

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

Rafael Pereira Lemos

**ENGENHARIA DE PROTEÍNAS COM BASE EM EVOLUÇÃO:
Interconversão Funcional Entre
5-Hidroxiosurato Hidrolases e Transtirretinas**

BELO HORIZONTE

2022

Rafael Pereira Lemos

**ENGENHARIA DE PROTEÍNAS COM BASE EM EVOLUÇÃO:
Interconversão Funcional Entre
5-Hidroxiosurato Hidrolases e Transtirretinas**

Versão final

Dissertação apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Mestre em Bioinformática.

Orientadora: Profa. Dra. Mariana T. Quezado de Magalhães

Co-Orientador: Prof. Dr. Lucas Bleicher

BELO HORIZONTE

2022

043

Lemos, Rafael Pereira.

Engenharia de proteínas com base em evolução: interconversão funcional entre 5-hidroxiiosurato hidrolases e transtirretinas [manuscrito] / Rafael Pereira Lemos. – 2022.

101 f. : il. ; 29,5 cm.

Orientadora: Profa. Dra. Mariana T. Quezado de Magalhães. Co-Orientador: Prof. Dr. Lucas Bleicher.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. Pré-Albumina. 3. Engenharia de Proteínas. 4. Modelagem computacional. I. Magalhães, Mariana Torquato Quezado de. II. Bleicher, Lucas. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 573:004



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

FOLHA DE APROVAÇÃO

Rafael Pereira Lemos

"ENGENHARIA DE PROTEÍNAS COM BASE EM EVOLUÇÃO: Interconversão Funcional Entre 5-Hidroxiiosurato Hidrolases e Transtirretinas"

Dissertação aprovada pela banca examinadora constituída pelos Professores: Profa. Mariana Torquato Quezado de Magalhães - Orientadora (Universidade Federal de Minas Gerais), Prof. Lucas Bleicher - Coorientador (Universidade Federal de Minas Gerais), Prof. Antonio José da Costa Filho (Universidade de São Paulo) e Prof. Luiz Eduardo Vieira Del Bem (Universidade Federal de Minas Gerais).

Belo Horizonte, 08 de dezembro de 2022.



Documento assinado eletronicamente por **Aristoteles Goes Neto, Coordenador(a) de curso de pós-graduação**, em 16/02/2023, às 13:32, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2090395** e o código CRC **15805AAB**.

AGRADECIMENTOS

Os dois últimos anos foram completamente atípicos. Encontrava-me em uma área de estudo e ambiente de trabalho novos, e devido à pandemia, passei mais da metade da duração do meu mestrado sem contato presencial com meus colegas de laboratório e orientadores. Os experimentos e análises não davam certo, a saúde mental se deteriorava, mas no final me encontro aqui, com a certeza que fiz o que foi possível e com o apoio das melhores pessoas. E é assim que começo meus agradecimentos.

Aos meus pais, Rogério e Gláucia, e meus irmãos, Gabriel e Raquel, por terem-me (re) acolhido em casa, e por terem-me feito companhia em um período tão difícil.

Ao restante dos meus familiares, pelas conversas e apoios, de todas as formas e em todos os momentos.

À Camila, por ter sido meu porto seguro no último ano, por ter ficado ao meu lado enquanto eu estava completamente frito pela dissertação e planos do futuro, e por lembrar-se de mim na aula de bioquímica de proteínas e na pizzaria, embora não estivesse em nenhum dos dois lugares. Mesmo me abandonando em Paris, te amo!

Ao Flávio, Gabriel e Júlia, diretamente envolvidos na parte experimental deste trabalho, pela ajuda na bancada e análises, e pelas fofocas do dia a dia.

À Ana, por ser minha consultora particular de química, pelas ajudas na elaboração de algumas figuras e por ter sofrido junto comigo nessa reta final. “No fim, vai dar tudo errado!”.

Ao Bruno Santos, pelas discussões proveitosas e análise (extremamente) crítica do manuscrito, e pelas demais ajudas ao longo desses anos. Juízo!

Aos outros colegas do laboratório, por terem ajudado de várias formas, ou apenas aturando meu humor insuportável e piadas infames. Como é dito por aí, “dou dinheiro, mas não dou intimidade”.

Aos meus amigos de Belo Horizonte, especialmente ZZ, Bruba e Julha, por terem mantido contato mesmo quando me encontrava distante, por aguentarem meus eternos furos nos rolês, e pelas noites de filmes de terror e conversas até de madrugada sobre os temas mais aleatórios possíveis.

Aos meus amigos de Divinópolis (e atualmente arredores), pois mesmo com nossos apertados horários “de adulto”, sempre se mantiveram presentes e mantendo o companheirismo que já dura quase duas décadas.

Ao Professor Ronaldo Nagem e seu aluno Arthur Lobato, por terem cedido de forma solícita equipamentos, ferramentas e conhecimento inestimáveis para a realização deste trabalho.

Ao Professor Miguel Ortega, Bruno Silva e Paulo Henrique, pelas ajudas com o uso de servidores, dinâmicas moleculares e demais “paranauês” computacionais.

Ao técnico Jamil Silvano, pelos auxílios na resolução de diversos “pepinos”, e pela atenção e carinho que sempre demonstrou com todos.

Ao programa de pós-graduação em Bioinformática, em especial aos secretários Tiago Silva e Sheila Santana, por terem sido extremamente prestativos e atenciosos com minhas dúvidas e problemas, desde o período da minha formação complementar durante a graduação, até o fim dessa jornada no mestrado.

Às agências de fomento, CAPES, CNPq e FAPEMIG, por fornecerem o auxílio financeiro para o desenvolvimento deste projeto.

Ao meu co-orientador, Lucas Bleicher, por ter me apoiado desde a graduação como meu tutor em formação complementar, e agora diretamente com meu trabalho. Obrigado pela ideia original do projeto, pelas discussões que tivemos, e pelas noites de jogos de tabuleiro sempre bastante diferentes e divertidas.

À minha orientadora, Mariana Quezado, por mesmo de brincadeira, ter insistido para me ter como aluno, e por sempre ter me aceitado com um sorriso no rosto quando eu aparecia matriculado em mais uma disciplina ofertada pelo trio das proteínas. A convivência às vezes um pouco conflituosa é apenas um reflexo de duas pessoas com gênios fortes, uma paixão pelo conhecimento, e que mandam mensagens “tarde da noite”, às 21h.

O que aprendi nos últimos anos é imensurável, e deixo aqui meu muito obrigado a todos.

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

Isaac Asimov

i. RESUMO

Transtirretinas (TTR) são uma família de proteínas que possuem como função principal a ligação e a distribuição de hormônios tireoidianos (T3 e T4). Apesar de serem exclusivas de vertebrados, a maior parte dos processos evolutivos destas proteínas ocorreu antes do surgimento dos mesmos, originando-se por meio da duplicação de genes envolvidos na via catabólica de purinas. Estes genes codificam a enzima 5-hidroxiisourato hidrolase (HIUase), que atua na conversão de 5-hidroxiisourato (HIU) a 2-oxo-4-hidroxi-carboxi-5-ureidoimidazolina (OHCU). Diferentemente das TTR, a subfamília de HIUase é ubíqua em ambos procariotos e eucariotos, com os hominídeos sendo uma importante exceção. *In vivo*, ambas as subfamílias se encontram em uma forma homotetramérica, com cada monômero sendo composto por oito folhas-beta conectadas por sete voltas, e apenas uma alfa-hélice. Os tetrâmeros são estabilizados por interações hidrofóbicas entre cada par de dímeros, levando à formação de uma cavidade interna carregada e com poder catalítico. Considerando a intrincada história evolutiva dessas proteínas, além de suas similaridades estruturais primárias e quaternárias, nós hipotetizamos que substituições em regiões e em resíduos específicos seriam capazes de interconverter suas funções. Nós engenheiramos duas sequências de proteínas hipotéticas, onde uma sequência de aminoácidos representativa de uma subfamília foi substituída em posições específicas conservadas localmente e correlacionadas entre si pela outra subfamília, e vice-versa. Aplicando modelagem molecular, nós almejamos entender melhor o processo de neofuncionalização destes parálogos. Utilizando os programas *Modeller* e *AlphaFold*, geramos modelos estruturais tridimensionais por homologia, enquanto empregávamos um alinhamento quimérico manual, entre as proteínas de referência e as engenheiradas, para a otimização dos resultados. Os melhores modelos foram então refinados, validados estatisticamente e estruturalmente, e as suas cavidades foram analisadas em relação às suas geometrias, volumes e potenciais eletrostáticos, utilizando os programas CAVER e APBS. Nossos resultados demonstraram que, como esperado, os volumes e geometrias diferem entre si, devido às inerentes diferenças entre os tamanhos e propriedades físico-químicas de seus ligantes. As mudanças observadas em relação à polaridade e cadeias laterais de grandes resíduos de aminoácidos também estão de acordo com a conservação estrutural das subfamílias.

Palavras-chave: Transtirretina; Modelagem computacional; Engenharia de proteínas.

ii. ABSTRACT

Transthyretins are a protein family that exhibits binding and distribution of thyroid hormones (T3 and T4) as primary functions. Despite being exclusive to vertebrates, a major part of the evolutionary process of these proteins occurred before the emergence of such organisms, originating via duplication of purine catabolic pathway genes. Those encode the 5-hydroxyisourate hydrolase (HIUase) enzyme, which acts in the conversion of 5-hydroxyisourate (HIU) to 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU). Differently from transthyretins, the HIUase subfamily is ubiquitous in both prokaryotes and eukaryotes, with hominids being an important exception. *In vivo*, both subfamilies are found in a homotetrameric form, with each monomer being formed by eight beta-strands connected by seven loops, and one alpha-helix. Tetramers are made stable by hydrophobic interactions between each dimer pair, leading to the formation of an internal charged catalytic cavity. Considering these proteins' intricate evolutionary history, as well as their high primary and quaternary structural similarity, we hypothesized that specific *in silico* substitutions would be able to switch their functions. We then engineered two putative protein sequences, where one subfamily representative sequence was substituted in specific and correlated locally conserved positions by the other, and vice-versa. Applying computational modeling, we aimed to better understand the neofunctionalization process of these paralogues. Using *Modeller* and *AlphaFold*, we generated 3D homology structural models, while also employing a chimeric manual alignment, between the reference and engineered proteins, to further improve the results. The best models were refined, validated, and then their cavities geometries, volumes and electrostatic potentials were analyzed using CAVER and APBS software suites. Our results indicate that, as expected, the volumes and geometries differ from one another, due to size and physicochemical differences between their ligands. The observed changes in polarity and large amino acid residue side chains fit the subfamilies' structural conservation.

Keywords: Transthyretin; Molecular modeling; Protein engineering.

iii. LISTA DE FIGURAS

Figura 1 - Padrão de migração eletroforético das principais proteínas plasmáticas.....	16
Figura 2 - Estruturas dos hormônios tireoidianos.....	17
Figura 3 - Estrutura tridimensional de uma TTR.	18
Figura 4 - Detalhe do bolsão hidrofóbico da interface forte.....	19
Figura 5 - Diagrama esquemático da ligação do T4 na estrutura da TTR humana.	21
Figura 6 - Mecanismo proposto de agregação da transtirretina.....	22
Figura 7 - Complexo Transtirretina-RBP-Retinol.	24
Figura 8 - Sítios de ligação à zinco na TTR.	25
Figura 9 - Representação esquemática da via de degradação do ácido úrico.....	27
Figura 10 - Sobreposição das estruturas tridimensionais de TTR e HIUase.	29
Figura 11 - Comparação entre as estruturas dos ligantes de HIUase.	30
Figura 12 - Modo de ligação da molécula de AZX na cavidade catalítica de uma HIUase.....	31
Figura 13 - Mecanismo proposto para a reação enzimática da HIUase.	32
Figura 14 - Estratégia experimental.	43
Figura 15 - Árvore filogenética circular da prevalência das subfamílias de Transtirretinas e 5-Hidroxiourato Hidrolases em organismos vivos.	44
Figura 16 - Árvore filogenética retangular da prevalência de HIUase em primatas (<i>primates</i>).	46
Figura 17 - Alinhamento de estruturas primárias entre HIUase de <i>Herbaspirillum seropedicae</i> e <i>Danio rerio</i> (<i>Zebrafish</i>).	47
Figura 18 - Alinhamento de estruturas primárias entre HIUase de <i>Herbaspirillum seropedicae</i> e TTR de <i>Homo sapiens</i>	47
Figura 19 - Estruturas secundárias de HIUase e TTR.	48
Figura 20 - Padrão de organização das estruturas secundárias.....	49
Figura 21 - Sobreposição das estruturas quaternárias de HIUase e TTR.	50
Figura 22 - Comparação entre as geometrias das cavidades de HIUase e TTR.....	51
Figura 23 - Comparação entre estruturas resolvidas experimentalmente na presença de ligantes para HIUase e TTR..	52
Figura 24 - Comparação entre características químicas de resíduos localizados no centro das cavidades de HIUase e TTR.	53
Figura 25 - Comparação entre características químicas de resíduos localizados nas extremidades das cavidades de HIUase e TTR.	55

Figura 26 - Potenciais eletrostáticos de HIUase e Transtirretinas.....	56
Figura 27 - Comparação entre os mutantes e suas respectivas proteínas de origem.	58
Figura 28 - Gráficos de ERRAT e Ramachandran das estruturas tridimensionais geradas <i>in silico</i>	61
Figura 29 - Cálculo da variação de RMSD entre os C α de cada resíduo das proteínas de referência e mutantes.....	65
Figura 30 - Geometria das cavidades das proteínas mutantes de TTR.....	70
Figura 31 - Geometria das cavidades das proteínas mutantes de HIUase.....	71
Figura 32 - Potenciais eletrostáticos em HbsHIU e todos os modelos de MutTTR.....	74
Figura 33 - Potenciais eletrostáticos em HsTTR e todos os modelos de MutHIU.....	75
Figura 34 - Mapeamento de resíduos expostos nas estruturas tridimensionais das proteínas de referência e mutantes.	77
Figura 35 - Gráfico de barras das áreas de interface entre os dímeros AB-CD das proteínas de referência e mutantes.	79

iv. LISTA DE QUADROS E TABELAS

Quadro 1 - Classificação dos resíduos de aminoácidos de acordo com suas características químicas.....	38
Tabela 2 - Comparação entre os volumes das cavidades de HIUase e TTR.	51
Tabela 3 - Parâmetros de validação e análises estatísticas dos modelos gerados <i>in silico</i>	60
Tabela 4 - Comparação entre os volumes das cavidades de HbsHIU e Mutantes de TTR.	68
Tabela 5 - Comparação entre os volumes das cavidades de HsTTR e Mutantes de HIUase. ..	70
Tabela 6 - Comparação entre o número de resíduos expostos entre as diferentes proteínas de referência e mutantes.	78

v. LISTA DE ABREVIATURAS

5-HIU = 5-Hidroxiisourato

Å = Angstrom

APBS = *Advanced Poisson-Boltzmann Solver*

apoA-I = Apolipoproteína A-I

ASA = Accessible Surface Area

AZX = 8-Azaxantina

BLAST = *Basic Local Alignment Search Tool*

BLOSUM = *BLOcks of aminoacid SUBstitution Matrix*

C α = Carbono-alfa

DAU = 5,6-Diaminouracila

DOPE = *Discrete Optimized Protein Energy*

EMBOSS = *European Molecular Biology Open Software Suite*

FAP = Polineuropatia Familiar Amilóide (do inglês, *Familial Amyloid Polyneuropathy*)

HBP = Bolsão de Ligação à Halógeno (do inglês, *Halogen-Binding Pocket*)

HbsHIU = HIUase de *Herbaspirillum seropedicae*

HDL = Lipoproteína de Alta Densidade (do inglês, *High Density Lipoprotein*)

HIUase = 5-Hidroxiisourato Hidrolase

HsTTR = TTR de *Homo sapiens*

HT = Hormônio Tireoidiano

iTOL = *Interactive Tree of Life*

kDa = Quilo-Dalton

LCR = Líquido Cefalorraquidiano

MutHIU = Proteína Mutante de HIUase

MutTTR = Proteína Mutante de Transtirretina

NRDB = Banco de Dado Não-Redundante (do inglês, *Non-Redundant Database*)

OHCU = 2-Oxo-4-Hidroxi-4-Carboxi-5-Ureídoimidazolina

PAE = *Predicted Aligned Error*

PDB = *Protein Data Bank*

PFAM = *Protein Families Database*

pLDDT = *Predicted Local Distance Difference Test*

PSIPRED = *PSI-blast based secondary structure PREDiction*

RBP = Proteína Ligadora de Retinol (do inglês, *Retinol-Binding Protein*)

RPE = Epitélio Pigmentar da Retina (do inglês, *Retinal Pigment Epithelium*)

RMSD = Raiz Quadrada do Desvio Quadrático Médio (do inglês, *Root Mean Square Deviation*)

SMART = *Simple Modular Architecture Research Tool*

SSA = Amiloidose Sistêmica Senil (do inglês, *Senile Systemic Amyloidosis*)

T3 = L-3,5,3'-Triiodotironina

T4 = L-3,5,3',5'-Tetraiodotironina

TBG = Globulina Ligadora de Tiroxina (do inglês, *Thyroxin-Binding Globulin*)

TRP = Proteína Relacionada à Transtirretina (do inglês, *Transthyretin-Related Protein*)

TTR = Transtirretina

VADAR = Volume Area Dihedral Angle Reporter

YRGS = Tetrapeptídeo C-terminal em HIUase (Tyr-Arg-Gly-Ser)

ZBS = Sítio de Ligação à Zinco (do inglês, *Zinc-Binding Site*)

SUMÁRIO

1. INTRODUÇÃO	16
1.1. Pré-Albuminas: um Breve Histórico	16
1.2. Transtirretinas	17
1.2.1. Estrutura	18
1.2.2. Amiloidogênese	22
1.2.3. Outras Funções	23
1.3. Processo Evolutivo das TTR	26
1.3.1. Proteínas Relacionadas às Transtirretinas	26
1.4. 5-Hidroxiisourato Hidrolases	28
1.4.1. Mecanismo de Ação	30
1.5. Engenharia de Proteínas Baseada em Evolução	33
1.5.1. Interconvertendo TTR e HIUase	33
2. OBJETIVOS	35
2.1. Objetivo geral	35
2.2. Objetivos específicos	35
3. MATERIAL E MÉTODOS	36
3.1. Análises Filogenéticas	36
3.2. Alinhamentos de Sequências	36
3.3. Estruturas Secundárias	36
3.4. Análise de Cavidades e Potenciais Eletrostáticos	37
3.5. Geração das Sequências Mutantes	37
3.6. Modelagem Molecular	39
3.6.1. <i>Modeller</i>	39
3.6.2. <i>AlphaFold</i>	40
3.6.2.1. Instalação Local.....	41
3.6.2.2. <i>ColabFold</i>	41
3.7. Validação Estrutural	42
3.8. Visualização e Geração de Figuras	42
4. RESULTADOS E DISCUSSÃO	44
4.1. Etapa 1 - Proteínas de Referência	44
4.1.1. Análises Filogenéticas da Prevalência de HIUase e TTR em Organismos Vivos	

4.1.2.	Comparações Estruturais entre HIUase e TTR	46
4.1.2.1.	Estrutura Primária	46
4.1.2.2.	Estrutura Secundária	48
4.1.2.3.	Estrutura Terciária e Quaternária	50
4.2.	Etapa 2 - Proteínas Mutantes	57
4.2.1.	Definição e Geração das Mutantes	57
4.2.2.	Geração de Modelos Tridimensionais	59
4.2.3.	Validações Estruturais	60
4.2.4.	Comparação Entre os Arcabouços das Proteínas de Referência e Mutantes.....	65
4.2.5.	Volumes e Geometria das Cavidades	68
4.2.7.	Resíduos Expostos.....	76
5.	CONCLUSÃO	80
6.	REFERÊNCIAS BIBLIOGRÁFICAS	82
7.	APÊNDICES.....	91
	Apêndice I - Mapeamento da nomenclatura de resíduos importantes para as cavidades, estruturas e outras funções de HIUase e TTR de diferentes organismos.....	91
	Apêndice II: Mapeamento da nomenclatura de resíduos importantes para as cavidades, estruturas e outras funções das proteínas mutantes e suas proteínas de referência.	94
	Apêndice III – Lista completa de resíduos expostos nas proteínas de referência e mutantes.	97

1. INTRODUÇÃO

1.1. Pré-Albuminas: um Breve Histórico

As pré-albuminas são um grupo de proteínas globulares plasmáticas, descobertas em 1942. A sua nomenclatura é uma referência ao padrão único de migração das mesmas em géis de eletroforese, sendo a única proteína plasmática mais rápida que a albumina, em pH 8,6 (Prapunpoj & Leelawatwattana, 2009; Kabat, Moore & Landow, 1942; Seibert & Nelson, 1942, Figura 1). Outras características físico-químicas marcantes incluem sua estabilidade e resistência à desnaturação mesmo em pHs variados e altas concentrações de uréia, em relação à outras proteínas globulares (Branch, Robbins & Edelhoeh, 1971).

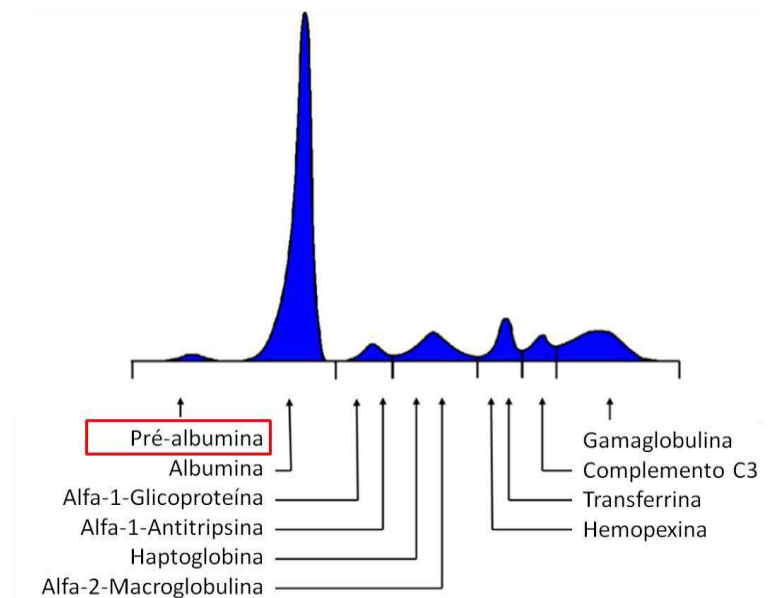


Figura 1 - Padrão de migração eletroforético das principais proteínas plasmáticas. A pré-albumina está destacada em vermelho. Adaptado com permissão de Regeniter & Siede, 2018 sob número de licença 5423640382852.

Mesmo com propriedades tão distintas, quase 20 anos se passaram para que suas funções fossem identificadas, sendo descoberto então um papel importante na ligação e no transporte dos hormônios tireoidianos (HT) e do retinol, também conhecido como vitamina A (Ingbar, 1958). Dessa forma, a nomenclatura destas proteínas foi alterada para melhor incluir suas funções biológicas, passando de pré-albumina para transtirretina (do inglês transthyretin, *transport protein for both thyroid hormone and retinol-binding protein*), nome este que perdura até os dias de hoje (Goodman, 1986).

1.2. Transtirretinas

As transtirretinas (TTR) são uma das três proteínas encontradas na circulação sanguínea de vertebrados, que tem como principal função a ligação e transporte dos HT, 3,5,3'-triiodo-L-tironina (T3) e 3,5,3',5'-tetra-L-iodotironina (Tiroxina, ou T4) (Prapunpoj & Leelawatwattana, 2009, Figura 2). Curiosamente, enquanto em vertebrados terrestres a afinidade da TTR ao T4 é muito maior, no restante dos vertebrados, a ligação ao T3 (forma ativa do hormônio) é preferencial (Yamauchi *et al.*, 1993).

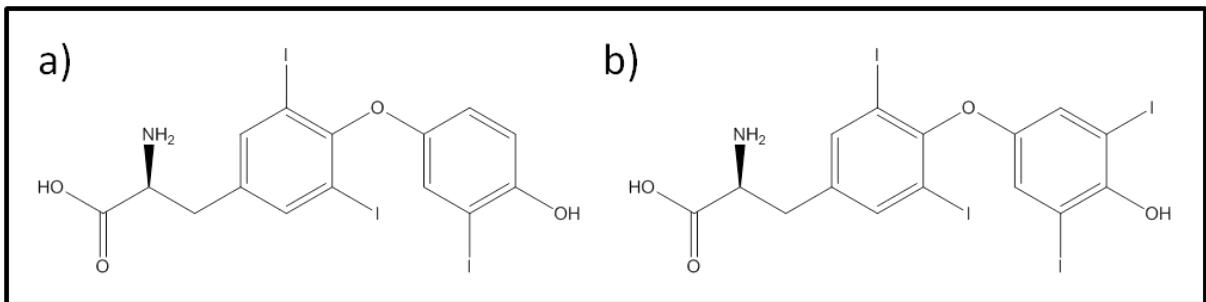


Figura 2 - Estruturas dos hormônios tireoidianos. A) Triiodotironina (T3). B) Tiroxina (T4). A única diferença entre as moléculas é a presença de um iodo adicional no carbono 5' da tiroxina.

Os HT são solúveis em membranas lipídicas, e por isso, quando presentes no plasma sanguíneo, necessitam de uma proteína carreadora. A alta afinidade desses hormônios às TTR permite que sempre exista uma fração suficiente dos mesmos circulando no sangue (Aldred, Brack & Schreiber, 1995; Mendel *et al.*, 1987).

Embora a principal proteína transportadora de HT no plasma sanguíneo seja a globulina ligadora de tiroxina (TBG), seguida de TTR e albumina sérica (65%, 10% e 20%, respectivamente), no líquido cefalorraquidiano (LCR) cerca de 60% do transporte se dá por meio da TTR (Bartalena & Robbins, 1993; Hagen & Elliot, 1973). Por ser circulante tanto no plasma sanguíneo quanto no LCR, que não possui contato direto com a circulação sistêmica, as TTR são sintetizadas em locais distintos nos organismos. Na corrente sanguínea a síntese se dá pelo fígado, e no sistema nervoso pelo plexo coróide (que também produz cerca de 70% do LCR) (Dickson *et al.*, 1987; Richardson, 2015). Outros tecidos também possuem expressão de TTR em níveis menores, como o epitélio pigmentar da retina (RPE) localizado no olho, e as ilhotas pancreáticas (Buxbaum & Reixach, 2009).

Uma das principais diferenças da TTR para as outras duas outras proteínas transportadoras majoritárias de HT se dá no fato das TTR exibirem diversos sítios não específicos de ligação a outros compostos, que permitem com que ela exerça funções adicionais (Cody & Wojtczak, 2009). Estas funções serão discutidas nos tópicos a seguir.

1.2.1. Estrutura

In vivo, a TTR pode ser encontrada como um tetrâmero constituído por quatro subunidades monoméricas idênticas de aproximadamente 14 kDa cada, totalizando cerca de 55 kDa (Figura 3a). Cada monômero é formado por 127 resíduos de aminoácido, dispostos em oito fitas- β (A-H) conectadas por sete voltas, e que se agrupam de forma antiparalela formando duas folhas- β com quatro fitas cada (DAGH e CBEF; Figura 3b). Além disso, é possível identificar uma única α -hélice de nove resíduos entre as voltas E e F, uma cauda C-terminal de cinco resíduos, e um N-terminal de dez resíduos sem estrutura definida (Blake *et al.*, 1974; Blake *et al.*, 1978).

A estabilidade em condições desnaturantes observadas nestas proteínas pode inclusive ser atribuída à grande concentração de folhas- β que possuem. Devido a essas estruturas, há a presença de um número maior de interações de hidrogênio entre as interfaces, as quais aumentam a área de interação e, conseqüentemente, a estabilidade da molécula (Branch, Robbins & Edelhoch, 1972; Blake *et al.*, 1978).

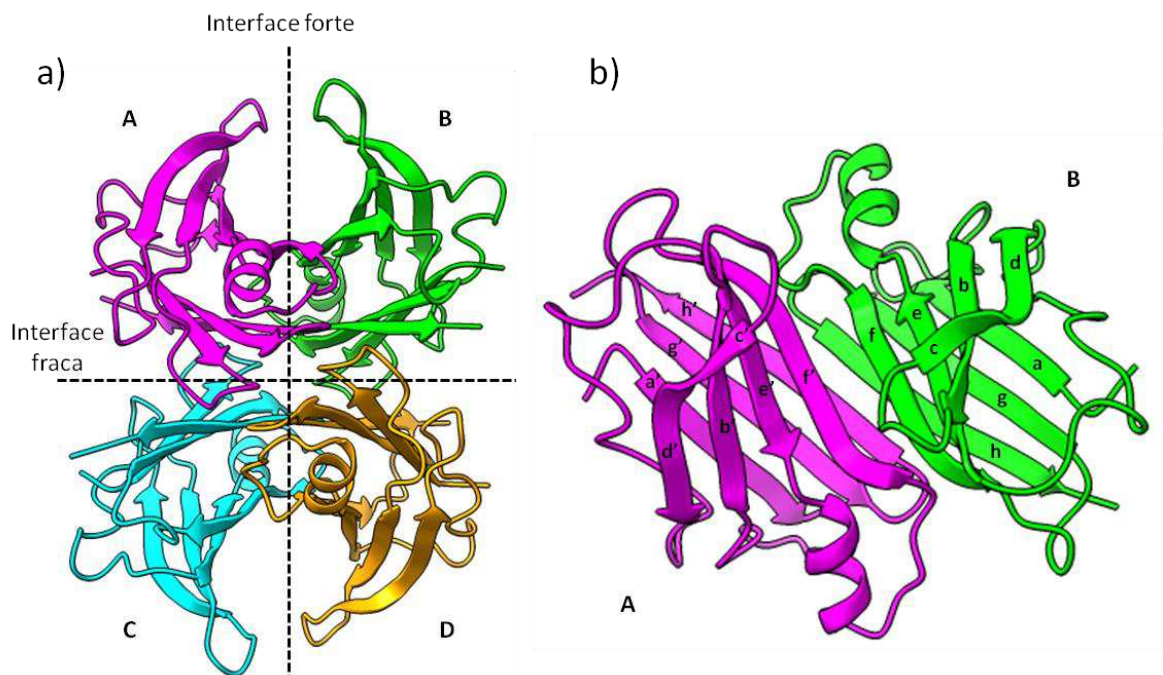


Figura 3 - Estrutura tridimensional de uma TTR. A) Estrutura quaternária do homotetrâmero da proteína evidenciando as interfaces fortes e fracas entre os dímeros. B) Detalhe do dímero AB evidenciando as oito folhas- β que compõem cada monômero, assim como a única α -hélice. Letras em maiúscula representam cada uma das cadeias monoméricas da proteína. O eixo da cavidade (horizontal) percorre a interface fraca. Letras em minúsculo representam cada folha- β . Figuras preparadas pelo autor utilizando ChimeraX v1.2. Código PDB da estrutura: 1F41 (Hörnberg *et al.*, 2000).

Os dímeros de TTR são compostos por duas folhas- β de oito fitas cada, sendo quatro de cada monômero. Dessa forma, a ligação se dá de forma DAGHH'G'A'D' (folha interna) e CBEFF'E'B'C' (folha externa). O arranjo antiparalelo das fitas H e H', e F e F' formam uma estrutura β -sanduíche estendida ao longo da interface, estabilizada por ligações de hidrogênio entre as fitas H e H'. Por outro lado, as fitas F e F' são estabilizadas pela inserção de um resíduo de fenilalanina (Phe87) em um bolsão hidrofóbico na subunidade oposta (Figura 4). Este arranjo também é chamado de interface forte (Figura 3a, eixo vertical), e ocorre no contato entre os monômeros A e B, e C e D (Leach *et al.*, 2019).

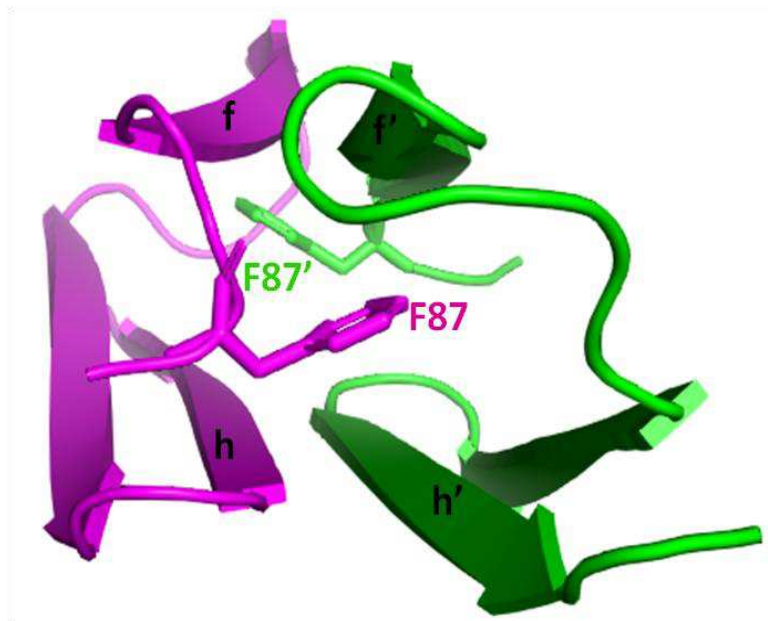


Figura 4 - Detalhe do bolsão hidrofóbico da interface forte. Em magenta está mostrado o monômero A da proteína, e em verde o monômero B. Os resíduos de fenilalanina estão destacados, apontando em direção à cadeia oposta. As fitas- β de cada cadeia são identificadas por letras minúsculas. Figura gerada pelo autor utilizando PyMOL v2.5.2. Código PDB da estrutura: 1F41 (Hörnberg *et al.*, 2000).

Já a interação dos dímeros, na chamada interface fraca (Figura 3a, eixo horizontal), leva à formação do tetrâmero da proteína. Esta interação é feita pela associação das voltas AB e GH das subunidades diagonalmente opostas (Figura 3a, monômeros A com D, e B com C). A estabilização dessa interface se dá por meio de interações hidrofóbicas e ligações de hidrogênio intermoleculares (Blake *et al.*, 1974; Blake *et al.*, 1978; Leach *et al.*, 2019).

Na estrutura quaternária final, uma grande cavidade é formada no interior do tetrâmero, possuindo cerca de 8Å de diâmetro e 50Å de comprimento (Figura 3a, ao longo da interface fraca). A cavidade é composta tanto por resíduos hidrofílicos (grupos hidroxila dos resíduos Ser117 e Thr119), hidrofóbicos (grupos metila dos resíduos Leu17, Thr106, Ala108, Leu110, Thr119 e Val121), além de resíduos carregados (grupos laterais dos resíduos Lys15,

Glu54 e His56). Não há presença de nenhum dos 48 resíduos aromáticos e hidrofóbicos presentes na estrutura primária da proteína (Blake & Oatley, 1977). Dentro da cavidade encontram-se dois sítios de ligação aos HT, dispostos de maneira oposta e separados por uma constrição central de 4Å, causada pelas cadeias laterais dos resíduos Leu110, Ser115 e Ser117 (Blake & Oatley, 1977).

Cada um dos sítios de ligação contém três pares de bolsões de ligação a halógenos (HBP), onde se ligam os átomos de iodo dos HT. Cada um dos pares é relativo a um dos monômeros e seus pares de simetria, sendo eles HBP1-HBP1', HBP2-HBP2', e HBP3-HBP3' (Figura 5). Cada molécula dos HT faz interações com dez resíduos presentes nos HBPs, incluindo contatos polares com as cadeias laterais dos resíduos Lys15, Glu54 e Thr106; e contatos hidrofóbicos com os resíduos Leu17, Ala108, Ala109, Leu110. Outros resíduos como Ser117, Thr119 e Val121 também estão envolvidos. Todos estão localizados nos arredores da cavidade da proteína (Wojtczak *et al.*, 1996; Tomar *et al.*, 2012).

Apesar do duplo sítio de ligação aos hormônios, apenas uma molécula pode se ligar à proteína. Isso se deve a uma interação cooperativa negativa causada por fatores eletrostáticos e estruturais, ligados à assimetria dos dois sítios de ligação. A ligação do hormônio ao sítio de maior diâmetro leva a mudanças conformacionais na proteína, por meio de mudanças na rede de ligações de hidrogênio da molécula. Dessa forma, o segundo sítio é colapsado e impedido de se ligar ao hormônio. Além disso, o posicionamento de moléculas de água nas cavidades, e o movimento do resíduo Ser117 após ligação aos TH, também são relacionados com a cooperatividade negativa da ligação (Neumann, Cody & Wojtczak, 2001; Prapunpoj & Leelawatwattana, 2009).

