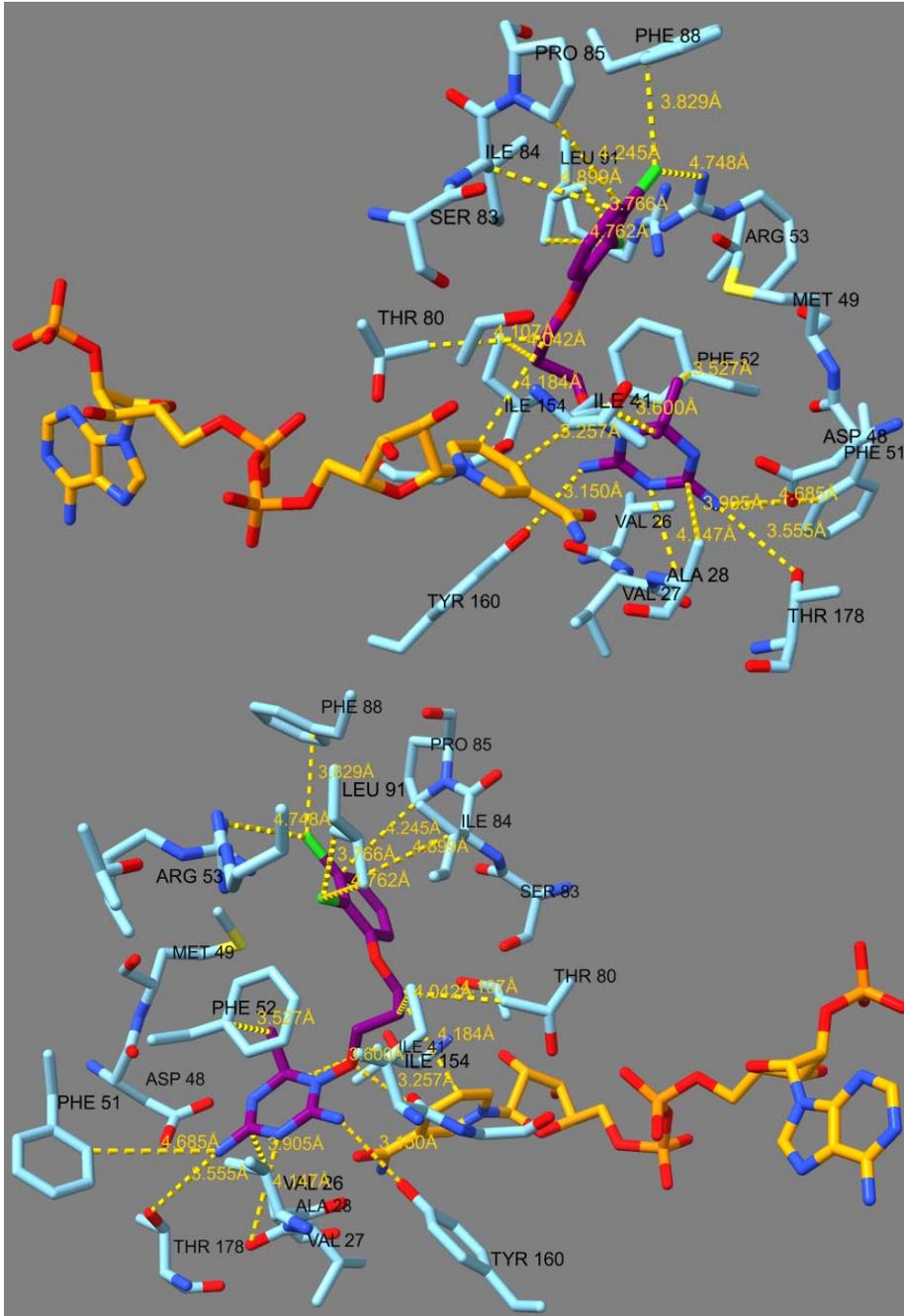
3HBB	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3CLB	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3CL9	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3KJS	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
2H2Q	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3IRN	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3IRM	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3IRO	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
3INV	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240
5T70	RASESSCSVFFRVPESGTEAAAGIEWQRETISEELTSANGNETKYYFEKLIPRNREEEQY	240

3HBB	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3CLB	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3CL9	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3KJS	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
2H2Q	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3IRN	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3IRM	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3IRO	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
3INV	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300
5T70	LSLVDRIIREGNVKHDRTGVGTLSIFGAQMRFSLRNNRPLLLTTKRVFWRGVCEELLWFL	300

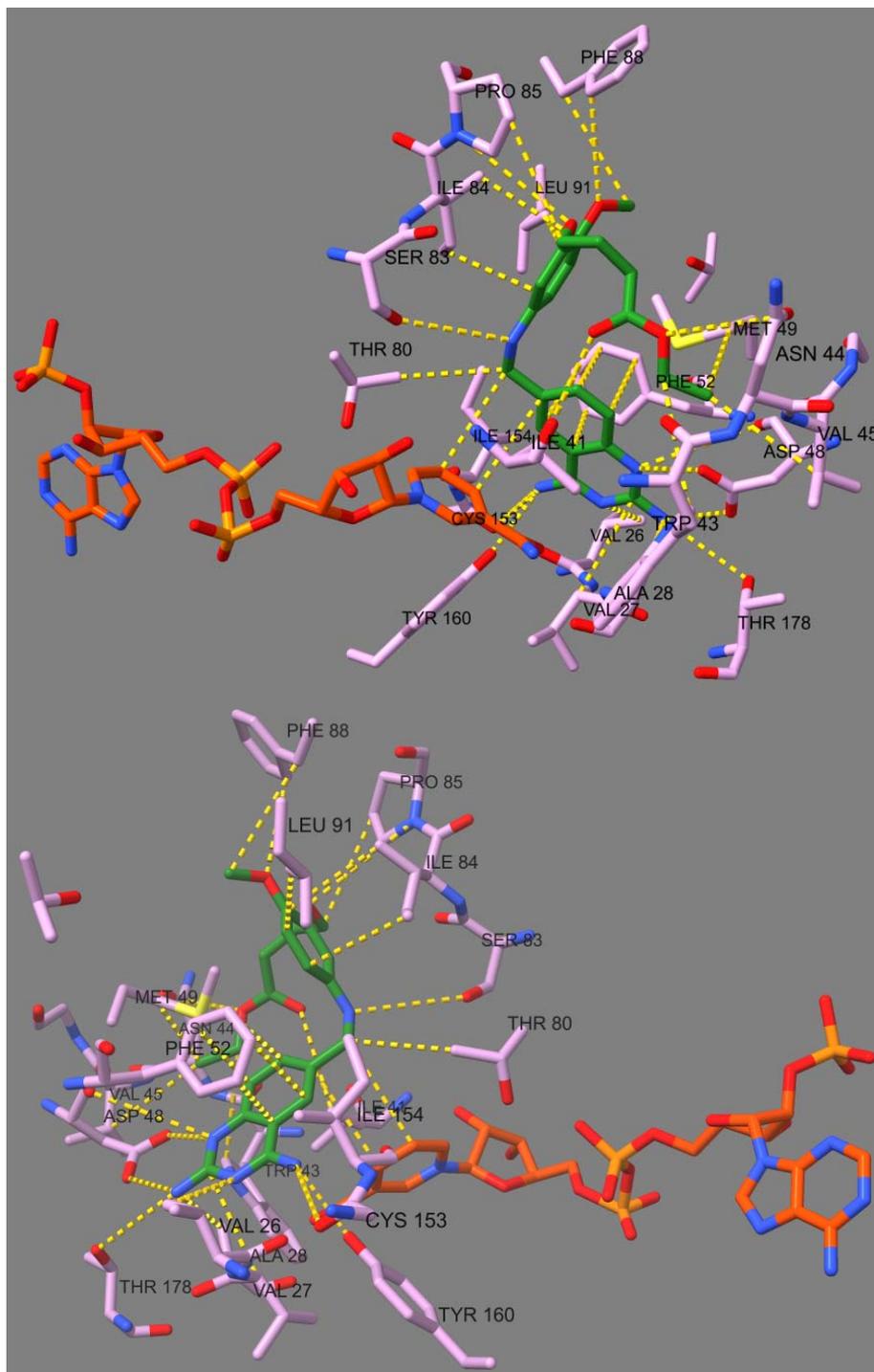
3HBB	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3CLB	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3CL9	