Em ensaios cristalográficos, a unidade assimétrica do estudo é caracterizada pela estrutura dimerizada, ao longo da interface forte, uma vez que a interação entre os monômeros individuais é muito próxima (Blake *et al.*, 1978). Isso se reflete nas 3971 estruturas¹ depositadas no *Protein Data Bank* (PDB), que majoritariamente seguem este padrão. O conjunto biológico (do inglês, “*biological assembly*”) caracteriza então o tetrâmero, formado pela união dos dímeros ao longo do eixo de simetria da interface forte.

¹ Até o dia 14/11/2022, existem 3971 estruturas depositadas com a palavra-chave “*transthyretin*” no PDB. Porém, uma parte desse número se refere a proteínas relacionadas, como as HIUases (que serão mencionadas posteriormente neste trabalho), mas que também mantém o mesmo padrão estrutural das TTR.

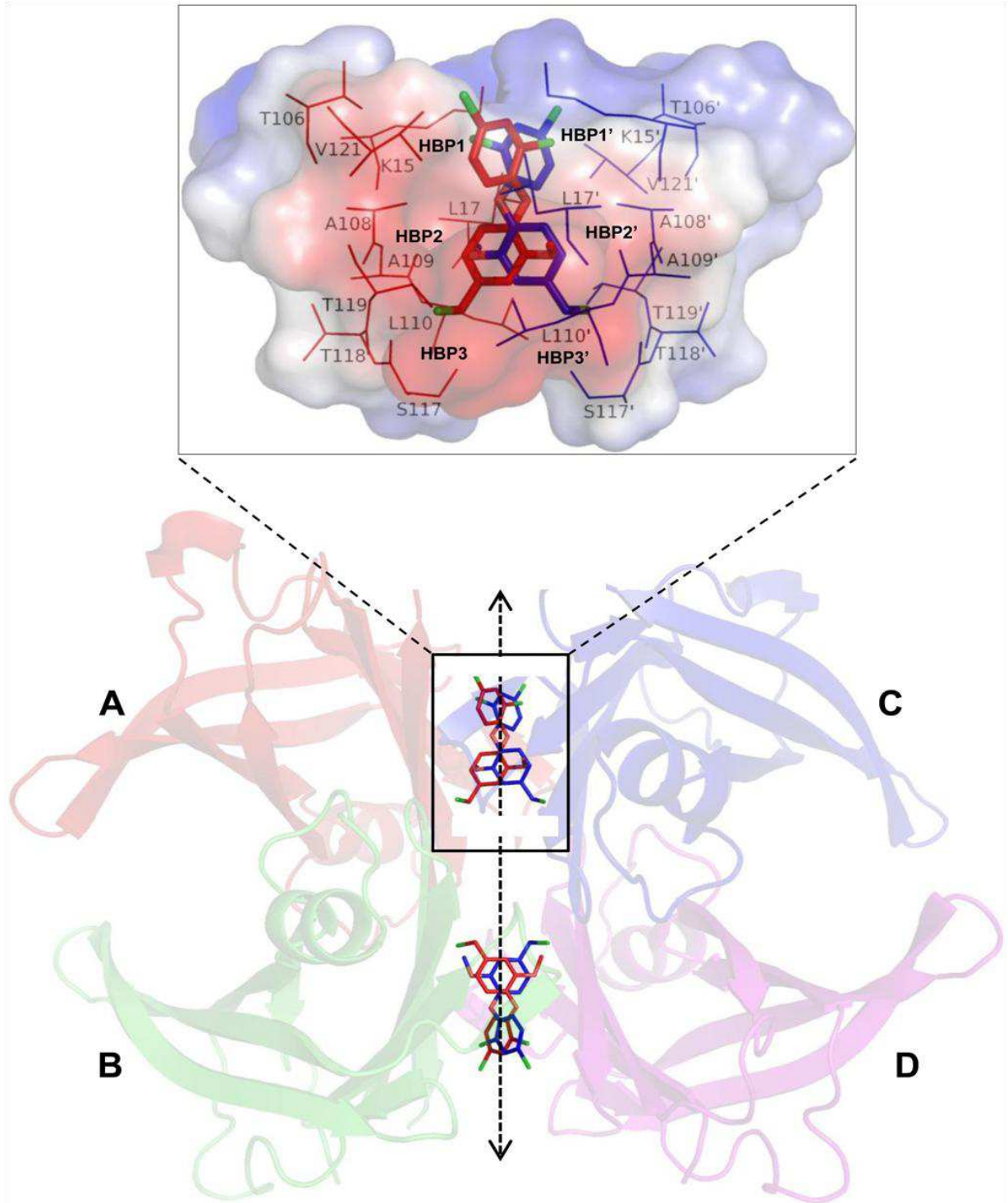


Figura 5 - Diagrama esquemático da ligação do T4 na estrutura da TTR humana. No painel superior são evidenciados os bolsões de ligação a halógenos, além dos resíduos mais importantes que estão em contato com o hormônio. No painel inferior é mostrada a estrutura completa da proteína, evidenciando a posição de ligação nas cavidades. Devido à simetria da proteína ao longo do eixo da cavidade (seta vertical), são mostrados os dois modos de ligação do hormônio. Letras em maiúscula indicam cada uma das unidades do homotetrâmero. Adaptado com permissão de Tomar *et al.*, 2012, sob *Creative Commons Attribution License*.

1.2.2. Amiloidogênese

Uma das principais causas do grande interesse no estudo das transtirretinas, especialmente em humanos, é a grande variedade de implicações fisiológicas e patológicas que a mesma está envolvida. As áreas mais estudadas, sem dúvidas, são aquelas que abordam as condições fisiopatológicas nas quais há a formação de fibras amilóides causadoras de uma série de doenças.

Os amilóides são depósitos de proteínas mal enoveladas e insolúveis, que podem variar de nanômetros a micrômetros, ocorrer em diferentes órgãos e tipos celulares, e serem hereditárias ou adquiridas (Planté-Bordeneuve & Said, 2011; Al Hamed *et al.*, 2021). Amilóides também são considerados, por alguns autores, como um padrão de estrutura ordenada (Lima & Icart, 2022). As TTR são intrinsecamente amiloidogênicas, podendo levar a um quadro chamado de amiloidose sistêmica senil (SSA), que acomete indivíduos idosos e é fatal sem o tratamento adequado (Yee *et al.*, 2019; Pinney *et al.*, 2013).

O mecanismo de agregação da TTR ainda não é totalmente compreendido, porém os modelos mais aceitos atualmente propõem um processo bilateral entre o estado nativo tetramérico, e estados nativos intermediários diméricos e monoméricos (Figura 6, Ruberg *et al.*, 2019). Uma vez dissociado e desestabilizado, os monômeros tendem a se desenovelar e começam a se agregar. Esses pequenos oligômeros amorfos vão se juntando até que ocorra a formação de fibras conhecidas como amilóides, podendo desencadear as doenças descritas (Ruberg *et al.*, 2019).

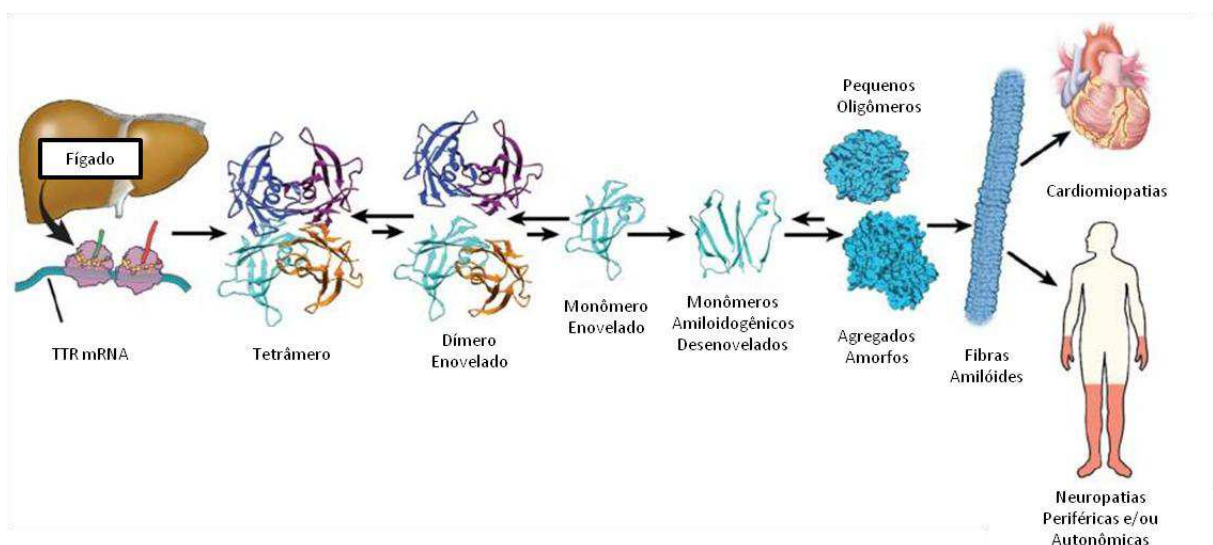


Figura 6 - Mecanismo proposto de agregação da transtirretina. São mostrados os processos desde a síntese hepática da proteína, passando pela via bilateral de enovelamento e desenovelamento, até a formação de agregados amorfos e fibras amilóides, levando à patologias órgão-específicas. Adaptado com permissão de Ruberg *et al.*, 2019 sob número de licença 5422101425084.

As mutações em diversos resíduos das TTR podem levar a uma desestabilização do tetrâmero da proteína, causando agregação. Dos 127 resíduos de aminoácido da proteína madura, cerca de 150 mutações já foram identificadas² (Rowczenio et al., 2014). A maioria das mutações descritas resulta em polineuropatia familiar amilóide (FAP), uma doença transmitida de forma autossômica dominante e fatal (Planté-Bordeneuve & Said, 2011).

Diversos órgãos podem ser afetados por agregados amilóides de TTR, especialmente os sistemas cardiovascular e nervoso. Das mutações patológicas mais importantes do ponto de vista médico, destacam-se: Val¹²²Ile (Jacobson, Gorevic & Buxbaum, 1990), variante cardiopática mais comum; Val³⁰Met, variação neuropática mais comum, fortemente amiloidogênica (Saraiva *et al.*, 1984); e Leu⁵⁵Pro, variação mais agressiva, capaz de ser secretada e acometer diversos órgãos (Jacobson *et al.*, 1992).

Entretanto, existem também mutações não-amiloidogênicas, como a Ala¹⁰⁸Ile e Thr¹¹⁹Met, que estabilizam o tetrâmero da proteína e impedem a agregação (Sant'Anna *et al.*, 2017). Estas mutações podem servir como base para o estudo de tratamentos contra as doenças causadas pela agregação de TTR.

1.2.3. Outras Funções

Além de serem proteínas carreadoras de hormônios tireoidianos, as TTR também possuem diversas outras funções. A mais antiga e bem elucidada é a ligação com a proteína ligadora de retinol (RBP), por meio da formação de um complexo macromolecular entre as duas proteínas. O retinol se liga à RBP, que por sua vez, se liga à TTR (Figura 7; Monaco, 2000). Devido ao mecanismo de transporte do composto, a afinidade à forma holo-RBP pela TTR é maior, o que evita a perda de retinol e RBP pela filtração glomerular renal. Ainda assim, ambas as formas holo- e apo- RBP podem formar o complexo com a TTR (Fex, Albertsson & Hansson, 1978; Naylor & Newcomer, 1999). Outro ponto importante é que, ao contrário de vertebrados terrestres, as transtirretinas de peixes não formam o complexo com a RBP, devido principalmente à ausência do tetrapeptídeo C-terminal Arg-Asn-Leu-(Ser)-Leu (RNL(S)L) (Santos, Anjos & Power, 2002; Liz *et al.*, 2020).

² Lista completa disponível em: <http://amyloidosismutations.com>

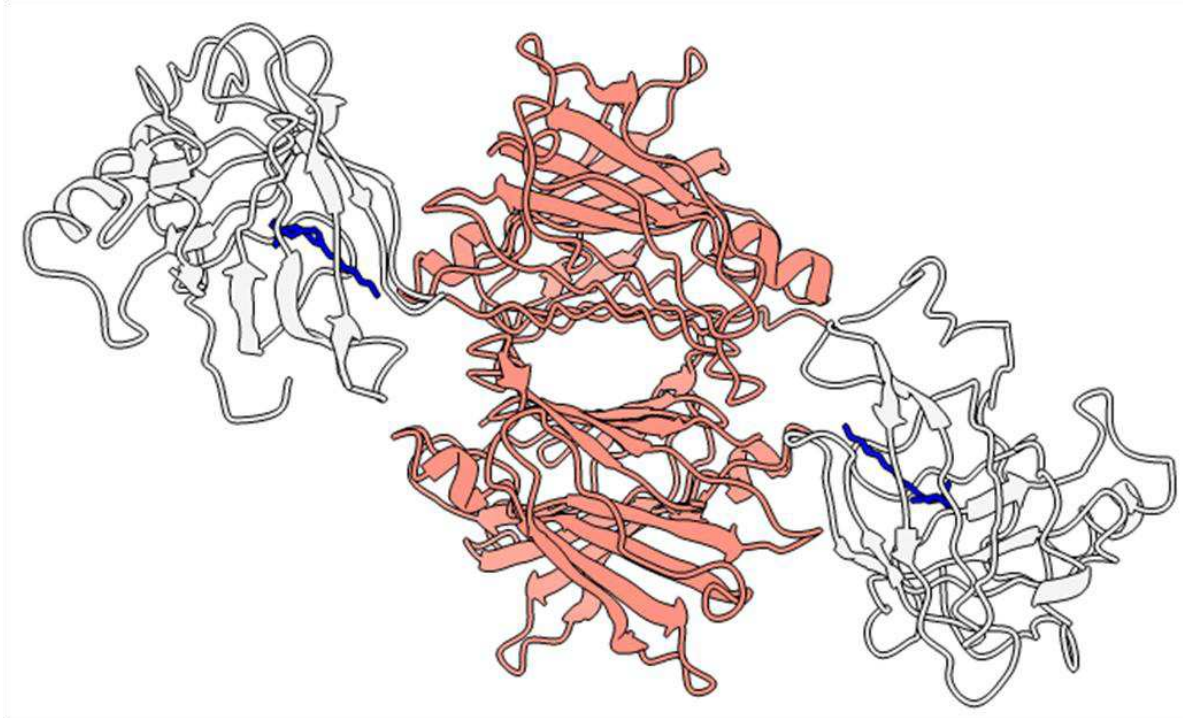


Figura 7 - Complexo Transtirretina-RBP-Retinol. A TTR é mostrada em salmão, ligada a duas moléculas de RBP, em branco, cada uma em um dos dímeros da proteína. As moléculas de retinol são mostradas em azul. Figura preparada pelo autor utilizando ChimeraX v1.2. Código PDB da estrutura: 1QAB (Naylor & Newcomer, 1999).

Outra propriedade das TTR é a de ligação ao zinco. Já é sabido que altas concentrações de Zn^{2+} poderiam induzir a formação de fibras amilóides por essa proteína, o que também evidenciaria uma possível interação entre este íon metálico e a TTR (Wilkinson-White & Easterbrook-Smith, 2007). Ao analisar a estrutura da TTR humana em diferentes condições de pH, Palmieri e colaboradores conseguiram identificar três sítios de ligação a zinco na proteína, especialmente em condições acídicas, presentes na mesma região de ligação à RBP, também fora da cavidade de ligação aos HT (Figura 8; Palmieri *et al.*, 2010).

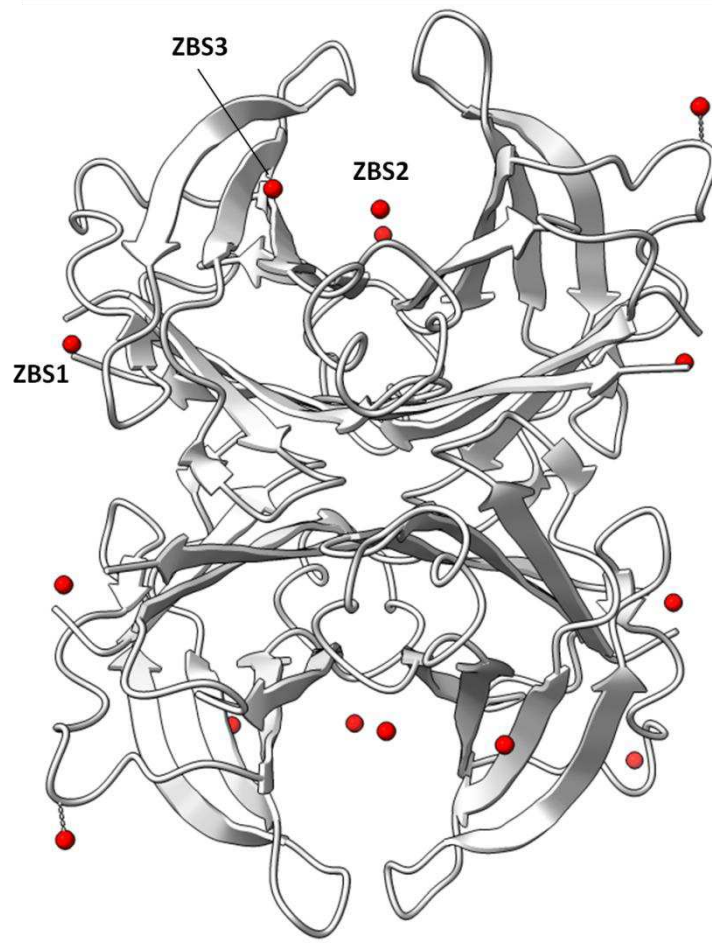


Figura 8 - Sítios de ligação a zinco na TTR. A estrutura tetramérica de TTR humana é mostrada com as moléculas de zinco ligadas, identificadas em vermelho. Os três sítios principais de ligação à zinco (ZBS) também são mostrados. Figura preparada pelo autor utilizando ChimeraX v1.2. Código PDB da estrutura: 3DGD (Palmieri *et al.*, 2010).

Em 2005, foi proposta, por Liz e colaboradores, uma função de protease críptica para a TTR (Liz & Sousa, 2005). O mecanismo proposto diz que a TTR atua principalmente na clivagem C-terminal da apolipoproteína AI (apoA-I), após o resíduo Phe225. A apoA-I é uma proteína plasmática capaz de se ligar e carregar, dentro de vesículas de lipoproteínas de alta densidade (HDL), uma pequena fração de TTR livre na corrente sanguínea (Sousa, Berglund & Saraiva, 2000; Liz *et al.*, 2004). Outros substratos foram posteriormente identificados, como neuropeptídeo Y e peptídeo β -amilóide, sendo a deposição deste último, diminuída pela capacidade proteolítica da TTR, uma das principais causas da doença de Alzheimer (Liz *et al.*, 2009; Schwarzman & Goldgaber, 1996). Em 2012, os mesmos autores que inicialmente descreveram a função de protease para a TTR a estabeleceram definitivamente como uma metalopeptidase zinco-dependente (Liz *et al.*, 2012).

Outras funções também já foram descritas, como a ligação ao ácido retinóico (Smith *et al.*, 1994) e à flavonóides (Lueprasitsakul *et al.*, 1990), participação na função motora e melhoria da regeneração de nervos (Riisoen, 1988), aumento na capacidade exploratória e na redução de sinais depressivos em ratos TTR-nocaute (Sousa *et al.*, 2004), além de manutenção da memória durante o envelhecimento (Brouillette & Quirion, 2008).

1.3. Processo Evolutivo das TTR

Apesar de serem extremamente importantes no metabolismo de vertebrados, e capazes de desencadear diversos processos patogênicos como amiloidogêneses, importantes em especial na área da saúde humana, o processo evolutivo das TTR demorou a ser compreendido. Apenas na década de 1990, mais de 50 anos depois da descoberta inicial das TTR, Aldred e colaboradores, analisando a porção N-terminal da proteína, identificaram que esta região era a que mais se modificava ao longo do curso da evolução, já que o restante da sequência primária se encontrava surpreendentemente conservado (Aldred, Prapunpoj & Schreiber, 1997).

A região N-terminal é uma das componentes da entrada da cavidade central, e possui uma estrutura tridimensional desordenada, contrastando com os resíduos do canal central, que não sofreram alterações sequenciais e estruturais nos últimos 400 milhões de anos (Hamilton *et al.*, 1993; Prapunpoj & Leelawatwattana, 2009). Além disso, diferentes polaridades e tamanhos eram encontrados entre espécies, sendo a região N-terminal distintamente mais hidrofóbica e maior em aves do que em mamíferos placentários. Ao serem analisadas as regiões codificadoras de TTR no genoma de diversos organismos, a porção mais variável se encontra justamente na borda dos éxons 1 e 2 do RNA mensageiro. Sendo assim, a pressão seletiva sobre o gene da TTR levou a um movimento progressivo do sítio de *splicing* entre os éxons 1 e 2 na direção 3' (Aldred, Prapunpoj & Schreiber, 1997). Logo após, foi postulado que as características do N-terminal da proteína determinariam a preferência de ligação aos HT, onde proteínas com regiões menores e mais hidrofílicas, presentes em vertebrados terrestres mais derivados, como mamíferos placentários, teriam maior afinidade ao T4 (Chang *et al.*, 1999).

1.3.1. Proteínas Relacionadas às Transtirretinas

Sequências relacionadas às TTR foram identificadas por Prapunpoj e colaboradores (2000), em janelas abertas de leitura no genoma de *Caenorhabditis elegans* (um nematódeo) e diversos outros microorganismos. Sendo às TTR presentes apenas em vertebrados, os autores

sugerem que a origem da família teria sido a partir de genes ancestrais³ relacionados à transtirretina, e concluíram que a maior parte da história evolutiva das TTR teria ocorrido antes da diferenciação de peixes e anfíbios, ou seja, antes do surgimento dos vertebrados (Prapunpoj *et al.*, 2000). Estas proteínas ancestrais foram denominadas Proteínas Relacionadas às Transtirretinas (TRP), e teriam dado origem às TTR por um evento de duplicação gênica anterior à separação de peixes cartilagosos (*chondrichthyes*) e ósseos (*osteichthyes*), ocorrida aproximadamente 400 milhões de anos atrás (Zanotti *et al.*, 2006).

Pouco depois, foram descobertos na bactéria aeróbia facultativa *Bacillus subtilis*, a presença de diversos novos genes envolvidos no metabolismo de purinas. Um desses genes, o *pucM*, codifica a enzima uricase nestes organismos, e possui 55% de identidade de aminoácidos com as TTR (Jung *et al.*, 2006). Além disso, diversos procaríotos, incluindo vertebrados, possuem genes similares ao *pucM* (Schultz, Nygaard & Saxild, 2001). Em 2005, pela primeira vez, foi observado que as TRP estariam envolvidas na hidrólise de 5-hidroxiisourato (5-HIU) em 2-oxo-4-hidroxi-4-carboxi-5-ureídoimidazolina (OHCU), um dos passos até então desconhecidos da via de degradação do ácido úrico, praticamente ubíqua em seres vivos, mas ausente em homínídeos (Lee *et al.*, 2005; Ramazzina *et al.*, 2006) (Figura 9).

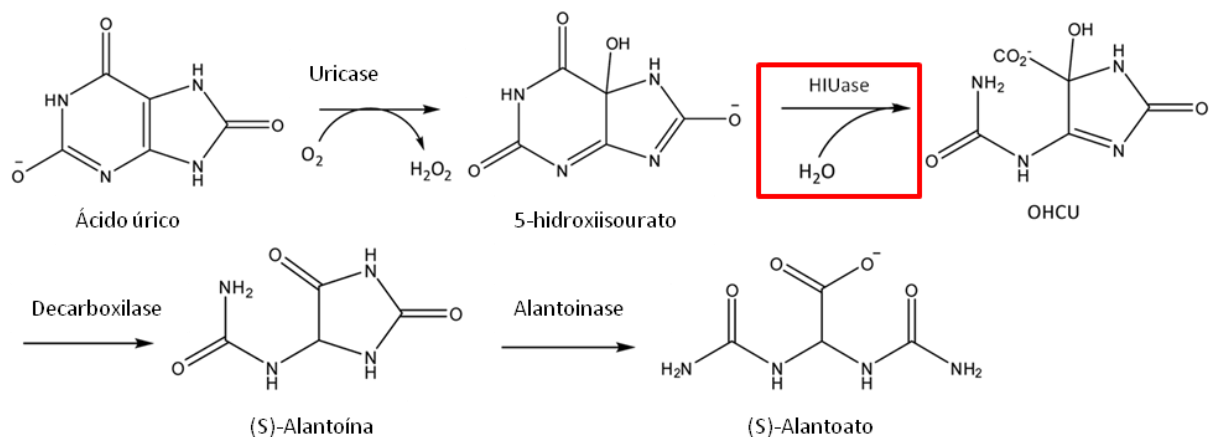


Figura 9 - Representação esquemática da via de degradação do ácido úrico. A enzima HIUase está destacada em vermelho. Adaptado de Jung *et al.*, 2006 sob *Creative Commons Attribution License*.

A via completa de degradação do ácido úrico é composta por cinco espécies químicas e quatro enzimas (Figura 9). A reação inicia-se com o ácido úrico propriamente dito, que é convertido em 5-HIU pela uricase (ou urato oxidase). Logo após, ocorre a hidrólise do 5-HIU em OHCU pela HIUase, seguida da descarboxilação do OHCU em (S)-Alantoína. Por fim, a

³ Genes originários de organismos filogeneticamente basais, como no caso dos vertebrados primitivos.

(S)-Alantoína é também hidrolisada, pela enzima Alantoinase, sendo convertida em (S)-Alantoato, produto final da via (Kahn, Serfozo & Tipton, 1997; Jung *et al.*, 2005).

Dessa forma, embora sequencialmente relacionadas às TTR, as TRP não são funcionalmente relacionadas às mesmas (Lee *et al.*, 2005). Com a função definida, as TRP foram então renomeadas para 5-hidroxiisourato hidrolases (HIUase).

1.4. 5-Hidroxiisourato Hidrolases

Após terem suas funções identificadas e seus nomes modificados, as HIUase foram extensivamente estudadas nos anos seguintes, por se tratarem de um excelente exemplo de evolução divergente (Zanotti *et al.*, 2006). As estruturas das HIUase de *Bacillus subtilis* (bactéria gram-positiva presente no solo; Jung *et al.*, 2006), *Salmonella dublin* (enterobactéria gram-negativa; Hennebry *et al.*, 2006), *Escherichia coli* (enterobactéria gram-negativa; Lundberg *et al.*, 2006) e *Zebrafish* (*Danio rerio*, peixe teleosteo; Zanotti *et al.*, 2006), foram determinadas por difração de raios X de maneira independente, e apresentam diferenças marcantes em relação à TTR.

Mesmo também se organizando de maneira homotetramérica e possuindo um padrão de organização tridimensional semelhante às TTR, as HIUase não são capazes de se ligar aos hormônios tireoidianos (Hennebry *et al.*, 2006; Figura 10). Apesar disso, o padrão de organização secundária em HIUase se dá da mesma forma que em TTR, onde cada monômero também apresenta oito folhas- β e uma α -hélice, que dimerizam por ligações de hidrogênio entre as fitas-H de cada monômero. O tetrâmero é então formado por contatos entre os dímeros, especialmente nas voltas AB e GH (Hennebry *et al.*, 2006; Zanotti *et al.*, 2006).

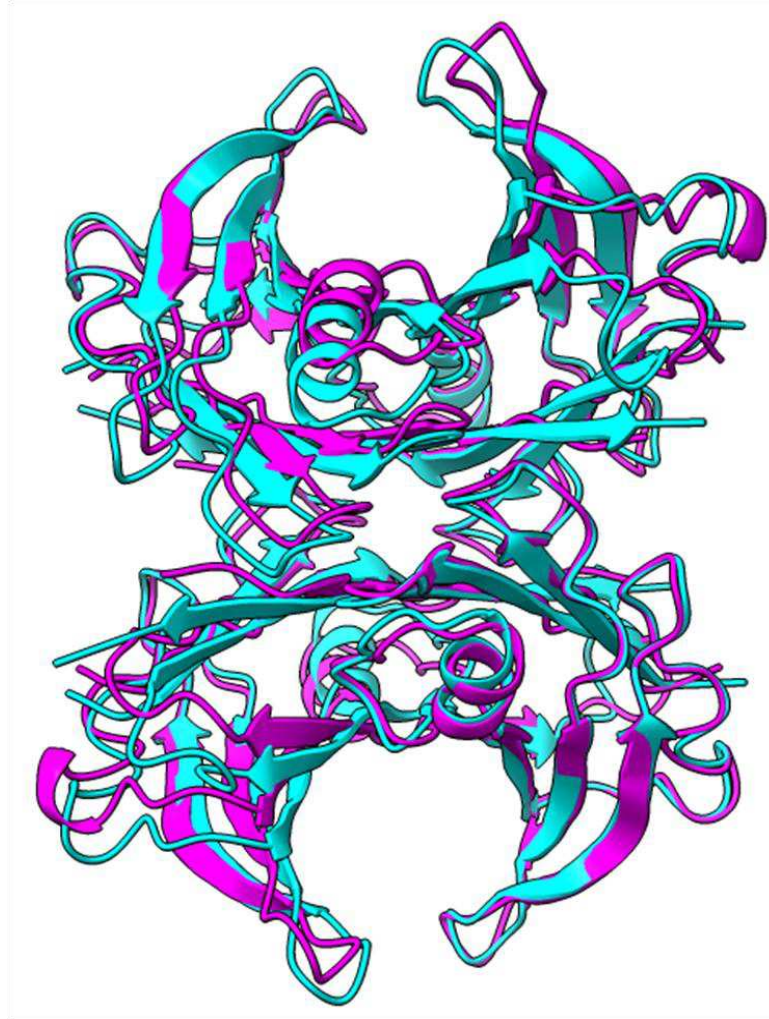


Figura 10 - Sobreposição das estruturas tridimensionais de TTR e HIUase. TTR de *Homo sapiens* é mostrada em ciano (ID PDB: 1F41; Hörnberg et al., 2000). HIUase de *Zebrafish* é mostrada em magenta (ID PDB: 2H1X; Zanotti et al., 2006). Figura preparada pelo autor utilizando ChimeraX v1.2.

A cavidade central, também presente em HIUase, serve como centro de encontro dos bolsões catalíticos da enzima, e é onde ocorre a conversão do 5-HIU em OHCU. A cavidade possui uma série de resíduos altamente conservados entre todas as proteínas da subfamília, e únicos em relação à TTR (Lundberg et al., 2006). Entretanto, quando comparada a uma TTR, a cavidade das HIUase é significativamente menos profunda (cerca de 8Å de profundidade) devido à presença de um grande resíduo de tirosina, que causa um “estrangulamento” no meio da cavidade, impedindo o acesso ao bolsão central (Jung et al., 2006; Zanotti et al., 2006). Além disso, em HIUase, a cavidade exhibe principalmente cargas negativas no interior do bolsão e positivas nas extremidades, sendo que em TTR as características são exatamente opostas (Lundberg et al., 2006).

De forma contrária às TTR, as HIUase possuem um tetrapeptídeo C-terminal, Tyr-Arg-Gly-Ser, formado pelos quatro últimos resíduos da proteína. Esse tetrapeptídeo é

absolutamente conservado em todas as proteínas da subfamília, e que exhibe uma conformação altamente estável (Jung *et al.*, 2006). Estes resíduos se localizam na interface entre os dímeros, na mesma posição do sítio de ligação a halógenos em TTR, e explicam, juntamente com a diminuição do volume da cavidade, a inabilidade destas proteínas de se ligarem aos HT (Hennebry *et al.*, 2006).

1.4.1. Mecanismo de Ação

O 5-HIU é um dos intermediários da via de degradação do ácido úrico, e sendo um composto altamente instável e reativo, deve ser metabolizado rapidamente. Sua meia vida em pH fisiológico é de cerca de trinta minutos, e sua degradação não enzimática também leva à formação de OHCU, outro composto extremamente instável e reativo. Porém, ao ser degradado espontaneamente, o OHCU é convertido em uma mistura racêmica de alantoína, e não somente (S)-alantoína, que é o composto presente em células vivas (Kahn & Tipton, 1998; Lee *et al.*, 2005; Ramazzina *et al.*, 2006) (Figura 9). Dessa forma, o acoplamento rápido das reações desta via se faz necessário, a fim de se evitar a formação de compostos indesejáveis no ambiente celular.

Para estudar o mecanismo da reação de hidrólise do 5-HIU, Jung e colaboradores co-cristalizaram a enzima HIUase de *Bacillus subtilis* com dois substratos (Figura 11): 8-Azaxantina (AZX), um análogo do ácido úrico e inibidor competitivo da uricase, primeira enzima da via (Colloc'h *et al.*, 1997); e 5,6-diaminouracila (DAU), um substrato mais estável e de degradação lenta pela enzima (Raychaudhuri & Tipton, 2003). Nenhuma das estruturas revelou mudanças estruturais significativas em relação à proteína nativa, tanto na conformação monomérica ou tetramérica (Jung *et al.*, 2006).

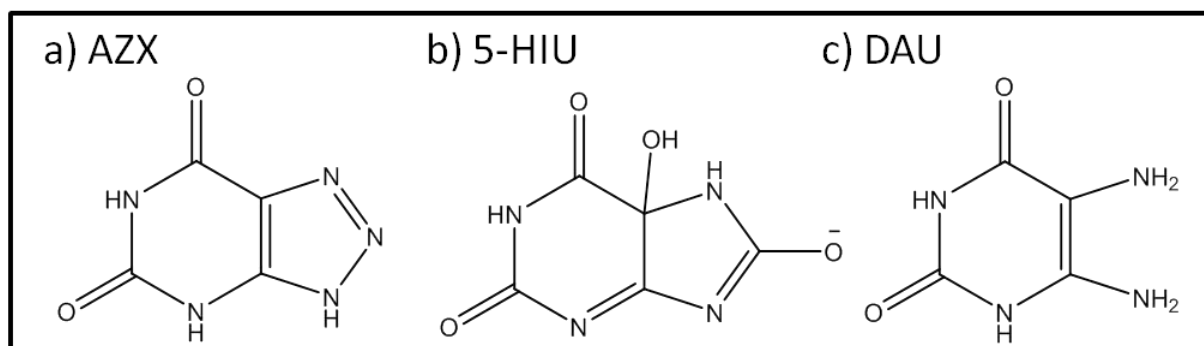


Figura 11 - Comparação entre as estruturas dos ligantes de HIUase. A) AZX, análogo de ácido úrico e um inibidor competitivo da uricase. B) 5-HIU, substrato preferencial da enzima. C) DAU, substrato mais estável e de degradação lenta.

Olhando para a cavidade catalítica, assim como em TTR, a presença de um eixo de simetria ao longo da cavidade faz com que existam dois sítios de ligação por dímero, permitindo assim o posicionamento das moléculas no sítio catalítico em duas conformações distintas, porém espelhadas (Lee *et al.*, 2006).

Em relação ao modo de ligação, o tetrapeptídeo C-terminal conservado em HIUase (YRGS) se alinha na porção lateral da cavidade, com os resíduos Arg49 e His105 se posicionando logo na entrada do sítio ativo⁴, que tem na porção central um par de His14 simétricos (Jung *et al.*, 2006; Zanotti *et al.*, 2006). Dessa forma, a molécula de AZX se posiciona a uma distância de 5-6Å do sítio ativo, com dois átomos de oxigênio interagindo com as aminas das cadeias laterais dos resíduos Arg49. Os resíduos Tyr118 interagem com a tripla de átomos de nitrogênio na molécula, e os resíduos His14 e His105 fazem contato com os outros átomos do AZX (Figura 12).