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3KJS	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
2H2Q	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3IRN	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3IRM	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3IRO	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
3INV	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360
5T70	RGETYAKKLSDKGVHIWDDNGSRAF LDSRGL TEYEEMDLGPVYGFQWRHFGAAYTHHDAN	360

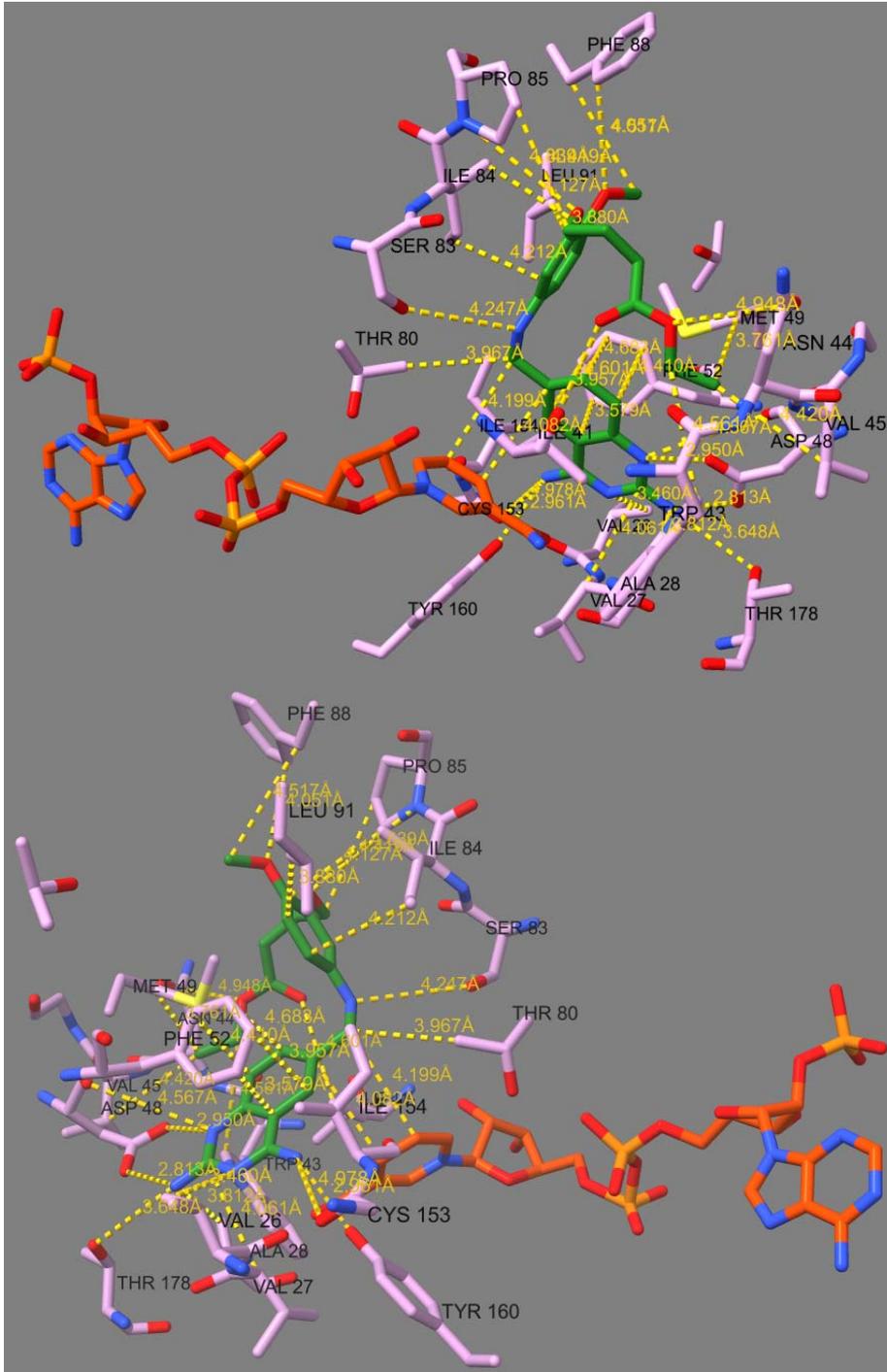
APÊNDICE B - Modelo cristalográfico 3INV (C50) de *T. cruzi* com os pontos de interação com o ligante destacados pela linha tracejada amarela e a distância entre os átomos.