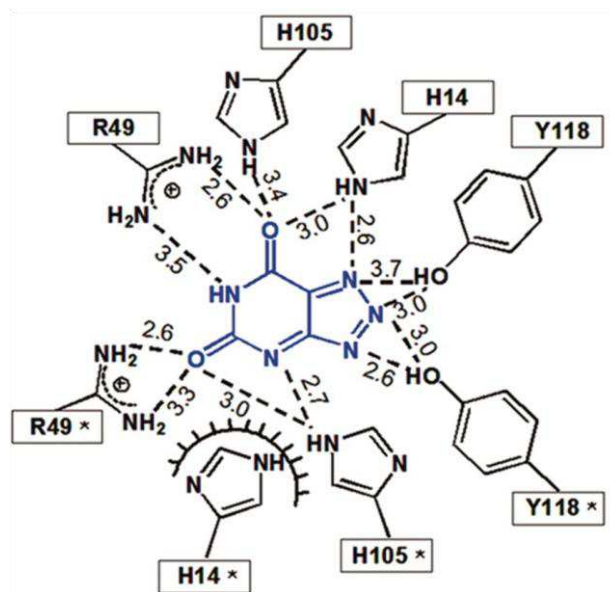


Figura 12 - Modo de ligação da molécula de AZX na cavidade catalítica de uma HIUase. A molécula de AZX é mostrada em azul, e a proteína em preto. O número do resíduo é mostrado nas caixas, e resíduos com asterisco indicam o equivalente de simetria na interface oposta. Linhas tracejadas indicam ligações de hidrogênio e suas distâncias previstas, e arcos indicam interações de Van der Waals. Numeração dos resíduos de acordo com HIUase de *Bacillus subtilis*, Adaptado com permissão de Jung *et al.*, 2006 sob *Creative Commons Attribution License*.

A reação completa desempenhada pela enzima foi proposta diversas vezes ao longo dos anos, utilizando abordagens experimentais *in vitro* com compostos similares ao instável 5-HIU (Raychaudhuri & Tipton, 2003; Jung *et al.*, 2006), ou modelagem computacional *in silico* (French & Ealick, 2011). O primeiro mecanismo proposto seria igual ao de glicosilases,

⁴ Nomenclatura de HIUase de *Bacillus subtilis*.

que utilizam resíduos de glutamato como potentes nucleófilos iniciadores da reação. Porém, as HIUase não apresentam nenhum tipo de resíduo nucleófilo na região da cavidade enzimática. A reação então é iniciada por uma molécula de água que é desprotonada pela dupla de resíduos de histidina localizados nas proximidades (Figura 13, painel superior esquerdo). A ligação de hidrogênio entre um resíduo de aspartato proximal à histidina leva ao aumento de basicidade deste resíduo, permitindo que ocorra a desprotonação da molécula de água.

As interações de hidrogênio formadas orientam a molécula de água de forma que ocorra um ataque ao carbono C6 do anel purínico do 5-HIU, levando à formação de um intermediário oxianion tetraédrico. A carga do oxianion é então estabilizada por um resíduo de arginina, de carga positiva (Figura 13, painel superior direito). O rearranjo dos elétrons do oxianion leva então à abertura do anel purínico do 5-HIU, e o último próton necessário também é doado pelo resíduo de arginina. A regeneração do sítio catalítico se dá pela doação de um próton da histidina inicial (vindo da molécula de água catalítica) para a arginina, completando o ciclo da enzima (Figura 13, painel inferior).

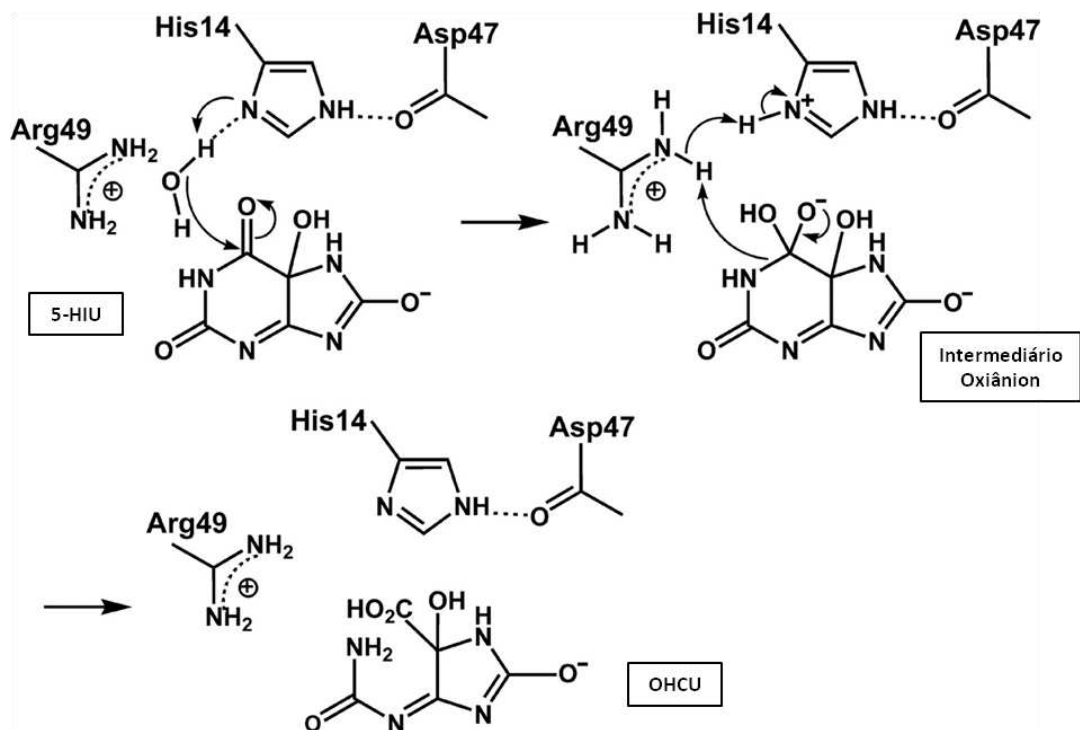


Figura 13 - Mecanismo proposto para a reação enzimática da HIUase. Linhas pontilhadas indicam ligações de hidrogênio. A mudança do substrato a cada passo é indicada nas caixas de texto. Numeração dos resíduos de acordo com HIUase de *Bacillus subtilis*. Adaptado com permissão de Jung *et al.*, 2006 sob *Creative Commons Attribution License*.

Íons de Zn^{2+} e Br^- também foram encontrados co-cristalizados com a HIUase de *Escherichia coli*, em uma região próxima ao sítio catalítico. Porém, a presença destes íons foi atribuída à composição do meio de cristalização, não à necessidade dos mesmos para a atividade catalítica da enzima, como ocorre com as TTR (Lundberg *et al.*, 2006).

1.5. Engenharia de Proteínas Baseada em Evolução

As abordagens de engenharia de proteínas são utilizadas desde o advento da tecnologia do DNA recombinante (Swint-Kruse, 2016). Hoje em dia, além de experimentos de manipulação em bancada, foram incorporadas técnicas computacionais visando alterar diversas características, desde afinidade a ligantes, atividade enzimática, estabilidade e até mesmo a criação de novas estruturas e funções (Brender *et al.*, 2017). Limitadas no passado apenas a pequenas proteínas e de uma maneira racional⁵, a incorporação de conhecimentos evolutivos possibilitou o estudo de objetos mais complexos, como proteínas maiores, e de uma forma não direcionada (Smith & Hecht, 2011).

De uma perspectiva evolutiva, as proteínas são constituídas de regiões conservadas e pouco variáveis, e regiões não conservadas, passíveis de sofrer menor pressão seletiva (Gu *et al.*, 2021). A identificação destas regiões permite então a proposição de alterações pontuais que podem levar a alterações significativas nas proteínas de estudo, como a potencialização ou criação de funções, aumento da afinidade ao substrato, ou até um melhor entendimento das relações evolutivas entre as mesmas.

1.5.1. Interconvertendo TTR e HIUase

Entre a HIUase e a TTR, uma série de mutações teriam sido capazes de alterar a função das proteínas, passando de uma enzima do metabolismo do ácido úrico para uma transportadora de hormônios tireoidianos. Estas mutações não alteraram o arcabouço principal das mesmas, mas afetaram profundamente regiões específicas e que são altamente conservadas nas proteínas, como as cavidades catalíticas centrais (Zanotti *et al.*, 2006). Desta forma, estas duas proteínas são um bom modelo para o estudo de evolução de proteínas, suas funções e estruturas (Richardson & Cody, 2009).

⁵ Aqui definida como de maneira planejada e direcionada, utilizando conhecimentos prévios de estrutura e função.

Estudos de engenharia de proteínas envolvendo HIUase e TTR já foram feitos, como a geração de um duplo mutante Y116T/I16A⁶ de HIUase capaz de realizar a ligação de hormônios tireoidianos na enzima (Cendron *et al.*, 2011). Porém, pequenas alterações como essa não seriam capazes de alterar todas as outras modificações ocorridas, principalmente para adaptação da estrutura da proteína em relação à mudança de substratos, e um número maior de mutações seria necessário para otimização das mesmas (Romero & Arnold, 2009; Cendron *et al.*, 2011).

Em 2021, Carrijo de Oliveira e colaboradores recriaram as proteínas ancestrais de momentos chave nos processos evolutivos das subfamílias de HIUase e TTR, por meio de uma reconstrução filogenética de sequências. Os resultados obtidos indicam que mesmo um parálogo resultante da duplicação gênica de uma HIUase em cordados já seria capaz de se tornar um transportador funcional de tiroxina (Carrijo de Oliveira *et al.*, 2021).

Sendo assim, neste trabalho buscamos por meio da engenharia de proteínas com base em evolução molecular *in silico*, realizar a interconversão funcional entre as proteínas destas duas subfamílias. Para alcançar os nossos objetivos, duas sequências foram geradas, aqui denominadas mutantes, onde uma sequência representativa de uma família foi substituída em posições específicas por resíduos conservados da outra. Deste modo, espera-se compreender melhor o processo de neofuncionalização destes parálogos, além de futuramente aplicar a mesma abordagem para outros grupos de proteínas com importâncias biotecnológicas.

⁶ Nomenclatura de HIUase de Zebra fish.

2. OBJETIVOS

2.1. Objetivo geral

O objetivo deste trabalho foi realizar a interconversão funcional entre as HIUase e as TTR por meio da geração de sequências mutantes, aplicando a abordagem de engenharia de proteínas com base em evolução molecular *in silico*.

2.2. Objetivos específicos

- Comparar filogeneticamente e estruturalmente as proteínas representativas das subfamílias de HIUase e TTR;
- Gerar sequências mutantes de HIUase e TTR com base na conservação de resíduos co-evoluídos entre as proteínas;
- Gerar, *in silico*, os modelos tridimensionais das proteínas mutantes;
- Analisar os volumes e geometrias das cavidades das proteínas mutantes;
- Analisar os potenciais eletrostáticos e a distribuição de cargas das proteínas mutantes;
- Analisar os resíduos expostos ao solvente e as áreas de interface entre os monômeros das proteínas mutantes.

3. MATERIAL E MÉTODOS

3.1. Análises Filogenéticas

Para realização das análises filogenéticas das proteínas pertencentes à família das TTR, foi utilizado o servidor SMART, no modo Normal (Simple Modular Architecture Research Tool. Letunic & Bork, 2018; Letunic, Khedkar & Bork, 2021). Foram feitas pesquisas utilizando tanto a sequência da transtirretina humana (Uniprot ID: P02766) como entrada, quanto por palavras-chave (“*transthyretin*”). As proteínas identificadas no SMART nrdb (*non-redundant database*) foram curadas manualmente, sendo selecionadas entidades representativas de todos os reinos da vida, de acordo com critérios de importância biológica e econômica.

De posse de proteínas representativas da família, as mesmas foram exportadas em formato Newick, onde árvores enraizadas e com seus vértices rotulados são representadas utilizando combinações aninhadas de parênteses e vírgulas. O arquivo foi então utilizado como entrada no servidor iTOL (Interactive Tree of Life. Letunic & Bork, 2021), para visualização e representação dos cladogramas (Figura 14a, ao final dos materiais e métodos). Foram mantidas as nomenclaturas e divisões cladísticas utilizadas pelo servidor SMART.

3.2. Alinhamentos de Sequências

Todos os alinhamentos de sequências descritos neste trabalho foram realizados utilizando ferramentas disponibilizadas pelo Instituto Europeu de Bioinformática (EMBL-EBI), sendo o algoritmo EMBOSS Needle para alinhamentos par-a-par, e Clustal Omega para alinhamentos múltiplos (Sievers *et al.*, 2011; Madeira *et al.*, 2022). Os alinhamentos foram tratados e visualizados por meio da ferramenta BoxShade v3.31, disponível para linha de comando em Linux (Figura 14a).

3.3. Estruturas Secundárias

A análise das estruturas secundárias das proteínas estudadas foi feita por meio dos servidores web PSIPRED Workbench (McGuffin, Bryson & Jones, 2000; Buchan & Jones, 2019) e PDBsum (Laskowski, 2007). O primeiro teve como entrada as sequências das proteínas deste estudo, utilizando os seguintes métodos de predição: PSIPRED 4.0 (predição de estruturas secundárias), DISOPRED3 (predição de desordem) e DomPred (predição de domínios proteicos). Já o segundo utilizou como entrada os arquivos PDB contendo as

estruturas tridimensionais das proteínas, buscando identificar a geometria de elementos estruturais, como folhas- β , α -hélices e *random coils* (Figura 14a).

3.4. Análise de Cavidades e Potenciais Eletrostáticos

A análise das cavidades foi feita por meio do *plugin* CAVER 3.0.3 para PyMOL (Chovancova *et al.*, 2012). As coordenadas x , y e z iniciais foram selecionadas de acordo com os resíduos identificados, em busca e revisão manual na literatura, como catalíticos e/ou estruturais em relação às cavidades (Figura 14a).

O volume das cavidades foi determinado utilizando o programa CAVER Analyst 2.0 (Jurcik *et al.*, 2018), com a sonda maior constituindo uma esfera de 3Å, e a menor uma esfera de 1.4Å (equivalente à uma molécula de água). Os volumes identificados como pertencentes à cavidade foram refinados no próprio programa, a fim de aumentar a precisão das medidas. O volume dos ligantes foi calculado por meio do servidor online Molecular Volume Calculator (Jayaram *et al.*, 2012).

Os potenciais eletrostáticos das estruturas foram calculados utilizando o pacote de programas online APBS-PDB2PQR (Jurrus *et al.*, 2018). Primeiramente o módulo PDB2PQR foi empregado para preparar as estruturas para cálculos de solvatação, por meio da adição de estados de protonação, parâmetros de carga e raios atômicos utilizando campos de força. Os arquivos PQR gerados foram então submetidos ao módulo APBS (do inglês, Adaptive Poisson-Boltzmann Solver), onde as equações da eletrostática contínua foram resolvidas com base nas contribuições dos solventes e íons calculados na estrutura. Todos os parâmetros padrão foram utilizados (Figura 14c).

Após os cálculos, os arquivos contendo os potenciais (extensão “-pot.dx”) foram analisados por meio do programa ChimeraX, onde foram sobrepostos com as estruturas de referência e as cavidades geradas no passo anterior. As estruturas foram então visualizadas em modo de superfície, e coloridas de acordo com os valores de potencial.

3.5. Geração das Sequências Mutantes

A geração das sequências das mutantes foi realizada por nosso grupo de pesquisa, pelo então doutorando Lucas Carrijo de Oliveira, sob a orientação do Prof. Lucas Bleicher (Figura 14b). Um alinhamento múltiplo entre as famílias das TTR e HIUase foi obtido a partir do PFAM (ID PF00576: *HIUase/Transthyretin Family*. Mistry *et al.*, 2021). A matriz BLOSUM (Henikoff e Henikoff, 1992) foi utilizada para atribuir peso às sequências, visando mitigar os efeitos da presença de eventuais sequências redundantes e calcular a frequência dos

aminoácidos de maneira mais assertiva. Considerando que a soma dos pesos de todas as sequências é sempre igual a 1, para calcularmos a frequência de um determinado resíduo basta somar os pesos das sequências que apresentam um aminoácido em sua determinada posição.

A seguir, uma árvore filogenética de TTR e HIUase de vertebrados foi utilizada para definir os dois subconjuntos de sequências, e para cada aminoácido, suas frequências relativas foram calculadas dividindo-se as frequências pela soma dos pesos do subconjunto específico, de acordo com a fórmula:

$$P_{aa} | \text{subfamília} = P_{aa} / P_{\text{subfamília}}$$

Para cada subfamília, os resíduos com frequência relativa superior a 90% foram selecionados. Estes mesmos resíduos foram denominados “globalmente conservados” quando presentes em ambas as subfamílias, e “localmente conservados” quando presentes em apenas uma.

As sequências das mutantes foram então geradas a partir de uma proteína de uma subfamília substituída, em posições específicas, por resíduos localmente conservados da outra. Dessa forma, foram geradas duas sequências, sendo elas: a) MutTTR: sequência de transtirretina de *Homo sapiens* (Uniprot: P02766) com substituições correspondentes a resíduos específicos da subfamília de HIUase e; b) MutHIU: sequência de HIUase de *Herbaspirillum seropedicae* (Uniprot: D8J1G5) com substituições correspondentes a resíduos específicos da subfamília de TTR.

Posições marginalmente conservadas foram resolvidas por meio do agrupamento de aminoácidos de acordo com suas propriedades físico-químicas (Quadro 1).

Quadro 1 - Classificação dos resíduos de aminoácidos de acordo com suas características químicas.

Característica	Aminoácidos
Amídicos	Asn, Gln
Alifáticos	Gly, Ala, Val, Leu, Ile
Básicos	His, Lys, Arg
Contendo Hidroxila	Ser, Thr, Tyr
Não Polares	Phe, Gly, Val, Leu, Ala, Ile, Pro, Met, Trp
Polares	Tyr, Ser, Asn, Thr, Gln, Cys

Muito Hidrofóbicos	Leu, Ile, Phe, Trp, Val, Met
Hidrofílicos	Arg, Lys, Asn, Gln, Pro, Asp
Carregados Positivamente	Lys, Arg
Carregados Negativamente	Asp, Glu
Muito Pequenos	Gly, Ala, Ser
Pequenos	Cys, Asp, Pro, Asn, Thr
Médios	Glu, Val, Gln, His
Grandes	Met, Ile, Leu, Lys, Arg
Aromáticos	Phe, Tyr, Trp
Similares	Asn, Asp, Gln, Glu

Fonte: Próprio autor.

Grupos com frequência superior a 90% para uma determinada posição foram utilizados para substituir o resíduo equivalente na outra subfamília, quando esta difere do aminoácido original. Dentre os aminoácidos contidos no mesmo agrupamento, foi selecionado aquele que fosse mais comum dentro de sua respectiva subfamília.

3.6. Modelagem Molecular

Para a geração de modelos estruturais tridimensionais das mutantes, foram utilizadas diversas abordagens computacionais, a fim de evitar quaisquer limitações individuais das técnicas escolhidas (Figura 14c).

3.6.1. *Modeller*

A primeira abordagem computacional utilizada foi a de modelagem por homologia (ou comparativa). Para isto, foi escolhido o programa *Modeller* (V9.23, Webb & Sali, 2016). As estruturas de referência foram obtidas por meio de uma análise BLAST das sequências (BLASTp utilizando banco de dados do *Protein Data Bank*). Foram selecionadas as estruturas depositadas com menor *e-value*, maior porcentagem de identidade, e que haviam sido depositadas contendo um tetrâmero como unidade biológica.

Para a geração dos modelos de ambas mutantes, foram utilizadas as estruturas de referência, sendo o alinhamento das estruturas primárias realizadas pelo script *alignment.salign()*⁷. O alinhamento resultante foi posteriormente refinado manualmente por meio de alinhamentos locais e *threading*, a fim de potencializar as regiões conservadas na estrutura primária das proteínas mutantes.

Sendo assim, os arquivos de entrada continham três sequências, onde a maior parte do alinhamento correspondia à relação entre a SpMut de interesse com sua sequência de origem (MutHIU com HbsHIU, e MutTTR com HsTTR, respectivamente). Entretanto, nas regiões onde houve troca de resíduos, a relação se dava entre a mutante e a sequência de “destino” (MutHIU com HsTTR, e MutTTR com HbsHIU, respectivamente). Com isso, uma maior cobertura e fidelidade foi obtida nos alinhamentos, evitando *gaps* e regiões não relacionadas.

Os modelos foram gerados por meio do script *model-multiple()*⁸, modificado com o módulo *special_restraints()*⁹ para restrição da movimentação de átomos presentes em múltiplas cadeias simétricas.

Em cada passo, foram geradas 20000 estruturas de cada mutante. A estrutura com menor pontuação DOPE (do inglês, *Discrete Optimized Protein Energy*. Shen & Sali, 2006) foi selecionada para os passos subsequentes. Cada estrutura selecionada foi minimizada de acordo com a energia global por meio do programa Chimera (V1.15, Pettersen *et al.*, 2004), utilizando campo de força AMBER ff14SB. Os parâmetros utilizados para minimização foram: 200 *steepest descent steps*, 0.02 *steepest descent step size* (Å), 20 *conjugate gradient steps*, e 0.02 *conjugate gradient steps* (Å).

As voltas que apresentaram conformações instáveis nas estruturas geradas foram refinados por meio do script *loop_refinement()*¹⁰. Cem modelos de cada volta foram gerados, sendo as quatro cadeias refinadas simultaneamente, a fim de evitar variações estocásticas.

3.6.2. AlphaFold

Para modelagem tridimensional sem a utilização de estruturas homólogas, foi utilizado o programa *AlphaFold* (Jumper *et al.*, 2021), que utiliza ferramentas de inteligência artificial, redes neurais e aprendizado profundo, e possui como entrada apenas a sequência a ser modelada. Foram empregadas duas metodologias diferentes utilizando os recursos do *AlphaFold*.

⁷ Adaptado de <https://salilab.org/modeller/9v1/manual/node270.html>

⁸ Adaptado de <https://salilab.org/modeller/manual/node21.html>

⁹ Adaptado de <https://salilab.org/modeller/examples/automodel/model-multichain-sym.py>

¹⁰ Adaptado de <https://salilab.org/modeller/manual/node35.html>

3.6.2.1. Instalação Local

O servidor para análises computacionais de biomoléculas para Ressonância Magnética Nuclear - NMRbox (Maciejewski *et al.*, 2017) dispõe de diversas máquinas virtuais baseadas na nuvem com alto poder computacional e de uso gratuito para a comunidade científica.

Foi então utilizada a versão 2.2.0 do *AlphaFold*, capaz de modelar cadeias oligoméricas (Evans *et al.*, 2022), e instalado de forma local nas máquinas disponibilizadas pelo NMRbox. As especificações da máquina escolhida foram: Selenium (selenium.nmrbox.org), contendo 20 núcleos de processamento e 40 threads Intel Gold 6230, 376 GB RAM, e ~16 GB GPU NVIDIA Tesla T4.

Os arquivos de entrada foram preparados em formato FASTA, contendo apenas a sequência a ser modelada repetida quatro vezes para identificação do estado oligomérico como tetrâmero. Todos os parâmetros padrões de instalação foram mantidos.

Foram gerados cinco modelos originais, e para cada um, cinco predições diferentes e independentes, totalizando 25 estruturas finais. Os modelos foram então relaxados de acordo com o procedimento padrão utilizando campo de força AMBER99SB, até que todas as violações e conflitos estruturais fossem resolvidos (Jumper *et al.*, 2021). Por fim, as estruturas foram ranqueadas em função dos valores de pLDDT (do inglês, *Predicted Local Distance Difference Test*. Tunyasuvunakool *et al.*, 2021), baseando-se na avaliação resíduo-específica da distância local entre todos os átomos, e como o ambiente de uma estrutura de referência é reproduzido em uma proteína modelo (Mariani *et al.*, 2013). A escala é representada por valores entre 0 e 100, com regiões superiores a 90 sendo consideradas de alta confiança, e abaixo de 50 como confiança muito baixa (Tunyasuvunakool *et al.*, 2021). Apenas a estrutura com maior valor de pLDDT em cada experimento foi escolhida para análises posteriores. Além disso, a posição relativa das cadeias da proteína foi verificada por análise visual do gráfico de PAE (do inglês, *Predicted Aligned Error*. Mirdita *et al.*, 2022).

3.6.2.2. ColabFold

Buscando promover maior acessibilidade de acesso ao *AlphaFold*, Mirdita e colaboradores publicaram, em 2022, uma versão do programa aberta, gratuita e disponível em nuvem, denominada *ColabFold*¹¹ (Mirdita *et al.*, 2022). A versão utiliza recursos computacionais do *Google Colaboratory* e uma metodologia mais rápida de busca por sequências homólogas, denominada MMseqs2 (Steinegger & Söding, 2017), ao invés do padrão *Jackhammer*, além de também possuir suporte à modelagem de cadeias multiméricas.

¹¹ Disponível em: <https://github.com/sokrypton/ColabFold>

As sequências a serem modeladas foram inseridas em formato FASTA, repetida quatro vezes (separadas por “:”) para identificação do estado oligomérico como tetrâmero. Foram gerados cinco modelos por sequência, e de maneira similar à versão local, todos foram minimizados utilizando campo de força AMBER99SB, até que todas as violações e conflitos estruturais fossem resolvidos. As estruturas foram ranqueadas de acordo com o valor de pLDDT, e apenas o modelo com maior valor foi escolhido para análises posteriores. O gráfico de PAE também foi utilizado para inspeção visual da orientação das cadeias da proteína.

3.7. Validação Estrutural

Todas as estruturas selecionadas nos passos anteriores foram analisadas e validadas por meio da plataforma *online* SAVES v6.0 (Figura 14c). Foram utilizados os seguintes programas: (a) ERRAT (Colovos & Yeates, 1993) para a verificação da estabilidade e confiança estatística da conformação dos resíduos na estrutura tridimensional gerada, por meio das interações átomo-átomo não ligadas e comparação com estruturas de referência; (b) PROCHECK (Laskowski *et al.*, 1993) para identificação de resíduos em posições desfavoráveis e/ou não permitidas em relação à seus ângulos diedros ϕ (ϕ) e ψ (Ψ), por meio de gráficos de Ramachandran (Ramachandran, Ramakrishnan & Sasisekharan, 1963); (c) VERIFY3D (Eisenberg, Lüthy & Bowie, 1997), para determinação da compatibilidade das estruturas tridimensionais analisadas em relação à suas estruturas primárias (3D/1D), por meio de uma análise resíduo-específica de suas localizações e ambientes químicos, e comparação com estruturas controle; (d) PROVE (Pontius, Richelle & Wodak, 1996), para análise da qualidade por meio da variação de volumes atômicos em relação à estruturas controle (Z-score).

Além disso, a ferramenta *online* VADAR (do inglês, *Volume Area Dihedral Angle Reporter*. Willard *et al.*, 2003) foi utilizada a fim de analisar parâmetros como a área acessível ao solvente (ASA, do inglês *Accessible Surface Area*), levando à determinação de resíduos expostos ou enterrados na estrutura tridimensional.

3.8. Visualização e Geração de Figuras

Todas as estruturas tridimensionais geradas e analisadas foram visualizadas por meio dos programas PyMOL (Schrödinger, LLC.), Chimera e ChimeraX (Pettersen *et al.*, 2021).

Os gráficos e análises estatísticas foram realizadas no programa RStudio (V. 2021.09.0), utilizando as bibliotecas “*ggplot2*” e “*dplyr*”.

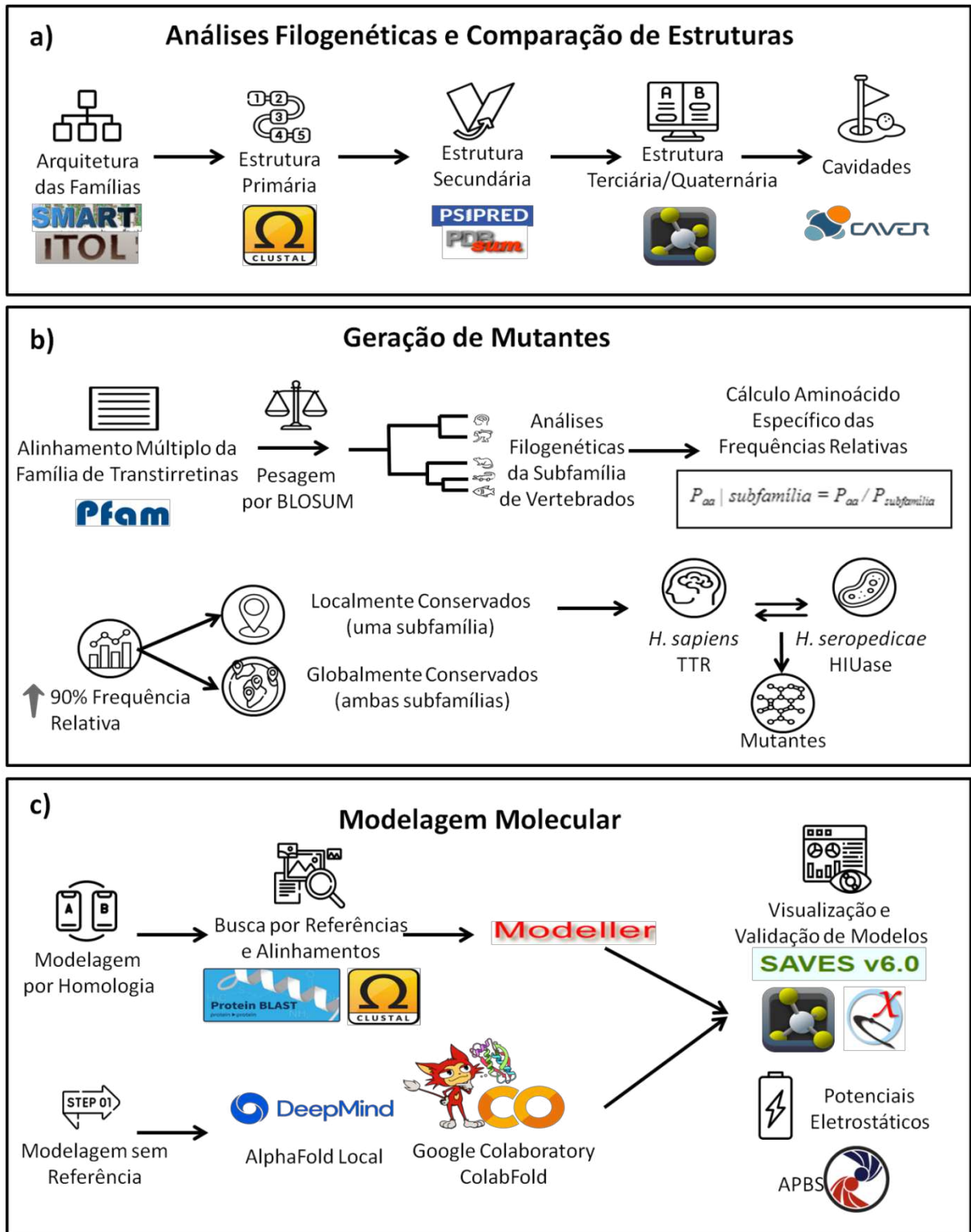


Figura 14 - Estratégia experimental. Neste painel são evidenciadas as etapas de análises filogenéticas e comparação de estruturas (a), geração das sequências de proteínas mutantes (b), e modelagem molecular (c).

4. RESULTADOS E DISCUSSÃO

4.1. Etapa 1 - Proteínas de Referência

4.1.1. Análises Filogenéticas da Prevalência de HIUase e TTR em Organismos Vivos

Com o intuito de demonstrar a intrincada história evolutiva das transtirretinas e da 5-hidroxiisourato hidrolases, inicialmente foi feita uma análise da filogenia das proteínas utilizadas neste estudo (Figura 15).

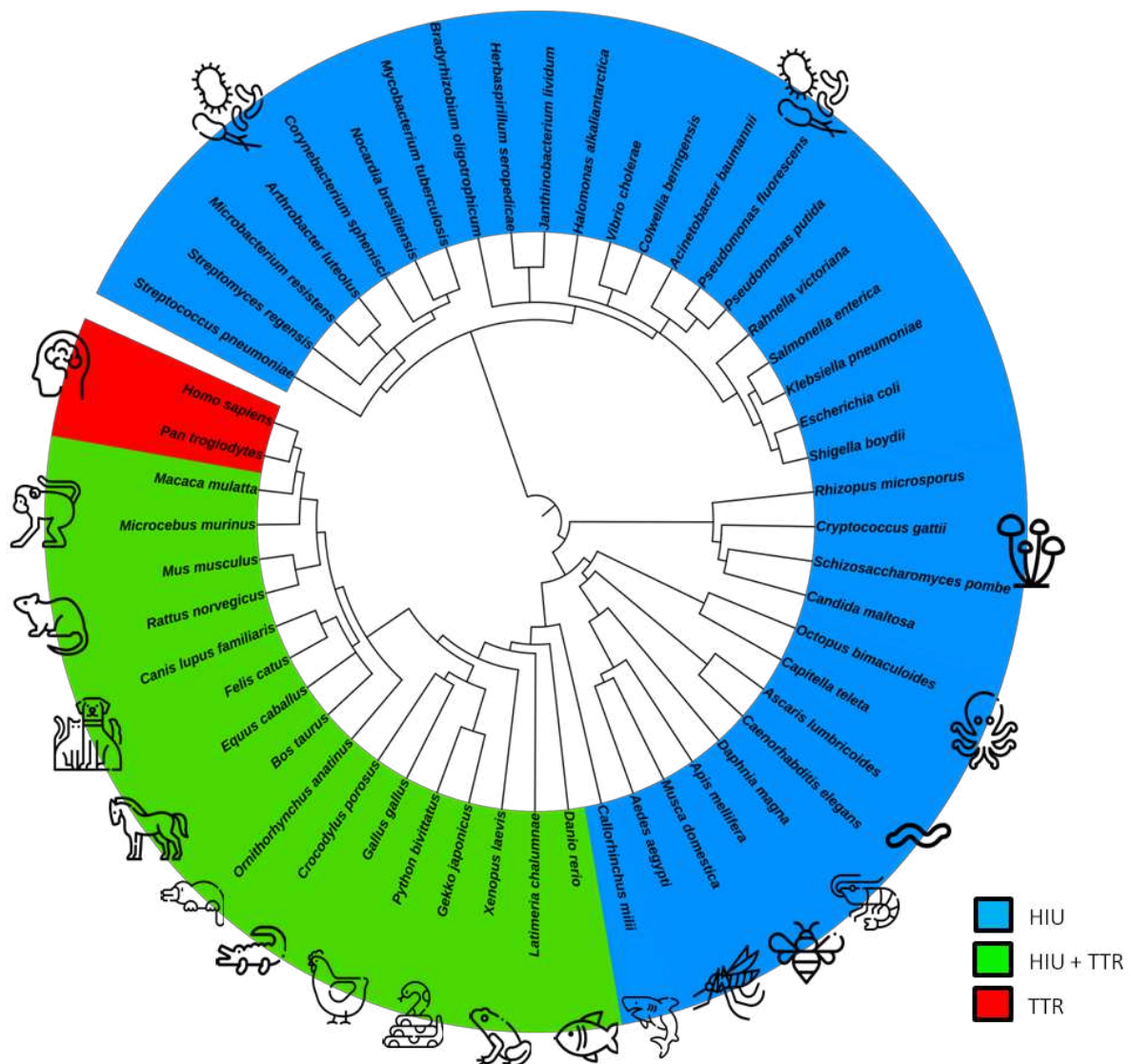


Figura 15 - Árvore filogenética circular da prevalência das subfamílias de Transtirretinas e 5-Hidroxiisourato Hidrolases em organismos vivos. Azul: organismos que possuem apenas HIUase. Verde: organismos que possuem ambas as enzimas. Vermelho: organismos que possuem apenas TTR, com perda do gene da HIUase. Grupos que possuem representantes com e sem perda secundária da HIUase, como répteis e aves, são mostrados apenas em verde para melhorar a visualização. Apenas algumas espécies representativas de cada táxon são mostradas.

As HIUase, ubíquas em grande parte dos táxons de organismos vivos, se mostram presentes desde grupos como archaeas (não representadas na imagem) até vertebrados mais derivados, como mamíferos. Tanto bactérias gram-positivas quanto gram-negativas também possuem exclusivamente essas enzimas, além de fungos, cefalópodes, artrópodes, platelmintos, peixes cartilagosos, entre outros. Na divergência evolutiva entre peixes cartilagosos e ósseos, posterior duplicação gênica e surgimento das transtirretinas ancestrais, ambas as proteínas passaram a coexistir em cordados, porém com funções biológicas diferentes (Zanotti *et al.*, 2006. Ramazzina *et al.*, 2006).

Apesar de alguns grupos de organismos possuírem uma perda secundária da enzima uricase de forma independente, como algumas aves e répteis, apenas hominídeos possuem uma deterioração¹² completa dos genes e proteínas da via de degradação de ácido úrico (Keilin, 1959; Keebaugh & Thomas, 2010). Isso se deve ao fato de que a perda de um único gene (como no caso da Uricase em hominídeos), pode levar à posterior falta de pressão seletiva nos subsequentes de uma via metabólica (Tanaka, Tateno & Gojobori, 2005).

Sendo assim, quatro gêneros (*Homo*, *Pan*, *Gorilla* e *Pongo*) de primatas, compondo o táxon dos hominídeos, apresentam apenas o gene da transtirretina, que é expressa nestes organismos (Figura 16).

A ausência da via de degradação de ácido úrico em hominídeos levou a um subsequente acúmulo deste composto nos indivíduos, culminando em diversas consequências sistêmicas. A gota é um exemplo, causada pelo acúmulo de cristais de urato monossódico em tecidos, e cujo surgimento foi considerado um “acidente evolutivo”, já que humanos não são capazes de degradá-los (Ragab, Elshahaly & Bardin, 2017; Keilin, 1959). Existem também consequências potencialmente benéficas, como os efeitos estimulatórios e antioxidantes do ácido úrico, que podem ter levado a uma maior capacidade cognitiva, redução das taxas de câncer e aumento da sobrevivência de indivíduos (Keebaugh & Thomas, 2010; Cendron *et al.*, 2011; Ames *et al.*, 1981).

¹² Deterioração é aqui definida como um acúmulo de mutações deletérias causadas por falta de pressão seletiva.

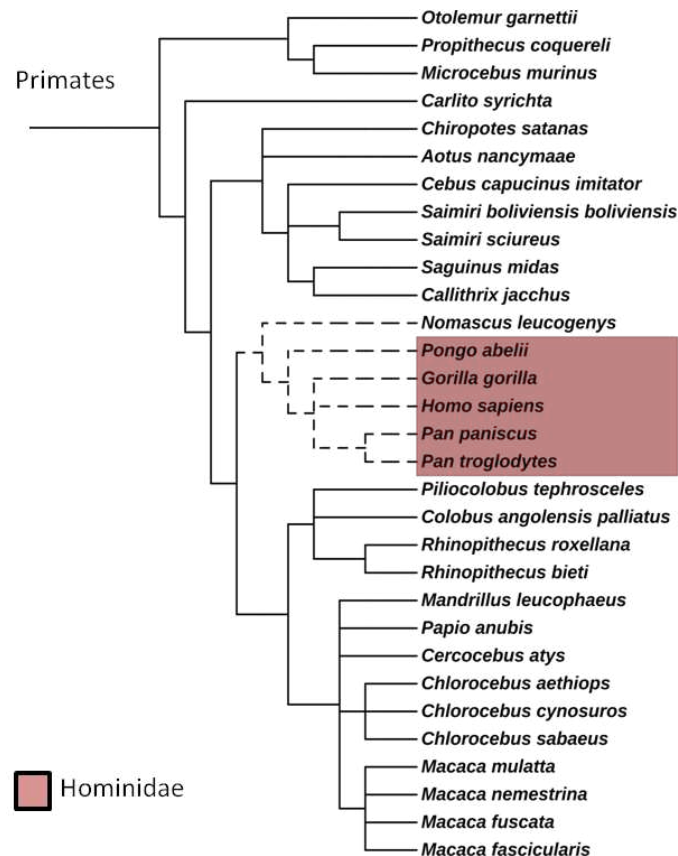


Figura 16 - Árvore filogenética retangular da prevalência de HIUase em primatas (*primates*). Todos os organismos neste grupo possuem ambas TTR e HIUase, com exceção dos quatro gêneros que compõem os hominídeos (destacados em vermelho). Apenas algumas espécies representativas de cada táxon são mostradas.

4.1.2. Comparações Estruturais entre HIUase e TTR

Com o intuito de avaliar semelhanças e divergências entre as TTR e HIUase, nesta seção foram feitas comparações, de forma sequencial, entre as estruturas primária, secundária, terciária e quaternária de TTR e HIUase. Para isto, foram escolhidas proteínas de *Herbaspirillum seropedicae* para HIUase (Uniprot ID: D8J1G5; Lee *et al.*, 2005), e *Homo sapiens* (Uniprot ID: P02766) para TTR. Um mapeamento completo dos resíduos importantes para a estrutura e função de HIUase e TTR de diferentes organismos está disponível no apêndice I, a fim de auxiliar na compreensão e no entendimento da discussão a ser apresentada nesta dissertação.

4.1.2.1. Estrutura Primária

As estruturas primárias de proteínas fazem referência às suas respectivas sequências de resíduos de aminoácidos. Ao alinhar duas ou mais estruturas primárias de diferentes proteínas, padrões de conservação ou divergência de resíduos podem ser identificados, e a identidade sequencial pode ser calculada. É importante frisar que proteínas com identidades

baixas entre si não necessariamente refletem em diferenças tridimensionais nas estruturas, ou mesmo funções distintas.

Primeiramente, foram alinhadas as sequências de HIUase de *Herbaspirillum seropedicae*, uma β -proteobactéria diazotrófica endofítica obrigatória¹³ (Baldani *et al.*, 1996), e HIUase de *Zebrafish* (*Danio rerio*), um peixe teleosteo de água doce bastante utilizado como organismo modelo, cuja enzima foi uma das primeiras a ter sua estrutura tridimensional elucidada (Bailone *et al.*, 2020; Zanotti *et al.*, 2006)(Figura 17).

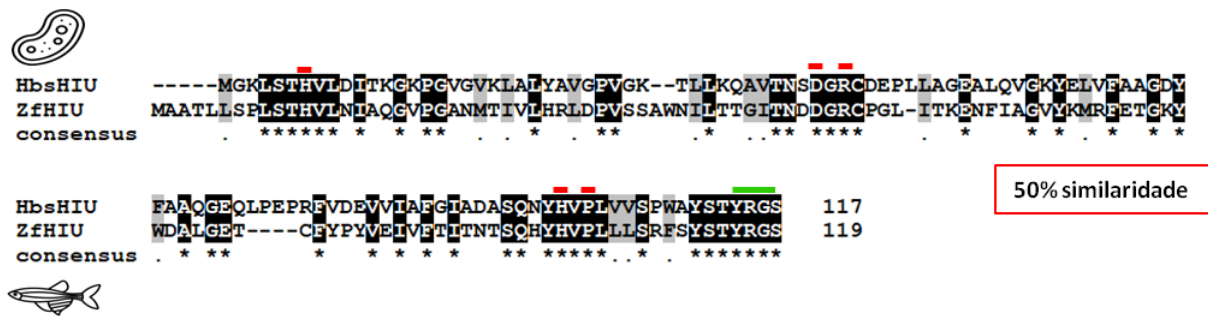


Figura 17 - Alinhamento de estruturas primárias entre HIUase de *Herbaspirillum seropedicae* e *Danio rerio* (*Zebrafish*). Resíduos conservados são mostrados com contorno em preto (linha de consenso: asterisco), e resíduos com características químicas semelhantes são mostrados com contorno em cinza (linha de consenso: ponto). Os principais resíduos catalíticos estão destacados com uma linha vermelha, e o tetrapeptídeo C-terminal está destacado com uma linha verde.

É possível perceber na figura 17, que mesmo com uma similaridade de 50% entre as duas sequências, regiões importantes para a estrutura e função da proteína são mantidas, como o tetrapeptídeo C-terminal (YRGS), além dos resíduos catalíticos.

Logo após, foi realizado o alinhamento das estruturas primárias de HIUase de *Herbaspirillum seropedicae* e TTR de *Homo sapiens*, que devido à importância médica desta proteína em humanos, possui uma literatura científica vasta e as estruturas mais bem elucidadas e estudadas (Liz *et al.*, 2020) (Figura 18).

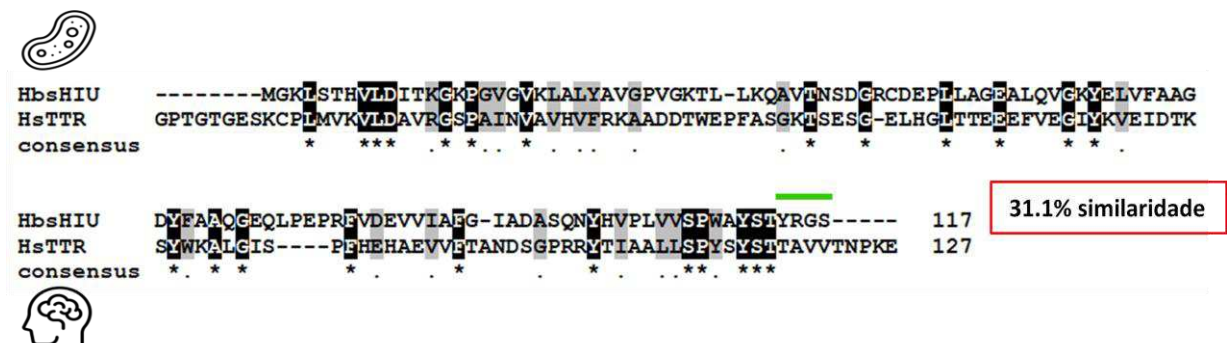


Figura 18 - Alinhamento de estruturas primárias entre HIUase de *Herbaspirillum seropedicae* e TTR de *Homo sapiens*. Resíduos conservados são mostrados com contorno em preto (linha de consenso: asterisco), e

¹³ Bactérias diazotróficas endofíticas obrigatórias são aquelas capazes de fixar nitrogênio atmosférico no interior de tecidos vegetais sem causar dano às mesmas, e que sobrevivem apenas se estiverem nestas condições (Melloni, 2004).

resíduos com características químicas semelhantes são mostrados com contorno em cinza (linha de consenso: ponto). O tetrapeptídeo C-terminal, ubíquo em HIUase, é destacado com uma linha verde

Ao se comparar os alinhamentos, é possível perceber as diferenças sequenciais entre as duas proteínas, incluindo a perda do tetrapeptídeo de resíduos C-terminais em TTR, assim como a presença de um número maior de resíduos na proteína madura de humanos. Além disso, com uma identidade de apenas 31.1%, a troca e/ou perda de resíduos catalíticos é visível, embora o arcabouço da estrutura tridimensional seja mantido.

4.1.2.2. Estrutura Secundária

As estruturas secundárias de proteínas fazem referência às organizações locais formadas pela estabilização das sequências primárias por meio de ligações de hidrogênio, e são comumente divididas em α -hélices, folhas- β e espirais aleatórias (Lyu *et al.*, 2021). Estas três conformações servem como base para a organização tridimensional das estruturas, e apresentam informações estruturais importantes.

Com o intuito de comparar as estruturas secundárias de HIUase e TTR, foi feita a análise das estruturas secundárias de HIUase de *Herbaspirillum seropedicae* e TTR de *Homo sapiens* (Figura x).

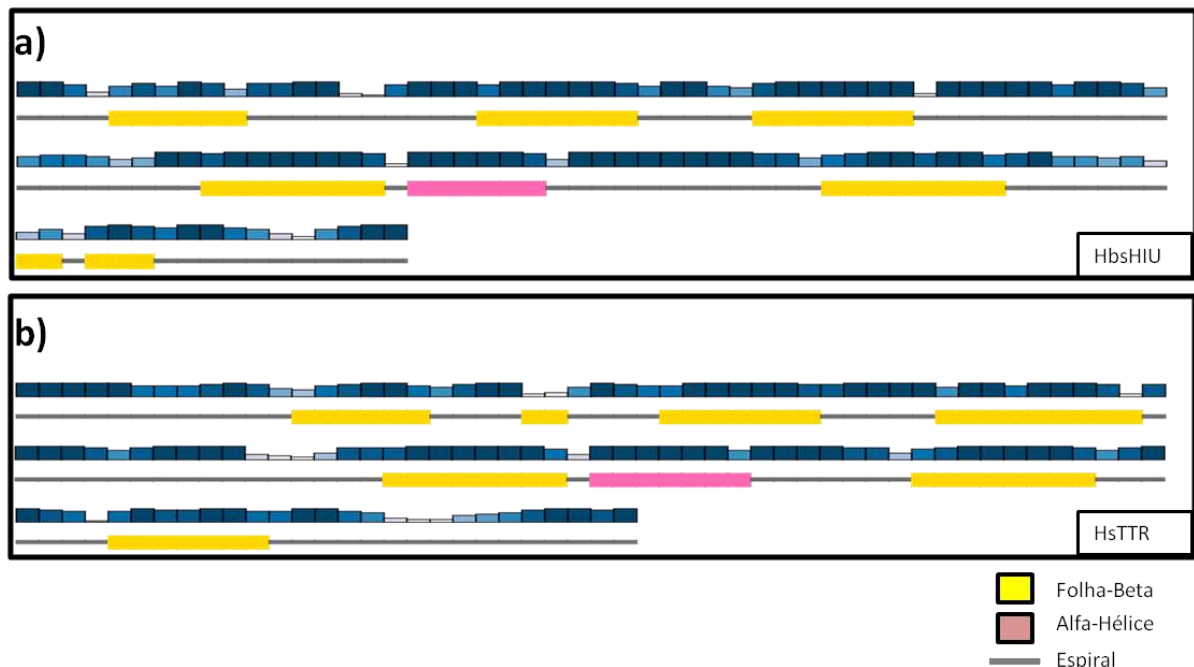


Figura 19 - Estruturas secundárias de HIUase e TTR. A) HIUase de *Herbaspirillum seropedicae*. B) TTR de *Homo sapiens*. São evidenciadas as três conformações clássicas, com folhas- β marcadas em amarelo, α -hélices em rosa, e espirais como linhas cinza. As barras azuis representam a confiança de predição do programa PSIPred.

É possível identificar a presença do mesmo padrão de organização nas duas proteínas, com uma presença majoritária de folhas- β , e apenas uma α -hélice, como já é descrito para a família de transtirretinas (Blake *et al.*, 1978).

Comparando o padrão de dobramento e organização das estruturas secundárias, a mesma organização é obtida, incluindo as conexões entre as folhas- β das proteínas (Figura 20).

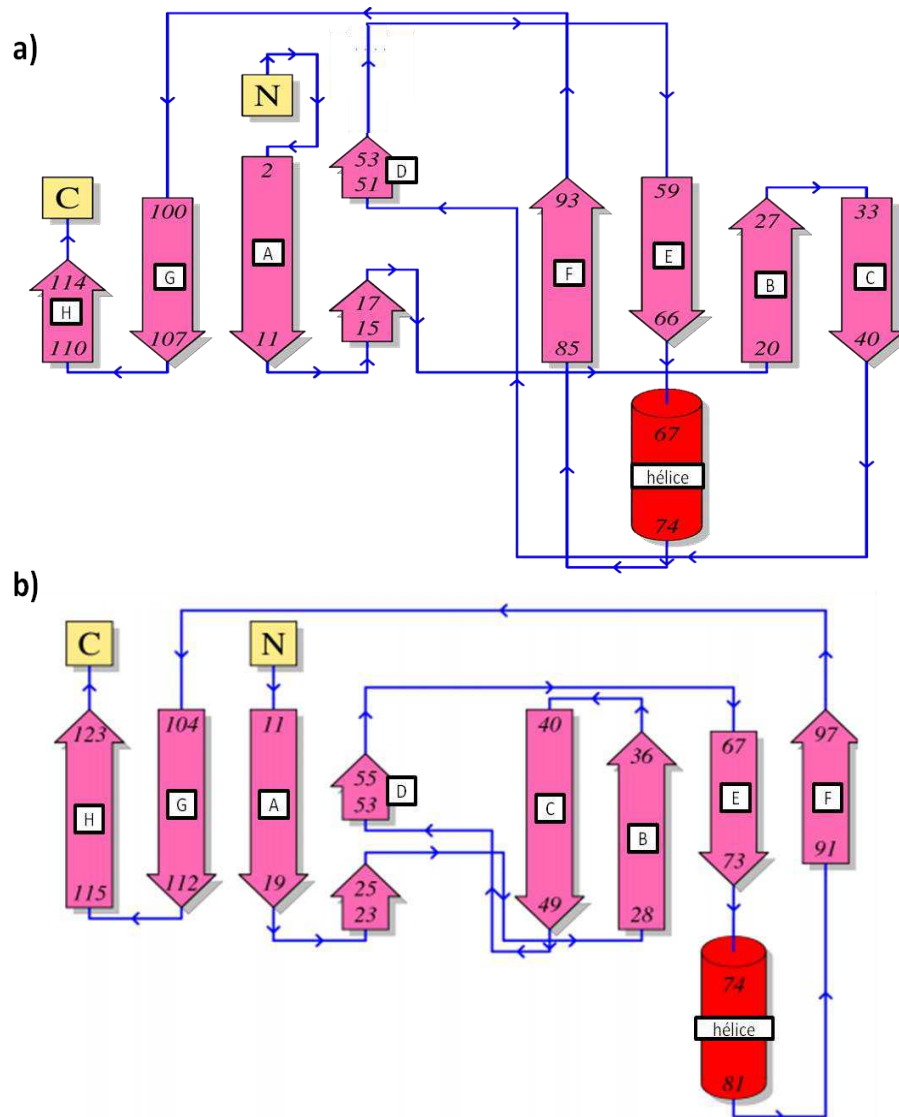


Figura 20 - Padrão de organização das estruturas secundárias. A) HIUase de *Herbaspirillum seropedicae*; B) TTR de *Homo sapiens*. São evidenciadas folhas- β em setas rosa, α -hélices em cilindros vermelhos, e a conexão entre as conformações por setas azuis. Os resíduos que compõem cada conformação também são identificados no interior de suas respectivas estruturas, assim como suas nomenclaturas canônicas.

4.1.2.3. Estrutura Terciária e Quaternária

As estruturas terciárias de proteínas fazem referência à conformação tridimensional e espacial de cadeias polipeptídicas únicas, enquanto estruturas quaternárias são formadas pelas associações entre diferentes cadeias (subunidades).

Ambas as HIUase e as TTR são formadas por quatro subunidades idênticas, levando à formação de um homotetrâmero em suas unidades biológicas e funcionais. Mesmo tendo funções biológicas bem distintas, o padrão de organização tridimensional das mesmas é conservado (Figura 21).

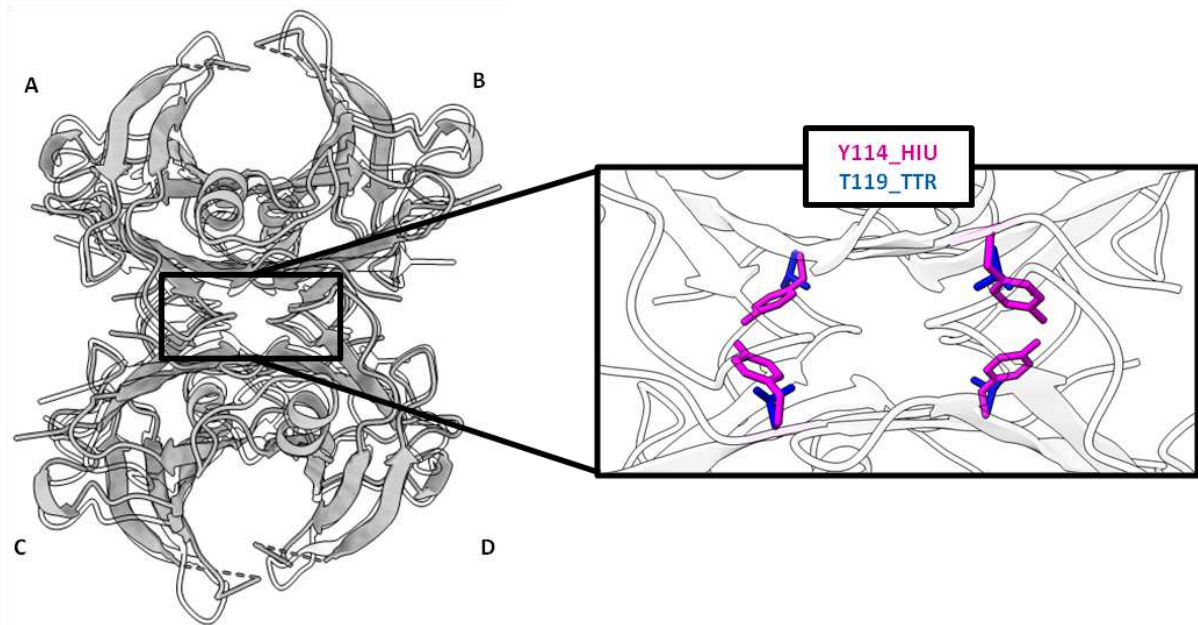


Figura 21 - Sobreposição das estruturas quaternárias de HIUase e TTR. À esquerda é evidenciada a alta semelhança estrutural entre HIUase de *Herbaspirillum seropedicae* (cinza) e TTR de *Homo sapiens* (branco). Letras maiúsculas referem-se a cada uma das quatro cadeias do homotetrâmero. À direita, é mostrado um aumento na região da cavidade das duas proteínas sobrepostas, com um dos resíduos mais importantes para a diferenciação das mesmas em evidência (Y114 em HIUase e T119 em TTR).

Com um RMSD (desvio quadrático médio das posições atômicas, do inglês, *Root Mean Square Deviation of Atomic Positions*) de 2.21Å entre as quatro cadeias polipeptídicas, fica explícita uma elevada conservação estrutural entre as duas proteínas, mesmo com a baixa identidade sequencial (31.1%). As principais diferenças se dão a nível atômico, onde substituições de resíduos cruciais para o funcionamento enzimático das HIUase ocorreram para acomodar os hormônios tireoidianos, transportados pelas TTR.

Talvez a mais importante substituição tenha ocorrido entre um resíduo de tirosina em HIUase (Tyr114) para uma treonina em TTR (Thr119). O resíduo de tirosina, mais volumoso que a treonina, bloqueia completamente uma parte da cavidade em HIUase, além de estar envolvido com a reação enzimática que essa proteína desempenha (Cendron *et al.*, 2011; Jung

et al., 2006). Em TTR, a presença do resíduo Thr119 libera a porção da cavidade anteriormente inacessível, e conseqüentemente, aumenta consideravelmente o volume acessível ao solvente. Este resíduo também está associado ao bolsão de ligação a halógenos 3 (HBP3), um dos sítios onde os hormônios tireoidianos se ligam (Tomar *et al.*, 2012).

Ao analisar o espaço topológico e os volumes das cavidades, é possível observar a diferença entre as duas proteínas, especialmente na região onde os resíduos Tyr114 e Thr119 se encontram em HIUase e TTR, respectivamente (Figura 22).

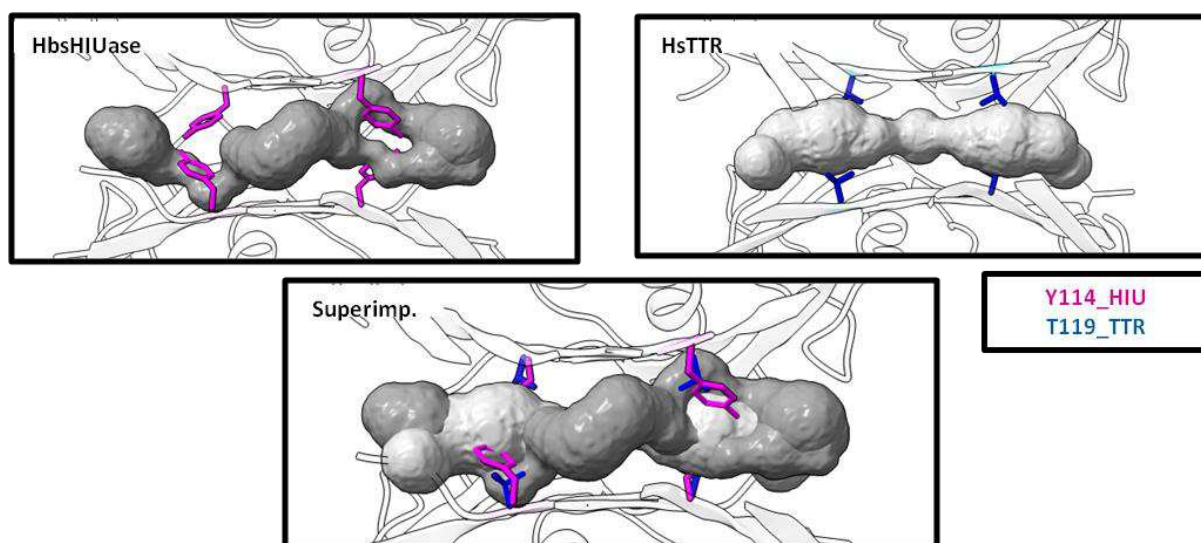


Figura 22 - Comparação entre as geometrias das cavidades de HIUase e TTR. Painéis superiores: à esquerda, HIUase de *Herbaspirillum seropedicae*, com o resíduo Y114 evidenciado em rosa e o espaço ocupado pela cavidade em cinza escuro. À direita, TTR de *Homo sapiens*, com o resíduo T119 evidenciado em azul e o espaço ocupado pela cavidade em cinza claro. Painel inferior: sobreposição das duas geometrias.

Em HIUase, o resíduo Tyr114 causa um “estrangulamento” da cavidade, levando a uma terminação precoce do espaço disponível. Já em TTR, o resíduo Thr119, muito menos volumoso, não é capaz de exercer a mesma função, e a cavidade se estende até o centro da proteína, aumentando a profundidade e espaço disponível (Jung *et al.*, 2006).

Estas variações na geometria e no volume da cavidade explicam, pelo menos em parte, a seletividade de cada proteína por seu ligante, já que as moléculas de T3 e T4 possuem um volume bem superior à molécula de 5-HIU, substrato para a enzima HIUase (Tabela 1).

Tabela 2 - Comparação entre os volumes das cavidades de HIUase e TTR.

Proteína	Volume da Cavidade (Å ³)	Ligante	Volume do Ligante (Å ³)
HbsHIU	1274,8	5-HIU	119,7
HsTTR	1737,9	T3 e T4	296,7 - 321,6

Fonte: próprio autor.

Nota: são apresentados os volumes das cavidades das proteínas, seus respectivos ligantes e volumes dos ligantes.

Analisando as estruturas já resolvidas experimentalmente e depositadas com os respectivos ligantes no PDB, a associação entre a diferença na geometria e volume das cavidades fica ainda mais evidente. É possível observar nitidamente a diferença estrutural dos ligantes, assim como a propagação da cavidade até uma porção mais central na proteína, no caso da TTR (Figura 23).

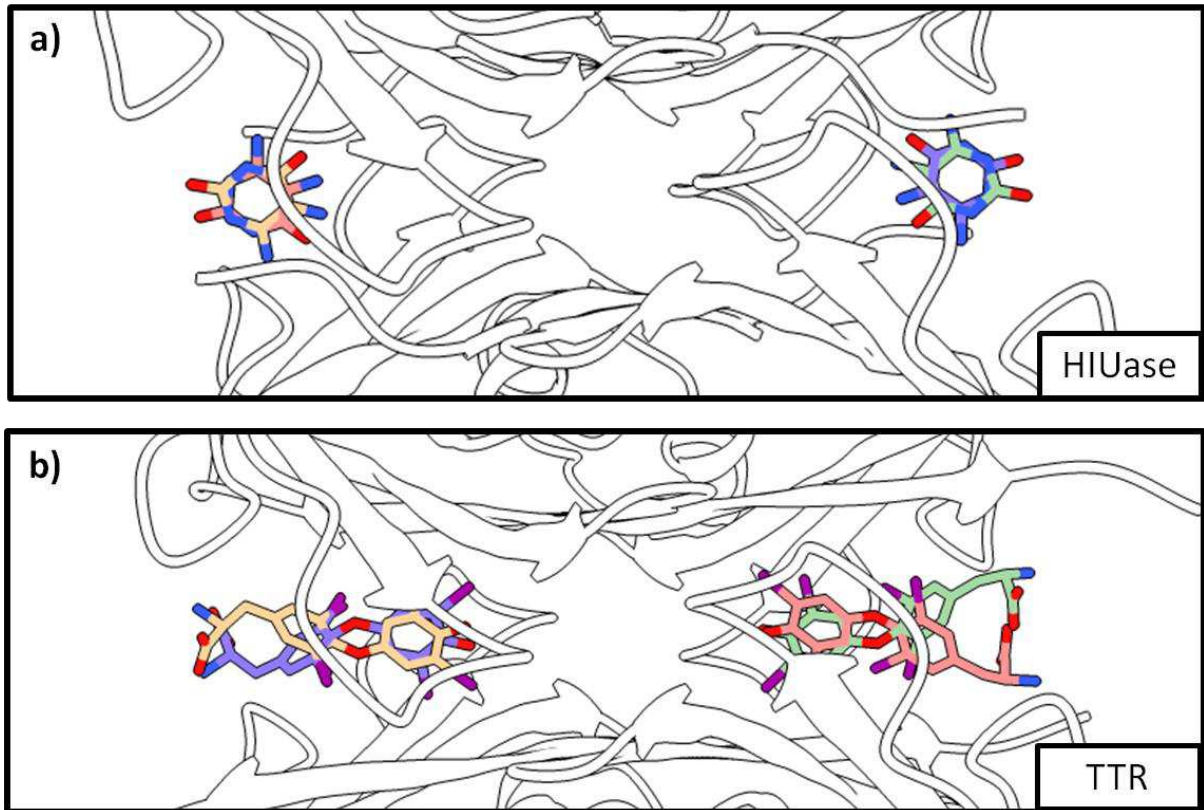


Figura 23 - Comparação entre estruturas resolvidas experimentalmente na presença de ligantes para HIUase e TTR. A) HIUase de *Bacillus subtilis* complexada com 5,6-diaminouracila, um substrato alternativo e mais estável ao 5-HIU (PDB ID: 2H0J; Jung *et al.*, 2006; Raychaudhuri & Tipton, 2003). B) TTR de *Homo sapiens* complexada com Tiroxina (T4) (PDB ID: 2ROX; Wojtczak *et al.*, 1996). Os ligantes são mostrados como estruturas coloridas no interior das cavidades das proteínas.

Além das diferenças de geometria e volume nas cavidades, existem também diferenças significativas nas características químicas de alguns resíduos importantes para a atividade das proteínas. Ao olhar apenas aqueles que compõem as cavidades, é possível perceber que todos aqueles localizados na porção central possuem características que se mantêm, evidenciando a função majoritariamente estrutural que eles exercem (Figura 24).

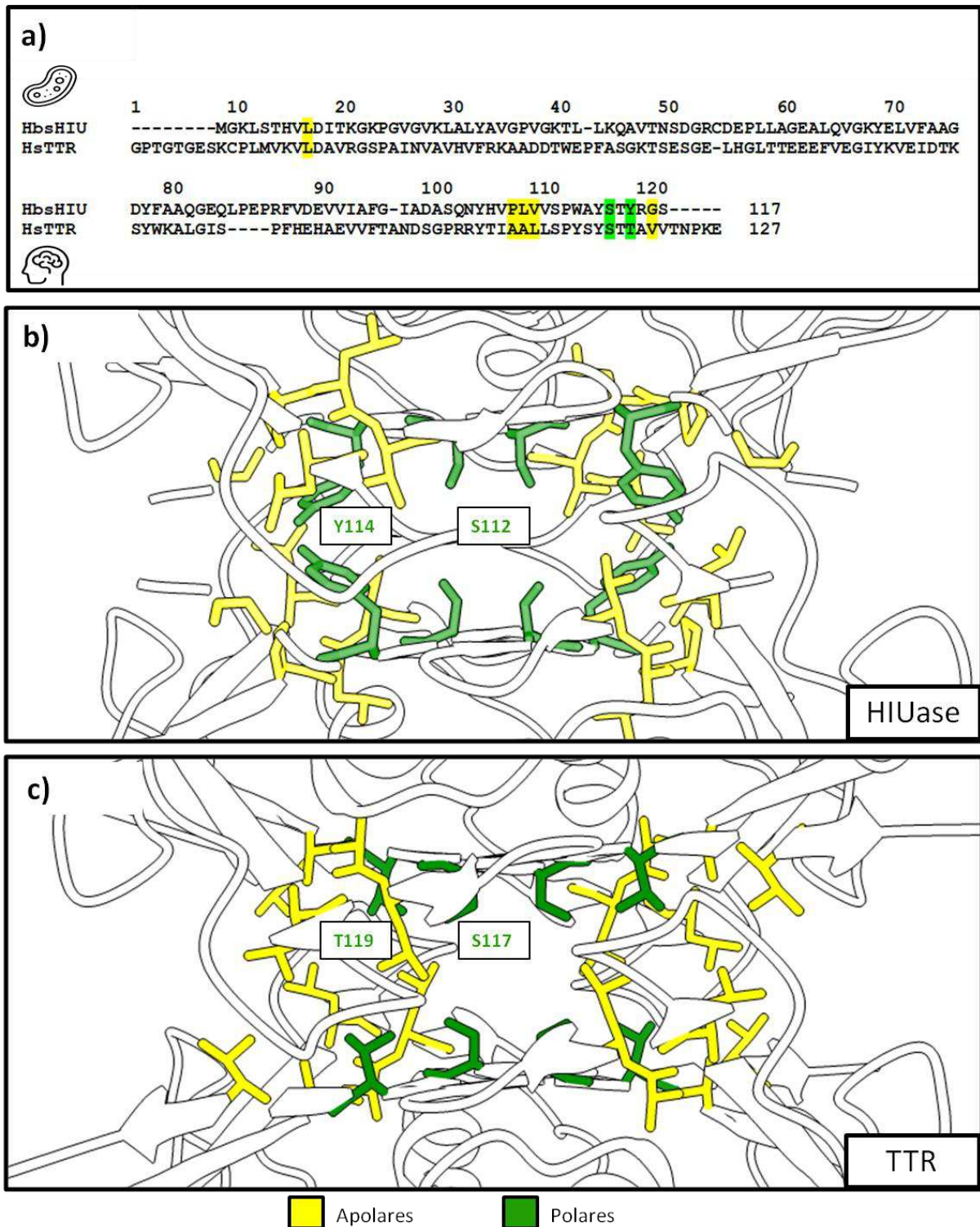


Figura 24 - Comparação entre características químicas de resíduos localizados no centro das cavidades de HIUase e TTR. A) Alinhamento entre HIUase de *Herbaspirillum seropedicae* e TTR de *Homo sapiens*, destacando apenas resíduos que compõem a porção central da cavidade de ambas proteínas. A numeração dos resíduos se refere à TTR. B e C) Representação tridimensional dos resíduos destacados em (A) para HIUase (B) e TTR (C). Resíduos polares (em verde) estão identificados na figura, incluindo o par Y114/T119, já discutido anteriormente.

Porém, ao analisar apenas as porções externas das cavidades, ocorrem a troca de um resíduo ácido para um resíduo polar (Asp43 para Ser52, utilizando as nomenclaturas de HbsHIU e HsTTR, respectivamente), um básico para um ácido (Arg45 para Glu54), um básico para um polar (His101 e Thr106) e um polar para um apolar (Ser117 para Val121). Estas mudanças nas características químicas de resíduos chave, envolvendo especialmente inversão de cargas e polaridade, pode ter afetado profundamente a função destas proteínas, já que estes resíduos estão diretamente relacionados à função enzimática e de ligação aos hormônios (Lundberg *et al.*, 2006; Cendron *et al.*, 2011; Tomar *et al.*, 2012) (Figura 25).