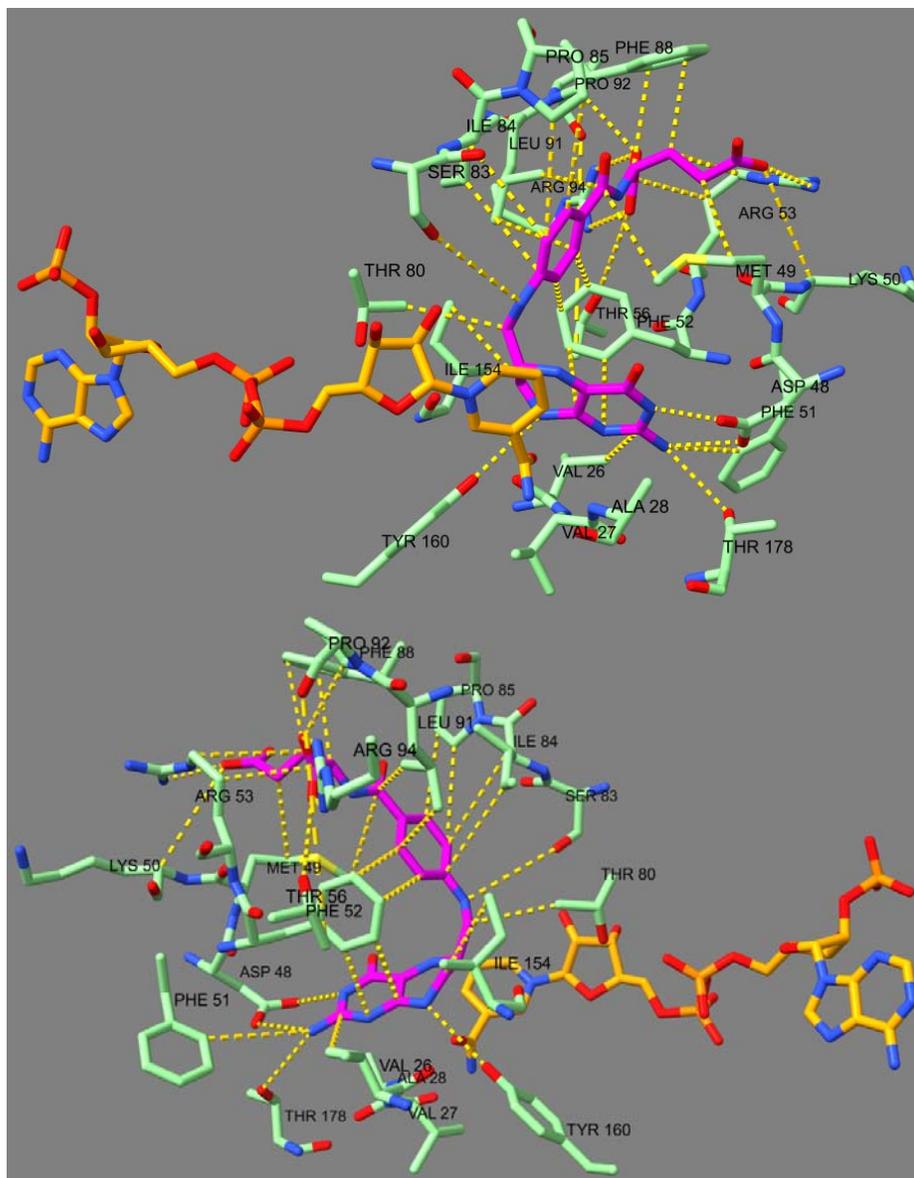


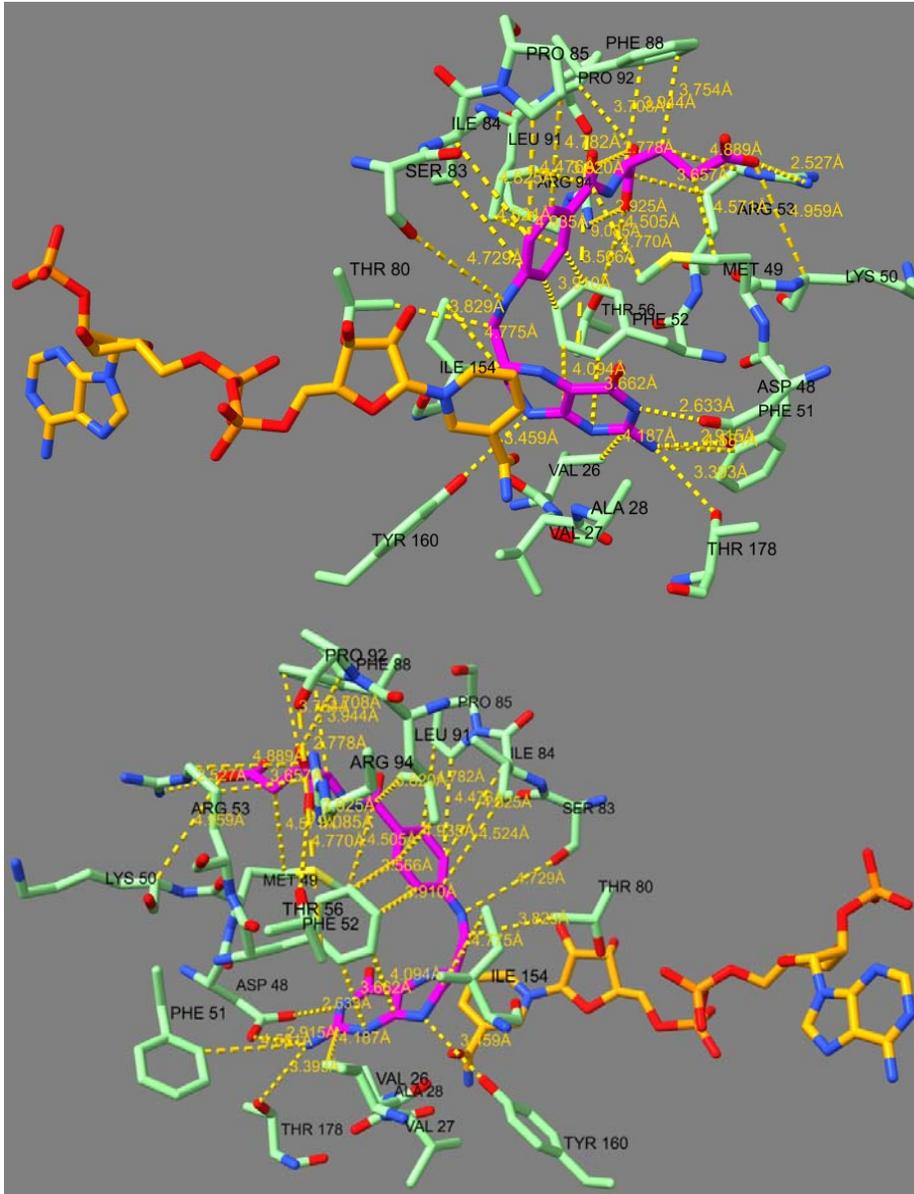
APÊNDICE C - Modelo cristalográfico 3KJS (DQ1) de *T. cruzi* com os pontos de interação com o ligante destacados pela linha tracejada amarela e a distância entre os átomos.