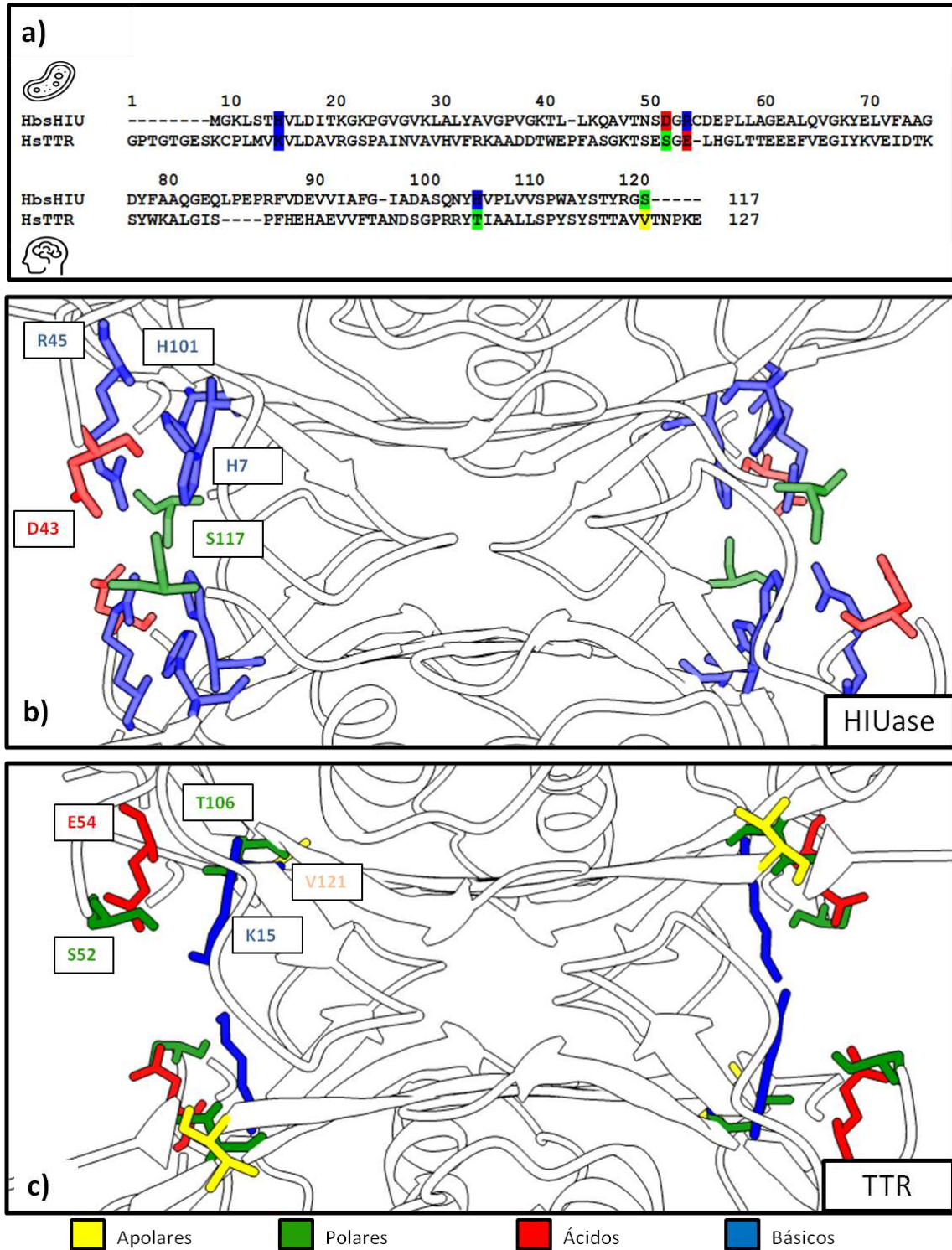


Figura 25 - Comparação entre características químicas de resíduos localizados nas extremidades das cavidades de HIUase e TTR. A) Alinhamento entre HIUase de *Herbaspirillum seropedicae* e TTR de *Homo sapiens*, destacando apenas resíduos que compõem as porções externas das cavidades de ambas proteínas. A numeração dos resíduos se refere à TTR. B e C) Representação tridimensional dos resíduos destacados em (A) para HIUase (B) e TTR (C). Os resíduos cujas características químicas diferem entre as duas proteínas, como discutido no texto, estão identificados na porção superior esquerda (cadeia A das proteínas). Os resíduos estão coloridos de acordo com o seguinte esquema: Amarelos: apolares; Verdes: polares; Azuis: básicos; Vermelhos: ácidos. Para um mapeamento completo dos resíduos em diferentes proteínas, ver apêndice I.

As diferenças nas porções externas das cavidades, em especial em relação às cargas líquidas, se refletem no potencial eletrostático no interior das mesmas, que em HIUase é predominantemente negativo, enquanto em TTR, de maneira oposta, se encontra majoritariamente positivo (Figura 26).

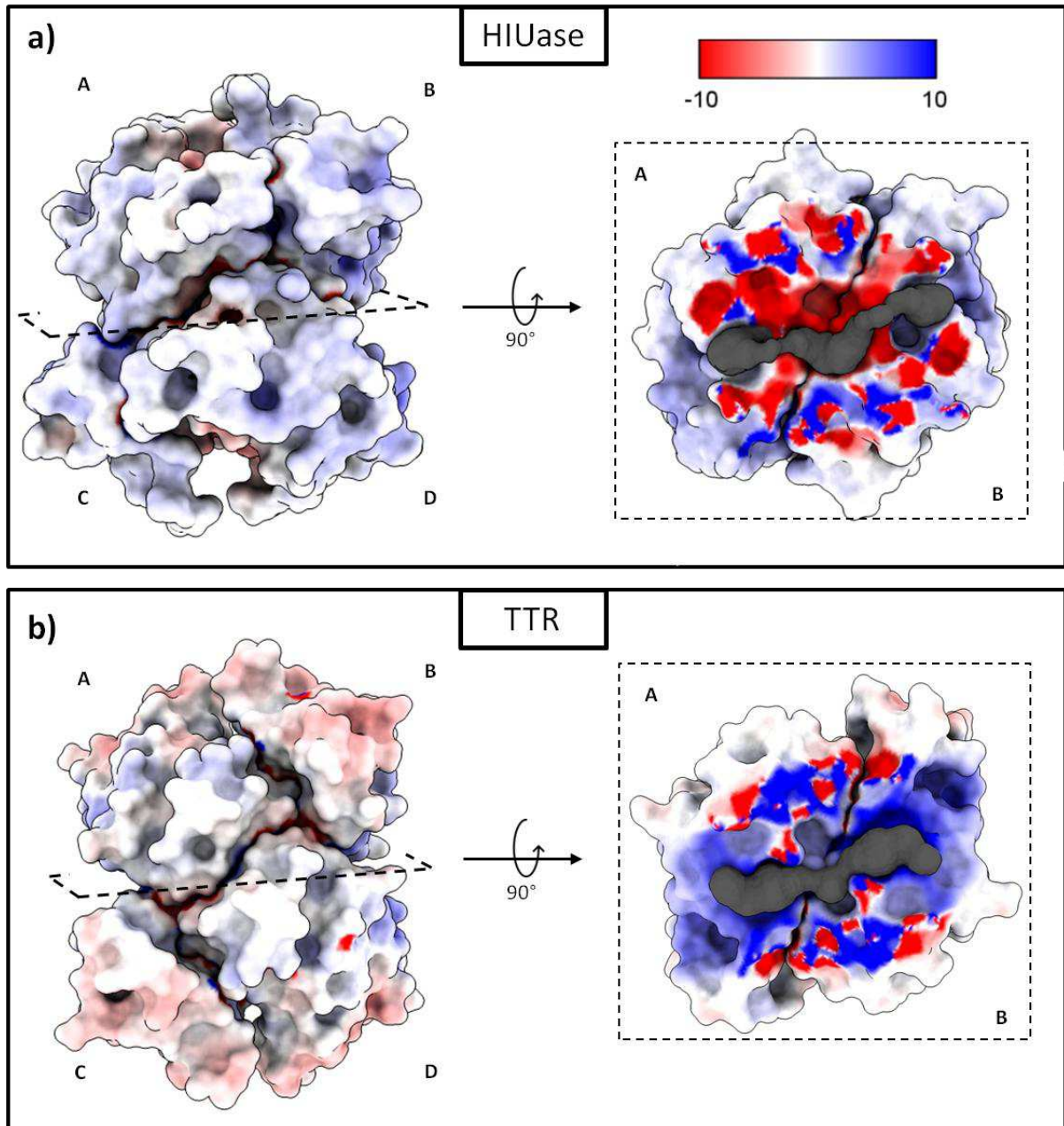


Figura 26 - Potenciais eletrostáticos de HIUase e Transtirretinas. À esquerda, são mostradas as estruturas quaternárias homotetraméricas de HIUase de *Herbaspirillum seropedicae* (a) e TTR de *Homo sapiens* (b), com cada monômero sendo identificado por uma letra maiúscula. À direita, em uma rotação de 90° do dímero superior (A-B), são mostrados detalhes das cavidades internas das proteínas, evidenciando os potenciais inversos entre as duas. Em cinza escuro é mostrada a geometria da cavidade, para melhor interpretação.

4.2. Etapa 2 - Proteínas Mutantes

4.2.1. Definição e Geração das Mutantes

Continuando uma das hipóteses desenvolvidas no trabalho do Dr. Lucas Carrijo de Oliveira em sua tese de doutorado (defendida em 2021 pelo Programa Interunidades de Pós-Graduação em Bioinformática da UFMG, em processo de depósito), foram propostas, *in silico*, duas sequências de proteínas com base nos resíduos localmente conservados nas famílias de HIUase e TTR, como descrito no material e métodos.

Estas sequências seriam capazes de mimetizar as funções principais de uma dessas proteínas, tendo seus arcabouços mantidos com base na outra. Em suma, a hipótese é que proteínas evolutivamente relacionadas seriam capazes de terem suas funções interconvertidas pela troca “cega” (ou não supervisionada) de alguns resíduos específicos e conservados em uma das subfamílias, e vice-versa.

A escolha dessas duas subfamílias de proteínas pode levar a utilização de TTR engenheiradas à realizar a função de degradação de metabólitos da via do ácido úrico, o que serviria no tratamento de gota, além de outras doenças reumáticas. Uma segunda implicação seria o melhor entendimento dos processos amiloidogênicos destas proteínas, já que HIUase não possuem a mesma propensão à agregação que TTR, mesmo compartilhando grande parte do arcabouço estrutural.

Utilizando como base as proteínas HIUase de *Herbaspirillum seropedicae* e TTR de *Homo sapiens*, as sequências propostas foram aqui denominadas: MutanteHIUase (MutHIU), para aquela cujos resíduos localmente conservados de uma TTR foram substituídos no arcabouço de uma HIUase; e MutanteTTR (MutTTR), para aquela, de forma análoga, cujos resíduos localmente conservados de uma HIUase foram substituídos no arcabouço de uma TTR (Figura 27). Dessa forma, a MutHIU, hipoteticamente, seria capaz de realizar as funções de uma TTR, e a MutTTR seria capaz de realizar as funções de uma HIUase.

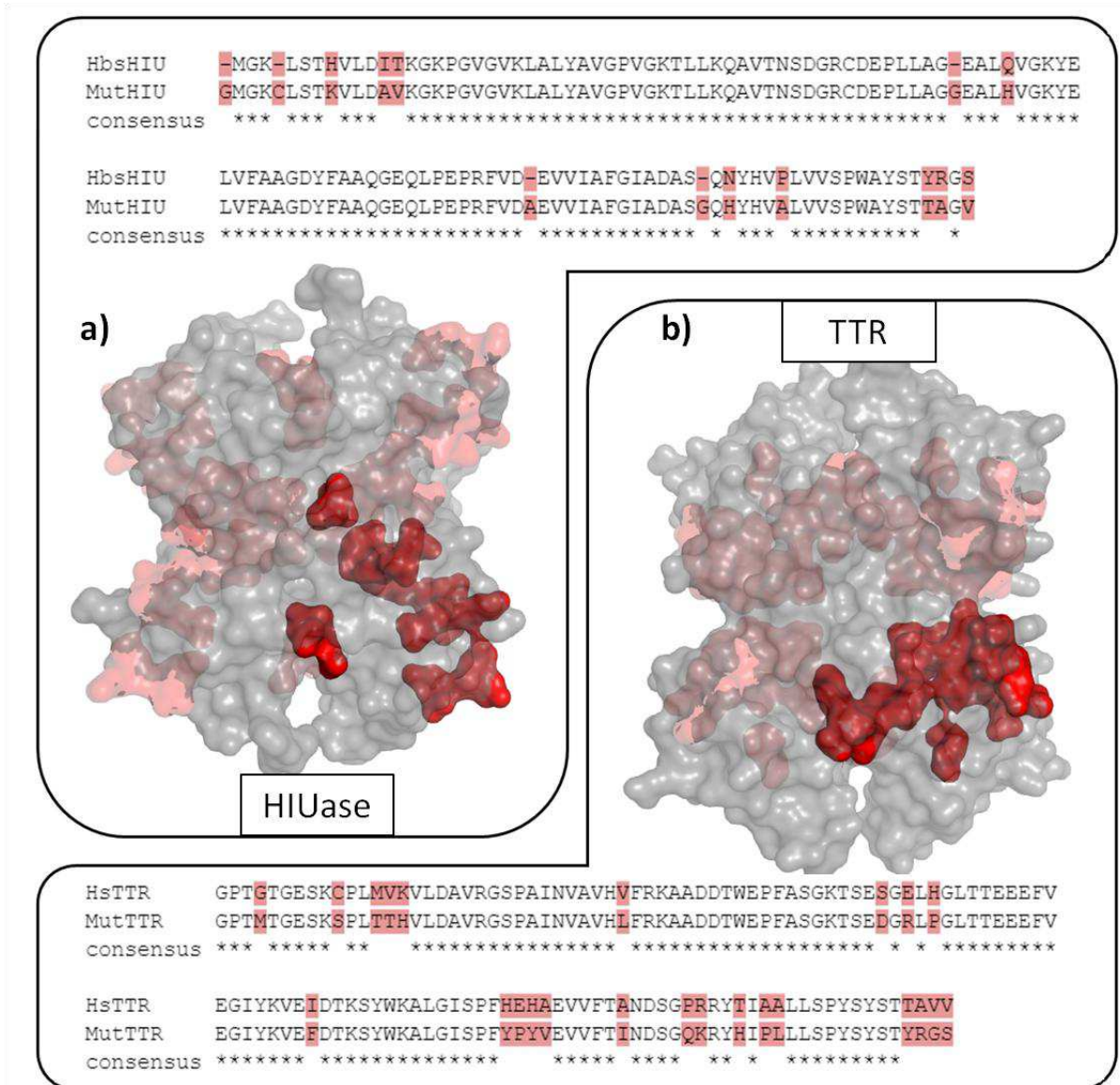


Figura 27 - Comparação entre os mutantes e suas respectivas proteínas de origem. A) Mutante de HIUase. Esta proteína foi gerada pela troca de resíduos localmente conservados de TTR de *Homo sapiens* na sequência de HIUase de *Herbaspirillum seropedicae*. B) Mutante de TTR. De forma análoga à mutante de HIUase, esta proteína foi gerada pela troca de resíduos localmente conservados de HIUase de *Herbaspirillum seropedicae* na sequência de TTR de *Homo sapiens*. Os resíduos trocados estão identificados em vermelho nos alinhamentos de sequências, sendo que nenhum outro resíduo foi alterado. As figuras centrais se referem aos resíduos alterados mapeados nas estruturas tridimensionais de suas proteínas correspondentes (HIUase de *Herbaspirillum seropedicae* para MutHIU (a); e TTR de *Homo sapiens* para TTR (b), respectivamente). Por questões de visualização, apenas uma cadeia do homotetrâmero foi destacada em vermelho escuro, enquanto as outras foram mostradas em vermelho claro.

Como pode ser observado, apenas cerca de 15% dos resíduos de cada proteína foram alterados (14/117 para MutHIU e 22/127 para MutTTR, em cada monômero), sendo estes identificados como localmente conservados em cada proteína. Para a MutHIU, as substituições de resíduos se deram tanto na porção da cavidade da proteína quanto em porções mais externas (Figura 27a, imagem central). Já para a MutTTR, todos os resíduos substituídos ficaram restritos à região da cavidade (Figura 27b, imagem central).

Os resíduos importantes para a função única de cada proteína foram identificados pela abordagem evolutiva empregada neste estudo, como o tetrâmero C-terminal exclusivo de HIUase (Eneqvist *et al.*, 2003), presente também na MutTTR e perdido na MutHIU. A lista completa de resíduos: (a) catalíticos e/ou componentes das cavidades; (b) estruturais; (c) que desempenham outras funções, e suas respectivas nomenclaturas nas proteínas mutantes, assim como suas características químicas e relações com as proteínas de origem, podem ser consultadas na tabela disponível no apêndice II.

Dos sete resíduos conservados ubiquamente entre as HIUase e as TTR¹⁴ (Leu9/17, Gly14/22, Val20/28, Leu24/Val32, Gly44/53, Ile89/Val93, Val102/Ile107), apenas o resíduo Val32 foi trocado para Leu32 na MutTTR. Porém, mesmo com a troca do resíduo identificado como localmente conservado, a característica química apolar foi mantida, o que corrobora as variações entre os outros resíduos. É interessante notar que, embora estes resíduos conservados nas proteínas não terem sido substituídos, outros resíduos em suas vizinhanças diretas sofreram modificações, evidenciando a importância dos sete para a estrutura tridimensional das proteínas, assim como o poder de resolução resíduo-específico da técnica aqui empregada.

4.2.2. Geração de Modelos Tridimensionais

A elucidação e o estudo de estruturas tridimensionais de proteínas é uma das principais formas de se entender os processos desempenhados pelas mesmas em uma escala molecular (Santhoshkumar & Yusuf, 2020). Muitas informações podem ser obtidas, como a organização das cadeias monoméricas e das cadeias laterais de resíduos de aminoácidos, além das estruturas serem passíveis de experimentos de dinâmica e docagem molecular.

De posse das sequências das proteínas mutantes, o próximo passo foi realizar a modelagem molecular das mesmas, a fim de obter os modelos tridimensionais que pudessem ser utilizados para estudos posteriores. Os modelos estruturais gerados *in silico* são úteis especialmente quando as estruturas das proteínas ainda não foram determinadas experimentalmente, ou quando um número grande de modelos com pequenas modificações é necessário.

¹⁴ Utilizando as nomenclaturas de *Herbaspirillum seropedicae* e *Homo sapiens*, respectivamente. Adaptado de Zanotti *et al.*, 2006, que utiliza nomenclatura de HIUase de *Danio rerio*.

4.2.3. Validações Estruturais

Validações estruturais de estruturas tridimensionais de proteínas geradas *in silico* podem ser realizadas tanto por testes estatísticos, comparações com outras estruturas já identificadas como de boa qualidade, ou inspeção visual. Os principais parâmetros e análises de validação estão apresentados na Tabela 3.

Tabela 3 - Parâmetros de validação e análises estatísticas dos modelos gerados *in silico*.

Proteína	Obtenção	ERRAT	Regiões Ramachandran				Verify3D	Z-score
			Mais Favorecidas	Adicionalmente Permitidas	Generosamente Permitidas	Não Permitidas		
HbsHIU	Experimental	92,69	86,8%	13,2%	0%	0%	100%	-5.31
HsTTR		92,74	92,5%	7,5%	0%	0%	95,02%	-7.04
MutHIU	Modeller	92,43	87,4%	11,6%	1%	0%	100%	-6.1
	AlphaFold	88,13	92,2%	7,8%	0%	0%	100%	-6.74
	ColabFold	87,59	89,4%	10,6%	0%	0%	100%	-6.58
MutTTR	Modeller	76,57	85%	14,7%	0,3%	0%	100%	-6.09
	AlphaFold	84,39	88,4%	11,6%	0%	0%	98,89%	-6.4
	ColabFold	88,01	87,9%	12,1%	0%	0%	98%	-6.21

Fonte: próprio autor.

Nota: apenas algoritmos selecionados da plataforma SAVES foram escolhidos e mostrados na tabela. As duas proteínas de referência (HbsHIU e HsTTR) foram obtidas experimentalmente (ID PDB para HsTTR: 1F41 (Hörnberg et al., 2000); HbsHIU: não depositada). As duas proteínas mutantes foram geradas *in silico* por meio de técnicas de modelagem por homologia (*Modeller*) e modelagem sem referência (*AlphaFold* e *ColabFold*).

A seguir, duas métricas importantes para a validação de estruturas, ERRAT e Ramachandran, foram analisadas individualmente. Os resultados para cada modelo das proteínas mutantes são mostrados na figura 28, assim como as validações para as proteínas usadas como moldes na modelagem por homologia.

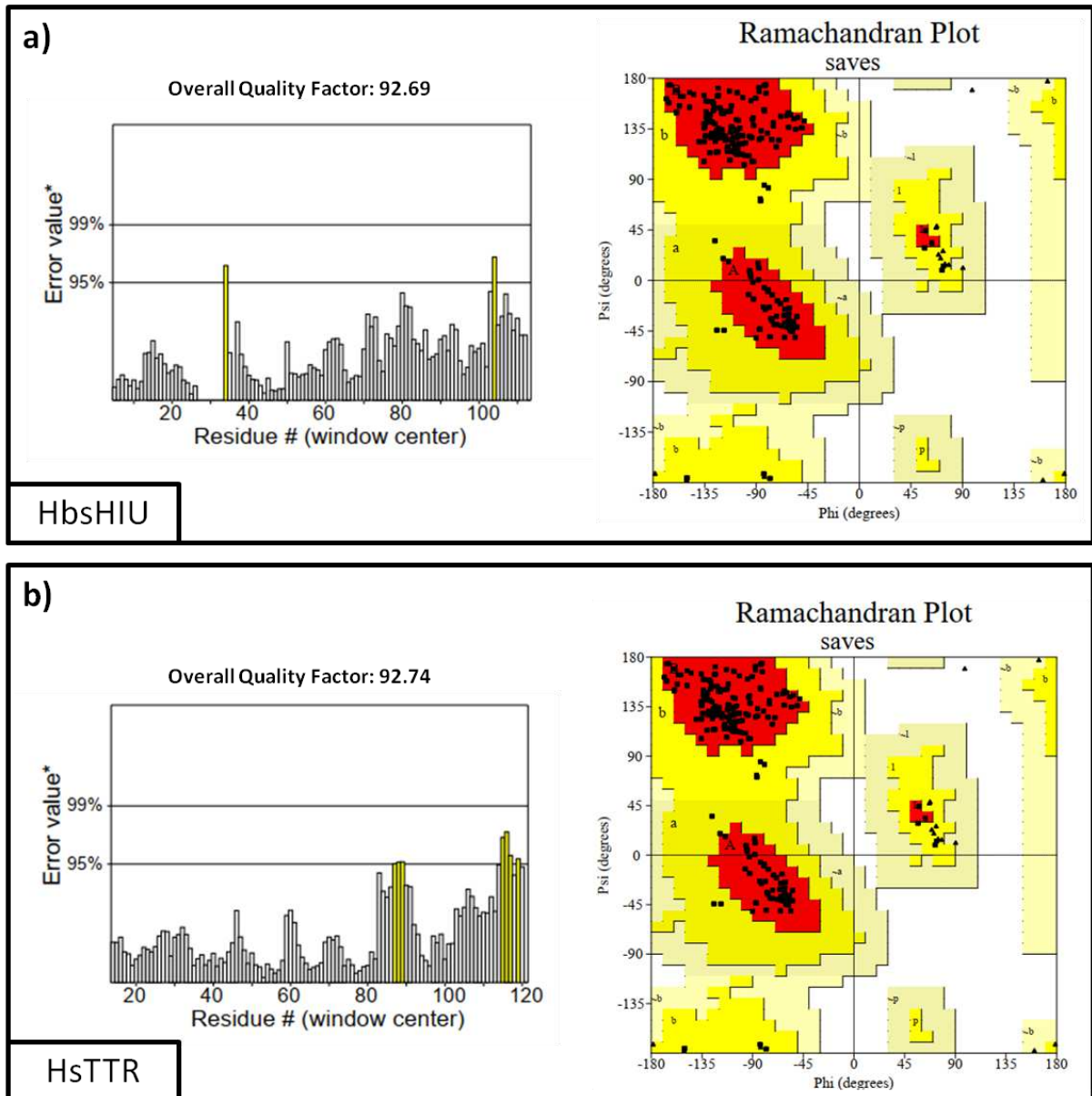
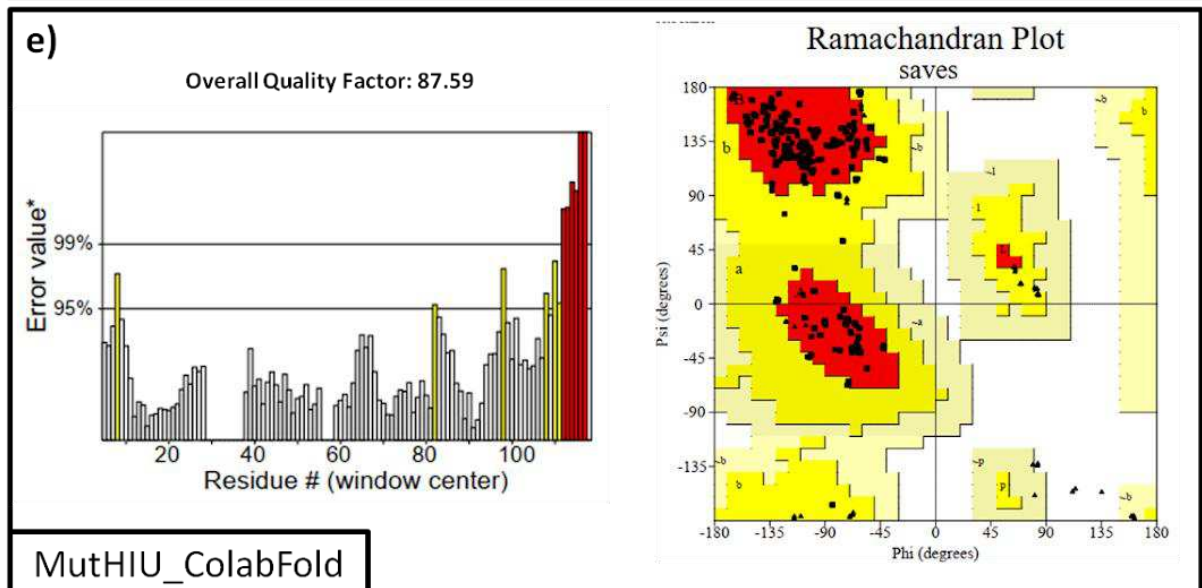
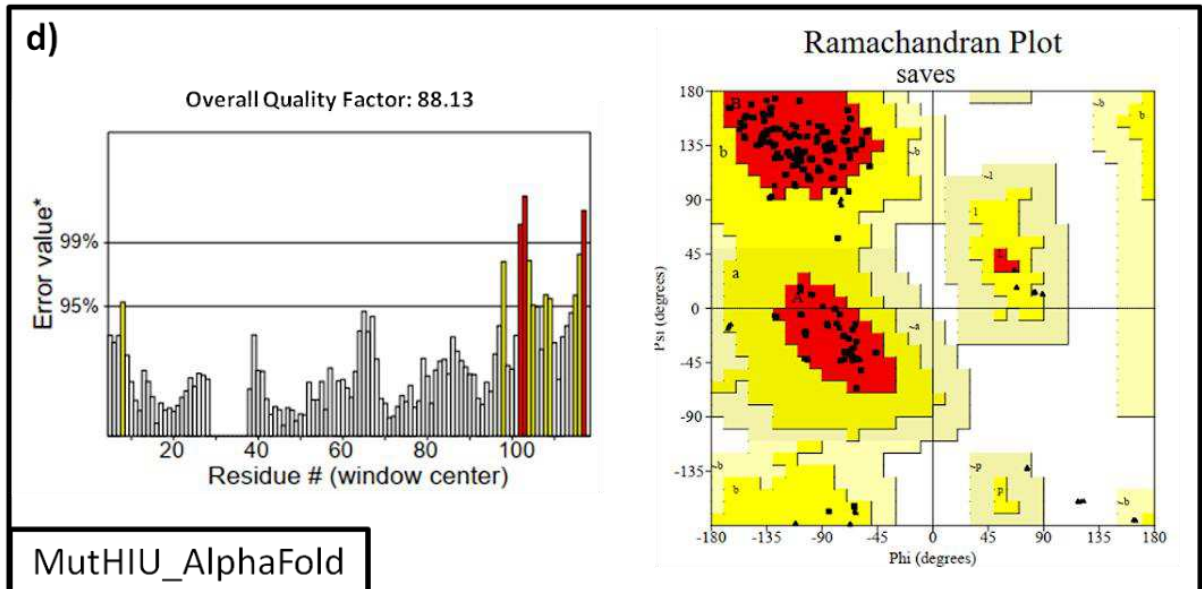
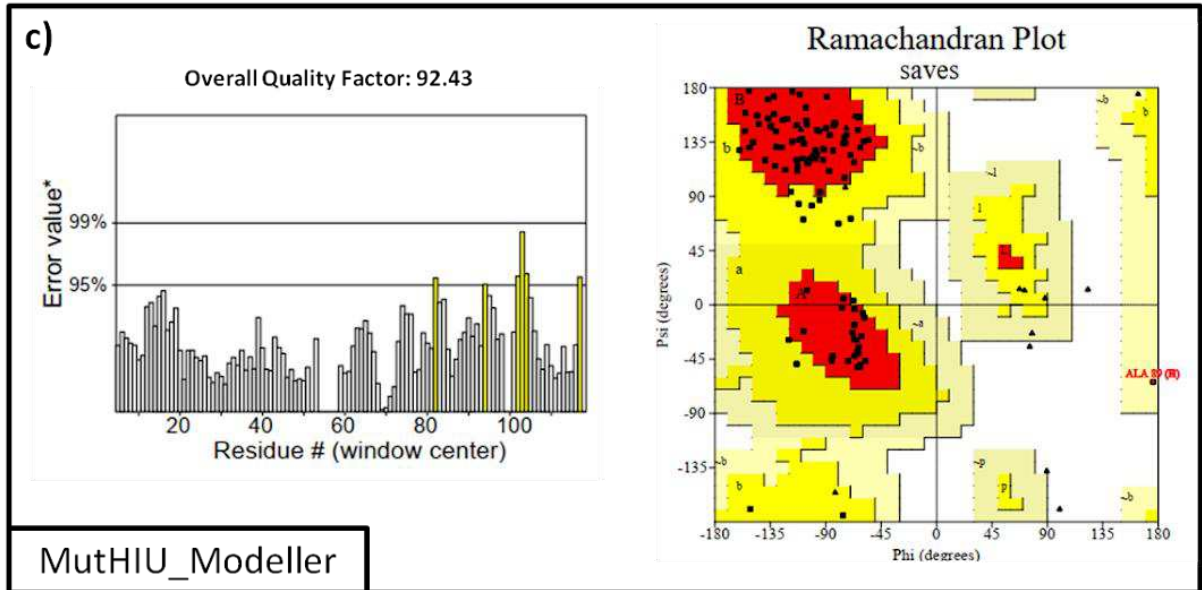
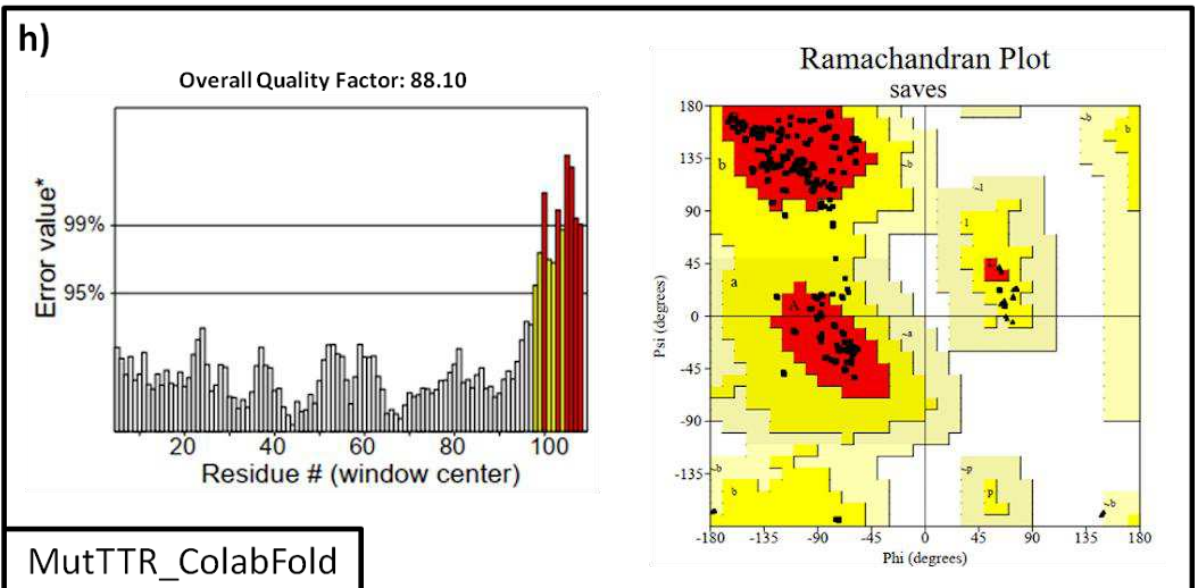
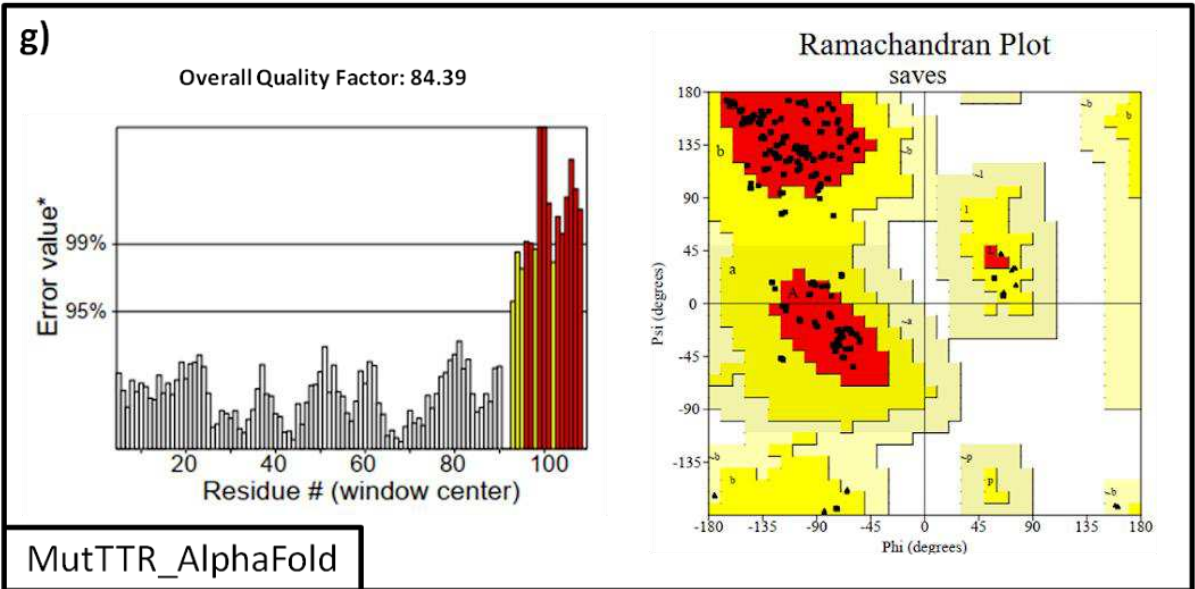
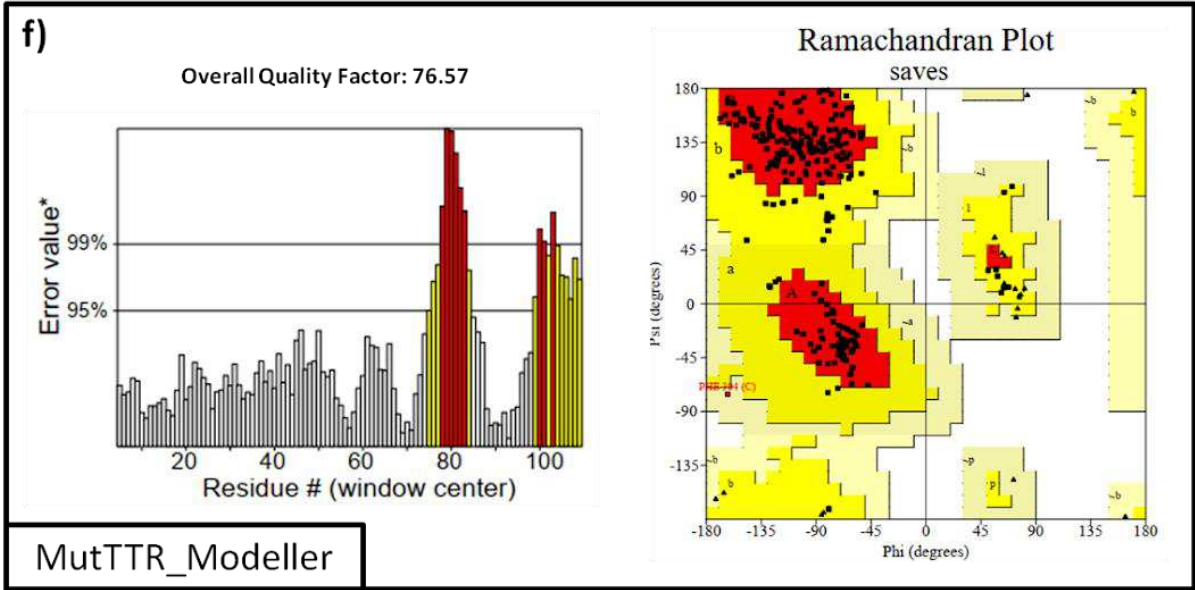


Figura 28 - Gráficos de ERRAT e Ramachandran das estruturas tridimensionais geradas *in silico*. A e B) HIUase de *Herbaspirillum seropedicae* (não depositada) e TTR de *Homo sapiens* (ID PDB: 1F41; Hörnberg *et al.*, 2000), usadas como moldes para a modelagem por homologia, e como controle das análises. C, D, E) Modelos *in silico* da proteína Mutante de HIUase (MutHIU) utilizando *Modeller* (C), *AlphaFold* (D), e *ColabFold* (E). F, G, H) Modelos *in silico* da proteína mutante de TTR (MutTTR) utilizando *Modeller* (F), *AlphaFold* (G), e *ColabFold* (H). Os gráficos de ERRAT são mostrados à esquerda da figura. Regiões em amarelo se encontram acima do intervalo de 95% de confiança de rejeição. Regiões em vermelho se encontram acima do intervalo de 99% de confiança de rejeição. Os valores gerais de qualidade da estrutura são mostrados acima do gráfico. Apenas a cadeia A de cada proteína é mostrada. Os gráficos de Ramachandran são mostrados à direita nas figuras. Pontos em preto identificam os resíduos de aminoácidos individualmente em cada proteína. Regiões em vermelho são as mais favorecidas, em amarelo adicionalmente permitidas, bege generosamente permitidas, e branco não permitidas. Os eixos identificam os ângulos *phi* (eixo x) e *psi* (eixo y), variando de -180 a 180°.





Como pode ser observado, ambas as proteínas cujas estruturas tridimensionais foram resolvidas experimentalmente (Figura 28a e 28b; Tabela 2) apresentam valores abaixo do erro de ERRAT, com apenas dois resíduos ultrapassando o intervalo de 95% de rejeição, e nenhum atingindo o limiar de 99% em HbsHIU. Em HsTTR, sete resíduos ultrapassaram o intervalo de 95%, e nenhum ultrapassou o limiar de 99%.

Por outro lado, todas as proteínas mutantes apresentam regiões que ultrapassam os intervalos de 95 e 99% de confiança de rejeição (Figura 28c-h), de acordo com valores esperados para proteínas resolvidas experimentalmente *in vitro*. Essas regiões contêm conformações provavelmente instáveis ou incorretas em suas estruturas. Estas regiões se encontram especialmente em voltas e nas porções terminais das proteínas, que são em sua maioria flexíveis e irregulares, dificultando assim a determinação de suas conformações por técnicas experimentais e computacionais (Tang, Zhang & Liang, 2014).

Porém, como em ambas a HIUase e a TTR não existem resíduos localizados nestas voltas que exercem as funções catalíticas ou relacionadas à ligação com os hormônios tireoidianos, pode-se atribuir estas possíveis instabilidades a limitações das técnicas utilizadas, já que as sequências modeladas para as mutantes são essencialmente quiméricas entre as duas proteínas de referência. Para as regiões terminais, que de forma oposta possuem diversos resíduos catalíticos, incluindo o tetrapeptídeo YRGS em HIUase, essas instabilidades poderiam potencialmente influenciar na abertura das cavidades e na posição dos resíduos em seu interior. Dessa forma, outras técnicas de refinamento e validação se mostram necessárias para determinar os melhores modelos. Essas técnicas e análises serão discutidas posteriormente nesta dissertação.

A única exceção às instabilidades se dá na MutHIU, gerada pelo método de modelagem por homologia (Figura 28c), que obteve valores semelhantes às proteínas de referência, não demonstrando nenhuma região ultrapassando o limiar de 99% de rejeição. A possível razão para este resultado é que a MutHIU sofreu menos alterações de resíduos que a outra mutante (12% contra 17%, em MutHIU e MutTTR, respectivamente)(Figura 27), conservando assim maior similaridade estrutural com sua proteína de origem, o que satisfaria melhor a modelagem por preenchimento de restrições espaciais empregada pelo *Modeller*.

Quanto aos gráficos de Ramachandran, é possível observar que nenhuma das estruturas, tanto experimentais quanto obtidas computacionalmente, apresentam resíduos em posições não permitidas, com apenas uma parcela em regiões generosamente permitidas em dois dos modelos obtidos (Tabela 2). Analisando a figura 28, também se nota que as distribuições dos ângulos diedros de todos os resíduos e de todos os modelos, seguem o

mesmo padrão, reforçando a similaridade estrutural entre as proteínas mutantes e as de referência.

4.2.4. Comparação Entre os Arcabouços das Proteínas de Referência e Mutantes

Como as HIUase e as TTR possuem uma alta sobreposição entre suas estruturas quaternárias (Figura 21), uma análise da variação espacial do arcabouço das proteínas evidenciaria as regiões que são menos conservadas, assim como regiões que possuem baixa variação estrutural e são mais conservadas evolutivamente.

Para isso, foi feita a comparação das taxas de deslocamento espacial dos arcabouços das proteínas sobrepostas, por meio do cálculo do RMSD entre o carbono-alfa ($C\alpha$) de cada resíduo individualmente (Figura 29).

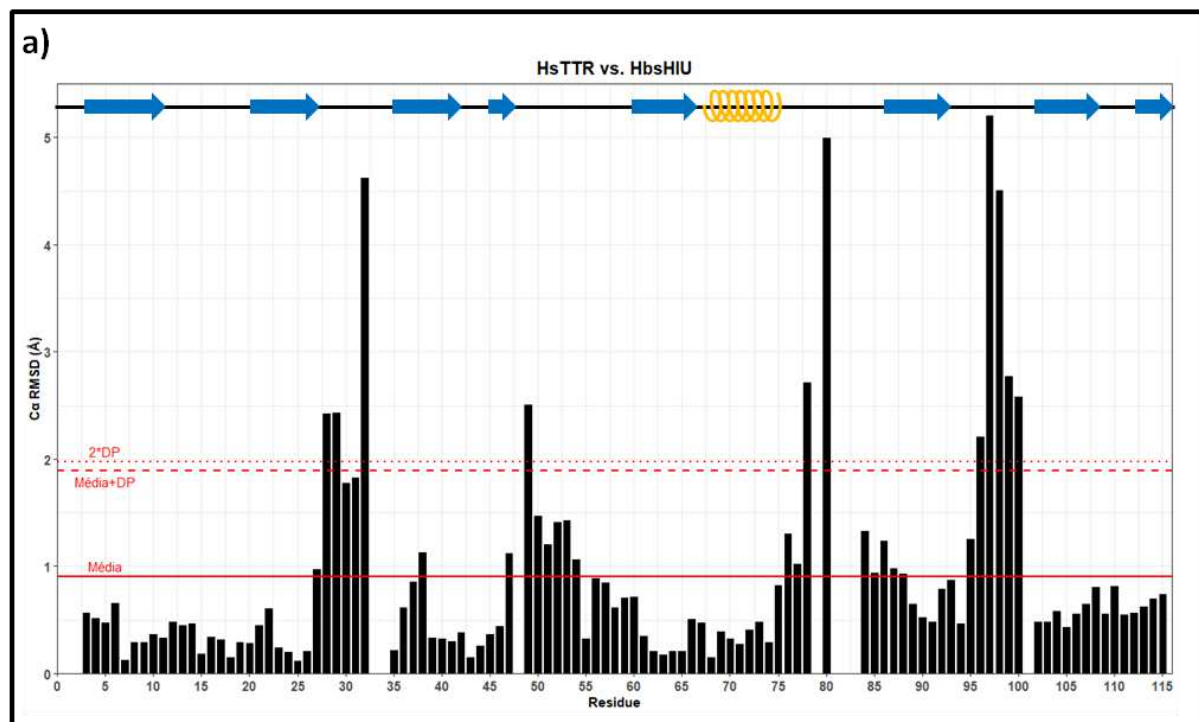
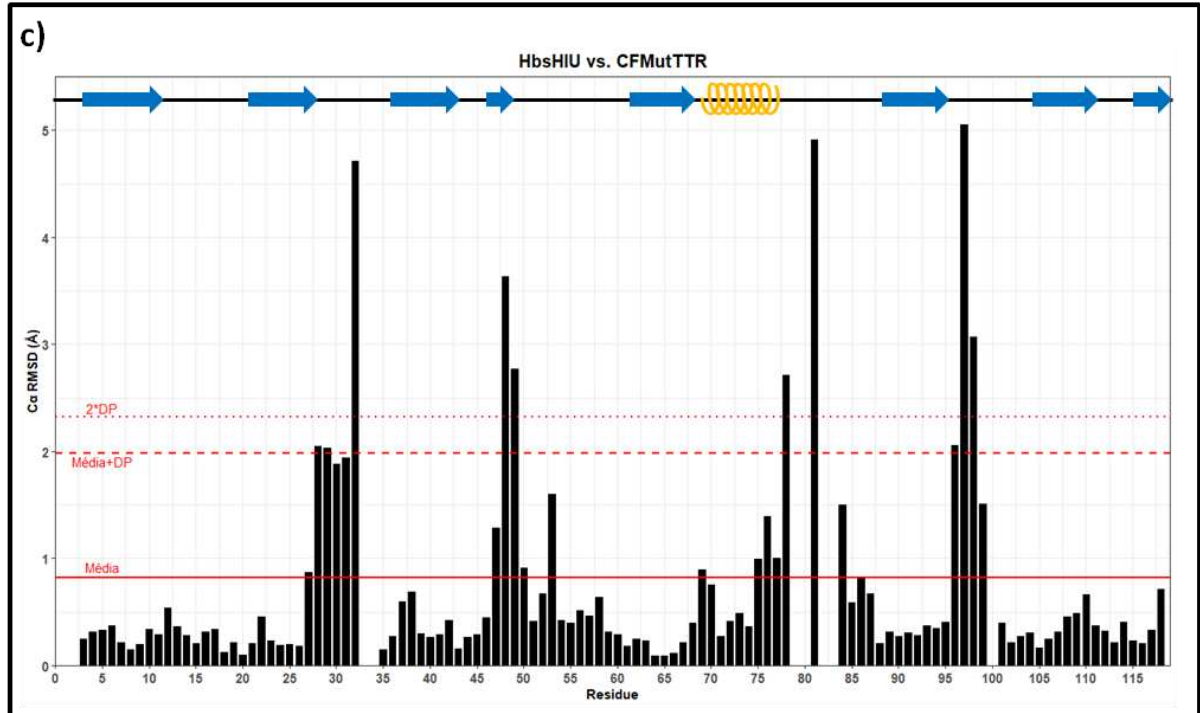
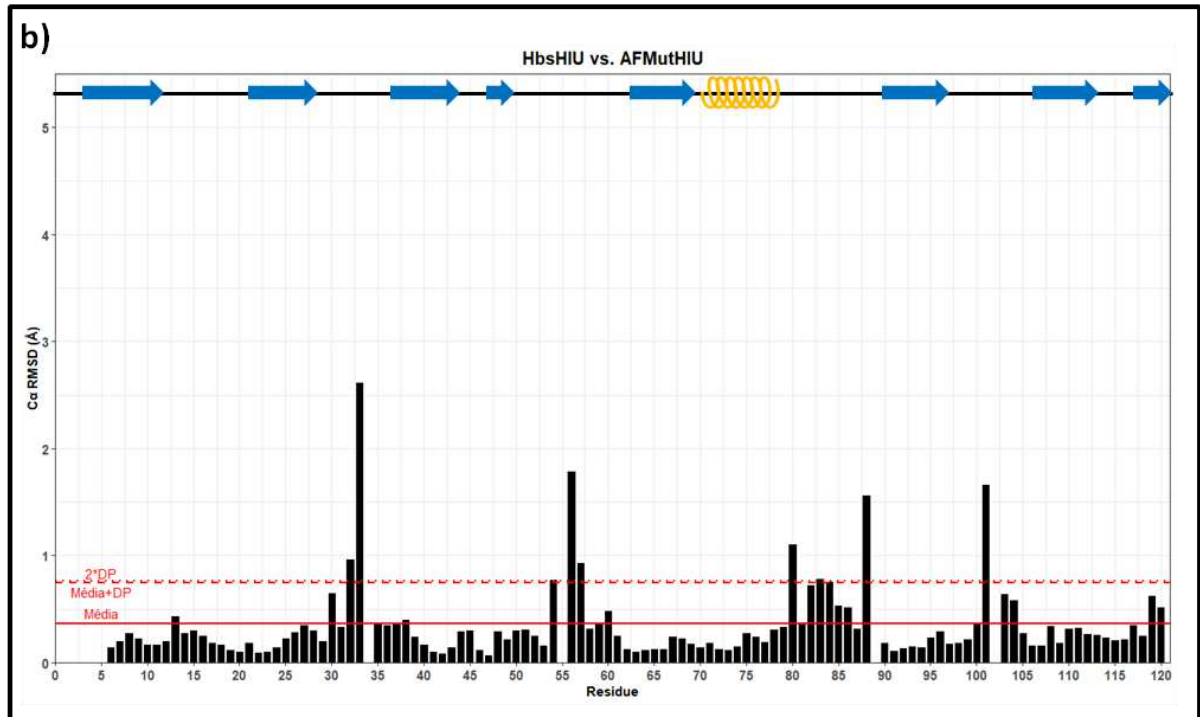
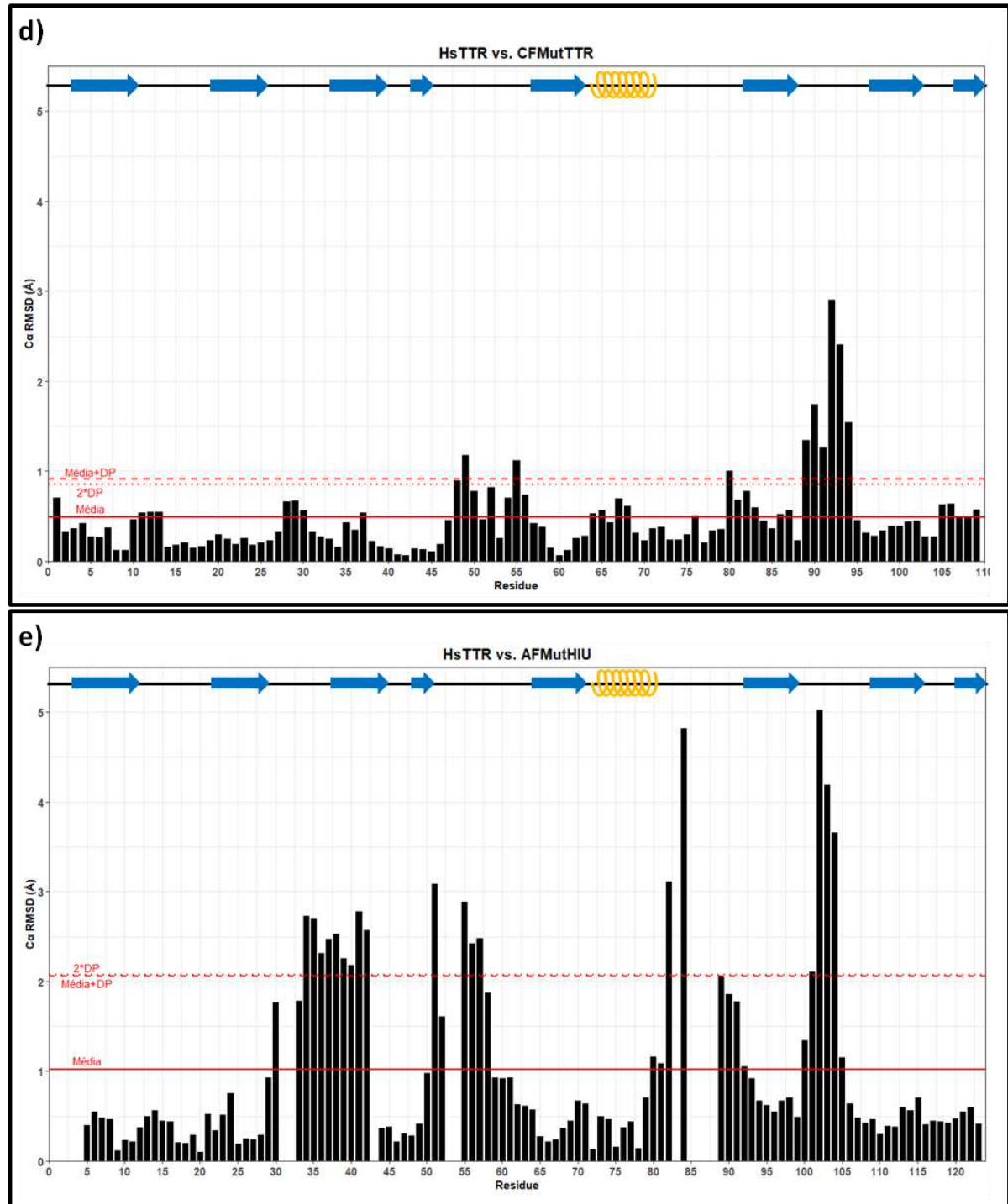


Figura 29 - Cálculo da variação de RMSD entre os $C\alpha$ de cada resíduo das proteínas de referência e mutantes. A) HbsHIU em comparação com HsTTR. B) HbsHIU em comparação com MutTTR gerada pelo *ColabFold*. C) HbsHIU em comparação com MutHIU gerada pelo *AlphaFold*. D) HsTTR em comparação com MutTTR gerada pelo *ColabFold*. E) HsTTR em comparação com MutHIU gerada pelo *AlphaFold*. O eixo X representa cada resíduo das proteínas sobrepostas, após alinhamento. O eixo Y representa o RMSD entre os $C\alpha$ de cada resíduo. A linha vermelha contínua identifica a média entre todos os resíduos, linha tracejada a média somada do desvio padrão, e linha pontilhada duas vezes o desvio padrão. O mapeamento das estruturas secundárias está mostrado abaixo do título, com setas azuis representando folhas- β e a espiral amarela representando α -hélices.





Ao analisar individualmente cada um dos resultados, é possível perceber que as duas proteínas de referência, quando alinhadas, possuem regiões com variações entre as posições espaciais de seus $C\alpha$ acima de duas vezes o valor do desvio padrão estabelecido para toda a estrutura (Figura 29a). Não surpreendentemente, todos os pontos com um RMSD acima dos pontos de corte estão localizados em regiões onde não há estrutura secundária definida, como as oito folhas- β e a α -hélice presentes nas estruturas de ambas. Este resultado já era esperado, pois enquanto as estruturas secundárias são conservadas evolutivamente, as regiões de

conexão tem menor pressão seletiva e com isso possuem maior variação (este resultado também pode ser observado entre os alinhamentos das duas proteínas, na Figura 18).

Já entre a comparação das proteínas de referência com as mutantes que utilizam o mesmo arcabouço (HbsHIU e MutHIU, figura 29b; e HsTTR e MutTTR, figura 29d), quase não existem regiões onde há uma grande variação entre os C α das estruturas, o que é evidenciado pelos valores de média e desvio padrão bem mais baixos, do que na comparação anterior. Mesmo nos pontos mais variáveis, todos estes estão localizados em regiões menos estruturadas, como voltas. Este resultado já era esperado, já que estas proteínas mantêm grande parte dos resíduos, visto que o arcabouço das mutantes é mantido.

Por último, na comparação entre as proteínas mutantes e aquelas cujas funções estão sendo interconvertidas (HbsHIU e MutTTR, figura 29c; e HsTTR e MutHIU, figura 29e), os valores se assemelham à comparação entre as duas proteínas de referência, com os valores de média e o desvio padrão mais altos, além de diversas regiões com variações acima dos pontos de corte. Mais uma vez, todas essas regiões estão localizadas em regiões onde não há estrutura secundária definida, já que mesmo com arcabouços diferentes, a organização espacial das proteínas é bastante similar, exceto em regiões de volta que conectam as folhas- β .

4.2.5. Volumes e Geometria das Cavidades

Após a geração e validação estrutural dos modelos tridimensionais a partir das sequências mutantes, o próximo passo foi analisar a geometria e volume das cavidades internas dos mesmos, visando identificar variações causadas pelas mudanças de resíduos, e compará-las com as proteínas de referência (Figura 22 e Tabela 1).

Iniciando a análise pela MutTTR, dentre todos os modelos *in silico*, os que mais se aproximaram do volume obtido para a cavidade de HIUase de *Herbaspirillum seropedicae* foram aqueles gerados pelo *AlphaFold* e *ColabFold*, com uma diferença de menos de 5Å³ entre eles, e uma diferença de cerca de 250Å³ para a proteína de referência (Tabela 4).

Tabela 4 - Comparação entre os volumes das cavidades de HbsHIU e Mutantes de TTR.

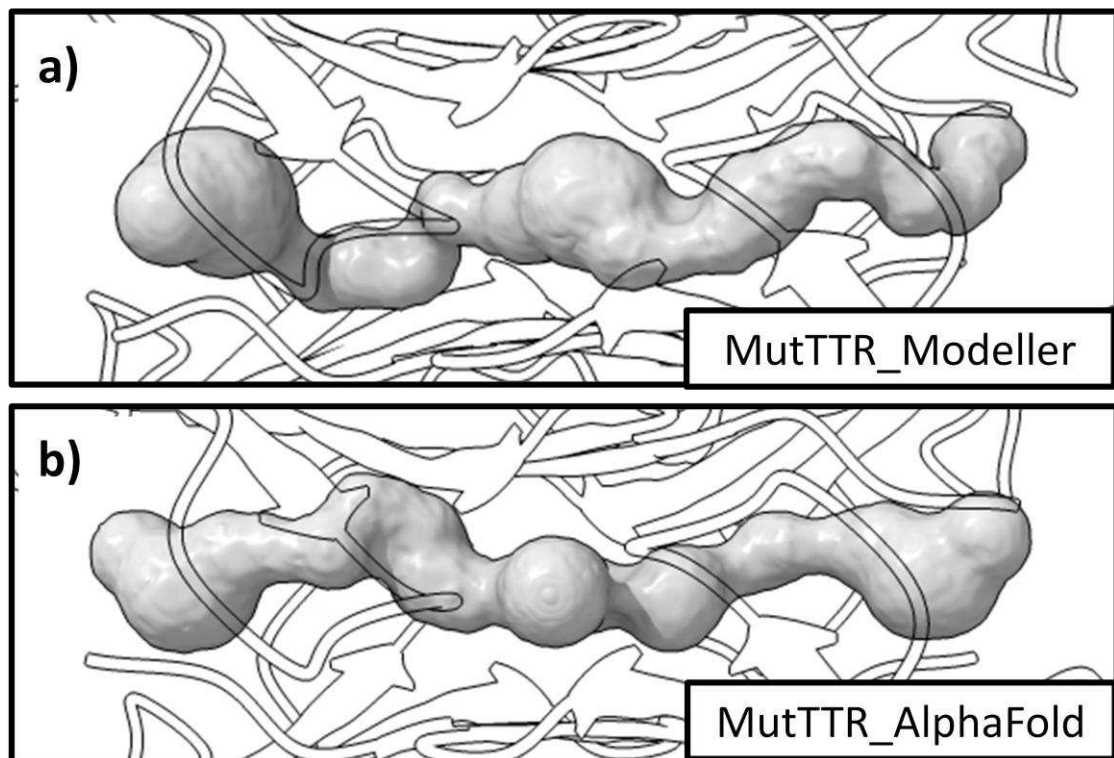
Proteína	Obtenção	Volume da Cavidade (Å ³)	Ligante	Volume do Ligante (Å ³)
HbsHIU	Experimental	1274,8	5-HIU	119,7
	Modeller	838,2		
MutTTR	AlphaFold	1003,9		
	ColabFold	1008,3		

Fonte: próprio autor.

Nota: são apresentadas as ferramentas utilizadas para a geração de cada modelo *in silico*, além do volume da cavidade de cada proteína e do provável ligante.

O volume cerca de 25% menor para as proteínas mutantes, em relação à proteína de referência, pode ser explicado pela variação dos modelos em relação às porções terminais das sequências. Como já discutido anteriormente, os resíduos nessas regiões se localizam espacialmente na entrada das cavidades das proteínas, e como não há referência para modelagem de resíduos terminais, estes podem ter se orientado de modo a diminuir o volume disponível ao solvente.

Estes resultados também podem ser observados na geometria das cavidades das proteínas mutantes, que mesmo possuindo o arcabouço de uma TTR, possuem uma cavidade bem mais estreita, comparável a uma HIUase (Figura 30). O “estrangulamento” causado pela presença do resíduo Tyr114, não presente em TTR, também é visível em todas as proteínas.



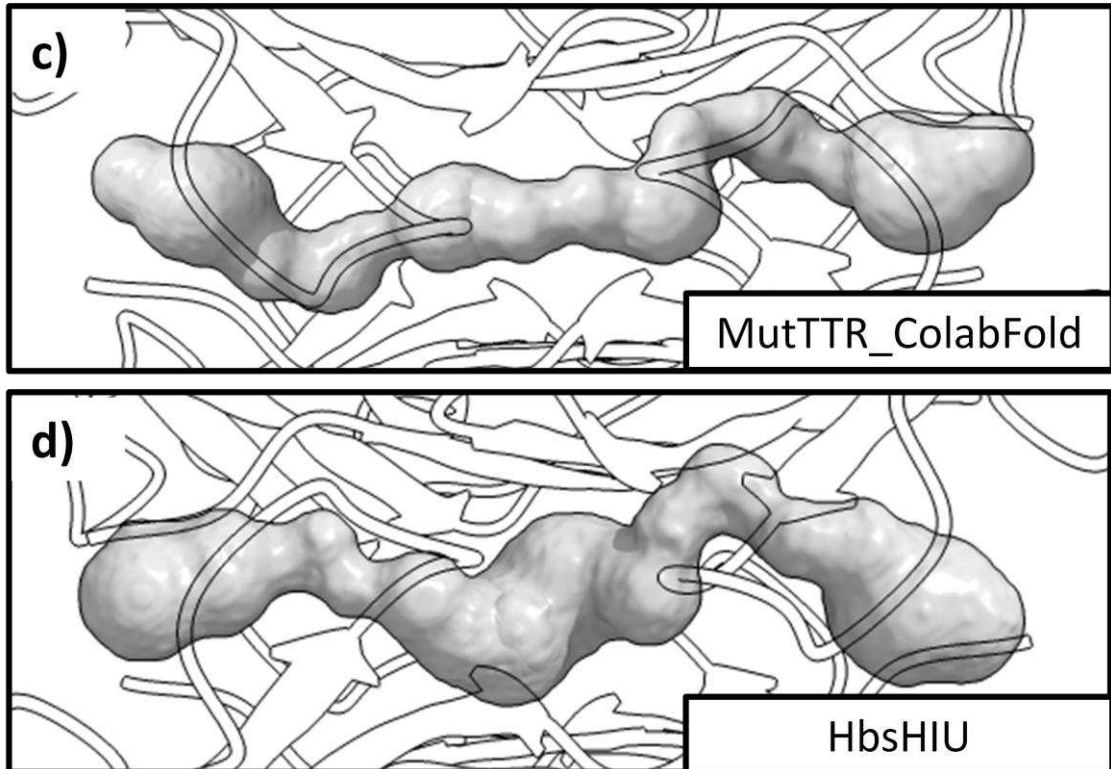


Figura 30 - Geometria das cavidades das proteínas mutantes de TTR. A) Modelo gerado pelo *Modeller*. B) Modelo gerado pelo *AlphaFold*. C) Modelo gerado pelo *ColabFold*. D) HIUase de *Herbaspirillum seropedicae*, proteína de referência. Todas as figuras estão mostrando em detalhe apenas a região das cavidades das proteínas.

Na outra proteína mutante, MutHIU, os volumes das cavidades se apresentam basicamente da maneira oposta em relação à análise anterior. Ambos os modelos do *AlphaFold* e *ColabFold* apresentam valores similares, mas cerca de 400\AA^3 de volume superior à HsTTR. Por outro lado, o modelo gerado pelo *Modeller* possui um volume abaixo do esperado, mas com uma variação menor (Tabela 4).

Tabela 5 - Comparação entre os volumes das cavidades de HsTTR e Mutantes de HIUase.

Proteína	Obtenção	Volume da Cavidade (\AA^3)	Ligante	Volume do Ligante (\AA^3)
HsTTR	Experimental	1737,9	T3 e T4	296,7 - 321,6
	Modeller	1536,5		
MutHIU	AlphaFold	2245,4		
	ColabFold	2067,4		

Fonte: próprio autor.

Nota: são apresentadas as ferramentas utilizadas para a geração de cada modelo *in silico*, além do volume da cavidade de cada proteína e do provável ligante.

Mais uma vez, estes resultados podem ser atribuídos aos resíduos terminais das proteínas, que se encontram logo na entrada das cavidades. Ao contrário da HbsHIU, estes podem ter se orientado de modo a gerar uma abertura maior na extremidade da proteína. Isso

pode ser observado na geometria das cavidades das proteínas mutantes de HIUase, quando comparadas à HsTTR (Figura 31).

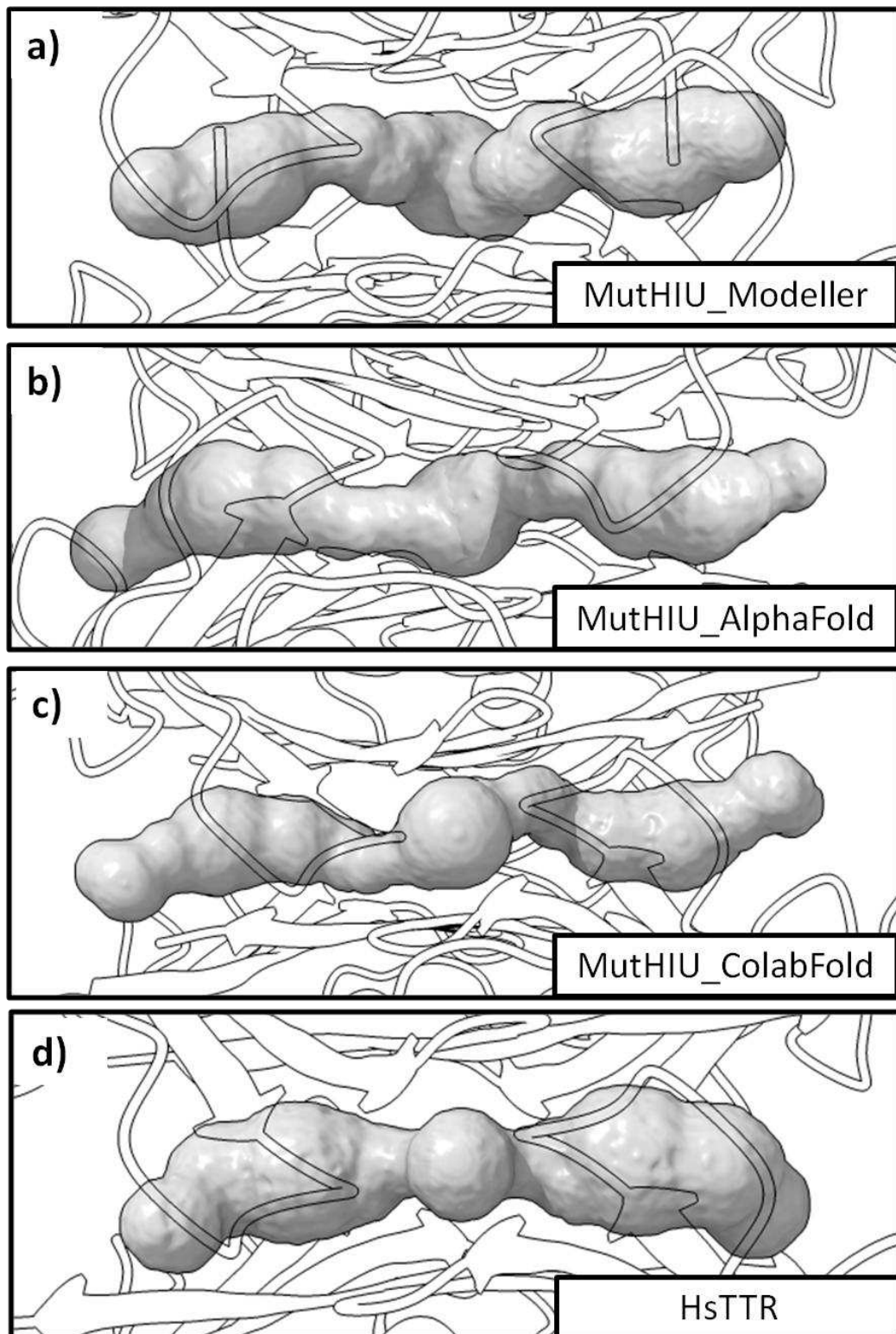


Figura 31 - Geometria das cavidades das proteínas mutantes de HIUase. A) Modelo gerado pelo *Modeller*. B) Modelo gerado pelo *AlphaFold*. C) Modelo gerado pelo *ColabFold*. TTR de *Homo sapiens*, proteína de referência. Todas as figuras estão mostrando em detalhe apenas a região das cavidades das proteínas.

Estas proteínas mutantes, mesmo possuindo o arcabouço de uma HIUase, possuem uma cavidade ampla e bem mais volumosa quando comparadas à sua proteína de origem. A troca do resíduo de tirosina por um de treonina, como presente em TTR, também impede que ocorra o estrangulamento das cavidades, permitindo assim o acesso mais profundo de um composto maior, como é o caso dos hormônios tireoidianos.

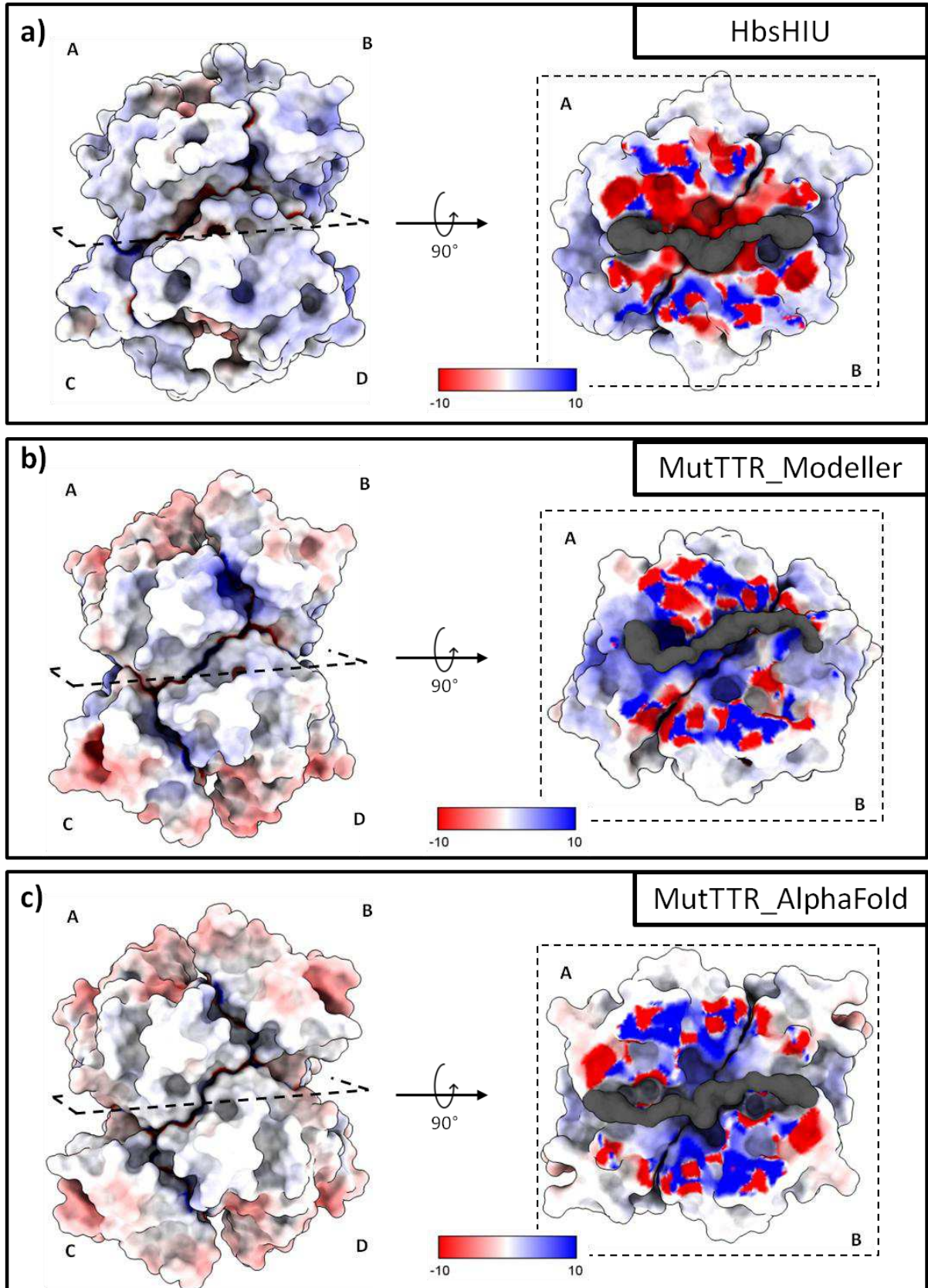
Também é possível perceber, no caso dos modelos gerados pelo *AlphaFold* e *ColabFold*, a maior extensão lateral das cavidades, o que evidencia o maior volume destes modelos, como mostrado na Tabela 4.