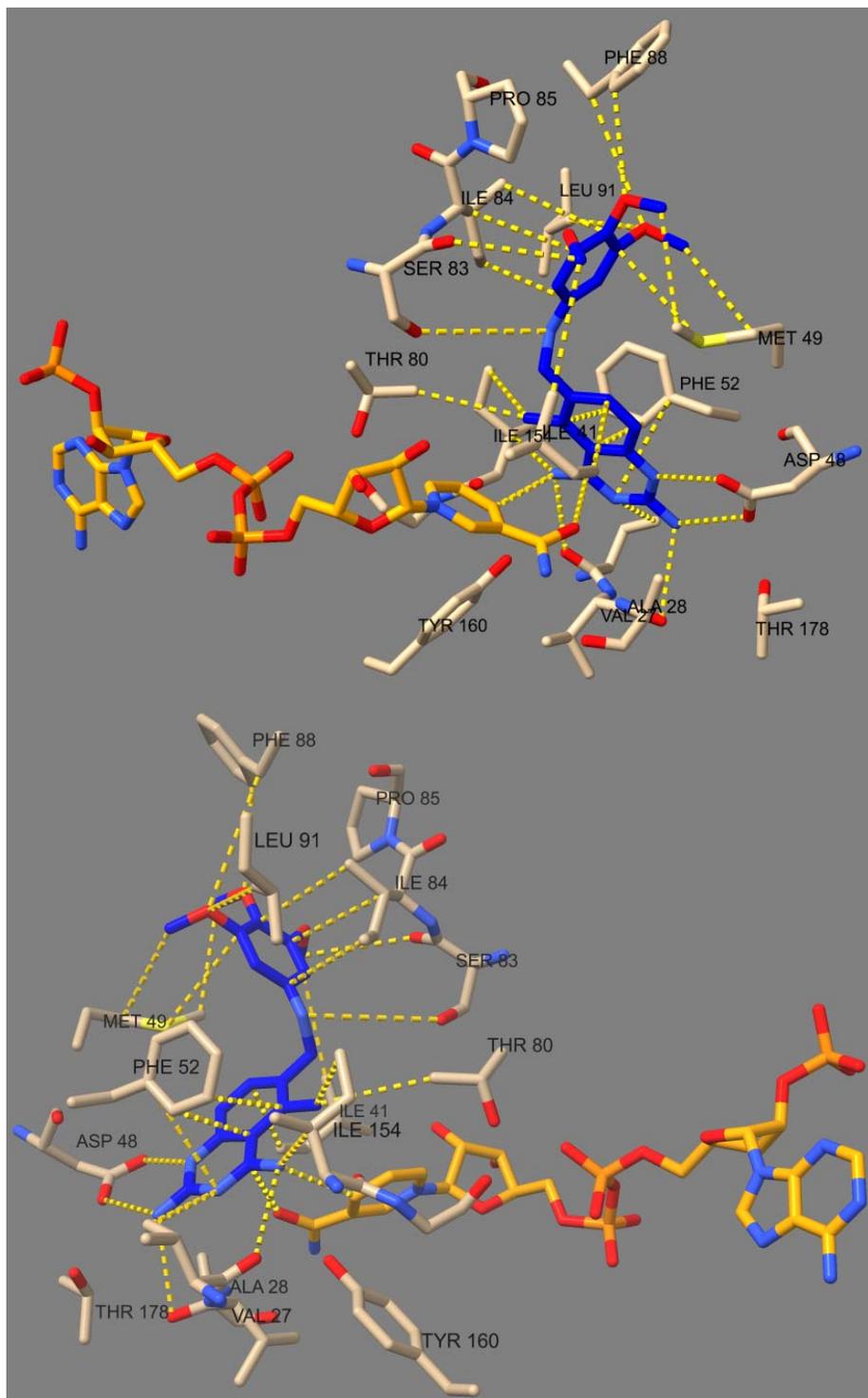


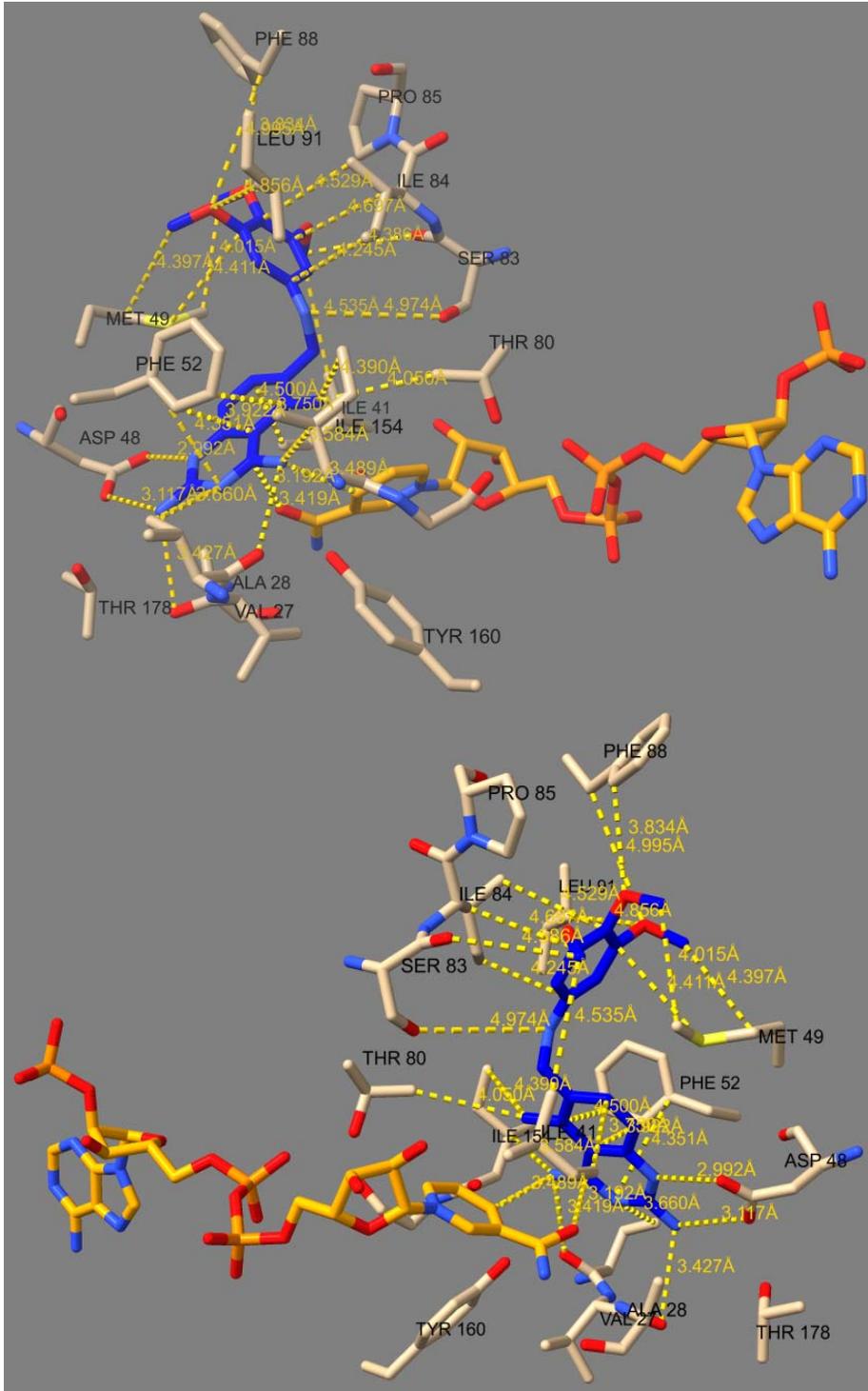
APÊNDICE D - Modelo cristalográfico 5T7O (THG) de *T. cruzi* com os pontos de interação com o ligante destacados pela linha tracejada amarela e a distância entre os átomos.





APÊNDICE E - Modelo cristalográfico 3CLB (TMQ) de *T. cruzi* com os pontos de interação com o ligante destacados pela linha tracejada amarela e a distância entre os átomos.





APÊNDICE F - Modelo cristalográfico 1U72 (MTX) de *H. sapiens* com os pontos de interação com o ligante destacados pela linha tracejada amarela e a distância entre os átomos.