4.2.6. Potenciais Eletrostáticos e Distribuição de Cargas nas Cavidades

Além do volume e geometria das cavidades, a distribuição de cargas no interior das proteínas também é importante para a caracterização e função das HIUase e TTR. Sendo assim, o potencial eletrostático de todas as mutantes geradas foi analisado, tanto nas regiões internas pertencentes às cavidades quanto nas porções externas das proteínas.

Primeiramente, para a HIUase de *Herbaspirillum seropedicae* (Figura 32a), é possível identificar que embora a parte externa da proteína possua um potencial eletrostático predominantemente positivo (cores azuladas), o interior da cavidade apresenta o potencial oposto, sendo predominantemente negativo (cores avermelhadas logo ao redor da cavidade, identificada em cinza escuro). A escala de cores também revela que os potenciais eletrostáticos internos e externos variam bastante, com um interior que possui carga líquida mais acentuada. Este resultado encontra-se em conformidade com dados encontrados na literatura para HIUase de *Danio rerio* e *Bacillus subtilis*, e TTR de *Homo sapiens* (Jung *et al.*, 2006; Zanotti *et al.*, 2006).

Porém, analisando as proteínas mutantes de TTR (Figura 32b-d), é possível perceber que os padrões encontrados na proteína de referência não se mantêm, já que em todos os três modelos gerados *in silico*, os potenciais apresentados são predominantemente negativos nas superfícies e positivos nos interiores das cavidades, o oposto do observado para a HbsHIU.



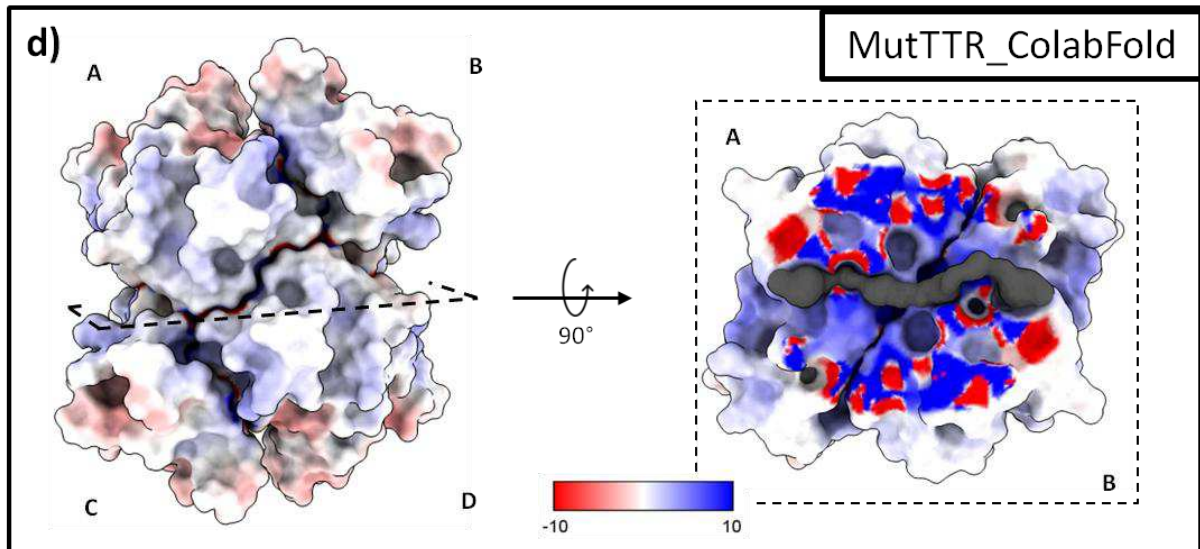


Figura 32 - Potenciais eletrostáticos em HbsHIU e todos os modelos de MutTTR. A) HIUase de *Herbaspirillum seropedicae*, utilizada como proteína de referência. B-D) Mutantes de TTR, geradas por diferentes metodologias de modelagem molecular *in silico*: *Modeller* (B), *AlphaFold* (C), e *ColabFold* (D). À esquerda, são mostradas as estruturas quaternárias homotetraméricas das proteínas, com cada monômero sendo identificado por uma letra maiúscula. À direita, em uma rotação de 90° do dímero superior (A-B), são mostrados detalhes das cavidades das proteínas, evidenciando a distribuição de cargas internas. Em cinza escuro é mostrada a geometria da cavidade, para melhor interpretação.

Esta divergência entre os potenciais eletrostáticos não era esperada, já que as mutantes de TTR teoricamente mimetizam as condições de uma HIUase, incluindo a distribuição de cargas, que é essencial para a reação enzimática catalisada por esta proteína. Curiosamente, a MutTTR sofreu o maior número de alterações em resíduos localizados na cavidade, em relação à MutHIUase (Figura 27). Estes resultados necessitam de uma investigação mais profunda, para identificar a possível causa dessa divergência, que inicialmente é atribuída à falta de refinamento da abordagem utilizada.

Por outro lado, para a TTR de *Homo sapiens* (Figura 33a), é possível identificar um padrão também oposto à HbsHIU, com a porção externa da proteína apresentando caráter predominantemente negativo na porção superior e inferior, e positivo na porção central. O interior é predominantemente positivo ao redor da cavidade. Levando em consideração o resultado anterior, fica visível que a MutTTR, ao invés de apresentar o padrão de potenciais eletrostáticos de uma HIUase, como era esperado, apresenta o padrão de TTR. Uma conclusão que pode ser obtida deste resultado é que a troca de resíduos presente nas MutTTR, embora tenha servido para alterar o volume e a geometria da cavidade, não levou à alteração das cargas internas da proteína.

Analisando agora a outra proteína mutante, MutHIU, os padrões observados se encontram diferentes entre os modelos gerados *in silico*. Para a proteína gerada utilizando o *Modeller* (Figura 33b), a parte externa se assemelha a uma TTR, porém o interior possui um

potencial completamente oposto, sendo majoritariamente negativo. Além disso, as cargas não estão bem distribuídas neste modelo, evidenciando um potencial erro de modelagem na abordagem por homologia.

Por outro lado, os modelos gerados pelo *AlphaFold* e *ColabFold* (Figura 33c-d), o padrão interior das cavidades é extremamente similar à proteína de referência, com a porção diretamente em contato com a cavidade apresentando caráter positivo.

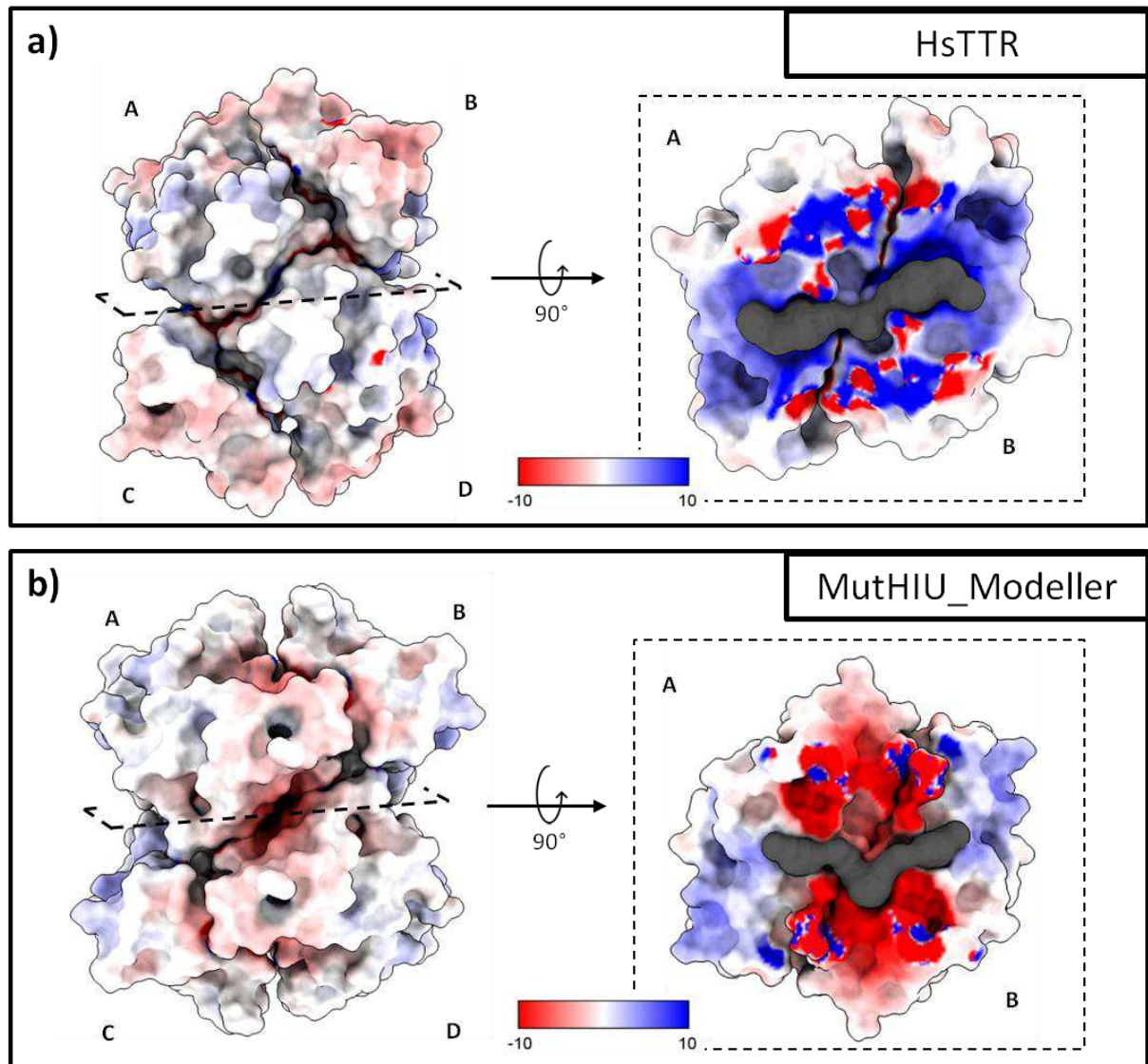
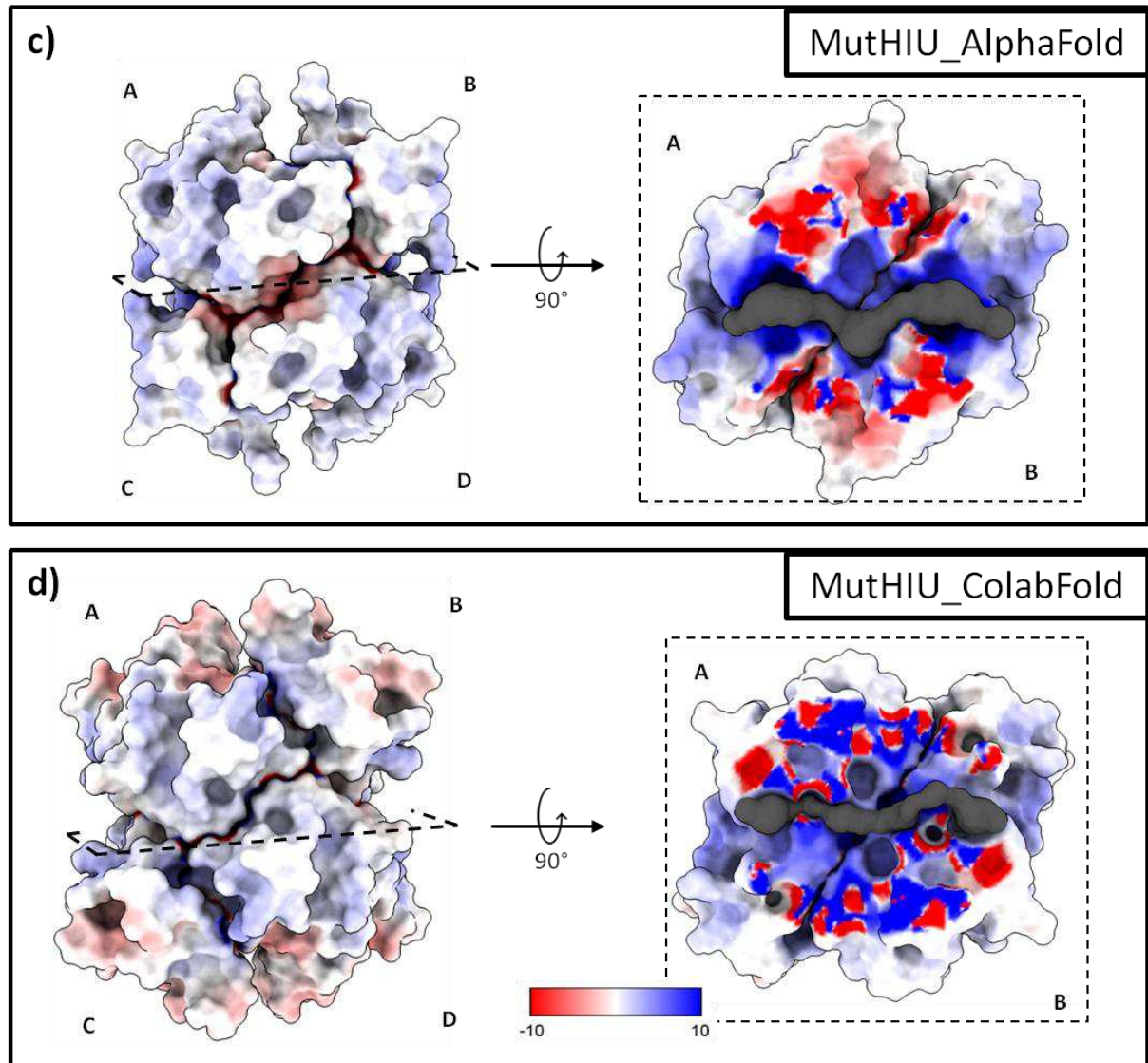


Figura 33 - Potenciais eletrostáticos em HsTTR e todos os modelos de MutHIU. A) TTR de *Homo sapiens*, utilizada como proteína de referência. B-D) Mutantes de HIUase, gerados por diferentes metodologias de modelagem molecular *in silico*: *Modeller* (B), *AlphaFold* (C), e *ColabFold* (D). À esquerda, são mostradas as estruturas quaternárias homotetraméricas das proteínas, com cada monômero sendo identificado por uma letra maiúscula. À direita, em uma rotação de 90° do dímero superior (A-B), são mostrados detalhes das cavidades das proteínas, evidenciando a distribuição de cargas internas. Em cinza escuro é mostrada a geometria da cavidade, para melhor interpretação.



De forma oposta ao resultado obtido com a HbsHIU e a mutante de TTR, a comparação entre a HsTTR e a mutante de HIUase evidenciou, em dois dos três modelos gerados *in silico* (*AlphaFold* e *ColabFold*), o que já era esperado. Ao apresentarem o mesmo padrão de cargas no interior da cavidade que uma TTR, essas mutantes são, potencialmente, capazes de acomodar de forma satisfatória os hormônios tireoidianos, já que outras métricas como o volume e geometria das cavidades também corroboram estes resultados.

4.2.7. Resíduos Expostos

A maioria das modificações de resíduos nas proteínas mutantes ocorreu na região das cavidades, onde também ocorrem os processos catalíticos e de transporte de HIUase e TTR, respectivamente. Pensando nisso, foi analisada a distribuição de resíduos que estariam expostos ao solvente nessas proteínas, a fim de corroborar a conservação estrutural das mutantes (Figura 34). Apenas os resíduos com valor de ASA fracional das cadeias laterais

acima de 0.5 foram considerados como expostos (Ständer *et al.*, 2021). A lista completa de resíduos expostos e seus valores de ASA são mostrados no apêndice III.

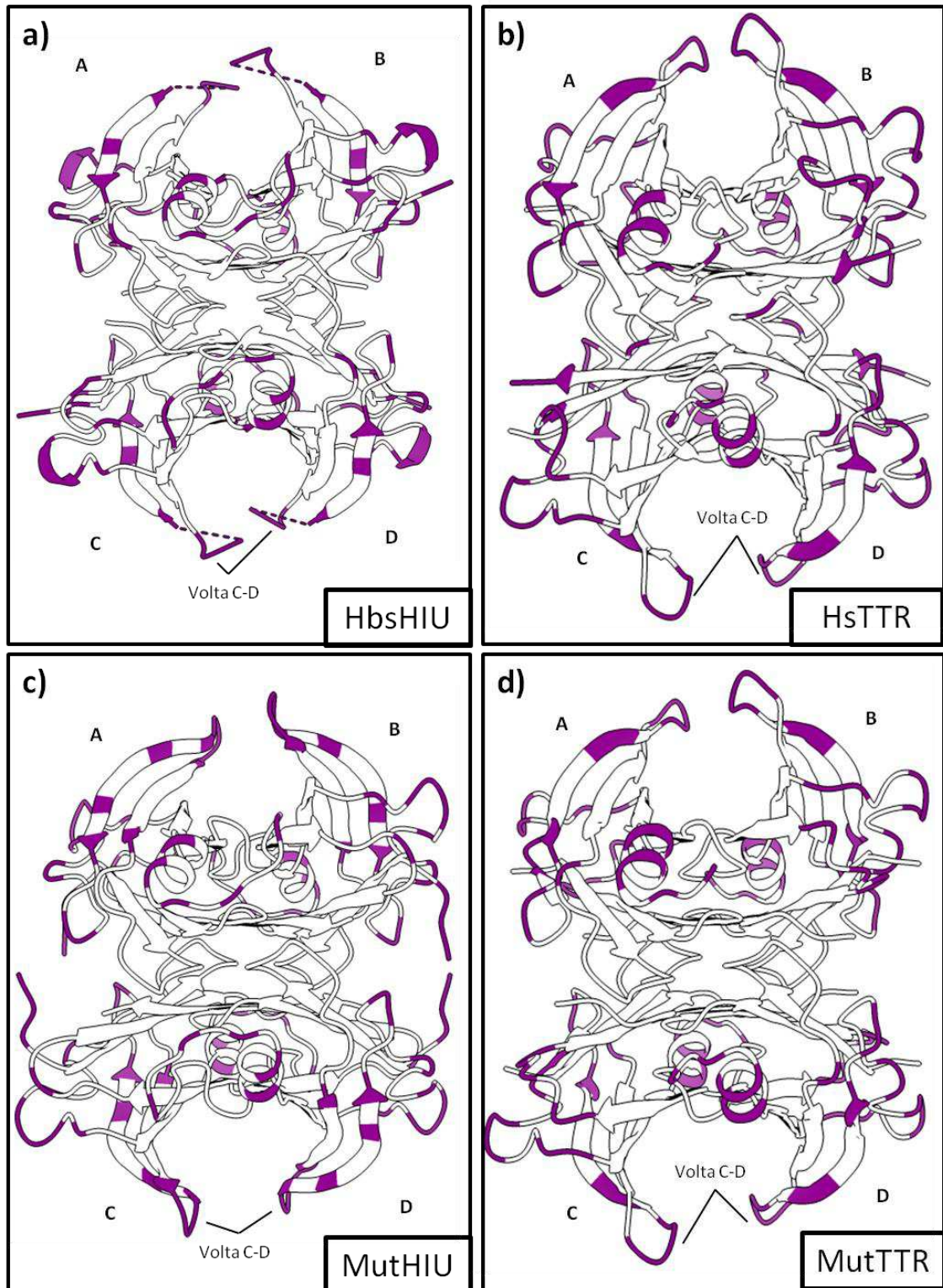


Figura 34 - Mapeamento de resíduos expostos nas estruturas tridimensionais das proteínas de referência e mutantes. A) HIUase de *Herbaspirillum seropedicae*. B) TTR de *Homo sapiens*. C) Proteína mutante de HIUase (MutHIU). D) Proteína mutante de TTR (MutTTR). São mostrados os homotetrâmeros de cada proteína, com cada monômero identificado por uma letra maiúscula, os resíduos expostos ao solvente identificados em roxo. A volta C-D também está identificada nas estruturas.

Como pode ser observado, os resíduos expostos nas quatro proteínas se encontram majoritariamente em regiões de volta, nas folhas- β B e C, e na porção externa da α -hélice, com nenhum resíduo das cavidades exposto diretamente. Além disso, a distribuição espacial das regiões segue o mesmo padrão entre as proteínas.

Apesar disso, analisando os números absolutos e as porcentagens de resíduos expostos entre as proteínas, observa-se uma diferença entre os percentuais das HIUase e TTR (Tabela 5).

Tabela 6 - Comparação entre o número de resíduos expostos entre as diferentes proteínas de referência e mutantes.

Proteína	Tamanho	Resíduos Expostos	%
HbsHIU	117	26	22,2
HsTTR	116	32	27,6
MutHIU	122	26	21,3
MutTTR	113	28	24,8

Fonte: próprio autor.

Nota: são mostradas as diferentes proteínas analisadas, seus tamanhos (em resíduos de aminoácidos), a quantidade absoluta de resíduos expostos e a porcentagem dos mesmos em relação ao total.

De acordo com a tabela, a HIUase e a MutHIU (que mantém o arcabouço de uma HIUase) apresentam uma porcentagem consideravelmente menor de resíduos expostos ao solvente, quando comparadas com a TTR e a MutTTR (que mantém o arcabouço de uma TTR).

As mutações na cavidade de ligação à tiroxina em TTR que levam ao aumento da área de interação entre o tetrâmero costumam conferir uma maior estabilidade à proteína, levando a uma menor propensão de agregação e o desencadeamento de doenças amiloidogênicas, como demonstra o trabalho de Sant'Anna e colaboradores (Sant'Anna *et al.*, 2017). O mecanismo de agregação destas proteínas é atribuído principalmente a estados intermediários parcialmente desenovelados que desencadeiam o processo (Yee *et al.*, 2019). Esses estados se originam especialmente na volta C-D do monômero, que está exposta ao solvente (Figura 34).

Seria plausível então que as HIUase, que também possuem uma propensão muito menor à agregação, também exibam uma maior área de interação entre as unidades oligoméricas, em relação à TTR. Isso explicaria a menor distribuição de resíduos expostos ao solvente, nas regiões externas da proteína.

De fato, ao analisar a área de interação dos dímeros AB e CD (Interface forte, Figura 3), é possível perceber que em HIUase, os valores são cerca de 15% superiores à TTR (Figura

35). Esse padrão também se repete com as proteínas mutantes, onde a MutHIU possui uma área aproximadamente 15% maior que a MutTTR, e 20% maior que a HsTTR.

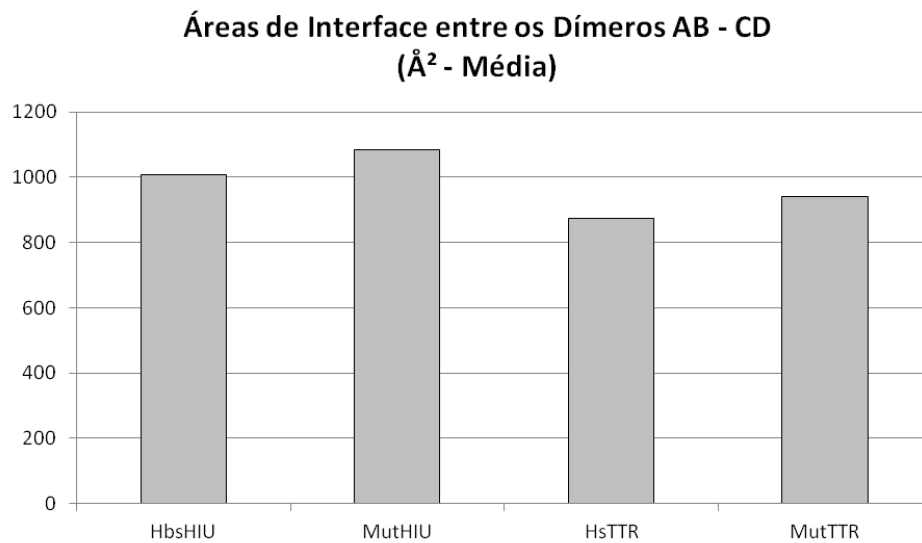


Figura 35 - Gráfico de barras das áreas de interface entre os dímeros AB-CD das proteínas de referência e mutantes. São mostradas em ordem: HIUase de *Herbaspirillum seropedicae*; Mutante de HIUase; TTR de *Homo sapiens*; Mutante de TTR. Os valores mostrados são a média dos dois dímeros.

5. CONCLUSÃO

A função das HIUase e TTR é bastante diferente, sendo que a HIUase é uma enzima e a TTR é majoritariamente uma carreadora de hormônios. Os volumes e as geometrias das cavidades, tão importantes para a distinção dessas proteínas, possuem características distintas. Uma das principais é a presença de um resíduo de tirosina em HIUase, substituído por uma treonina em TTR, que abre caminho para a ligação de uma molécula significativamente maior nas proteínas carreadoras.

Em relação às proteínas mutantes, foram geradas duas sequências, uma teoricamente capaz de simular a função da outra, utilizando como referência a HIUase de *Herbaspirillum seropedicae* e a TTR de *Homo sapiens*. Das duas sequências mutantes, diversos modelos tridimensionais foram gerados *in silico*, utilizando duas abordagens distintas. Todos os modelos foram validados por técnicas estatísticas e visuais, e avaliados em relação às mesmas características anteriormente analisadas nas proteínas de referência.

De fato, as duas proteínas mutantes possuem características que as distinguem de suas proteínas de origem (HbsHIU para MutHIU, e HsTTR para MutTTR, respectivamente), e que as colocam próximas das proteínas opostas. Algumas regiões já descritas anteriormente como variáveis, como algumas voltas e as regiões terminais, se mostraram particularmente desafiadoras, especialmente utilizando a modelagem por homologia, feita por meio de um alinhamento quimérico entre as duas proteínas de referência. Para melhor entendimento destes resultados, é necessária a aplicação de técnicas complementares em biologia estrutural, como dinâmicas e docagens moleculares *in silico*. Análises iniciais já estão em andamento, mas por questões de tempo, não puderam ser extensivamente discutidas e apresentadas aqui.

As perspectivas e perguntas futuras deste trabalho estão relacionadas primeiramente à utilização das técnicas complementares citadas acima, e também com o estudo *in vitro* das proteínas mutantes. Por mais que técnicas computacionais ofereçam diversas respostas em questões de biologia estrutural, a complementaridade de ensaios experimentais não pode ser deixada de lado. Os plasmídeos das proteínas mutantes já foram adquiridos, e o processo de obtenção das proteínas recombinantes já foi iniciado. Devido a complicações experimentais, estes dados não foram aprofundados e discutidos nesta dissertação. Porém, ensaios enzimáticos preliminares com a enzima HIUase já foram realizados e indicam forte correlação com os dados disponíveis na literatura.

Este estudo se encerra com a perspectiva de aplicação da técnica de obtenção de proteínas engenheiradas *in silico* com base no processo evolutivo das mesmas. Os resultados

aqui mostrados evidenciam que uma abordagem de seleção de resíduos localmente conservados é capaz de identificar aqueles que são importantes para a diferenciação de duas proteínas. Porém, para que isto ocorra, estudos mais profundos devem ser realizados, assim como o refinamento do processo.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- AL HAMED, R. et al. Comprehensive Review of AL amyloidosis: some practical recommendations. **Blood Cancer Journal**, v. 11, n. 5, p. 97, maio 2021.
- ALDRED, A. R.; BRACK, C. M.; SCHREIBER, G. The cerebral expression of plasma protein genes in different species. **Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology**, v. 111, n. 1, p. 1–15, maio 1995.
- ALDRED, A. R.; PRAPUNPOJ, P.; SCHREIBER, G. Evolution of Shorter and More Hydrophilic Transthyretin N-Termini by Stepwise Conversion of Exon 2 into Intron 1 Sequences (Shifting the 3' Splice Site of Intron 1). **European Journal of Biochemistry**, v. 246, n. 2, p. 401–409, 1997.
- AMES, B. N. et al. Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 78, n. 11, p. 6858–6862, nov. 1981.
- BAILONE, R. L. et al. Zebrafish as an alternative animal model in human and animal vaccination research. **Laboratory Animal Research**, v. 36, n. 1, p. 13, dez. 2020.
- BALDANI, J. I. et al. Emended Description of *Herbaspirillum*; Inclusion of [*Pseudomonas*] *rubrisubalbicans*, a Mild Plant Pathogen, as *Herbaspirillum rubrisubalbicans* comb. nov.; and Classification of a Group of Clinical Isolates (EF Group 1) as *Herbaspirillum* Species 3. **International Journal of Systematic Bacteriology**. p. 9, 1996.
- BARTALENA, L.; ROBBINS, J. Thyroid Hormone Transport Proteins. **Clinics in Laboratory Medicine**, v. 13, n. 3, p. 583–598, set. 1993.
- BLAKE, C. C. F. et al. Structure of human plasma prealbumin at 2.5 Å resolution. **Journal of Molecular Biology**, v. 88, n. 1, p. 1–12, set. 1974.
- BLAKE, C. C. F. et al. Structure of prealbumin: Secondary, tertiary and quaternary interactions determined by Fourier refinement at 1.8 Å. **Journal of Molecular Biology**, v. 121, n. 3, p. 339–356, maio 1978.
- BLAKE, C. C. F.; OATLEY, S. J. Protein–DNA and protein–hormone interactions in prealbumin: a model of the thyroid hormone nuclear receptor? **Nature**, v. 268, n. 5616, p. 115–120, jul. 1977.
- BRANCH, W. T.; ROBBINS, J.; EDELHOCH, H. Thyroxine-binding Prealbumin. **Journal of Biological Chemistry**, v. 246, n. 19, p. 6011–6018, out. 1971.
- BRENDER, J. R. et al. An Evolution-Based Approach to De Novo Protein Design. **Methods in molecular biology (Clifton, N.J.)**, v. 1529, p. 243–264, 2017.
- BROUILLETTE, J.; QUIRION, R. Transthyretin: A key gene involved in the maintenance of memory capacities during aging. **Neurobiology of Aging**, v. 29, n. 11, p. 1721–1732, nov. 2008.
- BUCHAN, D. W. A.; JONES, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. **Nucleic Acids Research**, v. 47, n. W1, p. W402–W407, 2 jul. 2019.

BUXBAUM, J. N.; REIXACH, N. Transthyretin: the servant of many masters. **Cellular and Molecular Life Sciences**, v. 66, n. 19, p. 3095–3101, out. 2009.

CENDRON, L. et al. Probing the Evolution of Hydroxyisourate Hydrolase into Transthyretin through Active-Site Redesign. **Journal of Molecular Biology**, v. 409, n. 4, p. 504–512, jun. 2011.

CHANG, L. et al. Evolution of thyroid hormone binding by transthyretins in birds and mammals. **European Journal of Biochemistry**, v. 259, n. 1–2, p. 534–542, jan. 1999.

CHOVANCOVA, E. et al. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. **PLoS Computational Biology**, v. 8, n. 10, p. e1002708, 18 out. 2012.

CODY, V.; WOJTCZAK, A. Mechanisms of Molecular Recognition: Structural Characteristics of Transthyretin Ligand Interactions. **In: Recent Advances in Transthyretin Evolution, Structure and Biological Functions**, edited by Richardson, S.J. & Cody, V. Heidelberg, Germany: Springer, 2009, p. 311–328

COLLOC'H, N. et al. Crystal Structure of the protein drug urate oxidase-inhibitor complex at 2.05 Å resolution. **Nature Structural & Molecular Biology**, v. 4, n. 11, p. 947–952, nov. 1997.

COLOVOS, C.; YEATES, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. **Protein Science**, v. 2, n. 9, p. 1511–1519, set. 1993.

DICKSON, P. W. et al. Thyroxine transport in choroid plexus. **Journal of Biological Chemistry**, v. 262, n. 29, p. 13907–13915, out. 1987.

EISENBERG, D.; LÜTHY, R.; BOWIE, J. U. [20] VERIFY3D: Assessment of protein models with three-dimensional profiles. Em: **Methods in Enzymology**. [s.l.] Elsevier, 1997. v. 277p. 396–404.

ENEQVIST, T. et al. The transthyretin-related protein family. **European Journal of Biochemistry**, v. 270, n. 3, p. 518–532, fev. 2003.

EVANS, R. et al. Protein complex prediction with AlphaFold-Multimer. [s.l.] Bioinformatics, 4 out. 2021. Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2021.10.04.463034>>. Acesso em: 13 jul. 2022.

FEX, G.; ALBERTSSON, P.-Å.; HANSSON, B. Interaction between Prealbumin and Retinol-Binding Protein Studied by Affinity Chromatography, Gel Filtration and Two-Phase Partition. **European Journal of Biochemistry**, v. 99, n. 2, p. 353–360, 1979.

FRENCH, J. B.; EALICK, S. E. Structural and kinetic insight into the mechanism of 5-hydroxyisourate hydrolase from *Klebsiella pneumoniae*. **Acta Crystallographica Section D: Biological Crystallography**, v. 67, n. Pt 8, p. 671–677, 12 jul. 2011.

GOODMAN, D. S. Statement Regarding Nomenclature for the Protein Known as Prealbumin, Which is also (Recently) Called Transthyretin. Em: GLENNER, G. G. et al. (Eds.). **Amyloidosis**. Boston, MA: Springer US, 1986. p. 287–288.

GU, J. et al. Evolution-Based Protein Engineering for Antifungal Peptide Improvement. **Molecular Biology and Evolution**, v. 38, n. 11, p. 5175–5189, 27 out. 2021.

HAGEN, G. A.; ELLIOTT, W. J. Transport of Thyroid Hormones in Serum and Cerebrospinal Fluid ¹. **The Journal of Clinical Endocrinology & Metabolism**, v. 37, n. 3, p. 415–422, set. 1973.

HAMILTON, J. A. et al. The x-ray crystal structure refinements of normal human transthyretin and the amyloidogenic Val-30→Met variant to 1.7-Å resolution. **Journal of Biological Chemistry**, v. 268, n. 4, p. 2416–2424, fev. 1993.

HENIKOFF, S.; HENIKOFF, J. G. Amino acid substitution matrices from protein blocks. **Proceedings of the National Academy of Sciences**, v. 89, n. 22, p. 10915–10919, 15 nov. 1992.

HENNEBRY, S. C. et al. The Crystal Structure of the Transthyretin-like Protein from *Salmonella dublin*, a Prokaryote 5-Hydroxyisourate Hydrolase. **Journal of Molecular Biology**, v. 359, n. 5, p. 1389–1399, jun. 2006.

HÖRNBERG, A. et al. A comparative analysis of 23 structures of the amyloidogenic protein transthyretin. **Journal of Molecular Biology**, v. 302, n. 3, p. 649–669, set. 2000.

INGBAR, S. H. Pre-Albumin: A Thyroxine-Binding Protein Of Human Plasma. **Endocrinology**. v. 68, p. 4, 1958.

JACOBSON, D. R.; GOREVIC, P. D.; BUXBAUM, J. N. A homozygous transthyretin variant associated with senile systemic amyloidosis: evidence for a late-onset disease of genetic etiology. **American Journal of Human Genetics**, v. 47, n. 1, p. 127–136, jul. 1990.

JACOBSON, DANIELR. et al. Transthyretin Pro55, a variant associated with early-onset, aggressive, diffuse amyloidosis with cardiac and neurologic involvement. **Human Genetics**, v. 89, n. 3, maio 1992.

JAYARAM, B. et al. Sanjeevini: a freely accessible web-server for target directed lead molecule discovery. **BMC Bioinformatics**, v. 13, n. S17, p. S7, dez. 2012.

JUMPER, J. et al. Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, n. 7873, p. 583–589, 2021.

JUNG, D.-K. et al. Structural and functional analysis of PucM, a hydrolase in the ureide pathway and a member of the transthyretin-related protein family. **Proceedings of the National Academy of Sciences**, v. 103, n. 26, p. 9790–9795, 27 jun. 2006.

JURCIK, A. et al. CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. **Bioinformatics**, v. 34, n. 20, p. 3586–3588, 15 out. 2018.

JURRUS, E. et al. Improvements to the APBS biomolecular solvation software suite. **Protein Science : A Publication of the Protein Society**, v. 27, n. 1, p. 112–128, jan. 2018.

KABAT, E. A.; MOORE, D. H.; LANDOW, H. AN ELECTROPHORETIC STUDY OF THE PROTEIN COMPONENTS IN CEREBROSPINAL FLUID AND THEIR RELATIONSHIP TO THE SERUM PROTEINS 1. **Journal of Clinical Investigation**, v. 21, n. 5, p. 571–577, 1 set. 1942.

- KAHN, K.; TIPTON, P. A. Spectroscopic Characterization of Intermediates in the Urate Oxidase Reaction. **Biochemistry**. p. 9, 1998.
- KAHN, K.; SERFOZO, P.; TIPTON, P. A. Identification of the True Product of the Urate Oxidase Reaction. **Journal of the American Chemical Society**. v. 119, p. 5435-5442, 1997.
- KEEBAUGH, A. C.; THOMAS, J. W. The Evolutionary Fate of the Genes Encoding the Purine Catabolic Enzymes in Hominoids, Birds, and Reptiles. **Molecular Biology and Evolution**, v. 27, n. 6, p. 1359–1369, 1 jun. 2010.
- KEILIN, J. The Biological Significance Of Uric Acid And Guanine Excretion. **Biological Reviews**. p. 30, 1959.
- LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283–291, 1 abr. 1993.
- LASKOWSKI, R. A. Enhancing the functional annotation of PDB structures in PDBsum using key figures extracted from the literature. **Bioinformatics**, v. 23, n. 14, p. 1824–1827, 15 jul. 2007.
- LEACH, B. I. et al. NMR measurements reveal the structural basis of transthyretin destabilization by pathogenic mutations. **Biochemistry**. p. 25, 2019.
- LEE, Y. et al. Transthyretin-related proteins function to facilitate the hydrolysis of 5-hydroxyisourate, the end product of the uricase reaction. **FEBS Letters**, v. 579, n. 21, p. 4769–4774, 29 ago. 2005.
- LETUNIC, I.; BORK, P. 20 years of the SMART protein domain annotation resource. **Nucleic Acids Research**, v. 46, n. D1, p. D493–D496, 4 jan. 2018.
- LETUNIC, I.; BORK, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. **Nucleic Acids Research**. p. 4, 2021.
- LETUNIC, I.; KHEDKAR, S.; BORK, P. SMART: recent updates, new developments and status in 2020. **Nucleic Acids Research**, v. 49, n. D1, p. D458–D460, 8 jan. 2021.
- LIMA, L. M. T. R.; ICART, L. P. Amyloidogenicity of peptides targeting diabetes and obesity. **Colloids and Surfaces B: Biointerfaces**, v. 209, p. 112157, jan. 2022.
- LIZ, M. A. et al. Transthyretin, a New Cryptic Protease. **Journal of Biological Chemistry**, v. 279, n. 20, p. 21431–21438, maio 2004.
- LIZ, M. A. et al. Substrate specificity of transthyretin: identification of natural substrates in the nervous system. **Biochemical Journal**, v. 419, n. 2, p. 467–474, 15 abr. 2009.
- LIZ, M. A. et al. Transthyretin is a metallopeptidase with an inducible active site. **Biochemical Journal**, v. 443, n. 3, p. 769–778, 1 maio 2012.
- LIZ, M. A. et al. A Narrative Review of the Role of Transthyretin in Health and Disease. **Neurology and Therapy**, v. 9, n. 2, p. 395–402, 1 out. 2020.
- LIZ, M. A.; SOUSA, M. M. Deciphering cryptic proteases. **CMLS Cellular and Molecular Life Sciences**, v. 62, n. 9, p. 989–1002, maio 2005.

LUEPRASITSAKUL, W. et al. Flavonoid Administration Immediately Displaces Thyroxine (T_4) from Serum Transthyretin, Increases Serum Free T_4 , and Decreases Serum Thyrotropin in the Rat*. **Endocrinology**, v. 126, n. 6, p. 2890–2895, jun. 1990.

LUNDBERG, E. et al. The transthyretin-related protein: Structural investigation of a novel protein family. **Journal of Structural Biology**, p. 13, 2006.

LYU, Z. et al. Protein Secondary Structure Prediction With a Reductive Deep Learning Method. **Frontiers in Bioengineering and Biotechnology**, v. 9, p. 687426, 15 jun. 2021.

MACIEJEWSKI, M. W. et al. NMRbox: A Resource for Biomolecular NMR Computation. **Biophysical Journal**, v. 112, n. 8, p. 1529–1534, abr. 2017.

MADEIRA, F. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. **Nucleic Acids Research**, v. 50, n. W1, p. W276–W279, 5 jul. 2022.

MARIANI, V. et al. IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. **Bioinformatics**, v. 29, n. 21, p. 2722–2728, 1 nov. 2013.

MCGUFFIN, L. J.; BRYSON, K.; JONES, D. T. The PSIPRED protein structure prediction server. **Bioinformatics**, v. 16, n. 4, p. 404–405, 1 abr. 2000.

MELLONI, R. et al. Densidade e diversidade fenotípica de bactérias diazotróficas endofíticas em solos de mineração de bauxita, em reabilitação. **Revista Brasileira de Ciência do Solo**, v. 28, n. 1, p. 85–93, fev. 2004.

MENDEL, C. M. et al. Thyroid Hormone-Binding Proteins in Plasma Facilitate Uniform Distribution of Thyroxine within Tissues: A Perfused Rat Liver Study*. **Endocrinology**, v. 120, n. 5, p. 1742–1749, maio 1987.

MIRDITA, M. et al. ColabFold: making protein folding accessible to all. **Nature Methods**, v. 19, n. 6, p. 679–682, jun. 2022.

MISTRY, J. et al. Pfam: The protein families database in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D412–D419, 8 jan. 2021.

MONACO, H. L. The transthyretin-retinol-binding protein complex. **Biochimica et Biophysica Acta**, p. 8, [s.d.].

Mutations in Hereditary Amyloidosis. Disponível em: <<http://amyloidosismutations.com/>>. Acesso em: 29 out. 2022.

NAYLOR, H. M.; NEWCOMER, M. E. The Structure of Human Retinol-Binding Protein (RBP) with Its Carrier Protein Transthyretin Reveals an Interaction with the Carboxy Terminus of RBP. **Biochemistry**, v. 38, n. 9, p. 2647–2653, 1 mar. 1999.

NEUMANN, P.; CODY, V.; WOJTCZAK, A. Structural basis of negative cooperativity in transthyretin. **Acta Biochimica Polonica**, v. 48, n. 4, p. 867–875, 31 dez. 2001.

PALMIERI, L. DE C. et al. Novel Zn^{2+} -binding Sites in Human Transthyretin. **Journal of Biological Chemistry**, v. 285, n. 41, p. 31731–31741, out. 2010.

PETTERSEN, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605–1612, out. 2004.

PETTERSEN, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. **Protein Science: A Publication of the Protein Society**, v. 30, n. 1, p. 70–82, jan. 2021.

PINNEY, J.H., et al. Senile systemic amyloidosis: clinical features at presentation and outcome. **Journal of the American Heart Association**, 2013.

PLANTÉ-BORDENEUVE, V.; SAID, G. Familial amyloid polyneuropathy. **The Lancet Neurology**, v. 10, n. 12, p. 1086–1097, dez. 2011.

PONTIUS, J.; RICHELLE, J.; WODAK, S. J. Deviations from Standard Atomic Volumes as a Quality Measure for Protein Crystal Structures. **Journal of Molecular Biology**, v. 264, n. 1, p. 121–136, nov. 1996.

PRAPUNPOJ, P. et al. Evolution of structure, ontogeny of gene expression, and function of *Xenopus laevis* transthyretin. **American Journal of Physiology-Regulatory, Integrative and Comparative Physiology**, v. 279, n. 6, p. R2026–R2041, 1 dez. 2000.

PRAPUNPOJ, P.; LEELAWATWATTANA, L. Evolutionary changes to transthyretin: structure-function relationships: Structure-function relationships of transthyretin. **FEBS Journal**, v. 276, n. 19, p. 5330–5341, out. 2009.

RAGAB, G.; ELSHAHALY, M.; BARDIN, T. Gout: An old disease in new perspective – A review. **Journal of Advanced Research**, v. 8, n. 5, p. 495–511, set. 2017.

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.

RAMAZZINA, I. et al. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. **Nature Chemical Biology**, v. 2, n. 3, p. 144–148, mar. 2006.

RAYCHAUDHURI, A.; TIPTON, P. A. A Familiar Motif in a New Context: the Catalytic Mechanism of Hydroxyisourate Hydrolase[†]. **Biochemistry**, v. 42, n. 22, p. 6848–6852, jun. 2003.

REGENITER, A.; SIEDE, W. H. Peaks and tails: Evaluation of irregularities in capillary serum protein electrophoresis. **Clinical Biochemistry**, v. 51, p. 48–55, jan. 2018.

RICHARDSON, S. J. Tweaking the Structure to Radically Change the Function: The Evolution of Transthyretin from 5-Hydroxyisourate Hydrolase to Triiodothyronine Distributor to Thyroxine Distributor. **Frontiers in Endocrinology**, v. 5, p. 245, 11 fev. 2015.

RICHARDSON, S. J.; CODY, V. (EDS.). Recent advances in transthyretin evolution, structure and biological functions. Dordrecht ; New York: **Springer**, 2009.

RIISØEN, H. Reduced prealbumin (transthyretin) in CSF of severely demented patients with Alzheimer's disease. **Acta Neurologica Scandinavica**, v. 78, n. 6, p. 455–459, dez. 1988.

ROMERO, P. A; ARNOLD, F. H. Exploring protein fitness landscapes by directed evolution. **Nature Reviews Molecular Cell Biology**. p. 11, 2009.

ROWCZENIO, D. M. et al. Online Registry for Mutations in Hereditary Amyloidosis Including Nomenclature Recommendations. **Human Mutation**, v. 35, n. 9, p. E2403–E2412, 2014.

RUBERG, F. L. et al. Transthyretin Amyloid Cardiomyopathy. **Journal of the American College of Cardiology**, v. 73, n. 22, p. 2872–2891, jun. 2019.

SANT'ANNA, R. et al. Cavity filling mutations at the thyroxine-binding site dramatically increase transthyretin stability and prevent its aggregation. **Scientific Reports**, v. 7, n. 1, p. 44709, abr. 2017.

SANTOS, C. R. A.; ANJOS, L.; POWER, D. M. Transthyretin in Fish: State of the Art. **Clinical Chemistry and Laboratory Medicine**, v. 40, n. 12, 10 jan. 2002.

SANTHOSHKUMAR, R.; YUSUF, A. In silico structural modeling and analysis of physicochemical properties of curcumin synthase (CURS1, CURS2, and CURS3) proteins of *Curcuma longa*. **Journal of Genetic Engineering and Biotechnology**, v. 18, n. 1, p. 24, dez. 2020.

SARAIVA, M. J. et al. Amyloid fibril protein in familial amyloidotic polyneuropathy, Portuguese type. Definition of molecular abnormality in transthyretin (prealbumin). **Journal of Clinical Investigation**, v. 74, n. 1, p. 104–119, jul. 1984.

SCHULTZ, A. C.; NYGAARD, P.; SAXILD, H. H. Functional Analysis of 14 Genes That Constitute the Purine Catabolic Pathway in *Bacillus subtilis* and Evidence for a Novel Regulon Controlled by the PucR Transcription Activator. **Journal of Bacteriology**, v. 183, n. 11, p. 3293–3302, jun. 2001.

SCHWARZMAN, A.L. AND GOLDGABER, D. Interaction of transthyretin with amyloid beta-protein: binding and inhibition of amyloid formation. **Ciba Found. Symp.** 199, 146–160., 1996

SEIBERT, F. B.; NELSON, J. W. ELECTROPHORETIC STUDY OF THE BLOOD PROTEIN RESPONSE IN TUBERCULOSIS. **Journal of Biological Chemistry**, v. 143, n. 1, p. 29–38, mar. 1942.

SHEN, M.; SALI, A. Statistical potential for assessment and prediction of protein structures. **Protein Science: A Publication of the Protein Society**, v. 15, n. 11, p. 2507–2524, nov. 2006.

SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, v. 7, p. 539, 11 out. 2011.

SMITH, B. A.; HECHT, M. H. Novel proteins: from fold to function. **Current Opinion in Chemical Biology**, v. 15, n. 3, p. 421–426, jun. 2011.

SMITH, T. J. et al. Retinoic acid inhibition of thyroxine binding to human transthyretin. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1199, n. 1, p. 76–80, jan. 1994.

SOUSA, J. C. et al. Transthyretin is involved in depression-like behaviour and exploratory activity. **Journal of Neurochemistry**, v. 88, n. 5, p. 1052–1058, 2004.

SOUSA, M. M.; BERGLUND, L.; SARAIVA, M. J. Transthyretin in high density lipoproteins: association with apolipoprotein A-I. **Journal of Lipid Research**, v. 41, n. 1, p. 58–65, jan. 2000.

STÄNDER, S. et al. Epitope Mapping of Polyclonal Antibodies by Hydrogen–Deuterium Exchange Mass Spectrometry (HDX-MS). **Analytical Chemistry**, v. 93, n. 34, p. 11669–11678, 31 ago. 2021.

STEINEGGER, M.; SÖDING, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. **Nature Biotechnology**, v. 35, n. 11, p. 1026–1028, nov. 2017.

SWINT-KRUSE, L. Using Evolution to Guide Protein Engineering: The Devil IS in the Details. **Biophysical Journal**, v. 111, n. 1, p. 10–18, jul. 2016.

TANAKA, T.; TATENO, Y.; GOJOBORI, T. Evolution of Vitamin B6 (Pyridoxine) Metabolism by Gain and Loss of Genes. **Molecular Biology and Evolution**. p. 8, 2005.

TANG, K.; ZHANG, J.; LIANG, J. Fast Protein Loop Sampling and Structure Prediction Using Distance-Guided Sequential Chain-Growth Monte Carlo Method. **PLoS Computational Biology**, v. 10, n. 4, p. e1003539, 24 abr. 2014.

TOMAR, D. et al. Crystallographic Study of Novel Transthyretin Ligands Exhibiting Negative-Cooperativity between Two Thyroxine Binding Sites. **PLoS ONE**, v. 7, n. 9, p. e43522, 4 set. 2012.

TUNYASUVUNAKOOL, K. et al. Highly accurate protein structure prediction for the human proteome. **Nature**, v. 596, n. 7873, p. 590–596, 26 ago. 2021.

WEBB, B.; SALI, A. Comparative Protein Structure Modeling Using MODELLER. **Current protocols in bioinformatics**, v. 54, p. 5.6.1-5.6.37, 20 jun. 2016.

WILKINSON-WHITE, L. E.; EASTERBROOK-SMITH, S. B. Characterization of the Binding of Cu(II) and Zn(II) to Transthyretin: Effects on Amyloid Formation. **Biochemistry**, v. 46, n. 31, p. 9123–9132, 1 ago. 2007.

WILLARD, L. VADAR: a web server for quantitative evaluation of protein structure quality. **Nucleic Acids Research**, v. 31, n. 13, p. 3316–3319, 1 jul. 2003.

WOJTCZAK, A. et al. Structures of Human Transthyretin Complexed with Thyroxine at 2.0 Å Resolution and 3',5'-Dinitro- *N* -acetyl- L -thyronine at 2.2 Å Resolution. **Acta Crystallographica Section D Biological Crystallography**, v. 52, n. 4, p. 758–765, 1 jul. 1996.

YAMAUCHI, K. et al. Purification and characterization of a 3,5,3'-L-triiodothyronine-specific binding protein from bullfrog tadpole plasma: a homolog of mammalian transthyretin. **Endocrinology**, v. 132, n. 5, p. 2254–2261, 1 maio 1993.

YEE, A. W. et al. A molecular mechanism for transthyretin amyloidogenesis. **Nature Communications**, v. 10, n. 1, p. 925, dez. 2019.

ZANOTTI, G. et al. Structure of Zebra fish HIUase: InsigHT into Evolution of an Enzyme to a Hormone Transporter. **Journal of Molecular Biology**, v. 363, n. 1, p. 1–9, out. 2006.

7. APÊNDICES

Apêndice I - Mapeamento da nomenclatura de resíduos importantes para as cavidades, estruturas e outras funções de HIUase e TTR de diferentes organismos. São apresentadas quatro proteínas: HIUase de *Herbaspirillum seropedicae*, *Danio rerio* e *Escherichia coli*; e TTR de *Homo sapiens*, a nomenclatura utilizada para os mesmos resíduos em cada proteína, suas funções em HIUase e TTR, além da característica química dos resíduos. Tabela vermelha: resíduos pertencentes às cavidades e/ou catalíticos. Tabela azul: resíduos estruturais não relacionados às cavidades. Tabela verde: resíduos que desempenham outras funções.

Cavidade/Catalíticos							
Proteína				Função em HIUase ^e	Característica Química em HIUase ^f	Função em TTR ^e	Característica Química em TTR ^f
HbsHIU ^a	ZfHIU ^b	EcHIU ^c	HsTTR ^d				
H7	H12	H9	K15	Bolsão Central ZBS1	Básico	HBP1/HBP2 Contato com Tiroxina	Básico
L9	L14	L11	L17	Fundo da Cavidade	Apolar	HBP1 Contato com Tiroxina	Apolar
D43	D50	D45	S52	Entrada da Cavidade	Ácido	-	Polar
R45	R52	R47	E54	Entrada da Cavidade	Básico	Contato com Tiroxina	Ácido
H101	H103	H98	T106	Entrada da Cavidade ZBS1	Básico	HBP1 Contato com Tiroxina	Polar
P103	P105	P100	A108	Irregularidades em Fita β	Polar	HBP3/HBP2 Contato com Tiroxina	Apolar
L104	L106	L101	A109	Fundo da Cavidade	Apolar	HBP2 Contato com Tiroxina	Apolar
V105	L107	L102	L110	-	Apolar	HBP3/HBP2 Contato com Tiroxina	Apolar
Y111	Y113	Y108	Y116	-	Aromático	-	Aromático

S112	S114	S109	S117	-	Polar	HBP3	Polar
T113	T115	T110	T118	-	Polar	-	Polar
Y114	Y116	Y111	T119	Encerra a Cavidade	Aromático	HBP3	Polar
G116	G118	G113	V121	-	Apolar	HBP1	Apolar
S117	S119	S114	V122	Centro do Sítio Ativo	Polar	-	Apolar

Estruturais							
Proteína				Função em HIUase ^e	Característica Química em HIUase ^f	Função em TTR ^e	Característica Química em TTR ^f
HbsHIU ^a	ZfHIU ^b	EcHIU ^c	HsTTR ^d				
I11	I16	Q13	A19	Liga a L107 (em ZfHIU)	Apolar/Polar	-	Apolar
A90	V92	E87	V94	Dimerização	Apolar/Ácido	Dimerização	Apolar
W109	F111	Y106	Y114	Dimerização	Aromático	Dimerização	Aromático
R115	R117	R112	A120	Dimerização	Básico	Dimerização	Apolar

Outras Funções							
Proteína				Função em HIUase ^e	Característica Química em HIUase ^f	Função em TTR ^e	Característica Química em TTR ^f
HbsHIU ^a	ZfHIU ^b	EcHIU ^c	HsTTR ^d				
G2	S7	N4	C10	-	Apolar/Polar	ZBS1	Polar
H7	H12	H9	K15	ZBS1 Bolsão Central	Básico	HBP1/HBP2 Contato com Tiroxina	Básico
A23	V28	T25	H31	-	Apolar/Polar	ZBS3/ZBS4	Básico
K32	W39	W34	W41	Ligação à Sulfato	Básico/Aromático	-	Aromático
D47	P54	K49	H56	-	Ácido/Polar/Básico	ZBS1	Básico
G53	K59	P53	E61	-	Apolar/Básico/Polar	ZBS4	Ácido
E54	E60	E54	E62	-	Ácido	ZBS3	Ácido

E62	K68	R63	K70	Ligação à Sulfato	Ácido/Básico	-	Básico
V64	R70	V65	E72	-	Apolar/Básico	ZBS3	Ácido
A66	E72	K67	D74	-	Apolar/Ácido/Básico	ZBS3/ZBS4	Ácido
V84	Y86	F81	H88	-	Apolar/Aromático	ZBS2	Básico
E86	Y88/E90	E83	H90	Dimerização	Ácido	Dimerização ZBS2	Básico
V88	E90	P85	E92	-	Apolar/Ácido/Polar	ZBS2	Ácido
H101	H103	H98	T106	ZBS1 Entrada da Cavidade	Básico	HBP1 Contato com Tiroxina	Polar

^a*Herbaspirillum seropedicae* HIUase: estrutura não depositada, cedida pelo Prof. Lucas Bleicher (ICB – UFMG).

^b*Zebrafish (Danio rerio)* HIUase: ID PDB: 2H1X. Zanotti *et al.*, 2006.

^c*Escherichia coli* HIUase: ID PDB: 2G2N. Lundberg *et al.*, 2006.

^d*Homo sapiens* TTR: ID PDB: 1F41. Tomar *et al.*, 2012.

^eFunção atribuída com base em busca na literatura e análise visual das estruturas tridimensionais.

^fCaracterísticas químicas atribuídas de acordo com a seguinte classificação: Apolares – G, A, V, L, I, M; Polares – S, T, C, P, N, Q; Aromáticos – F, Y, W; Básicos – K, H, R; Ácidos – D, E.

Apêndice II: Mapeamento da nomenclatura de resíduos importantes para as cavidades, estruturas e outras funções das proteínas mutantes e suas proteínas de referência. São discutidas quatro proteínas: HIUase de *Herbaspirillum seropedicae*, Mutante de TTR, TTR de *Homo sapiens* e Mutante de HIUase, a nomenclatura utilizada para os mesmos resíduos em cada proteína, suas características químicas nas proteínas de referência e a comparação das características químicas entre as proteínas de referência e as mutantes. Tabela vermelha: resíduos pertencentes às cavidades e/ou catalíticos. Tabela azul: resíduos estruturais não relacionados às cavidades. Tabela verde: resíduos que desempenham outras funções.

Cavidade/Catalíticos							
Proteína		Característica Química em HIUase ^c	Característica Química da MutTTR é Mantida Quando Comparada com HbsHIU? ^d	Proteína		Característica Química em TTR ^c	Característica Química da MutHIU é Mantida Quando Comparada com HsTTR? ^d
HbsHIU ^a	MutTTR ^b			HsTTR ^e	MutHIU ^f		
H7	H15	Básico	Sim	K15	K9	Básico	Sim
L9	L17	Apolar	Sim	L17	L11	Apolar	Sim
D43	D52	Ácido	Sim	S52	D45	Ácido	Não
R45	R54	Básico	Sim	E54	R47	Básico	Não
H101	H106	Básico	Sim	T106	H106	Básico	Não
P103	P108	Apolar	Sim	A108	A108	Polar	Sim
L104	L109	Apolar	Sim	A109	L109	Apolar	Sim
V105	L110	Apolar	Sim	L110	V110	Apolar	Sim
Y111	Y116	Aromático	Sim	Y116	Y116	Aromático	Sim
S112	S117	Polar	Sim	S117	S117	Polar	Sim
T113	T118	Polar	Sim	T118	T118	Polar	Sim
Y114	Y119	Polar	Sim	T119	T119	Aromático	Sim
G116	G121	Apolar	Sim	V121	G121	Apolar	Sim
S117	S122	Apolar	Sim	V122	V122	Polar	Sim

Estruturais							
Proteína		Característica Química em HIUase ^c	Característica Química da MutTTR é Mantida Quando Comparada com HbsHIU? ^d	Proteína		Característica Química em TTR ^c	Característica Química da MutHIU é Mantida Quando Comparada com HsTTR? ^d
HbsHIU ^a	MutTTR ^b			HsTTR ^e	MutHIU ^f		
I11	A19	Apolar	Sim	A19	A13	Apolar	Sim
A90	V94	Apolar	Sim	V94	A94	Apolar	Sim
W109	Y114	Aromático	Sim	Y114	W114	Aromático	Sim
R115	R120	Apolar	Sim	A120	A120	Básico	Sim

Outras Funções							
Proteína		Característica Química em HIUase ^c	Característica Química da MutTTR é Mantida Quando Comparada com HbsHIU? ^d	Proteína		Característica Química em TTR ^c	Característica Química da MutHIU é Mantida Quando Comparada com HsTTR? ^d
HbsHIU ^a	MutTTR ^b			HsTTR ^e	MutHIU ^f		
G2	S10	Polar	Não	C10	C5	Polar	Sim
H7	H15	Básico	Sim	K15	K9	Básico	Sim
A23	H31	Apolar	Não	H31	A25	Básico	Não
K32	W41	Básico	Não	W41	K34	Aromático	Não
D47	P56	Ácido	Não	H56	D49	Básico	Não
G53	E62	Ácido	Não	E61	E57	Ácido	Sim
E54	E63	Apolar	Sim	E62	A58	Ácido	Não
E62	K70	Ácido	Não	K70	E65	Básico	Não
V64	E72	Apolar	Não	E72	V67	Ácido	Não
A66	D74	Apolar	Não	D74	A69	Ácido	Não
V84	Y88	Apolar	Não	H88	V87	Básico	Não
E86	P89	Ácido	Não	H90	E90	Básico	Não

V88	E92	Apolar	Não	E92	V92	Ácido	Não
H101	H106	Básico	Sim	T106	H106	Polar	Não

^a*Herbaspirillum seropedicae* HIUase: estrutura não depositada, cedida pelo Prof. Lucas Bleicher (ICB – UFMG).

^bProteína Mutante de TTR.

^cCaracterísticas químicas atribuídas de acordo com a seguinte classificação: Apolares – G, A, V, L, I, M; Polares – S, T, C, P, N, Q; Aromáticos – F, Y, W; Básicos – K, H, R; Ácidos – D, E.

^dFoi comparada a característica química da proteína de referência (HbsHIU ou HsTTR) com a característica química do mesmo resíduo nas proteínas mutantes (MutTTR e MutHIU).

^e*Homo sapiens* TTR: ID PDB: 1F41. Tomar *et al.*, 2012.

^fProteína Mutante de HIUase.

Apêndice III – Lista completa de resíduos expostos nas proteínas de referência e mutantes. A) HIUase de *Herbaspirillum seropedicae*. B) TTR de *Homo sapiens*. C) Mutante de HIUase (MutHIU). D) Mutante de TTR (MutTTR). São mostrados, em ordem, o número dos resíduos, nomenclatura, estrutura secundária predita, valores de ASA das cadeias laterais, ASA fracional das cadeias laterais (0 – 1), ASA das cadeias principais, ASA fracional das cadeias principais (0 – 1), somatório dos valores de ASA das cadeias laterais e principais, e classificação. Resíduos com ASA fracional da cadeia lateral superior a 0.5 foram considerados expostos. Resíduos enterrados não são mostrados. Estruturas secundárias são classificadas em: C: *coil*. B: folha- β . H: α -hélice.

A)

HbsHIU

Número Resíduo	Nome Resíduo	Estrutura Secundária	ASA Lateral	ASA Lateral Fracional	ASA Principal	ASA Principal Fracional	ASA Total (Lateral + Principal)	Classificação
1	MET	C	119.8	0.66	140.4	0.64	260.2	exposto
3	LYS	B	104.2	0.59	104.8	0.49	209	exposto
18	VAL	B	67.5	0.51	76.7	0.45	144.2	exposto
29	PRO	C	131.7	0.99	152.6	0.99	284.3	exposto
30	VAL	C	77.1	0.58	102.4	0.6	179.5	exposto
31	GLY	C	39	0.74	91	1	130	exposto
33	THR	B	77.7	0.69	83.6	0.55	161.3	exposto
37	GLN	B	103.8	0.69	124.8	0.66	228.6	exposto
39	VAL	B	72.2	0.54	84.7	0.49	156.9	exposto
42	SER	C	79.5	0.85	90.7	0.69	170.2	exposto
47	ASP	C	94.7	0.8	118.9	0.76	213.6	exposto
48	GLU	C	123.7	0.79	126.7	0.67	250.4	exposto
53	GLY	C	32	0.61	33.1	0.36	65.1	exposto
54	GLU	C	149.6	0.96	155.1	0.82	304.7	exposto
55	ALA	C	44.2	0.52	48.5	0.39	92.7	exposto
57	GLN	C	110.8	0.74	110.8	0.59	221.6	exposto
69	ASP	H	68	0.58	71.5	0.45	139.5	exposto
73	ALA	H	62	0.72	87.3	0.71	149.3	exposto
75	GLY	C	38.6	0.73	58.8	0.65	97.4	exposto
77	GLN	C	144.1	0.96	166.3	0.88	310.4	exposto
79	PRO	C	103.9	0.78	105.3	0.68	209.2	exposto
80	GLU	C	102.5	0.66	123	0.65	225.5	exposto
86	GLU	B	84.9	0.54	87.6	0.46	172.5	exposto
95	ASP	C	62.2	0.53	62.2	0.4	124.4	exposto
98	GLN	B	79.2	0.53	79.2	0.42	158.4	exposto
99	ASN	B	74	0.58	84.2	0.51	158.2	exposto

B)

HsTTR

Número Resíduo	Nome Resíduo	Estrutura Secundária	ASA Lateral	ASA Lateral Fracional	ASA Principal	ASA Principal Fracional	ASA Total (Lateral + Principal)	Classificação
21	ARG	C	114.6	0.56	129.1	0.53	243.7	exposto
27	ASN	C	85.7	0.67	100	0.6	185.7	exposto
37	ALA	C	92.8	1.08	118.3	0.96	211.1	exposto
38	ASP	C	86	0.73	99.5	0.63	185.5	exposto
39	ASP	C	86.6	0.73	98.4	0.63	185	exposto
40	THR	B	79	0.7	79	0.52	158	exposto
42	GLU	B	89.4	0.57	89.4	0.47	178.8	exposto
43	PRO	B	96.4	0.73	111.9	0.72	208.3	exposto
48	LYS	B	113.6	0.64	126.3	0.59	239.9	exposto
50	SER	C	56.8	0.6	56.8	0.43	113.6	exposto
51	GLU	C	135.7	0.87	142	0.75	277.7	exposto
52	SER	C	57.1	0.61	58.2	0.44	115.3	exposto
56	HIS	B	86.4	0.52	101	0.5	187.4	exposto
57	GLY	C	37.6	0.72	64.8	0.71	102.4	exposto
60	THR	C	82.7	0.73	82.7	0.55	165.4	exposto
61	GLU	C	99.4	0.64	110.9	0.58	210.3	exposto
62	GLU	C	147.8	0.95	159.1	0.84	306.9	exposto
63	GLU	C	83.4	0.53	86.5	0.46	169.9	exposto
65	VAL	C	81.3	0.61	81.3	0.47	162.6	exposto
66	GLU	B	96.7	0.62	99.7	0.53	196.4	exposto
77	SER	H	59.7	0.63	62.9	0.48	122.6	exposto
80	LYS	H	99.6	0.56	114	0.53	213.6	exposto
81	ALA	H	68.1	0.8	94.9	0.77	163	exposto
83	GLY	C	41.4	0.79	72.4	0.8	113.8	exposto
85	SER	C	82.7	0.88	94	0.71	176.7	exposto
99	ASP	C	66	0.56	92.1	0.59	158.1	exposto
100	SER	C	72.7	0.77	86.2	0.65	158.9	exposto
101	GLY	C	28.4	0.54	30.7	0.34	59.1	exposto
102	PRO	C	76.8	0.58	77.2	0.5	154	exposto
104	ARG	B	120.7	0.59	121.3	0.5	242	exposto
123	THR	B	58.2	0.51	58.2	0.38	116.4	exposto
124	ASN	B	80.9	0.63	100.3	0.61	181.2	exposto

C)

MutHIU

Número Resíduo	Nome Resíduo	Estrutura Secundária	ASA Lateral	ASA Lateral Fracional	ASA Principal	ASA Principal Fracional	ASA Total (Lateral + Principal)	Classificação
1	GLY	C	60.7	1.16	127.7	1.41	188.4	exposto
2	MET	C	110.4	0.61	122.8	0.56	233.2	exposto
3	GLY	C	47.8	0.91	54.9	0.6	102.7	exposto
23	LYS	B	128.4	0.73	128.9	0.6	257.3	exposto
30	GLY	C	28.6	0.54	36.7	0.4	65.3	exposto
31	PRO	C	122.8	0.93	153.2	0.99	276	exposto
32	VAL	C	136.7	1.03	160	0.93	296.7	exposto
33	GLY	C	42.2	0.8	47.7	0.53	89.9	exposto
35	THR	B	91.1	0.8	91.2	0.6	182.3	exposto
39	GLN	B	104	0.69	122.4	0.65	226.4	exposto
41	VAL	B	71.5	0.54	85.6	0.5	157.1	exposto
44	SER	C	79.8	0.85	91.7	0.7	171.5	exposto
45	ASP	C	63	0.53	63	0.4	126	exposto
49	ASP	C	94.6	0.8	124.7	0.79	219.3	exposto
50	GLU	C	102.2	0.65	103.4	0.54	205.6	exposto
54	ALA	B	56.7	0.66	58	0.47	114.7	exposto
56	GLY	C	44.6	0.85	53.7	0.59	98.3	exposto
57	GLU	C	160.6	1.03	176.1	0.93	336.7	exposto
58	ALA	C	44.7	0.52	46.9	0.38	91.6	exposto
60	HIS	C	84.8	0.51	84.8	0.42	169.6	exposto
72	ASP	H	72.3	0.61	74.4	0.47	146.7	exposto
76	ALA	H	65	0.76	91.6	0.74	156.6	exposto
78	GLY	C	40.8	0.78	66.4	0.73	107.2	exposto
80	GLN	C	161.1	1.07	184.7	0.98	345.8	exposto
83	GLU	C	120.2	0.77	138.1	0.73	258.3	exposto
101	SER	C	75.2	0.8	85.8	0.65	161	exposto

D)

MutTTR

Número Resíduo	Nome Resíduo	Estrutura Secundária	ASA Lateral	ASA Lateral Fracional	ASA Principal	ASA Principal Fracional	ASA Total (Lateral + Principal)	Classificação
18	ASN	C	84.2	0.66	98.5	0.6	182.7	exposto
28	ALA	C	93	1.09	121.6	0.99	214.6	exposto
29	ASP	C	87.3	0.74	102.8	0.65	190.1	exposto
30	ASP	C	86.3	0.73	96.9	0.62	183.2	exposto
31	THR	B	77.1	0.68	77.1	0.51	154.2	exposto
33	GLU	B	93	0.6	93	0.49	186	exposto
34	PRO	B	93.7	0.71	114.3	0.74	208	exposto
39	LYS	B	112.5	0.64	123.7	0.58	236.2	exposto
41	SER	C	53.6	0.57	53.6	0.41	107.2	exposto
42	GLU	C	136.4	0.87	143.7	0.76	280.1	exposto
47	PRO	B	79	0.6	85.1	0.55	164.1	exposto
48	GLY	C	27.1	0.52	51.3	0.56	78.4	exposto
51	THR	C	74	0.65	74	0.49	148	exposto
52	GLU	H	125.7	0.8	139	0.73	264.7	exposto
53	GLU	H	138.1	0.88	148.9	0.78	287	exposto
56	VAL	C	69.7	0.52	69.7	0.41	139.4	exposto
57	GLU	C	101.5	0.65	106.8	0.56	208.3	exposto
67	LYS	H	100.7	0.57	100.7	0.47	201.4	exposto
68	SER	H	58.4	0.62	61.1	0.46	119.5	exposto
71	LYS	H	137.6	0.78	147.6	0.69	285.2	exposto
72	ALA	H	63.7	0.74	91.3	0.74	155	exposto
74	GLY	C	45.3	0.86	78.1	0.86	123.4	exposto
76	SER	C	86.9	0.92	97.7	0.74	184.6	exposto
89	ASN	B	69.8	0.55	74.9	0.45	144.7	exposto
90	ASP	C	101.5	0.86	127.9	0.81	229.4	exposto
92	GLY	C	49.1	0.94	64	0.7	113.1	exposto
94	LYS	C	162.6	0.92	166.5	0.78	329.1	exposto
95	ARG	B	140.3	0.68	141.4	0.58	281.7	exposto